Healthcraft Frontiers
https://www.acadlore.com/journals/HF

Unlocking Minds: An Adaptive Machine Learning Approach for Early Detection of Depression

Hafiz Burhan Ul Haq[1*], Muhammad Nauman Irshad[1], Muhammad Daniyal Baig[2]

[1] Faculty of Engineering-Department of Electronics and Telecommunication Engineering, King Mongkut's University of Technology Thonburi (KMUTT), 10140 Bangkok, Thailand
[2] Department of Cyber and Applied Computing, University of Gloucestershire, GL50 2RH Gloucestershire, United Kingdom

[*] Correspondence: Hafiz Burhan Ul Haq (hafiz.burh@kmutt.ac.th)

Received: 03-29-2024 **Revised:** 06-12-2024 **Accepted:** 06-23-2024

Citation: Haq, H. B. U., Irshad, M. N., & Baig, M. D. (2024). Unlocking minds: An adaptive machine learning approach for early detection of depression. *Healthcraft Front.*, *2*(2), 107-117. https://doi.org/10.56578/hf020205.

Abstract: Depression, a prevalent and severe medical condition, significantly impairs emotional well-being, cognitive functions, and behavior, often leading to substantial challenges in daily functioning and, in severe cases, an increased risk of suicide. Affecting approximately 264 million individuals worldwide across diverse age groups, depression necessitates effective and timely detection for intervention. In primary healthcare, the Patient Health Questionnaire-9 (PHQ-9) serves as a crucial tool for screening depression. This study leverages the PHQ-9 dataset, comprising 12 features and 534 samples, to evaluate depression levels using advanced machine learning (ML) techniques. A comparative analysis of the Support Vector Classifier (SVC) and AdaBoost Classifier (ABC) was conducted to determine their efficacy in classifying depression severity on a scale from 0 to 4. The SVC emerged as the superior model, achieving an accuracy of 94%. This research contributes to the early detection and prevention of depression by proposing an interactive interface designed to enhance user engagement. Future work will focus on expanding the dataset to improve model generalization and robustness, thereby facilitating more accurate and widespread applications in clinical settings.

Keywords: Machine learning; Depression detection; Psychology; Mental health; Patient health questionnaire-9; Ensemble learning

1. Introduction

Depression, often known as mental illness, is a serious illness that affects a person's emotions, thoughts, and behavior on a daily basis. Depressed individuals may have decreased interest in daily activities, feelings of guilt or low self-esteem, disrupted sleep or meals, exhaustion, difficulty focusing, and physical issues. Globally, mental illness is the primary cause of chronic illness, according to Dey et al. (2022). The WHO estimates that 3.8% of people worldwide suffer from depression, with adults accounting for 5.0% and people over 60 for 5.7%. Over 280 million individuals worldwide suffer from depression (Dey et al., 2022).

ML is widely used in academia for its ability to learn from large amounts of data and provide valuable insights (Gupta et al., 2022). In this study, the PHQ-9 dataset was used, which was collected from Kaggle. This dataset contains five different classes of depression, which are normal/happy, mild depression, moderate depression, moderately severe depression, and severe depression. As this is the classification problem, ML classifiers are used to train models that detect these levels and produce generalized results. It is a robust method compared to the manual calculation of the results of questionnaires based on multiple-choice questions (MCQs).

ML classifiers, i.e., SVC, logistic regression classifier, and ABC, were used to train the model. In the comparative analysis of these classifiers, the SVC performed well on unseen sample data and produced generalized results. These models were evaluated by accuracy, precision, recall, and F1-score.

Early detection and prevention of depression play an important role in treatment. Depression is the primary factor behind about two-thirds of suicides that occur each year (Jagtap et al., 2021). Previous studies did not provide an interface for the user to detect their depression level using these questionnaires. Some inventions have been used to detect depression, but they require patients to open cameras or use their voices. This makes patients feel uncomfortable, as their mental health is already disturbed. It is necessary to offer a user-interactive platform to the patients for their self-monitoring using MCQ questionnaires without making them feel uncomfortable. This application is helpful for depressed patients because they are not comfortable using cameras or voice for detection.

1.1 Research Background and Motivation

Major depressive disorder is among the most destructive illnesses affecting individuals and society and is present in more than 280 million individuals globally. It hampers normal functioning and serves as the leading precursor to suicides, accounting for two-thirds of them. This point presents it as one of the causes of the occurrence of chronic diseases; therefore, it is crucial to create effective diagnosis and treatment strategies.

Early detection and intervention are required for several reasons. Proper stepped-care intervention in depression and suicidal thoughts leads to a reduction in the severity of childhood symptoms. Traditional methods often have a certain degree of prejudice and consume a lot of time. Therefore, there is a need for advanced methodologies like ML.

1.2 Previous Techniques and Their Flaws

Even though there have been great improvements in various techniques for depression diagnosis, most of the current techniques have the following shortcomings:

a) Invasive techniques: The use of cameras or voice recorders may result in embarrassment for patients with many health complications, mostly psychological issues.
b) Manual processing: If manual methods are used, it would be very tiresome to give out and mark large questionnaires and clinical assessments. Hence, this may not be efficient for carrying out large screenings.
c) Lack of user-friendly platforms: Almost all of the currently available mobile application forms for depression analysis are rather heuristic and do not adequately align with the clinical requirements for accurate analysis and depression identification.

1.3 Motivation for This Study

Based on these challenges, the ML techniques were used in this study to develop a robust, easy-to-navigate system for early detection of depression. Using Kaggle's PHQ-9 dataset, this study aims to determine diverse levels of depression with the applied ML classifiers, including SVC, Logistic Regression Classifier, and ABC. From the classifiers, the SVC presents high values of the performance parameters in the sample of unseen data and guarantees high-level accuracy of the obtained results, which have been received in perspective for generalization.

This study has the following key contributions:

a) High-accuracy classification: The study shows how ML has the potential to accurately diagnose depression after attaining 94% accuracy in the classification of depression levels on the PHQ-9 dataset.
b) User-friendly interface: Responding to the issues related to the invasiveness of the approach, the study presents a concept of an easy-to-use interface that enables a person to manage his or her depression indicators by completing a set of questions.
c) Feature engineering and data preparation: This study also underlines how, due to their characteristics, the dataset and the corresponding features should be cleaned, normalized/standardized, and engineered to improve the models and their stability.
d) Early diagnosis and prevention: In this respect, the goal of the study is to help identify early signs of depression, prevent the disease from getting worse, and minimize the danger of suicidal thoughts, thereby enhancing the quality of people's psychological well-being.
e) Scalable and accessible solutions: The continuous enhancement of the model along with the growing dataset pave the way to accessible and large-scale options for mental health assessment and effective methods for intervention.

Apart from its core contributions, the rest of this study offers a thorough examination of relevant literature, providing insights into current approaches and their advantages and disadvantages. The study places its methodology into a larger framework of depression detection research through a thorough literature review, emphasizing a variety of approaches, from natural language processing (NLP) analysis of social media postings to the creation of mobile applications for early identification. Moreover, this study provides a detailed description of the suggested approach, which includes data collection, preprocessing, feature engineering, classification, and model assessment. The meticulous explanation of every stage emphasizes how crucial methodological rigor is to producing accurate and significant findings. The study provides readers with a roadmap for reproducing and expanding the research in future studies by outlining a precise and organized approach. With all things being considered, the thorough review of the literature and the methodological approach highlights the validity and importance of the research findings, establishing "unlocking minds" as a significant addition to the area of mental health intervention and depression detection.

2. Literature Review

Several studies have adopted ML and NLP strategies to identify depression on social media and other feeds. Hence, each study presents different methods and results for the comprehension and identification of depression.

Facebook posts of 1778 users were employed in the study accomplished by Dey et al. (2022) that sought to identify the subjects' depression rate using NLP and ML algorithms. The process incorporated data gathering, preparation, scrubbing, as well as the assessment of the interest features. Out of all the applied depression classification models, the Random Forest (RF) achieved the highest accuracy value of 0.74, which is an F1-score of 0.60. The PHQ-9 dataset was employed to measure the level of depression, showing the drawbacks of the tool in identifying depressive status in different populations and regions.

Stephen et al. (2022) proposed an Android application for early identification and prevention of depression based on the PHQ-9 depression set. Although the application for mobile devices has logical viability, the presented app is to some extent deficient and has weaknesses. For example, the Java implementation binds the application to particular platforms, the user interface lacks creativity and design, and the intelligent performance provided by artificial intelligence (AI) solutions is missing. Kim et al. (2021) mainly discussed the suicidal assessment with PHQ-9 and analyzed the college students' 8760 feasible data by comparing the outcomes of PHQ-2, PHQ-8, PHQ-9, and PHQ-10 and utilizing different ML algorithms. The RF model achieved 94% accuracy, but the study is limited to physician use because of the absence of end-user applications.

In a study by Glavin et al. (2022), although PHQ-2, PHQ-4, PHQ-8, and PHQ-9 all have datasets for identifying depression, the PHQ-9 obtained higher results. The results align with other studies showing that an automatic operation of the patient health questionnaire can enhance its effectiveness and efficiency by using ML algorithms. However, additional PHQ pairing tests with several populations are required. Méndez et al. (2021) suggested a three-segmented model to carry out pre-diagnosis procedures for depression employing the PHQ-9 questionnaire in smart homes along with Alexa and HMI with fuzzy logic distribution systems. Restricted by the firm's focus on household customers, camera-based activity monitors were complemented, which is pioneering.

Multi-score data on depression was then proposed by Gupta et al. (2022), who introduced a cluster-based federated learning model. Specifically, the client data obtained from various hospitals was partitioned into similar sets with the same HRV dataset. The Convolutional Neural Network (CNN) stood out with the highest accuracy of the four classes used in the study. Mohammad & Siddiqui (2021) conducted a brief study about the detection of the depression level using three datasets, namely the MHI-5, BDI, and PHQ-9, and RF classifiers and regressors. Through the application of Optuna optimization, an accuracy of 82% was achieved for the MHI-5 dataset, with a classification accuracy of 61% observed, while the classification of patient groups based on their PPS yielded 83% accuracy. The PHQ-9 evaluations, however, only reached an accuracy of 33%. It is noteworthy that their study did not extend to implementation for end-users.

Chiong et al. (2021) proposed a social media text classification on sadness using various single classifiers and ensemble classifiers. Depression was diagnosed only in labeled datasets. Jagtap et al. (2021) used ensemble learning techniques and NLP to identify sadness in messages on social media. This approach is an impressive success even with irrelevant testing datasets, which underlines the solidity of the authors' work. Nadeem et al. (2022) discussed the deep learning (DL) models such as CNN, Gated Recurrent Unit (GRU) and ML techniques for depression detection through the Twitter data. SCCL, the proposed framework, consists of a combination of Long Short-Term Memory (LSTM), CNN, GRU, and self-attention, which helps enhance the accuracy along with the F1-score. Zainudin et al. (2021) obtained Internet of Things (IoT) sensor data (ECG and GSR) for stress detection to create stress pattern data and classify data with several ML techniques. From these findings, it was evident that the decision tree classifier outperformed the other models evaluated.

Walambe et al. (2021) developed a multimodal AI-supported system for monitoring work habits and stress levels based on the collected data from sensors. In this regard, the method saves a lot of time while successfully indicating the existence of workload-related stress, contributing to the analysis of behavioral patterns that cause mental stress. Jyothirmy et al. (2023) also presented a review of stress detection models, highlighting that the Support Vector Machine (SVM) provides high accuracy. Nazeer et al. (2024) proposed a non-invasive stress detection gadget based on GSR sweat sensors and Extreme Gradient Boosting (XG-Boost) algorithms and recommended optimization for the wearable gadgets. Ding et al. (2023) proposed a sensible and relevant approach, and successfully obtained high accuracy when implementing RF and Gradient Boosting Machine (GBM) algorithms.

A multimodal dataset, WESAD, was employed by Garg et al. (2021) to detect stress. Different ML algorithms were tested and high accuracy was achieved with the help of the RF model. Manimeghalai et al. (2022) used ECG data for stress estimation and classification with an integration of ML techniques. Short-term ECG data was used more to determine the crucial role of stress in the correct diagnosis (Haq et al., 2023). Ecker et al. (2024) talked about the progress in psychotherapy in relation to memory reconsolidation neuroscience in the context of variegated alterations in managing several mental health disorders. Sun et al. (2023) discussed the difficulties in monitoring major depressive disorder through passive digital phenotypes obtained from wearables and smartphones, considering objective monitoring to improve treatments. Thabit (2024) worked on emotion detection. After analyzing the transformer-based models, promising results were obtained to categorize emotional data from the Twitter datasets. The research also explored the ethical issues concerning emotion analysis.

Table 1. A brief overview of existing methodologies

Author(s)	Approach	Dataset	Strength	Limitations
Dey et al. (2022)	ML with NLP	PHQ-9 with social media posts	The RF depression classification model achieved the highest accuracy of 0.74 with a 0.60 F1-score	The data collected for the PHQ9 questionnaire was limited to detecting depression among various populations and was limited to two areas of a country
Stephen et al. (2022)	Java application	PHQ-9	Manual calculation of the result	No ML model for robust performance
Kim et al. (2021)	K-nearest neighbors, linear discriminant analysis and RF	PHQ-9	94% accuracy of RF	Limited to clinics
Glavin et al. (2022)	Comparative analysis	PHQ-9, PHQ-2, PHQ-4, and PHQ-8	Comparison of PHQ-9, PHQ-2, PHQ-4, and PHQ-8	Limited sample of population
Méndez et al. (2021)	A three-step framework using fuzzy logic in smart homes	PHQ-9	Pre-diagnosis of depression using Alexa	Limited to house holders
Gupta et al. (2022)	Cluster-based federated learning model based on multi-score data	HRV dataset	CNN achieved the highest level of accuracy	
Mohammad & Siddiqui (2021)	RF regressor and RF classifiers	MHI-5, BDI and PHQ-9	MHI-5 was evaluated using Bayesian with an accuracy of 95.83%, PHQ-9 with an accuracy of 82.61%, and BDI using Bayesian obtained an accuracy of 83.33%	Not available to end users
Chiong et al. (2021)	ML single and ensemble classifier	Two public, labeled Twitter datasets and three non-Twitter depression-class only datasets	Using a ML classifier to predict depression effectively	Limited for labeled dataset
Nadeem et al. (2022)	ML and DL models include LSTM, CNN, and GRU	Twitter depression datasets	F1-score of 94.4% and an accuracy of 97.4% and 82.9%, respectively, for binary and ternary labeled data	Used an imbalanced dataset

Bhanot et al. (2023) said that COVID-19 is the most emergent economic crisis affecting the global economy, and proposed the use of ML techniques to identify the strategies for reviving the affected global economy with the help of the MII program. Schaub (2023) focused on the effectiveness of a conscious and subconscious state of a particular individual in the process of self-improvement, providing methods of eradicating certain prejudices in order to select a goal and reach it. Sleem (2022) discussed about disease detection using machine intelligence mechanism. Mandour (2023) addressed neutrosophic issue and discussed its application. Alenizi & Alrashdi (2023) discussed blockchain based mechanism existed in healthcare for preventing ransomware attacks. Bhathena (2021) wrote on mindfulness and meditation that self-empowerment covers serenity and clear-headedness. Wani et al. (2022) suggested an approach for the syndication of depression on social media via CNN and LSTM models that involves using some hybrid feature-based signals. Lorenzoni et al. (2024) investigated the application of ML and NLP methods in diagnosing depression and dealing with the problems regarding feature extraction, data preprocessing, and model identification. Korti & Kanakaraddi (2022) suggested an algorithm to detect depression from the data collected from the Twitter platform, highlighting the user tweet to understand the depression state. The Burns Depression Checklist is a self-report measure. Sharif et al. (2022) used ML algorithms (SVM, decision tree, and light gradient boosting) for depression detection and found out that the proposed decision tree classifier has a very high accuracy.

Squires et al. (2023) underscored that in the context of mental health, AI can be useful in identifying disorders and individualizing interventions. However, model validation and enhancement of patients' lives remain a significant concern. El-Latif et al. (2023) proposed a lightweight model for recognizing Alzheimer's disease with good accuracy with fewer layers compared with previous models. Ilias et al. (2023) suggested applying transformer-based models, i.e., Bidirectional Encoder Representations from Transformers (BERT) and MentalBERT, for stress and sadness identification on social media and discussed the enhancements with the linguistic factors. Munthuli et al. (2023) assessed XLM-RoBERTa concerning the classification of depression from the transcriptions of the speech responses, with high accuracy, completing the key terms related to classification.

Altogether, these studies reveal that diverse methods have been used for screening depression and stress with ML and NLP. However, many of the methods have inherent challenges, such as limited datasets, limited practical applications at the end-user level, and limitations in using AI to achieve optimum performance. To fill these gaps, this study seeks to design an integrated and exhaustive model for the early detection of depression with the help of AI and modern approaches to collecting data. The above-mentioned approaches are briefly described in Table 1.

3. Proposed Methodology

This study area involves the data file, which contains different datasets, the annotation of the classifiers, and the training process of the classifier. Out of all the methods, data preparation and model training and evaluation are the most crucial.

This study comprises three main modules: the input stage, the modeling stage, and the output stage of the training process of the ML. The approach followed entails three consecutive steps to provide a well-organized procedure to subsequently develop and evaluate classification models intended for depression-level recognition through the PHQ-9 dataset. As mentioned earlier, the flow of the model is depicted in Figure 1.

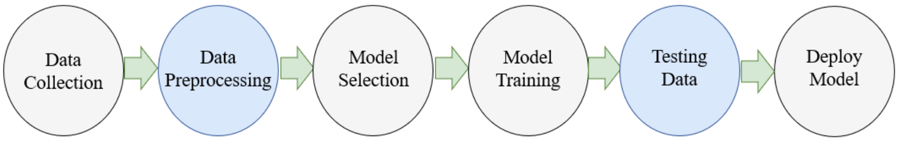

Figure 1. Process of the model

3.1 Data Preparation

3.1.1 Data collection

Data collection is part of the process which takes place during acquaintance with the defined problem. In this study, the PHQ-9 dataset from Kaggle was used, which is an open-source home to various datasets. This dataset was applied to identify the degree of depression in patients and make relevant classifications. It is a structured dataset that classifies depression levels (0-4) and comprises various fields, including scales (PHQ-1, PHQ-2, PHQ-3, PHQ-4, PHQ-5, PHQ-6, PHQ-7, and PHQ-9), age and sex of patients, and period time. Thus, this dataset is essential for research on depression management and classification. Figure 2 explains the classification of gender data.

3.1.2 Feature engineering

Feature engineering also involves data preprocessing, data cleaning and feature selection.

- Data cleaning: This is a crucial step in deep data preprocessing before feeding data to the ML algorithms. It involves entries on missing values, outliers, and NaN values in a given dataset.
- Handling missing values: Forward and backward fill methods were used to handle the missing values in the PHQ-9 dataset. Of course, rows and columns that contain missing values were deleted.
- Scaling values: Scaling was performed to bring all the attributes of the dataset onto a similar range or ground. In this study, the age column was rescaled using the standard scale to avoid any influence on the model's learning process.
- Encoding: Encoding is the conversion of the categorical data into a different form that is basically numerical. One hot encoding was applied for the "sex" column, while ordinal encoding was applied for the "period-time" column.
- Feature selection: Feature selection is important in training a model. When tested with other unknown samples, it makes the model generate the right outcome. Only features with a significant impact on the

model's performance were selected for training. Thus, PHQ-1, PHQ-2, PHQ-3, PHQ-4, PHQ-5, PHQ-6, PHQ-7, PHQ-9, age, sex, and period time were chosen as the variables because they enable the identification of the depression level.

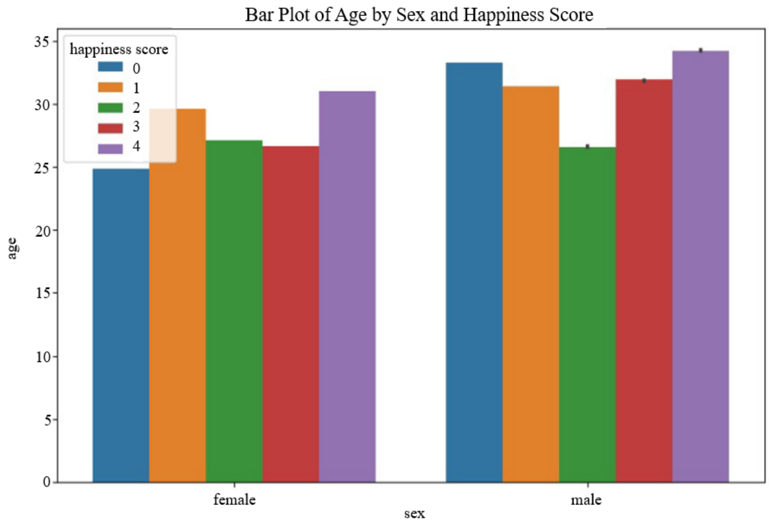

Figure 2. Bar plot of age by sex and happiness score

3.2 Model Training

In this step, the preprocessed data was given to the ML algorithms to train the model. The small-scale dataset containing PHQ-9 was split into training data (80%) and testing data (20%). Two well-known algorithms were used for training, i.e., SVC and ABC.

$$f\left(x\right)=sign\left(\sum_{i=1}^{nsv}\alpha iyiK\left(xi,x\right)+b\right) \tag{1}$$

3.2.1 SVC

SVC is a supervised learning model which can predict the category of unseen data with the help of training data. It works with the use of a hyperplane that separates classes. The main intention of SVC is to maximize the margin of the classes on the hyperplane in order to contain fewer classification errors. In this study, the linear kernel (LSVM) was chosen as it is mostly suggested for category classification of the target column.

3.2.2 Rationale for selecting SVC

SVC was chosen owing to its utilization for operating in more dimensions and its performance in the case of higher dimensionality than the number of instances. Furthermore, it is flexible with regard to the choice of kernel functions used in the decision function.

3.2.3 Parameter settings

Kernel: Linear
Regularization parameter (C): 1.0
Gamma: 'scale'

3.2.4 ABC

AdaBoost is an iterative method carried out in rounds, in which it accumulates weak classifiers. At the beginning, the training data had the same weights. Thereafter, along with an increase in the weights of misclassified data points, a new classifier is built with new weights.

$$H\left(x\right)=sign\left(\sum_{t=1}^{T}\alpha t\times ht\left(x\right)\right) \tag{2}$$

3.2.5 Rationale for selecting AdaBoost

AdaBoost was selected because of its feature of using several weak classifiers to develop one strong classifier. It is most helpful for increasing the performance of models with barely more than ideal predictive rates.

3.2.6 Parameter settings

The base estimator is a Decision Tree Classifier, with 50 estimators and a learning rate of 1.0.

4. Model Evaluation

In the last step, the effectiveness of the created model was reviewed using unseen data to make the assessment. The evaluation was conducted using several key metrics, including accuracy, precision, recall, and the F1-score, each of which provides a distinct perspective on the model's generalizability and reliability. The target column has five scores, which are interpreted in Table 2. By considering the values according to classes (0-4), both Figure 3 and Figure 4 provide the results.

Table 2. Labels and scores of the target column

Category	Score (Target)
Normal/happy/non-depressive	4
Mild depression	3
Moderate depression	2
Moderately severe depression	1
Severe depression	0

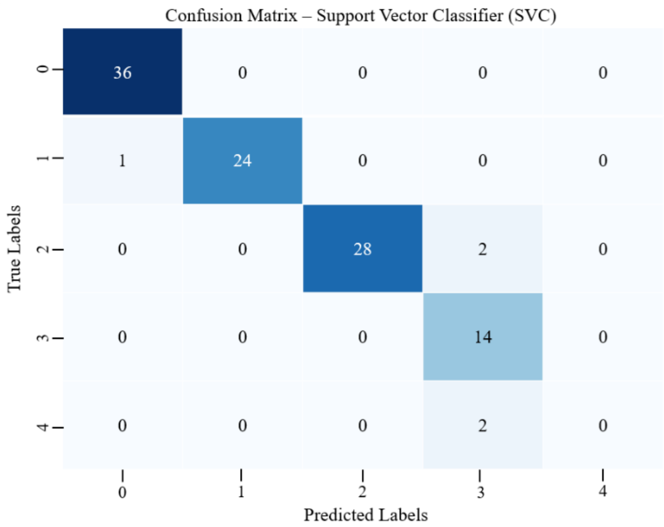

Figure 3. Confusion matrix of SVC

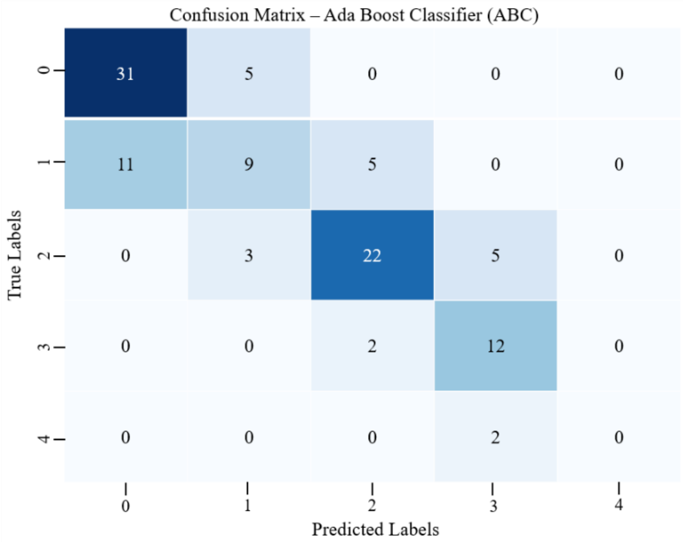

Figure 4. Confusion matrix of ABC

4.1 Accuracy

Accuracy is a measure of the overall correctness of the model. It calculates the ratio of correctly predicted instances to the total number of instances in the dataset (Zainudin et al., 2021).

$$Accuracy = \frac{Number\ of\ Correct\ Predictions}{Total\ Number\ of\ Predictions} \tag{3}$$

4.2 Precision

The precision of the model's positive predictions is a measure of their accuracy. The ratio of accurately predicted positive observations to the total number of expected positives is computed (Zainudin et al., 2021).

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives} \tag{4}$$

4.3 Recall

A model's recall is a measurement of its capacity to identify every pertinent incident. The ratio of accurately anticipated positive observations to actual positives is computed (Zainudin et al., 2021).

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Positives} \tag{5}$$

4.4 F1-score

The harmonic mean of recall and accuracy is the F1-score. It offers a compromise between recall and accuracy, accounting for both false negatives and false positives (Zainudin et al., 2021).

$$F1\ Score = 2\ X\ \frac{Precision\ x\ Recall}{Precision + Recall} \tag{6}$$

4.5 Confusion Matrix

A confusion matrix is frequently used to assess how well a classification system performs. It is very helpful in the field of ML when evaluating a model's performance based on its predictions on a dataset (Mohammad & Siddiqui, 2021). The confusion matrices of trained models are as follows:

However, the results show that SVM performs much better and achieves an accuracy of 95% compared to ABC with an accuracy of 69%, as mentioned in Table 3.

Table 3. Model evaluation

Algorithm	Accuracy	Precision	Recall	F1-score
SVC	0.95	0.94	0.95	0.94
ABC	0.69	0.66	0.69	0.67

5. Comparative Analysis

This section contains a comparative study of the existing approaches regarding depression detection for the PHQ-9 dataset. Table 4 summarizes the key aspects, such as methods used, parameters, accuracy, and limitations.

Table 4 shows the proposed method performs much better with an accuracy of 95% compared to the existing methods. In addition, it overcomes existing issues.

Table 4. Model evaluation

Author(s)	Method	Accuracy	Limitations
Dey et al. (2022)	RF	74%	Analyses are based only on texts on social media platforms, limiting the results to the community understudy, which cannot be generalized to different populations

Kim et al. (2021)	RF	94%	Primarily only used by physicians without application by end users
Glavin et al. (2022)	-	-	Need for further performance of PHQ pairing tests with other populations
Méndez et al. (2021)	Alexa, HMI, fuzzy logic	-	Limited to only household use without versatility in its usage
Mohammad & Siddiqui (2021)	RF, Optuna optimization	83%	Observed failure to incorporate it for end users, while it mainly deals with improving efficiency
Chiong et al. (2021)	Single and ensemble classifiers	-	Suitable only for labeled datasets without the user-interface of an end product
Ecker et al. (2024)	-	-	Determination of suicidal risk in patients without a focus on PHQ-9 or depression
This study	SVC	95%	-

6. Discussion

Amid the rapid evolution of technological advancements in mental health, the "Depression Identification" initiative stands as a significant milestone, embarking on its mission to enhance early diagnosis and treatment of depression. Consequently, the study opens a new chapter in depression diagnosis by providing appropriate treatments and care at the earliest possible time. The SVC model has demonstrated exceptional performance, achieving an accuracy of 94%, underscoring its effectiveness in this context. Although many important achievements have been made, a number of issues require further considerations as follows:

a) Data privacy issues: Work with the employee's personal and potentially pathological data, including data in the development of the PHQ-9 dataset, is the main privacy concern. Safety for all data submitted by users should be a major concern. It is imperative that future research focus on enhancing data protection measures to ensure that such information is adequately safeguarded against potential breaches.
b) Risk of misdiagnosis: As it has been established, the model's ability to classify patient enrollments is accurate. However, there is room for the incorrect classification of patients whom physicians may later misdiagnose. Over- or under-classifications may be a result of filtering the dataset or the general prejudices of the model. To reduce these risks and improve the model's dependability, constant verification and reference to clinical evaluation are required.
c) Dataset limitations: Some drawbacks can be pointed out in terms of generalizability energy, which might be associated with the PHQ-9, which has 534 samples. A limited number of observations can be significantly insufficient to represent all the possible expression forms of a disease in different population groups and in various conditions. As mentioned above, the need to increase the dataset and use multiple data sources is essential in enhancing the model's reliability.

It is crucial to recognize that the overall approach of the study is user-centered, which is a key strength. In particular, this study fills the gaps in existing literature and the zest of depression treatment by proposing a method to create a simple and clear self-monitoring interface for depression levels. "Unlocking minds" is a ray of hope, showing that people can work hand in hand with technology to make the world a better place to live, especially finding ways to empathize with one another to seek much-needed mental health.

7. Conclusion

A person suffering from depression deals with a major mental health issue. For this reason, it is essential to recognize and prevent depression early. The PHQ-9 dataset, which has 534 samples, was used in this study. However, SVC and ABC are only a few of the ML models that were trained using these samples. In a similar vein, comparison analysis, revealing the top classifier, was carried out after assessments. SVC obtained 94% accuracy on this dataset. Following evaluation and training, the trained model was implemented to offer the users an interactive interface, allowing them to monitor their state of depression and feel more at peace. The patient dataset could be expanded in the future to train ML algorithms. Moreover, the functionality of the application was significantly enhanced, leading to the development of a more robust and generalizable model. This model, designed to operate invisibly, accounts for the dynamic nature of depressive patients' environments and lifestyles, offering a more realistic approach to mental health monitoring.

In terms of future directions, more data could be incorporated, such as instances and an increased variety of the population, to make the model more stable and versatile. Moreover, the classification performance could be enhanced by researching and implementing more complex ML and DL algorithms. Lastly, the development of features providing real-time data collection and analysis within the specified interface would also be useful for providing the users with immediate feedback, enhancing the overall efficacy of the presented tool in mental health issues.

Data Availability

The data used to support the research findings are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflict of interest.

References

Alenizi, J. & Alrashdi, I. (2023). SFMR-SH: Secure framework for mitigating ransomware attacks in smart healthcare using blockchain technology. *Sustain. Mach. Intell. J.*, *2*(4), 1-19. https://doi.org/10.61185/SMIJ.2023.22104.

Bhanot, N., Ahuja, J., Kidwai, H. I., Nayan, A., & Bhatti, R. S. (2023). A sustainable economic revival plan for post-COVID-19 using machine learning approach–A case study in developing economy context. *Bench. Int. J., 30*(6), 1782-1805. https://doi.org/10.1108/BIJ-09-2021-0564.

Bhathena, D. (2021). Leveraging unlabeled data in supervised learning to objectively assess depression. (Doctoral Dissertation, Massachusetts Institute of Technology). https://hdl.handle.net/1721.1/139306

Chiong, R., Budhi, G. S., Dhakal, S., & Chiong, F. (2021). A textual-based featuring approach for depression detection using machine learning classifiers and social media texts. *Comput. Biol. Med., 135*, 104499. https://doi.org/10.1016/j.compbiomed.2021.104499.

Dey, S., Nawshin, A. T., Hossain, A., & Rahman, R. M. (2022). Machine learning combined with PHQ-9 for analyzing depression in Bangladeshi metropolitan areas. In 2022 IEEE 11th International Conference on Intelligent Systems (IS), pp. 1-8. https://doi.org/10.1109/IS57118.2022.10019625.

Ding, C., Zhang, Y., & Ding, T. (2023). A systematic hybrid machine learning approach for stress prediction. *PeerJ Comput. Sci.*, *9*, e1154. https://doi.org/10.7717/peerj-cs.1154.

Ecker, B., Ticic, R., & Hulley, L. (2024). *Unlocking the Emotional Brain: Memory Reconsolidation and the Psychotherapy of Transformational Change.* Taylor & Francis.

El-Latif, A. A. A., Chelloug, S. A., Alabdulhafith, M., & Hammad, M. (2023). Accurate detection of Alzheimer's disease using lightweight deep learning model on MRI data. *Diagnostics*, *13*(7), 1216. https://doi.org/10.3390/diagnostics13071216.

Garg, P., Santhosh, J., Dengel, A., & Ishimaru, S. (2021). Stress detection by machine learning and wearable sensors. In Companion Proceedings of the 26th International Conference on Intelligent User Interfaces, pp. 43-45. https://doi.org/10.1145/3397482.3450732.

Glavin, D., Maekawa, E., Grua, E. M., Nakamura, C. A., Scazufca, M., Araya, R., ..., & van de Ven, P. (2022). Other PHQ-9 item pairings are better than the PHQ-2: A machine learning analysis. *Procedia Comput. Sci., 206*, 101-110. https://doi.org/10.1016/j.procs.2022.09.089.

Gupta, A., Maurya, C., Dhere, K., & Chaurasiya, V. K. (2022). Wellness detection using clustered federated learning. In 2022 IEEE 6th Conference on Information and Communication Technology (CICT), pp. 1-5. https://doi.org/10.1109/CICT56698.2022.9997827.

Haq, H. B. U., Suwansantisuk, W., & Chamnongthai, K. (2023). An optimized deep learning method for video summarization based on the user object of interest. *Int. J. Adv. Comput. Sci. Appl.*, *14*(10), 244-256.

Ilias, L., Mouzakitis, S., & Askounis, D. (2023). Calibration of transformer-based models for identifying stress and depression in social media. *IEEE Trans. Comput. Soc. Syst.*, *11*(2), 1979-1990. https://doi.org/10.1109/TCSS.2023.3283009.

Jagtap, N., Shukla, H., Shinde, V., Desai, S., & Kulkarni, V. (2021). Use of ensemble machine learning to detect depression in social media posts. In 2021 Second International Conference on Electronics and Sustainable Communication Systems (ICESC), pp. 1396-1400. https://doi.org/10.1109/ICESC51422.2021.9532838.

Jyothirmy, S., Geethika, G., Sai, S. B., Saiteja, B., & Reddy, V. P. (2023). Machine learning algorithms based detection and analysis of stress–A review. In 2023 Second International Conference on Electronics and Renewable Systems (ICEARS), pp. 1456-1463. https://doi.org/10.1109/ICEARS56392.2023.10084933.

Kim, S., Lee, H. K., & Lee, K. (2021). Which PHQ-9 items can effectively screen for suicide? Machine learning approaches. *Int. J. Environ. Res. Public Health*, *18*(7), 3339. https://doi.org/10.3390/ijerph18073339.

Korti, S. S. & Kanakaraddi, S. G. (2022). Depression detection from Twitter posts using NLP and machine learning techniques. In 2022 Fourth International Conference on Emerging Research in Electronics, Computer Science and Technology (ICERECT), pp. 1-6. https://doi.org/10.1109/ICERECT56837.2022.10059773.

Lorenzoni, G., Tavares, C., Nascimento, N., Alencar, P., & Cowan, D. (2024). Assessing ML classification algorithms and NLP techniques for depression detection: An experimental case study. https://doi.org/10.48550/arXiv.2404.04284.

Mandour, S. (2023). An exhaustive review of neutrosophic logic in addressing image processing issues.

Neutrosophic Syst. Appl., 12, 36-55. https://doi.org/10.61356/j.nswa.2023.110.
Manimeghalai, P., Jayalakshmi, P. K., Chandran, R. R., Krishnan, S., & Shiny, S. (2022). ECG based stress detection using machine learning. In 2022 Second International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT), pp. 1-5 https://doi.org/10.1109/ICAECT54875.2022.9807877.
Méndez, J. I., Meza-Sánchez, A. V., Ponce, P., McDaniel, T., Peffer, T., Meier, A., & Molina, A. (2021). Smart homes as enablers for depression pre-diagnosis using PHQ-9 on HMI through fuzzy logic decision system. *Sensors, 21*(23), 7864. https://doi.org/10.3390/s21237864.
Mohammad, B. N. S. & Siddiqui, K. (2021). Random forest regressor machine learning model developed for mental health prediction based on MHI-5, PHQ-9 and BDI scale. In Proceedings of the 4th International Conference on Advances in Science & Technology (ICAST2021). https://doi.org/10.2139/ssrn.3867416.
Munthuli, A., Pooprasert, P., Klangpornkun, N., Phienphanich, P., Onsuwan, C., Jaisin, K., Pattanaseri, K., Lortrakul, J., & Tantibundhit, C. (2023). Classification and analysis of text transcription from Thai depression assessment tasks among patients with depression. *PLoS One, 18*(3), e0283095. https://doi.org/10.1371/journal.pone.0283095.
Nadeem, A., Naveed, M., Islam Satti, M., Afzal, H., Ahmad, T., & Kim, K. I. (2022). Depression detection based on hybrid deep learning SSCL framework using self-attention mechanism: An application to social networking data. *Sensors, 22*(24), 9775. https://doi.org/10.3390/s22249775.
Nazeer, M., Salagrama, S., Kumar, P., Sharma, K., Parashar, D., Qayyum, M., & Patil, G. (2024). Improved method for stress detection using bio-sensor technology and machine learning algorithms. *MethodsX, 12*, 102581. https://doi.org/10.1016/j.mex.2024.102581.
Schaub, F. (2023). *The Empowerment Solution: Six Keys to Unlocking Your Full Potential with the Subconscious Mind.* Simon and Schuster.
Sharif, M. S., Zorto, A., Kareem, A. T., & Hafidh, R. (2022). Effective machine learning based techniques for predicting depression. In 2022 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT), pp. 366-371. https://doi.org/10.1109/3ICT56508.2022.9990824.
Sleem, A. (2022). Empowering smart farming with machine intelligence: An approach for plant leaf disease recognition. *Sustain. Mach. Intell. J., 1*(3), 1-11. https://doi.org/10.61185/SMIJ.2022.1013.
Squires, M., Tao, X., Elangovan, S., Gururajan, R., Zhou, X., Acharya, U. R., & Li, Y. (2023). Deep learning and machine learning in psychiatry: A survey of current progress in depression detection, diagnosis and treatment. *Brain Inf., 10*(1), 10. https://doi.org/10.1186/s40708-023-00188-6.
Stephen, D., Attallah, A. A., & Ayu, M. A. (2022). Improving early detection and prevention of depression using an interactive mobile app. In 2022 IEEE 8th International Conference on Computing, Engineering and Design (ICCED), pp. 1-6. https://doi.org/10.1109/ICCED56140.2022.10010529.
Sun, S., Folarin, A. A., Zhang, Y., Cummins, N., Garcia-Dias, R., Stewart, C., ..., & Dobson, R. J. B. (2023). Challenges in using mHealth data from smartphones and wearable devices to predict depression symptom severity: Retrospective analysis. *J. Med. Internet Res., 25*, e45233. https://doi.org/10.2196/45233.
Thabit, S. A. (2024). Perspective chapter: Embracing the complexity of human emotion. In Advances in Sentiment Analysis-Techniques, Applications, and Challenges. https://doi.org/10.5772/intechopen.112701.
Walambe, R., Nayak, P., Bhardwaj, A., & Kotecha, K. (2021). [Retracted] Employing multimodal machine learning for stress detection. *J. Healthc. Eng., 2021*(1), 9356452. https://doi.org/10.1155/2021/9356452.
Wani, M. A., ELAffendi, M. A., Shakil, K. A., Imran, A. S., & Abd El-Latif, A. A. (2022). Depression screening in humans with AI and deep learning techniques. *IEEE Trans. Comput. Soc. Syst., 10*(4), 2074-2089. https://doi.org/10.1109/TCSS.2022.3200213.
Zainudin, Z., Hasan, S., Shamsuddin, S. M., & Argawal, S. (2021). Stress detection using machine learning and deep learning. *J. Phys.: Conf. Ser., 1997*(1), 012019. https://doi.org/10.1088/1742-6596/1997/1/012019.