JOURNAL
Science and Technology

Volume 30 - No. (6) - 2025

An Open Access Peer-Reviewed Journal Published Monthly by Faculty of Engineering and Computing - University of Science & Technology - Main Campus Aden - Yemen.

Online ISSN: 2410-5163 Print ISSN: 1607-2073

جامعة العلوم والتكنولوجيا

University of Science & Technology

المركز الرئيس - عـــدن

https://doi.org/10.20428/jst.v30i6

Journal of Science and Technology

Vol. 19 No. 6 (2025)

Journal of Science and Technology

https://doi.org/10.20428/jst.v30i6

Editorial Team

Editor In Chief

Prof. Dr. Abdulaziz Al kabab, Faculty of engineering, University of Science and Technology, Aden. Email: Editor.jst@ust.edu , Personal Email: aalkabab@gmail.com.

Deputy Editors

Assis. Prof. Dr. Nasr Al-Sakkaf, Faculty of Engineering, University of Science and Technology, Aden. Email: Editor.jst@ust.edu , Personal Email: nasrhamed1972@gmail.com.
Assis. Prof. Dr. Nabeel Monasar, Faculty of Engineering, University of Science and Technology, Aden.

Associate editors

Associate Prof. Dr. Abdulqader Alabadi, Aden, Yemen.
Assis. Prof. Dr. Adnan Abdullah Zain - University of Aden.
Assis. Prof. Dr. Lutfi Khanbari - University of Aden.
Assis. Prof. Dr. Muammar Al-Qutaibi - University of Science and Technology, Aden.
Assis. Prof. Dr. Khaled Omar Salem Baselim - University of Aden.
Assis. Prof. Dr. Mustafa Ahmed Shoaib.
Assis. Prof. Dr. Ali Belaid - University of Aden.

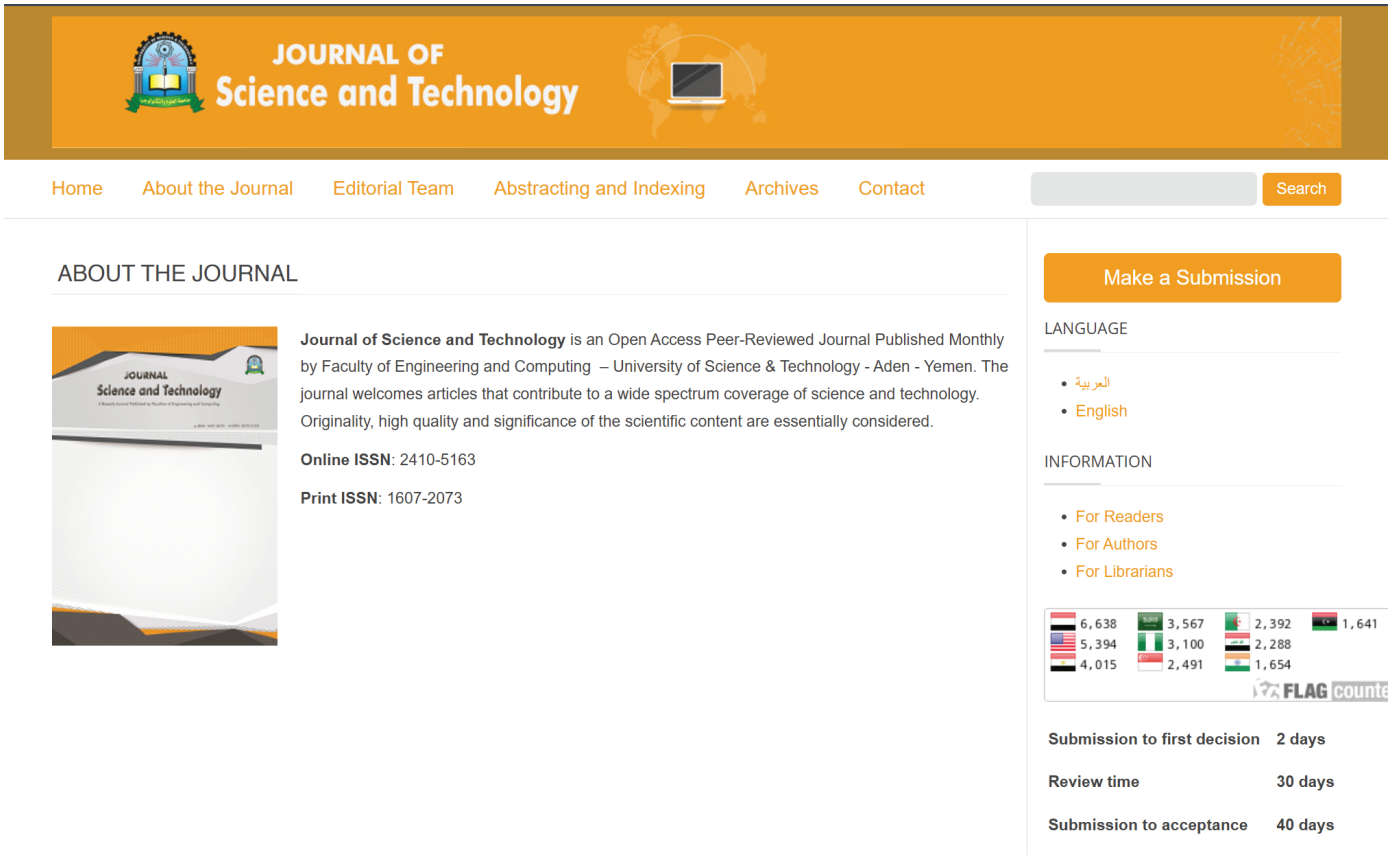
Assistant Editor

Sumaya Al-Badani, University of Science and Technology, Aden.

For Contact
Journal of Science and Technology
Faculty of Engineering and Faculty of Computing and Information Technology, University of Science and Technology, Main Campus - Aden - Yemen
+967718032009
jst@ust.edu

https://doi.org/10.20428/jst.v30i6

Website:



https://journals.ust.edu/index.php/JST

https://doi.org/10.20428/jst.v30i6

Publication Rules and Procedures

General Rules:

Research papers sent for publication in the Journal of Science and Technology at the University of Science and Technology – Aden, Yemen, should meet the following requirements:

1. The journal publishes research papers in Arabic and English, in the following categories (but not limited to): Physics, Chemistry, Mathematics, s Engineering, and Computer Science. if they comply with the following conditions:

a. Title:

A title should summarize the main idea of the manuscript. It should identify the variables under investigation and the relationship between them. It should be concise and fully explanatory for readers when standing alone. It is recommended to be no more than 12 words and with no abbreviations. It should be typed in uppercase and lowercase letters, centered in the upper half of the page.

b. Author's name and institutional affiliation:

Every manuscript should include the name of the author and the institutional affiliation of the author when the research was conducted.

Author's name:

The preferred form of an author's name is first name, middle initial(s), and last name, with no titles (e.g., Dr., Professor) or degrees (e.g., PhD, PsyD, EdD). If the manuscript is done by more than one author, the names of the authors should be in the order of their contributions, centered between the side margins.

Institutional affiliation:

The affiliation (i.e. institution) should be centered under the author's name on the next line. If an author has no institutional affiliation, list the city and state of residence below the author's name. The emails for all authors should be provided.

c. Abstract:

An abstract in both Arabic and English must be included, with the former version **in the language of the manuscript.** It should be in a form of a brief, comprehensive summary of the contents of the manuscript written in a single paragraph with no paragraph indentation. It must not exceed 200 words.

The abstract should clearly describe the problem under investigation, in one sentence if possible; identify the purpose of the research, the essential features of study method; the participants' relevant characteristics such as age, sex, and ethnic and/or racial group; the basic findings including statistical significance levels; and the conclusions and the implications or applications. Keywords (3 - 5) should follow the abstract to increase user's ability to find useful information in the manuscript.

Introduction:

The body of a manuscript starts with an introduction that frames the problem under study and explores the importance of the problem (why the problem deserves new research). The statement about importance might involve the need to resolve any inconsistency in results of past work and/or extend the reach of a theoretical formulation and/ or investigate a practical problem that people suffer. The introduction ends in concluding the statement of the problem with a brief but formal statement of the purpose of the research that summarizes the material preceding it.

The author needs to discuss the relevant related literature in the introduction. A scholarly description of the earlier work will provide a summary of the most recent directly related work and recognize the priority of the work of others. The description of relevant literature will present what other aspects of this study have been investigated in the previous studies and how the current study differs from the earlier ones. For summarizing earlier works, focus should be on the topic (research syntheses of the topic), methodological issues, relevant findings, and main conclusions.

The discussion of related literature should demonstrate the logical continuity between previous and present work (**demonstration of gaps**); and the development of the problem should have enough breadth and clarity that make it easy to understand by a wide range of professionals.

After developing the theoretical background and the problem of the study, the author has to state the objectives and the hypotheses or specific questions. The introduction should be 10 to 15% of the manuscript. It starts on a separate page (i.e. p. 3).

https://doi.org/10.20428/jst.v30i6

d. Method:

This section describes in detail how the study was conducted, including conceptual and operational definitions of the variables used in the study. A comprehensive description of the methods used enables the reader to evaluate the appropriateness of the methods and the reliability and the validity of the results. It may include participant characteristics, sampling procedures, sample size, measures and research design.

e. Results:

Results section summarizes the collected data and the analysis done on those data. It should provide sufficient detail about the data to justify the study conclusions. The results should also include details that may not match the study expectation; or even small effect sizes (or statistically non-significant findings) particularly when theory predicts large (or statistically significant) ones. Uncomfortable results should not be omitted. The data can be presented in tables or figures (data presented in tables should not be represented in figures). Tables should be numbered in order of mention in the text. Tables can be single-spaced and should not contain any lines. Asterisks may be used to indicate significant findings. Symbols, acronyms or abbreviations should be used sparingly. Explanatory footnotes should be used whenever possible rather than overlong titles. Images should be submitted as high-resolution files (300 dpi or higher) in TIFF format (LZW compression) or JPEGs.

f. Discussion:

After presenting the results, their implications should be evaluated and interpreted, especially with respect to the original hypotheses. The author needs to examine, interpret, and draw inferences and conclusions from the results emphasize any theoretical or practical consequences of the results (Results and discussion can be combined in one section). Similarities and differences between the results and the work of others should be used to contextualize, confirm, and clarify the conclusions. Each new statement should contribute to the interpretation and to the reader's understanding of the problem.

The interpretation of the results should discuss the limitations or weaknesses of the study, and address alternative explanations of the results. It also discusses the generalizability of the findings. This critical analysis should take into account differences between the target population and the accessed sample.

g. Conclusion:

This concluding section presents a brief, reasoned and justifiable commentary on the importance of the findings. It is tightly reasoned, self-contained, and not overstated. In this section, the importance of the problem (as stated in the introduction) should be discussed; what larger issues might depend on the findings; and what propositions are confirmed or disconfirmed.

Acknowledgements:

This **section** refers to the aid received by the author from other relevant parties. Reference should also be made to any financial assistance received to conduct the research. Any extraordinary assistance received by the author in word processing, data collection, data analysis, and so on, should be acknowledged. The acknowledgements should not exceed 60 words.

References:

References start on a separate page.

Authors should acknowledge the work of previous scholars by citing references to document statements in their manuscripts. JST uses IEEE standard format. For accurate, complete, and useful citations, authors can use a reference management tool.

Publication Procedures:

Manuscript acceptance, rejection, or acceptance with revision:

The editor decides whether the manuscript is accepted, rejected or needs to be revised based on the reviewers' reports.

Manuscript acceptance: Accepted manuscripts will undergo copy-editing and production phases of publication process. The authors will not be allowed to make further changes to the manuscript except for those recommended by the copyeditors. The authors remain responsible for the completion of any amendments required by the journal.

Manuscript Rejection: A manuscript is rejected if it falls outside the domain of the journal, has serious defects in design, methodology, analysis or interpretations, lack of contribution to the field, or has a low-quality.

Manuscript acceptance with revision:
A manuscript may be conditionally accepted. This takes place when the manuscript has a high potential for final acceptance and publication in the journal, and the author adheres to all the essential modifications required by the journal (e.g. gathering essential data, conducting new experiments, reanalyzing the data, etc.). The author has to attend to the editor's recommendations for revision. The revised manuscript should be resubmitted with an enclosed cover letter that contains a table explaining in detail how and where (in the manuscript) amendments have been done based on the reviewers' comments.

Publication Ethics:
The **JST** complies with the recommendations of the Committee on Publication Ethics (COPE) to promote the integrity of its published articles. The **JST** considers the following topics during the publishing process:

- **Originality & Source Acknowledgement:** The **JST** scans all submitted manuscripts before the peer-reviewing process using Turnitin®. The **JST** is zero-tolerant to plagiarism, self-plagiarism, copyright infringement, dual publication, text recycling and salami slicing. When any of these is identified after publishing, an announcement of retraction of the published material is highlighted in the journal's website. The authors are asked to provide appropriate references for published/unpublished cited texts. The corresponding author should confirm that the submission has not been previously published and is not being considered for publication elsewhere.
- **Research Misconduct:** The **JST** editorial team struggles to counter any possibility for data fabrication, manipulation and falsification. In case of suspected misconduct, the **JST** editors act in accordance to the COPE guidelines with this respect.
- **Conflicts of interest:** Authors should disclose potential conflicts of interest and indicate financial agreements or affiliations with any product or services used in the manuscript (as well as any potential bias against another product or service).
- **Authors:** Authors should disclose (in an author note) activities and relationships that if known to others might be viewed as a conflict of interest, even if the authors do not believe that any conflict or bias exists (e.g. an author has his stock in a company that manufactures a drug used in his study).
- **Reviewers:** Reviewers should also reveal their potential conflicts of interest (if any) to the action editor. They have an ethical obligation to be open and fair in assessing a manuscript without bias. They should not review a manuscript from a colleague or collaborator, a close personal friend, or a recent student. Reviewers should maintain the confidentiality of a manuscript. They should not discuss the manuscript with other individuals.
- **Using materials under copyrights:** The author should obtain letters of permission from copyright holders to reproduce (or adapt) copyrighted material and enclose copies of these letters with the accepted manuscript. Examples of material that require permission include reprinted figures and tables, tests and scale items, questionnaires, vignettes, etc.

Correction notices: If an error is detected in the published manuscript, the author can submit a proposed correction notice to the journal's editor. The notice should indicate the full title of the journal, the year of publication, the volume no., issue no., and the page nos. of the article, the precise location of the error(s) (e.g. page, line, column, exact quotation of the error, or paraphrasing of lengthy errors).

Publication Fees:
JST does not require any article processing charges or any article submission charges

Sponsorship:
JST is sponsored by the University of Science and Technology.

Opinions expressed in the journal are merely those of their authors and do not reflect those of the journal or the University.

https://doi.org/10.20428/jst.v30i6

Content:

Journal of Science and Technology

https://doi.org/10.20428/jst.v30i6

Microbial Contamination in Vegetables and Fruits from Aden Governorate, Yemen: Pathogen Isolation and Analysis of Phytochemical and Physicochemical Properties in Mixed-Herb (Turmeric, Ginger, Indian Costus) Extracts

K. N. Q. Salem [1]
N. M. S. Fartwte [2]
A. F. A. Moqbel [1]
O. S. A. Yousef [1]
A. A. A. Ahmed [1]
W. K. A. Waked [1]
A. Z. A. Abdulhameed [1]
W. G. M. Saleh [1]
D. N. A. Mohammed [1]
Y. K. A. Ali [1]
M. N. N. Mohammed [1]

Received: 25/02/2025
Revised: 23/03/2025
Accepted: 24/03/2025

© 2025 University of Science and Technology, Aden, Yemen. This article can be distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

© 2025 جامعة العلوم والتكنولوجيا، المركز الرئيس عدن، اليمن. يمكن إعادة استخدام المادة المنشورة حسب رخصة مؤسسة المشاع الإبداعي شريطة الاستشهاد بالمؤلف والمجلة.

[1] Department of Biology, Faculty of Sciences, University of Aden, Aden, Yemen
[2] Department of Biology, Faculty of Education, University of Aden, Aden, Yemen
*Corresponding Author's Email: mohamedkhaled73388@gmail.com

Microbial Contamination in Vegetables and Fruits from Aden Governorate, Yemen: Pathogen Isolation and Analysis of Phytochemical and Physicochemical Properties in Mixed-Herb (Turmeric, Ginger, Indian Costus) Extracts

K. N. Q. Salem
Department of Biology, Faculty of Sciences, University of Aden,
Aden, Yemen

N. M. S. Fartwte
Department of Biology, Faculty of Education, University of Aden,
Aden, Yemen

A. F. A. Moqbel
Department of Biology, Faculty of Sciences, University of Aden,
Aden, Yemen

O. S. A. Yousef
Department of Biology, Faculty of Sciences, University of Aden,
Aden, Yemen

A. A. A. Ahmed
Department of Biology, Faculty of Sciences, University of Aden,
Aden, Yemen

W. K. A. Waked
Department of Biology, Faculty of Sciences, University of Aden,
Aden, Yemen

A. Z. A. Abdulhameed
Department of Biology, Faculty of Sciences, University of Aden,
Aden, Yemen

W. G. M. Saleh
Department of Biology, Faculty of Sciences, University of Aden,
Aden, Yemen

D. N. A. Mohammed
Department of Biology, Faculty of Sciences, University of Aden,
Aden, Yemen

Y. K. A. Ali
Department of Biology, Faculty of Sciences, University of Aden,
Aden, Yemen

M. N. N. Mohammed
Department of Biology, Faculty of Sciences, University of Aden,
Aden, Yemen

Abstract— This study aimed to isolate and characterize *Staphylococcus aureus* (*S. aureus*) and *Escherichia coli* (*E. coli*) from vegetable and fruit samples collected in Aden Governorate, Yemen. Additionally, it sought to analyze the phytochemical and physicochemical properties of aqueous and ethanolic extracts from mixed-herb {turmeric, *Curcuma longa (C. longa*), ginger, *Zingiber officinale (Z. officinale*), and Indian costus, *Saussurea costus* (*S costus*)}. A cross-sectional study was conducted between August and November 2023, with samples obtained from central markets. Bacterial isolation was performed using selective media: Mannitol Salt Agar (MSA) for *S. aureus* and Eosin Methylene Blue (EMB) agar for *E. coli*. Results revealed a significantly higher prevalence of *S. aureus* (32.07% in vegetables; 14.29% in fruits) compared to *E. coli* (15.09%) in vegetables; undetected in fruits). Phytochemical screening confirmed the presence of bioactive constituents, including flavonoids, alkaloids, terpenoids, and tannins, which likely contributed to antimicrobial effects and emphasizing their therapeutic potential. The total ash content was measured at 5.78%, indicating suitable mineral content for therapeutic applications. The moisture content was 10.76%, which helps reduce microbial growth. Extractive values showed that water is more efficient in extracting polar compounds, while ethanol isolates non-polar compounds. The pH values of both aqueous and ethanolic extracts were slightly acidic, which helps preserve the stability of bioactive compounds. These findings highlight the need for improved agricultural hygiene and storage practices to minimize bacterial contamination in fresh produce. Additionally, the study advocates for further research into the bioactive compounds of plant extracts, optimization of extraction protocols, and exploration of synergistic combinations to enhance their application in food safety and natural antimicrobial therapies. This study underscores the potential of plant-derived compounds in combating foodborne pathogens and supports their integration into targeted strategies for public health, pharmaceutical development, and sustainable agriculture.

Keywords— *Staphylococcus Aureus*, *Escherichia* Coli, Ethanolic and Aqueous Extracts, Phytochemicals, Physicochemical.

I. Introduction

Staphylococcus aureus and *E. coli* are two significant bacterial pathogens that pose considerable health risks to contaminate food products, particularly vegetables and fruits. These bacteria are associated with a wide range of infections and foodborne illnesses, making them a critical focus for public health interventions [1].

Staphylococcus aureus is a Gram-positive bacterium that is part of the normal flora of the skin and mucous membranes in humans and animals [2]. Despite its commensal nature, it can cause severe infections when it breaches host defenses, leading to conditions ranging from mild skin infections to life-threatening diseases like septicemia [3]. The increasing prevalence of antibiotic-resistant strains, such as methicillin-resistant *S. aureus* (MRSA), exacerbates its threat to public health [4].

Escherichia coli, a Gram-negative bacterium predominantly residing in the intestinal tract of warm-blooded animals, can also lead to serious health complications. Pathogenic strains like *E. coli* O157:H7 produce Shiga toxins, causing severe gastrointestinal distress and systemic complications such as hemolytic uremic syndrome (HUS), particularly in vulnerable populations [5].

The contamination of vegetables and fruits with these pathogens often arises from poor agricultural practices, including the use of contaminated irrigation water or improper handling during harvest and distribution. *S. aureus* is commonly introduced through human contact, while *E. coli* can infiltrate food through exposure to animal waste or untreated water [6, 7]. Such contamination is a leading cause of foodborne illnesses, necessitating robust prevention and mitigation strategies.

Natural antimicrobial agents derived from plants offer a promising alternative to conventional antibiotics, particularly against antibiotic-resistant strains. Medicinal plants like (turmeric, ginger and Indian costus) have been widely recognized for their antimicrobial properties, attributed to bioactive compounds such as alkaloids, flavonoids, and glycosides [8]. These natural compounds provide a basis for exploring plant-based antibacterial treatments that could improve food safety and reduce dependency on synthetic antibiotics.

This study investigates the microbial contamination of vegetables and fruits in Aden Governorate, Yemen, focusing on the isolation and identification of *S. aureus* and *E. coli*. Additionally, it analyzes the phytochemical and physicochemical properties of mixed-herb extracts from (turmeric, ginger and Indian costus). The research aims to develop natural, plant-based solutions to reduce bacterial contamination in food products, thereby enhancing food safety, improving agricultural practices, and mitigating public health risks associated with foodborne illnesses.

II. Materials and Methods

A. Microbiology Methods and Procedures

1. *Study Area*: Aden, a port city in southern Yemen on the Arabian Peninsula, lies near the eastern entrance to the Red Sea. It is located about 170 km (110 mi) east of the Bab-el-Mandeb strait and to the north of the Gulf of Aden. Due to its strategic coastal position, Aden acts as a vital maritime link between the Red Sea and the Arabian Sea, facilitating connections across Africa, Asia, and the Middle East. Historically, Aden has been a major trading hub since ancient times and was a key port during British colonial rule from 1839 to 1967. Its natural harbor, one of the world's largest, has played a significant role in the region's maritime commerce. As of 2023, Aden's population stands at approximately 1,080,000, making it one of Yemen's largest cities [9].
2. *Sample Collection:* Vegetable and fruit samples, including lettuce, cabbage, Spinach, carrots, potatoes, beets, strawberries, grapes, and watermelon, were collected in the morning from the central market in Al-Shaikh. The samples were then transported to the **Yemen Standardization, Metrology, and Quality Control Organization** for analysis.
3. *Maintenance Media:* Bacterial isolates were preserved in a medium composed of 85% nutrient broth (NB) and 15% glycerol, stored at -20°C to ensure long-term viability.
4. *Gram Staining:* Gram staining was performed to differentiate *S. aureus* (Gram-positive) from *E. coli* (Gram-negative), facilitating initial identification. The procedure involved the following steps: Application of crystal violet dye. Fixation with Lugol's solution. Washing with alcohol. Counterstaining with safranin.
5. *Preparation of Culture Media:* General and selective media, including Blood Agar (BA), MacConkey Agar (MA), and (MSA), were prepared as follows: Media were sterilized via autoclaving to ensure aseptic conditions. Sterility was confirmed by incubating prepared media prior to use. Media were stored under sterile conditions for bacterial cultivation.
6. *Preparation of Fresh Vegetable and Fruit Samples:* Each sample (250 grams) was thoroughly washed with distilled water to remove surface debris. Samples were stored in sterile containers to prevent contamination. Tryptone Soy Broth (TSB) was sterilized by autoclaving and used as an enrichment medium. Serial dilutions of the samples were prepared and incubated at 37°C for 24 hours. Selective media, including MSA, (EMB), and Xylen Xylose Lysine Deoxycholate (XLD), were employed to isolate *S. aureus*, *E. coli*, and other Gram-negative bacteria [10].
7. *Biochemical Testing of Staphylococcus aureus and Escherichia coli:* Biochemical tests were performed to identify and characterize bacterial species based on their metabolic activities:
    a) Catalase Test, differentiates catalase-positive Staphylococcus species from catalase-negative Streptococcus species [11].
    b) Coagulase Test, identifies *S. aureus* by detecting coagulase enzyme activity, which causes plasma coagulation [11].
    c) Oxidase Test, detects cytochrome c oxidase activity, distinguishing oxidase-positive bacteria such as Pseudomonas sp., from oxidase-negative *E. coli* [11].
    d) Methyl Red (MR) Test, confirms stable acid production from glucose fermentation, a characteristic of *E. coli* [12].
    e) Indole Test, detects indole production from tryptophan hydrolysis, with a positive result confirming *E. coli* [13].
    f) Citrate Utilization Test, determines the ability to use citrate as a sole carbon source, distinguishing citrate-negative *E. coli* from other enteric bacteria [14].
    g) DNase Test, identifies DNase activity in *S. aureus*, indicated by a clear zone around the colony [15]. Capsule Staining, highlights bacterial capsules using copper sulfate, confirming capsule presence as a virulence factor [15].
    h) Urease Test, differentiates urease-positive Proteus mirabilis from urease-negative *E. coli* [13].
    i) Sugar Fermentation Test, evaluates glucose and fructose fermentation by *S. aureus* and *E. coli*, indicated by a yellow color change in the pH indicator [16].
    j) Kligler Iron Agar (KIA), assesses carbohydrate fermentation and hydrogen sulfide production, aiding in the identification of *E. coli* [17].

k) Voges–Proskauer (VP) Test, detects acetoin production from glucose fermentation, used to differentiate enteric bacteria [13].

B. Plant Sample Collection and Preparation

1. *Sample Collection*: Dried root samples of (turmeric, ginger and Indian costus) were collected from a herbalist in the Aden Crater area (**Al-Ekbar**). The plant materials were identified and authenticated by **Dr. Othman Al-Hawshabi**, Professor of Taxonomy of Flowering Plants and Flora, at the Biology Department, Faculty of Science, Aden University. After authentication, the roots were processed using traditional methods: Ground into fine powder. Purified to remove impurities. Combined into a homogeneous mixture for subsequent analysis.

2. *Extraction Procedure*: To maximize the range of isolated phytochemicals, a dual extraction method was employed:

   a) *Alcoholic Extraction:* Alcoholic extraction was performed using Soxhlet extraction with 70% ethanol to target alcohol-soluble components. Following established protocols [18, 19, 20]: 40 grams of dried powdered roots were subjected to Soxhlet extraction with 400 mL of 70% ethanol for 6–8 hours. The extract was filtered through sterile gauze and centrifuged, followed by filtration using Whatman No. 1 filter paper. The filtrate was dried in a Petri dish and stored at 4°C for later analysis.

   b) *Aqueous Extraction:* Aqueous extraction targeted water-soluble compounds using distilled water [18, 19], 40 grams of dried powdered roots were combined with 400 mL of distilled water and shaken on a magnetic stirrer at room temperature for 24 hours. The mixture was filtered, centrifuged, and dried in a laboratory oven at 50°C. The dried extract was stored at 4°C until required for further experiments.

3. *Phytochemical Screening:* Phytochemical screening identified bioactive compounds using standard colorimetric tests [21, 22]. Alkaloids were detected using Dragendorff's reagent, where a yellow or orange precipitate indicated their presence. Tannins were identified using 1% $FeCl_3$, resulting in a greenish-black coloration. Saponins were confirmed by the formation of persistent foam after shaking with water. Flavonoids were detected by a yellow color upon the addition of NaOH. Steroids and triterpenoids were identified by heating with H_2SO_4, where a red or pink coloration indicated steroids, while a reddish-brown ring confirmed the presence of triterpenoids.

4. *Physicochemical Analysis*: Determination of Extractive Value, the extractive value measures the quantity of phytochemicals soluble in a specific solvent, reflecting extraction efficiency:

$$Extractive\ value\ \% = \frac{Weight\ of\ the\ extract}{Weight\ of\ the\ air-\ dried\ drug} \times 100$$

Soxhlet extraction (70% ethanol) and distilled water were used for alcoholic and aqueous extracts, respectively [22].

5. *Loss on Drying (Moisture Content):* Moisture content impacts plant stability and preservation:

$$Moisture\ (\%) = \frac{Weight\ before\ dryin\ -\ Weight\ after\ drying\ Weight}{Weight\ of\ the\ sample} \times 100$$

Samples were dried in a pre-heated oven at 135°C, minimizing microbial contamination risks [23].
Ash Content Determination, ash content indicates the purity of plant materials:

$$Ash\ \% = \frac{Ash\ weight}{Weight\ of\ the\ sample} \times 100$$

Samples were heated in a muffle furnace at 500°C until complete combustion yielded white ash, reflecting minimal contamination [24]. Determination of pH, the pH was measured by mixing 5 grams of dried plant powder with 50 mL of distilled water, stirring for 10 minutes, filtering, and recording the pH with a calibrated pH meter [25].

III. Results

A. Total Number of Vegetable and Fruit Samples

The study focused on isolating *S. aureus* and *E. coli* from vegetable and fruit samples to evaluate the antimicrobial effects of aqueous and ethanol plant extracts. The results are summarized as follows in (**Table 1**): Vegetable Samples **(53):** 17 samples (32.07%) tested positive for *S. aureus*. 8 samples (15.09%) tested positive for *E. coli*. 28 samples (52.83%) were excluded. Fruit Samples **(21):** 3 samples (14.29%) tested positive for *S. aureus*. No *E. coli* was detected. 18 samples (85.71%) were excluded.

Table 1: Prevalence of Staphylococcus aureus and Escherichia coli in Vegetable and Fruit Samples

Sample Source	Vegetable Samples	Fruit Samples
Number of Samples	53 (100%)	21 (100%)
Positive Samples	25 (47.17%)	3 (14.29%)
Ignored Samples	28 (52.83%)	18 (85.71%)
S. aureus (Positive)	17 (32.07%)	3 (14.29%)
E. coli (Positive)	8 (15.09%)	0

B. Results of Bacterial Identification

The bacterial identification process confirmed the presence of *S. aureus* and *E. coli* in vegetable and fruit samples. Various growth media were used to isolate and differentiate the bacteria, as outlined below:

1. *Staphylococcus aureus*, (MSA), Yellow colonies indicating mannitol fermentation. Vogel & Johnson Agar (VOG), Black colonies due to tellurite reduction. Blood Agar, White colonies with beta hemolysis (clear zones around colonies). Nutrient Agar, Large white colonies.
2. *Escherichia coli*, (EMB) Agar, Purple to black colonies with a green metallic sheen, indicative of lactose fermentation. MacConkey Agar: Pink colonies with a darker halo due to lactose fermentation. (XLD) Agar: Circular yellow colonies, indicating acid production from xylose fermentation. Blood Agar: Gray colonies with partial hemolysis (alpha hemolysis).

C. Results of Biochemical Tests

Biochemical testing provided crucial insights into the physiological and metabolic characteristics of *S. aureus* and *E. coli* isolated from vegetable and fruit samples (**Table 2**). Capsule Test, *E. coli* was positive, indicating the presence of a polysaccharide capsule. *S. aureus* was negative. Catalase Test, both bacteria tested positive, confirming their ability to decompose hydrogen peroxide. Coagulase Test, *S. aureus* was positive, distinguishing it from other staphylococcal species. *E. coli* was negative. Hemolysis, *S. aureus* showed beta hemolysis (complete lysis of red blood cells). *E. coli* exhibited alpha hemolysis (partial lysis). Oxidase Test, both bacteria were negative, aligning with their facultative anaerobic nature. DNase Test, *S. aureus* was positive, indicating virulence. *E. coli* was negative. Gram Stain, *S. aureus* appeared as Gram-positive cocci (blue). *E. coli* appeared as Gram-negative rods (red). Citrate Utilization, *S. aureus* was positive, showing citrate usage as a carbon source. *E. coli* was negative. Indole Test, *E. coli* was positive, confirming indole production. *S. aureus* was negative. Voges-Proskauer (VP) Test, *S. aureus* tested positive for acetoin production. *E. coli* was negative. Methyl Red (MR) Test, both bacteria were positive, indicating strong acid production from glucose fermentation. Motility Test, *E. coli* was motile, whereas *S. aureus* was non-motile. Hydrogen Sulfide (H_2S) Production, both bacteria were negative for H_2S production.

Sugar Fermentation, both bacteria fermented glucose, but only *E. coli* produced gas. *S. aureus* fermented lactose, mannitol, and sucrose, while *E. coli* showed variability in sugar fermentation. Urease Test, both *S. aureus* and *E. coli* were negative.

Table 2. Biochemical Test Results for Identification of S. aureus and E. coli Isolated from Vegetables and Fruits

Biochemical Test	***Staphylococcus aureus***	***Escherichia coli***
Capsule Test	-	+
Catalase Test	+	+
Coagulase Test	+	-
Hemolysis	Beta (B)	Alpha (∝)
Oxidase Test	-	-
DNase Test	+	-
Gram Stain	Blue (Cocci)	Red (Rod)
Citrate Utilization	+	-
Indole Test	-	+
Voges-Proskauer (VP)	+	-
Methyl Red (MR) Test	+	+
Motility Test	-	+
Hydrogen Sulfide (H_2S)	-	-
Sugar Fermentation		
- Glucose	+	+ (with gas)
- Lactose	+	+
- Fructose	+	-
- Mannitol	+	+
- Xylose	-	+
- Sucrose	+	+
Urease Test	-	-

D. *Isolation of Staphylococcus aureus and Escherichia coli from Vegetables and Fruits*

The study aimed to isolate *S. aureus* and *E. coli* from various vegetable and fruit samples using a six-dilution method across three culture media: MSA, XLD, and EMB. The cultivation process was performed in Tryptone Water for 24 hours, and the procedure was repeated three times to enhance accuracy. The results showed variability in bacterial isolation across different samples and repetitions (**Table 3; 4**).

The isolation rates of *S. aureus* and *E. coli* varied significantly across vegetable types. **Cabbage** showed a *S. aureus* isolation rate of **19.04%** and the highest *E. coli* isolation rate among vegetables at **40.54%**, with undesirable bacteria at **4.21%**.

Lettuce had isolation rates of **16.67%** for *S. aureus* and **35.14%** for *E. coli*, with undesirable bacteria at **9.47%**. **Spinach** presented a moderate *S. aureus* isolation rate of **17.86%** and *E. coli* at **24.32%**, with undesirable bacteria at **12.63%**. **Beetroot** and **potatoes** had the highest *S. aureus* isolation rates at **20.24%**, but *E. coli* was not detected in either sample (0%), while undesirable bacteria were found at **20%**. **Carrots** had the lowest *S. aureus* isolation rate at **5.95%**, with no *E. coli* detected and the highest level of undesirable bacteria at **32.63%**. Fruit samples showed a different contamination pattern. **Watermelon** had the highest *S. aureus* isolation rate at **75%**, with no *E. coli* detected and undesirable bacteria at **30%**. **Grapes** had a *S. aureus* isolation rate of **25%**, with *E. coli* absent and undesirable bacteria at **34%**. **Strawberries** showed no contamination by *S. aureus* or *E. coli*, but undesirable bacteria were present at **36%**, the highest among fruits.

Table 3: Percentage of *Staphylococcus aureus* and *Escherichia coli* in Vegetable Samples

Total count of vegetable samples for all repeats	Type of vegetable sample	*Staphylococcus aureus*	*Escherichia coli*	Ignored bacteria
18 samples of vegetable	Cabbage	19.04%	40.54%	4.21%
	Lettuce	16.67%	35.14%	9.47%
	Spinach	17.86%	24.32%	12.63%
	Beets	20.24%	0	20%
	Carrots	5.95%	0	32.63%
	Potatoes	20.24%	0	20%

Table 4: Percentage of *Staphylococcus aureus* and *Escherichia coli* in Fruit Samples

Total count of fruit samples for all repeats	Type of fruit sample	*Staphylococcus aureus*	*Escherichia coli*	Ignored bacteria
9 samples of fruit	Grapes	25%	0	34%
	Strawberry	0	0	36%
	Watermelon	75%	0	30%

Prevalence of Bacteria in Vegetables. *S. aureus*: The highest prevalence was observed in beetroot, potatoes, and cabbage (~20%), while the lowest detection rate was in carrots (5.95%). This suggests potential contamination due to post-harvest handling practices. *E. coli*: The highest prevalence was recorded in cabbage (40.45%) and lettuce (35.14%), whereas no detection was observed in beetroot, carrots, or potatoes. This absence may be attributed to water contamination sources or inadequate washing procedures. Prevalence of Bacteria in Fruits. *S. aureus*: The highest prevalence was found in watermelon (75%), followed by grapes (25%). No detection was observed in strawberries, indicating better hygienic conditions or natural resistance to bacterial contamination (**Table 5**).

Table 5: Prevalence of *S. aureus* and *E. coli* in Vegetables and Fruits

Sample Type	*S. aureus* (%)	*E. coli* (%)	Other Bacteria (%)
Vegetables	19.04–20.24	24.32–40.54	4.21–32.63
Fruits	25–75	0	30–36

E. Mean Count of S. aureus in Vegetables

The mean count of ***S. aureus*** across dilutions showed a decreasing trend, with the highest mean at 10^{-1} (0.944), progressively declining to **0.389** at 10^{-6}. Significant differences were observed among vegetable types (Table 6). Specifically, beetroot and potatoes exhibited the highest mean counts (0.944), followed by cabbage (0.889) and spinach (0.833). The lowest mean count was found in carrots (0.278), indicating lower susceptibility to contamination.

Table 6: Mean Counts of *S. aureus* in Vegetables

Vegetables	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	Mean
Cabbage	1.000	1.000	1.000	1.000	1.000	0.333	0.889
Lettuce	1.000	1.000	1.000	0.667	0.667	0.333	0.778
Spinach	1.000	1.000	1.000	1.000	0.667	0.333	0.833
Beetroot	1.000	1.000	1.000	1.000	1.000	0.667	0.944
Carrots	0.667	0.667	0.333	0.000	0.000	0.000	0.278
Potatoes	1.000	1.000	1.000	1.000	1.000	0.667	0.944

F. Mean Count of Escherichia coli in Vegetables

The mean count of E. coli across dilutions remained stable at 0.444 up to dilution 10^{-4}, then significantly declined to 0.056 at 10^{-6} (**Table 7**). Vegetable-specific findings revealed the highest mean counts in cabbage (0.833) and lettuce (0.722), with moderate counts in spinach (0.500). No detectable E. coli was found in beetroot, carrots, or potatoes.

Table 7: Mean Counts of E. coli in Vegetables

Vegetables	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	Mean
Cabbage	1.000	1.000	1.000	1.000	0.667	0.333	0.833
Lettuce	1.000	1.000	1.000	1.000	0.333	0.000	0.722
Spinach	0.667	0.667	0.667	0.667	0.333	0.000	0.500
Beets	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Carrots	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Potatoes	0.000	0.000	0.000	0.000	0.000	0.000	0.000

G. Mean Count of S. aureus in Fruits

Trends Across Dilutions: The highest mean count was observed at dilutions **10^{-1}**, **10^{-2}**, **10^{-4}**, and **10^{-5}** in watermelon (**0.333**), then decreased to **0.000** from **10^{-3}** onward. Fruit-Specific Findings: Watermelon exhibited the highest mean count (**0.333**), followed by grapes (**0.111**). No detectable *S. aureus* was found in strawberries (**Table 8**).

Table 8: Mean Counts of *S. aureus* in Fruits

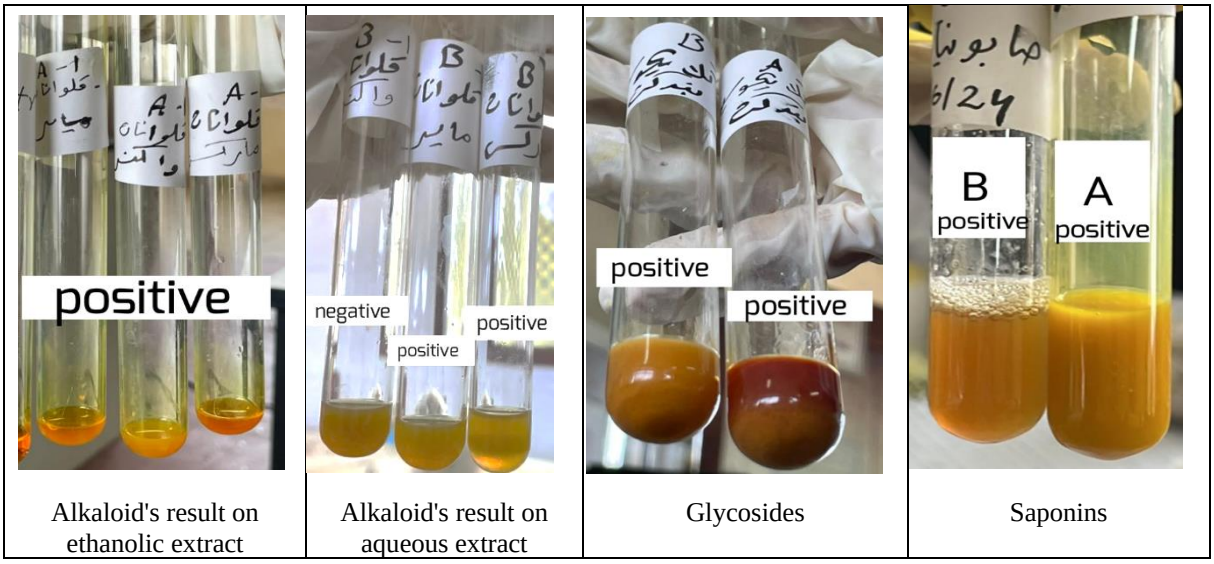
Fruits	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	Mean
Grapes	0.333	0.333	0.000	0.000	0.000	0.000	0.111
Strawberries	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Watermelon	0.667	0.667	0.333	0.333	0.000	0.000	0.333

H. Phytochemical Analysis of Ethanolic and Aqueous Extracts

Ethanolic extract (**Table 9; Fig. 1**), showed a broad range of bioactive compounds, including alkaloids, flavonoids, terpenoids, glycosides, tannins, resins, and carbohydrates. Amino acids and fucoidans were absent, indicating ethanol's inefficiency in extracting these polar compounds. Aqueous extract, contained polar compounds such as fucoidans and amino acids, which were absent in the ethanolic extract. Showed presence of glycosides, flavonoids, tannins, terpenoids, and carbohydrates, albeit at possibly lower concentrations.

Table 9: Phytochemical composition of ethanolic and aqueous extracts

Phytochemical	Ethanolic Extract	Aqueous Extract
Alkaloids	Positive (brown precipitate)	Positive (no brown precipitate)
Glycosides	Present	Present
Saponins	Present	Present
Flavonoids	Present	Present
Flavanones	Present	Present
Tannins	Present	Present
Resins	Present	Present
Fucoidans	Absent	Present
Terpenoids	Present	Present
Amino Acids	Absent	Present
Carbohydrates	Present	Present



Alkaloid's result on ethanolic extract | Alkaloid's result on aqueous extract | Glycosides | Saponins

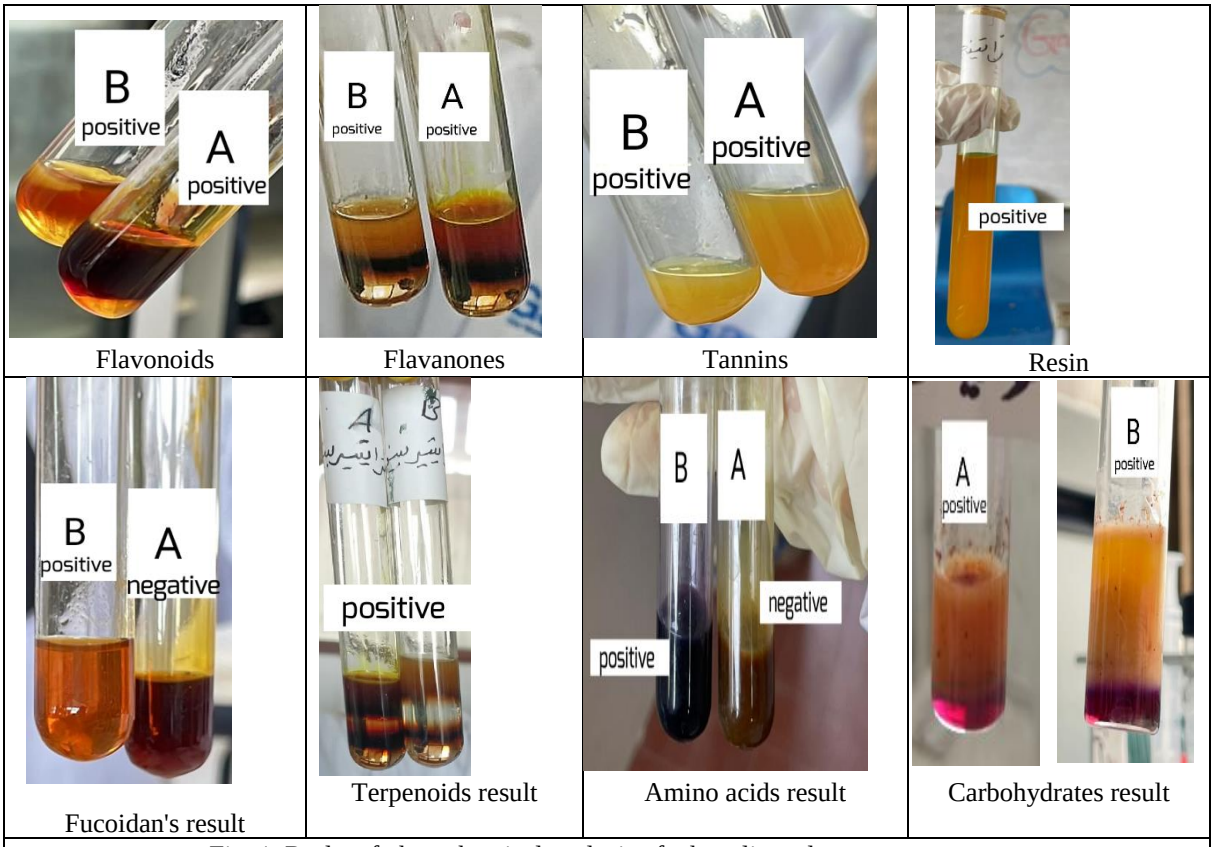


Fig. 1. Reslts of phytochemical analysis of ethanolic and aqueous extracts.
A= Ethanolic extract, B= Aqueous extract

I. Physiochemical Properties of Plant Extracts

Total Ash Content measured value: 5.78%. Indicates moderate mineral content and purity, suitable for therapeutic applications. Moisture content, measured value, 10.76%. Moderate moisture reduces microbial growth and preserves bioactive compounds. Extractive values, aqueous extract, 15.96%. Ethanolic extract, 14.87%. Highlights water's efficiency in extracting polar compounds and ethanol's ability to isolate non-polar compounds. pH values, aqueous extract, 5.113 (slightly acidic). Ethanolic extract, 5.37 (slightly acidic). Acidity preserves bioactive stability, beneficial for pharmaceutical formulations.

Table 10: Summary of Physiochemical Properties

Parameter	Value	Significance
Total Ash Content	5.78%	Indicates mineral content and purity.
Moisture Content	10.76%	Ensures stability and prevents microbial growth.
Aqueous Extract Value	15.96%	Efficient extraction of polar compounds.
Ethanolic Extract Value	14.87%	Efficient extraction of non-polar compounds.
Aqueous Extract pH	5.113	Suitable for preserving bioactive.
Ethanolic Extract pH	5.37	Suitable for stable formulations.

IV. Discussion

A. Importance of the Study

This study contributes to improving food safety in Yemen by exploring natural antimicrobial alternatives to combat bacterial contamination in vegetables and fruits. In a country facing challenging agricultural conditions, including limited access to clean water and proper sanitation, the findings emphasize the potential of plant-based extracts to reduce the presence of harmful pathogens like *S. aureus* and *E. coli* [26]. By utilizing locally available medicinal plants, this research offers a sustainable, cost-effective solution to enhance food safety, reduce reliance on synthetic antibiotics, and mitigate the public health risks associated with foodborne illnesses.

B. Bacterial Isolation and Identification

Mannitol Salt Agar effectively isolated *S. aureus*, as evidenced by the red-to-yellow color change due to mannitol fermentation. Similarly, (EMB) agar successfully isolated *E. coli*, with its characteristic black colonies and green metallic sheen. These results underscore the importance of selecting appropriate media for bacterial isolation, as MSA and EMB are reliable for *S. aureus* and *E. coli*, respectively [27, 28].

Biochemical tests revealed distinct metabolic characteristics of *S. aureus* and *E. coli*, aiding in accurate identification. Both bacteria tested positive for catalase, indicating their ability to survive in oxygen-rich

environments, while *S. aureus* was differentiated by the coagulase test, marking it as a pathogenic species [29]. The DNase test confirmed *S. aureus's* virulence [30], and sugar fermentation tests highlighted the metabolic versatility of both bacteria, with *E. coli* demonstrating motility and gas production [31]. The profiling of *E. coli* further confirmed its metabolic traits, including positive results for catalase, indole, and lactose fermentation [32], while *S. aureus* exhibited beta-hemolysis, indicating its pathogenic potential.

C. Contamination Levels in Vegetables and Fruits:

The study revealed significant variations in bacterial contamination across different produce types. Cabbage exhibited the highest *E. coli* contamination (40.54%), exceeding global averages reported in similar studies [33], while lettuce showed unexpectedly high *E. coli* levels (35.14%) compared to earlier findings [34]. In contrast, carrots and strawberries demonstrated lower contamination rates, likely due to differences in cultivation practices or surface morphology. Notably, *S. aureus* was detected in potatoes and watermelon, suggesting post-harvest handling as a critical contamination source [35, 36].

These discrepancies highlight the role of agricultural and environmental factors in shaping contamination risks. For instance, the use of untreated irrigation water—a common practice in regions like Yemen and Bangladesh [37, 38]—likely contributed to the elevated *E. coli* levels in leafy greens. Furthermore, Yemen's hot climate may accelerate bacterial proliferation, particularly when combined with poor storage conditions [39].

The WHO estimates that nearly 10% of the global population suffers from foodborne illnesses annually [40], with developing nations disproportionately affected due to inadequate sanitation infrastructure [41]. This aligns with our findings, where contamination rates in Yemeni produce mirror trends observed in other resource-limited settings.

We propose that, adopting treated irrigation water and organic fertilizers to minimize pathogen introduction [42, 43]. Training farmers on post-harvest hygiene protocols, including tool sanitation and temperature-controlled storage [44]. Implementing national food safety standards aligned with WHO guidelines to enhance export competitiveness [45].

D. Phytochemical Analysis of Plant Extracts

Phytochemical analysis of (turmeric, ginger, and Indian costus) extracts revealed various bioactive compounds, with solubility influenced by the solvent used. Ethanolic extracts were rich in alkaloids and glycosides, known for their pharmacological properties, while aqueous extracts effectively extracted polar compounds like amino acids. Phytochemical tests confirmed the presence of flavonoids, glycosides, saponins, tannins, carbohydrates, terpenoids, and resins in both extract types. Notably, alkaloids were present in the ethanolic extracts, while the aqueous extract was negative in Wagner's test, highlighting the solvent's role in alkaloid extraction. Amino acids were exclusive to aqueous extracts, confirming their water-soluble nature. Glycosides, present in both extracts, are beneficial for cardiovascular health, such as lowering blood pressure. Our results are consistent with previous studies identifying flavonoids, glycosides, and terpenoids in these plants [46].

Aqueous extracts showed higher extractive values, indicating a greater concentration of water-soluble compounds, including vitamins and glycosides, which are valuable for therapeutic applications targeting such compounds [47]. Ethanolic extracts, although lower in extractive value, were rich in secondary metabolites like alkaloids and flavonoids, responsible for distinct pharmacological effects [48]. Combining water and ethanol as extraction solvents can provide a more comprehensive profile of active compounds, enhancing therapeutic potential [49].

E. Physiochemical Properties of Extracts

The ash content of 5.78% indicates moderate mineral content and purity, aligning with acceptable standards [50]. The moisture content of 10.76% supports proper storage and handling while minimizing microbial growth [51]. The extractive values of 15.96% for aqueous extract and 14.87% for ethanolic extract highlight the efficiency of water in extracting polar compounds and ethanol in isolating non-polar compounds [52, 53]. Both extracts showed mildly acidic pH values, suitable for preserving bioactive stability [54].

F. Public Health and Economic Implications

Plant extracts can significantly improve food safety by reducing bacterial contamination in vegetables and fruits, offering a natural alternative to chemical preservatives. This approach not only enhances the quality and safety of agricultural products but also reduces the risk of foodborne illnesses, which are a major public health concern, especially in regions with limited access to clean water and proper sanitation [55].

Enhancing food safety yields significant economic benefits by reducing losses from bacterial contamination [56]. Implementing improved hygiene practices, along with safer storage and handling methods, minimizes spoilage of agricultural products, thereby reducing waste and extending the shelf life of fresh produce, which increases its market value. Additionally, improving food safety can enhance the competitiveness of agricultural exports [45]. Products that meet international safety standards are more likely to be accepted in global markets, leading to higher demand and better prices for exporters. Furthermore, reducing foodborne illnesses can lower healthcare costs and prevent productivity losses, contributing to overall economic stability and growth [57].

G. Limitations and Future Directions

One limitation of the current study is the relatively small sample size, which may affect the generalizability of the results. A larger sample size could provide more robust data and a clearer understanding of bacterial contamination in various agricultural products. Additionally, this study focused only on *S. aureus* and *E. coli*, and did not explore other pathogenic bacteria that may also pose risks to food safety, such as *Salmonella*, *Klebsiella* spp., or *Listeria*. Future studies should consider expanding the scope by including a

broader range of bacterial species and increasing the sample size to enhance the validity of the findings. This would help provide a more comprehensive understanding of bacterial contamination in produce and contribute to more effective food safety strategies.

V. Conclusion

This study highlights the potential of plant extracts (turmeric, ginger, and Indian costus) as natural antimicrobial agents to reduce bacterial contamination in Yemeni vegetables and fruits. Key pathogens (*S. aureus* and *E. coli*) were isolated using selective media (MSA and EMB) and differentiated via biochemical tests (e.g., catalase, coagulase). Contamination levels varied significantly, with leafy greens (cabbage, lettuce) showing the highest *E. coli* rates (40.54% and 35.14%), likely due to poor irrigation and handling practices. Phytochemical analysis revealed solvent-dependent extraction efficiencies: ethanolic extracts yielded alkaloids/flavonoids, while aqueous extracts contained polar compounds like amino acids. The extracts' mild acidity (pH 5.1–5.4) and stability support their use as safe, sustainable alternatives to synthetic preservatives. By mitigating contamination, these extracts can enhance food safety, reduce economic losses from spoilage, and improve Yemen's agricultural competitiveness, offering a scalable solution for regions facing similar challenges.

Acknowledgments

We sincerely thank **Prof. Dr. Othman Al-Houshabi** for his valuable collaboration in identifying the plant species and for his insightful contributions in refining certain formulations in our research. His expertise and guidance have greatly enriched our study.

References

[1] D. C. Zaharia, A. A. Muntean, M. G. Popa, A. T. Steriade, O. Balint, R. Micut, C. Iftene, I. Tofolean, V. T. Popa, C. Baicus, M. A. Bogdan, and M. I. Popa, "Comparative analysis of *Staphylococcus aureus* and *Escherichia coli* microcalorimetric growth," *BMC Microbiol.*, vol. 13, p. 171, 2013, doi: 10.1186/1471-2180-13-171.

[2] Centers for Disease Control and Prevention (CDC), "*Staphylococcus aureus* basics," [Online]. Available: https://www.cdc.gov/staphylococcus-aureus/about/index.html. [Accessed: 2024].

[3] T. A. Taylor and C. G. Unakal, "*Staphylococcus Aureus*," *StatPearls [Internet]*, Treasure Island (FL): StatPearls Publishing; 2024. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK441868/

[4] A. M. Algammal, H. F. Hetta, A. Elkelish, D. H. H. Alkhalifah, W. N. Hozzein, G. E. Batiha, N. El Nahhas, and M. A. Mabrok, "Methicillin-Resistant *Staphylococcus aureus* (MRSA): One Health Perspective Approach to the Bacterium Epidemiology, Virulence Factors, Antibiotic-Resistance, and Zoonotic Impact," *Infect. Drug Resist.*, vol. 13, pp. 3255–3265, 2020, doi: 10.2147/IDR.S272733.

[5] B. Pakbin, W. M. Brück, and J. W. Rossen, "Virulence Factors of Enteric Pathogenic *Escherichia coli*: A Review," *Int. J. Mol. Sci.*, vol. 22, no. 18, p. 9922, 2021, doi: 10.3390/ijms22189922.

[6] J. Kadariya, T. C. Smith, and D. Thapaliya, "*Staphylococcus aureus* and staphylococcal food-borne disease: an ongoing challenge in public health," *Biomed. Res. Int.*, vol. 2014, p. 827965, 2014, doi: 10.1155/2014/827965.

[7] S. M. Gambushe, O. T. Zishiri, and M. E. El Zowalaty, "Review of *Escherichia coli* O157:H7 Prevalence, Pathogenicity, Heavy Metal and Antimicrobial Resistance, African Perspective," *Infect. Drug Resist.*, vol. 15, pp. 4645–4673, 2022, doi: 10.2147/IDR.S365269.

[8] K. Nair, *Turmeric (Curcuma longa L.) and Ginger (Zingiber officinale Rosc.) - World's Invaluable Medicinal Spices: The Agronomy and Economy of Turmeric and Ginger*, 2019, doi: 10.1007/978-3-030-29189-1.

[9] Encyclopaedia Britannica, "Aden," *Britannica*, 25 Mar. 2023. [Online]. Available: https://www.britannica.com/place/Aden. [Accessed: 24 May 2023].

[10] A. O. Eni, L. A. Oluwawemitan, and O. U. Solomon, "Microbial quality of fruits and vegetables sold in Sango Ota, Nigeria," *Afr. J. Food Sci.*, vol. 4, no. 5, pp. 291–296, 2010.

[11] B. M. Al-Joda and A. H. Jasim, "Studying the Efficiency Dependence of CR-39 Detector on the Chamber Height," *J. Univ. Babylon Pure Appl. Sci.*, vol. 29, no. 2, pp. 68–176, 2021.

[12] S. Aryal, "Methyl Red (MR) Test - Principle, Procedure and Result Interpretation," *Microbiology Info*, 2018. [Online]. Available: https://microbiologyinfo.com/methyl-red-mr-test-principle-procedure-and-result-interpretation/

[13] J. F. MacFaddin, *Biochemical Tests for Identification of Medical Bacteria*, 3rd ed., Philadelphia: Lippincott Williams & Wilkins, 2000.

[14] C. R. Mahon, D. C. Lehman, and G. Manuselis, *Textbook of Diagnostic Microbiology*, 5th ed., Elsevier, 2015.

[15] W. C. Winn, *Koneman's Color Atlas and Textbook of Diagnostic Microbiology*, Lippincott Williams & Wilkins, 2006.

[16] S. Aryal, "*Staphylococcus aureus* - An Overview," *Microbe Notes*, 2022. [Online]. Available: https://microbenotes.com/staphylococcus-aureus-an-overview/

[17] P. Tille, *Bailey & Scott's Diagnostic Microbiology*, Elsevier Health Sciences, 2017.

[18] J. B. Harborne, *Phytochemical Methods: A Guide to Modern Techniques of Plant Analysis*, 2nd ed., London and New York: Chapman and Hall, 1984.

[19] S. S. Handa, P. S. Sumnan, L. Gennaro, and D. R. Dev, *Extraction Technologies for Medicinal and Aromatic Plants*, International Centre for Science and High Technology, UNIDO, 2008.

[20] K. Moothy, T. Punith, R. Vinodhini, and N. Thajuddin, "Antimicrobial activity and qualitative phytochemical analysis of *Punica granatum* Linn (Pericarp)," *J. Med. Plants Res.*, vol. 7, no. 9, pp. 474–479, 2012.

[21] N. R. Farnsworth, "Biological and Phytochemical Screening of Plants," *J. Pharm. Sci.*, vol. 55, no. 3, pp. 225–276, 1996, doi: 10.1002/jps.2600550302.
[22] H. O. Edeoga, D. E. Okwu, and B. O. Mbaebie, "Phytochemical constituents of some Nigerian medicinal plants," *Afr. J. Biotechnol.*, vol. 4, no. 7, pp. 685–688, 2005.
[23] G. Nasser and H. Ahmed, "Study of Chemical Composition, Antioxidant and Antimicrobial Properties of Two Selected Medicinal Plants," M.Sc. thesis, Univ. of Aden, Yemen, 2022, p. 157.
[24] Anonymous, *Pharmacopoeia of India*, Ministry of Health, Govt. of India, New Delhi, 1996.
[25] E. H. Mansour and A. H. Khalil, "Evaluation of antioxidant activity of some plant extracts and their application to ground beef patties," *Food Chem.*, vol. 69, no. 2, pp. 135–141, 2000.
[26] C. Ingrao, R. Strippoli, G. Lagioia, and D. Huisingh, "Water scarcity in agriculture: An overview of causes, impacts and approaches for reducing the risks," *Heliyon*, vol. 9, pp. e14194, 2023.
[27] S. E. Sharp and C. Searcy, "Comparison of mannitol salt agar and blood agar plates for identification and susceptibility testing of *Staphylococcus aureus* in specimens from cystic fibrosis patients," *J. Clin. Microbiol.*, vol. 44, no. 12, pp. 4420–4423, 2006.
[28] D. J. Leininger, J. R. Roberson, and F. Elvinger, "Use of eosin methylene blue agar to differentiate *Escherichia coli* from other gram-negative mastitis pathogens," *J. Vet. Diagn. Invest.*, vol. 13, no. 3, pp. 273–275, 2001.
[29] T. Foster, "Staphylococcus," in *Medical Microbiology*, 4th ed., S. Baron, Ed. Galveston, TX: Univ. of Texas Medical Branch, 1996, ch. 12.
[30] D. P. Kateete, C. N. Kimani, F. A. Katabazi, A. Okeng, M. S. Okee, A. Nanteza, M. L. Joloba, and F. C. Najjuka, "Identification of *Staphylococcus aureus*: DNase and mannitol salt agar improve the efficiency of the tube coagulase test," *Ann. Clin. Microbiol. Antimicrob.*, vol. 9, p. 23, 2010, doi: 10.1186/1476-0711-9-23.
[31] N. Sultana, D. Mridha, R. Sohel, M. Mahmud, M. Shahid, and H. Ahosanul, "Isolation, identification and antibiogram of verotoxin producing *Escherichia coli* from raw salad vegetables at Jashore, Bangladesh," *Afr. J. Microbiol. Res.*, vol. 15, pp. 401–407, 2021.
[32] P. Singh and A. Prakash, "Isolation of *Escherichia coli*, *Staphylococcus aureus*, and *Listeria monocytogenes* from milk products sold under market conditions at Agra Region," *Acta Agric. Slov.*, vol. 92, no. 1, pp. 83–88, 2008.
[33] B. V. Maikai and D. O. Akubo, "Coliform count and isolation of *Escherichia coli* in fresh fruits and vegetables sold at retail outlets in Samaru, Kaduna State, Nigeria," *Niger. Vet. J.*, vol. 39, no. 4, pp. 327–337, 2018.
[34] K. Holvoet, I. Sampers, B. Callens, J. Dewulf, and M. Uyttendaele, "Moderate prevalence of antimicrobial resistance in *Escherichia coli* isolates from lettuce, irrigation water, and soil," *Appl. Environ. Microbiol.*, vol. 79, no. 21, pp. 6677–6683, 2013.
[35] D. Devi and R. Gogoi, "Isolation and identification of microorganisms from some spoiled fruits and vegetables collected from Nagaon, Assam," *Plant Archives*, vol. 21, no. 1, pp. 1779–1782, 2021. [Online]. Available: https://doi.org/10.51470/PLANTARCHIVES.2021.v21.no1.242
[36] P. C. Obinna-Echem and L. Koanyie, "Quality and safety characteristics of minimally processed watermelon (*Citrullus lanatus*) sold in selected locations in Port Harcourt," *Eur. J. Agric. Food Sci.*, vol. 2, no. 3, pp. 1–6, 2020. [Online]. Available: https://doi.org/10.24018/ejfood.2020.2.3.57
[37] P. Mead, L. Slutsker, V. Dietz, L. Mccaig, J. Bresee, C. Shapiro, P. Griffin, and R. Tauxe, "Food-related illness and death in the United States," *Emerg. Infect. Dis.*, vol. 5, pp. 607–625, 1999, doi: 10.3201/eid0505.990502.
[38] A. N. Angelakis, G. Tchobanoglous, A. G. Capodaglio, and V. A. Tzanakakis, "The importance of nonconventional water resources under water scarcity," *Water*, vol. 16, no. 7, p. 1015, 2024, doi: 10.3390/w16071015.
[39] S. Bolan, L. P. Padhye, T. Jasemizad, M. Govarthanan, N. Karmegam, H. Wijesekara, D. Amarasiri, D. Hou, P. Zhou, B. K. Biswal, R. Balasubramanian, H. Wang, K. H. M. Siddique, J. Rinklebe, M. B. Kirkham, and N. Bolan, "Impacts of climate change on the fate of contaminants through extreme weather events," *Sci. Total Environ.*, vol. 909, p. 168388, Jan. 2024, doi: 10.1016/j.scitotenv.2023.168388.
[40] World Health Organization, "Food safety," 2024. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/food-safety.
[41] T. Bintsis, "Foodborne pathogens," *AIMS Microbiology*, vol. 3, no. 3, pp. 529–563, 2017, doi: 10.3934/microbiol.2017.3.529.
[42] S. Mishra, R. Kumar, and M. Kumar, "Use of treated sewage or wastewater as an irrigation water for agricultural purposes—Environmental, health, and economic impacts," *Total Environment Research Themes*, vol. 6, p. 100051, 2023, doi: 10.1016/j.totert.2023.100051.
[43] A. Sharma and R. Chetani, "A review on the effect of organic and chemical fertilizers on plants," *Int. J. Appl. Sci. Eng.*, vol. 5, pp. 677–680, 2020.
[44] H. Chen, A. J. Kinchla, N. Richard, A. Shaw, and Y. Feng, "Produce growers' on-farm food safety education: A review," *J. Food Prot.*, vol. 84, no. 4, pp. 704–716, 2021, doi: 10.4315/JFP-20-320.
[45] S. Henson and S. Jaffee, "Food safety standards and trade: Enhancing competitiveness and avoiding exclusion of developing countries," *Eur. J. Dev. Res.*, vol. 18, no. 4, pp. 593–607, 2006. [Online]. Available: https://doi.org/10.1080/09578810601070753
[46] E. M. Abdallah, K. A. Qureshi, A. M. Ali, and G. O. Elhassan, "Evaluation of some biological properties of *Saussurea costus* crude root extract," *Biosci. Biotechnol. Res. Commun.*, vol. 10, no. 4, pp. 601–611, 2017, doi: 10.21786/bbrc/10.4/2.
[47] S. S. Pawar and M. G. Jadhav, "Determination of extractive value and phytochemical analysis of *Bacopa*

monnieri (L)," 2016. [Online]. Available: https://www.researchgate.net/publication/296701173

[48] V. Varadharajan, U. K. Janarthanan, and V. Krishnamurthy, "Physicochemical, phytochemical screening and profiling of secondary metabolites of *Annona squamosa* leaf extract," *World J. Pharm. Res.*, vol. 1, no. 4, pp. 1143–1164, 2012.

[49] M. Jacotet-Navarro et al., "What is the best ethanol-water ratio for the extraction of antioxidants from rosemary? Impact of the solvent on yield, composition, and activity of the extracts," *Electrophoresis*, vol. 0, pp. 1–11, 2018.

[50] M. A. Buternio, C. Cavaliere, A. L. Capriotti, and C. Crescenzi, "Extending the applicability of pressurized hot water extraction to compounds exhibiting limited water solubility by pH control: Curcumin from the turmeric rhizome," *Anal. Bioanal. Chem.*, vol. 401, no. 9, pp. 2977–2985, 2011.

[51] A. K. Sachan and A. Kuma, "Stability testing of herbal products," *J. Chem. Pharm. Res.*, vol. 7, no. 12, pp. 511–514, 2015.

[52] M. C. Ali et al., "Effective extraction of flavonoids from *Lycium barbarum* L. fruits by deep eutectic solvents-based ultrasound-assisted extraction," *Talanta*, 2019. [Online]. Available: https://doi.org/10.1016/j.talanta.2019.05.012

[53] N. Zahed, R. Esmaeilzadeh Kenari, and R. Farahmandfar, "Effect of different extraction methods on antioxidant properties and encapsulation efficiency of anthocyanin of pomegranate peel," *Food Sci. Nutr.*, vol. 11, no. 7, pp. 3780–3787, 2023. [Online]. Available: https://doi.org/10.1002/fsn3.3362

[54] O. R. Alara, N. H. Abdurahman, and O. A. Olalere, "Ethanolic extraction of flavonoids, phenolics and antioxidants from *Vernonia amygdalina* leaf using two-level factorial design," *J. King Saud Univ. Sci.*, vol. 32, no. 1, pp. 7–16, 2020. [Online]. Available: https://doi.org/10.1016/j.jksus.2017.08.001

[55] K. Pacyga, D. Ronga, and I. Michalak, "Plant extracts—Importance in sustainable agriculture," *Ital. J. Agron.*, vol. 16, 2021, doi: 10.4081/ija.2021.1851.

[56] M. Focker and H. J. van der Fels-Klerx, "Economics applied to food safety," *Curr. Opin. Food Sci.*, vol. 36, pp. 18–23, 2020. [Online]. Available: https://doi.org/10.1016/j.cofs.2020.10.018

[57] PathSpot Technologies, "Economic impact of foodborne illnesses in the food industry and how to prevent them," 2024. [Online]. Available: https://www.pathspot.com.

Studying and Identifying the Challenges of Implementing Artificial Intelligence Systems in Yemen Based on Content Analysis and Fuzzy Cognitive Mapping

K. A. Al-Masouri

Received: 29/03/2025
Revised: 28/04/2025
Accepted: 29/04/2025

© 2025 University of Science and Technology, Aden, Yemen. This article can be distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

© 2025 جامعة العلوم والتكنولوجيا، المركز الرئيس عدن، اليمن. يمكن إعادة استخدام المادة المنشورة حسب رخصة مؤسسة المشاع الإبداعي شريطة الاستشهاد بالمؤلف والمجلة.

College of Computer and Information Technology, Saba University, Republic of Yemen
*Corresponding Author's Email: khalidahmed2009@gmail.com

Studying and Identifying the Challenges of Implementing Artificial Intelligence Systems in Yemen Based on Content Analysis and Fuzzy Cognitive Mapping

Khaled Ahmed Al-Masouri
College of Computer and Information Technology,
Saba University,
Republic of Yemen,
khalidahmed2009@gmail.com

***Abstract*—** Today, due to the expansion of knowledge and the increasing complexity of decision-making, the use of information systems, especially artificial intelligence (AI), has become increasingly important. Advanced and emerging technologies have undoubtedly played an undeniable role in the growing development of technological advancements across all industries. Most countries that have succeeded in achieving superior technologies have been able to transform ideas and knowledge into practical, high-selling products by adopting appropriate AI-related policies and strategies. Therefore, in this article, we address the significant challenges facing AI development in the Republic of Yemen. After identifying the challenges using content analysis and interview tools, we prioritized these challenges and the intensity of the relationship between them using fuzzy knowledge mapping. The statistical population for this study was selected from experts and specialists in the field of information and communications technology, particularly those specializing in artificial intelligence. In this study, 36 central challenges to AI development in the country were first identified based on existing strategic documents and content analysis. Then, based on four primary impact matrices—the fuzzy impact matrix, the impact strength matrix, and the final impact matrix—and using FCMapper and Pajek, fuzzy knowledge maps were drawn. The research results show that the five main challenges facing AI development in the country include unwillingness to invest economically in AI, lack of mechanisms for AI market growth and development, lack of necessary AI laws and regulations, lack of specialists and experts in AI-related fields, and lack of data exchange platforms.

Keywords— Artificial intelligence, challenges, Yemen, content analysis, fuzzy cognitive maps, information and communications technology.

the challenges facing AI development (Rao & Verweij, 2017).

Yemen, like other countries, should strive to develop an AI development program. Therefore, it is necessary to prioritize AI development challenges. In this regard, this research aims to answer the fundamental question, taking into account the specific characteristics of our country: What are the main challenges in AI development in the Republic of Yemen? Challenges that, if appropriate measures are not taken to address them, AI development will not be achieved. Therefore, in the next article, while listing the challenges facing AI development in the country, we will also develop a system for linking and prioritizing these challenges. In fact, in this research, we sought to organize these challenges. The 36 challenges were extracted from expert opinions without discovering the relationships between them. Therefore, in this step, we sought to identify the connections between these challenges using fuzzy cognitive maps. By organizing the extracted factors, the fuzzy cognitive maps, on the one hand, sorted these factors and, on the other hand, identified the relationships between the challenges. In fact, this process relies on four primary impact matrices: a fuzzy impact matrix, an impact strength matrix, and a final impact matrix.

Identifying the connections between challenges helps policymakers and decision-makers understand which other challenges might change if one challenge is resolved and whether this change and the changes occurring in the other challenges are related to the other challenges. Is this beneficial to the system?. Fuzzy cognitive mapping helps us determine the intensity of these relationships and the impacts and implications of these challenges on each other.

Therefore, the purpose of the current study is to prioritize AI development challenges to help IT and AI leaders and managers, based on variables such as financial and human resource constraints, business, time, etc., address the challenges that will be of greater importance and wider impact, as multiple constraints do not allow for solving all challenges at once. The next section reviews the relevant literature. The research methodology is then presented. Data analysis and research findings are then examined in detail. The final section provides discussion, conclusions, and future recommendations.

الملخص

تمكنت معظم الدول التي نجحت في تحقيق تقنيات متفوقة من تحويل الأفكار والمعرفة إلى منتجات عملية وعالية البيع من خلال اعتماد سياسات واستراتيجيات مناسبة تتعلق بالذكاء الاصطناعي. لذلك، نتناول في هذه المقالة التحديات الكبيرة التي تواجه تطوير الذكاء الاصطناعي في الجمهورية اليمنية. وبعد تحديد التحديات باستخدام تحليل المحتوى وأدوات المقابلة، حددنا أولويات هذه التحديات وشدة العلاقة بينها باستخدام رسم الخرائط المعرفية الضبابية. وقد تم اختيار المجتمع الإحصائي لهذه الدراسة من الخبراء والمتخصصين في مجال تكنولوجيا المعلومات والاتصالات، وخاصة المتخصصين في الذكاء الاصطناعي. وفي هذه الدراسة، تم تحديد 36 تحديًا رئيسيًا لتطوير الذكاء الاصطناعي في البلاد لأول مرة بناءً على الوثائق الاستراتيجية الموجودة وتحليل المحتوى. بعد ذلك، بناءً على أربع مصفوفات تأثير رئيسية: مصفوفة التأثير الضبابي، ومصفوفة قوة التأثير، ومصفوفة التأثير النهائية، وباستخدام برمجيتي FCMapper وPajek، رُسمت خرائط معرفية ضبابية. وتُظهر نتائج البحث أن التحديات الخمسة الرئيسية التي تواجه تطوير الذكاء الاصطناعي في البلاد تشمل: عدم الرغبة في الاستثمار الاقتصادي في الذكاء الاصطناعي، ونقص آليات نمو وتطوير سوق الذكاء الاصطناعي، ونقص القوانين واللوائح اللازمة للذكاء الاصطناعي، ونقص المتخصصين والخبراء في المجالات المتعلقة بالذكاء الاصطناعي، ونقص منصات تبادل البيانات.

الكلمات المفتاحية: الذكاء الاصطناعي، التحديات، اليمن، تحليل المحتوى، الخرائط المعرفية الضبابية.

I. Introduction

Artificial intelligence (AI) is one of the key technologies that has significantly contributed to administrative processes and solved many societal problems. It has become a focus of interest for many societies. Over time, the use and exploitation of AI technologies in factories, healthcare, banking, and security, as well as e-commerce, media, and mobile application platforms, has witnessed significant growth. Despite these developments, along with the technological advancements and solutions AI generates, more variables are emerging regarding how things are done and whether existing resources are sufficient to meet people's changing needs. Therefore, in addition to the unresolved issues, the challenges appear unsolvable, and experts are not sufficiently skilled to master their systems(Zhang et al, 2021).

Today, various economic, political, legal, social, security, and commercial sectors have demonstrated pioneering applications, and these will be integrated into human life in the not-too-distant future. Life would be difficult without AI. Overall, given its ability to process data and transform it into information, defined as goal-oriented behavior, and keeping pace with the rapid development of technology and the dissolution of organizations, human-centered companies will transform organizations and companies and enable their innovation management (Parida & Gassmann, 2021). This transformation is due to the multidisciplinary nature of AI, which has led to the emergence of diverse fields and activities. AI is a field of study, development, and innovation derived from diverse disciplines such as computing, mechanical engineering, and other industries that utilize characteristics of human intelligence such as mental abilities, learning, adaptability, acceptance, and decision-making (Chen, Chen, & Lin, 2020). At the policymaking level as well, decision-makers still need to understand the perspectives and expectations of societies and companies regarding AI from a future perspective so they can make decisions and formulate appropriate policies accordingly. This is because technology-related decisions impact the lives of millions of people and competition, and millions of companies will face long-term consequences (Lauterbach, 2019).

Yemen, like other countries, should strive to develop an AI development program. Therefore, it is necessary to prioritize AI development challenges. In this regard, this research aims to answer the fundamental question, taking into account the specific characteristics of our country: What are the main challenges in AI development in the Republic of Yemen? Challenges that, if appropriate measures are not taken to address them, AI development will not be achieved. Therefore, in the next article, while listing the challenges facing AI development in the country, we will also develop a system for linking and prioritizing these challenges. In fact, in this research, we sought to organize these challenges. The 36 challenges were extracted from expert opinions without discovering the relationships between them. Therefore, in this step, we sought to identify the connections between these challenges using fuzzy cognitive maps. By organizing the extracted factors, the fuzzy cognitive maps, on the one hand, sorted these factors and, on the other hand, identified the relationships between the challenges. In fact, this process relies on four primary impact matrices: a fuzzy impact matrix, an impact strength matrix, and a final impact matrix. Identifying the connections between challenges helps policymakers and decision-makers understand which other challenges might change if one challenge is resolved, and whether this change and the changes occurring in the other challenges are related to the other challenges. Is this beneficial to the system?. Fuzzy cognitive mapping helps us determine the intensity of these relationships and the impacts and implications of these challenges on each other.

Therefore, the purpose of the current study is to prioritize AI development challenges to help IT and AI leaders and managers, based on variables such as financial and human resource constraints, business, time, etc., address the challenges that will be of greater importance and wider impact, as multiple constraints do not allow for solving all challenges at once. The next section reviews the relevant literature. The research methodology is then presented. Data analysis and research findings are then examined in detail. The final section provides discussion, conclusions, and future recommendations.

II. Theoretical Foundations

Artificial intelligence (AI) is the key to achieving dramatic transformations in applications and operations in most organizations. Today, leveraging AI is considered one of the core activities of organizations. AI, as an academic discipline, was introduced and formally announced by John McCarthy (2006) at the Dartmouth Conference in 1956 and has since attracted the attention of many organizations. Artificial intelligence is a branch of computing or computational science that includes expert systems, algorithms, and slow computer programs. The true goal of AI is to mimic the human mind and make decisions like humans(Tolan et al. 2021).

Some thinkers believe that artificial intelligence (AI) is poised to disrupt the world. They argue that AI and intelligent machines capable of performing high-level mental processes such as reasoning, perception, learning, problem-solving, and decision-making, with advances in data collection, analysis, and computer processing power, have created opportunities to surpass human intelligence and enrich the way people live and work (Aayog, 2018). Such changes and developments brought about by the advancement of AI will naturally and inevitably create challenges, and if left unaddressed, they will inevitably lead to: If managed, corpses become an opportunity, but if left unchecked, they become a threat. Some of the challenges fall under the category of cultural challenges: ethical, psychological, social, and related to public trust. They arise when AI designers and engineers fail to consider human and social values in order to achieve their own goal of responsible AI products and services. A number of challenges arise when AI's relationship with legal issues, human rights, and regulations related to contracts, labor law, criminal law, etc. is neglected (Prec, Ozer, & Hojnik, 2019).

III. Research Background

AI research has made significant contributions both domestically and internationally. Among the important research with a long history, we can look back at some of the outstanding studies.

Boden (1998) presented a paper titled "Creativity and Artificial Intelligence." In this paper, he considers creativity to be a fundamental characteristic of human intelligence and believes that creativity poses a challenge to AI. The researcher argues that AI techniques can be used to generate new ideas in three ways: by creating clusters of similar ideas, exploring possible conceptual spaces, and creating changes and transformations that make previously impossible ideas possible (Boden, 1998).

In another study, Zhang et al. (2021) examined ethics and governance in AI. Researchers in the field of machine learning and artificial intelligence have worked on the ethics and governance of AI in situations such as job creation, job support, and job selection.

Tolan et al. (2021) studied the impact of AI on employment in society. By designing a framework for analyzing the impact of AI on employment, they categorized AI-related jobs, thereby improving people's awareness and perspectives on entering the labor market.

In a separate study, Miller (2018) focused on the social sciences to explain AI and believes that AI can be explained based on such studies and research on the subject of philosophy, mental sciences/psychology, and psychology.

In another noteworthy study, Cath (2018) examined AI governance frameworks across ethical, technical, and legal dimensions, and the challenges and opportunities related to these three dimensions. In his view, the wide-ranging impacts of AI on banking, social welfare, and human rights point to the proliferation of AI, or, in other words, the governance of AI, which must be considered to evaluate its fairness and transparency (Cath, 2018).

In another study (Saghiri et al., 2022), researchers focused on overcoming the challenges of AI. They argue that many sciences are organized based on observations of human intelligence, and therefore, numerous challenges arise in developing these sciences at the heart of AI-based systems. Some of the most significant challenges include the existence of systems for identifying problems, data integrity and imprecision, robustness, reliability, confidentiality, and controllability (Saghiri et al., 2022).

In another study focusing on the challenges of using AI in the education and higher education sectors, researchers addressed the challenge and divided the important issues in this field into three main sections: challenges related to research in education, challenges related to policy, and challenges and opportunities related to the relationship between educational institutions, universities, and industry (Luan et al., 2020). The challenges discussed in detail in the literature are those arising from the future of artificial intelligence. In other words, according to some researchers, if the production and application of AI in various fields grows unchecked, posing serious threats to medicine, transportation, and even weapons systems, it will create an overreliance on technology-based services, which is extremely costly and significantly increases risks (Pietikäinen & Silvén, 2021).

Responsible Artificial Intelligence: Requirements and Challenges is another study that addresses the risks and challenges facing responsible AI. It also addresses the social challenges of AI. The social risks of AI mentioned in this study include: decreased job opportunities, increased decision-making errors, negative behavioral changes in society, threats to military regimes, democracy, and economic challenges (Ghallab, 2019). Another study specifically on AI examines the status of AI use in cybersecurity in the United States. This study aims to assess the current challenges posed by the rise of AI to cybersecurity. This study considers various dimensions of security, validity, reliability, data, system reliability, and analytics, and addresses the relationship between AI and cyberspace (Soni, 2020). Foreign research in the field of AI can be seen in Table 1.

Table 1. Research in the field of artificial intelligence

Results	Subject	Reference	Number
Providing a collective vision for AI experts to highlight significant opportunities and realistically assess the impacts, challenges, and emerging research topics in the field of AI.	Artificial Intelligence: A Disciplinary Perspective on Challenges, Opportunities, Research, Application, and Policy Issues	& Dwivedi Williams, 2019	1
Demonstrate the important role of AI in productive employment.	Examining Ethics and Governance in AI	,Zhang, et.al 2021	2
.Categorize AI-related jobs	Impacts of AI on Employment in Society		3
Create explainable AI based on philosophy, psycholinguistics, and social psychology.	Artificial Intelligence with a Social Science Focus		4
How and how to guide AI development in the ethical, legal, and technical aspects of AI governance.	AI Governance: Ethical, Legal, and Technical Challenges and Opportunities		5
The necessity of using AI in digital pathology, leveraging opportunities, and overcoming challenges.	AI and Digital Pathology: Challenges and Opportunities		6
Categorize AI systems that are being used in health, and develop a global map of AI startups.	Challenges and Opportunities of AI in Health and Hygiene		7
The Role of Artificial Intelligence Technologies in Creating New Rules	Creativity and Artificial Intelligence	Boden, 1998	8
Important challenges facing AI include problem recognition, data integrity, invalidity, reliability and trust, confidentiality, controllability and predictability, etc.	What are the problems of artificial intelligence?	,Saghiri, et.al 2022	9
Among the most important challenges facing AI in the education sector are challenges related to educational research, challenges related to policymaking, and challenges related to the relationship between educational institutions and universities with industry.	Talks about the use of artificial intelligence in the education sector	,Luan, et.al 2020	10
Over-reliance on technology-based services is very costly and entails significant risks.	Threats posed by the unbridled growth of artificial intelligence	,Pietikäinen Silvén. 2021	11
Decreased employment opportunities, increased decision-making errors, negative societal behavioral changes, threats to military regimes, democracy, and economic challenges.	Responsible AI: Requirements and Challenges	Ghallab, 2019	12
The various dimensions of security, data integrity and validity, system reliability, and potential solutions for the relationship between AI and cyberspace.	The use of artificial intelligence in cybersecurity in the United States	Soni, 2020	13

https://doi.org/10.20428/jst.v30i6.2908

IV. Research Methodology

The current study is an applied descriptive study that uses content analysis to analyze AI development challenges and fuzzy cognitive mapping to prioritize AI development challenges. Figure 1 illustrates the research process.

Content Analysis

In addition to fieldwork, objective and experimental observation, long-term observations, library studies, and participation in meetings and group discussions, a semi-structured approach was used to collect information. Therefore, the central method of collecting information in this article is content analysis, which will be briefly explained below. The process of content analysis begins when the researcher seeks to identify patterns of meaning and themes in the data. This can be done through summarization. Data retrieval is then completed. The final stage of this process is to report on the content and meaning of the patterns and themes present in the data. Analysis is also a process that develops over time (Luan, et al., 2020). Data analysis using thematic analysis relies on the coding process. A theme or concept expresses the concept of a pattern present in the data and related to the research questions. This method is a textual data analysis process that transforms scattered and diverse data into rich and detailed data. The content grid organizes the following content based on specific procedures through four stages: "exposing the text," "capturing and understanding seemingly irrelevant information," "analyzing qualitative information," and finally "systematically observing a person, interaction, group, situation, organization, or culture": 1. Core themes (the codes and key points in the text), 2. Structured content (articles) obtained by combining and summarizing core themes, and 3. Overarching themes (higher-order themes that comprise the governing principles of the text as a whole) (Prec, Ozer & Hojnik, 2019).

In the first step of data analysis, interview outputs were analyzed using categorical content analysis, and the complications and harms caused by the development of artificial intelligence were identified. These complexities, in effect, constitute the input to the second step of analysis using fuzzy cognitive maps.

Fuzzy cognitive maps

Cognitive maps based on graph theory, by Axelrod (1994), were coined the title "A Structure for Evaluating Complex Social Relationships." Cognitive mapping is a technique used to model complex systems and identify causal relationships between them. Therefore, it has been applied in various technical and social sciences.

Recognizing that cognitive maps, including expert perspectives, relate to subjective, not real, reality. Objectively, and taking into account the quantitative ability of fuzzy logic, Kandasamy (1986) introduced weighted cognitive maps. Fuzzy cognitive maps are fuzzy graphical structures for representing causal relationships, representing the degree of ambiguity of the causal relationships between concepts with a number in the range [1, -1]. Fuzzy values, in addition to expressing the intensity of the relationship between variables, also indicate the direction of the relationship. For cognitive map analysis, we can count the number of variables and the number of correlations, but graph theory provides more indicators in addition to the number of variables (concepts and phrases) and correlations (relations and links) (Özesmi, 2018). The type of variables in a map is important because it shows how variables behave in relation to other variables. The presence of diverse variables in a cognitive map also makes understanding its structure easier. There are three types of variables: sender variables, receiver variables, and centrality variables. These variables are determined by their output and input scores. The output rank is the absolute value of the row sum of the variable values in the adjacency matrix and indicates the ability of the associative relationships extracted from the variable. The input rank of the column sum is the absolute value of the variables and indicates the ability to assimilate the relationships entered into the variable. The centrality (total influence) of a variable is the algebraic sum of its input score (hydro arrow) and its output score (output arrows).

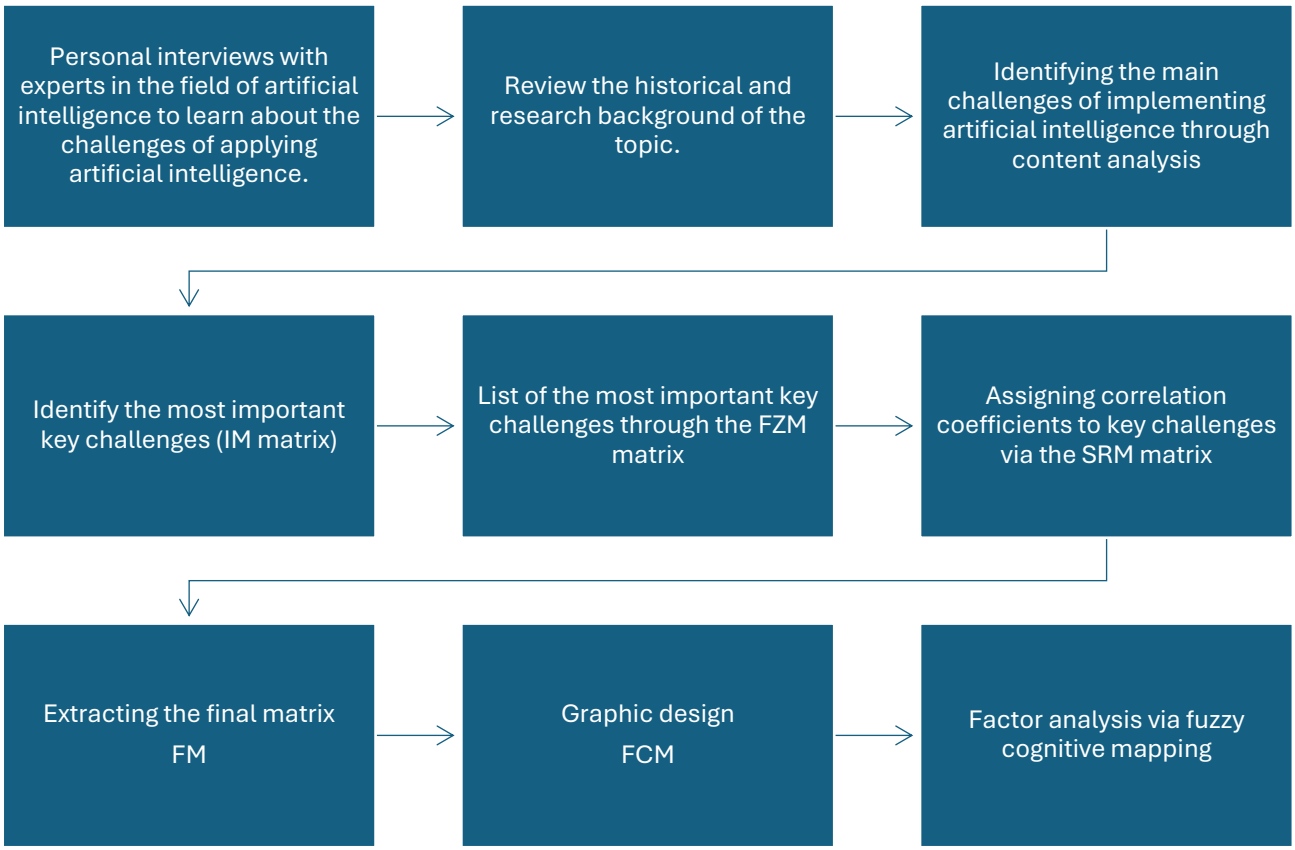


Figure 1. Research Implementation Process

Centrality refers to the contribution of each variable to the cognitive map and indicates how the variable relates to other variables and the strength of these relationships (Eden, 1992).

Research by Rodriguez-Repiso (2007) shows that the FCM method, a combination of the Analytical Hierarchy Process and the Critical Success Chain, works better. One of the most important advantages of spatial cognitive maps is the clear temporal variability. It is possible that the researcher does not have a solid understanding and knowledge of the problem, but with increasing experience and knowledge, new concepts can be easily added, eliminating the need to explain and analyze the problem from the beginning. This method can provide a highly strategic view of concepts and greatly assist the analyst in providing comprehensive and in-depth analyses. From an applied perspective, FCM can also provide managers with the ability to change strategies and monitor results. Fuzzy cognitive mapping and scenario design simulations and analysis help experts and decision makers identify different paths toward goals, key factors in achieving goals, and the feasibility of achieving goals based on experts' beliefs (Asadi et al., 2019).

It is worth noting that the statistical research community includes a number of academics, researchers, and experts relevant to the development of the ICT sector in the country. For the sample, we selected individuals who, in our opinion, have the most information on the topic compared to other authors. The reason for selecting these individuals was their experience and level of knowledge.

I. Data Analysis and Research Results

Text Coding and Extracting Key Themes

At this stage, to identify the most important challenges in AI development, 32 interviews were conducted with university academics, knowledge-based companies, experts from executive agencies, and policymakers, based on Table (2). Considering that this part of the research uses a qualitative approach, the 32 experts selected through cumulative sampling constitute an acceptable number, and the researcher achieved data sufficiency with this number of experts. In other words, unlike other methods, probability sampling, which aims to generalize the results of a representative sample of the population to the entire population under study, is not considered a goal in qualitative research. The criterion for describing or explaining a phenomenon is its most detailed form. Therefore, a criterion was introduced to achieve the maximum amount of information about the phenomenon. This criterion in this field of qualitative research is called saturation. Data saturation or theoretical saturation is one of the methods used in research. Qualitative criteria for determining sufficiency were used in the sample. It should be noted that the characteristic is that the number of members of the theoretical category or stratum has reached saturation. This situation allows for more data that can be expanded, modified, amplified, or added to existing theory.

Therefore, after a careful study of the texts, and in order to ensure that the challenges are realistic and useful, and that the models are actually practical, taking into account the harmful effects of the current situation, the challenges identified through interviews, literature, and documentary studies were examined, and through their analysis and synthesis, 86 main themes were identified. After extraction, 36 themes were finally selected. Table 3, for example, shows the extraction of some key themes.

Table 2. Demographic characteristics of the research sample

Number	verage Work Experience	ob Position
8	9	University faculty member specializing in artificial intelligence
9	6	CEO of knowledge and innovation companies developing artificial intelligence services
1	16	Expert in the Planning and Budgeting Authority
1	8	Expert in the Office of the Vice President for Science and Technology
13	19	IT experts in the executive branch

Among the experts, we selected three faculty members specializing in artificial intelligence, two CEOs of AI companies, a planning and budgeting expert, an IT expert, and three AI consultants working in FCM.

Collecting, summarizing, and classifying the coded texts and arriving at final factors

At this stage of the content analysis process, initial themes were extracted from the interviews and transcripts, ultimately arriving at 36 final themes. The results are shown in Table 4.

Table 3. Classification of AI Development Challenges

challenge title	general name of challenge	code
Lack of trust in artificial intelligence	Acceptance Artificial Intelligence	C1
Low public awareness and ambiguity about the concept of artificial intelligence and misconceptions about it		C2
Low level of knowledge (awareness) among managers in the field of artificial intelligence		3C
Lack of a comprehensive plan for AI development		C4
Perception of AI activities as time-consuming and costly		C5
A lack of urgency in solving problems through scientific methods and resistance from large private and public companies to the use of artificial intelligence		C6
Lack of scientific centers responsible for developing artificial intelligence	Research and Development in Artificial Intelligence	C7
Research-centric approach at universities rather than focusing on problem-solving in the field of artificial intelligence		C8
The uniqueness of AI for each project (meaning that after learning in one field, such as health or transportation, AI cannot be used in another field, and a new learning algorithm must be developed for each project).		C9
Lack of international collaboration between Yemeni researchers and specialists		C10
Weak university-industry relations in the field of AI and the absence of interaction mechanisms		C11
Lack of appropriate infrastructure and equipment for modeling, testing, and experimentation		C12

Lack of the necessary mechanisms for the growth and development of the AI market	Marketing, commercial and economic interest in artificial intelligence	C13
The country's unfavorable position regarding the establishment and maintenance of AI companies and jobs		C14
Unwillingness to invest economically in the field of AI		C15
Limited private sector access to powerful servers for processing big data		C16
Difficulty accessing non-native software		C17
Lack of bandwidth commensurate with AI activities		C18
Size and space limitations for storing AI-related data and information		C19
Poor data quality required for AI systems		C20
Lack of proper knowledge of existing data and the presence of dark (unknown) data		C22
High costs of data classification		C22
Lack of platforms for data sharing		C23
Lack of basic tools for Arabic language processing for AI activities		C24
Insufficient knowledge about AI capabilities	Educational and Research Interest in Artificial Intelligence Educational and Research Interest in Artificial Intelligence	C25
Difficulty teaching AI to people		C26
Weakness in AI specializations and graduates		C27
Lack of specialists and experts in AI-related fields		C28
Lack of specialized AI research centers		C29
Lack of necessary laws and regulations in the field of AI	Laws and Ethics Related to Artificial Intelligence	C30
Ethical challenges in AI		C31
Difficulty and confusion in integrating existing systems with AI systems		C32
Negative consequences of using AI (such as bias, unemployment, etc.)		C33
Lack of standards for accrediting AI-related activities in relevant institutions.		C34
The multiplicity and abundance of laws and IT policymakers in the country.		C35
Lack of sufficient knowledge among legislators about the capabilities and benefits of AI.		C36

Analyzing Cognitive Maps to Identify AI Development Challenges

To create a fuzzy cognitive map of AI development challenges, the first step is to compile a preliminary matrix. Therefore, the first step of the AI development challenge questionnaire was presented to 10 experts in the fields of AI, ICT, and industrial management. They were asked to rate the importance of each component on a five-point Likert scale, with categories of very high (VH), high (H), medium (M), low (L), and very low (VL). By collecting the experts' opinions, the preliminary matrix was created.

Fuzzy Matrix Calculations

Triangular fuzzy numbers are used to transform qualitative variables into quantitative variables. The mechanism for this transformation can be seen in Table 5.

Table 4. Triangular fuzzy numbers, five-point Likert scal

fuzzy number triangular	fuzzy value	abbreviation symbol	verbal variable
(0,0,0.25)	1	VL	Too little
(0,0.25,0.5)	2	L	Little
(0.75,0.5,0.25)	3	M	Average
(1,0.75,0.5)	4	H	A lot
(1,1,0.75)	5	VH	Too much

There are various methods for converting a fuzzy number into a specific number, including the center of gravity method, the surface center method, the maximum value **Relationship1**

$$x_m^1 = \frac{L + M + U}{3} : x_m^2 = \frac{L + 2M + U}{4} : x_m^3 = \frac{L + 4M + U}{6}$$

$$Crisp\ number = Z* = \max(x_{max}^1, x_{max}^2, x_{max}^3)$$

After converting the qualitative variables to fuzzy triangular numbers and converting the fuzzy triangular numbers to specific numbers, a fuzzy matrix is created. Given these relationships, the remaining elements take values between zero and 1. In this step, there must be an upper and lower bound. In this research, according to expert opinion, 5 represents the upper bound and 2 represents the lower bound. In Table 6, the fuzzy matrix summarizes the challenges of AI development:

Power Relationship Matrix Calculations

The power relationship matrix, or SRM, is a square matrix containing the same number of rows and columns as the number of AI development challenges. In this study, the row and column headings in the power relationship matrix represent the AI development challenges. In this section, the relationship between these components is studied as a binary relationship. Each element in this matrix represents the magnitude of the impact of challenge i on challenge j in the same matrix. This element can take values between 1 and -1. Each element of this matrix can take three values: positive, zero, or negative. If this value is positive, it indicates a direct connection between challenges i and j. If the value is negative, this indicates an inverse or negative connection between challenges i and j. If its value is zero, nothing happens between challenges i and j.

Forming the Final Matrix

When completing the SRM matrix, there may be some misleading data. This means that all the factors whose relationships were previously identified in the matrix may have illusory or incorrect relationships. In this situation, an expert in information analysis must be consulted to convert the SRM matrix to an FM matrix. In this situation, the FM matrix only displays numerically ambiguous elements, and method, the summation center method, and the weighted average center method. In this study, the center of gravity method was used with the help of Equation 1.

this matrix displays the causal relationship between the factors.

In this section, we consulted experts who have worked in the field of information technology and are also familiar with artificial intelligence topics. To identify the relationships between components, a cross-sectional table with the main column and row was provided to the experts as challenges for AI development. The experts were asked to select the relationship between the components from three scenarios: direct (positive), inverse (negative), and no relationship (zero). To ensure the identification of the type of relationship between the components, multiple experts in the field were consulted. An odd number of experts were selected and the relationship type was determined using the majority rule.

Graphic Representation of the Fuzzy Cognitive Map

In this stage, the graphical representation of the fuzzy mind map will be displayed to illustrate the challenges facing AI development. Figure 2 shows a fuzzy mind map of the challenges facing AI development. The final matrix of AI development challenges, after calculation and completion, becomes an input to the FCMapper program. After performing the calculations in the first stage, this program determines the values of the indicators for analyzing connections, the most important of which are: network density or density indicators, out degree, in degree, and centrality, which will be discussed in the next section. The second output of the FCMapper program is a file for drawing the connection network between components and factors as a network file. This file can be used as input for a network mapping program. In this research, the PEJK2 program was used to draw a map. In this map, each arrow between components i and j indicates the relationship between those two components. The degree and intensity of the relationship between two components is indicated by a number on the arrow. In the case of mutual communication between components, double-headed arrows are used.

Analysis of Fuzzy Cognitive Mapping Indicators

As shown in Table 6, the network density index is approximately 16 percent, indicating the ratio of the number of existing connections to the total possible connections. The number of map factors corresponding to the challenges of AI development in the current study was 36, and the number of regular connections was 213.

Table 5. Fuzzy Cognitive Mapping Indicators

density	Total number of factors	Total number of connections	number of disconnections	number of selfloops	number of regular connections	number of transmissions	number of receptions	number of ordinary
0.16	36	213	0	0	213	3	2	31

Analyzing External, Internal, and Centrality Indicators in Knowledge Maps

Three other main indicators are calculated by FCMapper: externality, internality, and centrality. The internality indicator indicates the total weight of the directional arrows entering an element, node, or network. In effect, this indicator demonstrates the influence of the element; the higher the indicator, the greater the influence of the node in the network. Table 7 illustrates the internal, external, and centrality challenges facing AI development.

As the table shows, the most effective approach focuses on the challenge of focusing university research on AI, which was assigned a score of 11.86. According to Figure 2, a knowledge map of the connections between the challenges, the three challenges with the greatest impact after this challenge were "Lack of proper knowledge of existing data and the presence of dark (unknown) data," "Unwillingness to invest economically in AI," and "Low quality of data required for AI systems," each with an impact of approximately 0.81. They have.

Following the main article challenge, the challenges of "lack of mechanisms for the growth and development of the AI market," "lack of platforms for data sharing," and "unwillingness to invest economically in AI" were significantly impacted, with scores of 10.96, 11.03, and 11.81, respectively.

The second indicator examined is the externality or influence index. As can be seen in Table 7, the highest degree of external influence or impact belongs to "Unwillingness to invest economically in AI," "Lack of necessary laws and regulations in the field of AI," and "Lack of specialized AI research centers," with scores of 11.68, 12.58, and 12.62.

The last indicator examined in this study is the centrality index. This index, which is derived from the sum of the external and internal indicators, indicates the importance of one component in its overall influence and impact on the other components. A component with the highest degree of centrality does not necessarily mean it has the highest degree of internality or externality. In Table 7, the centrality of the challenges facing AI development is arranged in descending order.

As can be seen in Table 7, the challenges of "Unwillingness to invest economically in AI," "Lack of necessary mechanisms for the growth and development of the AI market," and "Lack of necessary laws and regulations in the field of AI" have the highest centrality scores, with scores of 18.05 and 20.60. 23.59.

The challenge of "unwillingness to invest economically in AI," with a centrality of 23.59, has the highest overall impact. The challenges that have the greatest impact on the challenge of "unwillingness to invest economically in AI" are: "lack of the necessary mechanisms for the growth and development of the AI market," "lack of necessary laws and regulations in the field of AI," "lack of specialized AI research centers," and "low quality of data required for AI systems," with a weight of 0.81. Therefore, developing AI in the country requires focusing on the challenge of economic investment in AI as a top priority. To focus on this challenge, attention must be paid to the role and impact of four challenges: "Lack of necessary mechanisms for the growth and development of the AI market," "Lack of necessary laws and regulations in the field of AI," "Lack of specialized AI research centers," and "Low quality of data required for AI systems".

Table 7. Degree of externality, internality, and centrality of challenges

Challenge	Inner Degree	uter Degree	Central Degree	Challenge	Inner Degree	uter Degree	Central Degree
C15	10,96	12,62	23,59	C1	78/7	1,51	9,29
C13	11,81	8,79	20/6	C2	5,51	3/16	8,67
C30	5,47	12,58	18/05	C31	6,2	2,36	8.56
C28	6/15	10.04	16,18	C25	3,06	5,34	8,39
C23	11,03	4,68	15,72	C11	5,48	1,53	7,01
C6	5,38	10.16	15,54	C24	5,34	1,56	6,9
C34	4,68	10/16	14,84	C3	0	5,47	5,47
C8	11,86	1,57	13,42	C17	3,09	2,32	5,41
C29	1,56	11,68	13,24	C14	2,36	2,41	4,77
C21	8,7	2,35	11,05	C16	3,93	0,77	4,75
C20	6,26	4,78	11,04	C35	0	4,63	4,63
C26	4,67	6,12	10,79	C22	3,04	1,56	4,6
C19	7,11	3,26	10,37	C33	1,55	1,55	3,1
C4	0,79	9,29	10,08	C10	1,5	1,54	3,05
C27	3,84	6,21	10,05	C32	1,57	0,76	2,33
C7	2,3	7,73	10,03	C9	0,79	0	0,79
C18	3,96	5,56	9,51	C5	0,76	0	0,76
C12	7,91	1,58	9,49	C36	0	0,75	0,75

Prioritizing challenges through fuzzy knowledge mapping suggests that AI leaders and senior managers should first work to increase investor motivation to invest in AI by establishing knowledge-based companies, deploying AI technologies in large factories, and using AI algorithms to solve various problems in healthcare, transportation, communications, and other areas. The lack of a mechanism necessary for the growth and development of the AI market is another significant challenge, reflecting the inaccurate and sometimes repetitive administrative procedures that hinder AI development. In other words, given the novelty of AI technology, the infrastructure is still lacking, and the necessary conditions for its development in the local Yemeni market are not yet fully in place.

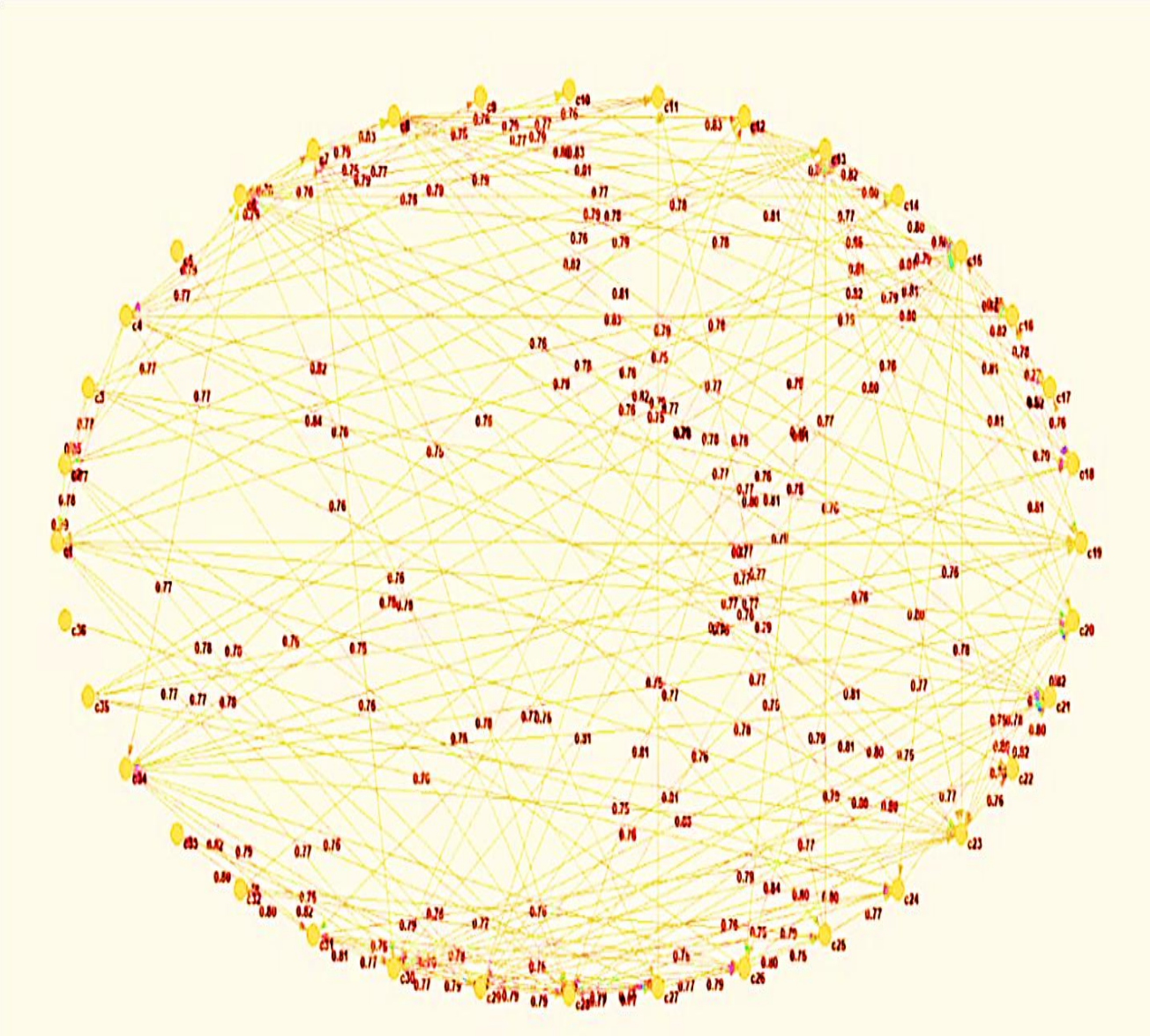


Figure 2. Research Knowledge Map

CONCLUSION

The lack of necessary laws and regulations in the field of artificial intelligence is another challenge and the third priority to be addressed. More specifically, legislative and regulatory authorities should work to reform and amend the legal implications of AI technology, reduce unnecessary administrative and bureaucratic barriers, facilitate business, increase tax exemptions and incentive subsidies, support AI growth and innovation centers, support and develop specialized laboratories for testing and experimenting with AI products to enable faster entry into the mass market, and assist in developing storage and processing infrastructure, intellectual property protection mechanisms, and government procurement processes. This is the most important step to mitigate the challenge of "investment hesitation."

Important limitations of this limited research into the challenges of AI development in the country include limited access to experts. For future research, we suggest that for each challenge, practical solutions, programs, and procedures be provided to AI development managers and officials to address these challenges. This issue is of great importance because many organizations responsible for emerging technologies, particularly AI, often face administrative and inexperienced challenges in determining appropriate procedures. To achieve this goal, a roadmap for AI development must be developed and an efficient and effective management system must be established to implement this roadmap.

REFERENCES

Boden, M. A. (1998). Creativity and artificial intelligence. Artificial intelligence, 103(1-2), 347-356.

Cath, C. (2022). Governing artificial intelligence: ethical, legal and technical opportunities and challenges. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 376(2133), 20180080.

Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. Ieee Access, 8, 75264-75278.

Ghallab, M. (2019). Responsible AI: requirements and challenges. AI Perspectives, 1(1), 1-7.

Haefner, N., Wincent, J., Parida, V., & Gassmann, O. (2021). Artificial intelligence and innovation management: A review, framework, and research agenda. Technological Forecasting and Social Change, 162, 120392.

Lauterbach, A. (2021). Artificial intelligence and policy: quo vadis? Digital Policy, Regulation and Governance.

Luan, H., Geczy, P., Lai, H., Gobert, J., Yang, S. J., Ogata, H., ... & Tsai, C. C. (2021). Challenges and future directions of big data and artificial intelligence in education. Frontiers in psychology, 2748.

Miller, T. (2022). Explanation in artificial intelligence: Insights from the social sciences. Artificial intelligence, 267, 1-38.

Perc, M., Ozer, M., & Hojnik, J. (2019). Social and juristic challenges of artificial intelligence. Palgrave Communications, 5(1), 1-7.

Pietikäinen, M., & Silven, O. (2022). Challenges of Artificial Intelligence--From Machine Learning and Computer Vision to Emotional Intelligence. arXiv preprint arXiv:2201.01466.

Rao, A. S., & Verweij, G. (2023). Sizing the prize: What's the real value of AI for your business and how can you capitalise. PwC Publication, PwC, 1-30.

Saghiri, A., Vahidipour, M., Jabbarpour, M., Sookhak, M., Forestiero, A. (2022). A Survey of Artificial Intelligence Challenges: Analyzing the Definitions, relationships, and Evolutions. Applied sciences, 12(8), 1-26.

Soni, V. D. (2020). Challenges and Solution for Artificial Intelligence in Cybersecurity of the USA. Available at SSRN 3624487.

Tizhoosh, H. R., & Pantanowitz, L. (2022). Artificial intelligence and digital pathology: challenges and opportunities. Journal of pathology informatics, 9, 38. doi: 10.4103/jpi.jpi_53_18

Tolan, S., Pesole, A., Martínez-Plumed, F., Fernández-Macías, E., Hernández-Orallo, J., & Gómez, E. (2021). Measuring the occupational impact of AI: tasks, cognitive abilities and AI benchmarks. Journal of Artificial Intelligence Research, 71, 191-236.

Zhang, B., Anderljung, M., Kahn, L., Dreksler, N., Horowitz, M. C., & Dafoe, A. (2020). Ethics and Governance of Artificial Intelligence Evidence from a Survey of Machine Learning Researchers. Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence (IJCAI-22) Journal Track.

فاعلية تصميم نظام اختبارات إلكترونية تكيفية على تقييم القدرة المعرفية وتعزيز التحصيل الدراسي

الاستـــلام: 28/فبراير/2025
التحكيم: 16/ مايو/2025
القبـــــول: 17/ مايو /2025

أحمد عبدالله بالحارث [1*]
ريم عبد عمر القامزي[1]

© 2025 University of Science and Technology, Aden, Yemen. This article can be distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

© 2025 جامعة العلوم والتكنولوجيا، المركز الرئيس عدن، اليمن. يمكن إعادة استخدام المادة المنشورة حسب رخصة مؤسسة المشاع الإبداعي شريطة الاستشهاد بالمؤلف والمجلة.

[1] قسم الحاسوب، كلية التربية، جامعة عدن، عدن، اليمن.
* عنوان المراسلة: balharethru@gmail.com

فاعلية تصميم نظام اختبارات إلكترونية تكيفية على تقييم القدرة المعرفية وتعزيز التحصيل الدراسي

أحمد عبدالله بالحارث
قسم الحاسوب، كلية التربية
جامعة عدن
عدن، اليمن
balharethru@gmail.com

ريم عبد عمر القامزي
قسم الحاسوب، كلية التربية
جامعة عدن
عدن، اليمن
reemabd88@gmail.com

الملخص:

استهدف البحث الحالي دراسة فاعلية تصميم نظام مقترح للاختبارات الإلكترونية التكيفية على تقييم القدرة المعرفية للطلبة ،ودورها في رفع التحصيل الدراسي لدى طلبة الجامعة ، من خلال المقارنة بين الاختبار الإلكتروني التكيفي، و الاختبار الإلكتروني (التقليدي) في تقدير القدرة باستخدام طريقة الأرجحية القصوى وقاعدة إنهاء الاختبار وفقًا لدرجة معيارية تم تحديدها سابقًا، بلغ عدد المشاركين في البحث (200) طالبًا من طلبة كليات جامعة العادل المستوى الأولى للعام الجامعي 2024/2023م. وقد توصلت نتائج الدراسة إلى وجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي الخطأ المعياري لتقدير القدرة بين الاختبار التكيفي والاختبار الخطي باستخدام طريقة الأرجحية القصوى لتقدير القدرة وقاعدة إنهاء الاختبار وفقًا لدرجة معيارية ولصالح الاختبار التكيفي، ووجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي درجات طلبة المجموعة التجريبية التي تم تقييمها باستخدام الاختبار التكيفي والمجموعة الضابطة التي تم تقييمها باستخدام الاختبار الخطي ولصالح الاختبار التكيفي، وتُعد النتيجة السابقة مؤشراً جيدًا على فاعلية الاختبار التكيفي في رفع التحصيل العلمي لدى الطلبة. واخيرًا خرجت الورقة بأهم التوصيات المتعلقة بالموضوع؛ حيث يسعى البحث إلى أن يقدم فكرة عامة لتطبيق التقنية الحديثة في التقييم، ودوروها في تطور أساليب التقويم باستخدام الاختبارات الإلكترونية التكيفية.

الكلمات المفتاحية : الاختبارات الإلكترونية التكيفية، الاختبارات الإلكترونية الخطية، التقييم، القدرة المعرفية، التحصيل الدراسي.

Abstract

The aim of the present study is to examine the effectiveness of designing a proposed adaptive electronic testing system in assessing students' cognitive ability and its role in increasing university students' academic achievement through a comparison between the adaptive electronic tests and the traditional electronic tests in estimating cognitive ability using the maximum likelihood method and the rule of completing the test according to a standard score previously determined. The participants of the study consisted of 200 first-year students from five colleges at Adel University for the academic year 2023/2024. The findings of the study revealed that statistically significant differences at the level of α=0.05 between the average standard error for estimating the ability between the adaptive test and the linear test using the maximum likelihood method. It is also observed that statistically significant differences at the level of $\alpha = 0.05$ between the mean scores of the experimental group, assessed using the adaptive test, and the control group, assessed using the linear test and in favor of the adaptive test, the previous findings is strong evidence of the effectiveness of adaptive test in improving students' academic achievement. Based on the findings, the study proposed some recommendations related to the topic, emphasizing the importance of integrating modern technology in assessment and highlighting the role of adaptive electronic test in advancing evaluation methods in higher education.

Keywords— *Adaptive electronic testing, linear electronic testing, assessment, cognitive ability, academic achievement.*

I. المقدمة

يشهد العصر الحالي تطورًا سريعًا في المعرفة كماً وكيفًا، مما يدفع المؤسسات التعليمية إلى البحث عن أدوات تمكنها من مواكبة هذا النمو المتتالي للمعرفة، لذلك تزايد الاهتمام بدعم بيئات التعلم ببرمجيات الحاسوب وتكنولوجيا المعلومات بشكل عام وبيئات التعلم التكيفية بشكل خاص، وذلك لما تقدمه من عدة مزايا، من أهمها سهولة الوصول إلى المعلومات واحتوائها على بيئة تفاعلية (بصرية وسمعية)، بالإضافة إلى أنها تضيف عوامل التشويق والإثارة إلى بيئة التعلم، مما يحسن الدافعية لدى المتعلمين.

ويرى الخزي والزكري [1] أن استخدام التقنيات الحديثة في التعليم يتوقف قبل عملية التقييم، وكأن التقييم لا يعتبر جزءاً من العملية التعليمية، بل إن التقنيات الحديثة يمكن أن تسهم بشكل فعال في تطوير التقييم، وليس من المنطقي أن يسعى التربويون إلى اكتشاف الاستراتيجيات الإلكترونية الحديثة في التعليم وهم يلتزمون بالطرق المعتادة في تقييم مخرجات ذلك التعليم.

لذا كان لا بد من زيادة الاهتمام ببناء الاختبارات حيث تعد من أهم المقاييس النفسية والتربوية التي تساعد في تقدير وتقييم القدرة المعرفية للطلبة والسعي إلى تطويرها لمساعدة متخذي القرار على التنظيم والتطوير والتحسين حتى يمكن اتخاذ القرارات التربوية بدقة أكبر، ولابد أن تتميز هذه الاختبارات والمقاييس العلمية بالصدق والاستقرار والموضوعية.

لذلك تعد الاختبارات بأنواعها من أهم أساليب التقويم في العملية التعليمية للتأكد من مدى تحقيق الأهداف التعليمية والتربوية المنشودة، ومن مدى التقدم الذي أحرزه المتعلمين.

وقد أكد ما سبق الرويس وويلز ووالد [2] في دراسة وجهت عدة انتقادات لعملية التقييم في صيغتها التقليدية المبنية على الاختبارات التقليدية، بما في ذلك مشكلاتها وأهمها عدم مراعاة خصائص الطلبة وما يصاحب ذلك من قلق وذاتية للمصحح؛ وهو ما ينعكس غالباً على عمليات التعليم والتعلم التي تتم بشكل متزايد إلكترونياً في جميع مراحل وقطاعات المنظومة التعليمية في ظل جائحة كورونا؛ مما فرض على عناصر المنظومة التعليمية واقع كيفية التعامل مع المناهج والأساليب الجديدة في التقييم وتفعيل استخدامات التكنولوجيا المتقدمة المستمدة من تطبيقات النظرية الحديثة في التقييم، وتمثل تكنولوجيا المعلومات والاتصالات العنصر الأساسي في تطبيق الاختبارات الإلكترونية أو الاختبارات الحاسوبية، وهي من الاتجاهات الحديثة في التقييم والتي من خلالها يمكن الحكم على مدى تحقيق الأهداف التعليمية، وفعالية أساليب واستراتيجيات التعليم والتعلم، وقدرات المتعلم وجاهزيته للتعليم، وكذلك فعالية مصادر التعلم.

أكد البربري ما تقدم، وأشار أيضاً إلى أن القاعدة الأساسية لهذه الاختبارات هي أنه ليس من المعقول قياس قدرة أو سمة مجموعة من الطلبة بنفس عدد كلمات المفردات ونفس مستوى السهولة والصعوبة، بل من الأجدى قياس أداء هؤلاء الطلبة بكلمات مفردات تختلف في مستوى صعوبتها حسب مستوى أداء الطالب في توليد الأسئلة من بنوك الأسئلة التي تحتوي على مئات كلمات المفردات أثناء الاختبار[3] .

لذلك ترى البياضة [4] إن ظهور الاختبارات التكيفية كان حلًا لمشكلة الاختبارات التقليدية والمتمثلة في تعريض المفحوصين للفقرات نفسها بغض النظر عن ملاءمة هذه الفقرات لقدراتهم حيث تشمل الاختبارات التقليدية بعض الفقرات السهلة التي يتعرض لها أفراد من ذوي القدرات المرتفعة، وهذا بحد ذاته يمثل مضيعة للوقت والجهد، ثم أن المفحوص قد يشعر بنوع من الملل من جراء اختباره بفقرات لا تتحد قدراته ؛ما يدفعه للإجابة عنها بشيء من عدم الاهتمام ،وكذلك عند اختبار الأفراد ذوي القدرات المنخفضة بفقرات صعبة، مما يشعرهم الإحباط لذا الإجابة عنها بشكل عشوائي.

يبدو من خلال استعراض الأدب النظري وجود إسناد نظري لموضوع الدراسة الحالية ولذلك يسعى هذا البحث إلى التغلب على مشكلات الاختبارات التقليدية من خلال استخدام الاختبارات التكيفية الإلكترونية، وقدرتها في تقييم القدرة المعرفية للطلبة، ورفع التحصيل العلمي للطلبة، وذلك كما أشار إلى أهميته كثير من الباحثين.

II. مشكلة البحث:

تتمثل المشكلة الرئيسية للبحث في كثرة مشكلات الاختبارات التي يتم إعدادها وتقديمها بصورة تقليدية، والتي من الممكن التغلب عليها باستخدام مزايا الاختبارات الإلكترونية التكيفية. وبسبب ندرة الدراسات والابحاث العربية التي كان لها اهتمام في تطبيق الاختبارات الإلكترونية التكيفية في قياس قدرات الطلاب في المقررات التعليمية، تبرز الحاجة إلى تصميم نظام اختبارات إلكترونية تكيفية وقياس فاعليته على تقييم القدرة المعرفية وتعزيز التحصيل الدراسي.

أسئلة البحث:

في ضوء مشكلة البحث المتمثلة في عدم وجود نظام اختبارات إلكترونية تكيفية، يمكن صياغة أسئلة الدراسة على النحو التالي:

1. السؤال الرئيسي:

- ما فاعلية تصميم نظام اختبارات إلكترونية تكيفية على تقييم القدرة المعرفية وتعزيز التحصيل الدراسي؟

2. الأسئلة الفرعية:

- ما معايير بناء وتصميم الاختبارات التكيفي الإلكترونية؟
- ما الفروق بين الاختبارات الخطية والاختبار التكيفي المحوسب في متوسط تقدير القدرة؟
- ما الفروق بين الاختبارات الخطية والاختبار التكيفي المحوسب في تعزيز التحصيل الدراسي؟

III. فرضيات الدراسة

1. لا توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي تقدير القدرة والخطأ المعياري لتقدير القدرة (كل على حده) للاختبار التكيفي والاختبار الخطي باستخدام طريقة الأرجحية القصوى لتقدير القدرة.
2. لا توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي درجات طلبة المجموعة التجريبية التي تم تقييمها باستخدام الاختبار التكيفي والمجموعة الضابطة التي تم تقييمها باستخدام الاختبار الخطي.
3. لا توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي درجات طلبة المجموعة التجريبية التي تم تقييمها باستخدام الاختبار التكيفي وبين المجموعة الضابطة التي تم تقييمها باستخدام الاختبار الخطي تعزى لمتغير الجنس (ذكور/أناث).
4. لا توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي درجات طلبة المجموعة التجريبية التي تم تقييمها باستخدام الاختبار التكيفي وبين المجموعة الضابطة التي تم تقييمها باستخدام الاختبار الخطي تعزى إلى متغير التخصص.

IV. أهداف البحث:

1. التعرف على الخلفية النظرية للاختبار الإلكتروني تكيفي.
2. الكشف عن اهم المعايير التي تساعد على بناء اختبار تكيفي الإلكتروني.
3. تصميم نظام مقترح لإعداد الاختبارات الإلكترونية التكيفية.
4. تقييم أثر استخدام الاختبار الإلكتروني التكيفي على التحصيل الدراسي.
5. الخروج ببعض التوصيات والاقتراحات المتعلقة بالاختبار الإلكتروني التكيفي.

V. أهمية البحث:

يمكن تحديد أهمية البحث في النقاط التالية:

1. يقدم البحث معلومات عن الاختبار الإلكتروني التكيفي الذي يتم بناءه وفق نظرية الاستجابة المفردة، موضحًا أهميتها كأحد أنواع الاختبارات الإلكترونية ودورها في معالجة القصور في جوانب الاختبار الإلكتروني التقليدية.
2. تنبع أهمية الدراسة من أهمية الاختبار الإلكتروني ودوره الهام في عملية تقييم العملية التعليمية.
3. توجيه نظر الباحثين إلى البحث واكتشاف طرق تقويم الإلكترونية تساعد على حل مشكلات الاختبارات الإلكترونية الخطية، وتتوافق مع قدرة الطالب والحد من عملية القلق والغش والتخمين.
4. توجيه اهتمام مصممي برمجيات الاختبارات الإلكترونية إلى تطوير طرق تقديم الاختبارات الإلكترونية مما يساعد في رفع تحصيل الطلاب وفق للنظريات الحديثة النفسية والتربوية.

VI. حدود الدراسة:

يتحدد البحث الحالي بالآتي:

- *الحدود البشرية:* عينة الدراسة (200) طالبًا وطالبة من طلبة جامعة العادل جميع الكليات المستوى الدراسي الأول
- *الحدود المكانية:* جامعة العادل (جميع كليات الجامعة).
- *الحدود الزمانية:* الفصل الدراسي الثاني للعام الدراسي 2024/2023م.
- *الحدود الموضوعية:* تم اختيار مقرر مهارات الحاسوب (2) وحدة الإنترنت.

VII. مصطلحات البحث:

في ضوء الاطلاع على التعريفات التي وردت في العديد من الأدبيات التربوية ذات العلاقة بمتغيرات البحث، ومراعاة طبيعة بيئة التعلم والعينة، وأدوات القياس بهذا البحث تم تعريف مصطلحات الدراسة إجرائيًا على النحو التالي:

- **الاختبار الالكتروني الخطي:**

ويعرف الاختبار الالكتروني الخطية إجرائيًا: هي إحدى الطرق التعليمية المنظمة التي تهدف لقياس مستوى أداء الطلاب بطريقة الإلكترونية سهلة مما يساعد على التغلب على صعوبات الاختبارات الورقية.

- **الاختبار الالكتروني التكيفي:**

الكلمات المفتاحية

- **ويعرف الاختبار الالكتروني التكيفي إجرائيًا:** هو الاختبارات الذي يقدم للطلبة المفردات التي تناسب مع قدراتهم ويختلف عدد المفردات ومراحلها وترتيبها من طالب إلى آخر حسب قدرة كل طالب.
- **التقييم الإلكتروني:**
- **ويعرف التقييم الإلكتروني إجرائيًا:** هو تطبيق الاختبارات باستخدام الكمبيوتر بدلًا من إجراء الاختبارات بشكل ورقي ويتم من خلالها توظيف البرامج وتقديم التغذية الراجعة الفورية، والتأكد من مدى اكتساب الطلبة للمعلومات.

القدرة المعرفية:

- **وتعرف القدرة المعرفية إجرائيًا:** هي مجموعة من العمليات الذهنية التي تمكن الطالب على التعامل مع المعلومات العلمية التي يحصل عليها عقلياً لتسهيل عملية التعلم.
- ***المعالجة الإحصائية المتبعة في البحث:***

تم استخدام الاختبارات الإحصائية التالية في تحليل البيانات من خلال برنامج الحزمة الإحصائية (spss25 Statistical Package for Social Sciences الإصدار الخامس والعشرون:

1. الإحصاءات الوصفية (الوسيط الحسابي –الانحراف المعياري-النسبة المئوية التكرار-الأهمية النسبية).
2. واستخدام اختبار t-test لعينتين.
3. للتحليل العاملي للتحقق من نظرية الاستجابة المفردة تم اخضاع بنك الأسئلة للتحليل العاملي وفق محك leckase.
4. تحليل التباين الأحادي (One Way ANOVA) لاختبار معنوية الفروق بين أفراد عينة الدراسة المتعلقة بالجنس وعدد مرات الاستخدام والتخصص.

VIII. الإطار النظري والدراسات السابقة:

الإطار النظري:

الاختبارات الإلكترونية الخطية:

لقد كانت الاختبارات الإلكترونية موجودة منذ زمن طويل، كما أشار كلاً من خان وجاويد [5] إلا أنها كانت تستخدم على نطاق محدود، بسبب بعض القضايا المتعلقة بموثوقية الاختبارات، إلى أن اجتاحت جائحة كورونا العالم، فحققت المؤسسات التعليمية قفزة نوعية على كافة مستويات التعليم الجامعي وما قبل الجامعي، وعلى المستويين الإقليمي والدولي في استخدام الاختبارات الإلكترونية، مع البحث عن طرق آمنة ومقبولة وموثقة وعادلة لتطبيق هذه الاختبارات.

تتسم الاختبارات الإلكترونية بمجموعة من الخصائص التي تتمثل فيما يلي:

- أقل تكلفة مقارنة بالاختبارات الورقية؛ حيث أنها توفر على المؤسسة تكاليف كل من الطباعة، والحفظ، والنقل.
- يمكن إعداد أنموذجات مختلفة من الاختبار الواحد في نفس الوقت.
- يسهل مراجعتها واكتشاف ما بها من أخطاء وتصحيحها [6].
- سهولة التصحيح تلقائيًا، ثَم يمكن للمعلمين معرفة درجة الواجبات المنزلية الأسبوعية للطلاب على الفور لذا تقديم ملاحظات بنائية لم تكن متاحة بواسطة التعليم اليدوي؛ ولا سيما مع الفصول ذات الأعداد الكبيرة.
- إن الاختبار يمكن الطلاب من رؤية نتيجة اختبارهم الذاتي لتقدمهم، بالإضافة إلى تلقي ردود فعل فورية على محاولاتهم. ومن السهل التأكد من حصول الطلاب على أسئلة مختلفة وقيم متغيرة في كل مرة[7] .

بالرغم من تلك المميزات للاختبارات الإلكترونية إلا أن قد أوضحت دراسة ويس[8] أن التطور النوعي في نظريات القياس من جانب والتطوير ونشر وتوزيع الاختبارات الإلكترونية من جانب آخر قد أظهر

العديد من جوانب القصور في الأنماط المختلفة من الاختبارات الإلكترونية التقليدية خاصة التي تعتمد على توليد العشوائية للمفردات من بنوك الأسئلة ذات قواعد بيانات هيكلية ثابتة دونما اعتبار لمستوى الأداء الفعلي للمتعلم في تحديد مستوى صعوبة أو سهولة المفردات

- آليات سحب المفردات منها وفقًا مستوى أداء المتعلم الأمر الذي قد أدى إلى ظهور قدر مرتفع من القلق الاختبار أثر بالسلب على أداء المتعلم وتكوين اتجاهات سلبية نحو التوسع في الاختبارات الإلكترونية عمومًا.

لذلك كان لظهور الاختبارات الإلكترونية التكيفية أهمية أكد عليها العديد من الدراسات منها دراسة [9] التي أوضحت الآتي:

إلى أنه في الاختبارات الورقة والقلم والاختبارات الإلكترونية الخطية، يتم تقديم نفس العناصر لجميع الأشخاص بغض النظر عن مستوى قدرتهم، وبغض النظر عما إذا كانت العناصر سهلة أو صعبة بالنسبة لهم، حيث لا توفر العناصر السهلة جدًا والصعبة جدًا معلومات عن مستوى أداء الشخص. وفي ضوء التناقض بين قدرة الشخص وصعوبة العنصر، سيؤدي هذا إلى انخفاض في كمية المعلومات، ومن ثم إلى انخفاض في دقة القياس.

الاختبارات الإلكترونية التكيفية:

تعد الاختبارات الإلكترونية التكيفية كما أوضح مارتن ولازنديك من الاتجاهات الحديثة لقياس القدرات العقلية وجوانب الإنجاز المختلفة، فالاختبارات الإلكترونية التكيفية هي اختبارات تتكيف مع قدرة المتعلم، بحيث لا يعرض على الطالب إلا الأسئلة أو المفردات المناسبة لمستواه، فلا يتعرض لمفردات سهلة جداً أو صعبة جداً، وبالتالي يتم الحصول على أقصى قدر ممكن من المعلومات عن قدرة الطالب بأقل عدد ممكن من كلمات المفردات، ومن ثم فإن عدد كلمات المفردات ومراحلها وترتيبها يختلف من طالب إلى آخر حسب قدرة كل طالب[10] .

أوضح كلًا من فون ديفر وماجس ين [11] وجود اختلاف بين الاختبارات المتكيفة مع الحاسوب (الإلكترونية التكيفية) بشكل واضح عن الاختبارات الخطية، وهي الاختبارات التقليدية الأكثر استخدامًا حتى الآن، وتُسمى أحيانًا اختبارات الورقة والقلم، وهي اختبارات يجيب فيها جميع الممتحنين على جميع الأسئلة في الاختبار بغض النظر عن صعوبة العناصر وبغض النظر عن مستوى الممتحن. تختلف الاختبارات المتكيفة مع الحاسوب أيضًا عن الاختبارات المعتمدة على الحاسوب (الإلكترونية الخطية) بشكل عام، حيث يمكن استخدام أجهزة الكمبيوتر لإدارة الاختبارات، ولكن أيضًا بطريقة خطية، كما هو الحال مع اختبارات الورقة والقلم.

التالية وكذا توقيت وعدد المفردات اللازمة لإنهاء الاختبار؛ مما يظهر وجود مشكلات في:

- تصميم بناء بنوك أسئلة الاختبارات الإلكترونية التقليدية.

أهمية الاختبارات الإلكترونية التكيفية في تنمية مهارات الطلبة:

أشار كلًا روبيرت وديفاد[12] إلى مجموعة الفوائد التي قد يحصل عليها الطالب بفضل تقييمه باستخدام هذا النوع من الاختبارات لخص بالنقاط الآتي:

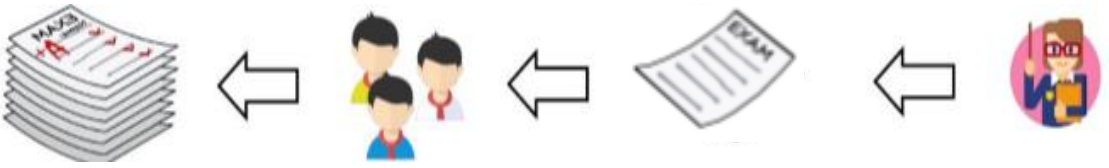
- تقييم الأداء بشكل دقيق وعادل: حيث يمكن تقييم أداء جميع الطلاب على نفس المستوى حتى إذا تفاعلوا مع مفردات مختلفة، ومن منظور نفسي فان هذا يزيد من الأنصاف في الدرجات وتقييم الأداء.
- زيادة الدافعية يتيح الاختبار التكيفي تحديًا مناسب لكل طالب يخضع للاختبار، وبذلك لا يشعر أصحاب الأداء المرتفع أو المنخفض بالملل في أثناء التقييم بالتفاعل مع المستويات المناسبة من التحدي.

مقارنة بين الاختبار التكيفي الإلكتروني والاختبار الخطي (التقليدي) الإلكتروني:

- وقد أشار كلًا من ماجس وين وديفر[11]بأن هناك اختلافات بشكل واضح بين الاختبارات الإلكترونية التكيفية والاختبارات الخطية Linear Tests أوضحه كالتالي:
- الاختبارات الخطية Linear Tests وهي الاختبارات التقليدية التي تعدّ الأكثر استخدامًا حتى الآن، ويطلق عليها أحيانًا اختبارات الورقة والقلم Paper-and-pencil Tests وهي تلك الاختبارات التي يجيب فيها جميع المفحوصين على جميع أسئلة الاختبار بغض النظر عن صعوبة تلك المفردات وكذلك بغض النظر عن مستوى المفحوص.
- وتختلف أيضًا الاختبارات الإلكترونية التكيفية عن الاختبارات المحوسبة Computer-based Tests عمومًا، حيث يمكن أن تستخدم الحواسيب في تطبيق الاختبارات ولكن بصورة خطية أيضًا كما هو الحال في اختبارات الورقة والقلم

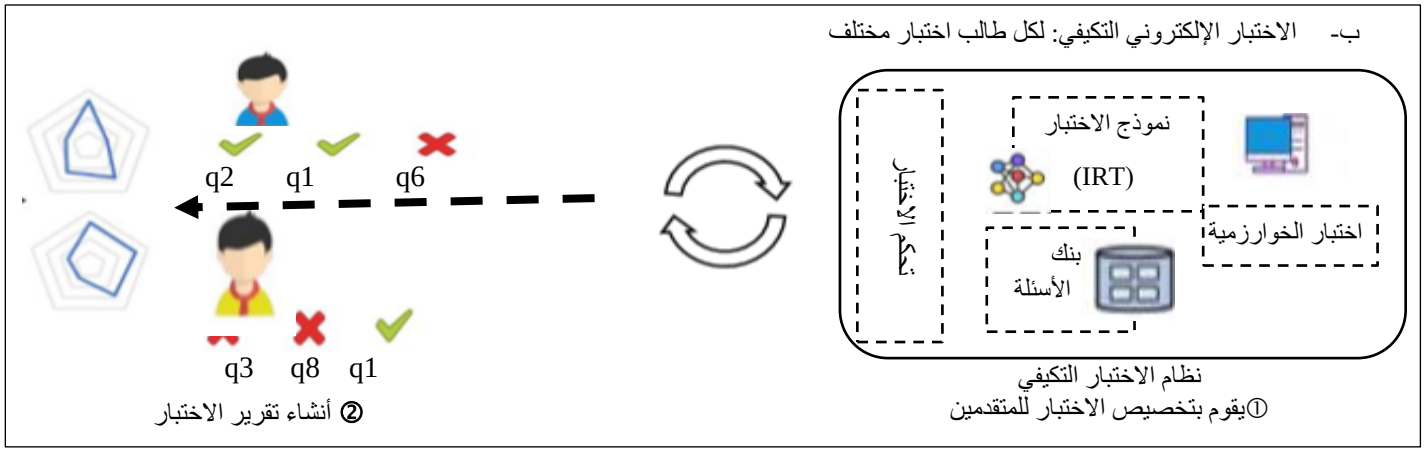
يمكن المقارنة بين الاختبارين من خلال الرسم التوضيحي الذي أشار إليه ليووآخرون [13] الذي يظهر في الأشكال التالية:

أ- الاختبار الإلكتروني الخطي (التقليدي): اختبار لكل طالب



①يقوم المعلم بتصميم الاختبار ②يقوم الطلبة بأداء نفس الاختبار ③يقوم النظام بإبلاغ الطلبة بالدرجة

الشكل 1: فكرة الاختبار الإلكتروني الخطي (التقليدي)



الشكل 2: فكرة الاختبار الإلكتروني التكيفي

وقد أضافة فرنك أن هذه الاختبارات تقيس قدرة الطلاب بكفاءة كونها تقدم سلسلة من الأسئلة التي تم تمييزها بدرجة صعوبة، ويتم اختيار الأسئلة لتتناسب مع مستوى قدرة الطالب، وعند نجاحه في السؤال المقدم له، يتم تقديم سؤال أكثر صعوبة بعد ذلك، وإذا أجاب عن السؤال بشكل غير صحيح يتم تقديم سؤال أقل تحديا، ويتوقف الاختبار عندما يتم تقدير قدرة المتقدم للاختبار بدقة المطلوبة. أي إلى أن يصل الطالب إلى قاعدة الإنهاء التي تم تحديدها من قبل المعلم[14] .

وصف الاختبار التكيفي:

وصف مارتينيز بلومي وآخرون [15] الاختبارات الإلكترونية التكيفية ترتكز على:

أساس نظري مستمد من نظرية الاستجابة للمفردة (item Response Theory) IRT التي عرفها محاسنة[16] نظرية تدور حول المفردة والأداء على الاختبار وكيفية ارتباط الأداء بالقدرة التي تقاس بالفقرات واستجابة المفردة يمكن أن تكون منفصلة أو متصلة أو متفرعة.

كما تعتمد هذه النظرية على القيمة الاحتمالية لاستجابة المتعلم للمفردة الاختبارية وتكون دالة لكل القدرة التي يفترض أن يقيسها الاختبار لدى المتعلم، وخصائص المفردة التي يحاول الإجابة عنها، وإن ذلك يتطلب الحصول على معلومات من مصدرين :

1. أحدهما يتعلق بالمتعلم.
2. والآخر يتعلق بالمفردة الاختبارية؛ وعادة نحتاج إلى قيمة عددية واحدة تتعلق بالمتعلم، وهي بارامتر القدرة المقاسة لدى المتعلم، وقيمة عددية أو أكثر تتعلق بالمفردة الاختبارية، وتمثل هذه النظرية أحد ركائز القياس النفسي الحديث، حيث ترى أن تساوي طول الاختبار من حيث عدد ونوعية المفردات لكل الممتحنين من الناحية المنهجية غير مقبول في قياس الأداء، وأن تقدير أداء المتعلم **يجب أن يكون قائماً على بعدين أساسيين هما:**
1. مدى امتلاك الفرد للقدرة أو السمة موضع القياس.
2. خصائص المفردة السيكومترية ومن ثم اعتمادها في القياس على الأداء الفعلي للمتعلم أثناء الاختبار دون اعتبار الفرضية اعتدالية خصائص العينة التي ينتمي إليها المتعلم كما في نظرية القياس التقليدية. ويتم تفسير العلاقة بين استجابات المفحوص وخصائص المفردة في رصد القدرة الفعلية للمتعلم بشكل مستقل عن العينة التي ينتمي اليها المفحوص.

وأوضح الدامغ والهاجري أنه عند استخدام الاختبارات الإلكترونية التكيفية يمكننا التحكم في مستويات الصعوبة والسهولة بما يتناسب مع مستوى الممتحن، وهذا يؤثر ضمناً على اتجاه الاختبار، حيث قد يرتفع أو ينخفض بناءً على إجابات المتعلم على الاختبار، وبالتالي يمكن الحكم على مستواه دون الحاجة إلى عرض جميع الأسئلة [17].

ومن خلال ما سبق يتضح لنا أهمية هذا النوع من الاختبارات والذي يعد من الأساليب الحديثة التي تساعد على القياس والتقويم وفق نظرية الاستجابة للمفردة، وتختلف الاختبارات الإلكترونية التكيفية في طريقة تطبيقها من حيث:

- يزود الممتحن بالفقرات التي تناسب مستواه فقط وذلك بعرض بعض الفقرات لتحديد قدرته الأولية.
- بناءً على قدرته الأولية المقدرة من خلال أدائه على تلك الفقرة، يزود بفقرة أخرى لاحقة من بنك الأسئلة تتوافق مع التقييم المستمر لمستواه.
- يمكن تقدير القدرة بطريقتين: الارجحية القصوى وطريقة بييزي.
- تتميز هذه النظرية أيضًا بتوفير الوقت والجهد في قياس قدرات الممتحن بأقل عدد من الفقرات.

مبادئ الاختبار الالكتروني التكيفي:

يقوم الاختبار الإلكتروني التكيفي على مبدأين أساسيين هما:

1. بناء بنك أسئلة

عند تصميم الاختبار الإلكتروني التكيفي لا بد من بناء بنك الأسئلة والذي يعد خطوة أولى، كما أوضح بيدر وبيرجونز لأنه يعتمد على مجموعة كبيرة نسبياً من المفردات، إذ يتطلب التقييم الدقيق للمتعلمين ذوي الخصائص المختلفة أن يحتوي هذا البنك على عدد كبير من مفردات الاختبار، مختلفة في معاملات صعوبتها وموزعة بشكل جيد في بنك الأسئلة، يسهل اختيارها في أي اختبار بمواصفات محددة، ومن هذه المعلومات مثلاً: الإجابة الصحيحة، معامل الصعوبة، معامل التمييز، معامل التخمين، وأن عدد (30) مفردة بهذه المواصفات يمكن أن يكون كافياً لبناء اختبار إلكتروني تكييفي [18].

2. إدارة الاختبار الإلكتروني التكيفي:

وقد تم الاعتماد على نوعين من الأسئلة، النوع الأول هو الأسئلة الثنائية التي تسجل الإجابة بصواب أو خطأ وتحدد الدرجة بـ (0,1)، والنوع الأخر هو أسئلة الاختيار من متعدد والتي تكون الإجابة فيها باختيار إجابة واحدة صحيحة من بين أربعة بدائل، وتحدد الدرجة أيضًا بـ (0,1).

تصميم الاختبارات التكيفية المحوسبة:

أوضحت عدد من الدراسات الخطوات الأساسية التي يمر بها تصميم الاختبار التكيفي منها ما أشار إليه Younyoung &Cayce [19]:

فقد قسمها إلى مجموعة من الخطوات الأساسية التي يمر بها تصميم الاختبار التكيفي تتمثل في:

1. تصميم بنك الأسئلة الخاص بالاختبار والذي يتم من خلاله توليد الأسئلة وفق خصائص وقدرات المتعلم.
2. يليها تحديد نقطة البداية والتي تتمثل في اختيار أول سؤال من بنك الأسئلة.

في حالة قيام المتعلم بالإجابة بشكل صحيح يتمكن النظام من تحديد مستوى المتعلم ويترتب عليه تحديد المفردة التالية والتي تتدرج في مستوى الصعوبة ويكون ذلك وفق قدرة المتعلم حتى يصل المتعلم إلى العدد المطلوب من حيث الكم أو وفق القدرة المحددة حسب المتعلم وعليه يتم إنهاء الاختبار أو تكليف المتعلم بمهام أخرى وفق قدرته.

وأشارت دراسة بوساكوك وآخرون إلى ضرورة [20] التصميم الجيد للاختبارات الإلكترونية التكيفية.

كما حدد إبراهيم خطوات بناء الاختبارات الإلكترونية التكيفية فيما يلي [21]:

1. بناء بنك الأسئلة لمفردات الاختبار.
2. اختيار النموذج المناسب للدراسة من نماذج نظرية الاستجابة للمفردة.
3. تحديد استراتيجية انتقاء المفردة وعرض المفردات التالية.
4. اختيار أسلوب تقدير القدرة المناسب.
5. تحديد قاعدة إنهاء الاختبار.

أضاف ونج ولي [22] بإن الحاسوب يقوم باختيار المفردة وفقًا لما يلي:

1. قاعدة محددة، غالبًا ما تكون متوسطة الصعوبة، لتقديمها للممتحن في البداية.
2. ثم استخدامها في تقدير مستوى القدرة الأولي للفرد.
3. وبناء على المستوى الأولي يتم اختيار المفردة التالية لتقديمها للممتحن، وبعد ذلك يتم إعادة تقدير مستوى قدرة الفرد.
4. وتستمر هذه الخطوات حتى يتحقق محك يتم اختياره لإنهاء تطبيق الاختبار وتحديد القدرة النهائية للفرد ومستوى دَقَّة تقديرها.

IX. الدراسات السابقة:

1. دراسات أوصت باستخدام الاختبار الإلكترونية بصورة عامة:

تضمنت العديد من الدراسات السابقة التي أوصت بضرورة استخدام الاختبارات الإلكترونية وفعاليتها في العملية التعليمية منها:

a. فعاليتها في تطوير أساليب التقييم مثل:

- دراسة دماس [23]: التي أسفرت نتائجها على وجود موقف إيجابي عن الاختبارات الإلكترونية حيث أكدت الدراسة إن الاختبار الإلكترونية حققت رضا الطالب من حيث التصحيح الفوري والصلاحية والشفافية. وخلصت الدراسة إلى أن الاختبارات الإلكترونية ستصبح هي المفضلة لدى الطلبة وإنها سوف تكتسب فعاليتها في سياق أساليب التقييم..

b. دور استخدام الاختبارات الإلكترونية في تعزيز مصداقية النظام التعليمي مثل:

- **دراسة خميس [24]:** وقد أظهرت أن الاختبارات الإلكترونية تتمتع بدرجة عالية من الثقة والمصداقية بين الطلبة وأن هذا النوع من الاختبارات يحظى بقبول وتفضيل كبير.

2. دراسات أوصت باستخدام الاختبارات الإلكترونية التكيفي:

قامت بعض الدراسات والابحاث بتطبيق الاختبارات التكيفي الإلكترونية والتحقق من فاعليتها في العديد من المجالات منها:

a. فاعليتها في تقدير القدرة العقلية بالاعتماد على أسلوب تقدير القدرة بطريقة الارجحية القصوى مثل:

- **دراسة عودة وعبيدات [25]:** التي أسفرت نتائجها إلى زيادة دقة تقديرات القدرة من خلال الاختبارات التكيفية مقارنة بنظيرتها الخطية مع اختلاف أسلوب تقدير القدرة المستخدمين، وزيادة في كمية المعلومات لأسلوب بطريقة الارجحية القصوى.
- **دراسة نور الدين [26]:** التي أوصت بضرورة اعتماد الجامعات للاختبارات الإلكترونية التكيفية والتوسع في تحويل الاختبارات التحصيلية الورقية والخطية المحوسبة إلى اختبارات تكيفية محوسبة، والتحقق من فاعلية الاختبارات التكيفية المحوسبة في تقدير قدرة الأفراد في السياق الثقافي للبيئة العربية ومدى الاختلاف بين التقدير البيزي والأرجحية القصوى في تقدير قدرة الأفراد في الاختبارات التحصيلية.
- **دراسة فرحات [27]:** وجود فروق ذات دلالة إحصائية بين متوسطات الخطأ المعياري لتقدير القدرة للاختبار الخطي المحوسب والاختبار التكيفي المحوسب باستخدام أسلوب الاحتمال الأقصى أو بيزية لتقدير القدرة وقاعدة إنهاء الاختبار بعدد محدد من الفقرات أو بأقل خطأ معياري، وكانت النتائج لصالح الاختبار التكيفي المحوسب.

b. فاعليتها في قياس التحصيل المعرفي لطلبة مثل:

- **دراسة مارتين ولازيندي [10]** وجاءت نتائج البحث لصالح الاختبار التكيفي الإلكتروني والذي يعتبر أكثر دقة في قياس التحصيل.
- **دراسة أجراها أوزورت وأوزورت [28]:** وقد أكدت نتائج لاختبار إمكانية وفعالية استخدام أنظمة الاختبار التكيفي في المناهج التركية.
- **دراسة نور لدين [26]:** حيث أجرى بحثًا هدف إلى التعرف على مدى فاعلية القياس التكيفي المحوسب في قياس تحصيل طلبة الجامعة توصل أن الاختبار التكيفي المحوسب فعال على كل من الاختبار التقليدي الخطي أو الورقي في قياس تحصيل طلبة الجامعة.
- **دراسة جرجس دميانة [29]:** هدف الدراسة إلى قياس أثر الاختبارات التكويني التكيفي في تنمية التحصيل الدراسي المتعلق بالبرمجة لدى طلاب تكنولوجيا التعليم وقد أسفرت النتائج عن وجود فرق دال إحصائياً لصالح التطبيق البعدي.

c. فعاليتها في تخفيف مستوى القلق مثل:

- **دراسة جيمس[30] :** التي أجمع فيها الطلبة على دور الاختبارات المحوسبة في تخفيف مستوى القلق لديهم؛ إضافة لقلة التكاليف المادية عند عقد هذه الاختبارات .

ويمكن تلخيص ما توصلت إليه الدراسات السابقة بالنقاط التالية:

1. أكدت الكثير من الدراسات على إيجابيات اعتماد الاختبارات التكيفية بديلًا للاختبارات الخطية.
2. تأكد معظم الدراسات السابقة على عمل المزيد من الدراسات بشأن الاختبارات الإلكترونية التكيفية.
3. كما أكدت الدراسات السابقة على دور الاختبارات الإلكترونية التكيفية في رفع التحصيل الدراسي لدى الطلبة.
4. الفجوة البحثية تظهر الأدبيات محاولات لتطبيق فكر الاختبارات الإلكترونية التكيفية في البيئة العربية، وتعدّ تلك المحاولات قليلة إذا ما قورنت بنظيرتها في البيئة الأجنبية؛ لذلك أن ما يميز هذه الدراسة بأنه الأولى -في حدود علم الباحثين - التي تجرى في الجامعات اليمنية كون الاختبارات التكيفي حديثًا نسبيًا.

X. الطريقة والإجراءات:

منهج البحث:

ينتمي هذا البحث إلى فئة البحوث التطويرية التي تستخدم بعض مناهج الدراسات الوصفية في مراحل الدراسة والتحليل والتصميم والمنهج الشبة تجريبي في الكشف عن فاعلية المتغير المستقل المتمثل في الاختبارات الإلكترونية التكيفية على المتغير التابع تقييم القدرة المعرفية للطلبة في التحصيل الدراسي لدى طلبة جامعة العادل المستوى الأول، ويتمثل المنهج الشبة تجريبي في استخدام اختبارين والمقارنة بينهما للتحقق من نتائج الفرضيات.

A. متغيرات البحث:

أشتمل البحث على المتغيرات التالية:

https://doi.org/10.20428/jst.v30i6.2815

- **المتغير المستقل:** المتمثل في الاختبارات الإلكترونية التكيفية.
  ويشمل متغيرين فرعيين: (الجنس، التخصص)
- **المتغير التابع:** المتمثل في دور الاختبارات الإلكترونية التكيفية تقييم القدرة المعرفية وتعزيز التحصيل الدراسي لدى طلبة جامعة العادل.

B. المنهج التجريبي للبحث:

- استخدم هذا البحث التصميم الشبة تجريبي ذو المجموعتين المتكافئتين الذي يعتمد على مقارنة نتائج تقييم الطلبة عند استخدامهم الاختبار الإلكتروني الخطي وبعد استخدامهم الاختبار الإلكتروني التكيفي. ويبين الجدول الآتي التصميم التجريبي لهذا البحث:

جدول (1) التصميم التجريبي للبحث

المجموعة التجريبية	التطبيق القبلي	مادة المعالجة التجريبية	الأدوات
المجموعة التجريبية	اختبار الإلكتروني	اختبار الإلكتروني التكيفي	اختبار تحصيل
المجموعة الضابطة	خطي		

C. مجتمع البحث وعينته:

مجتمع البحث:

تكون مجتمع الدراسة من جميع الطلبة المسجلين والدارسين من طلبة جامعة العادل، وذلك للعام الجامعي 2024/2023م ويبلغ عددهم (988) طالبًا طالبة، بواقع (383) طالبًا و (605) طالبة خلال الفصل الدراسي الثاني.

عينة البحث:

بناءً على حجم مجتمع الدراسة من الذكور والإناث تم تقسيم عينة البحث إلى مجموعتين مجموعة ضابطة ومجموعة تجريبية تحوي كل مجموعة (100) طالبًا وطالبة من طلبة جامعة العادل المستوى الأول لجميع لكليات.

جدول (2) توزيع عينة البحث الأساسية

المجموعة	النوع	العينة	النسبة المئوية
المجموعة التجريبية الاختبار الإلكتروني التكيفي	ذكور	57	29%
	إناث	43	21%
المجموعة الضابطة الاختبار الإلكتروني خطي	ذكور	57	29%
	إناث	43	21%
المجموع	**200**		**100%**

D. إجراءات البحث:

أداة البحث:

تحقيقًا لأهدا ف الدراسة تم تطوير واستخدام أداة لتقييم طلبة جامعة العادل - جميع الكليات للمستوى الأول - عن طريقة استخدام الاختبار الإلكتروني التكيفي على التقييم المعرفي في التعليم، ودوره في رفع التحصيل الدراسي لدى طلبة جامعة العادل. وفي هذه الخطوة تم تحديد وبناء أداة البحث المتمثلة في: **الاختبار التحصيلي**.

E. ضبط أداة البحث: (الاختبار التحصيلي)

صدق الاختبار:

للتحقق من صدق الاختبار من تم وضع الاختبار في صورته الأولية حيث يتكون من (100) سؤال وعرضه على السادة المحكمين لإبداء الرأي من حيث : مناسبة الاختبار لقياس ما وضع لقياسه، مناسبة لأسئلة للمقرر المختار، مناسبة الأسئلة للمستويات المعرفية التي تم ذكرها مسبقًا، وضوح تعليمات الاختبار، وبعد تلقي التعليمات المحكمين ،تم إعادة صياغة بعض الأسئلة ،و تعديل بعض الأفعال للتناسب مع المستوى التي وضعت من أجله حيث تم حذف بعض الأسئلة من قبل السادة المحكمين ،وأصبح الاختبار مكون من (91) سؤال جاهز لتطبيقه على العينة الاستطلاعية.

ثبات الاختبار

من خلال تطبيق الاختبار على عينة تجريبيه عددها (65) طالب من طلبة كلية العلوم الأدرية تم اختيارهم بطريقة عشوائية من بين أقسام الكلية. حيث تتراوح معاملات الصعوبة للمفردات البنك بين (0.27-0.95) والذي يشير إلى مستوى مناسب.

تحديد زمن الاختبار:

تم تقدير الزمن من خلال جمع أزمنة إجابات الطلبة وقسمتها على العدد الكلي للعينة فكان الزمن اللازم تقريبًا (45) دقيقة، مع الإشارة إلى عدم احتساب الوقت المنقض في التوجيه وإعطاء التعليمات.

تصميم الاختبار والمقياس (أدوات البحث):

تحقيقًا لأهداف الدراسة تم تطوير واستخدام أداة لتقييم طلبة جامعة العادل- جميع الكليات للمستوى الأول- عن طريقة استخدام الاختبارات الإلكترونية التكيفية ودورها في تقييم القدرة المعرفية للطلبة في التعليم، وفي هذه الخطوة تم تحديد وبناء أداة البحث المتمثلة في: تصميم اختبار تحصيلي الإلكتروني. وفقًا لطبيعة البحث الحالي والتي تستهدف قياس فاعلية نمط الاختبار الإلكتروني التكيفي ومقارنة بنمط الاختبار الإلكتروني الخطي.

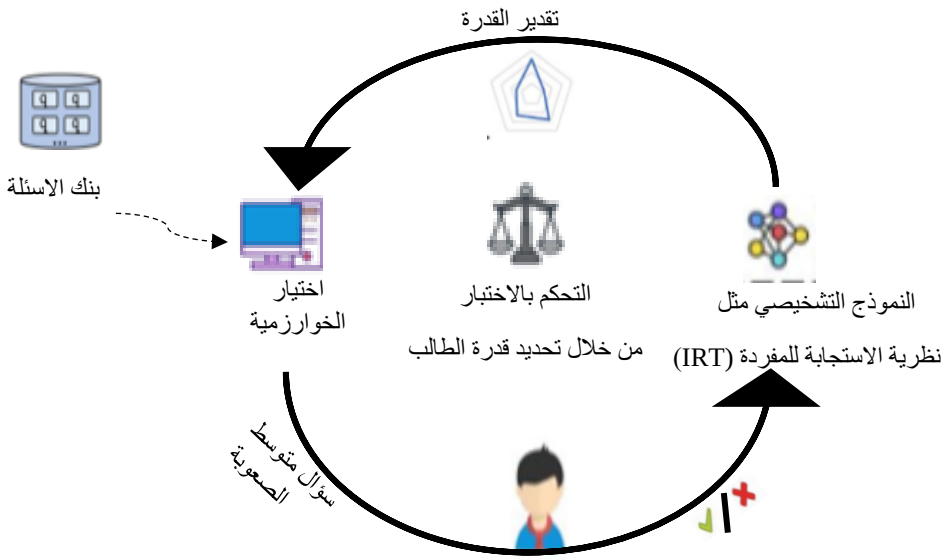
وبناء على ذلك تم تصميم النمطين:

a. تصميم الاختبار الإلكتروني التكيفي.
b. تصميم الاختبار الإلكتروني خطي.

أولًا: تصميم الاختبار الإلكتروني التكيفي.

1. مراحل تصميم الاختبار الإلكتروني التكيفي:

من خلال الاطلاع على الأدبيات ومنها ما أشار إليه كلًا من[19، 13] التي لخص مجموعة من الخطوات الأساسية التي يمر بها الاختبار الإلكتروني التكيفي. وقد تم وضع المعايير التي يقوم عليها تصميم الاختبار الإلكتروني التكيفي وفقًا لمخطط آليات السير البرنامج الذي أشار إليه ليو وآخرون كما هو موضح في الشكل (3):



رقم الشكل 3: مخطط بآليات سير البرنامج

ومن خلال ما تم عرضها أعلاه تم تحديد المراحل التي يمر بها تصميم الاختبار الإلكتروني التكيفي وهي كالتالي:

المرحلة الأولى: بناء مستودع (بنك)الاسئلة:

تم بناء بنك أسئلة يحوي (91) سؤالًا جاهز لتطبيقه على العينة بعد أن تم تقنين الأسئلة وتعديلها من قبل السادة المحكمين والتأكد من مناسبة الاختبار لقياس ما وضع لقياسه، مناسبة لأسئلة للمقرر المختار، مناسبة الأسئلة للمستويات المعرفية. وذلك كما أوصى به عدد من الباحثين منهم:

- علام [31] الذين أوصوا بكتابة أو جمع عدد كافٍ من المفردات الاختبارية تقيس مدى متسعًا من القدرة المعينة، وهذا يتطلب البدء بتحديد موصفات كل مفردة، ومحتواها، ونطاق المهارات التي تُمثلها، وكلما كانت هذه المواصفات واضحة، أمكن بناء عدد كافٍ ومتنوع من المفردات المتباينة في صعوبتها على مدى القدرة أو السمة المراد قياسها، كما ينبغي مراجعتها بدقة للتحقق من جودتها وعدم تحيزها لجماعة معينة دون غيرها.
- النظير الذي أكد أيضًا أن أهم مكون من مكونات الاختبار الإلكتروني التكيفي هو تحديد بنك الأسئلة حيث يمكن من خلاله تطبيق الاختبار التكيفي، ويجب أن يحتوي الاختبار على مجموعة كافية من الفقرات التي تتوزع صعوبتها حسب مستوى وخصائص الطلبة[32] .

حيث تم إجراء اختبار قبلي لتحديد درجة صعوبة كل سؤال من الأسئلة وتقسيم بنك الأسئلة إلى ثلاث مستويات (مستوى الأسئلة الصعبة – المتوسطة - السهلة) وتحديد مستوى الصعوبة لكل سؤال.

المرحلة الثانية: اختيار نموذج التحليل المناسب من نماذج نظرية الاستجابة للمفردة.

قبل أن نقوم بتطبيق الاختبار التكيفي لابد من اختيار أحد النماذج الرياضية المرتبطة بنظرية الاستجابة للمفردة.

نماذج الاستجابة للمفردة:

هناك عدد من النماذج في نظرية الاستجابة للمفردة LRT models لتحليل بيانات المفردات الاختبارية؛ غير أن هناك ثلاثة نماذج شائعة الاستخدام في البيانات ثنائية التقسيم التي تقدر بالدرجة (1) للإجابة الصواب، (0) للإجابة الخطأ، وهي النماذج (أحادية، وثنائية، وثلاثية) البارامتر، والتي أشار إليها عدد من الدراسات منها ما ذكرها التقي [33] حيث أظهرت العديد من الدراسات ومنها دراسة السباحة نتائج أوصت بضرورة استخدام نماذج الاستجابة للفقرات أو نموذج راش في بناء وتقنين مقاييس الاتجاه.

وقد تم استخدام نموذج راش (أحادي البارامتر) وهو النموذج الذي اتصل بالعينات صغيرة الحجم حيث يتطلب كحد أدنى (٢٠) مفردة و(٢٠٠) فرد. الذي يتميز ببساطته عن غيره من النماذج المتعلقة بنظرية استجابة الفقرة، لأنه يحتوي على معلمة واحدة فقط وهي الصعوبة، وبالتالي فهو يتجنب التعقيدات الإحصائية المتعلقة بعدة بارامترات التي تتعامل معها النماذج الأخرى، وبالتالي فهو يفسر الأداء وفق أقل عدد من المتغيرات، مما جعله أكثر استخداماً بين الباحثين.

المرحلة الثالثة: تحديد نقطة البداية:

وهو عبارة عن سؤال متوسط الصعوبة يتم اختياره عشوائيًا من بنك الأسئلة (من مستوى الأسئلة المتوسطة الصعوبة)، ومن خلال إجابة الطالب على السؤال يتدرج في الأسئلة التي تتناسب مع قدرته، وقد تم الاعتماد على أسلوب الأرجحية العظمى (أقصى معلومات):وهو بعد أن يتقدم المفحوص للفقرة أو الفقرات الأولية يتم تقدير القدرة له ثم يتم اختيار الفقرة ذات المعلومات الأكبر عند مستوى القدرة المقدر الحالي للمفحوص التي لم يتم تقديمها سابقًا وبعدها يتم تصحيحها وإيجاد تقدير جديد لقدرة المفحوص ثم يتم اختيار الفقرة التالية ذات المعلومات الأكبر عند مستوى القدرة الجديد وهكذا... حتى يتم إنهاء الاختبار. بناء إلى ما أشارت إليه عدد من الدراسات مثل دراسة كلًا من[26، 27،28،34،35] الذين أكدوا على فاعلية الاختبارات التكيفية المحوسبة في تقدير القدرة باستخدام طريقة الأرجحية العظمى.

المرحلة الرابعة: أساليب اختيار المفردات في الاختبارات الإلكترونية التكيفية:

خوارزمية اختيار المفردة التالية: تم اختيار خوارزمية لورد حيث تم تقدير قدرة الطالب المبدئية بالاستفادة من إجابتها على السؤال المتوسط الصعوبة، ثم تم جمع صعوبة السؤال مع أعلى صعوبة متوقعة.

حيث تم تصميم خوارزمية الاختيار للفقرات بطريقة بسيطة جدًا وواضحة، يتم تقدير قدرة أحد المتقدمين للاختبار ليكون له قياس قدرة معين واختياره باستخدام سؤال مستوى صعوبته يعادل مستوى قدرته، وإذا نجح في الاختيار يتم رفع مستوى تقدير المقدرة، وإذا خطأ في الإجابة عن هذا السؤال يتم خفض مستوى تقدير القدرة، وتقديم سؤال آخر إليه، وتكرر هذه العملية حتى يتم تقدير القدرة النهائي.

و تم استخدام إحدى استراتيجيات القياس التكيفي في اختيار المفردة التالية والتي أكد عليها كلًا [36] بأن المفردات التالية يتم اختيارها بما يتناسب مع قدرة كل ممتحن، حيث يتم اختيار درجة صعوبتها بحيث لا تكون صعبة جدًا أو سهلة جدا عليه، بمعنى أن احتمال إجابته عليها بشكل صحيح هو (50%)، ومع كل إجابة على المفردات التالية يصبح الحاسوب قادراً بشكل متزايد على تقييم معرفة الممتحنين وتقدير قدرته، لتحديد أفضل المفردات في التطبيق بعد ذلك، ولذلك فإن الاختبار التكيفي الإلكتروني هو نوع من الاختبارات التي تم تطويرها لزيادة كفاءة عملية تقدير معرفة الممتحنين، من خلال تعديل المفردات المقدمة لهم بناءً على إجاباتهم السابقة أثناء الاختبار.

وهذا ما أوضحت دراسة وايس بأن يتم تحديد العنصر التالي في المراحل الأولية من الاختبارات التكيفية المحوسبة من خلال قاعدة الخطوة. إذا تم الإجابة على العنصر الأول بشكل صحيح، يتم زيادة مقدار القدرة الأصلي السابق بمقدار، على سبيل المثال (0.50)، وإذا تم الإجابة عليه بشكل غير صحيح، يتم تقليل مقدار القدرة الأصلي السابق بنفس المقدار. بعد قياس كل عنصر وتسجيله، تُستخدم القدرة الجديدة لتحديد العنصر التالي، والذي يتم اختياره من بين جميع العناصر الموجودة في بنك الأسئلة التي لم يتم قياسها لهذا المختبر، من خلال تحديد العنصر الذي يوفر أكبر قدر من المعلومات للقدرة الحالية[37] .

وفقا لما أشار اليها الباحثين وبناء على ما سبق تم تحديد أليه انتقاء الفقرة التالية في البرنامج المقترح: حيث تم استخدام طريقة جمع صعوبة السؤال مع أكبر صعوبة إذا اجاب صح، والتي أكدت عليها بعض الدراسات ودعمت صحتها مثل دراسة كلًا لي وآخرون[38] حيث أوضحوا بأن هذه الطريقة تعطي تقديرات دقيقة بنفس القدر من الطرق التقليدية بالإضافة إلى ذلك، فإنها بسيطة وسهلة التنفيذ. هذا ما جعلها خيارًا جذابًا للاختبارات التكيفية. تعتمد على طرح سؤال متوسط الصعوبة ثم وفقًا لإجابة المتعلم يتم اختيار المفردة التالية وفقا للمعادلات التالية:

إذا كانت الإجابة صحيحة ينتقل الطالب إلى المفردة الأكثر صعوبة في نفس موضوع التعلم من خلال المعادلة التالية:

الإجابة صحيحة = (أعلى درجة صعوبة + درجة صعوبة السؤال السابقة) /2 (1)

وإذا كانت الإجابة خطأ ينتقل الطالب إلى السؤال الأقل صعوبة من خلال المعادلة التالية:

الإجابة خطأ = (أقل درجة صعوبة + درجة صعوبة السؤال السابقة) /2 (2)

وهكذا يستمر الطالب بين الانتقال بين الأسئلة حسب القاعدة المحددة حتى يصل إلى قاعدة الإنهاء، وقد استخدم البحث الاستراتيجية المحوسبة التي أشار إليها دعنا [39] في اختيار الفقرة التالية التي تقدم للطالب بعد تحديد مستوى القدرة لديهم حيث يقدم فقرات متتالية الصعوبة طبقا لصيغة رياضية تسمى دالة الخطوة التي تعني ان صعوبة الفقرة التي ترتيبها (ن+1) يساوي صعوبة الفقرة التي ترتيبها ن مضاف اليه أو مطروح منه قيمة معينة تسمى حجم الخطوة، واستخدم البحث قاعدة الإنهاء بأن تصل الدرجة المتوقعة للطالب ≤ 20

المرحلة الخامسة تحديد قاعدة إنهاء الاختبار:

بعد الاطلاع على عدد من الدراسات التي أوضحت عدد من الطرق يمكن من خلالها أنهاء الاختبار منها:

1. وفقًا لعدد محدد من المفردات.
2. وفقًا للخطأ المعياري.
3. عندما يكون قياس القدرة بعيدا جدا عن محك النجاح.
4. وفق درجة معيارية يتم تحديدها مسبقًا.

وقد أشار كلًا من مارتين وليزيندس [10]يمكن تمثل قاعدة الإنهاء المتمايز في عدد المفردات التي يحصل عليها الممتحنون، ويمكن أن تكون ذات الطول الثابت، والطول المتغير هما طريقتان لتحديد زمن وكيفية إنهاء الاختبار:

- يتطلب اختبار الطول الثابت أن يجيب جميع الممتحنين على نفس العدد من المفردات.
- أما اختبار الطول المتغير فيجيب الممتحنون على أعداد مختلفة من المفردات حتى يتم استيفاء مستوى تقدير القدرة في زمن محدد، أداء الممتحن بنفس درجة الاحكام.

وقد إشارة البياضة[4] إلى مجموعة من القواعد التي يمكن من خلالها إنهاء الاختبار التكيفي ومنها:

- التوقف عند استنفاذ الفقرات الموجودة في بنك الأسئلة وذلك في حالة البنوك الصغيرة.
- التوقف عند عدم كفاية الفقرات.
- في حالة الاختبارات الإلكترونية التكيفية عندما تكون إجابة المفحوص سريعة جدًا، أو بطيئة جدًا مما يدل على عدم جدية الممتحن.
- الوصول إلى الحد الأقصى لطول الاختبار المحدد سلفًا.
- الوصول الى مستوى معياري معين، يتم ايقاف الاختبار بمجرد وصول الطالب إلى الحد الأدنى للأداء المقبول.

وبناءً على ما سبق تم باختيار قاعدة لإنهاء وفق محك الطول المتغير كما أوصى بها ليروكس [40] والذي يعتمد على الوصول لمستوى معين من دِقة التقدير لقدرة المفحوص، ووفقًا لهذا المحك لا يتعرض جميع المفحوصين لنفس العدد من المفردات، ولكن بالقدر الذي يضمن تحقيق مستوى الدقة المحدد. حيث يتم ايقاف الاختبار بمجرد وصول الطالب إلى الحد الأدنى للأداء المقبول. وقد تم تحديدها بالوصول الطالب إلى مستوى معياري محدد، أي عندما تصل درجة الطالب الحقيقية ≤ 20كقاعدة لإنهاء الاختبار لتحديد قدرة المتعلم.

حساب الدرجة النهائية:

- ويتم حساب الدرجة النهائية للطالب من خلال الآتي:

- الدرجة المتوقعة: هي مجموع درجات الأسئلة المقدمة إلى الطالب على احتمال أن يجيب عليها جميعا بشكل صحيح.
- الدرجة الحقيقية: هي مجموع درجات الأسئلة التي أجاب عليها الطالب بشكل صحيح.

- حساب الدرجة الحقيقية في نموذج راش:

لا يوجد طريقة واحدة محددة لحساب الدرجة الحقيقية للطالب في نموذج راش، وقد تم استخدام طريقة "استخدام عدد الأسئلة التي أجاب عليها الطالب بشكل صحيح "حيث يتم فيها التالي:

- يتم حساب عدد الأسئلة التي أجاب عليها الطالب بشكل صحيح.
- يتم تحويل عدد الأسئلة الصحيحة إلى درجة على مقياس محدد.
- من المهم أن تكون طريقة حساب الدرجة الحقيقية للطالب عادلة لجميع الطلاب.
- يجب أن تأخذ طريقة الحساب بعين الاعتبار صعوبة الأسئلة التي أجاب عليها الطالب.
- حساب الدرجة الحقيقية للطالب = درجة صعوبة السؤال x الدرجة المخصصة للسؤال x احتمالية الإجابة (1,0) (3)
- حساب الدرجة المتوقعة للطالب= درجة صعوبة السؤال x الدرجة المخصصة للسؤال.

لإيجاد الدرجة النهائية استخدم البحث طريقة المجموع المرجح يتم حساب متوسط الدرجة المتوقعة والدرجة الحقيقية

- ولإيجاد الدرجة النهائية = (4)

نسبة الدرجة النهائية xالدرجة الكلية المحددة من قبل المعلم. (5)

https://doi.org/10.20428/jst.v30i6.2815

حيث تم برمجة الاختبار بحيث يختلف عرض الاسئلة الموضوعة من مختبر اخر.

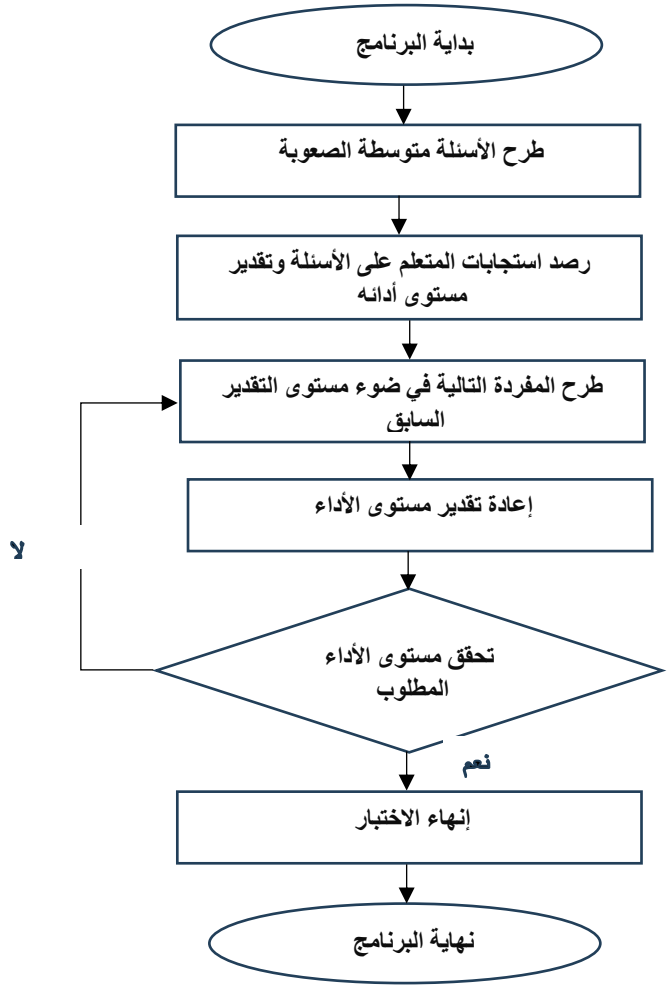
وأين كانت طريقة التوقف التي تم اختيارها فإن الاختبار التكيفي لا يقف إلا في الحالات التالية:

1. أن يصل الطالب إلى الحد الأدنى من عدد الأسئلة بحيث لا يشعر الطالب بأنه لم يتم قياسه بشكل دقيق إلا إذا أجاب على الأقل على 10 أو 20 سؤالاً. حيث روعي عند تصميم البرنامج هذه الخاصية.
2. حيث يمكن إنهاء الاختبار الإلكتروني لتكيفي وفقًا لعدد محدد من المفردات تم تحديدها ب (10) أسئلة من نوع الأسئلة الصعبة بدون الحاجة إلى عرض جميع فقرات الاختبار للطلبة ذو القدرة العالية وهذا أهم ما يميز هذا النوع من الاختبارات.
3. يغطي الاختبار جميع المواضيع، بمعنى أن الأسئلة المقدمة للطالب تمثل جميع مجالات الأسئلة التي يغطيها الاختبار، وخاصة في الاختبارات التي تشمل أكثر من مجال.

المرحلة السادسة تنفيذ الاختبار:

بناءً على ما سبق من خطوات لتصميم لاختبار تم تنفيذ الاختبار على عينة المتمثلة بطلبة جامعة العادل حيث تم تنصيب الاختبار في معمل الجامعة.

يمكن تلخيص الخطوات السابقة من خلال المخطط التالي الذي يوضح آليات عمل البرنامج حتى يصل بقواعد إنهاء الاختبار التكيفي كما هو موضح في الشكل (4).



رقم الشكل 4: مخطط يوضح

2. مميزات الاختبارات التكيفية:

من خلال تطبيق الاختبار التكيفي على عينة الدراسة – طلبة جامعة العادل – أظهر البحث عدة مميزات للاختبارات التكيفية مقارنة بالاختبارات الخطية أو التقليدية (اختبارات الورقة والقلم) وقد تطابقت تلك المميزات مع ما ارصدته الأدبيات، منها ما أشار إليه Linacre [41] ومن تلك المميزات التالي:

- سهولة إدارة المفردات الاختبارية.
- إتاحة درجة الفرد والتغذية الراجعة عن الأداء بشكل فوري.
- الحفاظ على مستوى دافعية المفحوص، حيث لا يتعرض لعدد كبير من المفردات.
- خفض معدلات القلق لديه كما أكد عليها عدد من الدراسات مثل [3، 28].

– بالإضافة إلى النقطة الأهم وهي دقّة تقدير مستوى القدرة للمفحوص أو مستوى السمة لديه.

– تمتع بأمان عال، وتجرى عند الطلب، ولا تحتاج إلى ورقة إجابة، وتلائم قدرة المختبر، ومرنة في اختيار المفردة، ومستوى تقنينها أعلى من الاختبار الخطي، ومدة المراقبة فيها قصيرة، وتعد فيها التقارير بصورة فورية ودقيقة وهذا ما أشار إليه[42،36] .

– كما تقوم بتوفير معلومات إحصائية عن أداء الممتحن في كل مفردة من مفردات الاختبار
حيث يمكن للطالب معرفة عدد الأسئلة التي أجاب عليها ودرجة صعوبة كل سؤال، والاجابة الصحيحة للسؤال في حالة إذا كان أجاب بشكل خاطئ على السؤال.

– تقدم عدد أقل من المفردات للحصول على المستوى المطلوب من الدقة والموثوقية مقارنة بالاختبارات الإلكترونية التقليدية. وقد أكد ذلك كلًّا من Huo, Magis and Raiche [44, 43].

3. مشكلات الاختبارات الإلكتروني التكيفي:

من خلال الاطلاع على بعض الدراسات والأبحاث السابقة التي أشارت إلى بعض المشكلات والعيوب للاختبارات الإلكتروني التكيفي منها دراسة علام [31] تم تلخيص تلك المشكلات كما يأتي:

1. عدم الاتفاق حول طريقة اختيار السؤال من بنك الأسئلة ما قد يؤدي إلى نتائج متباينة وقد تكون غير دقيقة.
2. تتطلب إجراءات تطبيقها إنشاء بنك أسئلة يشتمل على عدد كبير من المفردات تختزن في الحاسوب.
3. تحتاج المشكلات المتعلقة بتأثير سياق عرض المفردات وترتيبها، وتوازن محتواها إلى مزيد من الدراسات والبحوث التطويرية.
4. يتطلب أداء الاختبار التكيفي توافر بنية أساسية تدعم الاتصال بالشبكات المحلية وشبكة الإنترنت.
5. أن المختبر الذكي قد يجيب إجابة خاطئة عن الأسئلة الأولية بشكل متعمد، وسيفترض البرنامج بعد ذلك أن لديه قدرة منخفضة فيختار له سلسلة من الأسئلة السهلة.

4. معالجة مشكلات الاختبار الإلكتروني التكيفي:

تم التغلب على بعض المشكلات السابقة من خلال برنامج الاختبارات الإلكترونية التكيفية التي تم إنشاءه من خلال الآتي:

– تم التغلب على المشكلة الأولى من خلال: اختيار السؤال الأولي من بنك الأسئلة وهو سؤال متوسط الصعوبة حيث أكد على هذه الطريقة عدد من الدراسات: منها [25 , 38 ,27].

– تم التغلب على المشكلة الثانية من خلال: أنشاء بنك أسئلة يحوي (91) سؤال تم تقنينه من قبل المحكمين. حيث أشار عدد من الدراسات بأنه يمكن أنشاء بنك أسئلة مكون (30) سؤلًا يعد كافيًا ومنها دراسة نور لدين [26]: حيث أجرى بحثًا هدف إلى التعرف على مدى فاعلية القياس التكيفي المحوسب في قياس تحصيل طلبة الجامعة وقام بإنشاء بنك أسئلة مكونًا من (48) سؤالًا، ودراسة Istiyono et al. [42] تألفت أدوات البحث من بنك للأسئلة يحتوي على (62) فقرة.

– وقد تم التغلب على المشكلة الثالثة من خلال: اختيار خوارزمية لانتقاء المفردة التالية تم الإشارة إليها سابقًا.

– وقد تم التغلب على المشكلة الرابعة من خلال: تم تصميم برنامج مجاني بدون الحاجة إلى استخدام الإنترنت لتطبيق الاختبارات التكيفية.

– وقد تم التغلب على المشكلة الخامسة من خلال: وتم التغلب على مشكلة المختبر الذكي قد يجيب إجابة خاطئة عن الأسئلة الأولية بشكل متعمد من خلال إيجاد الدرجة المتوقعة والدرجة الحقيقية.

ثانيًا: الاختبار الإلكتروني الخطي:

تم تصميم اختبار الإلكتروني خطي مكونًا من (15) سؤالًا من نمط الاختيار من متعدد – الصواب والخطأ ذي خصائص سيكومترية جيدة مشتق من بنك الأسئلة.

– تم تقدير أداء المتعلم عن طريقة عرض الأسئلة على المتعلم بشكل خطي.

– تحديد درجة كل سؤال على حسب مستوى الصعوبة الخاص به، حيث تم تقسيم الأسئلة على حسب مستوى الصعوبة الى (أسئلة صعبة-متوسطة-سهلة).

وعند تصميم الاختبارين تم الالتزام بالآتي:

- **هدف الاختبار:** هدف الاختبار تقييم طلبة جامعة العادل للمستوى الأول من خلال بناء نوعين من الاختبارات الإلكترونية (الاختبار الإلكتروني الخطي - الاختبار الإلكتروني التكيفي) والمقارنة بينهم وفقا للمحتوى التعليمي المقرر.
- **تحديد مفردات الاختبار:** من خلال جدول المواصفات التي إعداده سابقًا.
- **تصميم بنك الاسئلة:** تم تصميم بنك أسئلة يحوي (91) سؤال من نوع أسئلة الصواب والخطأ والاختيار من متعدد حيث روعي عند بناء الأسئلة صياغتها صياغة لغوية سليمة، وتم وضع البرنامج في Server، وتم ربط بقية الأجهزة به بحيث نستطيع مراقبة سير الاختبار وتصحيح الأسئلة ألياً، علمًا بأن النظام يتيح التحكم في أنواع الأسئلة ومستوى صعوبته وتحديد الاستجابات والتفاعلات، وتم تحديد خصائصها السيكومترية.
- وقد تم بناء قاعد بيانات عن طريق برنامج الاكسس يحتوي على رقم السؤال ونوعه ومستوى الصعوبة.

المعالجة الإحصائية:

تمر الدراسة الحالية بمجموعة مختلفة من التحليلات الإحصائية من أجل الإجابة على الأسئلة البحثية وهي:

أ- التحقق من افتراضات نظرية الاستجابة للمفردة:

– التحقق من افتراضية أحادية البعد:

للتحقق من نظرية الاستجابة المفردة تم اخضاع بنك الأسئلة للتحليل العاملي بطريقة المكونات الأساسية باستخدام برنامج spss25 إذ تعد المكونات الأساسية الناتجة عن التحليل العاملي مؤشرًا مقبولًا احاديًا عندما لا تقل نسبة المؤشر للعامل المفسر بالتراكمي عن (20%) للعوامل المستخلصة وفق مقياس leckase.
وذلك للتعرف على مدى تحقق أحادية البعد كونها أحد الافتراضات الأساسية التي تقوم عليها نظرية الاستجابة للمفردة.

جدول (3) قيم التباين المفسر للعوامل الثمانية مقارنة للتفسير التراكمي للعامل الأخير

العامل	نسبة التباين المفسر	نسبة التباين التراكمي المفسر
1	15.286	15.286
2	9.816	25.102
3	7.692	32.794
4	8.878	41.673
5	8.297	49.970
6	4.539	54.509
7	6.569	61.078
8	5.444	66.523

المتوسط	8.315125	43.36688

ثم نجد المتوسط لكلا العمودين ثم نطرح المتوسط التباين التراكمي من التباين المفسر يكون الناتج (35.05%) أي أن بنك الأسئلة يحقق محك leckase.

– **التحقق من افتراض الاستقلال الموضعي:**

ويقصد الاستقلال الموضعي أن احتمال الإجابة الصحيحة للسؤال يكون مستقل عن ناتج إجابته وعن أي سؤال أخرى في الاختبار عند ضبط كل من القيمة التقديرية لقدرته وصعوبة السؤال وعليه فأن الاستقلال الموضعي يعني عدم ترابط الأسئلة في الاختبار وهذا ما هدفت إليه النظرية الحديثة. حيث يمكن التحقق من افتراض الاستقلال الموضعي من خلال مؤشر (Q3) وهوعن معاملات الارتباط بين البواقي لكل زوج من أزواج مفردات الاختبار، ويتحقق الاستقلال الموضعي إذا كانت جميع معاملات الارتباط بين أزواج مفردات الاختبار أقل من نقطة القطه وهي القيمة المطلقة للعدد (0.2)، وكان متوسط قيم مؤشر(Q3) أقل من القيمة المتوقعة للمتوسط الحسابي للتوزيع العيني لهذا المؤشر.

– **التحقق من افتراض التحرر من السرعة:**

لتحيق هذا الفرض تم إعطاء الوقت الكافي للطلاب للإجابة على أسئلة الاختبار بحيث لا يؤثر عامل السرعة على أدائهم بحيث يمكن ارجاع إجاباتهم الخطاء إلى قدرتهم المحدودة وليس إلى عدم كفاية الوقت. حيث أنهى جميع الطلاب الإجابة على الاختبار ضمن الوقت المحدد للإجابة وهو ساعة.

ب- التحقق من فرضيات البحث:

يسعى البحث للإجابة على الفرضيات التالية:

- لا توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي تقدير القدرة والخطأ المعياري لتقدير القدرة (كل على حده) للاختبار التكيفي والاختبار الخطي باستخدام طريقة الأرجحية القصوى لتقدير القدرة.
- لا توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي درجات طلبة المجموعة التجريبية التي تم تقييمها باستخدام الاختبار التكيفي والمجموعة الضابطة التي تم تقييمها باستخدام الاختبار الخطي.
- لا توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي درجات طلبة المجموعة التجريبية التي تم تقييمها باستخدام الاختبار التكيفي وبين المجموعة الضابطة التي تم تقييمها باستخدام الاختبار الخطي تعزى لمتغير الجنس (ذكور/أناث).
- لا توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي درجات طلبة المجموعة التجريبية التي تم تقييمها باستخدام الاختبار التكيفي وبين المجموعة الضابطة التي تم تقييمها باستخدام الاختبار الخطي تعزى إلى متغير التخصص.

3. عرض النتائج والمناقشة:

عرض ومناقشة نتائج البحث وتفسيرها:

سؤال البحث الرئيسي

- **ما فاعلية تصميم نظام اختبارات إلكترونية تكيفية على تقييم القدرة المعرفية وتعزيز التحصيل الدراسي؟**

تمت الإجابة عن سؤال البحث الرئيسي عن طريق الإجابة عن أسئلة البحث الفرعية الثلاثة التالية:

1. **ما معايير بناء وتصميم الاختبارات التكيفي الإلكترونية؟**
2. **ما الفروق بين الاختبارات الخطية والاختبار التكيفي المحوسب في متوسط تقدير القدرة؟**
3. **ما الفروق بين الاختبارات الخطية والاختبار التكيفي المحوسب في تعزيز التحصيل الدراسي؟**

❖ **الإجابة على سؤال البحث الأول:**

بعد الانتهاء من كتابة الإطار النظري للبحث وأيضًا بعد الانتهاء من بناء الاختبار وضبط المعايير يكون قد تم الإجابة على السؤال الأول.

❖ **الإجابة على سؤال البحث الثاني:**

استخدم البحث طريق الأرجحية القصوى لتقدير القدرة وقاعدة أنهاء تم تحديدها بدرجة معيارية، وخطأ معياري وقدره (0.25)، وللإجابة على هذا السؤال تم اختبار صحة الفرضية القائلة:

- **الفرضية الأولى:**

" لا توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي تقدير القدرة والخطأ المعياري لتقدير القدرة (كل على حده) للاختبار التكيفي والاختبار الخطي باستخدام طريقة الأرجحية القصوى لتقدير القدرة "

من خلال اختبار T لعينتين مرتبطين كما يتضح من الجدول (4)

جدول (4) يبين اختبار T لعينتين مرتبطين

الاختبار	القياس	المتوسط الحسابي	الانحراف المعياري	قيمة t	مستوى الدلالة	النتيجة الإحصائية
تقدير القدرة	الخطي	5.29	1.54	0.06	0.980	غير دالة
	التكيفي	10.57	3.52			
الخطأ المعياري لتقدير القدرة	الخطي	0.08	0.01	9.98	0.000	دالة
	التكيفي	0.29	0.001			

حيث يتضح من الجدول (4) الآتي:

1- أن قيمة ت المحسوبة للفروق بين متوسطي تقدير القدرة بين للاختبار الخطي والاختبار التكيفي كانت (0.06) وكانت مستوى الدالة لهما (0.98) وهي أكبر من مستوى الدالة (0.05) بالتالي تكون الفرضية غير دالة احصائيا، بالتالي قبول الفرضية الصفرية:

- لا توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي تقدير القدرة بين الاختبار التكيفي والاختبار الخطي باستخدام طريقة الأرجحية القصوى لتقدير القدرة. **أي ان مستوى القدرة بين الاختبارين لا فروق بينهما.**

2- أن قيمة ت المحسوبة للفروق بين مستوى الخطأ المعياري لتقدير القدرة بين الاختبار الخطي والاختبار التكيفي وهو (9.97) ومستوى الدالة (0.00) وهي أقل من مستوى الدالة (0.05) بالتالي تكون الدالة، بالتالي رفض الفرضية الصفرية وقبول البديلة:

- توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي الخطأ المعياري لتقدير القدرة بين الاختبار التكيفي والاختبار الخطي باستخدام طريقة الأرجحية القصوى لتقدير القدرة لصالح الاختبار التكيفي. **وقد أتفق ما تم التوصل مع إليه مع عدد من الدراسات منها:**

- دراسة نور الدين [26] حيث أتتفق إلى عدم وجود فروق ذات دلالة إحصائية بين متوسطي تقدير القدرة التحصيلية للاختبار الخطي الورقي والاختبار التكيفي المحوسب باستخدام طريقة الأرجحية القصوى لتقدير القدرة وتحديد قاعدة إنهاء الاختبار بـ (20) مفردة، كما أن متوسط الخطأ المعياري لتقدير القدرة للاختبار الإلكتروني التكيفي كان أقل بصورة دالة عند مستوى دلالة (0,05) عن متوسط الخطأ المعياري لتقدير القدرة للاختبار الخطي الورقي.دالة عند مستوى دلالة (0,05) عن متوسط الخطأ المعياري لتقدير القدرة للاختبار الخطي الورقي.
- دراسة. Samsudin et al [35] الذي توصل إلى كفاءة الاختبار التكيفي المحوسب في قياس أداء طلاب الصف الثامن على الاختبارات الدولية في مادة العلوم (TIMSS)استخدام طريقة الأرجحية القصوى لتقدير القدرة وتحديد قاعدة إنهاء الاختبار بـ (20) مفردة.
- ودراسة Ozyurt and Ozyurt [28] الذي توصل إلى أن الاختبارات التكيفي المحوسبة والتي تم إعدادها في وحدة الاحتمالات من مادة الرياضيات لطلاب الصف الحادي عشر تتمتع بثبات مرتفع باستخدام طريقة الأرجحية القصوى لتقدير القدرة وتحديد قاعدة إنهاء الاختبار بـ (15) أو (20) مفردة.

وقد دلت النتائج السابقة على أن:

استخدام طريقة الأرجحية القصوى لتقدير القدرة في للاختبار الإلكتروني التكيفي يعطي نفس مستوى تقدير القدرة التي يعطيها الاختبار الخطي إلا أن:

- الاختبار الإلكتروني التكيفي أكثر دقة في تقدير القدرة لأنه يقدم أخطاء معيارية أقل من الاختبار الخطي.
- تعتمد طريقة الأرجحية القصوى على اختيار المفردات في المراحل الأولية الاختبار الإلكتروني التكيفي من خلال قاعدة الخطوة، حيث تحدد مقدار الانتقال على متصل القدرة لتحديد قدرة المختبر لإعطائه مفردة جديدة تتناسب مع القدرة الجديدة له حتى يتم الوصول إلى نمط الاستجابة الصحيحة والخاطئة.
- **الإجابة على سؤال البحث الثالث:**

استخدم البحث طريق الأرجحية القصوى لتقدير القدرة وقاعدة أنهاء تم تحديدها بدرجة معيارية، وخطأ معياري وقدره (0.25)، وللإجابة على هذا السؤال تم اختبار صحة الفرضيات القائلة:

- **الفرضية الثانية:**

" لا توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي درجات طلبة المجموعة التجريبية التي تم تقييمها باستخدام الاختبار التكيفي والمجموعة الضابطة التي تم تقييمها باستخدام الاختبار الخطي".

ولتحقق من نتيجة هذه الفرضية تم استخدام اختبار t لعينتين مرتبطتين بين متوسطي درجات طلبة المجموعة التجريبية التي تم تقييمها باستخدام الاختبار التكيفي والمجموعة الضابطة التي تم تقييمها الاختبار الخطي كما يتضح في جدول(5):

الجدول (5) اختبار t لعينتين مرتبطتين

الاختبار	المتوسط الحسابي	الانحراف المعياري	الخطأ المعياري	قيمة t	درجة الحرية	مستوى الدلالة
التكيفي	6.62	1.54	0.11	5.308		
الخطي	6.09	1.81	0.07		199	0.000

حيث يتضح من الجدول (5) الآتي:

أن المتوسط الحسابي لدرجات طلبة المجموعة التجريبية التي تم تقييمها باستخدام الاختبار التكيفي بلغت (6.62) والمتوسط الحسابي للمجموعة الضابطة التي تم تقييمها باستخدام الاختبار الخطي بلغت (6.09) وكانت الفروقات لصالح الاختبار التكيفي بفارق (0.53).

وأن قيمة ت المحسوبة لعينين مرتبطتين بلغت (5.308) وهي أكبر من الجدولية وفق درجة الحرية (199) كون عدد العينة (200) طالب وطالبة والتي بلغت (1.96) وكذلك مستوى الدلالة (0.000) وهي اقل من مستوى الدلالة (0.05) **بالتالي نرفض الفرضية الصفرية ونقبل الفرضية البديلة وهي:**

- توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي درجات طلبة المجموعة التجريبية التي تم تقييمها باستخدام الاختبار التكيفي والمجموعة الضابطة التي تم تقييمها باستخدام الاختبار الخطي ولصالح الاختبار التكيفي.

ويفسر ذلك بأن الاختبارات الإلكترونية التكيفية قد توافقت بشكل كبير مع الطلبة باختلاف الفروق الفردية لديهم، حيث أشتمل بنك الأسئلة على مجموعة متنوعة ومتعددة المستويات، والتي بدورها تقدم الاسئلة في ضوء الفروق الفردية لكل طالب، فيشعر ذوي القدرة الأقل من المتوسط نتيجة استجابته الصحيحة على الأسئلة المقدمة ومن ثم التفاعل بشكل أكبر مع الاستفادة من التغذية الراجعة المقدمة من خلال الاختبارات الالكترونية التكيفية، بينما الطلبة ذوي القدرة المتوسط يجدون تنوع الأسئلة ومناسبة مستوى تعقيدها لقدرتهم ، كما أن ذوي القدرة الأعلى من المتوسط يشعرون بالارتياح وعدم الملل نتيجة وجود تحدي في الأسئلة المقدمة عبر الاختبارات وبالتالي يزيد حماسهم نحو الاستمرار في التحدي والتعلم.

وقد اشارات عدد من الدراسات إلى هذه النتيجة منها:

- ما إشارات إليه دراسة [45] حيث أكدت على وجود فرق دال إحصائياً عند مستوى (0.05) بين متوسط مستوى درجات الطلاب في التطبيق القبلي والبعدي لصالح التطبيق البعدي (الاختبار التكيفي).
- ودراسة [25] التي استهدفت أثر الاختلاف بين الاختبار التكيفية والخطية في دقة تقدير قدرات المتعلم، وقد أظهرت النتائج أن الاختبارات التكيفية أدق من الاختبارات الإلكترونية الخطية في قياس القدرة الفعلية.
- أيضاً دراسة [46] والتي اسفرت نتائجها على وجود فرق دال إحصائياً بين درجات الطلاب لصالح التطبيق البعدي للاختبار التحصيلي المعرفي لمقرر الحاسب الآلي وأمن البيانات .
- دراسة أحمد عبد النبي والتي أشارت نتائجها إلى فاعلية نمط الاختبارات البنائية التكيفية فيما يخص الجانب المعرفي[32] .

- **الفرضية الثالثة:**

"لا توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي درجات طلبة المجموعة التجريبية التي تم تقييمها باستخدام الاختبار التكيفي وبين المجموعة الضابطة التي تم تقييمها باستخدام الاختبار الخطي تعزى لمتغير الجنس (ذكور/أناث)"

اختبار معنوية الفروق بين أفراد عينة الدراسة يعزو لمتغير الجنس، تم استخدام اختبار(t)لاختبار الفروق بين متوسطات لعينتين مستقلتين (ذكور، اناث):

جدول رقم (6) اختبار (t) لاختبار الفروق بين متوسطات لعينتين مستقلتين

الاختبار	الجنس	المتوسط الحسابي	العينة	الانحراف المعياري	قيمة t	درجة الحرية	مستوى الدلالة Sig	النتيجة الإحصائية
التكيفي	ذكور	6.23	52	1.345	-3.938-	198	0.000	دالة
	إناث	7.06	47	1.636				
الخطي	ذكور	5.74	52	1.618	-3.022-	198	0.003	دالة
	اناث	6.50	47	1.939				

يتضح من الجدول (6) الآتي:

- توجد فروق ذات دلالة احصائية عند مستوى دالة (0.05) بين درجات طلبة المجموعة التجريبية التي تم تقييمها باستخدام الاختبار التكيفي بين الذكور والإناث حيث بلغت t المحسوبة (3.938) وهي أكبر من الجدولية (1.96) وكذلك مستوى الدلالة (0.000) وهي أقل من مستوى الدلالة (0.05) ولذلك توجد فروق وكانت الفروقات المعنوية لصالح الإناث كونهما أكثر متوسط حسابي وكانت بفارق متوسط (0.83)، وهذا يعزو أن الطالبات الاناث كن أكثر اهتماماً وحرصاً من الذكور.
- توجد فروق ذات دلالة احصائية عند مستوى دالة (0.05) بين درجات طلبة المجموعة الضابطة التي تم تقييمها باستخدام الاختبار الخطي بين الذكور والإناث حيث بلغت ت المحسوبة (3.022) وهي أكبر من الجدولية (1.96) وكذلك مستوى الدلالة (0.003) وهي أقل من مستوى الدلالة (0.05) وكانت الفروقات المعنوية لصالح الإناث كونهما أكثر متوسط حسابي وكانت بفارق متوسط (0.76) وهذا يعزو أن الطالبات الاناث كن أكثر اهتماماً وحرصاً من الذكور

لذلك نرفض الصفرية ونقبل بالبديلة وهي:

- توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي درجات طلبة المجموعة التجريبية التي تم تقييمها باستخدام الاختبار التكيفي وبين المجموعة الضابطة التي تم تقييمها باستخدام الاختبار الخطي تعزى لمتغير الجنس (ذكور/أناث).
- حيث كانت نتيجة الاختبارين لصالح الإناث كونهما أكثر متوسط حسابي، وهذا يعزو أن الطالبات الاناث كن أكثر اهتماماً وحرصاً من الذكور في التحصيل العلمي ويؤكد ذلك ظهور نفس الفارق في الاختبارين.

- **الفرضية الرابعة:**

"لا توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي درجات طلبة المجموعة التجريبية التي تم تقييمها باستخدام الاختبار التكيفي وبين المجموعة الضابطة التي تم تقييمها باستخدام الاختبار الخطي تعزى إلى متغير التخصص ".

لاختبار معنوية الفروق بين أفراد اجابات عينة الدراسة يعزو لمتغير القسم تم استخدام تحليل التباين الأحادي (One Way ANOVA)، حيث تم الاعتماد على مستوى دلالة ($\alpha = 0.05$)، وتعد الفروق دالة إحصائيًا إذا كانت قيمة مستوى دلالة الاختبار (Sig.) أصغر من مستوى الدلالة الذي اعتمد ها البحث في دراستها، والعكس صحيح.

جدول (7) يوضح دلالة الفروق بين إجابات عينة الدراسة وفق متغير القسم

نوع الاختبار	مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	قيمة ف	الدلالة	الدلالة الإحصائية
الخطي	بين المجموعات	53.940	3	17.980	5.877	0.001	دالة
	داخل المجموعات	599.651	196	3.059			
التكيفي	بين المجموعات	79.821	3	26.607	13.259	0.000	دالة
	داخل المجموعات	393.323	196	2.007			

يتضح من الجدول رقم (7) الآتي:

- توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي درجات طلبة المجموعة الضابطة التي تم تقييمها باستخدام الاختبار الخطي تعزو لمتغير التخصص كون مستوى الدلالة (0.001) وهي اقل من مستوى (0.05) وبالتالي تكون دالة وكانت الفروقات بحسب اختبار (LSD) لصالح قسم الحاسوب بمقدار (2.31).
- توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي درجات طلبة المجموعة التجريبية التي تم تقييمها باستخدام الاختبار التكيفي تعزو لمتغير التخصص كون مستوى الدلالة (0.000) وهي أقل من مستوى (0.05) وبالتالي تكون دالة وكانت الفروقات بحسب اختبار (LSD) لصالح قسم الحاسوب بمقدار (2.49) وقد يعزو ذلك ان طلبة التخصص الحاسوب لديهم معرفة ودراية أكثر من غيرهم في استخدام الحاسوب كون الاختبار التكيفي هو اختبار محوسب وايضًا لأن مقرر الاختبار كانت في مادة مهارات الحاسوب وهي من المواد التي لديها اتصال مباشر بهذا التخصص.

- **ولذلك نرفض الصفرية ونقبل بالبديلة وهي:**

توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي درجات طلبة المجموعة التجريبية التي تم تقييمها باستخدام الاختبار التكيفي وبين المجموعة الضابطة التي تم تقييمها باستخدام الاختبار الخطي تعزى إلى متغير التخصص.

4. الاستنتاجات والتوصيات والمقترحات:

تم عرض تحليل نتائج الدراسة وتفسيرها مستعين في ذلك بالمعالجة الإحصائية لاستجابات العينة للأدوات الدراسة وصولًا إلى الإجابة على أسئلة الدراسة وفرضياتها وسيتم عرض ملخص لعصارة ما توصل إليه البحث من النتائج في ضوء الفرضيات التي سعت لتحقيق أهداف الدراسة، وبناء على ذلك قدم البحث بعض التوصيات والمقترحات والتي

لها دور في تطوير أساليب التقييم الإلكتروني ورفع مستوى التحصيل لدى الطلبة.

4.1 ملخص الاستنتاجات:

توصلت الدراسة الحالية إلى النتائج التالية:

1. لا توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي تقدير القدرة بين الاختبار التكيفي والاختبار الخطي باستخدام طريقة الأرجحية القصوى لتقدير القدرة. أي ان مستوى القدرة بين الاختبارين لا فروق بينهما.
2. توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي الخطأ المعياري لتقدير القدرة بين الاختبار التكيفي والاختبار الخطي باستخدام طريقة الأرجحية القصوى لتقدير القدرة لصالح الاختبار التكيفي.
3. توجد فروق ذات دلالة إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي درجات طلبة المجموعة التجريبية التي تم تقييمها باستخدام الاختبار التكيفي، والمجموعة الضابطة التي تم تقييمها باستخدام الاختبار الخطي ولصالح الاختبار التكيفي.
4. توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي درجات طلبة المجموعة التجريبية التي تم تقييمها باستخدام الاختبار التكيفي، وبين والمجموعة الضابطة التي تم تقييمها باستخدام الاختبار الخطي تعزى لمتغير الجنس (ذكور/أناث).

حيث كانت نتيجة الاختبارين لصالح الإناث كونهما أكثر متوسط حسابي، وهذا يعزو أن الطالبات الاناث كن أكثر اهتماماً وحرصاً من الذكور.

5. توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي درجات طلبة المجموعة التجريبية التي تم تقييمها باستخدام الاختبار التكيفي، وبين المجموعة الضابطة التي تم تقييمها باستخدام الاختبار الخطي تعزى إلى متغير التخصص. وقد يعزو ذلك لأن مقرر الاختبار كان في مساق مهارات الحاسوب وعلى الرغم أن هذا المساق كان إحدى المتطلبات الجامعية إلا أن طلبة قسم الحاسوب لديهم معرفة ودراية أكثر من غيرهم فيه، وعلى الرغم من سهوله واجهة الاختبار إلا أن أيضًا طلبة قسم الحاسوب كانوا أكثر قدره على استخدام الاختبار مقارنةً بباقي التخصصات.

XI. التوصيات:

توصلت نتائج الدراسة إلى فاعلية لاختبارات الإلكترونية التكيفية في تقييم القدرة المعرفية للطلبة في التعليم، ودورها في رفع التحصيل لدى طلبة الجامعة، وفي ضوء هذه النتائج يقترح البحث التوصيات التالية:

1. ضرورة تطور أساليب التقويم باستخدام الاختبارات الإلكترونية في التعلم، متمثلة في الاختبارات التكيفية الإلكترونية، ويقترح البحث الحالي أن يتم البدء أولًا بتحويل الاختبارات التي تقدم لطلاب السنة الأول كمتطلبات للجامعة، مثل: مواد الثقافة الإسلامية، واللغة العربية العامة، اللغة الإنجليزية مهارات الحاسب إلى اختبارات الإلكترونية تكيفية ثم التوسع في ذلك وصولًا إلى إجراء الاختبارات التكيفية الإلكترونية لجميع المساقات الجامعية في جميع الجامعات، وذلك للاستفادة من مزايا الاختبار التكيفي.
2. الاستفادة من التصميم المستخدم في الاختبارات الإلكترونية التكيفية للتغلب على مشكلة القلق المصاحب للاختبارات والغش لدى الطلبة.
3. نشر ثقافة الاختبارات الإلكترونية التكيفية بين معلمي التعليم قبل الجامعي.
4. دعوة الخبراء في مجالي القياس والتقويم والتعلم الإلكتروني لإقامة ورش عمل في مجال الاختبارات الإلكترونية التكيفية وآليات الاستفادة منها بالعملية التعليمية.
5. توفير بيئة صفية مناسبة لعقد الاختبارات المحوسبة وجعله ملائمة لكل من الطلبة والمدرسين
6. تضمين مهارات تصميم وبناء الاختبارات الإلكترونية التكيفية ضمن برامج تنمية قدرات أعضاء هيئة التدريس.
7. التوسع في إجراء الدراسات حول الاختبارات الإلكترونية التكيفية في كافة المراحل التعليمية وفي مختلف الثقافات.

XII. المقترحات:

وختامًا حتى يمكن تفعل تطبيقات الاختبار التكيفي المحوسب لتطور أساليب التقويم للاختبارات الإلكترونية في التعلم الجامعي، اقترح البحث ما يلي:

1. عمل دراسات مشابهة على عدد من الكليات الأخرى والخروج باقتراحات ونتائج متفق عليها من أجل تحسين جودة تقديم الاختبارات المحوسبة لما فيه من مصلحة للطلبة والمدرسين أنفسهم.
2. أن تعمل الجامعات على إنشاء مراكز متخصصة في القياس والتقويم، تتوفر بها البنية التحتية، والبرامج الخاصة بتطوير وتحليل الاختبارات وتضم كفاءات متخصصة للإشراف على تطوير منظومة التقويم في كل جامعة.
3. عمل دراسات خاصة بالخصائص الفنية للاختبارات المحوسبة وفق النظرية الجديدة في القياس؛ لكي يتم تقديم اختبارات مناسبة للطلبة تتمتع بخصائص سيكومترية مناسبة.
4. إجراء مزيد من الدراسات حول مدى تقبل الطلاب مستحدثات تكنولوجية بما فيها الإلكترونية التكيفية.
5. إجراء دراسات لحل مشكلة التخمين في الأسئلة المحوسبة؛ وذلك أثناء القيام بغربلة الفقرات المقدمة للطلبة في كل مادة دراسية؛ حتى يزيد من مستوى فاعلية الفقرات التي تقدم للطلبة في الاختبارات المحوسبة.
6. إجراء دراسات للمقارنة بين نتائج الطلاب على الاختبارات الورقية والاختبارات الإلكترونية، والإلكترونية التكيفي.

XIII. المراجع العربية والأجنبية:

[1] ف. ع. الخزي و م. أ. الزكري، "تكافؤ الاختبارات الإلكترونية مع الاختبارات الورقية في التحصيل الدراسي: دراسة تجريبية على طلبة كلية التربية بجامعة الكويت"، مجلة دراسات الخليج والجزيرة العربية، مج. 37، ص. 167–198، 2011.

[2]N. Alruwais, G. Wills, and M. Wald, "Advantages and challenges of using e-assessment," International Journal of Information and Education Technology, vol. 8, no. 1, pp. 34-37, 2018.

[3] ر. س. البربري، "نمطا تصميم الاختبار الإلكتروني التكيفي الثابت والمتغير الطول وأثرهما على خفض مستوى قلق الاختبار وتنمية الاتجاهات نحو الاختبارات الإلكترونية لدى طلاب كلية التربية"، تكنولوجيا التعليم، 2020.

[4] آ. م. معزي البياضة، "بناء اختبار التكيفية للقدرة الرياضية للصف السابع الأساسي وفق الاستراتيجية الهرمية باستخدام نظرية الاستجابة للفقرة"، رسالة ماجستير، عمادة الدراسات العليا، جامعة مؤته، الأردن، 2011.

[5]Khan and M. Jawaid, "Technology Enhanced Assessment (TEA) in COVID 19 Pandemic," Pakistan Journal of Medical Sciences, vol. 36, no. COVID19-S4, pp. S108–S110. doi:

10.12669/pjms.36.COVID19-S4.2795, accessed on: 21 January 2022.
[6] ع. إسماعيل، "قلق الاختبارات الالكترونية وعلاقته بالأداء في نظر عينة من طالبات كلية الأميرة عالية بجامعة البلقاء التطبيقية - الأردن"، المجلة التربوية، كلية التربية، جامعة سوهاج، مج. 52، ص. 799–823، 2018.
[7] J. Rossiter, "Evaluation of software tools for formative assessment of control topics," IFAC Papers OnLine, vol. 52, no. 9, pp. 292–297, 2019.
[8] D. J. Weiss, "Item banking, test development, and test delivery," in The APA Handbook on Testing and Assessment, K. F. Geisinger, Ed. Washington, DC: American Psychological Association, 2013.
[9] Y. Zheng and H. Chang, "On-the-fly assembled multistage adaptive testing," Applied Psychological Measurement, vol. 39, no. 2, pp. 104-118, 2015.
[10] A.J. Martin and G. Lazendic, "Computer- adaptive testing: Implications for students' achievement, motivation, engagement, and subjective test experience," Journal of Educational Psychology, vol. 110, no. 1, pp. 27, 2018.
[11] D. Magis, D. Yan, and A. A. von Davier, Computerized Adaptive and Multistage Testing With R: using packages catr and mstr, Cham, Switzerland: Springer International Publishing, 2017.
[12] D. Robert, G. David, and B. Benjamin, "Computerized Adaptive Tests for Rapid and Accurate Assessment of Psychopathology Dimensions in Youth," Journal of the American Academy of Child & Adolescent Psychiatry, vol. 9, pp. 122–151, 2020.
[13] Q. Liu, Y. Zhuang, H. Bi, Z. Huang, W. Huang, J. Li, J. Yu, Z. Liu, Z. Hu, and Y. Hong, "Survey of Computerized Adaptive Testing: A Machine Learning Perspective," arXiv preprint arXiv:2404.00712, 2024. [Online]. Available: arxiv.org.
[14] A. Franco, "The Adaptive Quiz Activity," Retrieved from: https://moodle.org/plugins/mod_adaptivequiz, accessed on: 21 January 2022.
[15] F. Martinez-Plumed, R. B. Prudêncio, A. Martínez-Usó, and J. Hernández-Orallo, "Item response theory in AI: Analysing machine learning classifiers at the instance level," Artificial Intelligence, vol. 271, pp. 18-42, 2019.
[16] إ. محاسنة، "القياس النفسي في ظل النظرية التقليدية والنظرية الحديثة"، عمان: دار جرير للنشر والتوزيع، 2013.
[17] خ. الدامغ و هـ. الهاجري، "تصميم الاختبارات الإلكترونية لمتعلمي اللغة العربية -لغة ثانية-," مجلة كلية الدراسات الإسلامية والعربية للبنات بالإسكندرية, vol. 2, no. 35, pp. 823-875, 2019.
[18] Bider and E. Perjons, "Design science in action: developing a modeling technique for eliciting requirements on business process management (BPM) tools," Software & Systems Modeling, vol. 14, no. 3, pp. 1159–1188, 2015.
[19] C. Youn Young and M. Cayce, "Development of Adaptive Formative Assessment System Using Computerized Adaptive Testing and Dynamic Bayesian Networks," Applied Sciences, vol. 10, pp. 2-1, 2020.
[20] M. Boussakuk et al., "Design of Computerized Adaptive Testing Module into our Dynamic Adaptive Hypermedia System," International Journal of Emerging Technologies in Learning (iJET), vol. 16, no. 18, pp. 113-128, 2021. doi: 10.3991/ijet.v16i18.23841.
[21] أ. محمد إبراهيم، "الاختبارات المحبوكة المحوسبة: تطوير لبنوك الأسئلة," مجلة ابداعات تربوية, رابطة التربويين العرب, no. 3, ISSN: 2536-9407, أكتوبر 2017.
[22] C. Wang and H. Lu, "Mediating Effects of Individuals' Ability Levels on the Relationship of Reflective-Impulsive Cognitive Style and Item Response Time in CAT," Educational Technology & Society, vol. 21, no. 4, pp. 89–99, 2018.
[23] Dammas, "Investigate Students' Attitudes toward Computer Based Test (CBT) at Chemistry Course," Archives of Business Research, vol. 4, no. 6, pp. 58-71, 2016.
[24] م. ع. خميس، "التحول الرقمي: قضايا ومشكلات التعلم"، المؤتمر العلمي السادس عشر للجمعية العربية لتكنولوجيات التربية، كلية الدراسات التربوية، جامعة القاهرة، 2019.
[25] العودة و ع. عبيدات، "فاعلية الاختبار التكيفي المحوسب في تقدير القدرة، باستخدام مصفوفة رافن"، دراسات العلوم التربوية، مج. 40، ع. 3، ص. 1602–162، 2013.
[26] نور الدين، "فاعلية القياس التكيفي المحوسب في قياس تحصيل طلاب الجامعة"، المجلة السعودية للعلوم النفسية، جامعة الملك سعود، الجمعية السعودية للعلوم التربوية والنفسية، ع. 64، ص. 29–47، 2019.
[27] ر. فرحات، "الفروق بين الاختبارات المحوسبة الموائمة والخطية في تقدير القدرة وفقاً لنظرية الاستجابة للمفردة لطلبة كلية التربية جامعة الأزهر"، التربية، أكتوبر 2021.
[28] H. Özyurt and Ö. Özyurt, "Ability level estimation of students on probability unit via computerized adaptive testing," Eurasian Journal of Educational Research, vol. 58, pp. 27-44. doi: 10.14689/ejer.2015.58.7.
[29] د. جرجس، "أثر الاختبارات البنائية التكيفية في تنمية التحصيل الدراسي المرتبط بالبرمجة لدى طلاب تكنولوجيا التعليم"، كلية التربية النوعية، جامعة بنها، 2021.
[30] R. James, "Tertiary student attitudes to invigilated, online summative examinations,"

International Journal of Educational Technology, vol. 13, no. 19, pp. 2-13, 2016.
[31] ص. د. علام، "نماذج الاستجابة للمفردة الاختبارية أحادية البعد ومتعددة الأبعاد وتطبيقاتها في القياس النفسي والتربوي"، دار الفكر العربي، 2005.
[32] أ.عبد النبي عبد الملك النظير، "التفاعل بين نمط أسئلة الاختبار إلكتروني التكيفي (ثنائية/متعددة) الاستجابة ومركز التحكم (داخلي/ خارجي) عبر المنصات التعليمية الرقمية وأثره في تنمية مهارة تطوير البرمجيات التعليمية واتخاذ القرار والكفاءة الذاتية لطلاب تكنولوجيا المعلومات"، كلية التربية النوعية، جامعة عين شمس، الجمعية المصرية تكنولوجيا المعلومات، 2020.
[33] م. التقي، "النظرية الحديثة في القياس"، عمان: دار المبسرة للنشر والتوزيع، 2009.
[34] أ. علي الشهري، "تأثير قواعد التعليق ذات الطول الثابت على إدارة الاختبارات التكيفية"، المجلة العلمية لكلية التربية بجامعة الوادي الجديد، 2023.
[35] M. Samsudin, T. SomChut, and M. Ismail, "Evaluating computerized adaptive testing efficiency in measurement student's performance in science timss," Indonesian Journal of Science Education, vol. 8, no. 4, pp. 547-560, 2019.
[36] D. Cisar, D. Radosav, B. Markoski, R. Pinter, and P. Cisar, "Computer adaptive testing of student knowledge," Acta Polytechnic Hungarica, vol. 7, no. 4, pp. 139-152, 2010.
[37] D. Weiss, "Computerized adaptive testing for effective and efficient measurement in counseling and education," Measurement and Evaluation in Counseling and Development, vol. 37, no. 2, pp. 70-84, 2004.
[38] M. K. O. Lee, C. M. K. Cheung, and Z. Chen, "Acceptance of Internet Based learning medium: The role of extrinsic and intrinsic motivation," Information & Management, vol. 42, 2005.
[39] ز. يوسف دعنا، "بناء اختبار محبوك هرمي في الرياضيات للصف الثامن الأساسي وفق نموذج راش في نظرية السمة الكامنة"، دراسات العلوم، مج. 1، ع. 32، 2005.
[40] A. J. Leroux and B. G. Dodd, "A comparison of stopping rules for computerized adaptive screening measures using the rating scale model," Journal of Applied Measurement, vol. 15, pp. 213–226, 2014.
[41] J. M. Linacre, "Computer-Adaptive Testing: A Methodology Whose Time Has Come," Rasch Measurement Transactions, vol. 69, 2000. [Online]. Available: http://www.rasch.org/memo69.pdf
[42] E. Istiyono, W. Brams, R. Setiawan, and I. Megawati, "Developing computerized adaptive testing to measure physics higher order thinking skills of senior high school students and its feasibility of use," European Journal of Educational Research, vol. 9, no. 1, pp. 91-101, 2020.
[43] Y. Huo, "Variable-length Computerized Adaptive Testing: adaptation of the a-stratified strategy in item selection with content balancing," Ph.D. dissertation, Graduate College, University of Illinois at Urbana-Champaign, 2009.
[44] D. Magis and G. Raiche, "CatR: An R package for computerized adaptive testing," Applied Psychological Measurement, vol. 33, no. 7, pp. 576-577, 2011.
[45] أ. م. عبد الرحمن، تصميم الاختبارات, عمان: دار أسامة للنشر والتوزيع, 2011.
[46] د. ج. القس جرجس، "أثر الاختبارات البنائية التكيفية في تنمية التحصيل الدارسي المرتبط بالبرمجة لدى طلاب تكنولوجيا التعليم," كلية التربية النوعية, جامعة بنها, 2021.

EXPLAINABLE ARTIFICIAL INTELLIGENCE-BASED DIAGNOSIS ASSISTANT OF HEPATITIS C VIRUS

A. Oyedeji [1]*
M. Osifeko [1]
S. Iyiade [1]
T. Oke [1]
A. Olayiwola [1]

Received: 23/02/2025
Revised: 08/04/2025
Accepted: 09/04/2025

© 2025 University of Science and Technology, Aden, Yemen. This article can be distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

© 2025 جامعة العلوم والتكنولوجيا، المركز الرئيس عدن، اليمن. يمكن إعادة استخدام المادة المنشورة حسب رخصة مؤسسة المشاع الإبداعي شريطة الاستشهاد بالمؤلف والمجلة.

[1] Department of Computer Engineering, Olabisi Onabanjo University, Ago-Iwoye, Nigeria:
Corresponding Author's Email: oyedeji.ajibola@oouagoiwoye.edu.ng

Explainable Artificial Intelligence-Based Diagnosis Assistant of Hepatitis C Virus

A. Oyedeji
Department of Computer Engineering,
Olabisi Onabanjo University,
Ago-Iwoye, Nigeria
oyedeji.ajibola@oouagoiwoye.edu.ng

M. Osifeko
Department of Computer Engineering,
Olabisi Onabanjo University,
Ago-Iwoye, Nigeria
osifeko.martins@oouagoiwoye.edu.ng

S. Iyiade
Department of Computer Engineering,
Olabisi Onabanjo University,
Ago-Iwoye, Nigeria
omosubomiiyiade@gmail.com

T. Oke
Department of Computer Engineering,
Olabisi Onabanjo University,
Ago-Iwoye, Nigeria
okeracheal911@gmail.com

A. Olayiwola
Department of Computer Engineering,
Olabisi Onabanjo University,
Ago-Iwoye, Nigeria
olayiwola.abisola@oouagoiwoye.edu.ng

Abstract— Hepatitis C is a liver infection prevalent in developing countries, and early detection of this disease would significantly reduce the mortality rate. Advances in artificial intelligence have led to the development of medical diagnostics systems. However, the decisions gotten from these systems are not easily explainable to the end users. Data preprocessing, including feature scaling and oversampling using Synthetic Minority Oversampling Technique, was carried out on HCV data. Seven classifiers—logistic regression, decision tree, random forest, support vector machine, gradient boosting, K-nearest neighbor, and multilayer perceptron (MLP)—were implemented. The models were evaluated, and Shapley Additive Explanations (SHAP) values were employed for model interpretability. MLP with standard scaling has the best performance with an accuracy of 0.97 and a sensitivity and specificity of 1.00. The features with the most influence on the outcome are the albumin test, alkaline phosphatase, alanine transaminase, and aspartate aminotransferase, while sex and cholesterol had the least influence. A web-based diagnosis assistant was deployed for early diagnosis.

Keywords— Hepatitis C, Artificial Intelligence, Explainable AI, SHAP, Diagnosis Assistant.

I. INTRODUCTION

Hepatitis C is a prevalent liver infection caused by the hepatitis C virus (HCV) and is likewise a major cause of liver cancer and liver transplant. The virus can result in both acute and chronic hepatitis, with symptoms ranging from a mild infection to severe, lifelong contamination that includes malignancy, cirrhosis, and damage to the liver. Most hepatitis C infections result from exposure to infected blood through unsafe practices [1]. Approximately 1.5 million cases of the hepatitis C virus are diagnosed yearly, with about 58 million people across the globe having a persistent infection. This phenomenon is more prevalent in poorer developing nations, especially in Africa and Asia [2]. Approximately 3.2 million children and adolescents worldwide have chronic hepatitis C infection, of which the mortality is estimated to be about 290,000 individuals for the year 2019 because of cirrhosis and hepatocellular carcinoma (primary liver cancer) [1, 3].

Early discovery and diagnosis of the disease are essential for effective and timely treatment, as this can result in a reduction of morbidity and mortality rates worldwide. Delayed treatment of the disease can result in other life-threatening diseases such as cirrhosis, fibrosis, and many other complications [4]. The medical diagnosis of liver diseases most often involves a laboratory blood test to determine the levels of liver-related metabolites, more commonly referred to as the liver function tests, and hepatitis C antibody tests, which explain whether the patient is infected with HCV [4, 5, 6]. Most of these tests are not as easy to perform and thus negatively impact the early detection of the disease. The use of available, reliable, and verifiable HCV data for the diagnosis of the disease has therefore become a viable option using artificial intelligence [6].

The healthcare industry has historically been a pioneer in implementing new technologies such as the invention of new medical procedures and the management of chronic diseases [7]. Machine learning is now playing a significant part in disease prognosis and treatment [8], medical imaging and diagnostic services [9], the development of novel pharmaceuticals [10], and the management of medical records [11]. Recent research and advancements have led to the development of medical diagnostic systems utilizing the potential and decision-making ability of artificial intelligence (AI) and machine learning (ML). The remarkable progress experienced in diagnostics has been made possible by the ability of AI-based systems to analyze, break down, and comprehend information from complex data available [6, 12, 13].

As much as the performance of the model's predictions is important, the explainability and interpretability of the models are much more paramount to understanding how the conclusions have been reached. This is even more essential and expedient in data-driven healthcare AI assistants to give the end-users, particularly health workers and physicians, the necessary support to make informed decisions [14]. Understanding the basis for the diagnosis of a particular health condition based on the various inputs or features provided to the models is of great significance to reinforcing and validating the decision of the physicians. SHAP (Shapley Additive Explanations) is an efficient tool for explainable AI both for global and local model interpretability which is based on the game's theoretical Shapley values [15, 16, 17].

An analysis of the accuracy of existing ML methods and a novel AI-based ensemble model on the prediction of hepatitis C disease was the focus of the research [13]. It was observed that the proposed ensemble method had the best accuracy of 95.6%, followed by the Quick, Unbiased, Efficient, Statistical Tree and Bayesian network with accuracies of 94.6% and 94.5%, respectively. However, the study failed to consider the

sensitivity and specificity of the models to understand the kind of errors the models were making. In [18], Cascade RF-LR implemented with oversampling utilizing the artificial bee colony algorithm proposed to automatically detect the probability of HCV in multiclass data. The proposed model outperformed XGBoost, which was the second-best performing model, with a difference of 0.02 in terms of accuracy. A recent study employed k-means extreme machine learning for chronic hepatitis diagnosis with an accuracy of 72.36% [19]. However, these studies did not take into consideration the interpretability and explainability of the systems.

A performance evaluation study of various ML models was carried out on a dataset from the Jordan University Hospital [4]. Sequential forward selection and oversampling were applied, with the models including LR, KNN, DT, RF, and neural network all achieving an accuracy of approximately 82.0%. Furthermore, the study presented a global model interpretation by employing SHAP to show the feature importance and impact on the models' decisions.

The motivation for this project is the need for a diagnostic system for HCV detection and its interpretability. The objectives of this project are to collect HCV data, analyze and preprocess the data, design and develop classification models for the detection of hepatitis, and evaluate the models' performances. In addition, the interpretation and explanation of the models' outcomes will be presented using SHAP. Furthermore, a web application for HCV diagnosis was developed using React and FastAPI for the best-performing classification model.

II. Materials and Methods

The following sections describe the dataset collected, analysis and preprocessing steps, followed by the modelling steps for the prediction and diagnosis of hepatitis.

Brief and succinct descriptions of the ML and interpretability techniques are also provided. Data analysis, model development and model interpretation were carried out in Python [20], employing standard frameworks such as Pandas [21], NumPy, Matplotlib [22] and Scikit-learn [23], and SHAP [24] for global feature and local feature importance.

A. Patients Information and Preprocessing

The UCI Machine Learning Repository, a reputable open-access portal that offers benchmark datasets for machine learning research, is where the dataset used in this study was acquired [25, 26, 27]. Researchers specifically gathered the Hepatitis C Virus (HCV) dataset for use in both academic and clinical research. It consists of patient laboratory test results gathered from healthcare facilities to aid in the diagnosis of liver disease, including HCV and associated disorders.

The dataset has 14 attributes and 615 instances in total, where two of the features are categorical data (category and sex) and the remaining twelve features are numerical data. The Patient ID/No column was dropped as it contains a serial number of patients and provides no information relevant to the diagnosis of the patient's condition. The remaining 11 numerical data comprise the age (in years) of the patients and the results of laboratory liver function tests (LFTs) carried out, which include the albumin test (ALB), alkaline phosphatase (ALP), alanine transaminase (ALT), aspartate aminotransferase (AST), bilirubin (BIL), cholinesterase (CHE), cholesterol (CHOL), creatinine (CREA), gamma-glutamyl transferase (GGT), and protein test (PROT). The LFTs are blood tests used to measure different types of enzymes, proteins, and other substances made by the liver [4, 28, 29]. The category feature is the dependent variable having four (4) classes with the following labels: Hepatitis (Hep), Cirrhosis (Cir), Fibrosis (Fib), Blood donor (BD), and Suspect blood donor (SBD).

The dataset was refined to enable proper functioning of the model the data would be used on with the following steps: loading of the dataset, dropping of duplicates, dropping or filling null values, encoding, etc. No duplicate data was found, while there was a total of 26 missing data, which were removed, resulting in a total of 589 instances remaining. This choice was taken after determining that the missing values were not concentrated in specific features, rendering imputation unreliable without creating bias. The percentage of missing data (~4.2%) was low enough that statistical power would not be significantly impacted by deletion. In the context of medical data, imputation could lead to distorted values that may mislead the model, particularly for sensitive biochemical markers.

The sex attributes were encoded as 1 and 0 for the male and female genders, respectively. The descriptive statistical analysis of the cleaned data is shown in Table 1, while the distributions of the features are presented in Figure 1. The age of the patient population ranged from 23 to 77 years, with a mean of 47 years and a spread of years. There are quite a handful of outliers in various features such as ALB, ALP, ALT, AST, BIL, CREA, and GGT, but a choice to keep these records is made as we are dealing with medical data.

Feature engineering is a significant task in optimizing the correctness of a predictive algorithm on a dataset by transforming the feature space. The extraction of features is a very significant step in ML classification for selecting the best features or attributes with the best predictive power, which can be determined using the Pearson correlation coefficient (PCC) [30, 31]. The PCC is a measure of the degree of relationship between two features, X and Y, with values ranging between -1 and +1, and has been calculated as shown in Figure 2 using Equation 1 [32].

$$\rho_{X,Y} = \frac{\mathbb{E}[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y} \qquad (1)$$

Table 1. Descriptive Statistics of the Pre-processed Numerical

	Age	ALB	ALP	ALT	AST	BIL	CHE	CHOL	CREA	GGT	PROT
Mean	47.42	41.62	68.12	26.58	33.77	11.02	8.20	5.39	81.67	38.20	71.89
Std	9.93	5.76	25.92	20.86	32.87	17.41	2.19	1.13	50.70	54.30	5.35
Min	23	14.90	11.30	0.90	10.60	0.80	1.42	1.43	8.00	4.50	44.80
25%	39	38.80	52.50	16.40	21.50	5.20	6.93	4.62	68.00	15.60	69.30
50%	47	41.90	66.20	22.70	25.70	7.10	8.26	5.31	77.00	22.80	72.10
75%	54	45.10	79.90	31.90	31.70	11.00	9.57	6.08	89.00	37.60	75.20
Max	77	82.20	416.60	325.30	324.00	209.00	16.41	9.67	1079.10	650.90	86.50

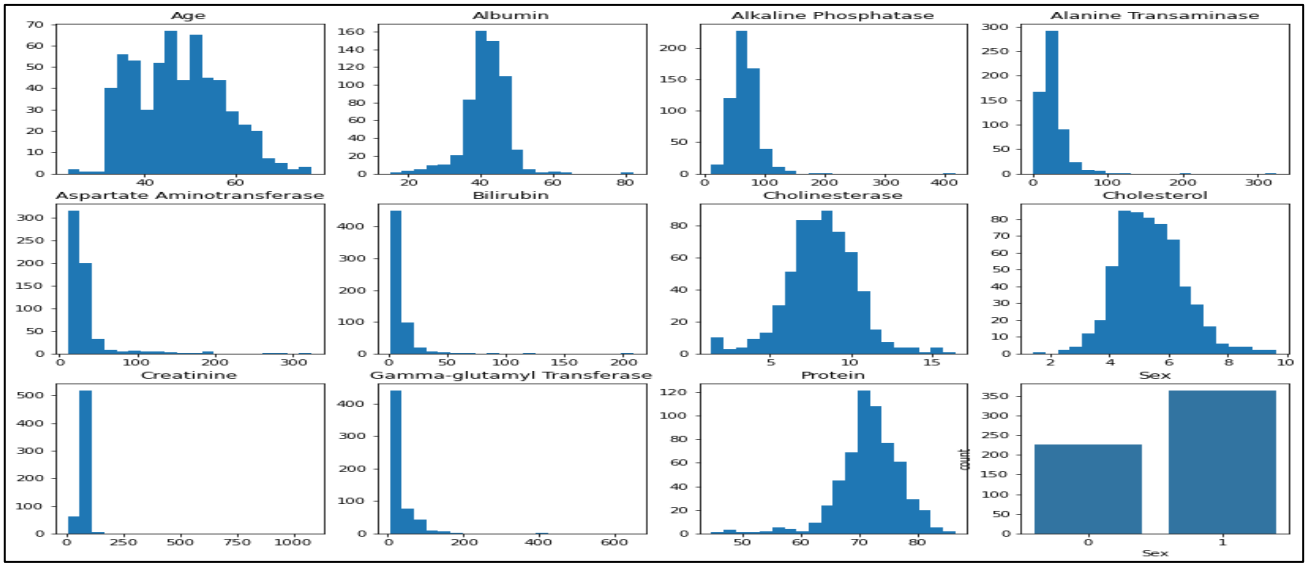


Fig. 1. Frequency Histograms of HCV Dataset Features.

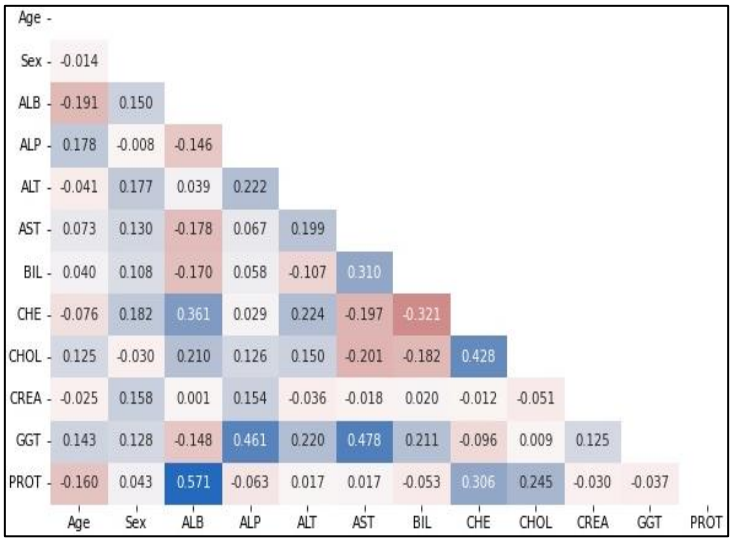


Fig. 2. Bivariate Correlation of the Numerical Features

Where μ_X and μ_Y are the means of X and Y respectively, σ_X and σ_Y are the standard deviations of X and Y respectively and E is the expectation.

Of the 12 variables or features, ALB was moderately correlated with PROT (ρ=0.571) and weakly correlated with CHE (ρ=0.361).

No feature was removed from the dataset as there were no highly correlated features, that is, features with correlation coefficients ranging between $0.7 \leqslant \rho \leqslant 1$.

Figure 3 shows an imbalanced distribution of the categories with the blood donor amounting to 526 (89.3%) instances. The eligible 589 patients' data were randomly split

into a training set (N = 471) and a test set (M = 118). Due to the highly imbalanced nature of the training set, oversampling was done to bring the minority categories to the same amount as the majority category (N_(Blood Donor) = 421), and this procedure is referred to as data augmentation [6]. Synthetic Minority Oversampling Technique (SMOTE) [33] was subsequently adapted to the training set, thereby bringing the total number of training samples to 2105 with all categories having a total number of 421 instances. SMOTE was selected because it effectively creates synthetic samples of the minority class by interpolating between existing samples, avoiding precise duplication and lowering the risk of overfitting. In comparison to random oversampling, SMOTE improves the model's generalization capabilities by maintaining the structure of the feature space and balancing the distribution of classes.

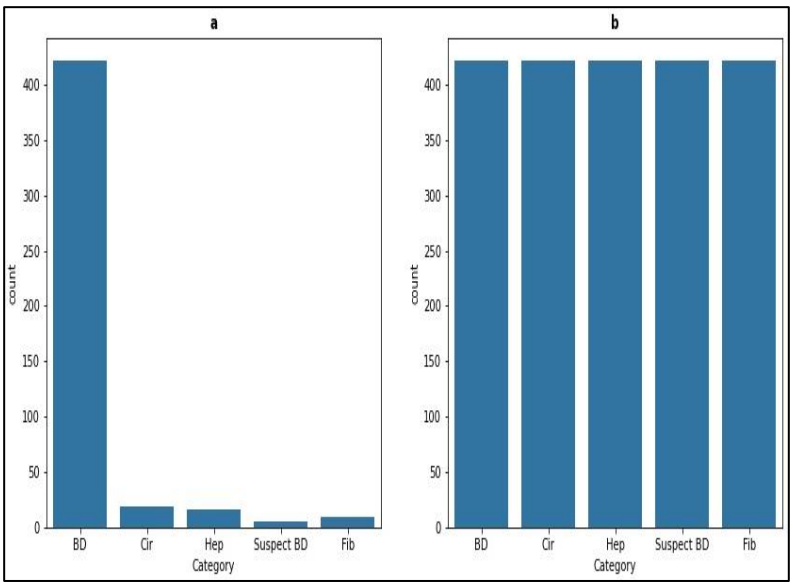


Fig. 3. Distribution of Target Classes in Train Set (a) before SMOTE (b) after SMOTE

The performance of many ML algorithms is improved by scaling the numerical input variable to a standard range, especially for models which utilize the weighted sum of the input or distance measures. The data standardization scales each input variable as stated in Equation 2. This shifts the distribution to be centered around the mean with a standard deviation of 1. The feature scaling (FS) was achieved using the StandardScaler library available in scikit-learn.

$$z = \frac{x - \mu}{\sigma} \quad (2)$$

where z is the standardized value of the original value x, μ and σ are the average and standard deviation of x respectively.

B. Classification Models' Development

ML algorithms are efficient tools for learning patterns from data and predicting future outcomes. There are several applications of ML and AI including classification, regression, clustering and anomaly detection. Classification is a supervised learning method for the prediction of the target label of unseen data from previous data; the target labels can be either binary or multiclass [34, 35].

Seven ML and ANN classification methods were used to build models to classify and predict the HCV status of the patient in this paper. The models employed are the multivariable logistic regression (LR) [36], decision tree classifier (DT) [37], random forest classifier (RF) [38, 39], gradient boosting classifier (GB) [40], k-nearest neighbor (KNN) [35], support vector machine classifier (SVC) [41] and multilayer perceptron (MLP) [42]. They were chosen to represent a broad spectrum of learning paradigms, including:

- Linear models (logistic regression) for baseline interpretability,
- Tree-based models (DT, RF, GB) for handling nonlinear interactions and feature hierarchies,
- Distance-based learners (KNN) for intuitive classification based on feature similarity,
- Margin-based classifiers (SVM) known for robust generalization in high-dimensional spaces,
- Neural networks (MLP) for capturing complex non-linear patterns.

This diverse selection enables a comprehensive performance comparison and facilitates model interpretability and robustness analysis.

The LR model is used for classification by estimating the probability of an event happening and is expressed mathematically using

$$P(y = 1|x_1, x_2, \dots, x_n) = \frac{1}{1+e^{-(\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_n x_n)}} \quad (3$$

Where, $x_1, x_2, \dots, x_n$ are the independent variables, and $\beta_1, \beta_2, \dots \beta_n$ are their respective coefficients. It maps any real-valued number into the [0, 1] interval, making it suitable for predicting probabilities. The DT is a tree data structure made up of several nodes and branches.

The decision tree algorithm utilizes a divide and conquer method and repeatedly partitions the input variables to classify or predict the output parameter while RF is an ensemble ML

technique containing numerous decision trees to predict a target variable. The RF uses the Bagging method where homogenous weak learners' models (in this case, decision trees) independently learn from one another in parallel. The final prediction is achieved by a model-averaging approach. The GB tree model is a stacked learning approach where a robust predictive model is formed by combining several individual weak learning trees (DT).

The KNN algorithm uses the distance between data points to make classifications. The class labels are selected based on the majority vote of the selected number of neighbours. The Euclidean distance is used to measure proximity. In SVM, each data point is plotted in n-dimensional space with n as the number of attributes or inputs. The model then performs classification by finding the optimal hyper-plane that differentiates the target labels using the decision rule specified in (4).

$$y = sign(wx + b) \quad (4)$$

Where, y is the predicted class label. sign is the sign function, which returns +1 or -1 depending on the input. The MLP is a type artificial neural network (ANN) commonly used for tasks like classification and regression. An MLP consists of at least three layers: an input layer, one or more hidden layers, and an output layer. The output of a single neuron in an MLP is specified in (5)

$$z = f(\sum_{i=1}^{n} w_i x_i + b) \quad (5)$$

Where z is the output of the neuron. f is the activation function. w_i are the weights associated with the inputs. x_i are the inputs to the neuron. b is the bias term.

The selected models were fitted and trained on the train data set while they were tested and evaluated using the test set. To optimize the performance and correctness of the models, the various algorithms' hyperparameters were tuned and set as depicted in Table 2. Hyperparameters were optimized via GridSearch Cross-Validation, which ensured a systematic and unbiased exploration of the model's parameter space. The GridSearch CV was used to avoid the models memorizing the data, thereby resulting in overfitting. For all computations, the random state was set to 42 to allow for the reproducibility of results.

The settings notably impacted:

- Model convergence (e.g., max_iter=3000 for LR and MLP ensured full convergence),
- Model complexity (e.g., max_depth=11 in DT and RF controlled overfitting),
- Learning dynamics (e.g., learning_rate=0.1 in GB balanced speed and stability),
- Neighbor selection in KNN (neighbors=2) enhanced sensitivity due to small class sizes.

Overall, proper hyperparameter tuning was crucial in attaining high accuracy and stability across all models, especially under imbalanced conditions.

Table 2. Selected Classification Algorithms' Hyperparameters

Model	Hyperparameters	Value
LR	max_iter	3000
	C	1.7
DT	max_depth	11
RF	n_estimators	350
	max_depth	10
GB	max_depth	4
	learning_rate	0.1
	n_estimators	200
KNN	n_neighbors	2
	weights	distance
SVC	C	1.9
	kernel	rbf
MLP	activation	relu
	solver	adam
	learning_rate	constant
	max_iter	200

C. SHAP for Model Interpretability

Explaining and interpreting the decision-making process of a model is oftentimes a difficult process as there is no standard metric for the measurement of explainability and interpretability [14]. Understanding why models make certain decisions is as important as the models' accuracy, especially in the medical field [15]. The interpretability of all the models used in this study was assessed using SHAP based on the Shapley values from game theory [16, 17]. SHAP is used to explain the outcome of ML models. It connects optimal credit allocation with local explanations and assigns each input or attribute an importance value for a particular prediction.

The KernelExplainer method in SHAP which is a flexible method capable of working with all types of models is used to determine the individual feature contributions using Shapley values. The KernelExplainer builds a weighted linear regression to compute the importance of each feature. Due to how computationally expensive the process is, a sample of 500 training sets was used as background data for the SHAP model inference. This was implemented on the models without feature scaling. The SHAP algorithm is presented as:

Algorithm 1: SHAP Algorithm

1. **Model Prediction**
   Obtain the prediction $\hat{y}$ for the instance to be explained
2. **Initialize Parameters**
   Let F be the set of all features.
   For each feature x_i, initialize its Shapley value Θ_i to 0.
3. **Calculate the Shapley Value**
   For each feature x_i:
   a. Consider all possible subsets S of features excluding x_i
   b. For each subset S:
      i. Calculate the model prediction using only the features in S:f(S)
      ii. Calculate the model prediction using the features in S plus feature x_i: $f(SU\{i\})$.
      iii. Compute the difference in predictions: $\Delta = f(SU\{i\} - f(S)$.

iv. Weight the difference by the combinatorial factor $\frac{|S|!(|F|-|S|-1)!}{|F|!}$

v. Accumulate the weighted difference to the Shapley value of feature x_i:

$$\Theta_i += \frac{|S|!\,(|F|-|S|-1)!}{|F|!}\Delta$$

4. Normalize Shapley Values

For each feature x_i, average its accumulated Shapley value over all subsets S to get the final Shapley value

$$\Theta_i = \sum_{S \sqsubseteq F \setminus \{i\}} \frac{|S|!\,(|F|-|S|-1)!}{|F|!}\Delta$$

III. Result and Discussion

A. Model Evaluation

In Table 3, the confusion matrices for the individual models with and without feature scaling are presented. The DT, KNN, MLP and SVC had a slightly better prediction accuracy when feature scaling was incorporated compared to the performance without feature scaling. The mean performance of the classification models was determined by metrics including accuracy, sensitivity and specificity as presented in Table 4 and Figure 4. All models performed better than the baseline accuracy (accuracy gotten by always predicting the majority class which is the blood donor group) of 0.893. For classification without applying feature scaling, GB narrowly performed the best with an accuracy of 0.97, followed by SVC, MLP, RF and KNN; all with an accuracy of 0.95, LR (0.94) and DT (0.91). The results obtained for training while applying feature scaling were similar except for MLP and DT improving their accuracy performance to 0.97 and 0.92 respectively.

In terms of the model's ability to correctly detect cases of patients that have an HCV-related disease or suspected donor, referred to as the sensitivity, SVC (with and without FS) and MLP (with FS) performed best with a sensitivity of 1.00, meaning that all patients with HCV were detected.

Table 3. Confusion Matrices for the Models

Methods	Actual Class	Predicted Class without FS					Predicted Class with FS				
		BD	CIR	FIB	HEP	SBD	BD	CIR	FIB	HEP	SBD
LR	BD	104	0	0	1	0	104	1	0	0	0
	CIR	1	4	0	0	0	1	4	0	0	0
	FIB	0	0	1	2	0	1	0	0	2	0
	HEP	0	1	2	1	0	0	0	2	2	0
	SBD	0	0	0	0	1	0	0	0	0	1
DT	BD	101	2	1	1	0	101	2	0	2	0
	CIR	1	4	0	0	0	1	4	0	0	0
	FIB	0	0	1	2	0	0	1	1	1	0
	HEP	2	0	1	1	0	1	0	2	1	0
	SBD	0	1	0	0	0	0	0	0	0	1
RF	BD	105	0	0	0	0	105	0	0	0	0
	CIR	1	4	0	0	0	1	4	0	0	0
	FIB	0	0	1	2	0	0	0	1	2	0
	HEP	2	0	1	1	0	2	0	1	1	0
	SBD	0	0	0	0	1	0	0	0	0	1
GB	BD	105	0	0	0	0	105	0	0	0	0
	CIR	0	5	0	0	0	0	4	1	0	0
	FIB	0	0	2	1	0	0	0	3	0	0
	HEP	1	1	1	1	0	1	1	1	1	0
	SBD	0	0	0	0	1	0	0	0	0	1
KNN	BD	105	0	0	0	0	105	0	0	0	0
	CIR	1	4	0	0	0	1	4	0	0	0
	FIB	0	0	2	1	0	1	0	0	2	0
	HEP	1	1	1	1	0	0	1	1	2	0
	SBD	1	0	0	0	0	0	0	0	0	1
SVC	BD	104	0	0	1	0	105	0	0	0	0
	CIR	0	4	0	1	0	0	4	1	0	0
	FIB	0	0	2	1	0	0	0	0	3	0
	HEP	0	1	1	2	0	0	1	1	2	0
	SBD	0	1	0	0	0	0	0	0	0	1
MLP	BD	103	0	0	0	2	105	0	0	0	0
	CIR	0	5	0	0	0	0	4	1	0	0
	FIB	0	0	2	1	0	0	0	2	1	0
	HEP	1	1	1	1	0	0	1	1	2	0
	SBD	0	0	0	0	1	0	0	0	0	1

Table 4: Performance Evaluation

	Without Feature Scaling			**With Feature Scaling**		
	Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity
LR	0.94	0.92	0.99	0.94	0.84	0.99
DT	0.91	0.77	0.96	0.92	0.84	0.96
RF	0.95	0.77	1.00	0.95	0.77	1.00
GB	0.97	0.92	1.00	0.97	0.92	1.00
KNN	0.95	0.77	1.00	0.95	0.84	1.00
SVC	0.95	1.00	0.99	0.95	1.00	1.00
MLP	0.95	0.92	0.98	0.97	1.00	1.00

Contrary to sensitivity, specificity measures the classifier's ability to detect patients without a condition. Three models (RF, GB and Figure KNN) with and without feature scaling, as well as scaled SVC and MLP had the best specificity having achieved a value of 1.00 (all 105 blood donors without any traces of HCV infection were correctly identified). These two terms are very important for consideration in medical diagnostics. Considering all the metrics and the confusion matrices, the MLP (with FS) has the best performance, achieving an accuracy of 0.97, and sensitivity and specificity of 1.00.

Using the same or comparable HCV datasets, earlier research has reported accuracy rates between 85.0% and 95.0%; however, most of these studies either did not address class imbalance or used interpretability methodologies that were too restrictive [43, 44, 45, 46]. This work stands out for the following reasons

- Using SMOTE to address imbalanced data improves sensitivity
- Evaluating seven different machine learning models under both scaled and unscaled conditions.
- In contrast to previous studies, SHAP is used for both local and global interpretability.
- Deploying a real-time diagnostic web app, making the solution immediately usable by clinicians.

In comparison to relevant literature, the study's top-performing model obtained 97.00% accuracy, 100.00% sensitivity, and 100.00% specificity, all of which are on the upper bound.

B. Model Interpretability using Shapley's Values

With the above evaluation results presented, the big questions remain unanswered. Questions such as: (i) How do the models make the decisions? (ii) What features or attributes are responsible for the decision? (iii) How much importance or impact do individual features have? (iv) Can the results be trusted? SHAP can answer these questions using both global and local model interpretability which is based on the addition or aggregation of Shapley values

C. Global Model Interpretability

The average behaviors of the models were explained using the global model interpretability function. SHAP not only shows the feature's importance but takes it a step further showing the feature's influence whether positive or negative. Figure 5 depicts the feature importance of the models for all the category classes (BD, Hep, Cir, Fib and Suspect BD). The various classes are color-coded according to the legend. The models found the Albumin test, Alkaline Phosphatase, Alanine Transaminase and Aspartate Aminotransferase as the most important features with the most influence on the outcome of the predictions with a few variations or exceptions.

Analyzing the feature importance of each model deeper reveals that the 5 most influential features (ALB, AST, ALP, BIL and CHE) of the DT and RF were similar, and since the RF is a stack of DTs, the feature influence would be similar. GB, however, surprisingly found PROT as the most important closely followed by AST and CHE. The SVC and MLP had similar features (AST, ALT, ALP, GGT and BIL) with the most influence on the predictions.

In general, all models found sex and cholesterol as the two least important or influential features for the prediction of HCV. This goes to show that the hepatitis C virus is not a respecter of gender. KNN and SVC models hardly utilize these features (sex and CHOL) in addition to Cholinesterase at all for prediction, while sex was the least important for all models except MLP where the least was CHOL.

Figure 6 depicts the influence of individual values on the outcome where each point represents an instance of each feature. The blue points represent low values of the feature while the red points represent high values as shown by the legend. The x-axis shows the value of the influence of each instance on the predictions. Generally, the lower values of AST and GGT have a positive influence while the higher values have a negative impact on the outcomes for all the seven models considered. This contrasts with the ALP, ALB and PROT, in which the higher values have a positive influence while the lower values are responsible for a negative impact on the model output. In addition, Sex and CHOL are clustered around 0.0 (i.e. they have low SHAP values). SHAP values for all models thereby verifying that the two features have minimal impact on the model's outputs. For clinical acceptance, SHAP values offer clear, feature-level insights into the model's decision-making process. From a medical perspective. Standard liver function tests, including albumin (ALB), ALP, ALT, AST, and BIL, were found to be the most influential markers. When screening for or diagnosing liver-related diseases, clinicians can give priority

to these markers. The negligible influence of sex and cholesterol aligns with known HCV pathology, suggesting clinicians can focus less on these factors in isolation. Local SHAP values enable individualized patient-level explanations, which are essential for precision medicine, allowing physicians to understand why a specific diagnosis was made for a particular patient, thereby improving trust and clinical usability.

D. Local Model Interpretability

SHAP values can be calculated for each prediction to know how the features contribute to that single prediction. All models were also evaluated using the local interpretability method. This is of great advantage and significance in understanding individual predictions. Figure 7 shows the feature impacts for an individual prediction with a base value (average prediction) of $E[f(x)] \approx 0.17$ and the model prediction probability value $f(x) \approx 0$. The plots show the individual feature contributions to arriving at the model prediction probability.

The feature impact of all the features in achieving the prediction probability for the LR is depicted in Figure 7(a). ALP has the most influence on the model's prediction with a Shapley value of +0.24 followed by GGT, ALT and CHE with values-0.15, -0.09 and -0.06 respectively. These are the most predictive features for the individual instance. CHOL, Sex and CREA have little to no impact on the outcome of the LR model. These features all have a Shapley value of ±0 and thus are the least predictive features. Interestingly, these three least impactful features are consistent with the same values across all seven models implemented. It can be concluded that the values of CHOL (5), sex (1 - Male) and CREA (74) do not have an impact on the final prediction

E. Web-Based Hepatitis C Virus Diagnosis Assistant

A web-based hepatitis C virus diagnosis assistant was built based on the best-performing model, the multilayer perceptron with feature scaling to enable clinicians to predict the risk of Hepatitis C in patients using react for the frontend and fastAPI was used for the backend. Vercel was used for the hosting of the frontend while render was used for the deploying of the fastAPI. The diagnosis assistant has been designed to be user-friendly, interactive and efficient for end users including physicians and laboratory scientists. The Hepatitis C virus diagnosis assistant is available at https://hcv-prediction-ui-y8yy.vercel.app/. Figure 8 depicts the test page and the result page.

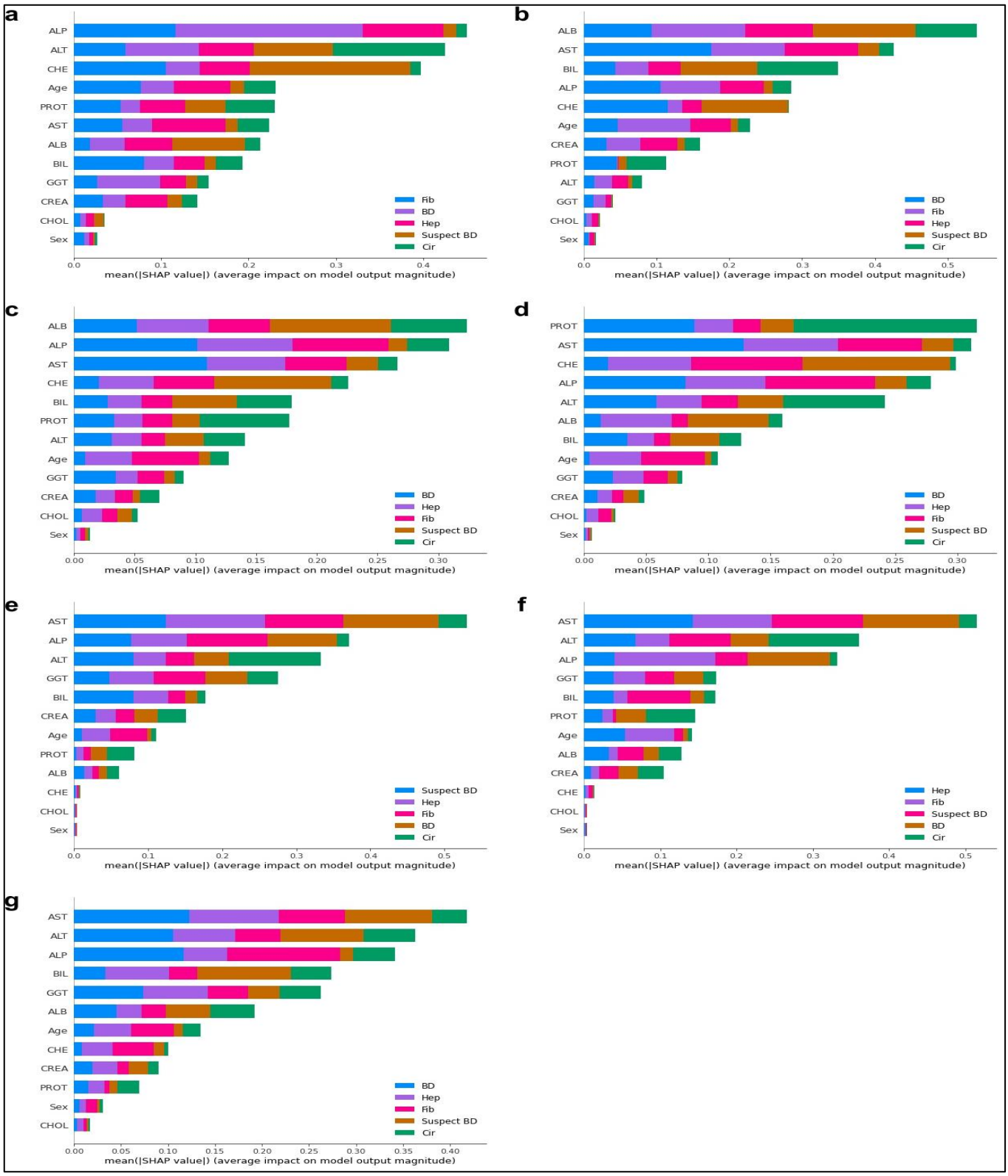


Fig. 4. Global Feature Importance using SHAP (Shapley Additive Explanations) for the following Models without Feature Scaling (a) LR (b) DT (c) RF (d) GB (e) KNN (f) SVC (g) MLP

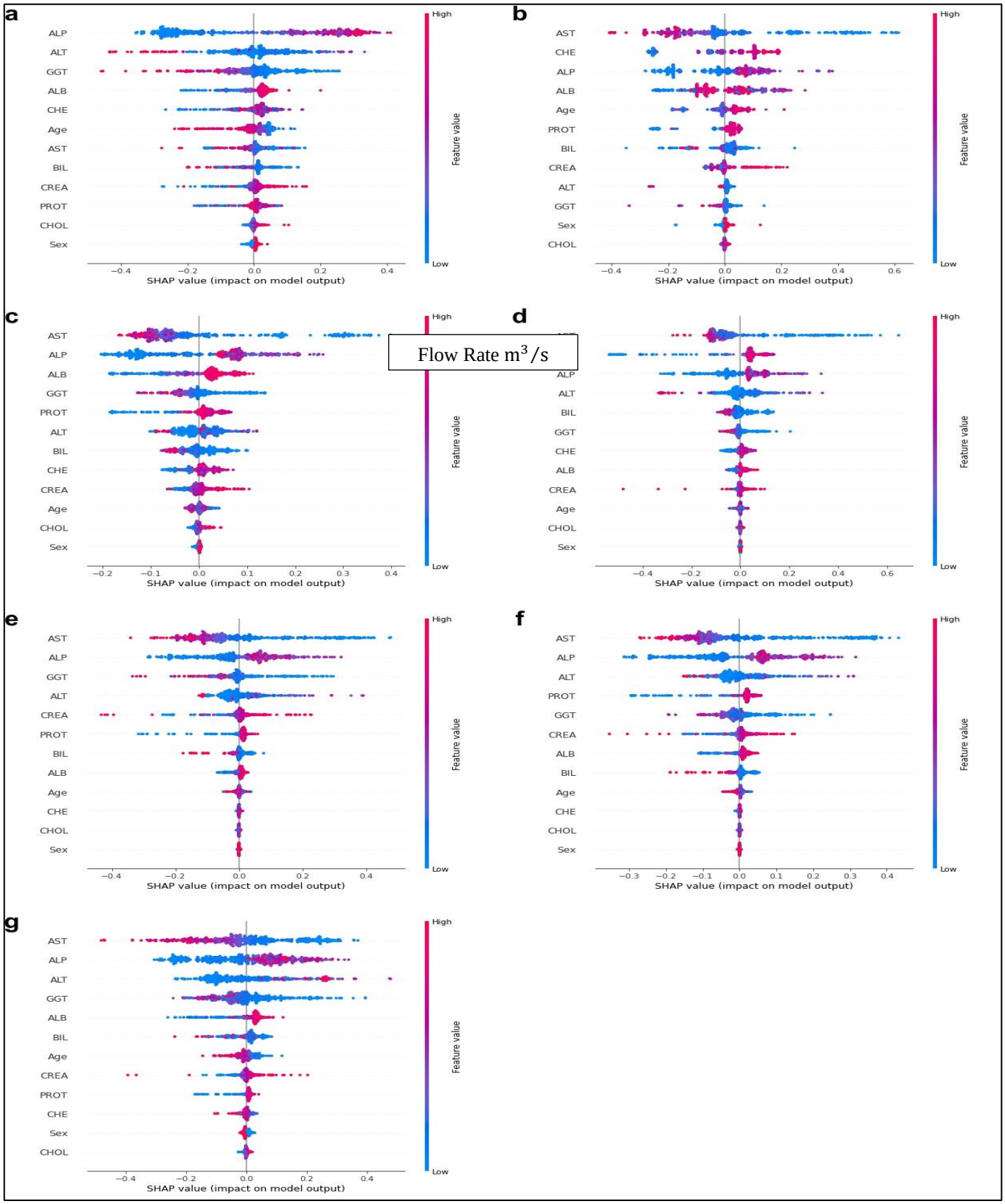


Fig. 5. Summary Plots using SHAP (Shapley Additive Explanations)
for the following Models without Feature Scaling (a) LR (b) DT (c) RF (d) GB (e) KNN (f) SVC (g) MLP

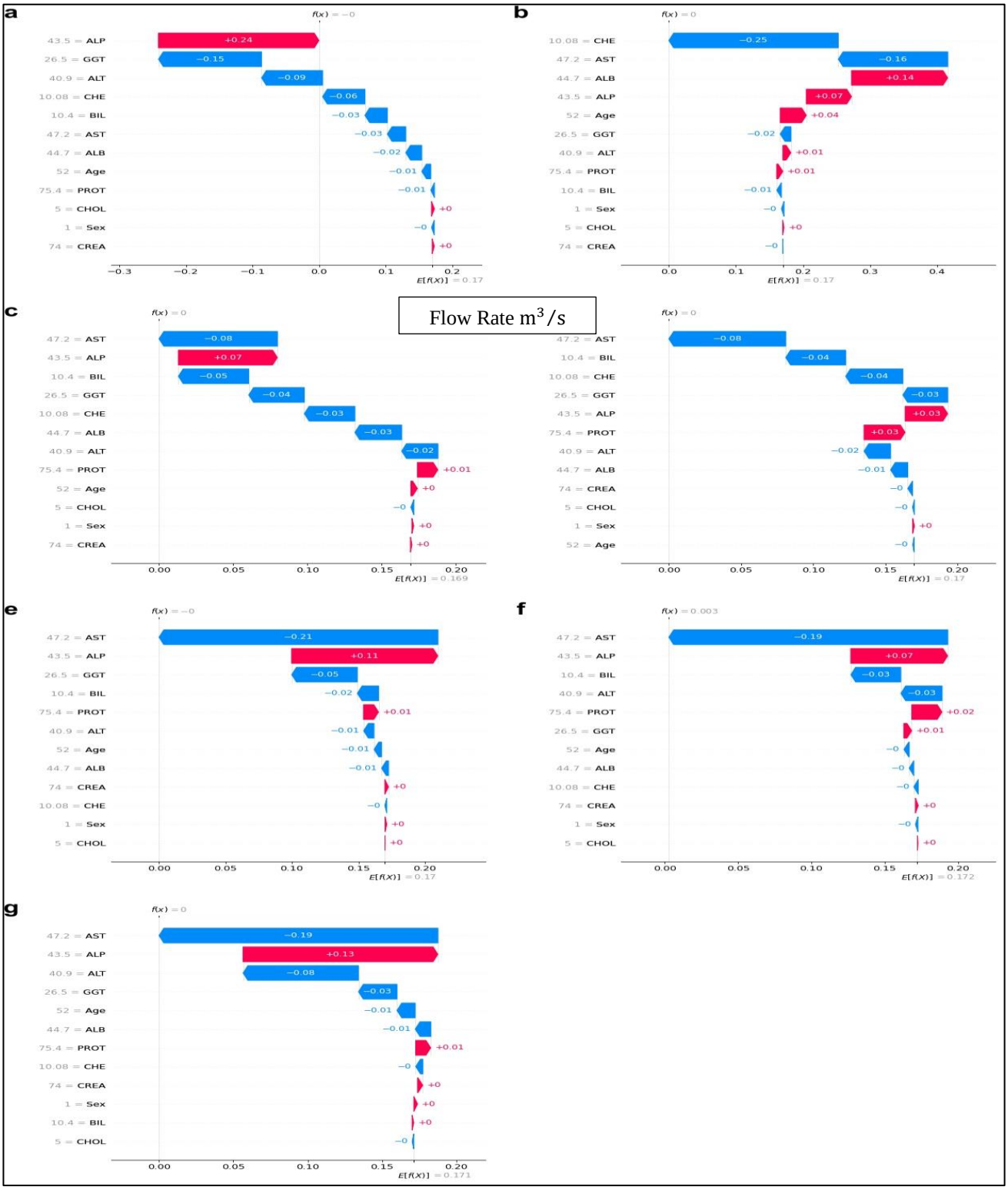


Fig. 6. Feature Importance for a Single Test Instance
for the following Models without Feature Scaling (a) LR (b) DT (c) RF (d) GB (e) KNN (f) SVC (g) MLP

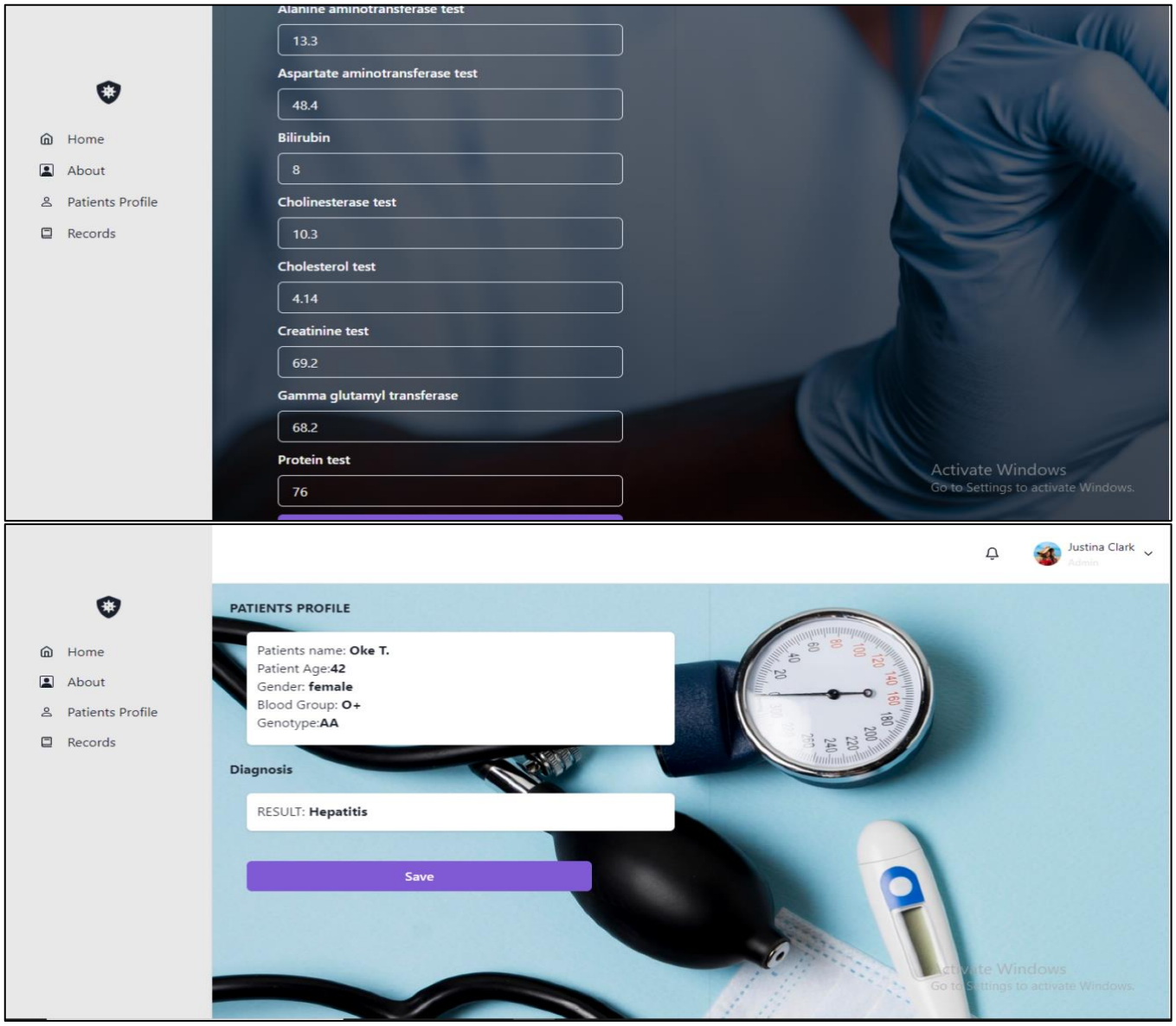


Fig. 7. HCV Diagnosis Assistant - (Top) Test Page (Bottom) Result Page

IV. Conclusion

Early detection and diagnosis of diseases and health conditions result in a significant reduction in the mortality rate among patients. The process of detection and diagnosis of hepatitis C infection can be improved by adapting the use of artificial intelligence and machine learning. Artificial intelligence algorithms are efficient tools for the early diagnosis of health conditions. This study conducted a comparative analysis of AI methods for the detection of HCV using accuracy, sensitivity and specificity as metrics. The study showed that the multilayer perceptron algorithm implemented with feature scaling had the best level of the performance metrics used at the same time.

Of great importance to the end-users is the understanding of how the models are making their decisions and what the feature contributions are. SHAP based on Shapley values was used for global and local model interpretability. The models found the albumin test, alkaline phosphatase, alanine.

Transaminase and Aspartate Aminotransferase are the most important features with the most influence while sex and cholesterol are the two least important features for the prediction of HCV.

Despite promising results, the study admits a number of shortcomings. Larger datasets are required to enhance external validity, even though this size is enough for model training. Analysis of population subgroups was limited because only age and gender were available.

In the future, we aim to explore improving model generalizability, larger and more varied datasets covering various ethnicities, comorbidities, and geographic populations. Additional clinical features, such as patient history, symptoms, or treatment data, will also be incorporated. Deploying the web-based assistant in real-time and clinically testing it in hospital settings for validation and feedback is also another direction.

REFERENCES

[1] World Health Organization, "Hepatitis C," accessed Apr. 20, 2023. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/hepatitis-c

[2] H. Mamdouh Farghaly, M. Y. Shams, and T. Abd El-Hafeez, "Hepatitis C Virus prediction based on machine learning framework: a real-world case study in Egypt," *Knowledge and Information Systems*, 2023, doi: 10.1007/s10115-023-01851-4.

[3] M. Cedolin, M. E. Genevois, and Z. Canbulat, "Hepatitis C Diagnosis Using Computational Intelligence Techniques," 2024, pp. 29–36, doi: 10.1007/978-3-031-67192-0_4.

[4] A. M. Ali et al., "Explainable Machine Learning Approach for Hepatitis C Diagnosis Using SFS Feature Selection," *Machines*, vol. 11, no. 3, p. 391, 2023, doi: 10.3390/machines11030391.

[5] S. Agrawal, R. K. Dhiman, and J. K. Limdi, "Evaluation of abnormal liver function tests," *Postgraduate Medical Journal*, vol. 92, no. 1086, pp. 223–234, 2016, doi: 10.1136/postgradmedj-2015-133715.

[6] L. Chen, P. Ji, and Y. Ma, "Machine Learning Model for Hepatitis C Diagnosis Customized to Each Patient," *IEEE Access*, vol. 10, pp. 106655–106672, 2022, doi: 10.1109/ACCESS.2022.3210347.

[7] A. O. Oyedeji, M. O. Osifeko, O. Folorunsho, O. R. Abolade, and O. O. Ade-Ikuesan, "Design and Implementation of a Medical Diagnostic Expert System," *Journal of Engineering Science*, vol. 10, no. 2, pp. 103–109, 2019.

[8] M. A. Myszczynska et al., "Applications of machine learning to diagnosis and treatment of neurodegenerative diseases," *Nature Reviews Neurology*, vol. 16, no. 8, pp. 440–456, 2020, doi: 10.1038/s41582-020-0377-8.

[9] S. K. Mun, K. H. Wong, S. C. B. Lo, Y. Li, and S. Bayarsaikhan, "Artificial Intelligence for the Future Radiology Diagnostic Service," *Frontiers in Molecular Biosciences*, vol. 7, 2021, doi: 10.3389/fmolb.2020.614258.

[10] S. Ekins et al., "Exploiting machine learning for end-to-end drug discovery and development," *Nature Materials*, vol. 18, no. 5, pp. 435–441, 2019, doi: 10.1038/s41563-019-0338-z.

[11] A. J. P. L et al., "Medical information retrieval systems for e-Health care records using fuzzy based machine learning model," *Microprocessors and Microsystems*, 2020, doi: 10.1016/j.micpro.2020.103344.

[12] H. Haga et al., "A machine learning-based treatment prediction model using whole genome variants of hepatitis C virus," *PLoS ONE*, vol. 15, no. 11 November, 2020, doi: 10.1371/journal.pone.0242028.

[13] M. O. Edeh et al., "Artificial Intelligence-Based Ensemble Learning Model for Prediction of Hepatitis C Disease," *Frontiers in Public Health*, vol. 10, 2022, doi: 10.3389/fpubh.2022.892371.

[14] C. E. Charlton, M. T. C. Poon, P. M. Brennan, and J. D. Fleuriot, "Development of prediction models for one-year brain tumour survival using machine learning: a comparison of accuracy and interpretability," *Computer Methods and Programs in Biomedicine*, p. 107482, 2023, doi: 10.1016/j.cmpb.2023.107482.

[15] S. M. Lundberg and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions," *31st Conference on Neural Information Processing Systems (NIPS 2017)*, vol. 32, no. 2, pp. 1208–1217, 2017.

[16] M. Sundararajan and A. Najmi, "The many shapley values for model explanation," *37th International Conference on Machine Learning, ICML 2020*, vol. PartF16814, pp. 9210–9220, 2020.

[17] T. R. Noviandy et al., "Interpretable machine learning approach to predict Hepatitis C virus NS5B inhibitor activity using voting-based LightGBM and SHAP," *Intelligent Systems with Applications*, vol. 25, p. 200481, Mar. 2025, doi: 10.1016/j.iswa.2025.200481.

[18] T. H. S. Li, H. J. Chiu, and P. H. Kuo, "Hepatitis C Virus Detection Model by Using Random Forest, Logistic-Regression and ABC Algorithm," *IEEE Access*, vol. 10, pp. 91045–91058, 2022, doi: 10.1109/ACCESS.2022.3202295.

[19] J. Cai, T. Chen, Y. Qi, S. Liu, and R. Chen, "Fibrosis and inflammatory activity diagnosis of chronic hepatitis C based on extreme learning machine," *Scientific Reports*, vol. 15, no. 1, p. 11, Jan. 2025, doi: 10.1038/s41598-024-84695-4.

[20] S. Raschka, J. Patterson, and C. Nolet, "Machine learning in python: Main developments and technology trends in data science, machine learning, and artificial intelligence," *Information (Switzerland)*, vol. 11, no. 4, 2020, doi: 10.3390/info11040193.

[21] J. Reback *et al.*, "pandas-dev/pandas: Pandas 1.0.5," *Zenodo*, Jun. 2020, doi: 10.5281/ZENODO.3898987.

[22] J. Ranjani, A. Sheela, and K. P. Meena, "Combination of NumPy, SciPy and Matplotlib/Pylab—A good alternative methodology to MATLAB—A Comparative analysis," *Proc. 1st Int. Conf. Innovations Inf. Commun. Technol. (ICIICT)*, 2019, doi: 10.1109/ICIICT1.2019.8741475.

[23] G. Varoquaux, L. Buitinck, G. Louppe, O. Grisel, F. Pedregosa, and A. Mueller, "Scikit-learn," *GetMobile: Mobile Comput. Commun.*, vol. 19, no. 1, pp. 29–33, 2015, doi: 10.1145/2786984.2786995.

[24] S. Mangalathu, S. H. Hwang, and J. S. Jeon, "Failure mode and effects analysis of RC members based on machine-learning-based SHapley Additive exPlanations (SHAP) approach," *Eng. Struct.*, vol. 219, 2020, doi: 10.1016/j.engstruct.2020.110927.

[25] R. Lichtinghagen and G. Hoffmann, "HCV data Data Set," *Machine Learning Repository, University of California, Irvine*, Accessed: Dec. 13, 2022. [Online]. Available: https://archive.ics.uci.edu/ml/datasets/HCV+data

[26] D. Dua and C. Graff, "UCI Machine Learning Repository," [Online]. Available: http://archive.ics.uci.edu/ml

[27] G. Hoffmann, A. Bietenbeck, R. Lichtinghagen, and F. Klawonn, "Using machine learning techniques to generate laboratory diagnostic pathways—a case study," *J. Lab. Precis. Med.*, vol. 3, p. 58, 2018, doi: 10.21037/jlpm.2018.06.01.

[28] B. R. Thapa and A. Walia, "Liver Function Tests and their Interpretation," *Indian J. Pediatr.*, vol. 74, pp. 67–75, 2007.

[29] Hepatitis NSW, "Hepatitis factsheet: Liver Function Tests," 2017.

[30] S. M. Abdelmoneim, M. Kayed, and S. A. Taie, "A Comparative study for Feature Extraction and Classification of Images," *Proc. ACCS/PEIT 2019 - 6th Int. Conf. Adv. Control Circuits Syst. and 5th Int. Conf. New Paradigms Electron. Inf. Technol.*, pp. 105–110, 2019, doi: 10.1109/ACCS-PEIT48329.2019.9062848.

[31] I. Guyon and A. Elisseeff, "An introduction to feature extraction," *Stud. Fuzziness Soft Comput.*, vol. 207, pp. 1–25, 2006, doi: 10.1007/978-3-540-35488-8_1.

[32] P. Schober and L. A. Schwarte, "Correlation coefficients: Appropriate use and interpretation," *Anesth. Analg.*, vol. 126, no. 5, pp. 1763–1768, 2018, doi: 10.1213/ANE.0000000000002864.

[33] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *J. Artif. Intell. Res.*, vol. 16, pp. 321–357, 2002, doi: 10.1613/jair.953.

[34] A. Oyedeji, A. Salami, O. Folorunsho, and O. Abolade, "Analysis and Prediction of Student Academic Performance Using Machine Learning," *J. Inf. Technol. Comput. Eng.*, vol. 4, no. 1, pp. 10–15, 2020, doi: 10.25077/jitce.4.01.10-15.2020.

[35] A. Olayiwola, A. Oyedeji, O. Omoyeni, O. Ayemimowa, and M. Olaoluwa, "Comparative Analysis of Machine Learning Models for Detection of Fake News: A Case Study of Covid-19," *J. Inf. Technol. Comput. Eng.*, vol. 7, no. 1, pp. 29–33, 2023, doi: 10.25077/jitce.7.01.29-33.2023.

[36] L. Buitinck *et al.*, "API design for machine learning software: experiences from the scikit-learn project," 2013. [Unpublished/Preprint if applicable].

[37] B. Charbuty and A. Abdulazeez, "Classification Based on Decision Tree Algorithm for Machine Learning," *J. Appl. Sci. Technol. Trends*, vol. 2, no. 1, pp. 20–28, 2021, doi: 10.38094/jastt20165.

[38] A. Chaudhary, S. Kolhe, and R. Kamal, "An improved random forest classifier for multi-class classification," *Inf. Process. Agric.*, vol. 3, no. 4, pp. 215–222, 2016, doi: 10.1016/j.inpa.2016.08.002.

[39] Z. Masetic and A. Subasi, "Congestive heart failure detection using random forest classifier," *Comput. Methods Programs Biomed.*, vol. 130, pp. 54–64, 2016, doi: 10.1016/j.cmpb.2016.03.020.

[40] U. Singh, M. K. Gourisaria, and B. K. Mishra, "A Dual Dataset approach for the diagnosis of Hepatitis C Virus using Machine Learning," *Proc. 2022 IEEE Int. Conf. Electron. Comput. Commun. Technol. (CONECCT)*, 2022, doi: 10.1109/CONECCT55679.2022.9865758.

[41] A. Tharwat, "Parameter investigation of support vector machine classifier with kernel functions," *Knowl. Inf. Syst.*, vol. 61, no. 3, pp. 1269–1302, 2019, doi: 10.1007/s10115-019-01335-4.

[42] H. Taud and J. F. Mas, "Multilayer Perceptron (MLP)," in *Geomatic Approaches for Modeling Land Change Scenarios*, pp. 451–455, 2018, doi: 10.1007/978-3-319-60801-3_27.

[43] D. B. Smith, B. Donald, J. Bukh, C. Kuiken, A. S. Muerhoff, C. M. Rice, J. T. Stapleton, and P. Simmonds, "Expanded classification of hepatitis C virus into 7 genotypes and 67 subtypes: updated criteria and genotype assignment web resource," *Hepatology*, vol. 59, no. 1, pp. 318–327, 2014, doi: 10.1002/hep.26744.

[44] R. Safdari, A. Deghatipour, M. Gholamzadeh, and K. Maghooli, "Applying data mining techniques to classify patients with suspected hepatitis C virus infection," *Intell. Med.*, vol. 2, no. 4, pp. 193–198, 2022, doi: 10.1016/j.imed.2021.12.003.

[45] O. O. Oladimeji, A. Oladimeji, and O. Olayanju, "Machine Learning Models for Diagnostic Classification of Hepatitis C Tests," *Front. Health Inform.*, vol. 10, no. 1, p. 70, 2021, doi: 10.30699/fhi.v10i1.274.

[46] A. Alotaibi, L. Alnajrani, N. Alsheikh, A. Alanazy, S. Alshammasi, M. Almusairii, S. Alrassan, and A. Alansari, "Explainable ensemble-based machine learning models for detecting the presence of cirrhosis in Hepatitis C patients," *Computation*, vol. 11, no. 6, p. 104, 2023, doi: 10.3390/computation11060104.

TRANSFORMING BIOMASS INTO ENERGY: A PATHWAY TO SUSTAINABLE ENERGY SECURITY IN NORTHERN NIGERIA

M. G. Jegawa [1]
M. L. Mohammed [2] *
A. B. Muhammad [2]
J. Sanib [2]
B. O. Agboola [3]

Received: 15/03/2025
Revised: 28/04/2025
Accepted: 29/04/2025

© 2025 University of Science and Technology, Aden, Yemen. This article can be distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

© 2025 جامعة العلوم والتكنولوجيا، المركز الرئيس عدن، اليمن. يمكن إعادة استخدام المادة المنشورة حسب رخصة مؤسسة المشاع الإبداعي شريطة الاستشهاد بالمؤلف والمجلة.

[1] IMT Atlantique Bretagne-Pays de la Loire, Nantes Campus, 4, rue Alfred Kastler - La Chantrerie, CS 20722, 44307 Nantes cedex 3, France
[2] Department of Energy and Applied Chemistry, Usmanu Danfodiyo University, Sokoto, Nigeria
[3] Department of Petroleum Chemistry, American University of Nigeria, Yola, Nigeria

*Corresponding Author's Email: misbahu.ladan@udusok.edu.ng

Transforming Biomass into Energy: A Pathway to Sustainable Energy Security in Northern Nigeria

M. G. Jegawa
IMT Atlantique Bretagne-Pays de la Loire, Nantes Campus, 4, rue Alfred Kastler - La Chantrerie, CS 20722, 44307 Nantes cedex 3,
France

M. L. Mohammed
Department of Energy and Applied Chemistry, Usmanu Danfodiyo University
Sokoto, Nigeria
misbahu.ladan@udusok.edu.ng

A. B. Muhammad
Department of Energy and Applied Chemistry, Usmanu Danfodiyo University
Sokoto, Nigeria

J. Sani
Department of Energy and Applied Chemistry, Usmanu Danfodiyo University
Sokoto, Nigeria

B. O. Agboola
Department of Petroleum Chemistry, American University of Nigeria
Yola, Nigeria

Abstract— The increasing global population has led to a significant surge in demand for fossil fuels. However, given the substantial environmental impact associated with their use, it has become essential to explore and develop alternative sources of fuel. Biomass, among other alternatives, is now becoming an important consideration as an alternative to fossil fuels. In this review, we made a critical assessment of the potentials of biomass as a source of fuel in Northern Nigeria. We took into account the opportunities and challenges of having biomass as a source for fuel, taking into consideration the geographical, climatic, and agricultural activities in Northern Nigeria. The review also explored the implications for Nigeria's bioenergy industry, as well as the challenges and limitations of establishing biomass and bioenergy industries. The role of government in terms of policy formulation was highlighted, along with suggestions for improvement. Although there will be challenges to encounter, the potential of having biomass as a source of fuel in Northern Nigeria is enormous and would be significantly beneficial socially and economically.

Keywords— Biomass, Energy, Fossil Fuel, Northern Nigeria, Pollution, Renewable Energy.

I. INTRODUCTION

In the 21st century, energy is surely one of the most pressing issues, particularly when considering its environmental impact. The heavy use of fossil fuels for our daily lives has greatly harmed the environment by contributing to global warming [1].

The functioning of a modern society relies heavily on sustainable energy, as it is the foundation for various essential activities, including transportation, manufacturing, healthcare, agriculture, and communication [2]. As evident in some research, energy is central to sustainable economic development; thus, many countries are providing special grants for energy research and development to achieve sufficiency and security. Sustainable energy is essential for achieving some of the United Nations' 17 Sustainable Development Goals (SDGs), including Goal 7, which focuses on energy transition from fossil fuels [3, 4]. Countries with growing economies, industries, and populations tend to consume more energy, resulting in an expected continued rise in global energy demand [5].

Hence, collective action and the political will to transition away from fossil fuels have gained increasing attention, primarily due to global warming caused by anthropogenic greenhouse gas emissions [6]. Through international cooperation and agreements such as the Kyoto Protocol and the Paris Agreement and the transition to renewable energy sources, technological innovations, including energy efficiency, are essential to a sustainable energy future for our planet [7]. However, the transition to alternative energy sources becomes a real challenge as the energy system of most developed and developing countries simply operates on fossil fuels [8, 9].

Renewable energy sources offer sustainable characteristics and numerous benefits, as the depletion of natural resources has made exploring alternatives an urgent necessity. They provide low environmental impact, prevent the exhaustion of fossil fuels and other natural resources, and allow for the utilization of greener alternatives, thereby improving energy security [10]. In addition to lower environmental impact, renewables also have a lesser production cost compared to the conventional fossil fuels [11]. For instance, wind and solar energy sources produce no emissions during operation, as for biomass, it releases CO_2 when burned, but this emitted CO_2 is reabsorbed by plant growth, and that is why biomass is considered a carbon-neutral energy source [12]. However, despite a recent surge in renewable energy integration, they are not a direct replacement for fossil fuels [13].

The average energy consumption per person in Nigeria (156 kWh) is significantly lower than that of some developing countries, such as Malaysia (4,114 kWh), South Africa (4,405 kWh), and Venezuela (3,413 kWh), which highlights Nigeria's heavy reliance on generators [14]. In addition to a significant portion of the citizens lacking access or connection to the grid, the country experiences frequent power outages, averaging 10-12 hours per day [15]. Furthermore, in rural areas, approximately 60% of the population remain unconnected to the grid [16].

Northern Nigeria possesses substantial renewable energy resources, such as solar, wind, hydro, and biomass. The development of renewable energy in this region can foster sustainable and inclusive socio-economic growth, generate job opportunities, attract new investments, and offer environmentally friendly energy solutions. However, despite this potential, the region's electricity generation remains heavily reliant on fossil fuels, causing energy insecurity and leaving many people without access to electricity. With vast agricultural and forest resources, Northern Nigeria has the potential to produce over 200 billion kg of biomass annually

[17]. Yet the region faces severe energy deprivation due to a lack of investment in modern energy infrastructure and favorable policies. This highlights the underutilization of biomass as a resource to improve energy access in the region. Biomass such as crop waste, forest residues, and animal manure could be converted into modern energy carriers, reducing reliance on imported fuels [18].

This study therefore assesses Northern Nigeria's biomass availability by agro-ecological zone, evaluates conversion efficiencies for localized energy solutions, and outlines policy gaps to optimize utilization.

II. Biomass Dominance in Nigeria's Energy Landscape

The majority of households in developing countries, including Nigeria, are using solid fuels such as wood or charcoal on open fires or inefficient stoves for cooking [19]. The use of such polluting fuels and technologies results in poor health and environmental pollution and contributes to climate change [12]. Other activities, including transportation, water treatment, manufacturing, and communication, all require energy, which mostly comes in the form of electricity, to spur socio-economic development [20]. Nigeria has a total electrical energy demand of 98,000 MW, but its current production stands at only 3,900 MW, despite an installed capacity of 12,667 MW [21]. However, only 45% of Nigerians have access to this limited electricity supply [22]. The national grid's expansion to remote areas is hindered by several challenges, including scattered population, bureaucratic bottlenecks, technical difficulties, and lack of planning, among others [23].

Although not used for electricity generation, biomass accounts for 79% of Nigeria's energy use, followed by oil at 15%, natural gas at 3%, electricity at 2%, and coal at 1% [22]. Factors such as limited access to the national grid, poverty, and cultural preference for biomass utilization for cooking underscore its dominance as a primary energy source in Nigeria [24]. The huge consumption of biomass energy in Nigeria is mainly dominated by wood and charcoal for cooking purposes. However, due to limitations in local charcoal production methods, wood fuel dominates the traditional biomass consumption [25].

The breakdown of Nigeria's energy consumption by source in 2021 is presented in Figure 1 [26].

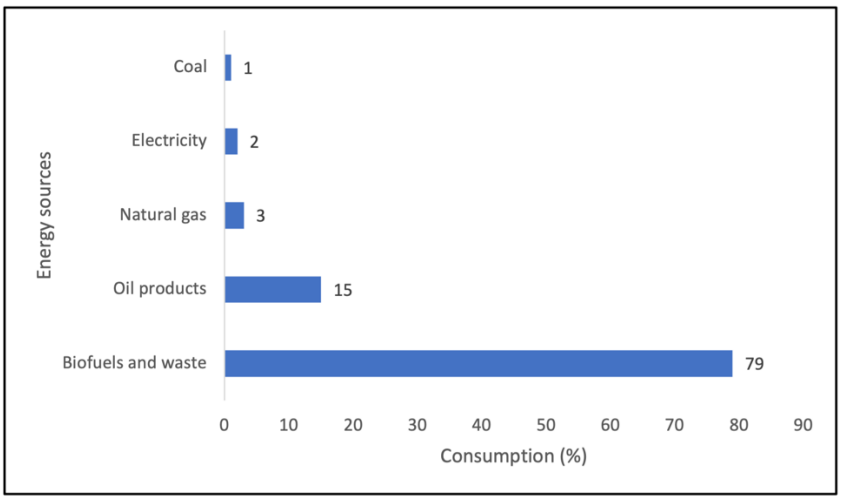


Fig. 1. Breakdown of energy consumption in Nigeria by source in the year 2021 [26]

The huge consumption of biomass energy in Nigeria is mainly dominated by wood and charcoal for cooking purposes. However, due to limitations in local charcoal production methods, wood fuel dominates the traditional biomass consumption [25]. The Northern region of the country is endowed with other renewable energy sources, which, if properly harnessed, could significantly enhance energy security and sustainability. For instance, Nigeria receives high solar irradiation, averaging around 20.1 MJ/m²/day, with the Northern regions reaching 4.1–5.8 kWh/m²/day [27]. This potential makes the region suitable for Concentrating Solar Power (CSP) projects, with estimated capacities of 88.7 GW for CSP and 210 GW for solar PV [28]. On the other hand, the wind energy potential in Nigeria is moderate and geographically concentrated, with the core Northwestern region showing a higher monthly average mean speed than other regions at 6.79 ms^{-1}. Specifically, Kano (9.27 ms^{-1}), Katsina (8.15 ms^{-1}), Sokoto (7.99 ms^{-1}), Gusau (6.52 ms^{-1}), and Kaduna (5.31 ms^{-1}) show the most potential for large and medium wind power generation, as they experience greater wind speeds at higher altitudes of 10 m and above [29]. Additionally, Nigeria possesses considerable small and large hydropower potential, which remains largely untapped. With over 278 potential small hydropower (SHP) sites with a combined capacity of 734.3 MW, only 37 MW has been explored [30].

III. Northern Nigeria's Agro-Ecological Zones and Biomass Potential

To understand biomass characteristics and its distribution in Nigeria, it is important to consider the ecology and its variation. Various phenomena affect the variation of the agroecological zones in Nigeria, which include landforms, geology, soils, land use, vegetation, hydrology, climate, and wildlife; thus, a diverse selection of food crops are easily cultivated, considering the economic, nutritional, and social importance across the zones [31].

According to 2018 estimates, up to 78% of land in Nigeria is used for agricultural activities, and when broken down into

different components to show the diversity of the agricultural land in the country, 37% of this agricultural land is arable and is utilized for cultivation of crops such as wheat, maize, and rice, which could be replanted every year [32]. However, in the Northern part of Nigeria, 52% of the arable land is used for farming or agriculture, which is higher than the average of 37% for the country [32]. Furthermore, 7.4% of agricultural land is dedicated to permanent crops, such as coffee, citrus, and rubber, which do not require replanting. This category also includes land for flowering shrubs, fruit trees, and nut trees. In contrast, 33.3% of the land is used for cultivation of naturally occurring crops that could take five or more years to mature, including herbaceous forage crops [33]. The different agro-ecological zones in Nigeria are as follows:

A. Sahel Savanna (Yobe, Jigawa, and Borno states)

Found in the farthest Northwest and Northeast regions of Nigeria, this zone is characterized by low rainfall of less than 600 mm (typically 3 to 4 months) and a dry season which could exceed 8 months [34]. This semi-arid zone has high evaporation rates due to heat and wind. The trees in this zone are drought-resistant and sparsely dispersed and could reach a height of 4 to 9 meters [34]. This zone is characterized by the presence of common tree species such as Acacia senegal, Commiphora africana, Acacia raddiana, and Acacia laeta. Additionally, the zone has a vast area of scarce grassland that is very short [35], such as Aristida stipoides, Schoenefeldia gracilis, and Chloris priean [36]. Crops cultivated in the Sahel region include millet, sorghum, cowpeas, pigeon peas, groundnuts, green grams, and chickpeas [37].

B. Sudan Savanna (Sokoto, Kaduna, Kano, and some parts of Borno States of Nigeria)

Covering more than 25% of Nigeria, this area exhibits higher precipitation within a range of 600 mm to 1000 mm, thereby producing more fertile soil and vegetation [34]. Although humidity in this region is usually under 40%, it could reach higher than 60% during periods of higher rainfall, and the dry season could last between 4 months and 6 months [36]. The zone is characterized by its agricultural activities, which produce vital crops such as cotton, millet, maize, and groundnuts, which are economically important crops [36]. The region has the largest population density as well as the highest number of cattle, which suggests high agricultural activities in general; alongside stunted trees, the natural vegetation is covered by short grasses of about 1 to 2 m high [36].

C. Guinea Savanna (Benue, Plateau, Nasarawa, Kwara, Niger, Kogi, and parts of Taraba States)

Located in the North Central region of the country and characterized by high annual rainfall between 1000 mm and 1500 mm, which could last for about 6 to 8 months [34]. This region has a large variation of both trees and grass, making it the most extensive vegetation in Nigeria. The region is blessed with tall trees that can reach up to 15 meters and grasses that are between 1 and 3 meters high; however, most of this vegetation is susceptible to being burnt by forest fires during the dry season [34]. Nevertheless, due to climate adaptation over time, the trees and vegetation have developed long taproots and thick barks, which make them resilient and enable them to survive forest fires [36]. Food crops such as millet, sorghum, maize, cowpea, and yam are some of the most cultivated crops in the region [38].

In addition to Nigeria's vegetation zones and features, another important aspect of understanding Nigeria is through its six geopolitical zones, presented as follows:

The North-West Zone has seven states (Jigawa, Kaduna, Kano, Katsina, Kebbi, Sokoto, and Zamfara);

The North-East Zone has six states (Adamawa, Bauchi, Borno, Gombe, Taraba, and Yobe);

The North-Central Zone has seven states, including Abuja (Benue, Kogi, Kwara, Nasarawa, Niger, Plateau, and the Federal Capital Territory (FCT));

The South-South Zone has six states (Akwa Ibom, Bayelsa, Cross River, Delta, Edo, and Rivers).

The Southeast Zone has five states (Abia, Anambra, Ebonyi, Enugu, and Imo) and

The South-West Zone has six states (Ekiti, Lagos, Ogun, Ondo, Osun, and Oyo).

Based on these geo-political zones, all the zones containing “North” in their description and name are regarded as Northern regions/states, while all the zones containing “South” in their description are regarded as Southern regions/states.

Biomass for energy consists of woody and woody biomass, herbaceous biomass, aquatic biomass; animal and human waste biomass, and biomass mixtures [18]. These biomass resources also differ based on their climatic region, with the rain forest zone expected to produce woodier biomass [39] and the savannah zones producing more crop residues [40]. This categorization of biomass is solely based on ecological or vegetation types found in Northern Nigeria as discussed earlier. Nigeria has different sources of biomass feedstock, which could be converted into energy. Figure 2 illustrates the different biomass sources and their overall contribution to Nigeria’s biomass potential [24][25].

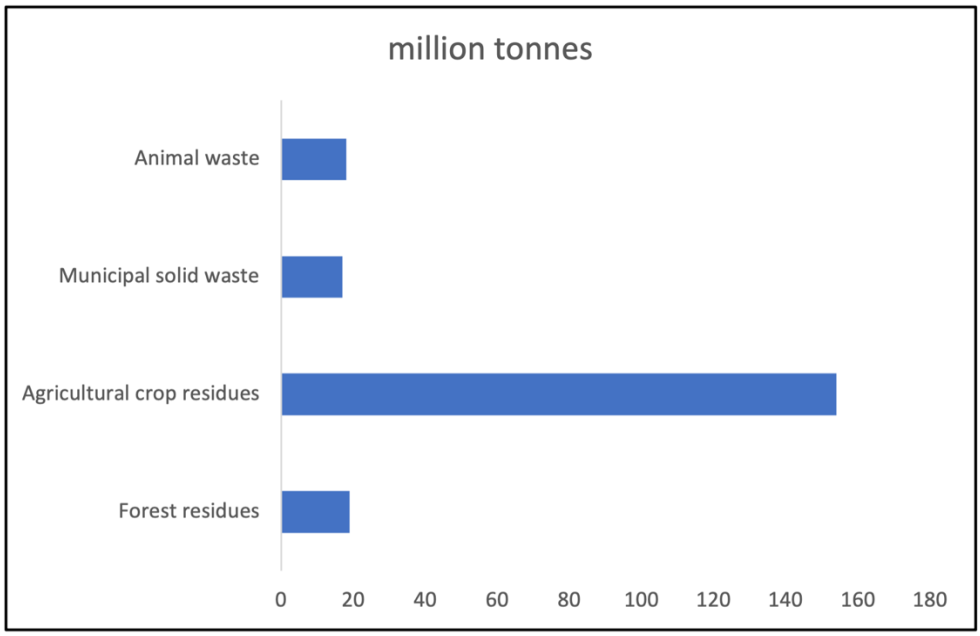


Fig. 2. Estimated annual biomass resource potential in Nigeria (Million Tonnes) [25]

IV. Distribution Of Biomass Resources In Northern Nigeria

Woody biomass consists of trees and other plantations with woody materials that are found and grown in the forest, rangelands, or woodlands [41]. These resources could include all but not limited to, the main trunk, leaves, branches, and other parts of the woody-like plant [42]. They indicate forests' ability to produce wood and sequester carbon and are essential for different applications, including paper production, construction, and energy [43]. Nigeria has established forest reserves covering an area of about 11,000,000 hectares, equivalent to 12.2% of the country's total land area [44]. With high potential for economic returns and energy production, forest residues—including logging residues (such as edgings, offcuts, sawdust, veneer log cores, slabs, bark, trimmings, and rejects)—as well as demolition wood and process residues, offer an alternative economic pathway for the country [18]. Nigeria's forest residue could be estimated at about 19 million tons, which is approximately equivalent to 8.68 Mtoe (363 PJ) [45]. Table 1 shows the forest reserves and plantations in selected Northern States of Nigeria [18].

Table 1: Forest reserves and plantations categorized by state [18]

State	Area of Forest Reserve (ha)	Area of Forest Plantation (ha)
Adamawa	10,011	2,374
Bauchi	840,280	1,200
Benue	60,175	2,234
Borno	582,820	432,052
Jigawa	92,000	3,000
Kaduna	613,484	6,146
Kano	77,702	2,186
Katsina	245,100	18,900
Kebbi	340,289	17,750
Kogi	540,360	5,000
Kwara	460,350	6,000
Niger	756,906	4,956
Plateau	402,500	6,800
Sokoto	602,631	10,943
Taraba	10,011	1,359

A. Agricultural Biomass

There are so many types of agricultural crops and their residues in Northern Nigeria, which could serve as biomass feedstock. In nature, agricultural biomass can be described as having no woody stems, and they live until the end of their growing season [46]. Most agricultural crops and grasses fall under this category, even though some crops are produced specifically as energy crops. These rain-fed crops are dispersed across Nigeria through traditional small-scale farming, feeding the country while also providing energy from their residues [47]. The most commonly produced crops in Northern Nigeria include maize, cocoyam, cassava, guinea corn, millet, rice, melon, yam, groundnut, soybeans, cotton, and cowpea. Table 2 highlights the states in Northern Nigeria with the highest production of food crops [47].

Northern Nigeria has huge potential for biomass resources from agricultural residues, and various conversion technologies could take advantage of these resources to provide cleaner energy that is cheap and reliable through a sustainable practice that conforms to international standards [12]. Agricultural residues produced in Northern Nigeria could be of two types: crop residues obtained from harvesting of target crops on the farm, such as leaves, straws, and stalks of cereals such as rice, maize, sorghum, millet, cocoa pods, cassava peelings, etc., and agro-industrial by-products derived from crop processing activities, which include coconut shells and husks, rice husks, oil seed cakes, cocoa husks, and sugarcane bagasse [48]. The crop residues, estimated at around 2.35 million tonnes, could produce an equivalent of 29 PJ of energy for Nigeria [25], and they do not provide competition for food production [49]. The majority of these residues, including maize, millet, and sorghum, occur within the months of November to March [50], while for rice, it is usually from October to December [51], and given the potential of harvest yield from agricultural crop production and residues, Northern Nigeria has a great potential to diversify its energy production by converting these residues.

Table 2: Statistics on the production of key agricultural products in Northern Nigeria [47]

Agricultural Resource	Production Area (thousand ha)	Total Production (thousand metric tons)	State with Highest Production	Highest Production (thousand metric tons)
Cowpea	2,860	3,368	Benue	428
Cassava	3,482	42,533	Benue	3,792
Maize	4,149	7,677	Kaduna	436
Cotton	399	602	Zamfara	155
Soybeans	291	356	Benue	79
Groundnut	2,785	3,799	Niger	547
Sorghum	4,960	7,141	Kano	746
Millet	4,364	5,171	Sokoto	714
Rice	2,433	4,473	Kaduna	732

B. Municipal Solid Waste (MSW)

Due to overpopulation in urban areas, MSW is generated in commercial, industrial, and household sectors. MSW could be plastics, glass, metal, paper, wood, textiles, and other biodegradables. Millions of tons of household waste are collected annually, and the majority of this waste is disposed of in landfills [52, 53]. There exists a variation of waste according to regions or cities; however, as a low-income country, the average waste generation in Nigeria is at 0.5 kg/capita/day [54]. Methane could be produced through incineration or anaerobic digestion/degradation. However, proper infrastructure and waste management systems are necessary, requiring better-designed landfills and a defined process for the collection, handling, and management of the waste [18].

C. Human and Animal Waste

There is a common cultural dislike from both public and regulatory bodies to utilize animal waste for energy production, as it can contaminate both ground and surface water resources [55, 56].

There are more widespread agricultural activities in both livestock and crop farming in the northern regions and households in Nigeria, as illustrated in Figure 3 [57]. By observation, the same is true for human waste production since the population of the Northern regions is higher that of the southern regions by almost 7 million people.

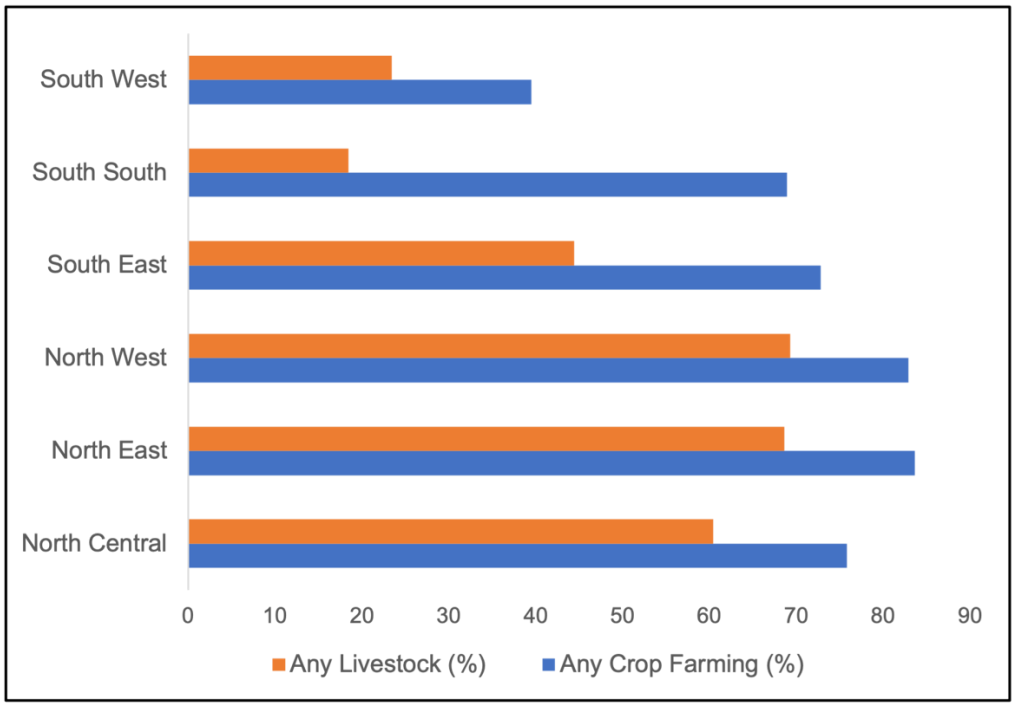


Fig. 3. The proportion of the populace engaging in agricultural pursuits in Nigeria [57]

D. Compositional Characteristics of Northern Nigeria's Crop Residues

It is noteworthy that the physiochemical properties of agricultural residues determine their energy potential. Table 3 summarizes key characteristics of Northern Nigeria's predominant crop residues, which aids in selecting more efficient conversion technologies [18].

From Table 3, maize cobs and groundnut shells offer the highest energy density, with 25,330 KJ/kg and 18,130 KJ/kg, respectively, which makes them ideal for thermochemical conversion, gasification. The average LHV of 17,279 for all the crops (Table 3) shows strong energy output from all the residues. On the other hand, rice husks and millet straws have around 14,000–16,410 kJ/kg and 15,400 kJ/kg, respectively, thus requiring supplemental fuels for efficient combustion. These findings inform policymakers to prioritize maize cobs and groundnut shells for bioenergy projects in order to maximize energy output in Northern Nigeria's existing biomass resource. Furthermore, this analysis entails strategic economic pathways for Northern Nigeria, as shown by high residue-to-product ratios for sorghum straws (RPR 0.85–7.40) and maize stalks (RPR 0.55–4.33), indicating low-cost feedstock for decentralized biomass energy systems.

V. BIOMASS TRANSFORMATION TECHNIQUES

Various techniques could be employed to convert biomass resources into biofuel; the techniques discussed in this section are chosen due to the maturity of the technology in the country, since different techniques could convert different biomass feedstock into either solid fuel like pellets, liquid fuels like biofuels, or gas fuel like biogas [18]. When it comes to biomass conversion technologies, there is no one perfect approach that solves every problem, as several factors must be analyzed to determine the best technology for deployment. These factors include the quantity of the resource, the type of biomass material that is available, and the form of energy (electricity, heat, fuel, etc.) that is required [58].

Pyrolysis is a thermal decomposition process that converts biomass into gas, char, and oil. This process has been shown to effectively utilize diverse feedstocks, including agricultural and municipal waste, to produce biochar, bio-oil, and gases, mostly methane [18]. On the other hand, thermochemical methods are recognized for their flexibility and efficiency in converting various biomass types into valuable liquid fuels, heat, and electricity [59]. The economic assessments indicate that these methods can be integrated into existing energy systems, enhancing their viability for large-scale production [59]. Gasification techniques have been used for partial oxidation of biomass at high temperature to generate syngas and other hydrocarbons such as propane or butane [18, 60]. Anaerobic digestion is a crucial biotechnological process for converting organic waste into biogas, a renewable energy source. Key factors influencing biogas yield include temperature, pH, and the composition of organic waste [61]. Biodiesel production using Jatropha curcas and vegetable oils has garnered significant attention due to its potential as a sustainable and renewable energy source. Jatropha curcas, a non-edible oilseed crop, is particularly promising since it does not compete with food resources, making it an attractive alternative to edible oils for biodiesel production [62]. Ethanol can be made by processing sugars and starches from biomass like energy crops or agricultural residues like cassava and sugarcane, which are grown in parts of Northern Nigeria [18]. The production process of ethanol typically involves pretreatment, hydrolysis, fermentation, and distillation. Yeast and bacteria are commonly used microorganisms for fermentation, with Saccharomyces cerevisiae being the most widely used yeast strain [63].

Table 3. Ultimate and Proximate Analysis of Major Crop Residues in Northern Nigeria, 2017 [18]

Crop	**Residue**	**RPR**	**Proximate Analysis (%)**				**Ultimate Analysis (%)**						**LHV (kJ/kg)**
			Ash (%)	**Volatiles (%)**	**Fixed (%)**	**Carbon**	**C (%)**	**H (%)**	**O (%)**	**S (%)**	**N (%)**	**Cl (%)**	
Cowpea	Shells	1.20–1.90	5.9	75.3	18.8		43	5.6	43.3	0.01	0.6	0.13	17,900
Cassava	Peels	0.36–0.91	11.7	59.4	28.9		22.1	13.5	37.3	1.82	2.38		16,400
	Stalks	0.20–1.00	5.7	76	18.3		48.8	6.7	43.4	-	1.1	-	17,000
Maize	Cobs	0.20–1.80	1.6	84.3	14.1		46.2	5.4	47.2	0.2	0.9	-	25,330
	Husk	0.20–0.30	34.4	55.2	10.4		31.1	3.6	32.6	0.5	1.1	-	16,370–19,900
	Stalks	0.55–4.33	6.3	73.4	20.3		41.9	5.4	51.3	0.1	1.4	-	17,740
Soybean	Straws	0.80–3.94	4	88.8	7.2		45	6.7	45.4	-	2.9	-	17,900
Groundnut	husks/shells	0.37–1.20	3.1	68.1	28.8		49.3	7.3	39.1	0.02	1.1	-	13,785–18,130
Sorghum	Straws	0.85–7.40	8.1	73.4	18.5		39.5	7.5	43	0.2	1.1	0.6	15,400
Millet	Straws	0.95–2.00	2.7	94.1	3.2		42.7	6	33	-	0.1	-	15,400
Rice	Husks	0.17–0.35	15.8	69.3	14.9		38.2	5.9	-	0.1	0.7	-	14,000–16,410
	Straws	0.40–3.96	21.5	62.6	15.9		28.6	4	65.7	0.6	1.2	-	12,440

Key: RPR = Residue-to-Product Ratio; LHV = Lower Heating Value;"-" = Data not available

VI. Energy efficiency of small-scale cassava ethanol production in Nigeria

The Cassakero Project is a Nigerian government initiative to replace kerosene with cassava-derived ethanol for house cooking [64]. This project offers critical insight into energy balance and practical challenges of small-scale biofuel production. The project reported a Net Energy Ratio (NER) of 1.20 with agrochemical and 1.34 without it, therefore indicating a marginal energy efficiency. Renewable energy sources, including wood fuel, organic fertilizer, and hydropower, contributed to 75-84% of total energy inputs, with a Net Energy Gain (NEG) ranging from 2.29 with agrochemicals to 3.52 MJ per liter of ethanol without agrochemicals, thus saving more energy per liter [64]. The outcome of the project showed that the ethanol conversion stage accounts for 74-83% of total energy demand because of inefficient distillation practices. NER seems to be higher without the agrochemicals at 12.25 MJ/l vs. 11.01 MJ/l, though it reduces cassava yields (10.8 vs. 20 tonnes/ha), thus highlighting a trade-off between energy efficiency and feedstock productivity. Sensitivity analysis of ±10% energy input variation validates the efficiency of the project [64].

In the case of Northern Nigeria, where cassava cultivation is limited, agricultural residues such as sorghum and millet stalks may offer higher energy returns due to lower processing demands. This reduces competition with food crops and aligns with the region's agro-ecological strengths, given limited use of agrochemicals in the region, which may result in lower crop yields but greater surplus energy, higher efficiency, and higher renewable energy contribution, as is the case with the Cassakero project.

VII. Economic viability of biomass versus fossil fuels

Biomass feedstocks are locally sourced and, when compared to fossil fuels, are inexpensive and easily accessible, thereby reducing import dependence and enhancing energy security [65]. However, due to limited economies of scale, small ethanol production plants usually incur higher production costs per liter than fossil fuels, which are produced in centralized, large-scale refineries. Biomass direct combustion and biogas power generation show favorable internal rates of return (19.16% and 13.49%, respectively) and shorter payback periods (7.71 years for direct combustion) compared to fossil fuels [66]. Moreover, biomass energy systems generate rural employment and boost the local economy by retaining income in the economy and offer lower greenhouse gas emissions when burnt [66].

A. Energy Crops and Food Crops—Conflict and Symbioses In Nigeria

Biofuels are simply fuels produced from living things or their waste products. They are seen as greener and more sustainable alternatives to fossil fuels [67]. In low-income countries, producing biofuels from edible foods poses risks to food security [68]. However, biomass energy from non-edible substrates has garnered increased attention as it provides benefits such as clean energy production, better use of agricultural waste, improved health, and more job opportunities [69, 70].

As highlighted in the preceding sections of this article, the Northern regions of Nigeria have enormous potential to produce bioenergy from agricultural residues, municipal solid wastes, forest residues, etc. Energy or agricultural crops produced as feedstock for biofuel production include crops such as oil plants (such as Jatropha), corn, sorghum, rapeseed, sugarcane, grasses (switchgrass, miscanthus, alfalfa, reed canary grass, etc.) [71], and trees (willow, poplar, eucalyptus, and paulownia), which can reproduce after being felled and can be harvested every 3-5 years [72]. It is noteworthy that Northern Nigeria has vast land areas suitable for intensive and mechanized cultivation of food and energy crops compared to the regions in the south [69]. While maize, millet, wheat, rice, potatoes, sugarcane, and sorghum could all be used for ethanol production, cassava is the most profitable feedstock, followed by sugarcane in Nigeria, mainly due to the country's established expertise in producing these crops [67].

Jatropha, a second-generation energy crop, could be cultivated in all ecological zones of Nigeria. It is non-edible and therefore does not compete with food; it has highly resilient growth capabilities, and it can grow very well even in harsh weather conditions and low rainfall as found in most Northern States in Nigeria [73].

A continuous large-scale production of important energy crops for bioenergy production will cause demand for food crops to rise; consequently, the prices of these crops will also rise [74]. This trend of energy crops for biofuel production must be carefully analyzed, as most Nigerian farmers struggle to cultivate enough food crops for themselves. Utilizing land and water resources for energy crops could exacerbate issues, increase poverty, and limit access to food [67]. After all, more than 60% of the country's population lives below the poverty level [75], primarily due to low income and poverty [76]. As biofuel production could compromise food security, it is essential to assess poverty and how biofuel production could impact the four pillars of food security: availability, access, stability, and utilization [67].

VIII. Environmental implications of bioenergy production

While biomass is often promoted as a renewable energy source, its production can lead to significant environmental degradation. For instance, expansion of energy crops or agro-residues for ethanol production may lead to the conversion of forests and grasslands, and this directly affects biodiversity and increases carbon emissions, vegetation loss, and soil disturbance [78]. Although the use of agrochemicals can boost harvest yields, such practices could degrade soil health over time and affect water resources [79].

Biomass energy production can result in substantial greenhouse gas emissions, particularly during the cultivation, harvesting, and processing stages [80]. A number of pollutants, including carbon dioxide (CO_2), carbon monoxide (CO), nitrogen oxides (NO_x), and particulate matter (PM), could be produced during combustion and upstream processes, though specific emissions depend on the type of biomass and combustion method used [81]. For example, studies on biomass electricity generation systems indicate that agricultural residues have a higher greenhouse gas emission with 291.25 gCO_2e/kWh; while forestry and

industrial wastes have lower emissions with 43 gCO2e/kWh and 45.93 gCO2e/kWh, respectively; and dedicated energy crops gave 208.41 gCO2e/kWh [82].

Therefore, long-term sustainability of bioenergy production requires careful strategic planning and management as it cuts across several dimensions, like environmental, economic, and social factors [83]. Hence, life cycle assessments (LCA) are crucial for evaluating the environmental footprint of bioenergy production, including greenhouse gas emissions, land use, and water consumption. These assessments help identify the potential for carbon dioxide removal, which can range from 45 to 99%, depending on the bioenergy process used [84]. Economic sustainability involves ensuring that bioenergy production is cost-effective and competitive with other energy sources. This requires careful management of supply chains and the use of optimization techniques to enhance efficiency [85]. While social sustainability focuses on the impact of bioenergy production on communities, including job creation and energy access. A multidimensional approach that considers cultural, institutional, and technical factors is necessary to address these social aspects effectively [86].

IX. Opportunities and Challenges

The Nigerian Government Biomass Program (2005) was a deliberate intervention aimed at encouraging the utilization of biomass resources for automotive applications. The program primarily focused on exploring potential biomass-derived fuels, such as bioethanol and biodiesel, as alternatives to conventional fuels in the automotive sector. Its objectives included reducing reliance on imported gasoline and minimizing environmental pollution from fossil fuels [77]. The policy aimed at raising domestic production of biofuel to 100% in 2020 along with other key projections, such as increasing 1.3 billion liters of ethanol (E-10 blend) from 2010 to 2 billion liters by 2020 and increasing 480 million liters of biodiesel from 2010 to 900 million liters by 2020.

In 2007, Nigeria launched an ambitious biofuel initiative to supply the domestic market with a 10% ethanol blend with gasoline (E10) [88] with an audacious amount of over US$ 3.86 billion investment for 19 ethanol biorefineries in the country [89]. Yet, there is not any operational large-scale commercial bioethanol plant in the country [90].

Northern Nigeria has huge potential for biomass resources from crops, forest residues, and various conversion technologies that could take advantage of these resources to provide cleaner energy. Hence, there should be a clear collaboration between the government and relevant institutions for the development of the biofuel industry in Northern Nigeria, since the region has the highest established potential for ethanol production in the country (Figure 4) [40]. The government can establish a clear framework and direction, while the institutions play a central role in research, development, and implementation. At both domestic and international levels, government involvement is crucial. Their strategies will help ensure the implementation and sustainable development of initiatives, tackling issues such as raw material shortages, high costs, and competition with traditional fossil fuels [90]. Meanwhile, institutions play a vital role in determining the industries' success, as they provide the necessary rules and structure for the bioenergy market to develop, promote collaboration between stakeholders, reduce uncertainties, and facilitate efficient resource allocation [91].

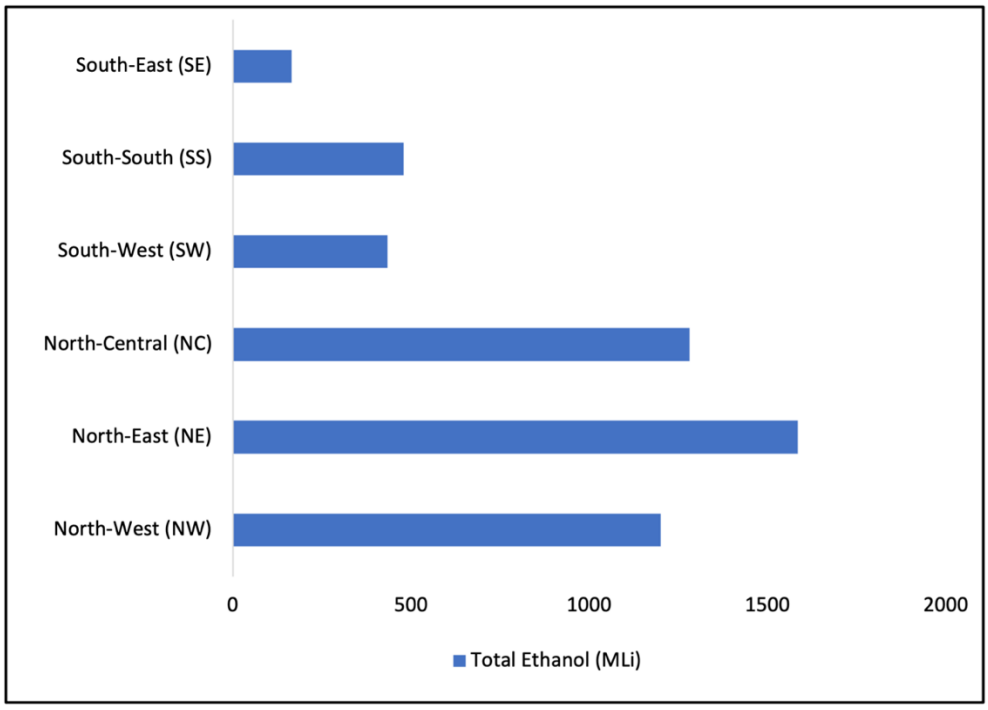


Fig. 4. Ethanol production potential by region – Nigerian Energy-Food Model [40]

X. Conclusion

Reliable and affordable energy is essential for economic growth and progress, and its importance to other aspects of modern civilization cannot simply be overstated. In Northern Nigeria, there is a significant potential in the exploration of biomass as a renewable energy source. With vast arable landmass and extensive agricultural activities, there exist numerous opportunities for sustainable energy development that could alleviate the poverty rate in the region, provide job employment, infrastructure development, and basic social needs such as electricity and water supply. Furthermore, following a sustainable strategy and cultivating energy crops presents a viable solution to meet energy demands without encroaching on food crop production, thereby eliminating conflicts between food and biofuel production.

Government policies and established relevant institutions play a pivotal role in accelerating and implementing the biofuel policy, actualizing it, and continuing its development to transition Nigeria's fossil-based energy systems and encourage its utilization. The current policies have laid the groundwork, but still, biofuels are not being used on a large scale, at least not even half their potential, and there exist several barriers and confusing frameworks for management and implementation of the initiatives. Other areas of improvement include the provision of relevant incentives and the reduction of bureaucratic processes to hasten the decision-making processes and investment in the sector.

References

[1] N. Scafetta, "Impacts and risks of 'realistic' global warming projections for the 21st century," *Geosci. Front.*, vol. 15, 2024, Art. no. 101774. [Online]. Available: https://doi.org/10.1016/j.gsf.2023.101774.

[2] H. Jie, I. Khan, M. Alharthi, M. W. Zafar, and A. Saeed, "Sustainable energy policy, socio-economic development, and ecological footprint: The economic significance of natural resources, population growth, and industrial development," *Util. Policy*, vol. 81, 2023, Art. no. 101490. [Online]. Available: https://doi.org/10.1016/j.jup.2023.101490.

[3] M. Sheraz, Q. Qin, and M. Z. Mumtaz, "Energy transition in OECD countries: Catalyzing governance quality for SDG 7 attainment," *Energy Policy*, vol. 194, 2024, Art. no. 114315. [Online]. Available: https://doi.org/10.1016/j.enpol.2024.114315.

[4] R. El-Araby, "Biofuel production: exploring renewable energy solutions for a greener future," *Biotechnol. Biofuels Bioprod.*, vol. 7, 2024, Art. no. 129. [Online]. Available: https://doi.org/10.1186/s13068-024-02571-9.

[5] U. Lawan, "Biofuels as the starring substitute to fossil fuels," *Petrol. Sci. Eng.*, vol. 2, 2018. [Online]. Available: https://doi.org/10.11648/j.pse.20180201.17.

[6] N. S. Chipangamate and G. T. Nwaila, "Assessment of challenges and strategies for driving energy transitions in emerging markets: A socio-technological systems perspective," *Energy Geosci.*, vol. 5, 2024, Art. no. 100257. [Online]. Available: https://doi.org/10.1016/j.engeos.2023.100257.

[7] M.-T. Huang and P.-M. Zhai, "Achieving Paris Agreement temperature goals requires carbon neutrality by mid-century with far-reaching transitions in the whole society," *Adv. Clim. Change Res.*, vol. 12, pp. 281–286, 2021. [Online]. Available: https://doi.org/10.1016/j.accre.2021.03.004.

[8] N. S. Caetano, T. M. Mata, A. A. Martins, and M. C. Felgueiras, "New trends in energy production and utilization," *Energy Procedia*, vol. 128, pp. 1–8, 2017. [Online]. Available: https://doi.org/10.1016/j.egypro.2016.12.122.

[9] Y. Sugiawan and S. Managi, "New evidence of energy-growth nexus from inclusive wealth," *Renew. Sustain. Energy Rev.*, vol. 103, pp. 40–48, 2019. [Online]. Available: https://doi.org/10.1016/j.rser.2018.12.044.

[10] S. Banerjee, R. B. Tasnim, F. Zhafran, S. Syafiqah, S. T. Thirugnana, and D. Ir., et al., "Sustainability of renewable energy systems with special reference to ocean thermal energy conversion schemes," *Non-Met. Mater. Sci.*, vol. 4, 2022. [Online]. Available: https://doi.org/10.30564/nmms.v4i2.5023.

[11] M. A. Abdelkareem, M. El Haj Assad, E. T. Sayed, and B. Soudan, "Recent progress in the use of renewable energy sources to power water desalination plants," *Desalination*, vol. 435, pp. 97–113, 2018. [Online]. Available: https://doi.org/10.1016/j.desal.2017.11.018.

[12] M. O. Ukoba, E. O. Diemuodeke, T. A. Briggs, K. E. Okedu, K. Owebor, M. Imran, et al., "Assessment of biomass energy potentials and appropriate sites in Nigeria using GIS computing strategy," *Pan Afr. Conf. Sci. Comput. Telecommun. (PACT)*, vol. 5, pp. 122–127, 2023. [Online]. Available: https://ictjournal.icict.org.zm/index.php/icict/article/view/289.

[13] R. York and S. E. Bell, "Energy transitions or additions?," *Energy Res. Soc. Sci.*, vol. 51, pp. 40–43, 2019. [Online]. Available: https://doi.org/10.1016/j.erss.2019.01.008.

[14] B. Salman, M. Y. Ong, S. Nomanbhay, A. A. Salema, R. Sankaran, and P. L. Show, "Thermal analysis of Nigerian oil palm biomass with sachet-water plastic wastes for sustainable production of biofuel," *Processes*, vol. 7, 2019, Art. no. 475. [Online]. Available: https://doi.org/10.3390/pr7070475.

[15] R. O. Yusuf, J. A. Adeniran, S. I. Mustapha, and J. A. Sonibare, "Energy recovery from municipal solid waste in Nigeria and its economic and environmental implications," *Environ. Qual. Manag.*, vol. 28, pp. 33–43, 2019. [Online]. Available: https://doi.org/10.1002/tqem.21617.

[16] B. Ugwoke, O. Gershon, C. Becchio, S. P. Corgnati, and P. Leone, "A review of Nigerian energy access studies: The story told so far," *Renew. Sustain.*

Energy Rev., vol. 120, 2020, Art. no. 109646. [Online]. Available: https://doi.org/10.1016/j.rser.2019.109646.

[17] K. E. Okedu, B. C. Oyinna, E. O. Diemuodeke, I. Colak, and A. Kalam, "Multicriteria GIS-based assessment of biomass energy and hydropower potentials in Nigeria," *Meas.: Sens.*, vol. 33, 2024, Art. no. 101243. [Online]. Available: https://doi.org/10.1016/j.measen.2024.101243.

[18] S. O. Jekayinfa, J. I. Orisaleye, and R. Pecenka, "An assessment of potential resources for biomass energy in Nigeria," *Resources*, vol. 9, 2020, Art. no. 92. [Online]. Available: https://doi.org/10.3390/resources9080092.

[19] Y. B. Aemro, P. Moura, and A. T. de Almeida, "Inefficient cooking systems a challenge for sustainable development: A case of rural areas of Sub-Saharan Africa," *Environ. Dev. Sustain.*, vol. 23, pp. 14697–14721, 2021. [Online]. Available: https://doi.org/10.1007/s10668-021-01266-7.

[20] I. Dunmade, "Potentials for sustainable power supply in Nigeria: An overview of energy resources in western Nigeria," *Int. J. Energy Power Eng.*, vol. 6, p. 34, 2017. [Online]. Available: https://doi.org/10.11648/j.ijepe.20170603.13.

[21] N. Newman, "Off-the-grid thinking to end Nigeria's blackouts," *Eng. Technol.*, vol. 14, pp. 68–71, 2019. https://doi.org/10.1049/et.2019.0209.

[22] K. A. Babatunde, O. O. Agbede, I. I. Olateju, S. D. Bamidele, O. M. Osuolale, and F. N. Osuolale, "Biomass gasification potential in Nigeria: A review," *LAUTECH J. Civ. Environ. Stud.*, vol. 3, 2019. https://doi.org/10.36108/laujoces/9102/20(0280).

[23] O. A. Jolaosebikan, "Epileptic electric power generation and supply in Nigeria: causes, impact and solution," *J. Contemp. Res. Soc. Sci.*, vol. 1, pp. 73–81, 2019. https://doi.org/10.33094/26410249.2019.13.73.81.

[24] S. O. Boluwaduro and E. Boluwaduro, "Fuelwood: energy, belief system and popular cultural practices in Nigeria," *Afr. J. Relig. Philos. Cult.*, vol. 4, 2023.

[25] F. O. Olanrewaju, G. E. Andrews, H. Li, and H. N. Phylaktou, "Bioenergy potential in Nigeria," *Chem. Eng. Trans.*, vol. 74, 2019. https://doi.org/10.3303/CET1974011.

[26] D. O. Obada, M. Muhammad, S. B. Tajiri, M. O. Kekung, S. A. Abolade, S. B. Akinpelu, and A. Akande, "A review of renewable energy resources in Nigeria for climate change mitigation," *CSCEE*, vol. 9, 2024, Art. no. 100669. https://doi.org/10.1016/j.cscee.2024.100669.

[27] C. Ogbonnaya, C. Abeykoon, U. M. Damo, and A. Turan, "The current and emerging renewable energy technologies for power generation in Nigeria: A review," *TSEP*, vol. 13, 2019, Art. no. 100390. https://doi.org/10.1016/j.tsep.2019.100390.

[28] A. A. Bisu, T. G. Ahmed, U. S. Ahmad, and A. D. Maiwada, "A SWOT Analysis Approach for the Development of Photovoltaic (PV) Energy in Northern Nigeria," *CLES*, vol. 9, 2024, Art. no. 100128. https://doi.org/10.1016/j.cles.2024.100128.

[29] B. Olomiyesan, O. Oyedum, P. Ugwuoke, and M. Abolarin, "Assessment of wind energy resources in Nigeria – a case study of north-western region of Nigeria," *Int. J. Phys. Res.*, vol. 5, pp. 83–91, 2017. https://doi.org/10.14419/ijpr.v5i2.8327.

[30] C. O. Ugwu, P. A. Ozor, and C. Mbohwa, "Small hydropower as a source of clean and local energy in Nigeria: Prospects and challenges," *JFUECO*, vol. 10, 2022, Art. no. 100046. https://doi.org/10.1016/j.jfueco.2021.100046.

[31] Geoinfotech Resources Limited, *Agro-ecological zones in Nigeria*, 2024. [Online]. Available: https://geoinfotech.ng/agroecological-zones-in-nigeria/ (accessed Dec. 5, 2024).

[32] F. Sedano, V. Molini, and M. A. K. Azad, *The state of land use in Northern Nigeria: a Landsat-based mapping framework*, 2020. https://doi.org/10.1596/1813-9450-9335.

[33] Index Mundi, *Nigeria land use*, 2018. [Online]. Available: https://www.indexmundi.com/nigeria/land_use.html (accessed Apr. 7, 2024).

[34] P. A. Adedibu, A. A. Opeyemi, A. J. Lawrence, J. I. Paul, and E. Oguntoye, "Savanna biomes in Nigeria: indicator species and plant adaptation strategies," 2022. https://doi.org/10.14293/S2199-1006.1.SOR-.PPDXHKL.v1.

[35] A. Ayanlade, O. D. Jeje, J. O. Nwaezeigwe, O. O. I. Orimoogunje, and O. S. Olokeogun, "Rainfall seasonality effects on vegetation greenness in different ecological zones," *Environ. Chall.*, vol. 4, 2021, Art. no. 100144. https://doi.org/10.1016/j.envc.2021.100144.

[36] Federal Department of Forestry, Federal Ministry of Environment, Federal Republic of Nigeria, *National Forest Reference Emission Level (FREL) for the Federal Republic of Nigeria*, Abuja, 2019. [Online]. Available: https://redd.unfccc.int/files/2019_submission_frel_nigeria.pdf.

[37] B. V. Bado, A. Whitbread, and M. L. S. Manzo, "Improving agricultural productivity using agroforestry systems: performance of millet, cowpea, and ziziphus-based cropping systems in West Africa Sahel," *Agric. Ecosyst. Environ.*, vol. 305, 2021, Art. no. 107175. https://doi.org/10.1016/j.agee.2020.107175.

[38] K. A. Nketia, T. A. Adjadeh, and S. G. K. Adiku, "Evaluation of suitability of some soils in the forest-Savanna transition and the Guinea Savanna Zones of Ghana for maize production," *West Afr. J. Appl. Ecol.*, vol. 26, pp. 61–73, 2018.

[39] K. Fan, P. Liu, P. Mao, J. Yao, and R. Zang, "The turnover dynamics of woody plants in a tropical lowland rain forest during recovery following anthropogenic disturbances," *J. Environ. Manage.*, vol. 342, 2023, Art. no. 118371. https://doi.org/10.1016/j.jenvman.2023.118371.

[40] Y. A. Yahya, M. S. Sadiq, and Y. Garba, "Socioeconomic characteristics and utilization of crop residues by crop-livestock farmers in (Northern Guinea Savanna)," in *Nigerian Society for Animal Production (NSAP) 46th Annual Conference–Dutsin-Ma 2021 Book of Proceedings*, Kaduna, 2021.

[41] A. Camia, J. Giuntoli, R. Jonsson, N. Robert, N. Cazzaniga, G. Jasinevičius, *et al.*, "The use of woody biomass for energy production in the EU," Publications Office, 2021. [Online]. Available: https://data.europa.eu/doi/10.2760/831621

[42] S. Arora, L. K. Sarao, and A. Singh, "Bioenergy from cellulose of woody biomass," in *Agroindustrial Waste for Green Fuel Application*, N. Srivastava, B. Verma, and P. Mishra, Eds. Singapore: Springer, 2023, pp. 89–120.

[43] S. Pandey and J. T. Erbaugh, "Driving sustainable uptake: a systematic review of global literature on policies governing woody biomass for energy," *Discover Sustainability*, vol. 5, 2024, Art. no. 28. https://doi.org/10.1007/s43621-024-00205-6

[44] Mongabay, "Statistics: Nigeria," 2023. [Online]. Available: https://worldrainforests.com/deforestation/forest-information-archive/Nigeria.htm#google_vignette (accessed Mar. 8, 2025).

[45] A. A. Adeleke, N. Petrus, S. Ayuba, A. M. Yahya, P. P. Ikubanni, I. S. Okafor, *et al.*, "Nigerian biomass for bioenergy applications: a review on the potential and challenges," *Journal of Renewable Materials*, vol. 11, 2023, pp. 4123–4141.

[46] N. Zandi Atashbar, N. Labadie, and C. Prins, "Modelling and optimisation of biomass supply chains: a review," *International Journal of Production Research*, vol. 56, 2017, pp. 3482–3506. https://doi.org/10.1080/00207543.2017.1343506

[47] M. Saleem, "Possibility of utilizing agriculture biomass as a renewable and sustainable future energy source," *Heliyon*, vol. 8, 2022, Art. no. e08905. https://doi.org/10.1016/j.heliyon.2022.e08905

[48] H. Yusuf, Y. Garba, and M. A. Adam, "Proximate composition, mineral contents and fatty acids profile of selected crop residues across savannah agro-ecological zones of northern Nigeria," *Fudma Journal of Agriculture and Agricultural Technology*, vol. 8, 2022, pp. 212–222. https://doi.org/10.33003/jaat.2022.0801.085

[49] N. I. Mohammed, N. Kabbashi, and A. Alade, "Significance of agricultural residues in sustainable biofuel development," in *Agricultural Waste and Residues*, A. Aladjadjiyan, Ed. IntechOpen, 2018. https://doi.org/10.5772/intechopen.78374

[50] B. Ugwoke, R. Tieman, A. Mill, D. Denkenberger, and J. M. Pearce, "Quantifying alternative food potential of agricultural residue in rural communities of sub-Saharan Africa," *Biomass*, vol. 3, 2023, pp. 138–162. https://doi.org/10.3390/biomass3020010

[51] T. O. Ajewole, F. B. Elehinafe, O. B. Okedere, and T. E. Somefun, "Agro-residues for clean electricity: A thermo-property characterization of cocoa and kolanut waste blends," *Heliyon*, vol. 7, 2021, Art. no. e08055. https://doi.org/10.1016/j.heliyon.2021.e08055

[52] A. A. Sokan-Adeaga and G. R. E. E. Ana, "A comprehensive review of biomass resources and biofuel production in Nigeria: Potential and prospects," *Reviews on Environmental Health*, vol. 30, 2015. https://doi.org/10.1515/reveh-2015-0015

[53] S. Nanda and F. Berruti, "Municipal solid waste management and landfilling technologies: a review," *Environmental Chemistry Letters*, vol. 19, 2021, pp. 1433–1456. https://doi.org/10.1007/s10311-020-01100-y

[54] C. C. Nnaji, "Status of municipal solid waste generation and disposal in Nigeria," *Management of Environmental Quality: An International Journal*, vol. 26, 2015. https://doi.org/10.1108/MEQ-08-2013-0092

[55] S. Gwara, E. Wale, A. Odindo, and C. Buckley, "Attitudes and perceptions on the agricultural use of human excreta and human excreta-derived materials: A scoping review," *Agriculture*, vol. 11, 2021, Art. no. 153. https://doi.org/10.3390/agriculture11020153

[56] N. Boyd Williams, R. S. Quilliam, B. Campbell, D. Raha, D. C. Baruah, M. L. Clarke, *et al.*, "Challenging perceptions of socio-cultural rejection of a taboo technology: Narratives of imagined transitions to domestic toilet-linked biogas in India," *Energy Research & Social Science*, vol. 92, 2022, Art. no. 102802. https://doi.org/10.1016/j.erss.2022.102802

[57] D. D. Sasu, "Share of households participating in agricultural activities in Nigeria in 2019, by zone," *Statista*, 2022. [Online]. Available: https://www.statista.com/statistics/1119593/households-participating-in-agricultural-activities-in-nigeria-by-type-and-area/#:~:text=According%20to%20the%20results%20of (accessed Dec. 6, 2024).

[58] P. McKendry, "Energy production from biomass (part 2): Conversion technologies," *Bioresource Technology*, vol. 83, 2002. https://doi.org/10.1016/S0960-8524(01)00119-5

[59] A. M. Elgarahy, A. S. Hammad, D. M. El-Sherif, M. Abouzid, M. S. Gaballah, M. S. Gaballah, *et al.*, "Thermochemical conversion strategies of biomass to biofuels, techno-economic and bibliometric analysis: A conceptual review," *Journal of Environmental Chemical Engineering*, vol. 9, 2021, Art. no. 106503. https://doi.org/10.1016/J.JECE.2021.106503

[60] H. N. Abubackar, M. C. Veiga, and C. Kennes, "Syngas fermentation for bioethanol and bioproducts," in *Sustainable Resource Recovery and Zero Waste Approaches*, 2019, pp. 207–221.

https://doi.org/10.1016/B978-0-444-64200-4.00015-3

[61] J. Zhao, S. Ren, C. Li, M. Jiao, G. Wu, and H. Chen, "Research progress and perspectives of biogas production from municipal organic solid waste," *Int. J. Chem. React. Eng.*, 2024. [Online]. Available: https://doi.org/10.1515/ijcre-2023-0082.

[62] J. V. L. Ruatpuia, G. Halder, M. Vanlalchhandama, F. Lalsangpuii, R. Boddula, N. Al-Qahtani, *et al.*, "*Jatropha curcas* oil a potential feedstock for biodiesel production: A critical review," *Fuel*, vol. 370, 2024, Art. no. 131829. [Online]. Available: https://doi.org/10.1016/j.fuel.2024.131829.

[63] H. Zhang, P. Zhang, T. Wu, and H. Ruan, "Bioethanol production based on *Saccharomyces cerevisiae*: Opportunities and challenges," *Fermentation*, vol. 9, 2023, Art. no. 709. [Online]. Available: https://doi.org/10.3390/fermentation9080709.

[64] E. I. Ohimain, "Energy analysis of small-scale ethanol production from cassava: A case study of the Cassakero Project in Nigeria," *J. Technol. Innov. Renew. Energy*, vol. 2, 2013. [Online]. Available: https://doi.org/10.6000/1929-6002.2013.02.02.4.

[65] D. Z. Hodges, "Economic assessment of biomass based power generation," *IntechOpen eBooks*, 2022. [Online]. Available: https://doi.org/10.5772/intechopen.103692.

[66] S. Chen, H. Feng, J. Zheng, J. Ye, Y. Song, H. Yang, and M. Zhou, "Life cycle assessment and economic analysis of biomass energy technology in China," *Processes*, vol. 8, 2020, Art. no. 1112. [Online]. Available: https://doi.org/10.3390/PR8091112.

[67] S. Matemilola, I. O. Elegbede, F. Kies, G. A. Yusuf, G. N. Yangni, and I. Garba, "An analysis of the impacts of bioenergy development on food security in Nigeria: Challenges and prospects," *Environ. Clim. Technol.*, vol. 23, 2019. [Online]. Available: https://doi.org/10.2478/rtuect-2019-0005.

[68] M. Cutazzo, "The impact of biofuels on the realization of the human right to food," in *Food Diversity Between Rights, Duties and Autonomies*, A. Isoni, M. Troisi, and M. Pierri, Eds., LITES – Legal Issues in Transdisciplinary Environmental Studies, Springer, Cham, 2018. [Online]. Available: https://doi.org/10.1007/978-3-319-75196-2_17.

[69] N. A. Dick and P. Wilson, "Analysis of the inherent energy-food dilemma of the Nigerian biofuels policy using partial equilibrium model: The Nigerian energy-food model (NEFM)," *Renew. Sustain. Energy Rev.*, vol. 98, pp. 500–514, 2018. [Online]. Available: https://doi.org/10.1016/j.rser.2018.09.043.

[70] O. Verma, M. Prashanth, A. Khosla, and N. Changotra, *Role of Bioenergy in Climate Change, Food, Energy and Rural Development*, 1st ed., CRC Press, New York, 2022.

[71] P. Yadav, P. Priyanka, D. Kumar, A. Yadav, and K. Yadav, "Bioenergy crops: Recent advances and future outlook," in *Prospects of Renewable Bioprocessing in Future Energy Systems*, A. Rastegari, A. Yadav, and A. Gupta, Eds., Biofuel and Biorefinery Technologies, Springer, Cham. [Online]. Available: https://doi.org/10.1007/978-3-030-14463-0_12.

[72] M. Abreu, L. Silva, B. Ribeiro, A. Ferreira, L. Alves, S. M. Paixão, *et al.*, "Low indirect land use change (ILUC) energy crops to bioenergy and biofuels—A review," *Energies*, vol. 15, 2022, Art. no. 4348. [Online]. Available: https://doi.org/10.3390/en15124348.

[73] P. Kumar, V. C. Srivastava, and M. K. Jha, "*Jatropha curcas* phytotomy and applications: Development as a potential biofuel plant through biotechnological advancements," *Renew. Sustain. Energy Rev.*, vol. 59, pp. 818–838, 2016. [Online]. Available: https://doi.org/10.1016/j.rser.2015.12.358.

[74] F. Taghizadeh-Hesary, E. Rasoulinezhad, and N. Yoshino, "Energy and food security: Linkages through price volatility," *Energy Policy*, vol. 128, pp. 796–806, 2019. [Online]. Available: https://doi.org/10.1016/j.enpol.2018.12.043.

[75] C. C. Okafor, C. A. Nzekwe, C. C. Ajaero, J. C. Ibekwe, and F. A. Otunomo, "Biomass utilization for energy production in Nigeria: A review," *Cleaner Energy Syst.*, vol. 5, 2022, Art. no. 100043. [Online]. Available: https://doi.org/10.1016/j.cles.2022.100043.

[76] C. Asadu, "Analytical overview of agricultural conditions in Nigeria," *Agro-Sci.*, vol. 14, p. 1, 2015. [Online]. Available: https://doi.org/10.4314/as.v14i1.1.

[77] I. Edeh and S. O. Okpo, "Biomass to biofuels: A sure sustainable energy strategy in Nigeria," *Petrol. Chem. Ind. Int.*, vol. 6, pp. 106–117, 2023.

[78] P. Yan, C. Xiao, L. Xu, G. Yu, A. Li, S. Piao, and N. He, "Biomass energy in China's terrestrial ecosystems: Insights into the nation's sustainable energy supply," *Renew. Sustain. Energy Rev.*, vol. 127, 2020, Art. no. 109857. [Online]. Available: https://doi.org/10.1016/j.rser.2020.109857.

[79] M. E. Hossain, S. Shahrukh, and S. A. Hossain, "Chemical fertilizers and pesticides: Impacts on soil degradation, groundwater, and human health in Bangladesh," 2022, pp. 63–92. [Online]. Available: https://doi.org/10.1007/978-3-030-95542-7_4.

[80] Z. J. Mather-Gratton, S. Larsen, and N. S. Bentsen, "Understanding the sustainability debate on forest biomass for energy in Europe: A discourse analysis," *PLoS One*, vol. 16, 2021, Art. no. e0246873. [Online]. Available: https://doi.org/10.1371/journal.pone.0246873.

[81] A. S. Tomlin, "Air Quality and Climate Impacts of Biomass Use as an Energy Source: A Review," *Energy Fuels*, vol. 35, pp. 14213–14240, 2021. https://doi.org/10.1021/acs.energyfuels.1c01523.

[82] A. Kadiyala, R. Kommalapati, and Z. Huque, "Evaluation of the Life Cycle Greenhouse Gas Emissions from Different Biomass Feedstock

Electricity Generation Systems," *Sustain.*, vol. 8, p. 1181, 2016. https://doi.org/10.3390/su8111181.

[83] N. Timothy, G. Joseph, O. Nellie, and K. Emily, "Sustainable biomass energy production and utilization in sub-Saharan Africa: A case study of Kenya," *J. Hortic. For.*, vol. 14, pp. 56–67, 2022. https://doi.org/10.5897/JHF2022.0689.

[84] A. I. Osman, B. Fang, Y. Zhang, Y. Liu, J. Yu, M. Farghali, *et al.*, "Life cycle assessment and techno-economic analysis of sustainable bioenergy production: A review," *Environ. Chem. Lett.*, 2024. https://doi.org/10.1007/s10311-023-01694-z.

[85] M. Roudneshin and A. Sosa, "Optimising Agricultural Waste Supply Chains for Sustainable Bioenergy Production: A Comprehensive Literature Review," *Energies*, 2024. https://doi.org/10.3390/en17112542.

[86] A. Padilla-Rivera, M. G. Paredes, and L. P. Güereca, "A systematic review of the sustainability assessment of bioenergy: The case of gaseous biofuels," *Biomass Bioenerg.*, vol. 125, pp. 79–94, 2019. https://doi.org/10.1016/j.biombioe.2019.03.014.

[87] P. Jha and S. Schmidt, "State of biofuel development in sub-Saharan Africa: How far sustainable?," *Renew. Sustain. Energy Rev.*, vol. 150, p. 111432, 2021. https://doi.org/10.1016/j.rser.2021.111432.

[88] E. I. Ohimain, "Emerging bio-ethanol projects in Nigeria: Their opportunities and challenges," *Energy Policy*, vol. 38, pp. 7161–7168, 2010. https://doi.org/10.1016/j.enpol.2010.07.038.

[89] A. Nwozor, G. Owoeye, O. Olowojolu, M. Ake, S. Adedire, and O. Ogundele, "Nigeria's quest for alternative clean energy through biofuels: An assessment," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 655, p. 012054, 2021. https://doi.org/10.1088/1755-1315/655/1/012054.

[90] S. N. Mishra, *Biofuel Development: Institutional Design Across the World*, 2021. https://doi.org/10.1007/978-3-319-71057-0_140-1.

[91] P. Papilo, M. Marimin, E. Hambali, and I. S. Sitanggang, "Institutional analysis on palm oil-based bioenergy for rural community electricity development in Indonesia: A hybrid of soft system and hard system approach," *Period. Polytech. Soc. Manag. Sci.*, vol. 29, 2021. https://doi.org/10.3311/PPSO.12995.

A Study of the M-Projective Curvature Tensor $\overline{W}^i_{jkh}$ in Generalized Recurrent and Birecurrent Finsler Spaces

A. Mohammed A. Al-Qashbari [1,2*]
F. A. M. AL-Ssallal [1,2]

Received: 29/03/2025
Revised: 28/04/2025
Accepted: 29/04/2025

© 2025 University of Science and Technology, Aden, Yemen. This article can be distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

© 2025 جامعة العلوم والتكنولوجيا، المركز الرئيس عدن، اليمن. يمكن إعادة استخدام المادة المنشورة حسب رخصة مؤسسة المشاع الإبداعي شريطة الاستشهاد بالمؤلف والمجلة.

[1] Dept. of Math's., Faculty of Educ. Aden, Univ. of Aden, Aden, Yemen
[2] Dept. of Med. Eng., Faculty of the Engineering and Computers, Univ. of Science & Technology, Aden, Yemen
*Corresponding Author's Email: adel.math.edu@aden.net a.alqashbari@ust.edu

A Study of the M-Projective Curvature Tensor $\overline{W}^{i}_{jkh}$ in Generalized Recurrent and Birecurrent Finsler Spaces

Adel Mohammed Ali Al-Qashbari[1,2]

[1] Dept. of Math's., Faculty of Educ. Aden, Univ. of Aden, Aden, Yemen
[2] Dept. of Med. Eng., Faculty of the Engineering and Computers, Univ. of Science & Technology, Aden, Yemen
Email: adel.math.edu@aden.net & a.alqashbari@ust.edu

Fahmi Ahmed Mothana AL-ssallal[3]

[3]Dept. of Math's., Faculty of Educ. Aden, Univ. of Aden,
Aden, Yemen
fahmiassallald55@gmail.com

Abstract— This paper aims to examine the properties of the M-projective curvature tensor in the context of generalized Finsler spaces, specifically within the framework of a -space. The study begins with the derivation of the M-projective curvature tensor, which is expressed as the sum of the standard M-projective curvature tensor and additional terms involving the Ricci tensor and scalar curvature. Through covariant differentiation, the behavior of this tensor under certain conditions is analyzed, leading to a set of conditions necessary for the space to exhibit generalized recurrent Finsler properties. The paper includes multiple theorems that explore these relationships, proving that the M-projective curvature tensor satisfies the conditions for generalized recurrent Finsler spaces. Additionally, the work introduces the concept of generalized birecurrent Finsler spaces and presents several characterization theorems. Finally, the results are corroborated with computational formulations that demonstrate the conditions under which the tensor relationships hold true.

Keywords— The $\mathcal{B}$ -covariant derivative of first and second orders, Generalized recurrent and birecurrent Finsler space, Weyl tensor W^{i}_{jkh} and the M-projective curvature tensor $\overline{W}^{i}_{jkh}$.

I. INTRODUCTION

The study of curvature tensors and their applications in Finsler geometry has been a subject of considerable interest due to their importance in understanding the geometric properties of spaces. The M-projective curvature tensor, denoted, is central to this paper, and its study within the context of -spaces provides new insights into the geometric structure of generalized Finsler spaces. These spaces are an extension of classical Finsler geometry, incorporating more general forms of curvature and torsion.

The main objective of this work is to investigate the properties of the M-projective curvature tensor and establish its connection to generalized recurrent Finsler spaces. A series of equations is derived that detail the behavior of the tensor under covariant derivatives, leading to several key results. These results are encapsulated in the form of theorems that provide necessary and sufficient conditions for the M-projective curvature tensor to define a generalized recurrent Finsler space. Additionally, we introduce and explore the concept of generalized birecurrent Finsler spaces, extending the classical notion of birecurrence to a more general context. The work culminates in the formulation of several important conditions that must be satisfied for the tensor to adhere to the structure of generalized Finsler spaces, thereby enriching the understanding of the geometric properties of these spaces.

The theoretical foundation laid in this study paves the way for future research into more complex geometric structures, where such tensors play a crucial role in understanding the curvature and torsion of higher-dimensional spaces. The results presented in this paper have potential applications in theoretical physics, particularly in the study of spacetime geometries and general relativity, as well as in differential geometry and mathematical physics.

The study of curvature tensors in Finsler geometry has garnered significant attention over the years, with several key contributions advancing our understanding of these geometric structures. One notable area of research focuses on the properties and applications of the w*-curvature tensor in relativistic spacetimes. Abu-Donia, Shenawy, and Abdehameed (2020) [1] investigate the w*-curvature tensor, shedding light on its role in relativistic geometries, which serves as a foundation for exploring more general tensorial structures in Finsler manifolds. This study has influenced subsequent work on curvature tensors in both general relativity and specialized geometric spaces.

Ahsan and Ali (2014, 2016) [2, 3] provide a detailed analysis of the w-curvature tensor, introducing fundamental properties and implications for spacetime in the framework of general relativity. Their work has been pivotal in understanding the behavior of curvature tensors within both classical and

relativistic contexts and has been further extended by researchers working on higher-order curvature structures.

Al-Qashbari et al. (2025) [4, 5] have made significant strides in the study of M-projective curvature tensors, focusing on their decomposition and the role of Lie derivatives in generalized Finsler spaces. Their contributions offer critical insights into higher-order tensors, providing an extensive analysis of the M-projective curvature tensor and its transvecting forms in the context of Finsler geometry. This has provided a rigorous framework for subsequent studies on generalizations of Finsler spaces, especially in terms of torsion and curvature structures.

The work of Al-Qashbari, Abdallah, and their collaborators (2024, 2025) [6, 7, 8] explores generalized recurrent Finsler spaces, utilizing both Berwald's and Cartan's covariant derivatives to define and study higher-order curvature tensors. These studies examine the interplay between torsion and curvature, further enriching the theoretical foundations of Finsler spaces, particularly in relation to the geometric structures defined by Cartan's covariant derivatives.

Further contributions by Misra et al. (2014) [9] and Goswami (2017) [10] have examined recurrent Finsler spaces with special curvature tensors, deepening the understanding of higher-order geometric structures in these spaces. Misra et al. analyze recurrent Finsler spaces with Berwald's curvature tensor field, while Goswami's systematic review provides a comprehensive overview of special Finsler spaces and their differential geometric properties. These works collectively provide a broad perspective on the evolution of curvature tensor research, especially in the context of Finsler spaces. Pandey, Saxena, and Goswami (2011) [11] focus on the generalized H-recurrent spaces, offering an in-depth exploration of their curvature tensors and their application in various geometric settings. This work complements the broader study of recurrent Finsler spaces, laying the groundwork for further investigations into generalized curvature structures and their implications in mathematical physics. Rund (1981) [12] presents a classical text on the differential geometry of Finsler spaces, offering foundational principles that underpin much of the modern study in this field. His work has influenced numerous subsequent studies, including those investigating the advanced properties of higher-order curvature tensors and their applications in both pure mathematics and theoretical physics.

In this study, several identities are explored to establish a relationship between Weyl's curvature tensor and the M-projective curvature tensor. Initially, a comprehensive introduction is provided to the foundational concepts of both Weyl's curvature tensor and the M-projective curvature tensor. Subsequently, a set of identities is driven that elucidate the connection between these two tensorial structures, contributing to the broader understanding of their geometric properties.

II. PRELIMINARIES

In this section, we establish several key conditions and definitions crucial for the development of the paper. The metric tensor and Berwald's connection coefficients are both positively homogeneous of degree 0 with respect to the directional arguments.

Consider two vectors y_i and y^i that satisfy the following relations

a) $y_i = g_{ij}\, y^j$, b) $y_i\, y^i = F^2$ and

c) $\delta_j^k y^j = y^k$. (2.1)

The metric tensors g_{ij} and g^{ij} are connected by the identity:

a) $g_{ij}\, g^{jk} = \delta_i^k = \begin{cases} 1 , & if \quad i = k \\ 0 , & if \quad i \neq k \end{cases}$,

b) $g_{ir}\, \delta_j^i = g_{rj}$ and c) $g^{jk}\delta_k^i = g^{ji}$. (2.2)

The torsion tensor C_{ijk} satisfies the condition:

$C_{ijk}\, y^i = C_{ijk}\, y^j = C_{ijk}\, y^k = 0$. (2.3)

The covariant derivative $\mathcal{B}_k T_j^i$ of any tensor T_j^i with respect to Berwald's connection is defined as

$\mathcal{B}_k T_j^i = \partial_k T_j^i - \left(\dot\partial_r T_j^i\right) G_k^r + T_j^r G_{rk}^i - T_r^i G_{jk}^r$. (2.4)

Additionally, the vector y^i and the metric function F vanish identically under Berwald's covariant derivative:

a) $\mathcal{B}_k F = 0$ and b) $\mathcal{B}_k y^i = 0$. (2.5)

The metric tensor g_{ij} is non-zero and satisfies

$\mathcal{B}_k g_{ij} = -2\, C_{ijk|h}\, y^h = -2\, y^h \mathcal{B}_h C_{ijk}$. (2.6)

The curvature tensor W_{jkh}^i , torsion tensor W_{jk}^i and deviation tensor W_j^i are defined as follows:

$$W_{jkh}^i = H_{jkh}^i + \frac{2\,\delta_j^i}{(n+1)} H_{[hk]} + \frac{2\,y^i}{(n+1)} \dot\partial_j H_{[kh]} + \frac{\delta_k^i}{(n^2-1)} \left(n\, H_{jh} + H_{hj} + y^r \dot\partial_j H_{hr} - \frac{\delta_h^i}{(n^2-1)} \left(n\, H_{jk} + H_{kj} + y^r \dot\partial_j H_{kr}\right) \right) , \quad (2.7)$$

$$W_{jk}^i = H_{jk}^i + \frac{y^i}{(n+1)} H_{[jk]} + 2\left\{ \frac{\delta_{[j}^i}{(n^2-1)} \left(n\, H_{k]} - y^r H_{k]\, r}\right) \right\} , \quad (2.8)$$

$$W_j^i = H_j^i - H\, \delta_j^i - \frac{1}{(n+1)} \left(\dot\partial_r H_j^r - \dot\partial_j H\right) y^i , \quad (2.9)$$

respectively.

Moreover, we assume that the tensor W_j^i satisfies the following identities:

a) $W_k^i\, y^k = 0$, b) $W_i^i = 0$, c) $W_k^i\, y_i = 0$,

d) $g_{ir}\, W_j^i = W_{rj}$, e) $g^{jk} W_{jk} = W$ and

f) $W_{jk}\, y^k = 0$. (2.10)

The M-projective curvature tensor $\bar W_{jkh}^i$, torsion tensor $\bar W_{kh}^i$, Ricci tensor $\bar W_{jk}$, curvature vector $\bar W_k$ and scalar curvature $\bar W$ satisfy the following relations:

a) $\bar W_{jkh}^i\, y^j = \bar W_{kh}^i$, b) $\bar W_{kh}^i\, y^k = \bar W_h^i$,

c) $\bar W_{jki}^i = \bar W_{jk}$, d) $\bar W_{ki}^i = \bar W_k$,

e) $\bar W_i^i = \bar W$ and f) $g_{ir}\, \bar W_{jkh}^i = \bar W_{rjkh}$. (2.11)

Cartan's third curvature tensor R_{jkh}^i , Ricci tensor R_{jk} , the vector H_k and scalar curvature H are defined as

a) $R_{jk}\, y^j = H_k$, b) $R_{jk}\, y^k = R_j$, c) $R_i^i = R$,

d) $g_{ir}\, R_j^i = R_{rj}$ and e) $H_k\, y^k = (n-1)H$. (2.12)

Al-Qashbari and AL-ssallal [6], Al-Qashbari, Haouse and AL-ssallal [7] introduced and studied of curvature tensor by using Berwald's and Cartan's first and second order derivatives in Finsler space which characterized by the conditions:

$$\mathcal{B}_m W_{jkh}^i = \lambda_m W_{jkh}^i + \mu_m \left(\delta_k^i g_{jh} - \delta_h^i g_{jk}\right) + \tfrac{1}{4}\gamma_m \left(W_k^i g_{jh} - W_h^i g_{jk}\right). \quad (2.13)$$

A Finsler space F_n , in which the curvature tensor W^i_{jkh} satisfies the condition (2.13), is referred to as the generalized $\mathcal{B}$W -recurrent space and denoted by G^{2nd} $\mathcal{B}$W - RF_n .

Taking the covariant derivative of (2.13) with respect to x^l in the context of Berwald's connection, we obtain:

$$\mathcal{B}_l\mathcal{B}_m W^i_{jkh} = (\mathcal{B}_l\lambda_m)W^i_{jkh} + \lambda_m(\mathcal{B}_l W^i_{jkh}) + (\mathcal{B}_l\mu_m)(\delta^i_k g_{jh} - \delta^i_h g_{jk}) + \mu_m\mathcal{B}_l(\delta^i_k g_{jh} - \delta^i_h g_{jk}) + \frac{1}{4}\gamma_m\big((\mathcal{B}_l W^i_k)g_{jh} + W^i_k(\mathcal{B}_l g_{jh}) - (\mathcal{B}_l W^i_h)g_{jk} - W^i_h(\mathcal{B}_l g_{jk})\big) + \frac{1}{4}(\mathcal{B}_l\gamma_m)(W^i_k g_{jh} - W^i_h g_{jk}). \quad (2.14)$$

By applying equations (2.6) and (2.13) to equation (2.14), we get

$$\mathcal{B}_l\mathcal{B}_m W^i_{jkh} = \lambda_{ml}W^i_{jkh} + \lambda_m\Big(\lambda_l W^i_{jkh} + \mu_l(\delta^i_k g_{jh} - \delta^i_h g_{jk}) + \frac{1}{4}\gamma_l(W^i_k g_{jh} - W^i_h g_{jk})\Big) + \mu_{ml}(\delta^i_k g_{jh} - \delta^i_h g_{jk}) - 2\mu_m\mathcal{B}_q y^q(\delta^i_k C_{jhl} - \delta^i_h C_{jkl}) + \frac{1}{4}\gamma_{ml}(W^i_k g_{jh} - W^i_h g_{jk}) + \frac{1}{4}\gamma_m\big((\mathcal{B}_l W^i_k)g_{jh} - (\mathcal{B}_l W^i_h)g_{jk}\big) - \frac{1}{2}\gamma_m\mathcal{B}_q y^q(W^i_k C_{jhl} - W^i_h C_{jkl}).$$

Or

$$\mathcal{B}_l\mathcal{B}_m W^i_{jkh} = (\lambda_{ml} + \lambda_m\lambda_l)W^i_{jkh} + (\mu_{ml} + \lambda_m\mu_l)(\delta^i_k g_{jh} - \delta^i_h g_{jk}) + \frac{1}{4}\gamma_m\big((\mathcal{B}_l W^i_k)g_{jh} - (\mathcal{B}_l W^i_h)g_{jk}\big) + \frac{1}{4}(\lambda_m\gamma_l + \gamma_{ml})(W^i_k g_{jh} - W^i_h g_{jk}) - 2\mu_m\mathcal{B}_q y^q(\delta^i_k C_{jhl} - \delta^i_h C_{jkl}) - \frac{1}{2}\gamma_m\mathcal{B}_q y^q(W^i_k C_{jhl} - W^i_h C_{jkl}) . \quad (2.15)$$

The equation (2.15), can be expressed as:

$$\mathcal{B}_l\mathcal{B}_m W^i_{jkh} = a_{ml}W^i_{jkh} + b_{ml}(\delta^i_k g_{jh} - \delta^i_h g_{jk}) + \frac{1}{4}c_{ml}(W^i_k g_{jh} - W^i_h g_{jk}) + \frac{1}{4}\gamma_m\big((\mathcal{B}_l W^i_k)g_{jh} - (\mathcal{B}_l W^i_h)g_{jk}\big) - 2\mu_m\mathcal{B}_q y^q(\delta^i_k C_{jhl} - \delta^i_h C_{jkl}) - \frac{1}{2}\gamma_m\mathcal{B}_q y^q(W^i_k C_{jhl} - W^i_h C_{jkl}) . \quad (2.16)$$

Where $a_{ml} = \lambda_{ml} + \lambda_m\lambda_l$, $b_{ml} = \mu_{ml} + \lambda_m\mu_l$ and $c_{ml} = \lambda_m\gamma_l + \gamma_{ml}$ are non-zero second order covariant tensor fields, γ_m and μ_m are non-zero first order covariant vector fields, respectively.

Definition 2.1. In a Finsler space where the Weyl's projective curvature tensor W^i_{jkh} satisfies the condition (2.16), it is referred to as a generalized BW-birecurrent space. The tensor is called a generalized B-birecurrent tensor. These spaces and tensors are denoted as $G^{2nd}\mathcal{B}W - BRF_n$ and $G^{2nd}\mathcal{B} - BR$, respectively.

Characterization of M-Projective Curvature $\overline{W}^i_{jkh}$ and Deviation Tensors in $G^{2nd}\mathcal{B}\overline{W} - RF_n$ Spaces and $G^{2nd}\mathcal{B}\overline{W} - BRF_n$ Spaces

This paper focuses on the characterization of the M-projective curvature tensor $\overline{W}^i_{jkh}$ and deviation tensors in the context of generalized Finsler spaces, specifically within the framework of $G^{2nd}\mathcal{B}\overline{W} - RF_n$ spaces. We aim to explore the geometric properties and structural relationships of these tensors, providing a comprehensive analysis of their roles in describing curvature and deviation phenomena in higher-dimensional Finsler geometries. The study contributes to the understanding of M-projective curvature in relation to generalized Finsler manifolds and offers insights into their potential applications in advanced differential geometry and theoretical physics.

Some properties of W^i_{jkh} curvature tensor was proposed by Al-Qashbari, Abdallah and Al-ssallal.

For $(n = 4)$ a Riemannian space, the M-projective curvature tensor $\overline{W}^i_{jkh}$ is given by

$$W^i_{jkh} = \overline{W}^i_{jkh} + \frac{1}{6}(\delta^i_k R_{jh} + \delta^i_h R_{jk}) - \frac{1}{6}(g_{jh}R^i_k + g_{jk}R^i_h) . \quad (3.1)$$

By taking the covariant derivative of (3.1), with respect to x^m in the context of Berwald we get

$$\mathcal{B}_m W^i_{jkh} = \mathcal{B}_m\overline{W}^i_{jkh} - \frac{1}{6}[(\mathcal{B}_m g_{jh})R^i_k + g_{jh}(\mathcal{B}_m R^i_k) + (\mathcal{B}_m g_{jk})R^i_h + g_{jk}(\mathcal{B}_m R^i_h)] + \frac{1}{6}\mathcal{B}_m(\delta^i_k R_{jh} + \delta^i_h R_{jk}). \quad (3.2)$$

Using (2.6), in the equation (3.2) can be written as

$$\mathcal{B}_m W^i_{jkh} = \mathcal{B}_m\overline{W}^i_{jkh} + \frac{1}{6}\mathcal{B}_m(\delta^i_k R_{jh} + \delta^i_h R_{jk}) - \frac{1}{6}(g_{jh}(\mathcal{B}_m R^i_k) + g_{jk}(\mathcal{B}_m R^i_h)) + \frac{1}{3}\mathcal{B}_q y^q(R^i_k C_{jhm} + R^i_h C_{jkm}). \quad (3.3)$$

By substituting equations (2.13) and (3.1) in to (3.3), we obtain:

$$\mathcal{B}_m\overline{W}^i_{jkh} + \frac{1}{6}\mathcal{B}_m(\delta^i_k R_{jh} + \delta^i_h R_{jk}) - \frac{1}{6}(g_{jh}(\mathcal{B}_m R^i_k) + g_{jk}(\mathcal{B}_m R^i_h)) + \frac{1}{3}\mathcal{B}_q y^q(R^i_k C_{jhm} + R^i_h C_{jkm}) = \lambda_m\overline{W}^i_{jkh} + \frac{1}{6}\lambda_m(\delta^i_k R_{jh} + \delta^i_h R_{jk}) - \frac{1}{6}\lambda_m(g_{jh}R^i_k + g_{jk}R^i_h) + \mu_m(\delta^i_k g_{jh} - \delta^i_h g_{jk}) + \frac{1}{4}\gamma_m(W^i_k g_{jh} - W^i_h g_{jk}) . \quad (3.4)$$

Alternatively, this can be expressed as:

$$\mathcal{B}_m\overline{W}^i_{jkh} = \lambda_m\overline{W}^i_{jkh} + \mu_m(\delta^i_k g_{jh} - \delta^i_h g_{jk}) + \frac{1}{4}\gamma_m(W^i_k g_{jh} - W^i_h g_{jk}) - \frac{1}{6}\mathcal{B}_m(\delta^i_k R_{jh} + \delta^i_h R_{jk}) + \frac{1}{6}(g_{jh}(\mathcal{B}_m R^i_k) + g_{jk}(\mathcal{B}_m R^i_h)) - \frac{1}{3}\mathcal{B}_q y^q(R^i_k C_{jhm} + R^i_h C_{jkm}) + \frac{1}{6}\lambda_m(\delta^i_k R_{jh} + \delta^i_h R_{jk}) - \frac{1}{6}\lambda_m(g_{jh}R^i_k + g_{jk}R^i_h). \quad (3.5)$$

This demonstrates that

$$\mathcal{B}_m\overline{W}^i_{jkh} = \lambda_m\overline{W}^i_{jkh} + \mu_m(\delta^i_k g_{jh} - \delta^i_h g_{jk})$$

$$+\frac{1}{4}\gamma_m\left(W_k^i g_{jh} - W_h^i g_{jk}\right)$$
$$-\frac{1}{3}\mathcal{B}_q y^q\left(R_k^i C_{jhm} + R_h^i C_{jkm}\right). \quad (3.6)$$

If and only if

$$\mathcal{B}_m\left(\delta_k^i R_{jh} + \delta_h^i R_{jk}\right) = \lambda_m\left(\delta_k^i R_{jh} + \delta_h^i R_{jk}\right),$$
$$\mathcal{B}_m R_k^i = \lambda_m R_k^i \text{ , where } g_{jh} \neq 0 \text{ , and}$$
$$\mathcal{B}_m R_h^i = \lambda_m R_h^i \text{ , where } g_{jk} \neq 0 . \quad (3.7)$$

In conclusion the proof of theorem is completed, we can determine

Theorem 3.1. In the space $\mathrm{G}^{2\mathrm{nd}}\mathcal{B}\bar{W} - \mathrm{RF_n}$, the M-projective curvature tensor $\bar{W}_{jkh}^i$ represents a generalized recurrent Finsler space, provided that the condition (3.7) is satisfied.

By transecting equation (3.5) with y^j and utilizing equations (2.11a), (2.5b), (2.1a), (2.3) and (2.12a), we obtain the following result:

$$\mathcal{B}_m\bar{W}_{kh}^i = \lambda_m\bar{W}_{kh}^i + \mu_m\left(\delta_h^i y_k - \delta_k^i y_h\right)$$
$$+\frac{1}{4}\gamma_m\left(W_k^i y_h - W_h^i y_k\right) - \frac{1}{6}\mathcal{B}_m\left(\delta_k^i H_h + \delta_h^i H_k\right)$$
$$+\frac{1}{6}\left(y_h(\mathcal{B}_m R_k^i) + y_k(\mathcal{B}_m R_h^i)\right)$$
$$+\frac{1}{6}\lambda_m\left(\delta_k^i H_h + \delta_h^i H_k\right) - \frac{1}{6}\lambda_m\left(y_h R_k^i + y_k R_h^i\right). \quad (3.8)$$

This demonstrates that

$$\mathcal{B}_m\bar{W}_{kh}^i = \lambda_m\bar{W}_{kh}^i + \mu_m\left(\delta_k^i y_h - \delta_h^i y_k\right)$$
$$+\frac{1}{4}\gamma_m\left(W_k^i y_h - W_h^i y_k\right). \quad (3.9)$$

If and only if

$$\mathcal{B}_m\left(\delta_k^i H_h + \delta_h^i H_k\right) = \lambda_m\left(\delta_k^i H_h + \delta_h^i H_k\right),$$
$$\mathcal{B}_m R_k^i = \lambda_m R_k^i \text{ , where } y_h \neq 0 \text{ , and}$$
$$\mathcal{B}_m R_h^i = \lambda_m R_h^i \text{ , where } y_k \neq 0 . \quad (3.10)$$

Therefore, the proof of theorem is completed, we conclude

Theorem 3.2. In the space $\mathrm{G}^{2\mathrm{nd}}\mathcal{B}\bar{W} - \mathrm{RF_n}$, the M-projective torsion tensor $\bar{W}_{kh}^i$ (M-projective curvature tensor $\bar{W}_{jkh}^i$) represents a generalized recurrent Finsler space, provided that the condition (3.10) is satisfied.

By transvecting (3.8) with y^k and applying $(n = 4)$, along with equations (2.11b), (2.5a), (2.5b), (2.1b), (2.10a), (2.1c) and (2.12e), we obtain the following result:

$$\mathcal{B}_m\bar{W}_h^i = \lambda_m\bar{W}_h^i + \mu_m\left(y^i y_h - \delta_h^i F^2\right)$$
$$-\frac{1}{4}\gamma_m\left[W_h^i F^2\right] + \frac{1}{6}\left(y_h(\mathcal{B}_m R_k^i)y^k + F^2(\mathcal{B}_m R_h^i)\right)$$
$$-\frac{1}{6}\mathcal{B}_m\left(y^i H_h + 3\delta_h^i H\right) - \frac{1}{6}\lambda_m\left(y_h R_k^i y^k + F^2 R_h^i\right)$$
$$+\frac{1}{6}\lambda_m\left(y^i H_h + 3\delta_h^i H\right). \quad (3.11)$$

This demonstrates that

$$\mathcal{B}_m\bar{W}_h^i = \lambda_m\bar{W}_h^i + \mu_m\left(y^i y_h - \delta_h^i F^2\right)$$
$$-\frac{1}{4}\gamma_m\left[W_h^i F^2\right] . \quad (3.12)$$

If and only if

$$\mathcal{B}_m\left(y^i H_h + 3\delta_h^i H\right) = \lambda_m\left(y^i H_h + 3\delta_h^i H\right),$$
$$\mathcal{B}_m R_k^i = \lambda_m R_k^i \text{ , where } y_h y^k \neq 0 \text{ , and}$$
$$\mathcal{B}_m R_h^i = \lambda_m R_h^i \text{ , where } F^2 \neq 0 . \quad (3.13)$$

Thus, the proof of theorem is completed, we conclude

Theorem 3.3. In the space $\mathrm{G}^{2\mathrm{nd}}\mathcal{B}\bar{W} - \mathrm{RF_n}$, the M-projective deviation tensor $\bar{W}_h^i$ represents a generalized recurrent Finsler space if the tensors $\left(y^i H_h + 3\delta_h^i H\right)$ *and* $\left(y_h R_k^i y^k + F^2 R_h^i\right)$ are generalized recurrent Finsler spaces.

By contracting the indices i and h in the equations (3.5), (3.8) and (3.11), and utilizing equations $(n = 4)$, (2.2a), (2.1a), (2.1b), (2.10b), (2.10c), (2.10d), (2.12c), (2.12d), (2.12e) and (2.1c), in conjunction with (2.11c), (2.11d) and (2.11e), we obtain the following result:

$$\mathcal{B}_m\bar{W}_{jk} = \lambda_m\bar{W}_{jk} + \mu_m(1-n)g_{jk} + \frac{1}{4}\gamma_m\left[W_{jk}\right]$$
$$+\frac{1}{6}\left((\mathcal{B}_m R_{jk}) + g_{jk}(\mathcal{B}_m \mathrm{R})\right) - \frac{1}{6}\mathcal{B}_m(1+n)R_{jk}$$
$$+\frac{1}{6}\lambda_m(1+n)R_{jk} - \frac{1}{6}\lambda_m\left(R_{jk} + g_{jk}\,\mathrm{R}\right)$$
$$-\frac{1}{3}\mathcal{B}_q y^q\left(R_k^i C_{jim} + R\, C_{jkm}\right). \quad (3.14)$$

This demonstrates that

$$\mathcal{B}_m\bar{W}_{jk} = \lambda_m\bar{W}_{jk} + \mu_m(1-n)g_{jk} + \frac{1}{4}\gamma_m\left[W_{jk}\right]$$
$$-\frac{1}{3}\mathcal{B}_q y^q\left(R_k^i C_{jim} + R\, C_{jkm}\right). \quad (3.15)$$

If and only if

$$\mathcal{B}_m R_{jk} = \lambda_m R_{jk} \text{ , and}$$
$$\mathcal{B}_m \mathrm{R} = \lambda_m \mathrm{R} \text{ , where } g_{jk} \neq 0 . \quad (3.16)$$

and

$$\mathcal{B}_m\bar{W}_k = \lambda_m\bar{W}_k + \mu_m(1-n)\,y_k$$
$$+\frac{1}{6}\left(y_i(\mathcal{B}_m R_k^i) + y_k(\mathcal{B}_m \mathrm{R})\right) - \frac{1}{6}\mathcal{B}_m(1+\mathrm{n})H_k$$
$$-\frac{1}{6}\lambda_m\left(y_i R_k^i + y_k \mathrm{R}\right) + \frac{1}{6}\lambda_m(1+\mathrm{n})H_k . \quad (3.17)$$

This demonstrates that

$$\mathcal{B}_m\bar{W}_k = \lambda_m\bar{W}_k + \mu_m(1-n)\,y_k . \quad (3.18)$$

If and only if

$$\mathcal{B}_m R_k^i = \lambda_m R_k^i \text{ , where } y_i \neq 0 \text{ ,}$$
$$\mathcal{B}_m R = \lambda_m R \text{ , where } y_k \neq 0 \text{ , and}$$
$$\mathcal{B}_m H_k = \lambda_m H_k . \quad (3.19)$$

In the last

$$\mathcal{B}_m\bar{W} = \lambda_m\bar{W} + \mu_m(1-n)F^2$$
$$+\frac{1}{6}\left(y_i y^k(\mathcal{B}_m R_k^i) + F^2(\mathcal{B}_m \mathrm{R})\right) - \frac{1}{2}(1+n)\mathcal{B}_m H$$
$$-\frac{1}{6}\lambda_m\left(y_i R_k^i y^k + F^2 R\right) + \frac{1}{2}\lambda_m(1+n)H. \quad (3.20)$$

This demonstrates that

$$\mathcal{B}_m\bar{W} = \lambda_m\bar{W} + \mu_m(1-n)F^2. \quad (3.21)$$

If and only if

$$\mathcal{B}_m R_k^i = \lambda_m R_k^i \text{ , where } y_i y^k \neq 0 \text{ ,}$$
$$\mathcal{B}_m R = \lambda_m R \text{ , where } F^2 \neq 0 \text{ , and}$$
$$\mathcal{B}_m H = \lambda_m H . \quad (3.22)$$

In conclusion the proof of theorem is completed, we can say

Theorem 3.4. In the space $\mathrm{G}^{2\mathrm{nd}}\mathcal{B}\bar{W} - \mathrm{RF_n}$, $\bar{W}$ − Ricci tensor $\bar{W}_{jk}$, vector $\bar{W}_k$ and scalar $\bar{W}$ are defined in equations (3.15), (3.18) and (3.21), respectively, if and only if the conditions in equations (3.16), (3.19) and (3.22) are satisfied.

We introduce a new class of Finsler spaces, namely, generalized birecurrent spaces. These spaces generalize the concept of birecurrence to a broader setting and exhibit geometric properties. We investigate the curvature tensor of these spaces and establish several characterization theorems.

Thus, this paper focuses on what is called the covariant derivative of second order.

By taking the covariant derivative of equation (3.3), with respect to in the sense of Berwald, the following results are obtained:

$$\mathcal{B}_l\mathcal{B}_m W^i_{jkh} = \mathcal{B}_l\mathcal{B}_m \bar{W}^i_{jkh} + \frac{1}{6}\mathcal{B}_l\mathcal{B}_m\left(\delta^i_k R_{jh} + \delta^i_h R_{jk}\right) - \frac{1}{6}\mathcal{B}_l\left(g_{jh}(\mathcal{B}_m R^i_k) + g_{jk}(\mathcal{B}_m R^i_h)\right) + \frac{1}{3} y^q \mathcal{B}_q\mathcal{B}_l\left(R^i_k C_{jhm} + R^i_h C_{jkm}\right). \quad (3.23)$$

Or

$$\mathcal{B}_l\mathcal{B}_m W^i_{jkh} = \mathcal{B}_l\mathcal{B}_m \bar{W}^i_{jkh} - \frac{1}{6}\left((\mathcal{B}_l g_{jh})(\mathcal{B}_m R^i_k) + g_{jh}(\mathcal{B}_l\mathcal{B}_m R^i_k) + (\mathcal{B}_l g_{jk})(\mathcal{B}_m R^i_h) + g_{jk}(\mathcal{B}_l\mathcal{B}_m R^i_h)\right) + \frac{1}{6}\mathcal{B}_l\mathcal{B}_m\left(\delta^i_k R_{jh} + \delta^i_h R_{jk}\right) + \frac{1}{3} y^q \mathcal{B}_q\mathcal{B}_l\left(R^i_k C_{jhm} + R^i_h C_{jkm}\right). \quad (3.24)$$

The equation (3.24) can be expressed as

$$\mathcal{B}_l\mathcal{B}_m W^i_{jkh} = \mathcal{B}_l\mathcal{B}_m \bar{W}^i_{jkh} - \frac{1}{6}\left(g_{jh}(\mathcal{B}_l\mathcal{B}_m R^i_k) + g_{jk}(\mathcal{B}_l\mathcal{B}_m R^i_h)\right) + \frac{1}{6}\mathcal{B}_l\mathcal{B}_m\left(\delta^i_k R_{jh} + \delta^i_h R_{jk}\right) - \frac{1}{6}\left((\mathcal{B}_l g_{jh})(\mathcal{B}_m R^i_k) + (\mathcal{B}_l g_{jk})(\mathcal{B}_m R^i_h)\right) + \frac{1}{3} y^q \mathcal{B}_q\mathcal{B}_l\left(R^i_k C_{jhm} + R^i_h C_{jkm}\right). \quad (3.25)$$

Using equation (2.6), equation (3.25) can be rewritten as follows:

$$\mathcal{B}_l\mathcal{B}_m W^i_{jkh} = \mathcal{B}_l\mathcal{B}_m \bar{W}^i_{jkh} - \frac{1}{6}\left(g_{jh}(\mathcal{B}_l\mathcal{B}_m R^i_k) + g_{jk}(\mathcal{B}_l\mathcal{B}_m R^i_h)\right) + \frac{1}{6}\mathcal{B}_l\mathcal{B}_m\left(\delta^i_k R_{jh} + \delta^i_h R_{jk}\right) + \frac{1}{3} y^q \mathcal{B}_q\mathcal{B}_l\left(R^i_k C_{jhm} + R^i_h C_{jkm}\right) + \frac{1}{3}\mathcal{B}_q y^q\left(C_{jhl}\left(\mathcal{B}_m R^i_k\right) + C_{jkl}\left(\mathcal{B}_m R^i_h\right)\right). \quad (3.26)$$

Similarly, by applying equations (2.16) and (3.1) in (3.26), we obtain the result:

$$\mathcal{B}_l\mathcal{B}_m \bar{W}^i_{jkh} + \frac{1}{6}\mathcal{B}_l\mathcal{B}_m\left(\delta^i_k R_{jh} + \delta^i_h R_{jk}\right) - \frac{1}{6}\left(g_{jh}(\mathcal{B}_l\mathcal{B}_m R^i_k) + g_{jk}(\mathcal{B}_l\mathcal{B}_m R^i_h)\right) + \frac{1}{3}\mathcal{B}_q y^q\left(C_{jhl}\left(\mathcal{B}_m R^i_k\right) + C_{jkl}\left(\mathcal{B}_m R^i_h\right)\right) + \frac{1}{3} y^q \mathcal{B}_q\mathcal{B}_l\left(R^i_k C_{jhm} + R^i_h C_{jkm}\right) = a_{ml}\bar{W}^i_{jkh} + \frac{1}{6}a_{ml}\left(\delta^i_k R_{jh} + \delta^i_h R_{jk}\right) + b_{ml}\left(\delta^i_k g_{jh} - \delta^i_h g_{jk}\right) - \frac{1}{6}a_{ml}\left(g_{jh}R^i_k + g_{jk}R^i_h\right) - \frac{1}{2}\gamma_m\mathcal{B}_q y^q\left(W^i_k C_{jhl} - W^i_h C_{jkl}\right) + \frac{1}{4}c_{ml}\left(W^i_k g_{jh} - W^i_h g_{jk}\right) + \frac{1}{4}\gamma_m\left(g_{jh}\left(\mathcal{B}_l W^i_k\right) - g_{jk}\left(\mathcal{B}_l W^i_h\right)\right) - 2\mu_m\mathcal{B}_q y^q\left(\delta^i_k C_{jhl} - \delta^i_h C_{jkl}\right). \quad (3.27)$$

Alternatively, the equation (3.27) can be expressed as:

$$\mathcal{B}_l\mathcal{B}_m \bar{W}^i_{jkh} = a_{ml}\bar{W}^i_{jkh} + b_{ml}\left(\delta^i_k g_{jh} - \delta^i_h g_{jk}\right) + \frac{1}{4}c_{ml}\left(W^i_k g_{jh} - W^i_h g_{jk}\right) + \frac{1}{4}\gamma_m\left(g_{jh}\left(\mathcal{B}_l W^i_k\right) - g_{jk}\left(\mathcal{B}_l W^i_h\right)\right) - 2\mu_m\mathcal{B}_q y^q\left(\delta^i_k C_{jhl} - \delta^i_h C_{jkl}\right) - \frac{1}{3} y^q\mathcal{B}_q\mathcal{B}_l\left(R^i_k C_{jhm} + R^i_h C_{jkm}\right) + \frac{1}{6}a_{ml}\left(\delta^i_k R_{jh} + \delta^i_h R_{jk}\right) - \frac{1}{6}a_{ml}\left(g_{jh}R^i_k + g_{jk}R^i_h\right) - \frac{1}{6}\mathcal{B}_l\mathcal{B}_m\left(\delta^i_k R_{jh} + \delta^i_h R_{jk}\right) + \frac{1}{6}\left(g_{jh}(\mathcal{B}_l\mathcal{B}_m R^i_k) + g_{jk}(\mathcal{B}_l\mathcal{B}_m R^i_h)\right) - \frac{1}{3}\mathcal{B}_q y^q\left(C_{jhl}\left(\mathcal{B}_m R^i_k\right) + C_{jkl}\left(\mathcal{B}_m R^i_h\right)\right) - \frac{1}{2}\gamma_m\mathcal{B}_q y^q\left(W^i_k C_{jhl} - W^i_h C_{jkl}\right). \quad (3.28)$$

This demonstrates that

$$\mathcal{B}_l\mathcal{B}_m \bar{W}^i_{jkh} = a_{ml}\bar{W}^i_{jkh} + b_{ml}\left(\delta^i_k g_{jh} - \delta^i_h g_{jk}\right) + \frac{1}{4}c_{ml}\left(W^i_k g_{jh} - W^i_h g_{jk}\right) + \frac{1}{4}\gamma_m\left(g_{jh}\left(\mathcal{B}_l W^i_k\right) - g_{jk}\left(\mathcal{B}_l W^i_h\right)\right) - 2\mu_m\mathcal{B}_q y^q\left(\delta^i_k C_{jhl} - \delta^i_h C_{jkl}\right) - 2\mu_m y^q\mathcal{B}_q\left(\delta^i_k C_{jhl} - \delta^i_h C_{jkl}\right) - \frac{1}{2}\gamma_m\mathcal{B}_q y^q\left(W^i_k C_{jhl} - W^i_h C_{jkl}\right) - \frac{1}{3} y^q\mathcal{B}_q\mathcal{B}_l\left(R^i_k C_{jhm} + R^i_h C_{jkm}\right) - \frac{1}{3}\mathcal{B}_q y^q\left(C_{jhl}\left(\mathcal{B}_m R^i_k\right) + C_{jkl}\left(\mathcal{B}_m R^i_h\right)\right). \quad (3.29)$$

If and only if

$$\mathcal{B}_l\mathcal{B}_m\left(\delta^i_k R_{jh} + \delta^i_h R_{jk}\right) = a_{ml}\left(\delta^i_k R_{jh} + \delta^i_h R_{jk}\right),$$
$$\mathcal{B}_l\mathcal{B}_m R^i_k = a_{ml}R^i_k \text{ , where } g_{jh} \neq 0 \text{ , and}$$
$$\mathcal{B}_l\mathcal{B}_m R^i_h = a_{ml}R^i_h \text{ , where } g_{jk} \neq 0 . \quad (3.30)$$

In conclusion the proof of theorem is completed, we can determine

Theorem 3.5. In the space $G^{2nd}\mathcal{B}\bar{W} - BRF_n$, the M-projective curvature tensor $\bar{W}^i_{jkh}$ defines a generalized birecurrent Finsler space if and only if the condition in equation (3.30) is satisfied.

By transvecting equation (3.28) with y^j and utilizing equations (2.11a), (2.5b), (2.1a), (2.3) and (2.12a), we obtain the following result:

$$\mathcal{B}_l\mathcal{B}_m \bar{W}^i_{kh} = a_{ml}\bar{W}^i_{kh} + b_{ml}\left(\delta^i_k y_h - \delta^i_h y_k\right) + \frac{1}{4}c_{ml}\left(W^i_k y_h - W^i_h y_k\right) - \frac{1}{6}\mathcal{B}_l\mathcal{B}_m\left(\delta^i_k H_h + \delta^i_h H_k\right) + \frac{1}{6}\left(y_h(\mathcal{B}_l\mathcal{B}_m R^i_k) + y_k(\mathcal{B}_l\mathcal{B}_m R^i_h)\right) + \frac{1}{6}a_{ml}\left(\delta^i_k H_h + \delta^i_h H_k\right) - \frac{1}{6}a_{ml}\left(y_h R^i_k + y_k R^i_h\right) + \frac{1}{4}\gamma_m\left(y_h(\mathcal{B}_l W^i_k) - y_k\left(\mathcal{B}_l W^i_h\right)\right). \quad (3.31)$$

This demonstrates that

$$\mathcal{B}_l\mathcal{B}_m \bar{W}^i_{kh} = a_{ml}\bar{W}^i_{kh} + b_{ml}\left(\delta^i_k y_h - \delta^i_h y_k\right) + \frac{1}{4}c_{ml}\left(W^i_k y_h - W^i_h y_k\right) + \frac{1}{4}\gamma_m\left(y_h(\mathcal{B}_l W^i_k) - y_k\left(\mathcal{B}_l W^i_h\right)\right). \quad (3.32)$$

If and only if

$$\mathcal{B}_l\mathcal{B}_m\left(\delta^i_k H_h + \delta^i_h H_k\right) = a_{ml}\left(\delta^i_k H_h + \delta^i_h H_k\right),$$

$$\mathcal{B}_l\mathcal{B}_m R_k^i = a_{ml}R_k^i \text{ , where } y_h \neq 0 \text{ , and}$$
$$\mathcal{B}_l\mathcal{B}_m R_h^i = a_{ml}R_h^i \text{ , where } y_k \neq 0 \text{ .} \quad (3.33)$$

Therefore, the proof of theorem is completed, we conclude Theorem 3.6. In the space $G^{2nd}\mathcal{B}\bar{W} - BRF_n$, the covariant derivative of the second orders for the M-projective torsion tensor $\bar{W}_{kh}^i$ (M-projective curvature tensor $\bar{W}_{jkh}^i$) defines a generalized birecurrent Finsler space if and only if the condition in equation (3.33) is satisfied.

By transvecting condition (3.31) with y^k and applying $(n = 4)$, along with equations (2.11b), (2.5a), (2.5b), (2.1b), (2.10a), (2.1c) and (2.12e), we obtain the following result:

$$\mathcal{B}_l\mathcal{B}_m\bar{W}_h^i = a_{ml}\bar{W}_h^i + b_{ml}\left(y^i y_h - \delta_h^i F^2\right)$$
$$-\frac{1}{4}c_{ml}\left(W_h^i F^2\right) - \frac{1}{4}\gamma_m F^2(\mathcal{B}_l W_h^i)$$
$$-\frac{1}{6}\left(y_h y^k(\mathcal{B}_l\mathcal{B}_m R_k^i) + F^2(\mathcal{B}_l\mathcal{B}_m R_h^i)\right)$$
$$+\frac{1}{6}\mathcal{B}_l\mathcal{B}_m\left(y^i H_h + 3\delta_h^i H\right)$$
$$+\frac{1}{6}a_{ml}\left(y_h R_k^i y^k + F^2 R_h^i\right)$$
$$-\frac{1}{6}a_{ml}\left(y^i H_h + 3\delta_h^i H\right). \quad (3.34)$$

This demonstrates that

$$\mathcal{B}_l\mathcal{B}_m\bar{W}_h^i = a_{ml}\bar{W}_h^i + b_{ml}\left(y^i\, y_h - \delta_h^i F^2\right)$$
$$-\frac{1}{4}c_{ml}\left(F^2 W_h^i\right) - \frac{1}{4}\gamma_m F^2(\mathcal{B}_l W_h^i)\,. \quad (3.35)$$

If and only if

$$\mathcal{B}_l\mathcal{B}_m\left(y^i H_h + 3\delta_h^i H\right) = a_{ml}\left(y^i H_h + 3\delta_h^i H\right),$$
$$\mathcal{B}_l\mathcal{B}_m R_k^i = a_{ml}R_k^i \text{ , where } y_h y^k \neq 0 \text{ , and}$$
$$\mathcal{B}_l\mathcal{B}_m R_h^i = a_{ml}R_h^i \text{ , where } F^2 \neq 0 \text{ .} \quad (3.36)$$

Therefore, the proof of theorem is completed, we can say Theorem 3.7. In the space $G^{2nd}\mathcal{B}\bar{W} - BRF_n$, the M-projective deviation tensor $\bar{W}_h^i$ represents a generalized birecurrent Finsler space if the tensors $\left(y^i H_h + 3\delta_h^i H\right)$ and $\left(y_h R_k^i y^k + F^2 R_h^i\right)$ are generalized birecurrent Finsler space.

By contracting the indices i and h in the equations (3.28), (3.31) and (3.34), and utilizing equations $(n = 4)$, (2.2a), (2.1a), (2.1b), (2.10b), (2.10c), (2.10d), (2.12c), (2.12d), (2.12e) and (2.1c), along with the relations in equations (2.11c), (2.11d) and (2.11e), we obtain the following result:

$$\mathcal{B}_l\mathcal{B}_m\bar{W}_{jk} = a_{ml}\bar{W}_{jk} + b_{ml}(1-n)g_{jk} + \frac{1}{4}c_{ml}W_{jk}$$
$$-\frac{1}{6}(1+n)\mathcal{B}_l\mathcal{B}_m R_{jk} + \frac{1}{6}\left((\mathcal{B}_l\mathcal{B}_m R_{jk}) + g_{jk}(\mathcal{B}_l\mathcal{B}_m R)\right)$$
$$-\frac{1}{3}\mathcal{B}_q y^q\left(C_{jil}(\mathcal{B}_m R_k^i) + C_{jkl}(\mathcal{B}_m R)\right)$$
$$+\frac{1}{6}a_{ml}(1+n)R_{jk} - \frac{1}{3}y^q\mathcal{B}_q\mathcal{B}_l\left(R_k^i C_{jim} + R C_{jkm}\right)$$
$$+\frac{1}{4}\gamma_m\mathcal{B}_l W_{jk} - \frac{1}{2}\gamma_m\mathcal{B}_q y^q W_k^i C_{jil}$$
$$-2\mu_m y^q\mathcal{B}_q(1-n)C_{jkl} - \frac{1}{6}a_{ml}\left(R_{jk} + g_{jk}R\right). \quad (3.37)$$

This demonstrates that

$$\mathcal{B}_l\mathcal{B}_m\bar{W}_{jk} = a_{ml}\bar{W}_{jk} + b_{ml}(1-n)g_{jk} + \frac{1}{4}c_{ml}W_{jk}$$
$$+\frac{1}{4}\gamma_m\mathcal{B}_l W_{jk} - \frac{1}{2}\gamma_m\mathcal{B}_q y^q W_k^i\, C_{jil}$$
$$-\frac{1}{3}y^q\mathcal{B}_q\mathcal{B}_l\left(R_k^i C_{jim} + R\, C_{jkm}\right)$$
$$-\frac{1}{3}\mathcal{B}_q y^q\left(C_{jil}(\mathcal{B}_m R_k^i) + C_{jkl}(\mathcal{B}_m R)\right)$$
$$-2\mu_m y^q\mathcal{B}_q(1-n)C_{jkl}\,. \quad (3.38)$$

If and only if

$$\mathcal{B}_l\mathcal{B}_m R_{jk} = a_{ml}R_{jk}\,, \text{ and}$$
$$\mathcal{B}_l\mathcal{B}_m R = a_{ml}R\,, \text{ where } g_{jk} \neq 0\,. \quad (3.39)$$

And $\mathcal{B}_l\mathcal{B}_m\bar{W}_k = a_{ml}\bar{W}_k + b_{ml}(1-n)y_k$

$$+\frac{1}{6}\left(y_i(\mathcal{B}_l\mathcal{B}_m R_k^i) + y_k(\mathcal{B}_l\mathcal{B}_m \mathrm{R})\right) - \frac{1}{6}(1+\mathrm{n})\mathcal{B}_l\mathcal{B}_m H_k$$
$$-\frac{1}{6}a_{ml}\left(y_i R_k^i + y_k \mathrm{R}\right) + \frac{1}{6}a_{ml}(1+\mathrm{n})H_k\,. \quad (3.40)$$

This demonstrates that

$$\mathcal{B}_l\mathcal{B}_m\bar{W}_k = a_{ml}\bar{W}_k + b_{ml}(1-n)\,y_k. \quad (3.41)$$

If and only if

$$\mathcal{B}_l\mathcal{B}_m\, H_k = a_{ml}H_k \text{ ,}$$
$$\mathcal{B}_l\mathcal{B}_m R_k^i = a_{ml}R_k^i \text{ , where } y_i \neq 0\,, \text{ and}$$
$$\mathcal{B}_l\mathcal{B}_m \mathrm{R} = a_{ml}\mathrm{R} \text{ , where } y_k \neq 0\,. \quad (3.42)$$

In the last

$$\mathcal{B}_l\mathcal{B}_m\bar{W} = a_{ml}\bar{W} + b_{ml}(1-n)F^2$$
$$-\frac{1}{6}\left(y_i y^k(\mathcal{B}_l\mathcal{B}_m R_k^i) + F^2(\mathcal{B}_l\mathcal{B}_m \mathrm{R})\right) + \frac{1}{2}(n+1)\mathcal{B}_l\mathcal{B}_m H$$
$$+\frac{1}{6}a_{ml}\left(y_i R_k^i y^k + F^2 \mathrm{R}\right) - \frac{1}{2}a_{ml}(n+1)H. \quad (3.43)$$

This demonstrates that

$$\mathcal{B}_l\mathcal{B}_m\bar{W} = a_{ml}\bar{W} + b_{ml}(n-1)F^2\,. \quad (3.44)$$

If and only if

$$\mathcal{B}_l\mathcal{B}_m H = a_{ml}H\,,$$
$$\mathcal{B}_l\mathcal{B}_m R_k^i = a_{ml}R_k^i \text{ , where } y_i y^k \neq 0 \text{ , and}$$
$$\mathcal{B}_l\mathcal{B}_m R = a_{ml}R \text{ , where } F^2 \neq 0\,. \quad (3.45)$$

In conclusion the proof of theorem is completed, we can say Theorem 3.8. In the space $G^{2nd}\mathcal{B}\bar{W} - BRF_n$, the M-projective Ricci tensor $\bar{W}_{jk}$, vector $\bar{W}_k$ and scalar $\bar{W}$ are defined in equations (3.38), (3.41) and (3.44), respectively, provided that the conditions in equations (3.39), (3.42) and (3.45) are satisfied.

By transvecting equation (3.28) with g_{ir} and utilizing equations (2.2b), (2.10d), (2.12d) and (2.11f), we obtain the following result:

$$\mathcal{B}_l\mathcal{B}_m\bar{W}_{rjkh} = a_{ml}\bar{W}_{rjkh} + b_{ml}\left(g_{rk}\, g_{jh} - g_{rh}g_{jk}\right)$$
$$-\frac{1}{6}\mathcal{B}_l\mathcal{B}_m\left(g_{rk}R_{jh} + g_{rh}\, R_{jk}\right)$$
$$+\frac{1}{6}\left(g_{jh}(\mathcal{B}_l\mathcal{B}_m R_{rk}) + g_{jk}(\mathcal{B}_l\mathcal{B}_m R_{rh})\right)$$
$$-\frac{1}{3}y^q\mathcal{B}_q\left(C_{jhl}(\mathcal{B}_m R_{rk}) + C_{jkl}(\mathcal{B}_m R_{rh})\right)$$
$$-\frac{1}{3}y^q\mathcal{B}_q\mathcal{B}_l\left(R_{rk}C_{jhm} + R_{rh}C_{jkm}\right)$$
$$+\frac{1}{6}a_{ml}\left(g_{rk}R_{jh} + g_{rh}\, R_{jk}\right) - \frac{1}{6}a_{ml}\left(g_{jh}R_{rk} + g_{jk}\, R_{rh}\right)$$
$$+\frac{1}{4}\gamma_m\left((\mathcal{B}_l W_{rk})g_{jh} - (\mathcal{B}_l W_{rh})g_{jk}\right)$$
$$+\frac{1}{4}c_{ml}\left(W_{rk}g_{jh} - W_{rh}g_{jk}\right)$$
$$-\frac{1}{2}\gamma_m\mathcal{B}_q y^q\left(W_{rk}C_{jhl} - W_{rh}C_{jkl}\right)$$
$$-2\mu_m y^q\mathcal{B}_q\left(g_{rk}C_{jhl} - g_{rh}C_{jkl}\right). \quad (3.46)$$

This demonstrates that

$$\mathcal{B}_l\mathcal{B}_m\bar{W}_{rjkh} = a_{ml}\bar{W}_{rjkh} + b_{ml}\left(g_{rk}\,g_{jh} - g_{rh}g_{jk}\right)$$
$$+\tfrac{1}{4}\gamma_m\left((\mathcal{B}_l W_{rk})g_{jh} - (\mathcal{B}_l W_{rh})g_{jk}\right)$$
$$+\tfrac{1}{4}c_{ml}\left(W_{rk}g_{jh} - W_{rh}g_{jk}\right)$$
$$-\tfrac{1}{2}\gamma_m\mathcal{B}_q y^q\left(W_{rk}C_{jhl} - W_{rh}C_{jkl}\right)$$
$$-2\mu_m y^q\mathcal{B}_q\left(g_{rk}C_{jhl} - g_{rh}C_{jkl}\right)$$
$$-\tfrac{1}{3}y^q\mathcal{B}_q\left(C_{jhl}(\mathcal{B}_m R_{rk}) + C_{jkl}(\mathcal{B}_m R_{rh})\right)$$
$$-\tfrac{1}{3}y^q\mathcal{B}_q\mathcal{B}_l\left(R_{rk}C_{jhm} + R_{rh}C_{jkm}\right). \qquad (3.47)$$

If and only if

$$\mathcal{B}_l\mathcal{B}_m\left(g_{rk}R_{jh} + g_{rh}R_{jk}\right) = a_{ml}\left(g_{rk}R_{jh} + g_{rh}R_{jk}\right),$$
$$\mathcal{B}_l\mathcal{B}_m R_{rk} = a_{ml}R_{rk} \text{ , where } g_{jh} \neq 0 \text{ , and}$$
$$\mathcal{B}_l\mathcal{B}_m R_{rh} = a_{ml}R_{rh}, \text{ where } g_{jk} \neq 0\ . \qquad (3.48)$$

Thus, the proof of theorem is completed, we conclude

Theorem 3.9. In the space $G^{2nd}\mathcal{B}\bar{W} - BRF_n$, M-projective associate tensor $\bar{W}_{rjkh}$ (M-projective curvature tensor $\bar{W}^i_{jkh}$) characterizes a generalized birecurrent Finsler space, provided that the condition (3.48) is satisfied.

CONCLUSION

This paper introduces and explores several new properties of the M-projective curvature tensor, $\bar{W}^i_{jkh}$, in the context of Finsler geometry, specifically within the framework of generalized recurrent Finsler spaces. By deriving a series of equations and theorems, we have demonstrated the underlying relationships between various curvature tensors and their conditions for being generalized recurrent Finsler spaces.

The key findings of the study are as follows:

1. Generalized Recurrent Finsler Spaces: We have established that in the space $G^{2nd}\mathcal{B}\bar{W} - BRF_n$, the M-projective curvature tensor $\bar{W}^i_{jkh}$ represents a generalized recurrent Finsler space, provided that certain conditions, such as equation (3.7), are satisfied. This is crucial as it extends the concept of recurrent Finsler spaces and provides a deeper understanding of their geometric properties.
2. M-projective Torsion Tensor: Through detailed derivations and the application of transvection techniques, we showed that the M-projective torsion tensor $\bar{W}^i_{jkh}$ in a similar manner also represents a generalized recurrent Finsler space, subject to specific conditions (3.10). This adds a new layer to the study of torsion tensors in Finsler geometry, highlighting their potential to model more complex geometric structures.
3. Deviations in Geometrical Properties: The study also presents a comprehensive approach to the M-projective deviation tensor $\bar{W}^i_h$. By incorporating deviations into the curvature analysis, we have established that the deviation tensor represents a generalized recurrent Finsler space under conditions (3.13), thus providing a more generalized framework for the classification of Finsler spaces.
4. Curvature Tensors and Their Interactions: By examining various covariant derivatives and transvecting equations, we have also shown how the covariant derivatives $\mathcal{B}_m\bar{W}^i_{jkh}$ and $\mathcal{B}_l\mathcal{B}_m\bar{W}^i_{jkh}$ contribute to understanding the interplay between different curvature tensors in the context of generalized recurrent Finsler spaces. The results underscore the importance of these interactions in establishing a more robust classification of Finsler spaces and contribute to the understanding of their broader geometric properties.
5. Implications for Birecurrent Finsler Spaces: The concept of generalized birecurrent Finsler spaces is introduced, providing an extension to the classical idea of birecurrence. This new class of spaces offers an enriched set of geometrical structures with applications in higher-order differential geometry and theoretical physics, particularly in the study of spaces with more intricate curvatures and torsions.
6. New Theorems and Classifications: The paper provides the proof and theorems associated with the space $G^{2nd}\mathcal{B}\bar{W} - BRF_n$ and the conditions under which the M-projective curvature tensor and its derivatives define generalized recurrent or birecurrent Finsler spaces. These results extend the existing framework of Finsler geometry by providing more generalized conditions for classification.

In conclusion, the results of this study open several new avenues for future research, particularly in understanding more complex and higher-dimensional Finsler spaces and their applications. The proposed definitions and conditions lay the groundwork for further investigation into generalized recurrent Finsler spaces, M-projective tensors, and their relations to curvature, torsion, and deviation. Moreover, the introduction of generalized birecurrent spaces offers a promising direction for the exploration of spaces with highly intricate geometric structures.

RECOMMENDATIONS

Based on the comprehensive derivations and proofs established within this study, the following recommendations are proposed for future research directions in the study of M-projective curvature tensors and generalized Finsler spaces:

1. Further Analysis of Generalized Finsler Spaces:

The results obtained for generalized recurrent Finsler spaces, specifically in relation to the M-projective curvature tensor $\bar{W}^i_{jkh}$, offer potential for further investigation. Future studies could explore additional conditions under which these spaces exhibit notable geometric properties such as curvature, torsion, and deviation. It would be beneficial to extend the characterization of these spaces to higher-dimensional settings (i.e., beyond $(n = 4)$ and analyze their applications in complex geometries.

2. Refinement of Covariant Derivatives in Second-Order Calculations:

In this work, we have introduced covariant derivatives such as $\mathcal{B}_l\mathcal{B}_m\bar{W}^i_{jkh}$ and related expressions. These derivatives offer a systematic approach to studying the geometry of M-projective tensors. However, more advanced techniques could be developed to refine the computational process, particularly for higher-order covariant derivatives. Exploring the use of alternative differential operators might yield more efficient methods for deriving related geometric invariants.

3. Potential Applications to Generalized Birecurrent Spaces:

The concept of generalized birecurrent spaces is introduced as a potential extension of classical birecurrent

Finsler spaces. This new class of spaces requires deeper exploration of its geometric and physical interpretations, especially in contexts where Finsler spaces with curvature constraints are applicable. We suggest investigating the applicability of these spaces in areas such as theoretical physics, where the structure of spacetime may be modeled using advanced differential geometry.

4. Numerical Simulations and Computational Methods:

While the theoretical framework presented in this paper provides a solid foundation, numerical simulations could offer insights into the practical implications of the results. Computational methods for solving systems of differential equations governing the M-projective curvature tensors may lead to new insights into the behavior of generalized Finsler spaces under various boundary conditions. This would further enhance the understanding of the relationship between curvature tensors and geometric properties in high-dimensional spaces.

5. Integration of Torsion and Curvature in Physical Models:

Torsion and curvature play a significant role in the study of differential geometry, particularly in the context of general relativity and other physical theories. Future research could focus on how these tensors interact in the context of Finsler spaces with torsion, especially when considering specific models of spacetime or material media. The results of this study, especially the properties of $\overline{W}^i_{jkh}$ and its transvecting forms, could serve as the foundation for developing new models of physical phenomena.

6. Extension to Non-Riemannian Geometries:

The framework of M-projective curvature tensors could also be extended to non-Riemannian geometries, such as those encountered in quantum gravity and string theory. By relaxing some of the assumptions inherent to classical Riemannian geometry, future studies could explore how curvature tensors behave in more generalized spaces and whether this yields new insights into the fundamental nature of space and time.

These recommendations aim to guide future investigations into the properties and applications of generalized Finsler spaces and M-projective curvature tensors, contributing to both theoretical advancements and practical implementations in geometric and physical models.

REFERENCES

[1] H. Abu-Donia, S. Shenawy, and A. Abdehameed, "The w*-curvature tensor on relativistic space-times," *Kyungpook Math. J.*, vol. 60, pp. 185–195, 2020.

[2] Z. Ahsan and M. Ali, "On some properties of w-curvature tensor," *Palestine J. Math.*, vol. 3, no. 1, pp. 61–69, 2014.

[3] Z. Ahsan and M. Ali, "Curvature tensor for the spacetime of general relativity," *Palestine J. Math.*, vol. 2, pp. 1–15, 2016.

[4] A. M. Al-Qashbari, A. A. Abdallah, and S. M. Baleedi, "A study of M-projective curvature tensor W̄_jkh^i in GBK-5RF_n via Lie derivative," *Int. J. Appl. Sci. Math. Theory*, vol. 11, no. 1, pp. 1–9, 2025.

[5] A. M. Al-Qashbari, A. A. Abdallah, and F. Abdallah, "Decomposition of generalized recurrent tensor fields of R^h-nth order in Finsler manifolds," *J. Sci. Technol.*, vol. 30, no. 2, pp. 99–105, 2025.

[6] A. M. A. AL-Qashbari, A. A. Abdallah, and K. S. Nasr, "Generalized of h-torsion and curvature structures in generalized recurrent space of third order by Cartan covariant derivatives," *GPH-Int. J. Math.*, vol. 8, no. 2, pp. 1–9, 2025.

[7] A. M. Al-Qashbari, A. A. Abdallah, and F. A. Al-ssallal, "Recurrent Finsler structures with higher-order generalizations defined by special curvature tensors," *Int. J. Adv. Res. Sci. Commun. Technol.*, vol. 4, no. 1, pp. 68–75, 2024.

[8] A. M. Al-Qashbari and F. A. Al-ssallal, "Study of curvature tensor by using Berwald's and Cartan's higher-order derivatives in Finsler space," *Technol. Appl. Humanit. Acad. J.*, vol. 1, no. 1, pp. 1–15, 2024.

[9] A. M. Al-Qashbari, M. Haoues, and F. A. Al-ssallal, "A decomposition analysis of Weyl's curvature tensor via Berwald's first and second order derivatives in Finsler spaces," *J. Innov. Appl. Math. Comput. Sci.*, vol. 4, no. 2, pp. 201–213, 2024.

[10] A. M. Al-Qashbari, A. A. Abdallah, and S. M. Baleedi, "Berwald covariant derivative and Lie derivative of conharmonic curvature tensors in generalized fifth recurrent Finsler space," *GPH-Int. J. Math.*, vol. 8, no. 1, pp. 24–32, 2025.

[11] A. M. A. AL-Qashbari and F. Y. A. Qasem, "Study on generalized BR-trirecurrent Finsler space," *J. Yemen Eng.*, vol. 15, pp. 79–89, 2017.

[12] A. M. A. AL-Qashbari, "On generalized for curvature tensors P_jkh^i of second order in Finsler space," *Univ. Aden J. Nat. Appl. Sci.*, vol. 24, no. 1, pp. 171–176, 2020.

[13] A. M. A. AL-Qashbari, "Some properties for Weyl's projective curvature tensors of generalized Wh-birecurrent in Finsler spaces," *Univ. Aden J. Nat. Appl. Sci.*, vol. 23, no. 1, pp. 181–189, 2019.

[14] A. M. A. AL-Qashbari, "Some identities for generalized curvature tensors in B-Recurrent Finsler space," *J. New Theory*, no. 32, pp. 30–39, 2020.

[15] A. M. A. AL-Qashbari, "Recurrence decompositions in Finsler space," *J. Math. Anal. Model.*, vol. 1, pp. 77–86, 2020.

[16] A. M. A. AL-Qashbari and A. A. S. AL-Maisary, "Study on generalized W_jkh^i of fourth order recurrent in Finsler space," *J. Yemeni Eng., Univ. Aden*, vol. 17, no. 2, pp. 72–86, 2023.

[17] A. M. A. AL-Qashbari, A. A. Abdallah, and K. S. Nasr, "On generalized trirecurrent space by using Gh-covariant derivative in Finsler geometry," *J. Math. Probl. Equ. Stat.*, vol. 6, no. 1, pp. 91–100, 2025.

[18] A. M. A. AL-Qashbari, S. Saleh, and I. Ibedou, "On some relations of R-projective curvature tensor in recurrent Finsler space," *J. Non-Linear Model. Anal.* (China), vol. 6, no. 4, pp. 1216–1227, 2024.

[19] S. M. S. Baleedi, "On certain generalized BK-recurrent Finsler space," M.S. thesis, Univ. Aden, Aden, Yemen, 2017.

[20] B. Misra, S. B. Misra, K. Srivastava, and R. B. Srivastava, "Higher order recurrent Finsler spaces with

Berwald's curvature tensor field," *J. Chem. Biol. Phys. Sci.*, vol. 4, no. 1, pp. 624–631, 2014.
[21] A. Goswami, "A study of certain types of special Finsler spaces in differential geometry: Systematic review," *J. Math. Appl. Sci. Technol.*, vol. 9, no. 1, pp. 23–30, 2017.
[22] P. N. Pandey, S. Saxena, and A. Goswani, "On a generalized H-recurrent space," *J. Int. Acad. Phys. Sci.*, vol. 15, pp. 201–211, 2011.
[23] H. Rund, *The Differential Geometry of Finsler Spaces*, 2nd ed. Berlin: Springer-Verlag, 1981. [In Russian], Nauka, Moscow.

Development of a Web-Based System for Predicting Depression and Suicide Attempt Using Ensemble Machine Learning Model

Akinyele O.T (1*)
Kayode A.A (1)
Adegoke-Elijah A (1)
Olowookere T.A(1)

Received: 13/03/2025
Revised: 17/05/2025
Accepted: 18/05/2025

© 2025 University of Science and Technology, Aden, Yemen. This article can be distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

© 2025 جامعة العلوم والتكنولوجيا، المركز الرئيس عدن، اليمن. يمكن إعادة استخدام المادة المنشورة حسب رخصة مؤسسة المشاع الإبداعي شريطة الاستشهاد بالمؤلف والمجلة.

[1]Department of Computer Science, Redeemer's University, Nigeria ,Nigeria
*Corresponding Author's Email: {akinyele11482, kayodeade, adegoke-elijaha, olowookereta}@run.edu.ng

Development of a Web-Based System for Predicting Depression and Suicide Attempt Using Ensemble Machine Learning Model

Akinyele O.T
Department of Computer Science,
Redeemer's University, Nigeria
Nigeria
akinyele11482@run.edu.ng

Adegoke-Elijah A.
Department of Computer Science,
Redeemer's University, Nigeria
Nigeria
adegoke-elijaha@run.edu.ng

Kayode A.A,
Department of Computer Science,
Redeemer's University, Nigeria
Nigeria
kayodeade@run.edu.ng

Olowookere T.A
Department of Computer Science,
Redeemer's University, Nigeria
Nigeria
olowookereta @run.edu.ng

Abstract— Depression is a mental illness that can make a person's life difficult and can eventually lead to suicide. Depressed individuals who do not receive timely attention develop worse conditions and may eventually commit suicide. Depression and suicide are becoming a global health concern which need to be adequately addressed. In this study, an ensemble learning model which make use of demographic data to detect depression and suicide attempt and also guide individuals from committing suicide through the web-based application system is proposed. The forever Alone demographic dataset which was downloaded from Kaggle online data repository was used, the dataset was imbalanced and was balanced using synthetic minority oversampling technique (SMOTE). The dataset was split into 60/40, 70/30 and 80/20 train/test percentage split, however, the 80/20 train/test split performed best and it was used and reported in this study. The study employs an ensemble machine learning model, specifically Adaboost with Extra trees as base estimators for prediction. Adaboost enhances model performance especially in handling class imbalance leading to excellent accuracy. Results obtained reveal that Adaboost ensemble model outperformed all other machine learning algorithms across all evaluation metrics with 82.00% recall and 78.69% accuracy for depression, and 93.85% recall and 90.60% accuracy for suicide attempt respectively on the balanced dataset. The uniqueness of Adaboost in sequential weighting of misclassified instances which enhances model performance, especially in handling class imbalance thus leading to an excellent accuracy. It was therefore used for the prediction system. The study affirmed the prowess of ensemble machine learning model for predicting depression and suicide attempt. Ethical issues were also discussed in the study.

Keywords— Ensemble Learning, Depression, Data split, Machine learning, Suicide attempt, Adaboost.

I. INTRODUCTION

It has been estimated that more than 300 million people worldwide suffer from depression, a mental condition that can make a person's life difficult and eventually lead to suicide [1]. Even though they display the symptoms through observable behavior, sad people typically are unaware of their condition. For example, depressed people often keep quiet, feel tired, and are socially isolated in real life, but on social media, they tend to be more active in expressing their emotions and ideas. One of the risk factors of suicide is depression [2]. Although neurological therapy or psychological counseling might reduce depression, early detection is important. Both depression and suicide are a global health concern. Depression sufferers who do not receive timely attention will develop worse conditions and may eventually commit suicide.

Suicide is a serious problem that affect millions of people worldwide and one of its main causes is depression [3]. It should be noted that early prediction is required for depression and suicide therapies to be successful. In the health sector, several ailments have been detected and diagnosed using machine learning (ML) techniques. Many researchers have worked on depression status and suicide attempt based on social media post but limited work has been done in using ensemble learning technique to detect depression and suicide attempt using demographic data with some risk factors like income, employment status, age, race, body weight, virginity and so on. The study therefore explores demographic predictors using ensemble method.

Angskun et al. [4], proposed a depression detection model on social networks using big data analytics. ML techniques which include Decision Tree, SVM, Random Forest, Naïve Bayes and Deep Learning were used for the depression detection model while the novel model was used to capture depressive moods of depression sufferers. From the result, it was discovered that the proposed model with feature selection using Random Forest outperformed all other existing models without feature selection.

Aldhyani et al. [5], worked on experimental research for building a suicidal ideation detection system using publicly available Reddit datasets, word-embedding approaches, for

text representation and hybrid deep learning and ML for classification with the accuracy of 95%.

Mahmud et al. [6], intended to identify the best ML model to forecast suicide risk among university students in Bangladesh by comparing using six popular ML models to identify the most efficient predictive model for suicidal behavior. SVM performance happened to be the best with 79% accuracy.

Dhasmana et al. [7], worked on ML model that can predict depression and suicide thought through data by the social media platform like Reddit. Post was extracted from social media platform called subreddits. Raw data about the two subreddits was taken and Natural Language Processing techniques was done to clean data suitable for the research purpose, then the model was trained and tested with the accuracy of 0.985.

In Saha et al. [8], suicidal ideation from social media using ML approach was investigated. The Authors made use of logistic regression, naïve bayes, SVM, k-nearest neighbor, decision tree, random forest, and gradient boosting in the study with the accuracy of 88.6%. The results from the study support the idea that several different ML algorithms may greatly enhance suicide detection and prevention efforts.

Many people have expressed their emotions and thoughts about a wide range of topics using different social media platforms such as Twitter and Facebook [9], with Twitter being the commonly used to identify depression [10,11]. However, social media post alone is not enough to predict depression as not every depressed people express their feelings online. Most of the depressed people are active online and on social media, therefore prediction based on social media post alone may not be enough. Some underlying risk factors like income, race, employment status, age, friends, social fear etc. can make an individual to be depressed.

Evidence has been shown that demographic data can be used to provide indicators for depression and suicide predictions. Cruz et al. [12], applied Naïve Bayes classifier to predict depression levels among students using demographic and psychological survey data. Their model achieved an accuracy of 78.03% and revealed that risk factors like income level, academic stress and sleep patterns are useful indications of prediction while our study focuses on prediction of depression and suicide attempt using demographic data.

Tulubas et al. [13], conducted a comprehensive investigation into the relationship between digital addiction and academic achievement among students. Their research adds value to the predictive discourse by identifying digital behavior as a mediating variable. It proffers that pattern of technology use may serve as early indicators of mental health risks in youth populations.

Papadakis et al. [14], worked on a technological perspective by integrating a computer simulation and cloud-based systems into educational environment. The authors emphasis on the usability, interactivity and accessibility of cloud platforms may provide basic support for the deployment of predictive mental health system in educational settings. Their research brings out the necessity of adaptive technologies that can also monitor user behavior in real time for modern depression and suicide risks model.

In our study, features were selected based on the feature importance. However, manual feature selection allows for fine tuning of the model. The features that are relevant for the specific task was selected thus leading to a more effective model. The most important feature for both depression and suicide are job title, friends, age, education level, social fear, income, body weight, employment, race, sexuality and so on. Date and time were removed from the features as they do not serve as contributing factor to the target variable, likewise 'what help from others feature' and "improve yourself how" was removed because it is not categorical in nature.

In this work, an Adaboost ensemble model that predicts depression and suicide attempt status was developed and a web-based system that assists experts in guiding individuals from committing suicide was developed. The main risk factors for depression includes job title, friends and age while the one mainly responsible for suicide is job title, age and friend.

II. METHODOLOGY

The method employed in the study include data collection, preprocessing, balancing dataset using SMOTE, splitting the dataset into various training/test ratio to know the best data split, applying ML algorithm for both depression and suicide attempt, applying extra trees for Adaboost ensemble, evaluating the model performance and developing the system. Fig. 1 represented the architecture of the model for prediction.

A. Dataset Description and Limitations

The dataset used in this study is named foreverAlone, obtained from an online data repository known as Kaggle, which consists of demographic data. It is a categorical data of survey that took place from May, 2016 to September, 2016. The dataset originally served as a platform to share the "forever alone" meme, but over time it evolved into both an identity and a community for individuals who have experienced prolonged loneliness to discuss and share their struggles.

The dataset contains 19 columns and 469 records which are used for the prediction of depression and suicide attempt. The dataset is small and imbalanced for both target variable column (depressed and attempt suicide). Due to limitation of the small dataset, appropriate alternative was made to make predictions accurate which include using SMOTE, adopting Adaboost ensemble and using several performance metrics. Imbalanced dataset can affect the predictive performance of the research, that is the reason the data was balanced using SMOTE (synthetic minority oversampling techniques. SMOTE works by oversampling the minority class or under sampling the majority class. However, smote may introduce bias thus one of the reasons for using Adaboost ensemble that reduces bias and overfitting. SMOTE also makes training a model easier as it helps in preventing the model from becoming biased towards one class.

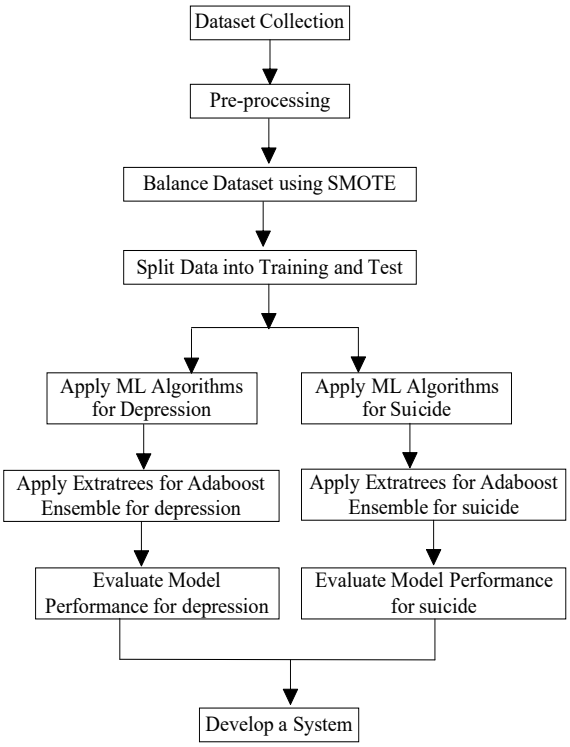


Fig. 1 Architecture of the Model

B. Data Preprocessing

The process of preparing the raw data and making it suitable for a ML model is known as data preprocessing. It is an important, of course the first step involved in creating a ML model. After downloading the dataset online, data preprocessing is the next step. Making the dataset suitable for a model will increase the accuracy and efficiency of the model.

In this study, missing values are first cleaned from the dataset. Some columns contained null values and incorrectly filled values. The job title column contains incorrectly filled values such as -, N/a, null values, *, and T. The correctly filled values were renamed as None. Categorical features were converted to numerical using label encoding. Age and income were preprocessed using min-max scaling method. The scaling ensured that all numerical features had values within a constant range.

C. Data Balancing and Limitation

The dataset downloaded from Kaggle was imbalanced, so, SMOTE was applied to balance both depression and suicide attempt column. The main benefit of SMOTE is its ability to improve the performance of machine-learning models [15].

A dataset is imbalanced if the classification categories are not approximately equally represented. The performance of the prediction will be improved when the dataset is balanced. It also makes training a model easier as it helps in preventing the model from becoming biased towards one class. In the depressed column, 312 participant was depressed while 157 was not depressed while the suicide attempt column has 85 participant that have made suicide attempt and 384 participant that have not made an attempt to commit suicide. Smote has limitation of bias and they are handled using multiple metrics and not relying on accuracy alone. Metrics like F1score, precision and recall are used to get a more balanced view of the model performance. Adaboost also plays a significant role in smote bias and overfitting due to synthetic redundancy.

Tables 1 and 2 illustrate the dataset description before and after applying smote for both depression and suicide attempt respectively. Smote increased the target variable in both the depressed and suicide attempt cases. Hyperparameter tuning was not applied in the study however future work will take it into consideration.

Table 1 Dataset for depression before and after SMOTE

	Depressed	**Not depressed**	**Total**
Before SMOTE	312	157	469
After SMOTE	303	303	606

Table 2 Representation of dataset for suicide attempt before and after SMOTE.

	Suicide Attempt	**No Suicide Attempt**	**Total**
Before SMOTE	85	384	469
After SMOTE	372	372	774

D. Data Splitting

The dataset was split into training and test set. Various data splitting percentage were used which includes (60/40), (70/30) and (80/20) training/test. This was done to compare result in order to choose data split which has the best accuracy. The machine learning algorithms were used on the training dataset while the test set was used to evaluate predictive performance of the models. However, the 80/20 training/test percentage split outperformed the other percentage splits and it was used and reported in this study. Table 3 shows the proportion of the dataset used for both training and testing in both cases.

E. Feature Importance Analysis

Feature importance was analyzed using SHAP values and Extra trees to identify the features that most influence prediction in both cases. The chart for both cases are shown in Figure 2 and 3.

Table 3 Proportion of data set used for data training and testing

	Training	Testing	Total
Depression	484	122	606
Suicide Attempt	595	149	744

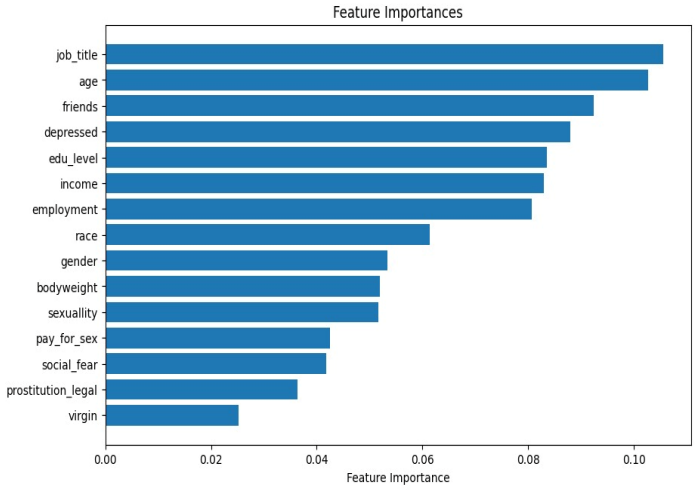


Figure 2: Suicide attempt Feature Importance Chart

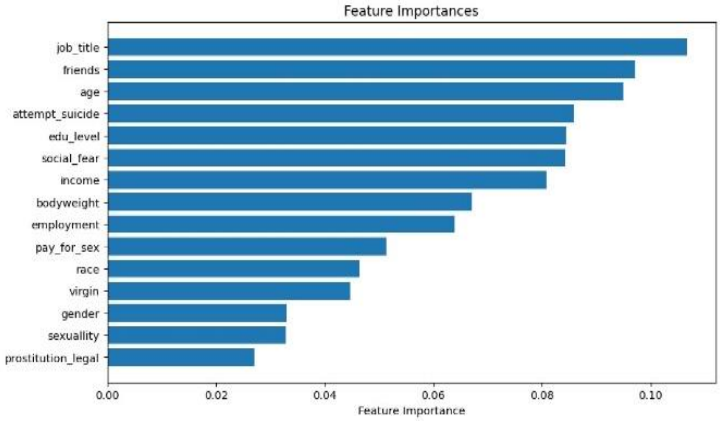


Figure 3: Depression Feature Important Chart

F. Algorithms used for Depression and Suicide Attempt prediction

Four Machine Learning algorithms used in this study include: SVM, extra trees, logistic regression and naïve bayes.

a) Support vector machine: SVMs are ML algorithm that identify data from classification and regression studies [16, 17]. SVM supervised learning approach is used for analyzing data and categorizing it into two groups. They are one of the robust prediction algorithms in ML. In this study, we considered using SVM as one of the methods for model training, because it is well suited for binary classification tasks. Also, it has shown good results in the medical diagnostics, electric load forecasting, optical character recognition and other fields. The data points (vectors) that are the closest to the hyperplane, which affect the position of the hyperplane are called support vector. Since these vectors support the hyperplane, hence called a Support vector.

b) Extra trees: As the name suggested, extra trees algorithm gets its name from the random way it constructs each decision tree and the way it processes the data, which really reduces the chance of overfitting the data. It belongs to the ensemble learning family. It combines the predictions from many decision trees. It is an extension of the random forest algorithm and shares many similarities with it. Extra trees algorithm randomly picks several columns of data without replacement, and try to determine which column best describes the data throughout the training of decision trees. Each column corresponds to an attribute which describes a feature of the data set. Its main strength lies in the computational efficiency [18].

c) Logistic Regression: Logistic regression otherwise known as logistic model or logit model, examines the relationship between a categorical dependent variable and multiple independent variables, and estimates the probability of occurrence of an event by fitting data to a logistic curve [19]. It can be used to classify an observation into one of two classes (like 'depressed' and 'not depressed, suicide attempt and no suicide attempt'), or into one of many classes.

d) Naïve bayes: Naive Bayes classifier assumes that the presence of individual feature in a class is unconnected to the presence of any other feature. Its main target is the text classification industry. Naive Bayes is mainly used for clustering and classification purpose and depends on the conditional probability of happenings [20]. Bayes' theorem also known as Bayes' Rule, is used to determine the probability of a hypothesis with prior knowledge and it depends on the conditional probability.

reduce bias that smote might have caused and to reduce variance. However, its strong performance with weak learners also serves as a contributing factor. It also outperformed all other boosting method when trying it with the dataset and performed excellently well. Boosting reduces bias, noise and variance error. Adaboost pays more attention to the misclassified samples by assigning more weight to the misclassified until they are all correctly classified. It is good for classification problem and also update weight iteratively until it finally reaches its desired aim before prediction.

In adaptive boosting, the base classifier updates the weight attached to each of the observations in the dataset. The datapoint with higher accuracy would be assigned a lower weight. The weighted dataset is used as a training set for the next weak learner. The resulting dataset is re-weighted for a misclassified sample and used as the input for the next classifier. The base algorithm read the data and assign equal weight to each sample observation. Extra trees were used as the base algorithm for the ensemble model due to its higher computational efficiency.

G. Applying Adaboost and Justification for its use for Prediction

Adaboost Ensemble model was applied for both depression and suicide attempt. In Boosting, trees are grown using the information from a previously grown tree one after the other. The adaptive boosting (Adaboost) combines weak classifiers into a single strong classifier. It is sequential in nature.

H. Performance Evaluation of the Model

The evaluation of the model was done using confusion matrix. It sums up the number of correct and incorrect predictions. It is a 2 X 2-dimensional matrix because we are dealing with binary classification. Some model performs more better than the other. The performance differences in the model can be attributed to their handling of non-linearities and feature contribution. For instance, Naïve Bayes underperform due to its assumption of feature independence while Extra trees and Adaboost better captures complex interactions between features. Tables 4 and 5 show the confusion matrix for both cases.

Table 4 Confusion Matrix for Depression.

Actual Class	**Predicted Class**	
	Depressed	Not depressed
Depressed	TP	FN
Not Depressed	FP	TN
Total	Positive	Negative

Table 5 Confusion Matrix for Suicide Attempt.

	Predicted Class	
Actual Class	Suicide Attempt	No Suicide Attempt
Suicide Attempt	TP	FN
No Suicide Attempt	FP	TN
Total	Positive	Negative

Abbreviations: TN = True Negative; FN = False Negative; FP = False Positive and TP = True Positive.

I. Terminologies of the confusion matrix

a) True Positives [TP]: These are the positive cases that the classifier properly classified as depressed and attempt suicide.
b) True Negatives [TN]: These are the negative cases that the classifier properly classified as not depressed and no suicide attempt.
c) False Positives [FP]: These are the negative cases that were wrongly classified as positive.
d) False Negatives [FN]: These are the positive cases that were wrongly classified as unfavorable.

J. Technology Used for the Development of the System

The implementation of the prediction system was done using ML libraries, Scikitlearn, Django, Google Collaboratory notebook, cascading style sheet (CSS)3, JavaScript, and hypertext markup language (HTML) 5.

III. Results and Discussion

A. Extra Tree Model Result

Out of the four ML models employed in this study, the extra trees model performance is the best across all evaluation metrics. The model's performance show how well it performs when addressing both cases across all evaluation metrics as affirmed by the confusion matrices obtained from the model (see Fig. 4 and Fig 5). So, due to its computational efficiency, with accuracies of 76.23% for depression and 89.26% for suicide attempt respectively, it was employed as the base estimator for the boosting ensemble method, however in Cruz et al. the performance metrics achieved is an accuracy of 78.03%, precision of 75.68%, recall of 80.39% and F1 score of 77.96%.

It can be deduced from Fig. 4 that the model achieved a true negative of 53 and a true positive of 40 for its prediction. However, 10 of its instances were wrongly predicted as "not depressed" (false negative) while 19 instances were wrongly predicted as "depression" (false positive). This implies that Extra trees classifier correctly predicted the absence of depression in 53 instances, wrongly predicted the absence of depression in 19 instances, correctly predicted the presence of depression in 40 instances and wrongly predicted 10 cases of depression.

Also, the performance metrics for the suicide attempt predictive model using the extra trees classifier produced excellent results. Fig. 4 shows the confusion matrix obtained from extra trees classifier on the dataset for suicide attempt model. From the diagram, it is deduced that the model achieved a true negative of 72 and a true positive of 61 of its prediction. However, 4 of its instances were wrongly predicted as "not attempt suicide" (false negative) while 12 were wrongly predicted as "suicide" (false positive) whereas the actual value is "not attempting suicide".

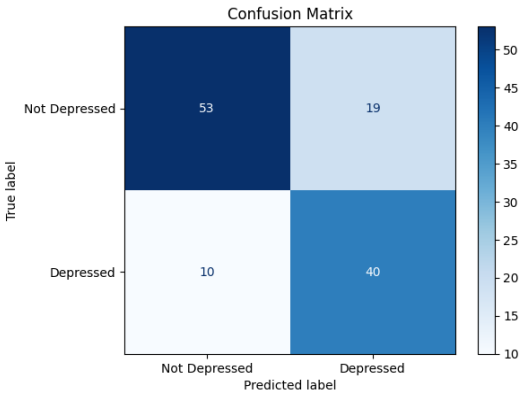


Fig. 4 Confusion Matrix obtained from Extra Trees Classifier for Depression

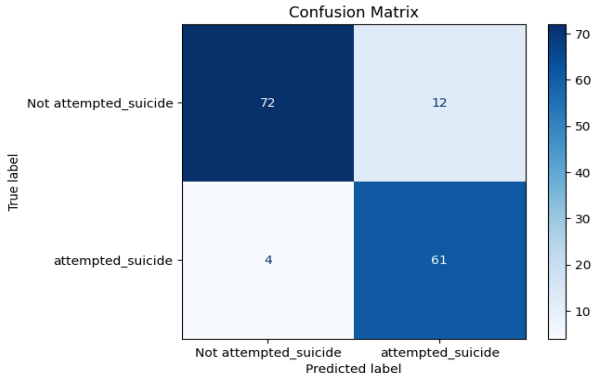


Fig. 5 Confusion Matrix obtained from Extra Trees for Suicide attempt.

B. Ensemble Model Results

As stated earlier, balanced dataset at 80/20 training/ test percentage split performed best when compared with other splits and was used in this study. The confusion matrices obtained from the ensemble model for both depression and suicide attempt are shown in Fig. 6 and Fig. 7 respectively.

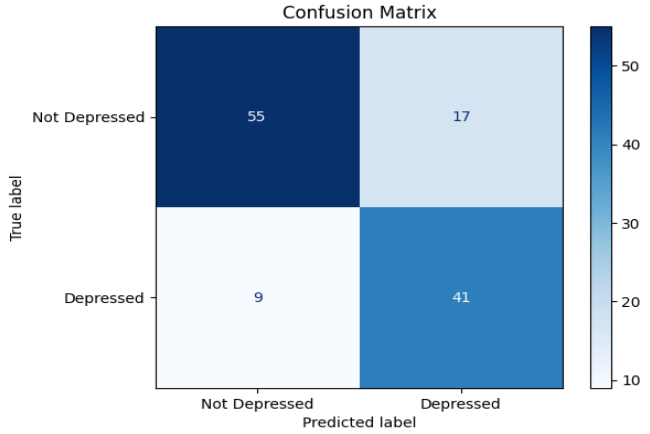


Fig. 6 Confusion matrix obtained from the ensemble model for depression.

The confusion matrix obtained from the ensemble model for predicting depression (Fig. 6) shows that the model achieved a true positive of 41 cases and true negative of 55 cases. However, 17 of its instances were wrongly predicted as "depressed" (false positive) while 9 instances were wrongly predicted as "not depressed" (false negative). This indicate that 55 instances were correctly predicted as not depressed while 41 was correctly predicted as depressed.

Also, Fig. 7, which is the confusion matrix obtained from the ensemble model for predicting suicide attempt, shows that the model achieved a true negative of 74 cases and a true positive of 61 cases of its prediction. However, 4 of its instances were wrongly predicted as "not suicide" (false negative) while 10 were wrongly predicted as "suicide" (false positive). This indicate that 74 instances were correctly predicted as no suicide attempt while 61 cases were predicted as suicide attempt.

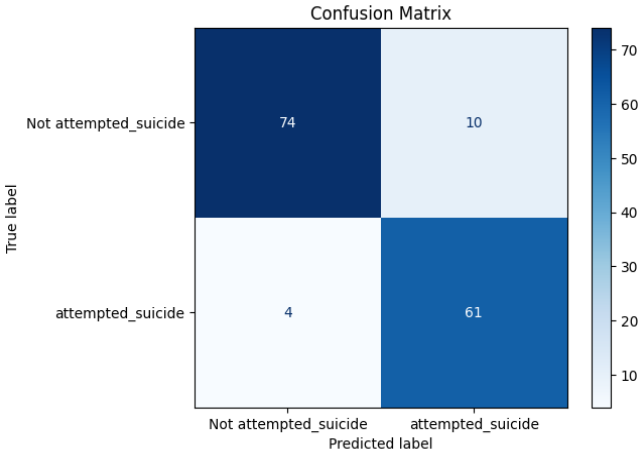


Fig. 7 Ensemble model for suicide attempt

IV. Results Discussion

Tables 6 and 7 summarize the performance of the classifiers and ensemble model used for predicting depression and suicide attempts. Among the models, Extra Trees achieved the highest performance across all evaluation metrics and was chosen as the base estimator for the Adaboost ensemble model. For depression prediction, the Extra Trees model achieved 76.23% accuracy, 67.80% precision, 80.00% recall, and a 73.39% F1-Score. Similarly, for suicide attempt prediction, the metrics were 89.26% accuracy, 83.56% precision, 93.85% recall, and an 88.41% F1-Score

Table 6 Summary of the models' performance for depression.

S.No	Algorithm	Accuracy	Precision	Recall	F1-score
1	SVM	50.82%	44.57%	82.00%	57.75%
2	**Extra Trees**	**76.23%**	**67.80%**	**80.00%**	**73.39%**
3	Logistic regression	75.41%	67.24%	78.00%	72.22%
4	Naïve Bayes	71.31%	65.31%	64.00%	64.65%
5	**Adaboost Ensemble**	**78.69%**	**70.69%**	**82.00%**	**75.93%**

Table 7 Summary of the models' performance for suicide attempt

S.No	Algorithm	Accuracy	Precision	Recall	F1-score
1	SVM	51.68%	47.33%	95.38%	63.27%
2	**Extra Trees**	**89.26%**	**83.56%**	**93.85%**	**88.41%**
3	Logistic regression	75.41%	67.24%	78.00%	72.22%
4	Naïve Bayes	71.14%	65.28%	72.31%	68.61%
5	**Adaboost Ensemble**	**90.60%**	**85.92%**	**93.85%**	**89.71%**

However, the Adaboost ensemble model demonstrated better performance than the Extra Trees model. Its metrics for depression prediction include 78.69% accuracy, 70.69% precision, 82.00% recall, and a 75.93% F1-Score. For predicting suicide attempts, the results were even higher: 90.60% accuracy, 85.92% precision, 93.85% recall, and an 89.71% F1-score, meaning that the ensemble model yielded remarkable accuracy in classifying a significant portion of the dataset across both scenarios. Its ability to identify positive cases while minimizing false positives is reflected in its precision scores of 70.69% for depression prediction and 85.92% for suicide attempt prediction.

Furthermore, the model showed exceptional sensitivity in detecting true positive instances, achieving recall rates of 82.00% and 93.85% for depression and suicide attempt respectively, which correspond to low false negative rates. The F1-score, representing a balance between precision and recall, gave 75.93% for depression and 89.71% for suicide attempt predictions respectively. These metrics highlight the ensemble model's robust predictive capabilities for the given task.

A. Implementing the Model in a Web-Based Application

The Adaboost Ensemble model was used to implement a web-based application since the predictive power of the ensemble model is better than those of individual models. This model can now be used to predict depression and suicide attempt status of new individuals/patients by integrating it in the real-life prediction process where the expert input the patient's data and get the required result. This implementation is categorized into two webpages. The home page and result page.

However, the aspect of user authentication does not occur because the system was developed for expert during screening and not for public use. Although access control will be deployed in future work. The user interface is more user friendly enough and self-explanatory to use without any difficulty for expert.

B. Experimental Result

The home page of the system gives room for the user to input the data required to check for depression status. The description of the home page is presented in Fig. 8. The data of the individual who needs help is entered and the user then click "Predict". The inputs to the depression model are: age, income, gender, sexuality, race, bodyweight, virgin, prostitution legal, pay for sex, social fear, attempt suicide, friends, employment, job title, education level while the Output is Depressed or Not Depressed.

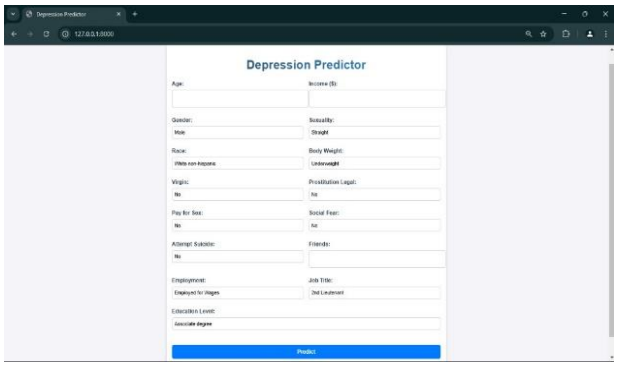


Fig. 8 Description of the Home page

Upon clicking predict, the system shows the status of the patient to be "not depressed' or "depressed". If the patient is not depressed, the system takes no action and the user goes back to home page as shown in Fig. 9.

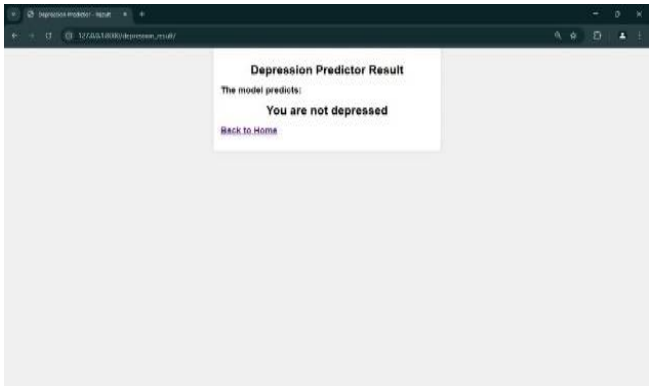


Fig. 9 Result Page: "Not Depressed"

However, if the patient is depressed as shown in Fig. 10, a suicide test is required and there is a link "Take suicide test" that takes the user to the suicide predictor page where suicide attempt prediction takes place. The inputs to the suicide attempt predictor are age, income, gender, sexuality, race, virgin, pay for sex, prostitution legal, social fear, bodyweight, depression, employment, friends, job title, education level, while the output is Suicide Attempt or No Suicide attempt.

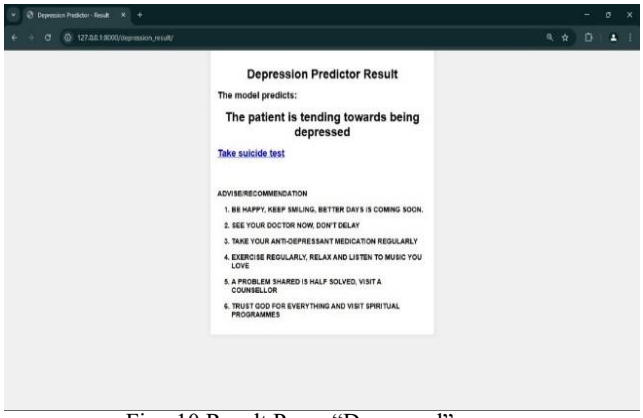


Fig. 10 Result Page: "Depressed".

The expert enters the input based on the patient response and the system predicts and displays the suicide attempt status of the patient as shown in Figs 11 and 12.

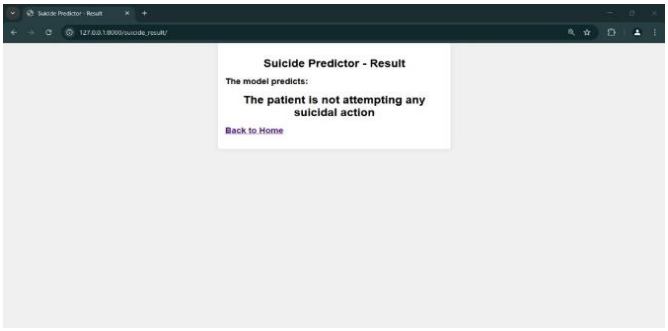


Fig. 11 Suicide Predictor Result Page: No suicidal attempt.

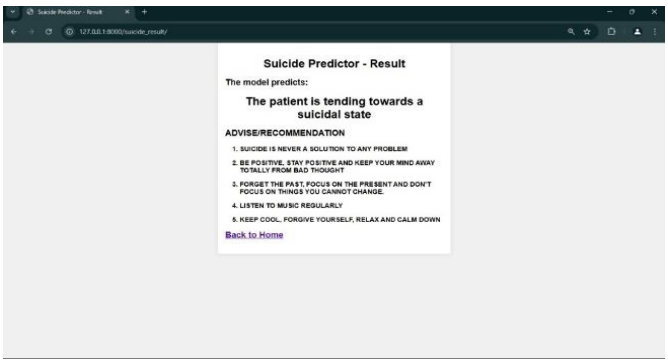


Fig. 12 Recommendation/Advise page.

V. Conclusion and Future Works

The study used four machine learning algorithms for predicting both depression and suicide attempt using the foreverAlone dataset. However, balancing the dataset before developing a model plays a significant role in the study by improving its predictive performance and accuracy. Adaboost ensemble approach, Logistic Regression, Naive Bayes, and Extra Tree classifier all demonstrated their distinctive advantages.

In summary, the study developed an Adaboost ensemble prediction model for depression and suicide attempt, demonstrating great accuracy of the Adaboost for mental health classification problem and deployed the model with a user-oriented web-based system with ethical safeguards has the lowest accuracy. Ensemble method also plays a significant role in predicting depression and suicide attempt with better accuracy than other ML algorithms.

The system developed act like a decision support system for expert and not a replacement for diagnostic evaluation. Comparing the training and testing data percentage split improved the study as 80/20 training/test outperformed the other training/test percentage splits. The role of data split in ML cannot be over emphasized. The accuracy of each model depends on the method used and the algorithm itself. However, it is crucial to take into account the particular aim and objectives when selecting a model because different algorithms may be more suited for different tasks. To get better outcomes, it is advised to continue optimizing and fine-tuning these models. The use of k-fold cross validation can also be incorporated in the future work. Ethical AI practices, interpretability and collaboration with clinical stakeholders and psychologists is very important in future work

Additionally, the generalization and robustness of the model could be enhanced by the incorporation of larger and more varied datasets from external health providers like WHO, NHANES for real world validation is highly important in future work.

A. Ethical Issues

The dataset used is publicly available on online dataset repository known as Kaggle and it is anonymized, therefore getting formal institution review board may not be necessary or required.

Based on the prediction models, real-time monitoring and intervention strategies might be created to help those who are at risk for depression and suicide. If the system operates without human intervention, it may lead to dangerous decision, violation of medical ethics and data protection.

B. Implication of Misclassification (FN &FP)

The risk of false negatives and false positives have ethical implications. It is therefore necessary to encourage safe deployment and use with expert human intervention.

False positive will cause unnecessary anxiety for the individual after receiving inaccurate prediction and also the expert limiting access for those in critical need and allocating time for those that doesn't need urgent care. However, all positive predictions should be reviewed by trained mental health before intervention and also fine tuning the model is highly important.

False negatives will cause a high-risk individual not to receive timely help due to error in diagnosis. This will increase the severity of depression and increase suicide attempt. This can be reduced by taking into consideration routine reassessment of individual especially those at higher risk level.

References

[1] World Health Organization, *Depression and Other Common Mental Disorders: Global Health Estimates*, Geneva: Licensed: CC BY-NC-SA 3.0 IGO, 2017.

[2] J. Reutfors, T. M. L. Andersson, A. Tanskanen, A. DiBernardo, G. Li, L. Brandt, and P. Brenner, "Risk factors for suicide and suicide attempts among patients with treatment-resistant depression: Nested case-control study," *Arch. Suicide Res.*, vol. 25, no. 3, pp. 424–438, 2021, doi: 10.1080/13811118.2019.1691692.

[3] R. Mahadik, S. Salunkhe, S. Sneha, and V. Sagvekar, "Analysis of suicide attempts and its prediction," *Int. Res. J. Eng. Technol.*, vol. 8, no. 4, pp. 2511–2517, Apr. 2021.

[4] J. Angskun, S. Tipprasert, and T. Angskun, "Big data analytics on social networks for real-time depression detection," *J. Big Data*, vol. 9, no. 1, 2022, doi: 10.1186/s40537-022-00622-2.

[5] T. H. H. Aldhyani, S. N. Alsubari, A. S. Alshebami, H. Alkahtani, and Z. A. T. Ahmed, "Detecting and analyzing suicidal ideation on social media using deep learning and machine learning models," *Int. J. Environ. Res. Public Health*, vol. 19, no. 19, 2022, doi: 10.3390/ijerph191912635.

[6] S. Mahmud, M. Mohsin, A. Muyeed, S. Nazneen, M. A. Sayed,

and N. Murshed, "Machine learning approaches for predicting suicidal behaviors among university students in Bangladesh during the COVID-19 pandemic," 2022, doi: 10.21203/rs.3.rs-2069873/v1.

[7] G. Dhasmana, A. Gujjar, and M. Sharma, "DSTP: Depression and suicidal ideation prediction using social media data," *AIP Conf. Proc.*, vol. 060002, 2022, doi: 10.1063/5.0105936.

[8] S. Saha, S. Dasgupta, A. Anam, R. Saha, S. Nath, and S. Dutta, "An investigation of suicidal ideation from social media using machine learning approach," *Baghdad Sci. J.*, vol. 20, no. 3, suppl., pp. 1164–1181, 2023, doi: 10.21123/bsj.2023.8515.

[9] R. Safa, P. Bayat, and L. Moghtader, "Automatic detection of depression symptoms in Twitter using multimodal analysis," *J. Supercomput.*, vol. 78, no. 4, pp. 4709–4744, 2022, doi: 10.1007/s11227-021-04040-8.

[10] S. Smys and J. S. Raj, "Analysis of deep learning techniques for early detection of depression on social media network—a comparative study," *J. Trends Comput. Sci. Smart Technol.*, vol. 3, no. 1, pp. 24–39, 2021, doi: 10.36548/jtcsst.

[11] D. Liu, X. L. Feng, F. Ahmed, M. Shahid, and J. Guo, "Detecting and measuring depression on social media using a machine learning approach: Systematic review," *JMIR Ment. Health*, vol. 9, no. 3, e27244, 2022, doi: 10.2196/27244.

[12] F. T. Cruz, C.R. Panganiban, M. R. Anthonio, and R. M. Lualhati, "Prediction of depression status in college students using a Naïve Bayes classifier-based machine learning model," arXiv preprint, arXiv:2307.14371, Jul.2023

[13] T. Tulubas, T. Karakose, and S. Papadakis, "A holistic investigation of the relationship between digital addiction and academic achievement among students," Eur. J. Investig. Health psychol. Educ., vol. 13, no. 10, pp.2006-2034, Oct. 2023, doi: 10.3390/ejihpe13100242.

[14] S. Papadakis et al., "Revolutionizing education: using computer simulation and cloud-based smart technology to facilitate successful open learning," in Proc. CosinEi and CSTOE 2022 Workshops co-located with ACNS Conf. on Cloud and Immersive Technologies, CEUR Workshop Proc., vol. 3358, pp. 1-18, 2023.

[15] K. Dhiraj, "How to handle unbalanced data with SMOTE," *Baeldung*, 2023. [Online]. Available: https://www.baeldung.com/cs/synthetic-minority-oversampling-technique

[16] B. Mahesh, "Machine learning algorithms – a review," *Int. J. Sci. Res.*, vol. 9, no. 1, pp. 381–386, 2020.

[17] A. A. Kayode, N. O. Akande, A. A. Adegun, and M. O. Adebiyi, "An automated mammogram classification system using modified support vector machine," *Med. Devices Evid. Res.*, vol. 12, pp. 275–284, 2019.

[18] E. K. Ampomah, Z. Qin, and G. Nyame, "Evaluation of tree-based ensemble machine learning models in predicting stock price direction of movement," *Inf.*, vol. 11, no. 332, 2020.

[19] D. R. Shah and R. Pradhananga, "Assessment of pedestrians' red light violation behavior at signalized crosswalks in Kathmandu, Nepal," *Transp. Res. Interdiscip. Perspect.*, vol. 24, 101035, 2024.

[20] B. Kundu and D. S. Naidu, "Classification and feature extraction of different hand movements from EMG signal using machine learning-based algorithms," in *Proc. Int. Conf. Electr. Commun. Comput. Eng. (ICECCE)*, 2021, pp. 1–5.

Authors' Contribution

The authors confirm contribution to the paper as follows: Akinyele O.T: Supervision, Conceptualization, Methodology, Project administration, Writing-Original draft preparation, Reviewing and Editing;
Kayode A.A: Supervision, Visualization, Data curation;
Adegoke-Elijah A.: Data curation, Methodology, Investigation, Resources,
Olowookere T.A: Investigation, Software, Validation, Writing – review & editing.
All authors reviewed the results and approved the final version of the manuscript.

Competing Interests

Authors declared that they have no competing interests.

Data Availability

Dataset used for this study is available at https://www.kaggle.com/datasets/kingburrito666/the-demographic-rforeveralone-dataset

Dataset generated during the study after smote are available from the corresponding author upon request

Funding

There is no funding for this work.

Acknowledgements

All authors whose works are referenced in this paper are sincerely acknowledged.

Design And Fabrication of Color Sorting Machine Based on Computer Vision

Ibrahim Ahmed Alrushidy [1, *]
Yousef Abobakr Bahumid [1]
Ramzi Abdulsalam Bin Shujaa [1]
Ahmed Ehab Abdullah [1]
Ammar Ali Kurd [1]
Mohammed Saeed Moqbel Abdullah [1]

Received: 21/04/2025
Revised: 26/05/2025
Accepted: 27/05/2025

© 2025 University of Science and Technology, Aden, Yemen. This article can be distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

© 2025 جامعة العلوم والتكنولوجيا، المركز الرئيس عدن، اليمن. يمكن إعادة استخدام المادة المنشورة حسب رخصة مؤسسة المشاع الإبداعي شريطة الاستشهاد بالمؤلف والمجلة.

[1] University of Science and Technology – Yemen
*Corresponding Author's Email: a.kurd@ust.edu

Design And Fabrication of Color Sorting Machine Based on Computer Vision

Ibrahim Ahmed Alrushidy
University of Science and Technology – Yemen
alrushidy.ibrahim@gmail.com

Yousef Abobakr Bahumid
University of Science and Technology – Yemen
yousef.abobakrb@gmail.com

Ramzi Abdulsalam Bin Shujaa
University of Science and Technology – Yemen
ramzi2014.99@gmail.com

Ahmed Ehab Abdullah
University of Science and Technology – Yemen
ahmedehabm31@gmail.com

Ammar Ali Kurd
University of Science and Technology – Yemen
a.kurd@ust.edu

Mohammed Saeed Moqbel Abdullah
University of Science and Technology – Yemen
m.saeed@ust.edu

Abstract— Manual sorting of objects by color can be exhausting and time-consuming, resulting in high labor costs and the potential for errors. At the same time, achieving accurate automation of this sorting process poses a challenging engineering task. This study presents a real-time method utilizing image processing of video frames collected by a camera and processed by an onboard Raspberry Pi computer. Unlike systems relying on color sensors or AI, this work achieves comparable accuracy using cost-effective, fast color identification with different colors' shades. An algorithm was specifically developed using the OpenCV package for this objective. The sorted objects are then transferred by a robotic gripper mechanism to their respective bins. The mechanical side will have a robotic arm and conveyor belt. The proposed design will be proven to be an effective and economical solution based on its capacity to pass tests set to evaluate its performance, where it exhibited a 90% accuracy under ideal operating conditions. The strategy being proposed is ideal for industrial applications that require color-sorting machinery, such as the fruit packaging and food processing sectors.

Keywords— Color-sorting, image processing, robotic arm, real-time processing, object detection, mechatronics, conveyor belt, computer vision, Raspberry Pi, OpenCV."

Introduction

In industrial automation, image processing has become crucial for many applications. This integration has paved the way for creative solutions that boost efficiency and precision. One significant application of this technology is categorizing things based on their color. For instance, in industrial automation applications requiring color detection such as sorting objects by color, image processing techniques and algorithms can be utilized to achieve this purpose. This will effectively speed up the sorting process, reduce costs and minimize human errors [1]. Most industrial applications demand sorting in real time, which in turn leads to employing cameras and using sophisticated algorithms and sensors [1]. This will streamline production line conditions for diverse products and as more companies adopt automation, the demand for image processing will continue to grow.

Although the color detection problem appears to be the same in all applications, there are differences in this problem that arise from the specific needs of the production process itself, such as the speed and accuracy of detection. This results in a number of detection methods, including the use of artificial intelligence [3, 4, 12], simple color thresholding techniques [2, 5], or specialized color sensors [6–8].

Studies on color detection with the use of AI or color thresholding approaches involve microprocessor-based solutions such as the usage of Raspberry Pi, generally with image processing software libraries such as OpenCV [2, 9-13]. On the other hand, investigations using dedicated color sensors adopt microcontroller-based solutions such as Arduino boards [6-8].

Dedicated color sensors offer cost advantages; however, they are inadaptable and have a limited accuracy, which renders them inadequate for dynamic industrial environments. AI-based solutions are more accurate and highly adaptable but demand higher computational performance, which leads to an increase in cost and limiting speed if operated on modest hardware. In this work we propose an automated production line (robotic arm and conveyor belt) control system using a color thresholding method based on hue, saturation, and value (HSV) color space thresholding on a microprocessor-based solution employing Raspberry Pi and OpenCV. The suggested system aims to establish a balance between cost, accuracy, and adaptability. Studies on color detection using Al or color thresholding approaches involve microprocessor-based solutions such as Raspberry Pi, typically with image processing libraries like OpenCV [2, 9-13].

I. METHODOLOGY

This section presents the materials, methods, and overall flow of the system development. The study outlines the core hardware and software components employed in the system's development, including the Raspberry Pi 4 Single-Board Computer (SBC), Arduino Uno microcontroller development board, Arducam Camera Module, Python programming language, and OpenCV library. Each component plays a critical role in enabling the functionality and integration of the sorting process.

The design methodology is divided into three principal phases:

1. Embedded System Design: Focused on electronic control and software integration.

2. Mechanical Design: Pertaining to the physical assembly and actuation mechanisms.
3. System integration: In this final phase the work from the previous two phases is integrated and tested.

The sequential flow of system design and integration is illustrated in Figure 1, which provides a structured overview of the developmental workflow.

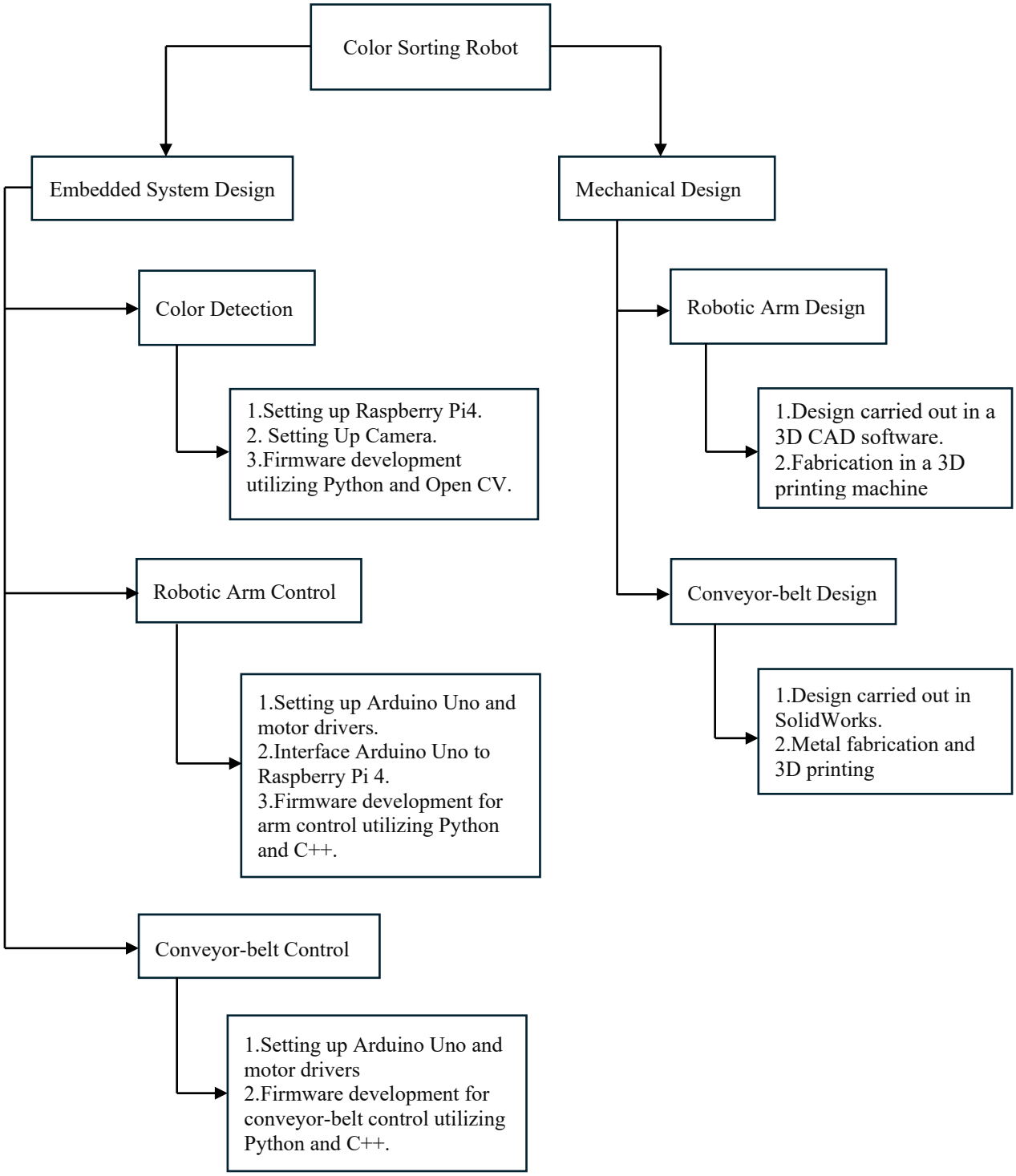


Figure 1 System design flow

Proposed embedded system

The embedded system architecture comprises a Raspberry Pi 4 serving as the master controller, with an Arduino Uno functioning as a slave controller to interface with motor drivers and regulate the mechanical actuation of the robotic arm and conveyor belt. As illustrated in Table 1, a comparative analysis of various Raspberry Pi models confirms that the Raspberry Pi 4 fulfills the computational demands necessary for real-time image processing tasks.

Table 1 Comprehensive comparison of various raspberry pi models [14]

Model	CPU	RAM	USB Ports	Ethernet	WiFi	Bluetooth	GPIO Pins	HDMI Ports	Size	Power
Raspberry Pi Model B	Single-core 700MHz	512MB	2	Yes	No	No	26	1	85.60mm x56mm	5V
Raspberry Pi 2	Quad-core 900MHz	1GB	4	Yes	No	No	40	1	85.60mm x56mm	5V
Raspberry Pi Zero	Single-core 1GHz	512MB	1	No	No	No	40	mini HDMI	65mmx30mm	5V
Raspberry Pi 3 Model B+	Quad-core 1.4GHz	1GB	4	Yes	Yes	Yes	40	1	85.60mm x56mm	5V
Raspberry Pi 4 Model B	Quad-core 1.5GHz	2-8GB	4	Yes	Yes	Yes	40	2 micro HDMI	85.60mm x56mm	5V

For image acquisition in color detection, a compatible camera module was selected to ensure seamless integration with the Raspberry Pi 4. This study employed a 5-megapixel (MP) camera from Arducam, which interfaces with the Raspberry Pi 4 via the Mobile Industry Processor Interface (MIPI). While a lower-resolution camera could theoretically suffice, the marginal cost difference rendered higher specifications a more practical choice.

Color recognition process

The system performs real-time color identification using HSV color space thresholding and contour detection algorithms implemented through OpenCV. As shown in Figure 2.

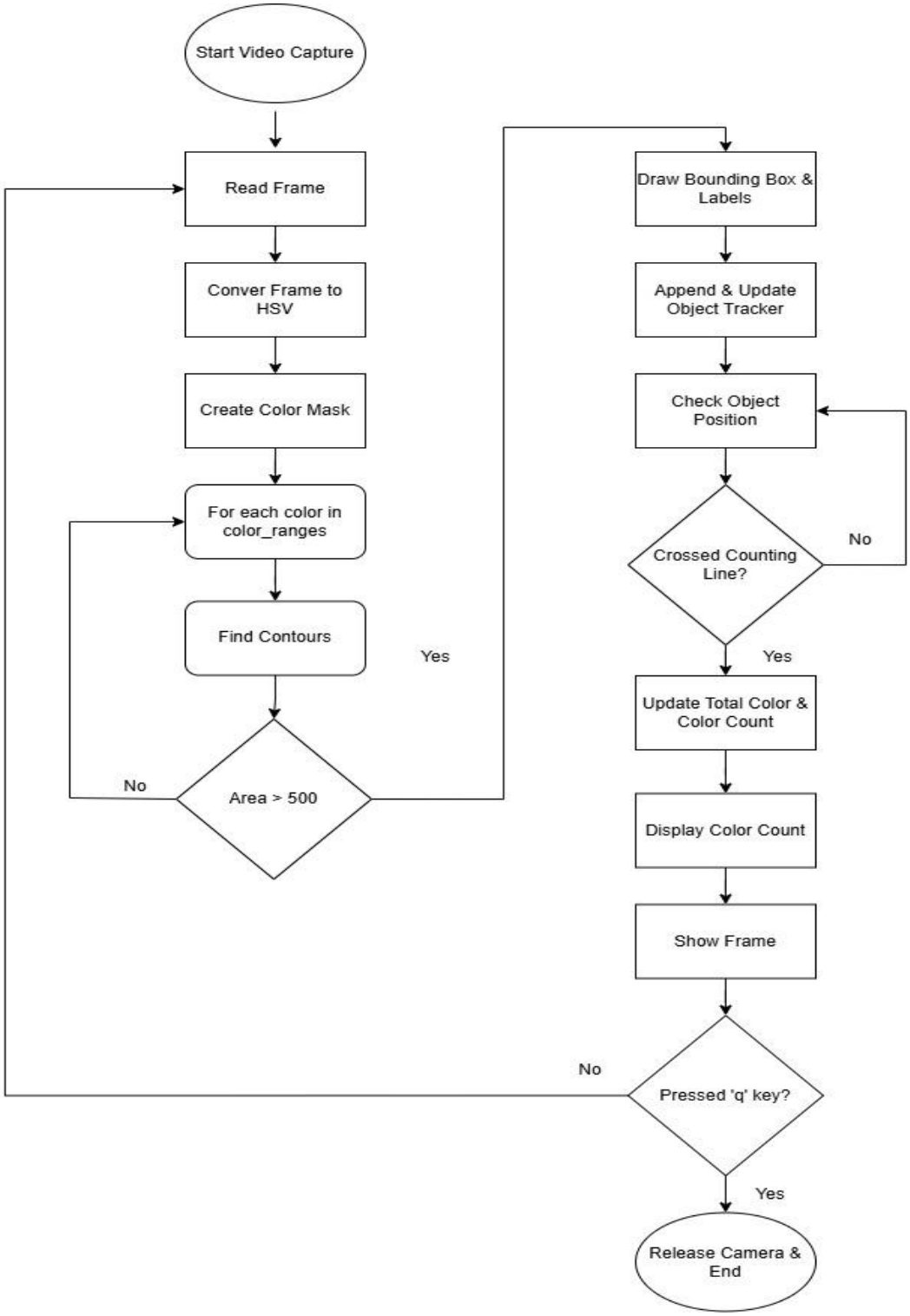


Figure 2 Color detection flowchart

Proposed mechanical structure

The mechanical assembly comprises two primary components:

1. The robotic arm designed using CAD software and manufactured via 3D printing.
2. The conveyor-belt system was CAD-modeled and fabricated via metalworking processes.

Robotic arm

The robotic arm used for this project is a standard programmable mechanical system composed of multiple rigid segments connected by motorized joints, forming a kinematic chain. Each joint—actuated by servo motors—enables rotational or translational movement, granting the arm its range of motion.

Conveyor Belt

The conveyor belt system transports materials sequentially for recognition in an assembly-line configuration. As illustrated in Figure 3, the complete system architecture is presented through detailed computer-aided design (CAD) schematics, which provide comprehensive visualization of the integrated components and their spatial relationships.

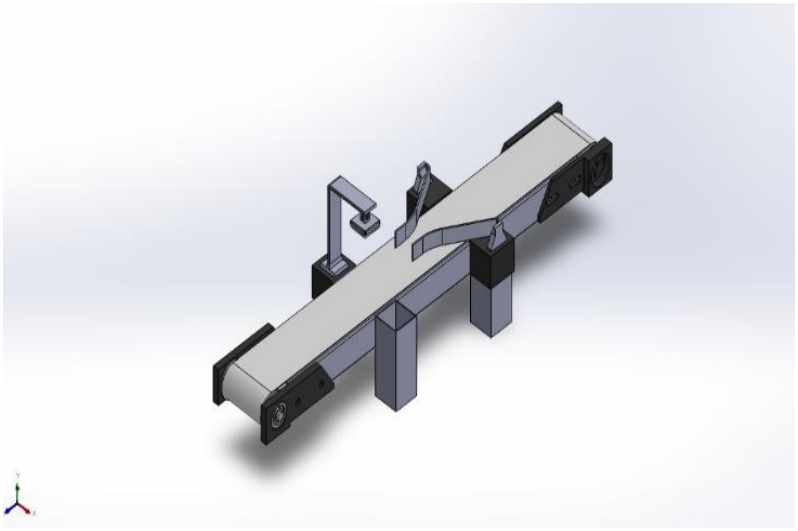


Figure 3 3D view of the complete conveyor belt

Its continuous motion is achieved through a closed-loop mechanism driven by two rollers positioned at opposite ends of the belt assembly. Figure 4 presents the computer-aided design (CAD) schematic of the roller component, illustrating its structural and functional details.

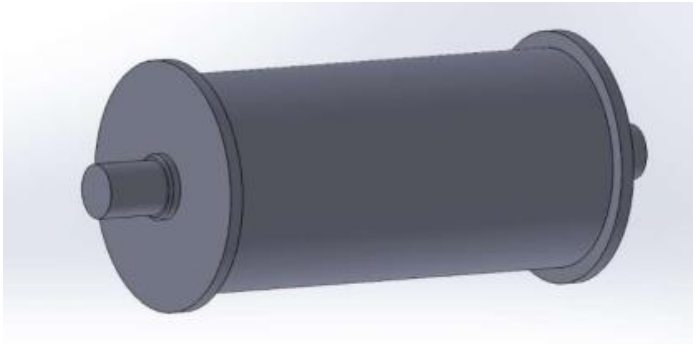


Figure 4 3D view of the roller design

System integration

To test the system, it was decided to sort based on three colors: yellow, green, and blue. However, the system can be adaptable to other colors easily. The overall flow of the system is depicted in Figure 5.

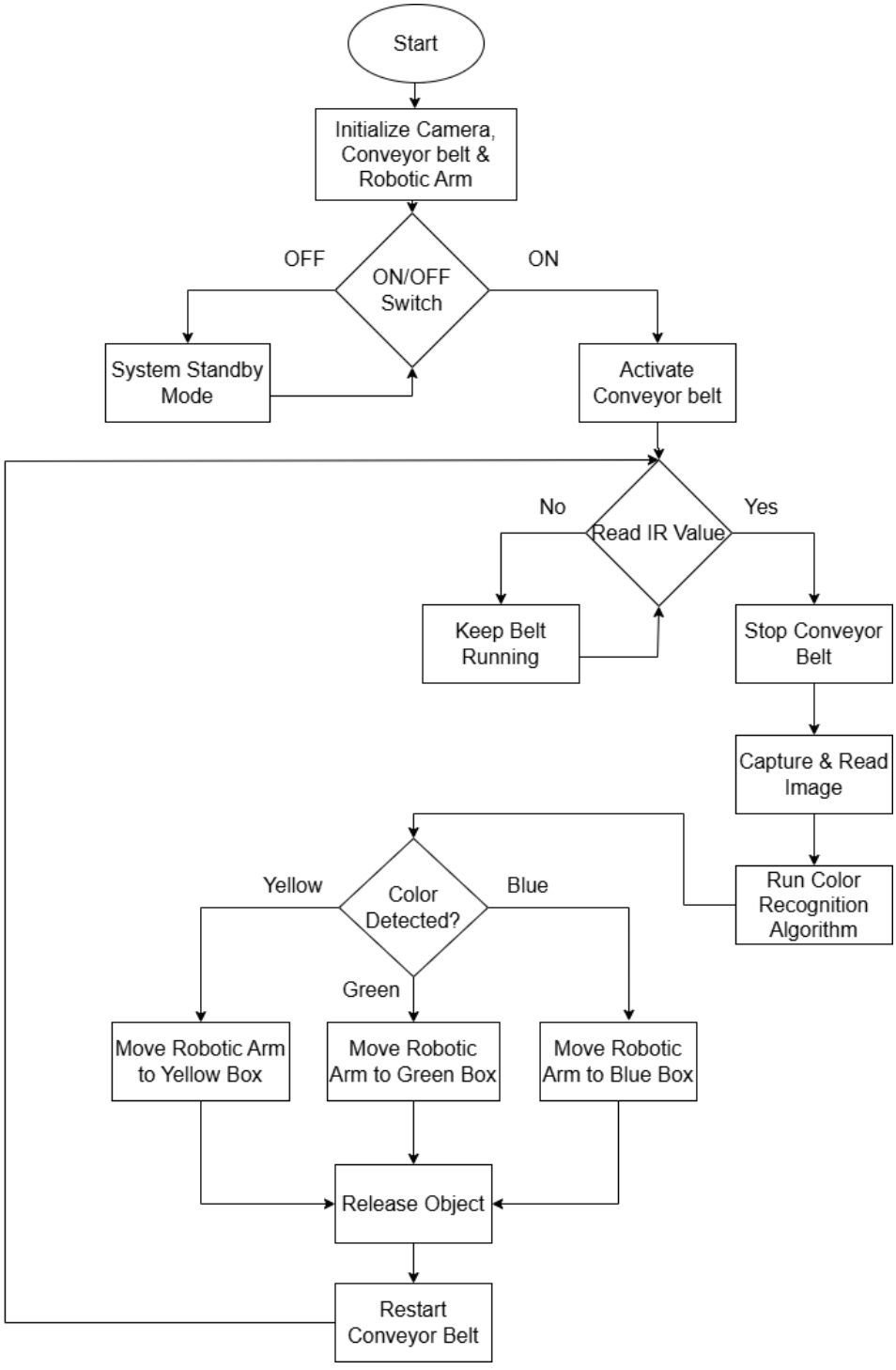


Figure 5 System integration flowchart

II. RESULTS

The color-based sorting robotic system was successfully designed, implemented, and evaluated. The complete system is shown in Figures 6.

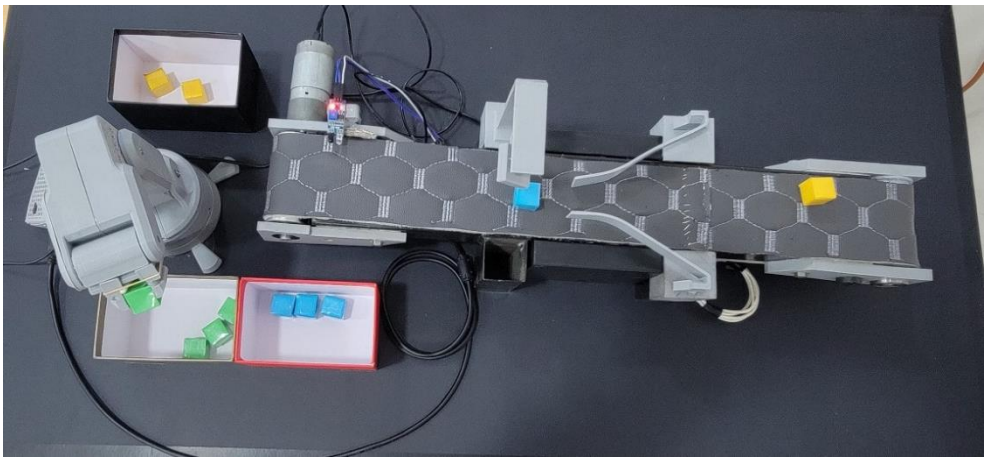


Figure 6 Prototype of the computer vision-based sorting system performing real-time color classification.

Through systematic testing, the system demonstrated the capability to autonomously classify and sort objects based on color with an average processing rate of one item every five seconds. The system exhibited low latency in processing images and executing sorting commands, with an average response time of 0.5 seconds from detection to giving an action command to the arm to pick up the object.

The color detection algorithm was evaluated under various lighting conditions. The algorithm demonstrated high accuracy, successfully identifying and categorizing ~90% of the items correctly under standard lighting conditions. Under challenging lighting scenarios, such as low light and varying light intensity, the accuracy might decrease to ~85%, indicating the algorithm's robustness and reliability.

The conveyor belt maintained a somewhat consistent speed, ensuring smooth movement of items to the detection area. Some mechanical failures were observed during testing where the conveyor belt had some aligning issues where the object would swing to the sides, which made the object not land where the robot arm grip was. The issue was solved by designing a funneling system on each side to keep the object in the middle, as seen in Figure 7, highlighting a more functional design.

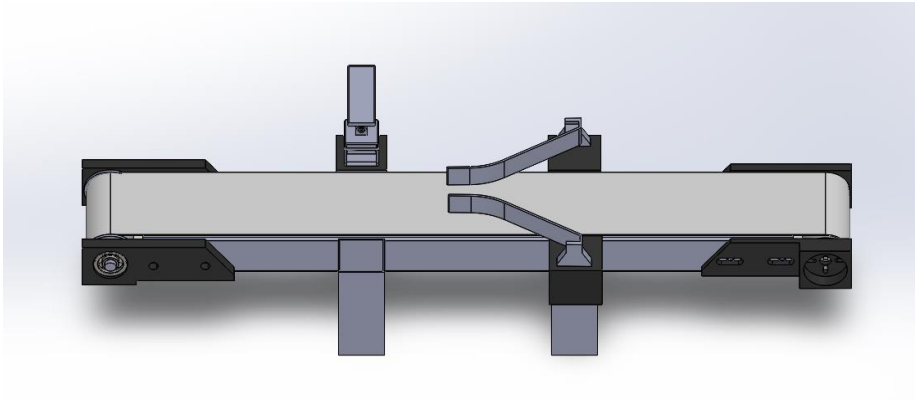


Figure 7 Top view of the convyer belt showing the implementation of the funneling system

III. Discussion

The primary objectives of this work were successfully achieved. The development and implementation of a functional color-based sorting robot, a reliable color detection algorithm, and the seamless integration between mechanical and embedded systems were all realized. It was demonstrated that it is possible to create an efficient and affordable sorting system using readily available and cost-effective components.

The high accuracy of the color detection algorithm under standard conditions validates the approach taken. The slight drop in performance under challenging lighting conditions suggests that further refinements could be made, such as implementing more advanced preprocessing techniques or adaptive thresholding to enhance robustness. The algorithm's performance indicates its suitability for industrial applications where lighting conditions can be controlled to some extent.

IV. Conclusion

This work demonstrated the design and implementation of a color-sorting machine based on a color thresholding algorithm; the design demonstrated a balance between accuracy, adaptability, and cost by using affordable computing boards and nonproprietary open-source computer vision libraries. The system successfully operated the conveyor belt and robotic arm to sort objects based on their color at a rate of 12 items per minute (one item every five seconds).

By employing a programmable solution based on image processing, this work can be adaptive to diverse industrial environments by updating the code base developed in this process. In any future effort, one could build upon this work to expand functionality to include multi-feature detection. Another area for improvement is in boosting system

throughput by employing a sorting mechanism other than a robotic arm, which would limit the adaptability of the system but increase its throughput.

In this project, the system was designed to detect and sort objects based on three colors: yellow, green, and blue. However, the same method can be improved to detect more colors by adjusting the HSV values in the code. Additionally, integrating shape recognition using contour approximation would enable the sorting of objects based not only on color but also on geometric shape (e.g., cubes, cylinders, spheres). This improvement will make the system more useful for different industrial applications that require sorting by more than one feature.

V. References

[1] Y. Jia, G. Yang and J. Saniie, "Real-time color-based sorting robotic arm system," 2017 IEEE International Conference on Electro Information Technology (EIT), Lincoln, NE, USA, 2017, pp. 354-358, doi: https://doi.org/10.1109/EIT.2017.8053385.

[2] V. D. Cong, L. D. Hanh, L. H. Phuong, and D. A. Duy, “Design and Development of Robot Arm System for Classification and Sorting Using Machine Vision,” FME Trans., vol. 50, no. 1, pp. 181–192, 2022, doi: https://doi.org/10.5937/fme2201181C.

[3] Medojevic, I., Veg, E., Joksimovic, A., & Ilic, J. (2022). Promotion of Color Sorting in Industrial Systems Using a Deep Learning Algorithm. *Applied Sciences*, *12*(24), 12817, doi: https://doi.org/10.3390/app122412817

[4] M. Haggag, S. Abdelhay, A. Mecheter, S. Gowid, F. Musharavati and S. Ghani, "An Intelligent Hybrid Experimental-Based Deep Learning Algorithm for Tomato-Sorting Controllers," in IEEE Access, vol. 7, pp. 106890-106898, 2019, doi: https://doi.org/10.1109/ACCESS.2019.2932730.

[5] M. Abdullah-Al-Noman, A. N. Eva, T. B. Yeahyea, and R. Khan, " Computer Vision-based Robotic Arm for Object Color, Shape, and Size Detection," Journal of Robotics and Control (JRC), vol. 3, no. 2, pp. 180–186, 2022, doi: https://doi.org/10.18196/jrc.v3i2.13906

[6] A. E. Elwakeel, Y. S. A. Mazrou, A. A. Tantawy, A. M. Okasha, A. H. Elmetwalli, S. Elsayed, and A. H. Makhlouf, "Designing, optimizing, and validating a low-cost, multi-purpose, automatic system-based RGB color sensor for sorting fruits," Agriculture, vol. 13, no. 9, p. 1824, 2023, doi: https://doi.org/10.3390/agriculture13091824

[7] A. R. M. Khairudin, M. H. A. Karim, A. A. Samah, D. Irwansyah, M. Y. Yakob, and N. M. Zian, "Development of colour sorting robotic arm using TCS3200 sensor," in 2021 IEEE 9th Conference on Systems, Process and Control (ICSPC), 2021, pp. 108–113, doi: https://doi.org/10.1109/ICSPC53359.2021.9689114

[8] S. Sinha, S. K. Suman, and A. Kumar, "Color Sensor-Based Object Sorting Robotic Arm," in Computing Algorithms with Applications in Engineering: Proceedings of ICCAEEE 2019, Springer, 2020, pp. 169–180, doi: https://doi.org/10.1007/978-981-15-2369-4_16

[9] S. N. Kulkarni and S. Kumar Singh, "Object Sorting Automated System using Raspberry Pi," 2018 3rd International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 2018, pp. 217-220, doi: https://doi.org/10.1109/CESYS.2018.8724056.

[10] R. R. Mhaski, P. B. Chopade and M. P. Dale, "Determination of ripeness and grading of tomato using image analysis on Raspberry Pi," *2015 Communication, Control and Intelligent Systems (CCIS)*, Mathura, India, 2015, pp. 214-220, doi: https://doi.org/10.1109/CCIntelS.2015.7437911.

[11] V. Pereira, V. A. Fernandes and J. Sequeira, "Low cost object sorting robotic arm using Raspberry Pi," 2014 IEEE Global Humanitarian Technology Conference - South Asia Satellite (GHTC-SAS), Trivandrum, India, 2014, pp. 1-6, doi: https://doi.org/10.1109/GHTC-SAS.2014.6967550.

[12] P. Yadav, M. Uikey, P. Lonkar, S. Kayande and A. Maurya, "Sorting of Objects Using Image Processing," 2020 IEEE International Conference for Innovation in Technology (INOCON), Bangluru, India, 2020, pp. 1-6, doi: https://doi.org/10.1109/INOCON50539.2020.9298360.

[13] I. B. Mustaffa and S. F. B. M. Khairul, “Identification of fruit size and maturity through fruit images using OpenCV-Python and Rasberry Pi,” Proceeding 2017 Int. Conf. Robot. Autom. Sci. ICORAS 2017, vol. 2018-March, pp. 1–3, 2018, doi: https://doi.org/10.1109/ICORAS.2017.8308068.

[14] Raspians. (n.d.). Raspberry Pi models comparison. Retrieved from https://raspians.com/raspberry-pi-models-comparison/

Integration of PLC and HMI Control with a Pneumatic Automation System

Hebah Marza[1,*]
Osama AL-Kutbi[1]
Najib Suhail[1]
Mohammed Fares[1]
Gehad A. A. Qasem[1]
Bassam G. N. Muthanna[1]

Received: 19/04/2025
Revised: 31/05/2025
Accepted: 01/06/2025

© 2025 University of Science and Technology, Aden, Yemen. This article can be distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

© 2025 جامعة العلوم والتكنولوجيا، المركز الرئيس عدن، اليمن. يمكن إعادة استخدام المادة المنشورة حسب رخصة مؤسسة المشاع الإبداعي شريطة الاستشهاد بالمؤلف والمجلة.

[1]Department of Mechatronics Engineering, Faculty of Engineering and Computing, University of Science and Technology, Aden, Yemen Email: hebahsocial@gmail.com

Integration of PLC and HMI Control with a Pneumatic Automation System

Hebah Marza
Department of Mechatronics Engineering, Faculty of Engineering and Computing, University of Science and Technology,
Aden, Yemen
hebahsocial@gmail.com

Osama AL-Kutbi
Department of Mechatronics Engineering, Faculty of Engineering and Computing, University of Science and Technology,
Aden, Yemen

Najib Suhail
Department of Mechatronics Engineering, Faculty of Engineering and Computing, University of Science and Technology,
Aden, Yemen

Mohammed Fares
Department of Mechatronics Engineering, Faculty of Engineering and Computing, University of Science and Technology,
Aden, Yemen

Gehad A. A. Qasem
Department of Mechatronics Engineering, Faculty of Engineering and Computing, University of Science and Technology,
Aden, Yemen

Bassam G. N. Muthanna
Department of Mechatronics Engineering, Faculty of Engineering and Computing, University of Science and Technology,
Aden, Yemen

***Abstract*— Pneumatic systems are extensively utilized in industrial applications due to their cleanliness, reliability, and the ease with which compressed air can be generated from the environment. These characteristics make them particularly advantageous in hygiene-sensitive industries such as food processing and packaging, where they are often favored over hydraulic alternatives. As the demand for automation continues to grow, systems like the Vertical Form Fill Seal (VFFS) machine have become essential for performing forming, filling, and sealing operations with minimal human intervention. This paper presents the experimental design and implementation of a simplified, automated pneumatic sealing system based on the VFFS concept. The system main controller is a Programmable Logic controller (PLC) and integrated with Human Machine Interface (HMI) for a real time monitoring and human interaction. The control logic evolved by, ladder diagrams, timers, sensor feedback, and strokes counter to ensure synchronization and repeated operation. The study shows off the probability of using accessible automation techniques to simulate industrial packing processes beneficially pointing up on its potential for educational and practical applications in latest manufacturing environments.**

Keywords— *Fluids control, PLC, solenoid valve, pneumatic circuits, industrial automation.*

I. INTRODUCTION

Nowadays since the Industry 3.0 and its growth into the Industry 4.0, organizations both large and small progressively focused on integrating their control systems with automated ones. This integration is leading through the need to conserve physical space, improve energy efficiency, and enhance troubleshooting capabilities [1-3]. As automation systems played an important role in modern manufacturing, especially in processes that require high repeatability. Systems like that significantly improve operational efficiency by empowering machines to operate in risky or dangerous environments that exceed human tolerance. Additionally, automation allows for dynamic adjustment of production speed based on real time demand, resulting in increased productivity. Key benefit of automation include reduce dependancy on manual labor, lessened material waste, and improved production accuracy today in the same environment. A significant portion of tasks are accomplished by machines highlighting automation's pivotal role in manufacturing.

control systems are vital component of automation systems, although the reverse is not always true, while control systems may function independently, automation systems naturally encompass a broader range of functionalities. The primary objective of control system is to ensure the output follows predefined setpoints. The core components, for typical control system are:

- **Input devices,** as sensors, which detect and monitor changes in the process.
- **Controllers,** analyze inputs and determine appropriate taken actions.
- **Output device,** as actuators, which implement those actions to achieve desired outcomes.

In contrast, by including great capabilities automation systems cover more of the basic controlled systems by many features such as computing setpoints, monitoring overall system performance, managing startup so shutdown procedures, and scheduling task and equipment usage [1,4,5].

Several studies have addressed this domain from various perspectives, including system design performance optimization and integration with control platforms, like Programmable Logic Controller (PLC). Fawzi and Salloom[6] proposed a method for designing multi actuator pneumatic circuits using Karnaugh Maps to simplify logic and integrate control via PLCs. They emphasized the advantage of pneumatic system, such as ease of setup and low maintenance, and provided detailed explanations of their components and operation. By minimizing Boolean expressions, the Karnaugh Map technique facilities efficient ON/OFF control design. A

case study involving a three station filling and capping machine demonstrated how control sequences can be derived and implemented using pneumatic components and PLCs programmed with ladder logic. Simulation tools such as FluidSIM, Automation Studio, and TIA Portal were employed to optimize automation performance. Santos and da Silva [7] similarly utilized Karnaugh Maps to simplify ON/OFF logic in pneumatic and electro-pneumatic systems, focusing on cylinder positioning and memory variable usage. Their approach supports PLC-based control implementation using Siemens STEP7 200 MicroWIN software.

Parikh et al. [8] explored cost-effective control of pneumatic systems using MATLAB Simulink and Arduino with MathWorks I/O packages. This approach simplifies control tasks through graphical tools, eliminating the need for extensive programming. However, their study is limited in scope and lacks comprehensive statistical analysis. Dobândă and Bordeașu [9] investigated the implementation of logic gates using fluid power components, presenting hydraulic equivalents for basic logic operations (AND, OR, NOT, etc.) along with corresponding truth tables. Their theoretical approach contributes to a deeper understanding of logic design in pneumatic and hydraulic automation systems. Panaitescu et al. [10] focused on the integration of magnetic proximity switches in pneumatic circuits. Through simulations using FluidSIM, they demonstrated how these sensors enhance piston position detection, improve energy efficiency, and extend component lifespan.

Papoutsidakis and Chatzopoulos [11] examined hydraulic systems used in automated production, highlighting their advantages in delivering high speed, linear motion, and significant power.

Their theoretical review compared control methods such as PID and Sliding Mode Control (SMC), noting SMC'S superior performance in resisting disturbances, They also employed the LuGre friction model to simulate friction dynamics in MATLAB/Simulink. Liu et al. [12].

Estimated three control strategies k+PWM, Bang Bang, one and a hybrid approach for regulating air pressure in pneumatic microfluidic systems, while, k+PWM improves precision with slower reactions, their composite method dynamically switches between the two based on error size, achieving a balance of speed accuracy and stability. This approach is particularly well suited for biomedical analytical microchip applications. Bang Bang control offers rapid response with higher error margins. Collectivity, these studies offer a comprehensive perspective on pneumatic control ,encompassing logic simplification, cost effective design , sensor integration ,and precision dynamics . while many contributions focus on theoretical framework or narrowly defined scenarios there is a growing trend towards hybrid systems that blend simplicity with adaptability This research aligns with that trajectory by investigating PLC controlled pneumatic systems through both simulation and practical case studies, aiming to bridge the gap between academic experimentation and real worlds industrial automation demands.

II. PROGRAMMABLE LOGIC CONTROLLERS (PLCS)

A programmable logic controller (PLC) is responsible for managing and regulating automated operation in real time. Its inherent flexibility and adaptability make it an essential component of modern industrial automation systems. Its a versatile and widely used controller in various process control applications. Positioned at the control level of the automation pyramid, as illustrated in Figure 1 the [1, 13].

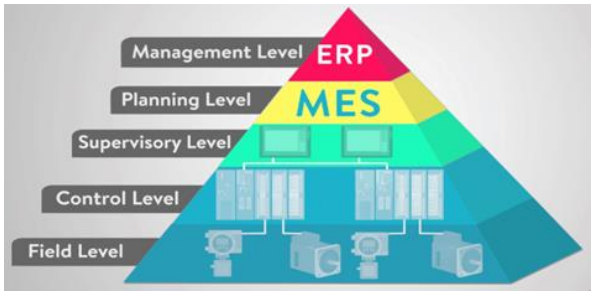


Figure 1: Automation pyramid is a pictorial example of the different levels of automation in a factory [13].

The word "programmable" give an emphasis to the one of the most powerful features in PLCs: its ability to be configured by end user to perform virtually any control function. Primely, it's introduced as electronic replacements for electronical relay systems, were designed to control starting and stopping of electric motors and other discrete output devices. In comparison, to traditional rely based systems, PLCs offer enhanced reliability, reduced system failures, and extended operational lifespans [1].

III. CONTROL OF PNEUMATIC CIRCUITS

pneumatic control has experienced a significant achievement over the past few decades. Primarily, driven by the growing demand for efficient energy saving systems in industrial processes. Actuators used in industrial automation are typically classified into three main types: electrical, hydraulic , and the pneumatic drive systems. Electrical drive systems are generally gifted for applications requiring high precision and flexible control, typically involving relativity low loads. In contrast, hydraulic systems are commonly employed in frameworks where handling heavy loads is essential. Pneumatic drive systems, on other hand, are more suitable for basic two position tasks where cost effectiveness and simple programming are key configuration[14].

Compressed air is one of the earliest forms of energy utilized by humans . Over time, its use and distribution have become widespread, making it a fundamental element in both manufacturing and service sectors. Historical records suggest that compressed air was first utilized around 250 BC. Its ability to replace manual labor has allowed industries to enhance productivity while ensuring consistent and uniform output. Moreover, compressed air can be effectively integrated with hydraulic, mechanical, electrical, and electronic systems, facilitating the simultaneous and efficient operation of various processes. The controlled use of this energy source is known as pneumatics, derived from the Greek word "pneuma," meaning "to blow" or "to breathe."

Pneumatic systems are used in a wide range of operations, such as drilling, riveting, deburring, and filling and capping, making them ideal for process automation. These systems can operate using purely pneumatic methods or can be integrated with electrical, electronic (such as PLC-based control), or hydraulic systems, either as a power source, a means of control, or both [6]. Many of the design principles applied to pneumatic systems are also relevant to hydraulic systems, with key differences depending on the nature of the fluid used [14].

IV. METHODS AND MATERIALS

This paper follows the steps outlined in Figure 2 and employs a combination of literature review, simulation, and case study to explore the control of pneumatic systems using PLCs. The research is divided into two phases:

1. **Theoretical Analysis:** A review of relevant studies focusing on the integration of PLCs with pneumatic components.
2. **Practical Case Study:** Using simulation tools to demonstrate a sample control system.

The study was conducted at the University of Science and Technology, Aden, Yemen, in the Mechatronics Lab. The hardware components used in the study are detailed in Table 1.

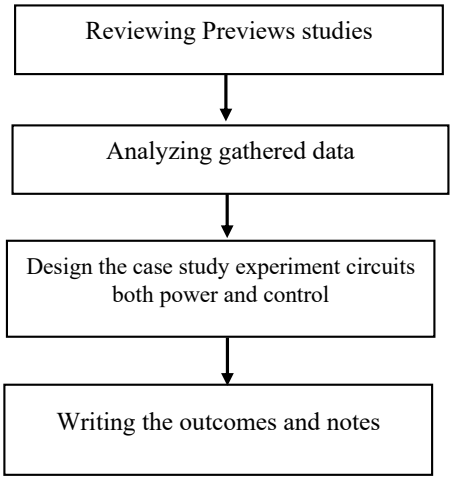


Figure 2: The paper followed steps

Table. 1 Hardware components

Name of the Hardware	Quantity	Specifications	Role in the Project
Cylinder	2	Single acting	Carry out a movement using compressed air [15]
Flow control valve	1	Mechanical valve	Control the compressed air flow rate.so, The speed of the cylinder can be controlled [16]
Single solenoid valve	1	5\2-way single 24V DC	Control the flow a form of on/off valve used for pneumatic and hydraulic systems [1]
Optical proximity sensor	1	3 wires proximity sensor	Measuring the cylinder strokes
FRL unit with pressure gauge	1	Withstand temperatures up to 50˚C[17]	Filters air, lubrication and measuring pressure
Compressor	1	Reciprocating air compressor ,2 bar working pressure	Ambient air suction and pressure measurement
PLC S7-1200,	1	CPU1215C DC\DC\DC	Controlling the solenoid valve
HMI screen	1	Siemens SIMATIC OP7	Graphical user interface
Power supply	1	220V\24V DC	Powering system

The software programs used in this study included FluidSIM-P (version 2010) for designing pneumatic circuits and TIA Portal V17 for creating the PLC ladder diagram code.

V. EXPERIMENTAL SETUP

Packaging is a critical stage in any production system, serving various functions such as kitting, wrapping, and filling products into containers. Among the most commonly used machines in the packaging industry is the Vertical Form Fill Seal (VFFS) machine, which efficiently packages products into bags or pouches. This study focuses on the pneumatic sealing unit of such a machine, as used for sealing products. The process is illustrated in Figure 3 [18,19].

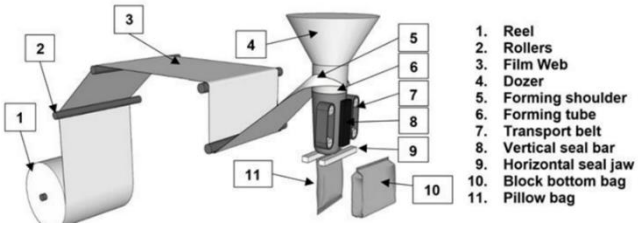


Figure 3: Schematic diagram of VFFS machine [19].

In the laboratory setup, two pneumatic cylinders are positioned face to face, with their shafts applying force to actuate the sealing mechanism, as shown in Figure 4. The primary control element is a 5/2 solenoid valve, which manages the extension and retraction of the cylinders. A modern PLC-based control system, integrated with a Human Machine Interface (HMI) screen, was implemented to enhance user interaction, monitoring, and system flexibility.

FluidSIM is a comprehensive software tool used for modeling, simulating, teaching, and researching electro-pneumatic, electro-hydraulic, pneumatic, digital, and electronic control systems. It offers an intuitive environment that combines diverse features and media sources, making it particularly useful for education and research. FluidSIM was used in this study to simulate the pneumatic circuit and control design, as shown in Figure 5 [6].

Implementing the system with a Programmable Logic Controller (PLC) significantly simplified construction, control, and monitoring. The control logic was developed using a ladder diagram, designed to operate the system in a cyclic manner. When the user presses the "Start" button on the HMI, Timer_0 is activated with a 5-second on-delay to energize the solenoid valve. At the same time, Timer_1 is triggered with the same delay, which in turn reactivates Timer_0, establishing a continuous operational loop.

An optical sensor is employed to count the strokes of one of the two synchronized cylinders, which are mechanically linked. The HMI screen displays the number of strokes counted and provides additional control buttons for "Reset" and "Stop" functions. The complete ladder diagram, implemented using TIA Portal V17, is shown in Figure 6.
Finally, the experimental setup was constructed using pneumatic components and simulated according to the designed schematic. The complete system is shown in Figure 7.

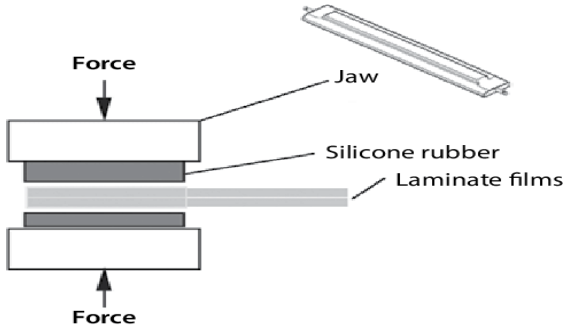


Figure 4: The pneumatic system is the force needed [20].

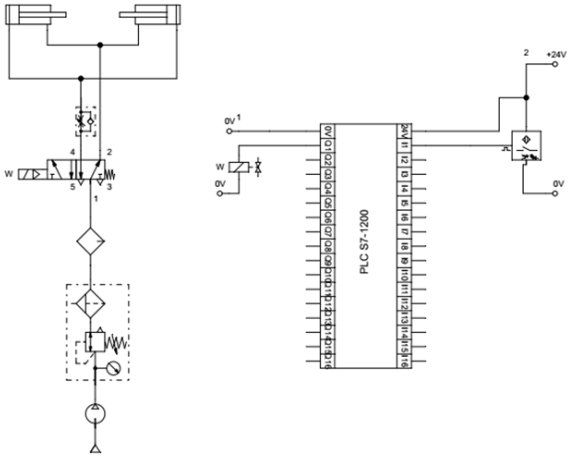


Figure 5: The schematic of the pneumatic control

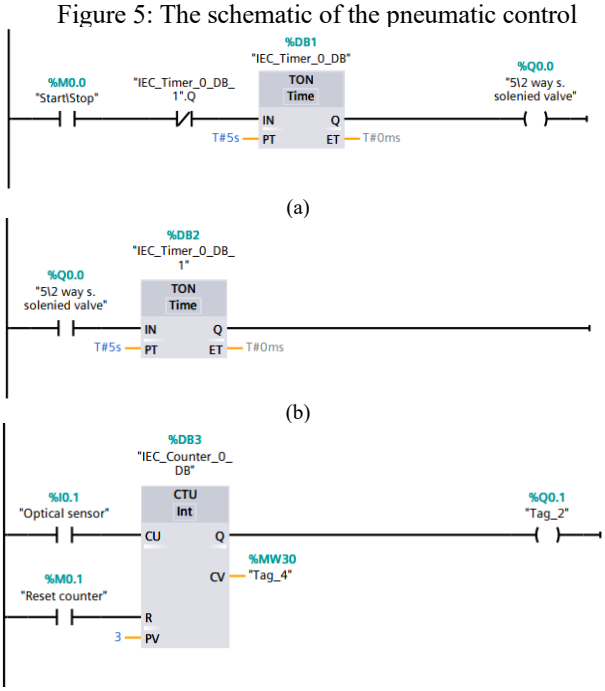


Figure 6: The ladder diagram, (a) Network 1, (b) Network 2, and (c) Network 3.

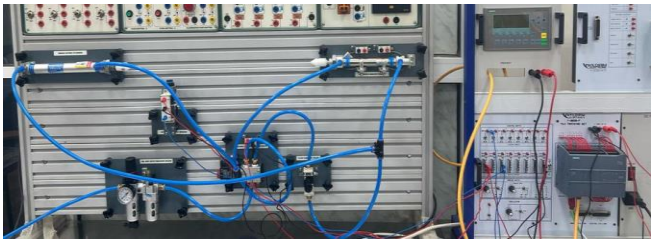


Figure 7: Experimental setup.

VI. Results and Discussions

The study demonstrated that implementing a PLC-based control system effectively automated the pneumatic sealing process. The ladder diagram facilitated a continuous operation cycle initiated via the HMI screen, eliminating the need for manual resets. Two timers managed the solenoid valve in a loop: Timer_0 triggered valve actuation after a 5-second delay, while Timer_1 reactivated Timer_0, ensuring continuous cylinder movement.

The face-to-face pneumatic cylinders operated in synchronization due to a mechanical linkage, enabling the sealing task to be performed with consistent force and motion. The sealing speed could be adjusted using a flow control valve by turning the adjustment screw in a counter-clockwise direction, as illustrated in Figure 8.

An optical sensor was used to count the piston strokes, confirming synchronization and enabling accurate real-time monitoring. Simulations conducted in FluidSIM closely matched the behavior of the actual system, validating the reliability and accuracy of the design. Overall, the setup could manifests that it is efficient, user friendly, and suits real world industrial packaging applications.

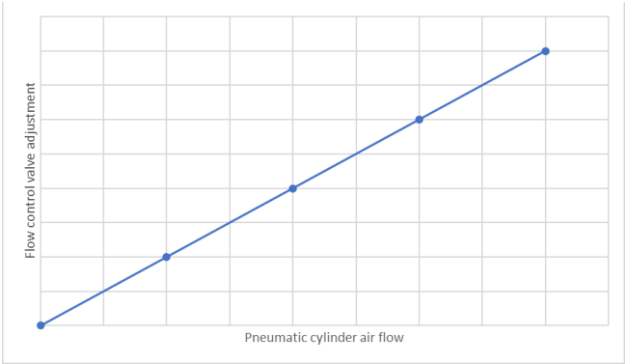


Figure 8: Proportional relation between air flow and adjusting the valve clock wise

VII. Conclusion

This study give a demonstration of the PLC controlled pneumatic sealing system effectiveness in achieving repeatable and synchronized operations in an automated packing process. The integration of timers and sensor provides reliable feedback and control, while Human Machine Interface (HMI) significantly enhance user interaction and system accessibility. reimplementation simulations conducted in FluidSIM confirmed the accuracy and feasibility of the design, ensuring confidence in the system's performance prior to physical deployment.

in spite of constrains in available materials, the experimental setup was successfully constructed and validated, underscoring the practicality and robustness of the proposed approach. The results confirm that such systems are not only efficient and user-friendly but also highly applicable to real-world industrial packaging environments.

For future advancements, it is recommended to integrate industrial communication protocols and Internet of Things (IoT) capabilities to enable real-time remote monitoring and diagnostics. Furthermore, the use of proportional control valves in place of mechanical flow regulators is advised to enhance precision, control flexibility, and adaptability to varying operational conditions, and future work will consider conducting more comprehensive practical experiments beyond simulation to further validate the system's performance in real industrial settings.

VIII. References

[1] Kuphaldt, T. R. (2022). Lessons In Industrial Instrumentation. https://nvkarta.com/project/library/uploads/engineering/kuphaldt/(en)_version_3.01_lessons_In_Industrial_Instrumentation.pdf

[2] "What is Industry 4.0? | IBM." Accessed: Apr. 06, 2025. [Online]. Available: https://www.ibm.com/think/topics/industry-4-0

[3] "The Rise of Industrial IoT: How PLCs Are Connected in the Age of Industry 4.0." Accessed: Apr. 06, 2025. [Online]. Available: https://balajiswitchgears.com/the-rise-of-industrial-iot-how-plcs-are-connected-in-the-age-of-industry-4-0/

[4] Fernandes, S., & Borkar, S. (2016). Development of PLC Based Industrial Automation with System Monitoring. In 4International conference on science, technology and management-16 (pp. 421-425).

[5] Pallavi, S. (2013). Lesson Introduction to Industrial Automation and Control. pp. 1–18, 2013.

[6] Fawzi, A. N., & Salloom, M. Y. (2023). Developing Multiple-Actuator Pneumatic Circuits Using the Karnaugh Maps Designing PLC Controlled. J. Eur. des Systèmes Autom, 56(2), 245-252. https://doi.org/10.18280/jesa.560209

[7] Santos, A. A., & Silva, A. F. D. (2017). Methodology for manipulation of Karnaugh maps designing for pneumatic sequential logic circuits. International Journal of Mechatronics and Automation, 6(1), 46-54. https://doi.org/10.1504/IJMA.2017.093307

[8] Parikh, P., Vasani, R., Sheth, S., & Gohil, J. (2016). Actuation of electro-pneumatic system using MATLAB simulink and arduino controller-A case of a mechatronics systems lab. In International Conference on Communication and Signal Processing 2016 (ICCASP 2016) (pp. 59-64). Atlantis Press. https://doi.org/10.2991/iccasp-16.2017.10

[9] E. Dobândă, E. & Bordeaşu, I. (2017). Using Logic Gates in Fluid Power Systems Control (Part 1). Proceedings of 2017 International Conference on Hydraulics and Pneumatics – HERVEX, pp. 42–47. https://fluidas.ro/hervex/proceedings2017/pp.42-47.pdf

[10] Panaitescu, M. Cojocaru, R. Axinte, T. Frăţilă, C. Damian, R. & Diaconu, M. (2021). Study of Electro-Pneumatic Circuits with Magnetic Proximity Switches. Hidraulica, 3, pp. 34–42.

[11] Papoutsidakis, M., & Chatzopoulos, A. (2018). Control Methods of Hydraulic Positioning Systems and a Promising Simulation Friction Model. Journal of Modern Technology & Engineering, 3(3). https://jomardpublishing.com/UploadFiles/Files/journals/JTME/V3N3/PapoutsidakisM.pdf

[12] Liu, X., Yang, X., Zuo, W., Bai, Y., Zhou, H., Wang, J., ... & Shang, Z. (2022). Research of control method for pneumatic control of pneumatic microchips. SLAS technology, 27(5), 290-301. https://doi.org/10.1016/j.slast.2022.06.002

[13] "What is the Automation Pyramid? - RealPars." Accessed: Jan. 25, 2025. [Online]. Available: https://www.realpars.com/blog/automation-pyramid

[14] Parker, J.K. and Schinstock, D., 1996, June. Introduction To Hydraulic & Logic Systems In A Controls Course. In 1996 Annual Conference (pp. 1-288).

[15] "Buy Pneumatic cylinders online | Festo GB." Accessed: Apr. 11, 2025. [Online]. Available: https://www.festo.com/gb/en/c/products/actuators/pneumatic-cylinders-id_pim135/

[16] "What is a flow control valve? | Festo GB." Accessed: Apr. 11, 2025. [Online]. Available: https://www.festo.com/gb/en/e/about-festo/blog/in-practice/what-is-a-flow-control-valve-id_1516985/

[17] "Fr + L Sets 2pc With Pressure Gauge - Airmax Pneumatics." Accessed: Apr. 11, 2025. [Online]. Available: https://www.airmaxindia.com/fr-l-sets-2pc-with-pressure-gauge/

[18] Sachin, M. & Rajendra, S. (2021). Efficiency Optimization of Continuous Motion Vertical Form Fill and Seal Machine. pp. 1118–1121, 2021. https://www.irjet.net/archives/V8/i8/IRJET-V8I8155.pdf

[19] Merabtene, M., Tanninen, P., Varis, J. & Leminen, V. (2022). Heat sealing evaluation and runnability issues of flexible paper materials in a vertical form fill seal packaging machine. BioResources, 17(1), pp. 223-242. https://doi.org/10.15376/biores.17.1.223-242

[20] Su, J. F. (2017). Biodegradable soy protein isolate/poly (vinyl alcohol) packaging films. Handbook of Composites from Renewable Materials, 30, 587. https://doi.org/10.1002/9781119441632.ch103.

Improving Detection and Classification Of Brain Tumors Using DenseNet201

Nashwan Amin Al-khulaidi [1, *]
Mohammed Hashem Almourish [1,2]
Emad Mahmood Ahmed Taleb [1,3]
Rasheed Abdullah Ali Farhan [1]
Mohammed Nabil Mohammed Latf [1]
Omar Mukhtar Ahmed Abdullah [1]

Received: 11/03/2025
Revised: 20/05/2025
Accepted: 21/05/2025

© 2025 University of Science and Technology, Aden, Yemen. This article can be distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

© 2025 جامعة العلوم والتكنولوجيا، المركز الرئيس عدن، اليمن. يمكن إعادة استخدام المادة المنشورة حسب رخصة مؤسسة المشاع الإبداعي شريطة الاستشهاد بالمؤلف والمجلة.

[1] Faculty of Engineering and Information Technology. Taiz University. Taiz, Yemen.
[2] Department of Computer Science, Isra University, Amman, Jordan.
[3] Faculty of Engineering and Computer Science, University of Science & Technology, Aden, Yemen.

https://doi.org/10.20428/jst.v30i6.2855

Improving Detection and Classification Of Brain Tumors Using DenseNet201

Nashwan Amin Al-khulaidi
Faculty of Engineering and Information Technology. Taiz University.
Taiz, Yemen
nashwan1email@gmail.com

Mohammed Hashem Almourish
Department of Computer Science, Isra University, Amman, Jordan
Amman, Jordan

Emad Mahmood Ahmed Taleb
Faculty of Engineering and Computer Science, University of Science & Technology,
Aden, Yemen
dr.temad@gmail.com

Rasheed Abdullah Ali Farhan
Faculty of Engineering and Information Technology. Taiz University.
Taiz, Yemen

Mohammed Nabil Mohammed Latf
Faculty of Engineering and Information Technology. Taiz University.
Taiz, Yemen

Omar Mukhtar Ahmed Abdullah
Faculty of Engineering and Information Technology. Taiz University.
Taiz, Yemen

Abstract— Brain tumors are among the leading causes of death worldwide. A brain tumor may originate in the brain or develop elsewhere in the body and metastasize to the brain, leading to secondary brain tumors. Thus, brain tumors can take many different forms. In this study, brain tumors were detected and classified based on magnetic resonance imaging (MRI) involving three different types of brain tumors and non-tumors. VGG-16, VGG-19 (Visual Geometry Group), DenseNet201 (Densely Connected Convolutional Networks), Inceptionv3, ResNet-50 (Residual Network with 50 layers), and EfficientNet-B0 were among the convolutional neural network (CNN) models that were employed and analyzed in order to determine the optimal model for detecting and classifying brain cancers. The best model was DenseNet201, which achieved accuracy, precision, and recall of 99.31%, 99.31%, and 99.25%, respectively.

Keywords— Brain Tumors , Magnetic resonance imaging, VGG-19, Densenet201, Inceptionv3, Resnet-50, EfficientNet-B0.Introduction

I. INTRODUCTION

One of the most difficult cancers to identify and treat is a brain tumor. They vary in size, shape, and aggression and can arise in various brain regions. Accurate brain tumor categorization and early diagnosis are essential for successful therapy and better patient outcomes. Traditional methods of detecting and diagnosing brain tumors, such as biopsy and histological analysis, can be invasive, time-consuming, and may have limited accuracy [1-2]. As a result, there is increasing interest in using machine learning (ML) methods for medical imaging-based brain tumor diagnosis and classification. Resolving diagnostic problems is greatly aided by the medical imaging technique used to diagnose brain tumors [3, 4, 5]. MRI and computed tomography (CT) are useful techniques for revealing crucial details regarding the existence and extent of aberrant brain tissue. Since MRI uses safe radiation to produce detailed images of the brain's internal structures, it is preferred over CT scans, which use dangerous radiation [6, 7]. The ability of machine learning (ML) systems to identify brain tumors early from MR images has been shown in a number of studies [8-10]. One of the most popular machine learning methods for classifying brain tumors is deep learning (DL) [10]. CNNs are among the most widely utilized deep learning algorithms, particularly in applications related to medical imaging [12-15].

CNNs can help in the automatic analysis of big MRI image collections, enabling precise classification and early brain tumor detection. Through the analysis of extensive datasets of MRI images, they can also spot distinctive features and patterns that differentiate between various kinds of brain cancers.

The six CNN architecture models used in this study (VGG-16, VGG-19, DenseNet201, InceptionV3, ResNet-50, and EfficientNet-B0) were trained to detect and classify three distinct kinds of brain tumors as well as MRI-obtained non-tumor pictures (glioma, meningioma, and pituitary).

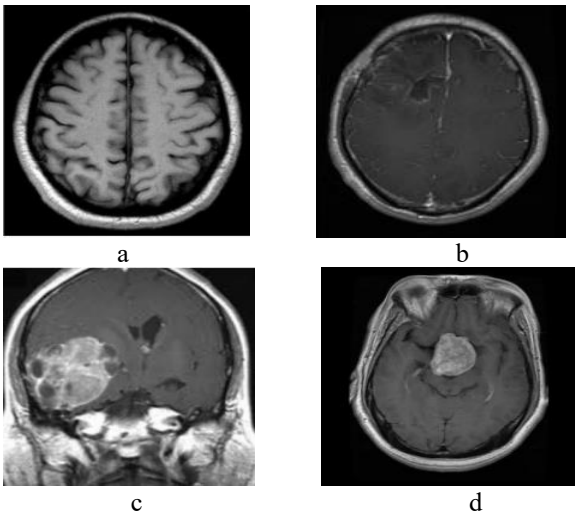


Figure 1: Non-tumorous and three types of brain tumors: a) Normal, b) Glioma, c) Meningioma, d) Pituitary

II. Literature Review

Recently, many techniques are being studied, and research into identifying and classifying brain tumors using MRI scans has progressed significantly. In this section, previous studies on the detection and classification of brain tumors will be reviewed, with a focus on those that employed DL techniques.

Francisco Javier Díaz-Pernas et al. [16] developed a multiscale CNN-based brain tumor segmentation and classification model that obtained a 97.3% accuracy rate for three distinct tumor types: pituitary, glioma, and meningioma. In Ref. [17], presented a DL technique that extracted features from brain MR images using a GoogLeNet model. They had a 98% accuracy rate for the same three kinds of brain tumors. A. Priya and V. Vasudevan [18] used a hybrid model involving AlexNet and GRU neural networks to identify and classify brain cancers from MRI data with a 97% accuracy rate. The BTC-f CNN model was proposed in Ref. [19] to classify three types of brain cancers (pituitary tumor, meningioma, and glioma) with a 98.86% accuracy rate.

In Ref. [20], ResNet (2+1)D and ResNet mixed convolution models were used to classify different types of brain tumors, with ResNet mixed convolution reaching a test accuracy of 96.98%.

Wen Jun & Zheng Liyuan [21] used a dataset of 3064 MR images to develop a brain tumor classification model that incorporates a multipath network and attention mechanism, achieving an accuracy of 98.61%. MuSunil Kumar & Dilip Kumar [22] classified three different kinds of brain cancers (glioma, meningioma, and adenoma) using a CNN model with the Adam optimizer had the best accuracy, 86.23%.

Ramin Ranjbarzadeh et al. [23] used CNN to detect cancer within a specific region of an image and achieved different accuracy during the image processing process with an average Dice score of 92% for the whole tumor, 91.1% for tumor enhancement, and 87.2% for tumor core segmentation. Ramdas Vankduthoo and Mohamed Abdel Hamid [24] suggested using CNN to identify various types of cancerous and non-cancerous brain tumors and the accuracy of this classification reached 95.17%.

Kakarla J, Isunuri BV, et al. [25] suggested an eight-layer CNN brain tumor classifier that had a 97.42% accuracy rate. Irsheidat, S., and Duwairi [26] utilized a CNN model to predict brain cancers from MR images, and achieved an accuracy of 96.7% on the evaluation data.

Ali Ari and Davut Hanbay [27] proposed a three-stage approach consisting of preprocessing, tumor classification, and tumor region extraction and obtained an accuracy of 97.18%.

Md Ishtyaq Mahmud, et al. [28] utilized CNN network to classify brain cancers and the model achieved an accuracy of 93.3%.

Obeidavi et al. [29] used the BRATS 2015 MRI dataset to create a CNN-based residual network for brain tumor detection. Their model is reportedly 97.05% accurate.

Hareem Kibriya et al. [30] suggested a CNN for classifying brain tumors based on a dataset of 3064 MRI images. where the accuracy of the study was 97.2%.

Vinayak K. Bairagi et al. [31] suggested CNN architecture to automatically detect brain tumors based on MR images. Several CNN architectures were explored and 98.67% accuracy was achieved using CNN AlexNet.

III. Methodology

As illustrated in Figure 2, the methodology utilized in this paper comprised five steps: data collection, preprocessing, dividing the dataset into training and test sets, model training, and model evaluation.

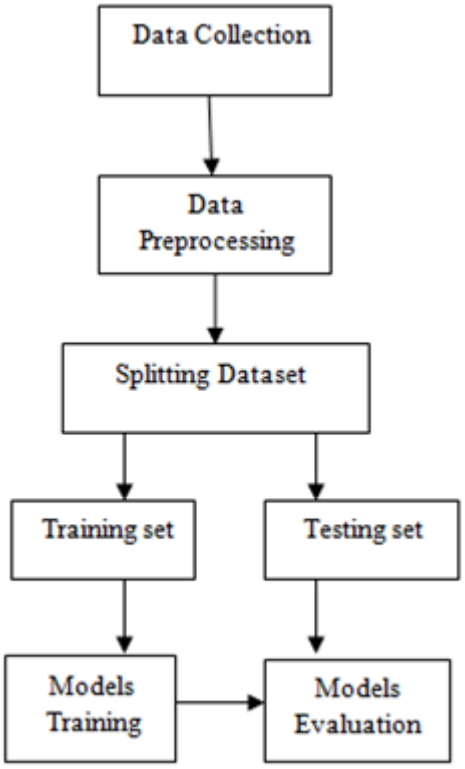


Figure 2 Steps of the proposed model for brain tumors detection and classification.

A. Data Collection

This step involved the collection of 7,023 MRI brain images, including 1,757 images of pituitary tumors, 1,645 photos of meningiomas, 1,621 images of gliomas, and 2,000 images of non-tumorous instances (Figure 3).

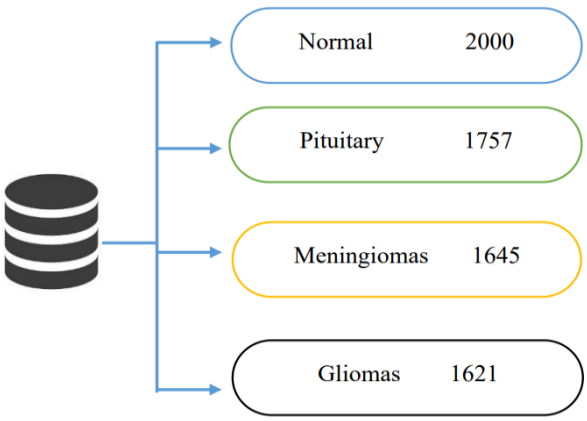


Figure 3: Collected Dataset

https://doi.org/10.20428/jst.v30i6.2855

B. Data Pre-processing

The collected MRI scans must be preprocessed to remove any noise and other image distortions that could potentially impact the accuracy of the CNN algorithms.

C. Splitting Dataset

As illustrated in Figure 4, the research dataset was split into two sections: 80% (60% for training, 20% for validation) and 20% for testing.

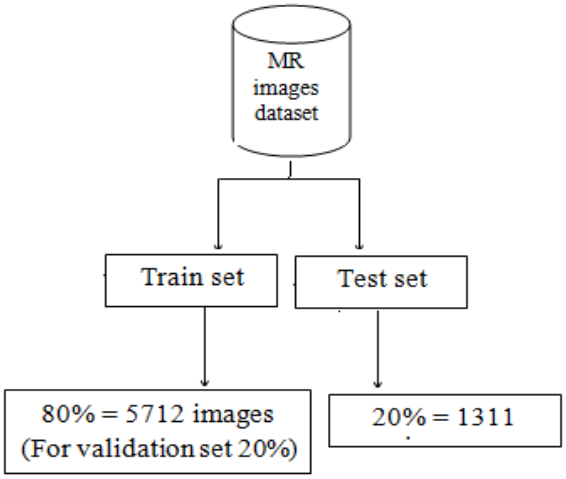


Figure 4: Dataset splitting

D. Models Training:

CNN models have been developed and compared for the purpose of detecting and classifying brain tumors in order to choose the most accurate model. As seen in Figure 5, convolutional, pooling, and fully linked layers are commonly seen in CNN architectural designs.

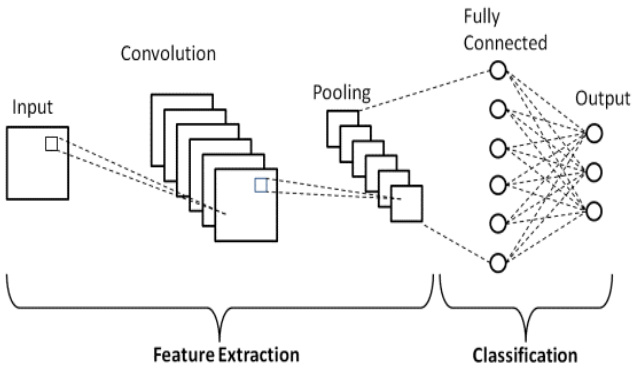


Figure 5: CNN Architecture

Six CNN models, including ResNet-50, VGG-16, VGG-19, DenseNet-201, InceptionV3, and EfficientNet-B0, were utilized to determine the best model. A brief explanation of these models is provided below:

ResNet50

The ResNet-50 model has 50 levels of depth. By avoiding some intermediary levels, the skip connection creates a residual block that links a layer's activations to later layers (Figure 6).

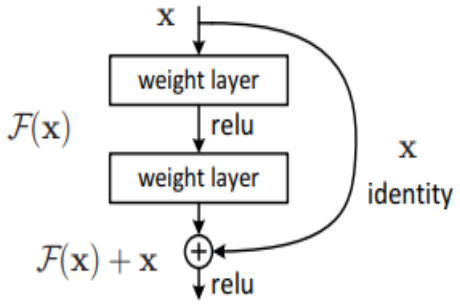


Figure 6: Residual learning: a building block.

VGG

VGG is a traditional CNN design that was created using an examination of growing network depth. The VGG-16 model has sixteen layers, comprising thirteen convolutional layers and three fully connected layers.

However, the VGG-19 model has a 19-layer network topology. However, the VGG19 model has a 19-layer network topology with 16 convolutional layers and 3 fully linked layers. In comparison to VGG-16, VGG-19 contains three extra convolutional layers.

DenseNet201

A CNN type known as a DenseNet makes use of dense blocks to create dense connections between layers, which let each layer access feature information from all layers that came before it. Every layer passes on its own feature maps to every layer after, it while also receiving new inputs from all layers before it in order to preserve the feed-forward nature.

InceptionV3

InceptionV3 is a CNN architecture that consists of convolutions, fully linked layers, concatenations, dropouts, average pooling, and max pooling. To compute the loss, Softmax is employed.

The EfficientNet B0

EfficientNet is a type of CNNs. The fundamental concept underlying Efficient Net is a novel scaling strategy that uses sophisticated coefficient scaling to scale models in a straightforward yet efficient manner while uniformly measuring all depth, width, and resolution variables.

E. Models Evaluation

The testing sets are used to assess the performance, accuracy, and generalization capacities of the trained models. After that, the model with the best performance is chosen for deployment.

IV. Result and Discussion

Results from the application of the suggested model are shown in this section. Three metrics (accuracy, recall, and precision) are used to illustrate the performance outcomes of the suggested models in Figure 7, where the performance of

the DenseNet-201 model was shown with the highest accuracy of 99.31%.

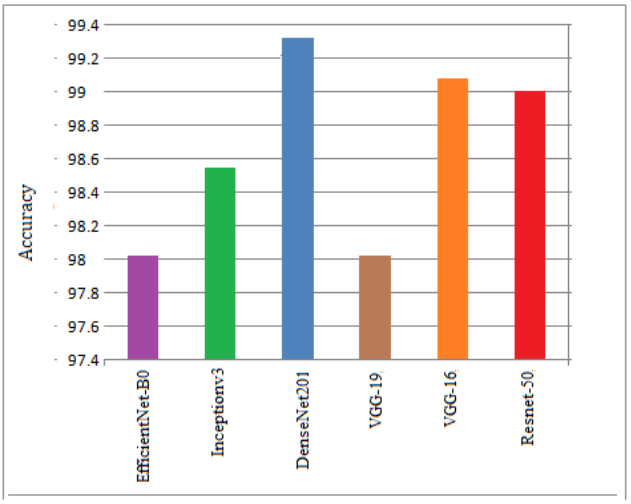


Figure 7: The accuracy of each model.

The confusion matrices for the DenseNet-201 model and the other models are displayed in Figure (8) and Figure (9), respectively.

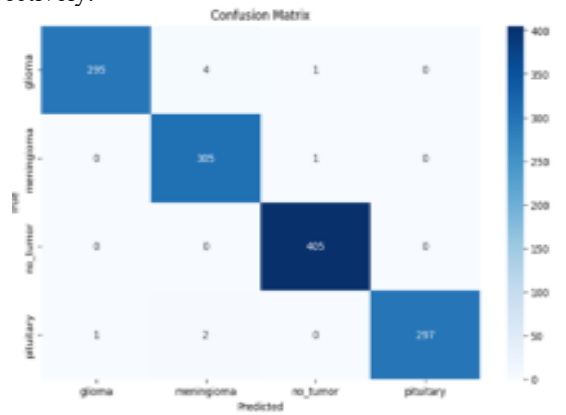
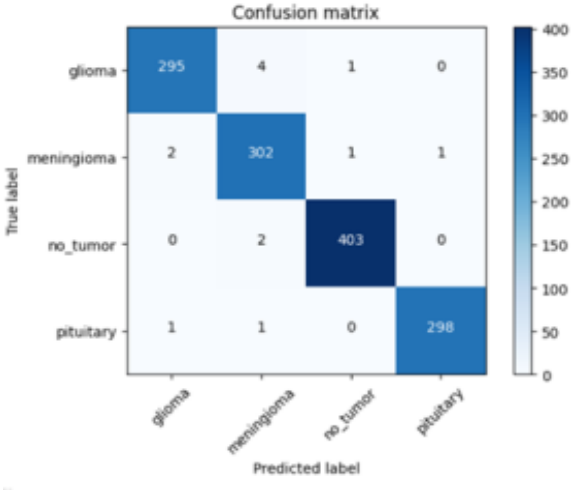
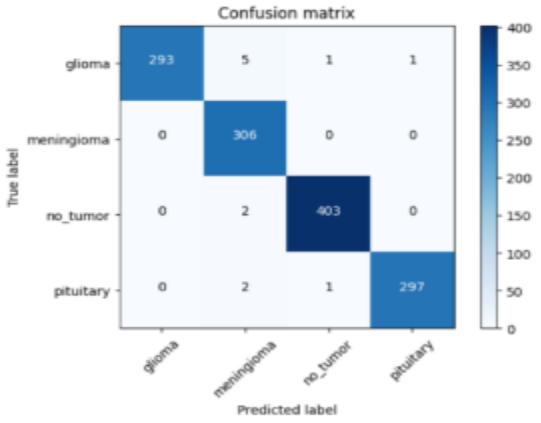


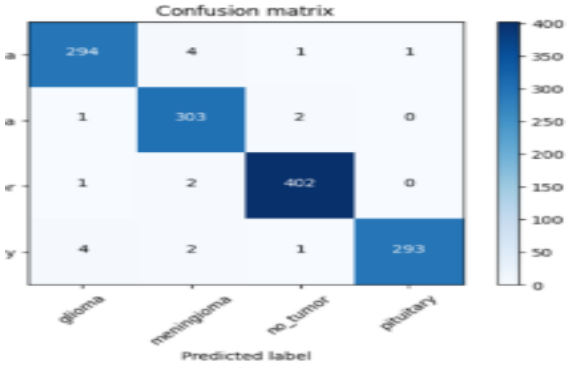
Figure 8: Confusion matrix of the DenseNet201 model.



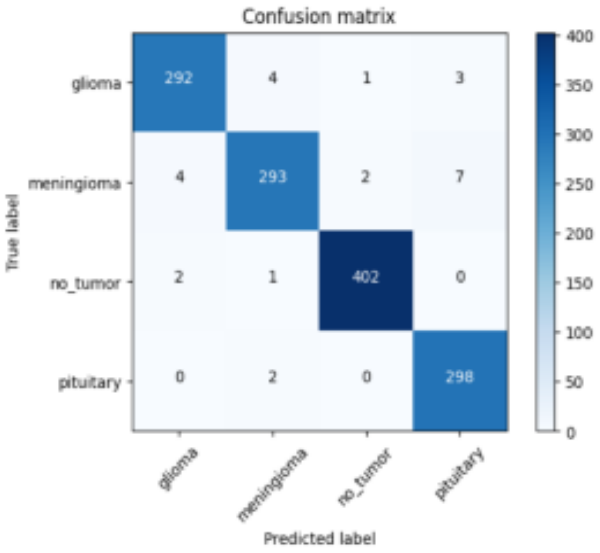
(a)



(b)

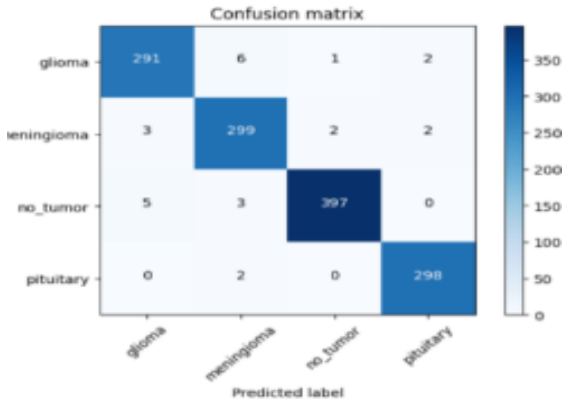


c)



d)

https://doi.org/10.20428/jst.v30i6.2855



e)

Figure 9: Confusion Matrix
a) ResNet50, b) VGG16, c) InceptionV3, d) Efficient_B0
e) VGG19

Table 1 shows the average precision and recall of the DenseNet-201 model, which are 99.31% and 99.25%, respectively.

Table 1. The performance of DenseNet201 model.

Category	Precision (%)	Recall (%)
Non-tumorous	99.78	99.79
Glioma	99.71	98.19
Meningioma	98.14	99.91
Pituitary	99.62	99.12
Average	**99.31**	**99.25**

The DenseNet201 model's accuracy and loss curves are displayed in Figures 10-11. The lowest error rate, which happened in epoch 11, was 0.038.

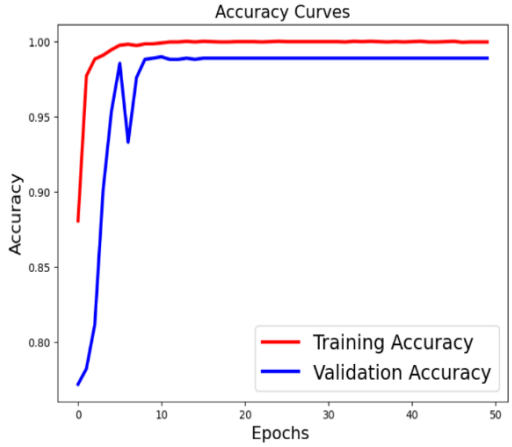


Figure 10: Accuracy Curves

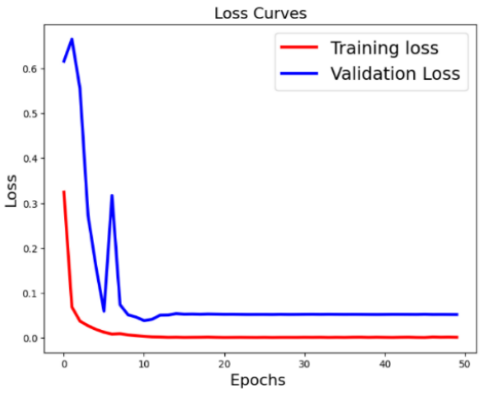


Figure 11: Loss Curves

Our DenseNet201-based method achieves a classification accuracy of 99.31%, surpassing previous approaches such as multiscale CNN (97.3% [16]), GoogLeNet (98% [17]), AlexNet-GRU hybrid (97% [18]), and BTC-fCNN (98.86% [19]), as detailed in Table 2. This superior performance can be attributed to DenseNet201's architectural features: dense connectivity that enhances feature reuse and mitigates vanishing gradients, more effective parameter utilization, and improved learning of complex hierarchical tumor features through deep supervision. Clinically, the accuracy enhancement of approximately 1.31–2.31% has the potential to reduce diagnostic errors in tumor identification and classification. Methodologically, the robustness of our results is supported by a thorough evaluation across six different CNN architectures, validation on a larger and more diverse dataset comprising 7,023 images from four classes, and detailed performance reporting, including high precision (99.31%) and recall (99.25%) metrics.

Table 2. Performance comparison with previous studies

Study	Method	Accuracy
[16]	Multiscale CNN	97.3%
[17]	GoogLeNet	98%
[18]	AlexNet-GRU	97%
[19]	BTC-fCNN	98.86%
Our work	**DenseNet201**	**99.31%**

V. Conclusion

Early detection and classification are essential for the successful treatment of brain cancers.
The most advanced networks for identifying and categorizing brain tumors are CNNs, which offer higher accuracy.

This study utilizes CNN models with six different architectures (VGG-16, VGG-19, DenseNet-201, InceptionV3, ResNet-50, and EfficientNet-B0) to classify brain tumors from MRI images.

The dataset comprises 7023 MRI brain images, including 2000 non-tumorous images, 1757 pituitary images, 1645 meningioma images, and 1621 glioma images.

While these results are promising, we acknowledge the need for future multi-center validation studies to assess generalizability across different MRI protocols and patient populations.

According to this study, the DenseNet-201 model performed the best, showing accuracy, precision, and recall of 99.31%, 99.31%, and 99.25%, respectively.

References

[1] V. Kumar, A. K. Abbas, and J. C. Aster, *Robbins and Cotran Pathologic Basis of Disease*, 9th ed. Elsevier, 2014.

[2] E. R. Gerstner and T. T. Batchelor, "Primary central nervous system lymphoma," *Arch. Neurol.*, vol. 67, no. 3, pp. 291–297, 2010.

[3] B. B. Vimala *et al*., "Detection and classification of brain tumor using hybrid deep learning models," *Sci. Rep.*, vol. 13, no. 1, p. 6425, 2023.

[4] N. V. Shree and T. N. R. Kumar, "Identification and classification of brain tumor MRI images with feature extraction using DWT and probabilistic neural network," *Brain Inform.*, vol. 5, no. 2, pp. 23–30, 2018.

[5] M. W. Nadeem *et al*., "Brain tumor analysis empowered with deep learning: A review, taxonomy, and future challenges," *Brain Sci.*, vol. 10, no. 2, p. 118, 2020.

[6] N. Abiwinanda, "Brain tumor classification using convolutional neural network," in *IFMBE Proc. World Congress on Med. Phys. Biomed. Eng.*, vol. 67, no. 1, pp. 183–189, 2018.

[7] R. Ranjbarzadeh *et al*., "Brain tumor segmentation of MRI images: A comprehensive review on the application of artificial intelligence tools," *Comput. Biol. Med.*, vol. 152, p. 106405, 2023.

[8] A. B. Ramakrishnan *et al*., "Optimizing brain tumor classification with hybrid CNN architecture: Balancing accuracy and efficiency through oneAPI optimization," *Inform. Med. Unlocked*, vol. 44, p. 101423, 2024.

[9] M. Z. Khaliki and M. S. Başarslan, "Brain tumor detection from images and comparison with transfer learning methods and 3-layer CNN," *Sci. Rep.*, vol. 14, p. 2664, 2024.

[10] N. A. Al-khulaidi *et al*., "Recognition of brain tumors using Radon transform and gray level co-occurrence matrix algorithm," *Al-Mansour J.*, vol. 53, no. 3, pp. 45–56, 2021.

[11] B. Srikanth and S. V. Suryanarayana, "Multi-class classification of brain tumor images using data augmentation with deep neural network," *Mater. Today: Proc.*, vol. 45, Part 2, pp. 1234–1240, 2021.

[12] M. H. Almourish *et al*., "Detection of pneumonia by using nine pre-trained transfer learning models based on deep learning techniques," *Iraqi J. Comput. Inform.*, vol. 47, no. 1, pp. 22–31, 2021.

[13] N. A. Al-khulaidi, M. H. Almourish, and A. M. Ali, "Detection of Covid-19 and chest pneumonia based on X-ray images using deep-transfer learning," *Iraqi J. Comput. Inform.*, vol. 50, no. 1, pp. 15–25, 2024.

[14] G. Yao, T. Lei, and J. Zhong, "A review of convolutional-neural-network-based action recognition," *Pattern Recognit. Lett.*, vol. 118, pp. 14–22, 2019.

[15] A. Dhillon and G. K. Verma, "Convolutional neural network: A review of models, methodologies and applications to object detection," *Prog. Artif. Intell.*, vol. 9, no. 2, pp. 85–112, 2020.

[16] F. J. Díaz-Pernas *et al*., "A deep learning approach for brain tumor classification and segmentation using a multiscale convolutional neural network," *Healthcare*, vol. 9, no. 2, p. 153, 2021.

[17] S. Deepak and P. M. Ameer, "Brain tumor classification using deep CNN features via transfer learning," *Comput. Biol. Med.*, vol. 111, p. 103345, 2019.

[18] A. Priya and V. Vasudevan, "Brain tumor classification and detection via hybrid AlexNet-GRU based on deep learning," *Biomed. Signal Process. Control*, vol. 89, p. 105710, 2024.

[19] B. S. A. El-Wahab *et al*., "BTC-fCNN: Fast convolution neural network for multi-class brain tumor classification," *Health Inf. Sci. Syst.*, vol. 11, no. 1, p. 12, 2023.

[20] S. Chatterjee *et al*., "Classification of brain tumours in MR images using deep spatiospatial models," *Sci. Rep.*, vol. 12, no. 1, p. 9025, 2022.

[21] W. Jun and Z. Liyuan, "Brain tumor classification based on attention guided deep learning model," *Int. J. Comput. Intell. Syst.*, vol.15, no. 35, pp. 1–12, 2022.

[22] M. Kumar and D. Kumar, "Human brain tumor classification and segmentation using CNN," *Multimed. Tools Appl.*, vol. 82, pp. 7599–7620, 2023.

[23] R. Ranjbarzadeh *et al*., "Brain tumor segmentation based on deep learning and an attention mechanism using MRI multi-modalities brain images," *Sci. Rep.*, vol. 11, p. 10930, 2021.

[24] R. Vankdothu and M. A. Hameed, "Brain tumor MRI images identification and classification based on the recurrent convolutional neural network," *Measurement: Sensors*, vol. 24, p. 100534, 2022.

[25] J. Kakarla *et al*., "Three-class classification of brain magnetic resonance images using average-pooling

https://doi.org/10.20428/jst.v30i6.2855

convolutional neural network," *Int. J. Imaging Syst. Technol.*, vol. 31, no. 4, pp. 2047–2060, 2021.

[26] S. Irsheidat and R. Duwairi, "Brain tumor detection using artificial convolutional neural networks," in *Proc. 2020 11th Int. Conf. Inf. Commun. Syst. (ICICS)*, pp. 324–329, 2020.

[27] A. Ari and D. Hanbay, "Deep learning based brain tumor classification and detection system," *Turk. J. Electr. Eng. Comput. Sci.*, vol. 26, no. 5, pp. 2275–2286, 2018.

[28] M. I. Mahmud, M. Mamun, and A. Abdelgawad, "A deep analysis of brain tumor detection from MR images using deep learning networks," *Algorithms*, vol. 16, no. 1, p. 45, 2023.

[29] M. R. Obeidavi and K. Maghooli, "Tumor detection in brain MRI using residual convolutional neural networks," in *Proc. 2022 IEEE Int. Conf. Mach. Vis. Image Process. (MVIP)*, pp. 1–6, 2022.

[30] H. Kibriya *et al*., "Multiclass classification of brain tumors using a novel CNN architecture," *Multimed. Tools Appl.*, vol. 81, pp. 29847–29863, 2022.

[31] V. K. Bairagi *et al*., "Automatic brain tumor detection using CNN transfer learning approach," *Med. Biol. Eng. Comput.*, vol. 61, no. 4, pp. 1431–1445, 2023.