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

[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.

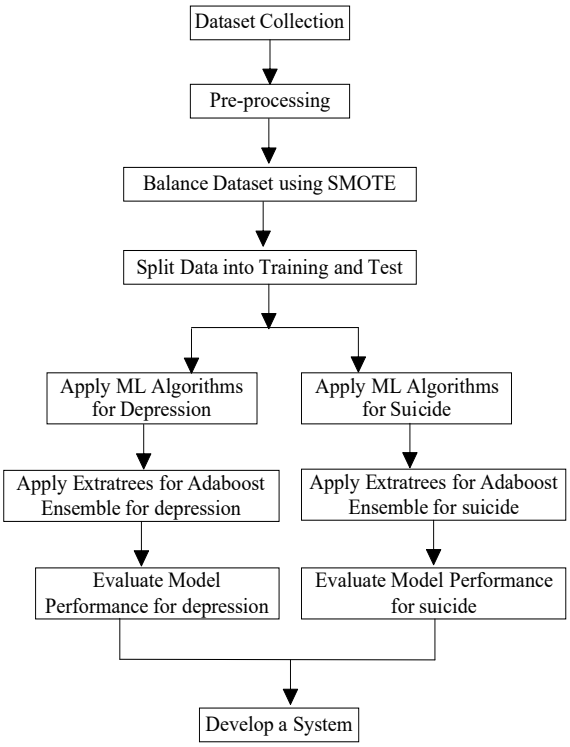

Fig. 1 Architecture of the Model

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

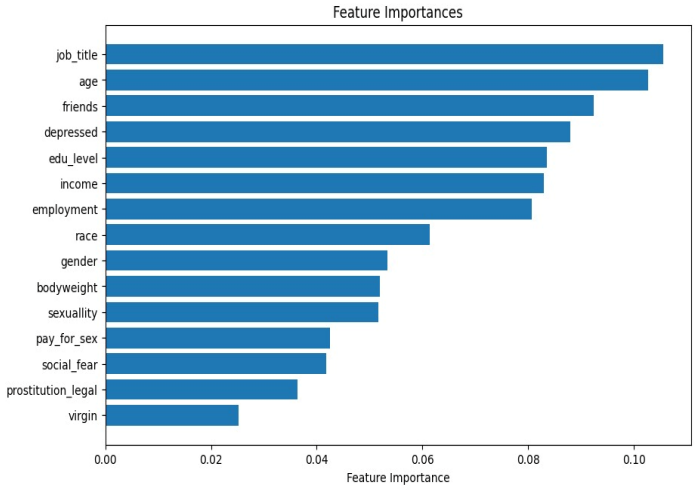

Figure 2: Suicide attempt Feature Importance Chart

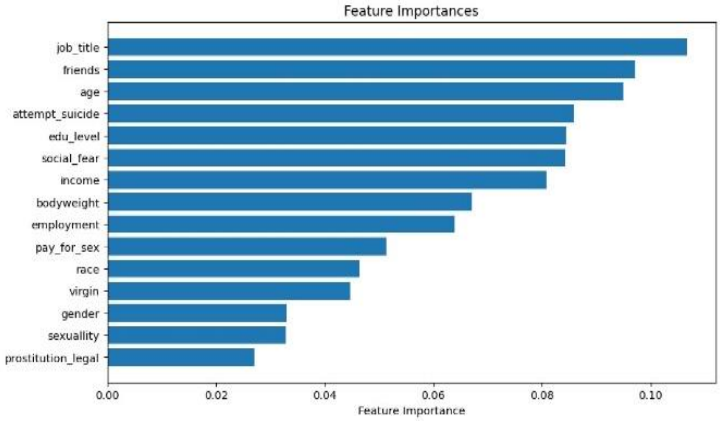

Figure 3: Depression Feature Important Chart

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".

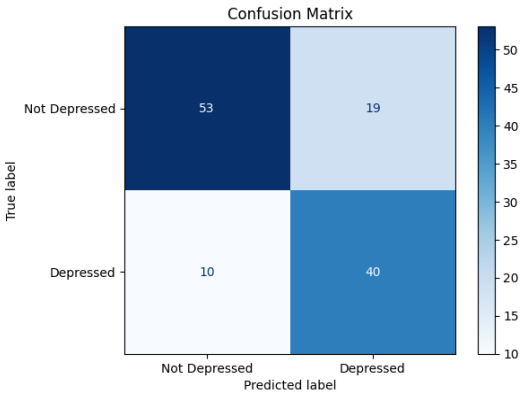

Fig. 4 Confusion Matrix obtained from Extra Trees Classifier for Depression

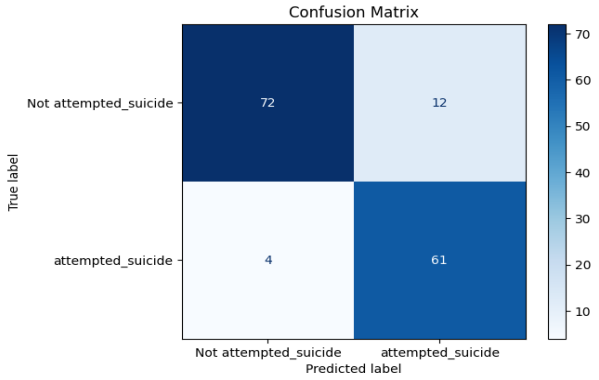

Fig. 5 Confusion Matrix obtained from Extra Trees for Suicide attempt.

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.

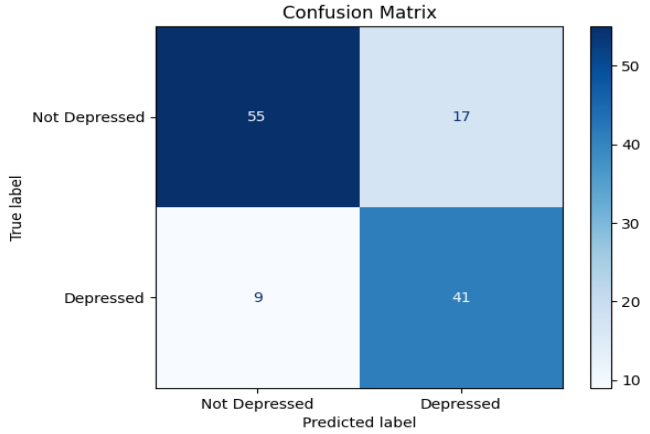

Fig. 6 Confusion matrix obtained from the ensemble model for depression.

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.

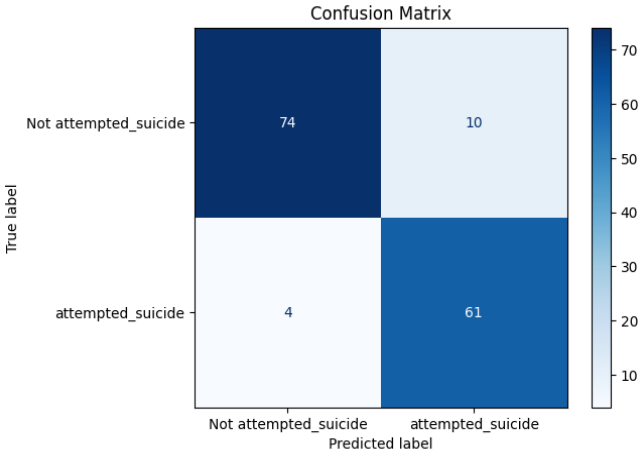

Fig. 7 Ensemble model for suicide attempt

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.

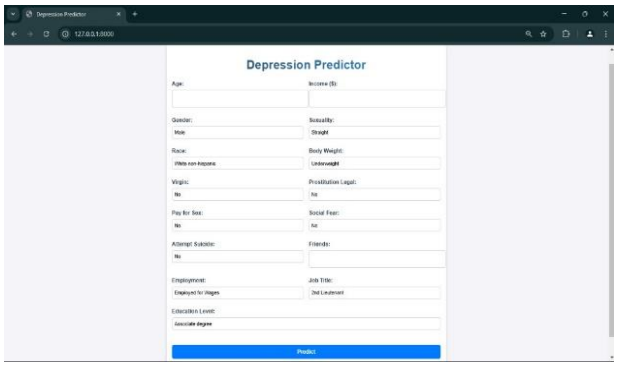

Fig. 8 Description of the Home page

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.

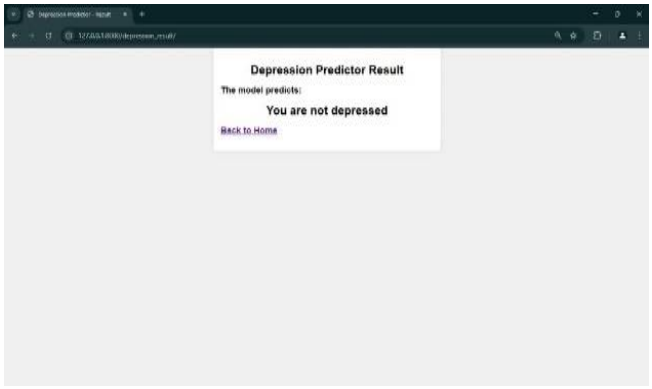

Fig. 9 Result Page: "Not Depressed"

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.

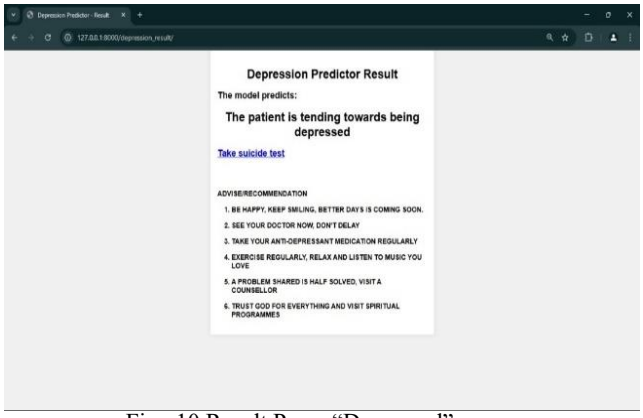

Fig. 10 Result Page: "Depressed".

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.

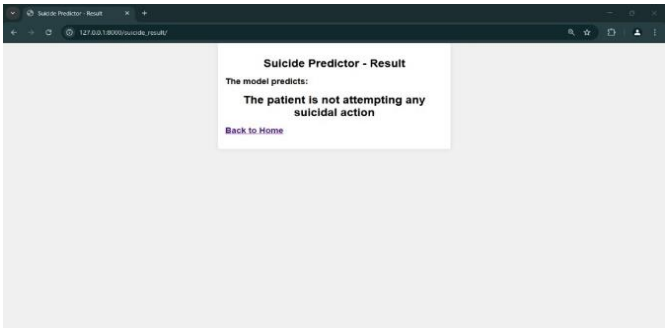

Fig. 11 Suicide Predictor Result Page: No suicidal attempt.

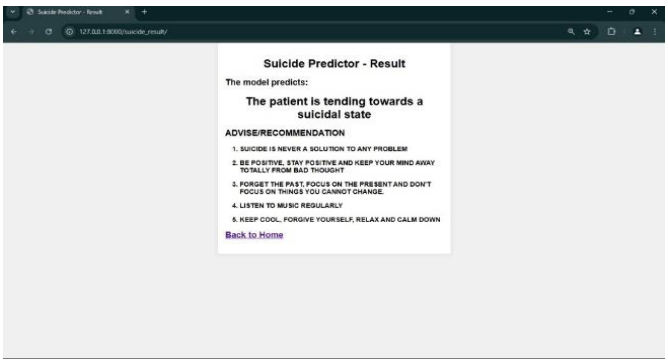

Fig. 12 Recommendation/Advise page.

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.