Metaheuristic Optimization Review

♦ ISSN: 3066-280X Vol: 06, No. 01, 2026 ♦ PP. 40–54

Article DOI: https://doi.org/10.54216/MOR.060104

RESEARCH ARTICLE

Hybrid Metaheuristic–Ensemble Pipeline for Student Mental Health: Waterwheel Plant Algorithm with Random Forest

Amel Ali Alhussan[1,*] **Abdelaziz A. Abdelhamid**[2]

[1] Department of Computer Sciences, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia

[2] Department of Computer Science, Faculty of Computer and Information Sciences, Ain Shams University, Cairo 11566, Egypt

Emails: aaalhussan@pnu.edu.sa · abdelaziz@cis.asu.edu.eg

Received: January 17, 2026 Revised: March 03, 2026 Accepted: May 03, 2026 ⋆ Corresponding author

ABSTRACT

The problem of depression among college students has become a burning research problem, as the number of psychosocial stress factors, academic loads, and lifestyle disorders that lead to the worsening of mental health has increased. Driven by the increasing need for innovative, data-driven, and interpretable diagnostic models, this paper presents a combined Machine Learning (ML) and metaheuristic optimization model for predicting depression using multidimensional psychosocial and academic data collected from 100 Computer Science students. The suggested hybrid model combines a Random Forest (RF) classifier with the Waterwheel Plant Algorithm (WWPA), a nature-inspired mechanism for optimizing hyperparameter settings and feature selection. Experimentation using the Random Forest baseline model yielded a baseline accuracy of 0.9081, a Sensitivity (True Positive Rate) of 0.8936, and an F-Score of 0.9032. The hybrid WWPA+Random Forest model showed significant gains after introducing the WWPA optimization method, achieving a high accuracy of 0.9577, a sensitivity of 0.9502, a specificity of 0.9644, and an F-score of 0.9553. These findings validate the high-quality performance of the proposed model in achieving balanced, high-precision classification and in resisting overfitting. The results highlight the potential to integrate ensemble learning with bio-inspired optimization to advance depression prediction, providing a scalable, explainable, and ethically appropriate framework for predicting depression early in a person's life. This work opens the way to creating a proactive digital mental health system that will enable educational organizations to identify at-risk students early and offer timely, individualized support for well-being and academic achievement.

Keywords: Depression Detection ▪ Machine Learning (ML) ▪ Waterwheel Plant Algorithm (WWPA) ▪ Random Forest Optimization ▪ Student Mental Health Analytics

1. INTRODUCTION

Depression is considered one of the most significant issues in the world, with millions of individuals being victims; students are the most vulnerable group. Academic stressfulness, a feature of academic activity, the personal, social, and emotional issues of students precondition the development of depressive symptoms in them [1]. Among the symptoms of depression that are observed in students are lack of concentration, lack of motivation and withdrawal behaviour and general poor performance in both their academic and personal lives. The existing stigma about mental health only exacerbates the existing problems by compelling some learners to hesitate to seek help [2]. In addition, conventional methods of depression testing such as self-report surveys and clinical interviews are usually inaccurate and time-consuming. These shortcomings indicate the necessity of precise, fast, and repeatable predictors of depression, particularly those in a student population, where early diagnosis has shown to be incredibly helpful in this context [3]. As it would be proposed in the present-day world, Artificial Intelligence (AI) and Machine Learning (ML) have already become effective in addressing the urgent problems in various spheres, such as medical and mental health application usage [4]. To detect depression, the technologies introduce novel ways of analysing large, multi-parametric data sets in less time and more accurately than the previously used statistical approaches have done so far [5]. Machine Learning and Artificial Intelligence can be applied to identify the correlation and patterns of data that humans cannot easily identify and therefore give new insights into how mental conditions change and develop. These technologies in mental health research can enable making models to predict the amount of risk of depression considering the psychosocial factors. Not only such models enhance the level of recognition of depression among students, but they also herald the opportunities of personalized recovery programs [6]. The effectiveness of any particular ML model is dependent significantly on the data analysis procedure, which occurs during model development or before it is developed. The primary elements of depression detection models are psychosocial data, comprising behavioral aspects, moods, academic stress, and interpersonal communication [7]. These data may be complex, very correlated, mutually dependent; thus, there is a need to resort to data preprocessing and feature extraction. This is what allows excluding the features that are irrelevant or redundant relative to the status of depression and retaining the ones that can be informative. Other than enhancing the efficiency and performance of ML models, it also presents the data in a format that facilitates the researcher and practitioner in the mental health field comprehend results and the factors leading to depression more easily and clearly [8]. In this way, the analysis of psychosocial data and high performance of ML models can be ensured through systematic analysis and processing of data, which is sufficient to diagnose and predict depression with high reliability [9]. In order to achieve more precise outcomes when using the ML models, the metaheuristic optimization method is more and more used to tune the parameters and to choose the features. These algorithms are nature inspired and can be used to solve large scale problems that have high dimensional search spaces [10]. One of these new metaheuristic methods is called Waterwheel Plant Algorithm (WWPA), which is based on the structure of the waterwheel plant system. As it is shown in the considered paper, the WWPA can be the tool of great utility when it is necessary to regulate the exploitation and exploration capacities of ML models. We adopt the proposed WWPA in this study and add it to the widely adopted ML algorithm, the Random Forest model, which is both very accurate and interpretable. This type of integration intends to enhance the status classification of depression by maximizing a number of parameters in the Random Forest model. Thus, the value of WWPA, as well as Random Forest, is used in this study to promote the accuracy of depression diagnosis, regarding the complexity and diversity of psychosocial data.

1.1 Paper Contribution

The main contributions of this paper are outlined as follows:

- **Integration of Psychosocial Dimensions:** This paper examines the complex interplay of psychosocial determinants and depression in university students, providing an in-depth data-driven view of the interaction of academic, social, and personal stressors to determine the impact on mental health.
- **Development of an Enhanced Predictive Framework:** To improve classification accuracy and interpretability, a strong predictive model is proposed by combining the Random Forest (RF) classifier with the new Waterwheel Plant Algorithm (WWPA), a successful blend of ensemble learning and biologically-inspired optimization.
- **Metaheuristic Optimization for Model Refinement:** The WWPA is used to optimize the Random Forest model via adaptive feature selection and hyperparameter optimization, demonstrating its ability to improve convergence stability and search effectiveness in multidimensional datasets.
- **Comparative Evaluation with State-of-the-Art Methods:** The suggested WWPA-RF framework is rigorously compared with modern machine learning algorithms and metaheuristic-based methods to demonstrate its functionality, cross-applicability, and superior performance compared to conventional optimization methods.
- **Contribution to Mental Health Informatics:** The current study will offer a scalable, interpretable method for early depression detection by incorporating machine learning and metaheuristic optimization in mental health. It will aid in the development of intelligent, evidence-based decision-support systems for student well-being.
- **Foundation for Future Research:** The suggested method not only provides methodological knowledge of data-based mental health assessment but also prepares the way for the application of analogous optimization-based ML schemas to other psychological and behavioral prediction problems.

The rest of the paper will be organized in the following way: Section 2 provides an in-depth overview of the previous literature in the area of machine learning implementation in the field of mental health, with a particular focus on outlining the gaps in current knowledge that this study will seek to provide. Section 3 includes the description of the Materials and Methods section, where the dataset, preprocessing, feature selection methods and the application of proposed Waterwheel Plant Algorithm (WWPA) based Random Forest framework are described. Section 4 describes the Results, where the performance of the proposed model is described along with the comparative analysis with other algorithms of the state-of-the-art. A Discussion of the findings is given in Section 5 as the result of the interpretation of the findings in the general framework of psychosocial factors and mental health prediction. Lastly, Section 6 provides the Conclusion, which summarizes the main findings, practical implications and recommendations for the future research in the field of integrating machine learning with metaheuristic optimization in mental health assessment.

2. LITERATURE REVIEW

Early childhood symptoms of depression and anxiety play an essential role in determining the mental health and cognitive development of children. As noted in the research by [11], the relationship between mental health problems and cognitive development has been debated for over 20 years, and the need for early diagnosis and treatment cannot be overestimated. The paper examined the data of 3,984 school children aged between 10 and 15 years, through a machine learning approach, to discover the factors that cause depression and anxiety. The findings indicated that school violence, bullying, family violence, academic performance, and family income are the most important factors that influence mental health. It was demonstrated that support vector machines and random forests are highly accurate at predicting these symptoms, underscoring their future use in improving mental health programs in schools and cognitive development.

Physical education (P.E.) in colleges must be changed to psychological education to ensure students' overall growth. According to the results [12], the conventional teaching approaches in P.E do not address the psychological aspects of learning, which significantly restrict the learning process and self-growth. The teaching model, based on the Internet of Things and deep learning, would be used for teaching psychology and the rest of the evaluation system. In this model, too, the increase in students' emotional self-control, self-challenge, and adaptability to adversity was also high, and average scores increased dramatically. Additionally, the sophisticated critical thinking and problem-solving skills also improved to a considerable extent, which is one of the indicators of higher learning. The strategy not only strengthened students' performance in training and theoretical courses but also improved their self-learning capacity and academic performance, which is extremely valuable for reforming P.E. teaching methods at the higher education level.

Stress in university students has increased as a result of COVID-19 because of the illness, mobile device dependency, decreased social interactions, and home isolation. According to recent discoveries [13], early stress detection is essential in academic success and mental health. A model based on machine learning and supervised algorithms was constructed and tested on the data of 444 students. Principal Component Analysis and hyperparameter optimization algorithms, such as Grid Search Cross-Validation, were among the techniques that improved the model. The best accuracy (80.5%) was obtained with the Multilayer Perceptron model, which detected high psychological (24.10%) and social stress (11.26%) levels. Although self-reported data have specific weaknesses, this research paper highlights the potential benefits of machine learning in informing interventions and enhancing students' well-being.

Mental health has been significantly impacted by the COVID-19 pandemic, especially among college students. After colleges reopened in China, controlling the psychological effects has also proven quite challenging. According to the latest study [14], 478 valid online questionnaires were investigated in a cross-sectional study to evaluate anxiety and depression among students and determine which factors influence these conditions using machine learning methods. Findings indicated that 15.5 percent of the students expressed anxiety symptoms, with 32.4 percent displaying depression. Some of the critical determinants of mental health were exercise rates, sleep quality, quarantine rates, the financial effect of quarantine on the family, and the conditions at campuses. The Synthetic Minority Oversampling Technique was used to address data imbalance, and multivariate logistic regression was used to identify significant predictors. The machine learning model showed good predictiveness as the AUC values of anxiety and depression were 0.885 and 0.806, respectively. The results of this research provide practical guidance on how to manage schools better and foster students' welfare.

One of the significant problems facing universities worldwide is student dropout, which refers to students who leave a higher education program without a degree. A recent study used machine learning models to predict first-year engineering student dropout at two Chilean universities and to identify predictive factors (as shown by recent research) [15]. It was found that results were higher when using the individual university models than when merging the datasets. Among eight machine learning models tested, gradient-boosted decision trees performed best. It was found that increases in all entrance exam scores, especially in mathematics, decreased the probability of dropout, and that increases in language test scores increased the probability of dropout. These results are practical for implementing specific interventions to improve the dropout rate.

One aspect that determines the quality of tertiary institutions is student performance, as academic performance is a primary indicator used to rank universities. In the study done by [16], machine learning techniques were critically examined to be able to predict student outcomes and overcome problems with the big picture of the educational data at hand. A literature review conducted in 2015–2021, focusing on recent articles, identifies six basic machine learning models: decision trees, artificial neural networks, support vector machines, K-nearest neighbors, linear regression, and Naive Bayes. Among them, artificial neural networks proved to be the most accurate in determining performance. The significant predictive variables were academic, demographic, internal, and family or personal characteristics. The results reveal the increased interest in using machine learning in education and emphasize the possibility of applying it to detect performance gaps and facilitate the process of improving academic performance.

Enhancing the efficiency of ideological and political education in Chinese colleges has become increasingly significant. The combination of educational psychology and ideological and political courses was examined in a recent analysis of research aimed at enhancing teaching effectiveness and stimulating intelligent information changes [17]. A complete online teaching form, based on the deep learning method of Single Shot MultiBox Detector networks, along with flipped classes and Massive Open Online Courses, was created. A survey of 100 first-year and sophomore students found that they had a high level of acceptance of such courses. The results showed an increase in students' values, ideology, morals, and knowledge, with most having a positive attitude and involvement in the classes. These findings demonstrate that combining educational psychology with innovative technologies in pedagogical methods can effectively engage students and improve the learning process, which could help develop a particular course in the future.

Recent studies examined 238 participants by comparing the traditional teaching approaches with deep learning-based methods in the educational process [18]. Findings revealed that although both groups improved their academic performance, the deep learning group showed a stronger association between academic self-efficacy and performance. Although the relationship between well-being and achievement was not strong, it was still significant. Altogether, the research suggests that deep learning technologies might help improve academic performance and self-efficacy, which can be helpful to universities and music schools that use digital tools in their curricula as the world changes dramatically due to digitalization in education.

Deep neural networks in artificial intelligence have demonstrated a high potential in the field of healthcare. The broad use of deep learning models in physical health monitoring can also be used to assess mental health, as demonstrated in recent studies [19]. A deep learning-oriented mental health monitoring scheme (DL-MHMS) has been developed to monitor the mental health of college students, using convolutional neural networks to classify mental status (positive, negative, or normal) based on EEG signals. The model showed impressive results, with classification performance and F1 scores of 97.54 and 98.35, respectively.

The recent research has pointed out that the support of self-efficacy, teacher autonomy, and peer support is significant in fostering deep learning in college students [20, 21].

The strain of population increases, survival and academic competition is rising and has led to anxiety and mental problems among college students. The authors of [22] have considered social networks like Weibo, QQ, and WeChat as a way of emotional expression and stress and as a form of detecting depression.

The mental health issues of students usually require intervention, but there are no generalized criteria for starting the assistance in relation to health surveys. After recent discoveries [23], a machine learning model was created to predict mental health disorders of students in the same year and in the ensuing year based on health survey results.

Machine learning algorithms have been effective in predicting various health-related issues among students at a university. According to the recent research description [24], physical, mental, and social health were assessed using several indicators, and the random forest model outperformed others with 99.40% accuracy.

Following the findings of the recent study [25], a longitudinal study of the prevalence and risk factors of probable anxiety and insomnia among machine learning students was conducted.

Machine learning offers promising solutions for identifying and anticipating depressive symptoms among college students by leveraging behavioral data from smartphones and fitness devices. The study by [26] used data about 138 students to find those who had depressive symptoms at the end of the semester.

The General Health Questionnaire-12 was recently reinforced with socio-demographic and career-related questions to make up a comprehensive screening tool [27].

Recent studies [28] examined mental health in 444 students applying machine learning to categorize depression levels.

The recent research article [29] proposed the Intelligent Feature Subset Selection with machine learning-based DAS predictive model to assess depression, anxiety, and stress.

A recent research article [30] defined the most critical causes of depression in Bangladeshi undergraduates and predicted referral cases to psychiatric services.

This discussion is evident in the literature review, indicating a growing tendency to apply machine learning and deep learning to explore the psychosocial and academic factors that impact student well-being. Overall, as summarized in Table 1, this literature shows that the concept of data-based solutions for early detection, individual intervention, and holistic student development is groundbreaking in schools.

3. MATERIALS AND METHODS

The research framework proposed, as shown in Figure 1, explains the systematic sequence of work that will be used in this study to develop, evaluate, and optimize machine learning models. It starts with the preprocessing of the data, whereby the data are cleaned, normalized, and feature encoded to maintain data quality and consistency. After preprocessing, several baseline machine learning models, such as Random Forest, Support Vector Machine, Naive Bayes, and Decision Tree, are deployed to define baseline performance metrics. The evaluation stage is used to measure the effectiveness of the model by using key performance indicators to provide information on predictive accuracy and stability. In an attempt to improve the performance of the models, several optimization techniques are used, and they include the Waterwheel Plant Algorithm (WWPA), Genetic Algorithm (GA), Particle Swarm Optimization (PSO), and the Grey Wolf Optimizer (GWO). The combination of this workflow makes sure that model training, performance evaluation and optimization are approached comprehensively.

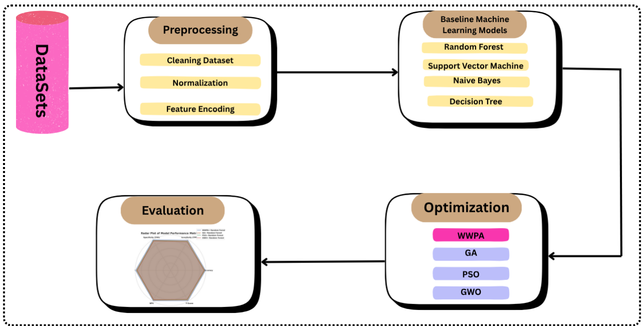

Figure 1. Proposed framework for machine learning model development, evaluation, and optimization.

3.1 Dataset Description

In this dataset, 100 Computer Science students were asked a questionnaire, and this is how they answered it. The first one is to determine the possible correlation between the level of depression, the academic performances, and the patterns of Attention Deficit Hyperactivity Disorder (ADHD), and the systematic analysis of the data will help with the task. The data include a broad set of psychosocial and scholarly variables that characterize students' demographic, behavioral,

Table 1. Summary of key literature: focus area, methodology, and findings.

Reference	Focus Area	Methodology	Key Findings and Contributions
[11]	Schoolchildren (10–15) depression/anxiety risk	ML (SVM, RF) on psychosocial data	Identified school/home violence, bullying, and academic pressure as predictors; high model accuracy for early detection.
[12]	Psychological education in physical education	IoT and DL-based teaching model	Enhanced emotional control, adaptability, and self-learning ability; improved course performance and engagement.
[13]	COVID-19 stress in university students	Supervised ML with PCA and grid search	MLP achieved 80.5% accuracy; revealed high psychological and social stress levels.
[14]	Post-pandemic anxiety and depression	Logistic regression with SMOTE	Anxiety (15.5%), depression (32.4%); key factors: exercise, sleep, quarantine, finance; AUC = 0.885/0.806.
[15]	Dropout prediction in engineering students	Gradient Boosting and comparative ML	Separate models outperform combined datasets; math scores lower dropout risk.
[16]	Student performance prediction	Review of DT, ANN, SVM, KNN, LR, NB	ANNs most accurate; key predictors: academic, demographic, and family features.
[17]	Ideological/political education reform	Deep learning (SSD), flipped classrooms, MOOCs	Improved moral, ideological, and cognitive engagement among students.
[18]	Music education and deep learning tools	Experimental comparison (DL vs. traditional)	Higher self-efficacy–performance correlation with DL; enhanced academic results.
[19]	EEG-based mental health assessment	CNN (DL-MHMS framework)	Accuracy 97.54%; reduced depression and sleep disorders; improved self-esteem.
[20]	Deep learning in higher education psychology	SPSS/AMOS mediation–moderation modeling	Self-efficacy mediates teacher autonomy–deep learning link; peer support moderates effect.
[21]	K–12 online learning analytics	Deep learning on interaction data	Enhanced prediction of learning behaviors via multimodal data.
[22]	Depression via social media	DISVM using Weibo text data	High detection stability and accuracy; supports early depression detection.
[23]	Annual student health survey	Ensemble ML (LightGBM best)	High MCC; campus-life questions key predictors; early mental health flagging tool.
[24]	Comprehensive health (physical, mental, social)	KNN, DT, NB, RF with PHQ-9, GAD-7, ISI, SBQ-R	RF achieved 99.4% accuracy; effective health assessment via ML.
[25]	COVID-19 anxiety and insomnia	Longitudinal XGBoost modeling	Accuracy: 97.3% anxiety, 96.2% insomnia; main predictors: relationships, sleep, ideation.
[26]	Depression via smartphone and wearable data	Longitudinal sensing + feature selection	85%+ accuracy up to 15 weeks prior; strong potential for early intervention.
[27]	Depression screening (GHQ-12, Bangladesh)	16 ML models; ExtraTrees best	60% depression prevalence; accuracy 90.26%; robust, scalable ML screening.
[28]	Mobile learning and depression detection	SPCA + CatBoost classification	44% not depressed; 31% severe; SPCA–CatBoost boosted accuracy 15%.
[29]	COVID-19 DAS (Depression–Anxiety–Stress)	IFSSML-DAS (G-GWO + BSO-LSSVM)	High predictive power; emphasized ML use for crisis mental health management.
[30]	University depression	ML (LR, KNN, SVM, RF, DT) + Android app	Best-performing ML identified; mobile app developed for student support.

and psychological features. Table 2 gives a detailed outline of each of the features of the dataset. The features depict specific aspects of students' academic or psychosocial portraits, and we can study the roles of different factors in demographic characteristics, mental health, learning patterns, and socialization relationships, which impact academic performance and well-being. The profiles of all these variables will be used to identify significant relations and trends, which will be utilized in predictive modelling and interventions for particular mental health problems.

The data will be aimed at facilitating a detailed analysis of the interactions among demographics, academic performance, psychological determinants, academic objectives, and the correlation among pupils who study Computer Science. These multidimensional features provide researchers with avenues to investigate meaning-based associations and predictive relationships that could help elucidate students' well-being and learning performance. The amount of time students spend in bed per day is a vital indicator of their health and academic success. Sleep is an imperative psychosocial determinant that affects the concentration of students, motivation, and emotional stability. Figure 2 indicates a sleeping average of 6 to 8 hours a day, which is within the recommended sleep patterns in young adults. Nonetheless, a significant percentage of students claimed to sleep fewer than 6 hours, which could be a pointer to high chances of insomnia and ineffective thinking. The relationship between the level of sleep and depression is an interesting topic to examine to comprehend the contribution made by mental health to the sleeping habits of students, as well as their general health. Figure 3 illustrates the mean hours of sleep for participants across the different depression states. In the visualization, it is possible to observe that the students who have reported depression have shorter and less regular sleep times than the students who have not reported any depression symptoms. Conversely, the same report on individuals who do not claim to be depressed shows that they are more likely to have normal cycles of sleep, and this implies that there are also strong connections between sleep behavior and mental health.

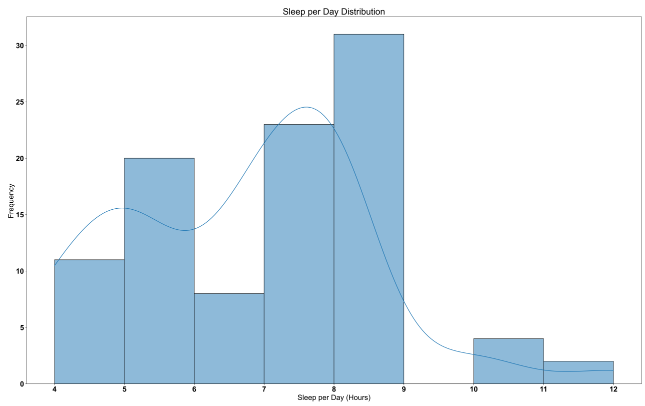

Figure 2. Sleep per Day Distribution

3.2 Machine Learning Models

This paper involved the use of four Machine Learning (ML) algorithms to classify and predict whether the students of

Table 2. Description of Dataset Features

Feature	Description
Age	Represents the age of each participant in the dataset, providing insight into the age distribution of the study population.
Gender	Indicates the gender of each individual, enabling exploration of gender-based variations and potential patterns within the dataset.
Academic Performance	Reflects the academic achievements of participants, serving as a critical variable for analyzing the relationship between mental health and learning outcomes.
Taking Notes in Class	Describes whether individuals take notes during lectures, offering insights into study habits, attention levels, and classroom engagement.
Depression Status	Indicates the presence or absence of depressive symptoms, contributing valuable information on the mental health profile of participants.
Facing Challenges in Completing Academic Tasks	Explores whether participants experience difficulties completing academic tasks, providing an understanding of academic stress and cognitive load.
Liking Presentations	Reflects participants' preferences for presentations, revealing insights into learning styles, communication preferences, and whether individuals display extroverted or introverted tendencies.
Sleep Hours per Day	Represents the average number of hours of sleep participants get per day, providing information on sleep patterns and their possible correlation with mental health and academic performance.
Number of Friends	Quantifies the social dimension by indicating the number of friends each participant has, thereby contributing to an understanding of their social interactions and support networks.
Liking New Things	Explores participants' openness to new experiences or ideas, offering insights into adaptability, curiosity, and innovation-oriented behavior.

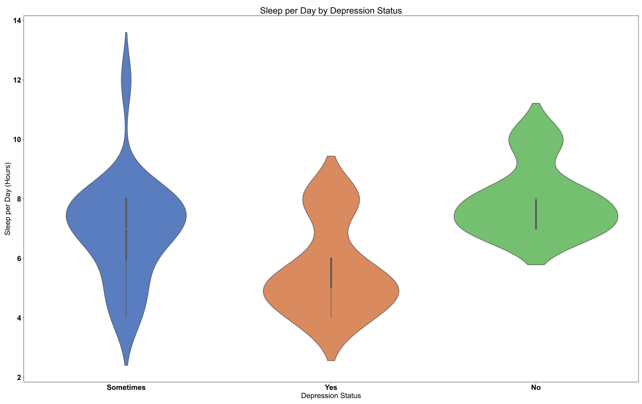

Figure 3. Sleep per Day by Depression Status

Computer Science were depressed or not, using psychosocial and academic predictors. These algorithms were selected because they are interpretable, robust, and relatively successful at classifying psychological and behavioral data. This section describes briefly and mathematically formulates each model.

3.2.1 Random Forest (RF)

The Random Forest algorithm is an ensemble-based learning technique that combines multiple decision trees to enhance predictive accuracy and reduce overfitting. Each tree in the forest is constructed from a random subset of the training data using a bootstrap sampling technique, while a random subset of features is selected for each split.

The prediction of a Random Forest classifier for a given input vector x is computed as the mode of the predictions from all individual trees:

$$\hat{y} = \text{mode}\{h_t(x) \mid t = 1, 2, \ldots, T\}$$

where $h_t(x)$ represents the class prediction from the t^{th} decision tree, and T denotes the total number of trees. The feature importance in Random Forest is evaluated using the Gini importance or mean decrease in impurity (MDI), given by:

$$I(f_j) = \frac{1}{T} \sum_{t=1}^{T} \sum_{n \in N_t(j)} p(n) \Delta i(n)$$

where $N_t(j)$ denotes the set of nodes in tree t that split on feature f_j, $p(n)$ is the proportion of samples reaching node n, and $\Delta i(n)$ represents the impurity decrease achieved by that split. The model is especially applicable for determining the most significant psychosocial characteristics for predicting depression.

3.2.2 Support Vector Machine (SVM)

The Support Vector Machine (SVM) is a supervised learning algorithm that aims to locate the best separating hyperplane between two classes of data. It aims to maximize the difference between the two classes, thereby enhancing classification accuracy and generalization.

Given a training dataset $\{(x_i, y_i)\}_{i=1}^{n}$ where $x_i \in \mathbb{R}^d$ and $y_i \in \{-1, +1\}$, the optimization problem for SVM can be formulated as:

$$\min_{\mathbf{w}, b, \xi_i} \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^{n} \xi_i$$

subject to:

$$y_i(\mathbf{w}^T \phi(x_i) + b) \geq 1 - \xi_i, \quad \xi_i \geq 0$$

where $\mathbf{w}$ is the weight vector, b is the bias, ξ_i are slack variables allowing misclassifications, and C is a penalty parameter controlling the trade-off between maximizing the margin and minimizing the classification error.

The kernel function $\phi(x_i)$ maps the input features into a higher-dimensional space to handle nonlinear relationships. The decision function is expressed as:

$$f(x) = \text{sign}\left(\sum_{i=1}^{n} \alpha_i y_i K(x_i, x) + b\right)$$

where $K(x_i, x)$ denotes the kernel function, and α_i are the Lagrange multipliers obtained during training.

3.2.3 Naive Bayes (NB)

A naive Bayes classifier is a probabilistic model based on the Bayes theorem that assumes conditional independence of the features given the class label. Although this is assumed, it turns out to be a very successful venture in most real-world scenarios due to its simplicity and efficiency. According to Bayes' theorem, the probability of:

$$P(C_k \mid X) = \frac{P(X \mid C_k)P(C_k)}{P(X)}$$

where:

- $P(C_k \mid X)$ is the posterior probability of class C_k given feature vector X,
- $P(X \mid C_k)$ is the likelihood of observing X given class C_k,
- $P(C_k)$ is the prior probability of class C_k, and
- $P(X)$ is the evidence or total probability of X.

Under the independence assumption, the likelihood $P(X \mid C_k)$ can be factorized as:

$$P(X \mid C_k) = \prod_{i=1}^{n} P(x_i \mid C_k)$$

The predicted class is then obtained by maximizing the posterior probability:

$$\hat{C} = \arg\max_{C_k} P(C_k) \prod_{i=1}^{n} P(x_i \mid C_k)$$

Naive Bayes is particularly effective for categorical or mixed-type data, making it suitable for analyzing survey-based mental health datasets.

3.2.4 Decision Tree (DT)

The Decision Tree algorithm is a non-parametric supervised learning model that divides the data into subsets based on feature values, creating a hierarchical tree structure. A test on an attribute is represented by each internal node, an outcome by each branch, and a class label by each leaf node. Gini impurity can be used to measure the impurity of a node t:

$$G(t) = 1 - \sum_{i=1}^{C} p(i \mid t)^2$$

where $p(i \mid t)$ is the probability of class i at node t and C is the total number of classes. Alternatively, entropy can be used to measure impurity:

$$H(t) = -\sum_{i=1}^{C} p(i \mid t) \log_2 p(i \mid t)$$

The goal at each split is to maximize information gain:

$$IG(t) = H(t) - \sum_{v \in \text{Values}(A)} \frac{|t_v|}{|t|} H(t_v)$$

In which attribute A is the attribute used to make the split, t_v is the subset of attribute A, which has value v, and $|t_v|/|t|$ is the proportion of samples in the subset. The decision trees are easy to comprehend and enable the illustration of the role of psychosocial and behavioral elements in the risk of depression among students. Together, these models will provide a robust comparative framework for forecasting the severity of depression and for explaining the psychosocial factors that influence students' well-being. Each algorithm has its advantages: both Random Forests and Decision Trees are highly interpretable, Naive Bayes has a probabilistic basis, and SVMs have strong classification performance in both linear and nonlinear settings.

3.3 Proposed Waterwheel Plant Optimization Algorithm (WWPA)

The Waterwheel Plant Optimization Algorithm (WWPA) is a new stochastic optimization technique inspired by the workings of natural systems. It is primarily driven by the way the waterwheel plants locate, capture, and consume their insect prey. Here, along with an explanation of the biological motivation, we present the mathematical formulation of the algorithm. Aldrovanda vesiculosa, or Waterwheel plant, has a long stem adorned with tiny, clear traps that resemble miniature flytraps. These traps are about 12 millimeters in diameter and are enclosed in hair-like bristles that prevent accidental firing or destruction. The hook-shaped teeth of the edges of its outside close with the closing of the trap, forming the exact interlock as that of the Venus flytraps. Each of them contains about forty long trigger hairs that detect movement; when touched, a trap closes very quickly. Moreover, digestive acids are secreted by glands within the trap, aiding nutrient digestion. An insect in captivity is trapped and gradually pulled to the base of the hinge with the help of the interlocking teeth, as well as the secretion of mucus. This continues till the prey is fully covered and digested. The Aldrovanda traps are active enough to trap and eat two to four of their prey before becoming inactive, like Venus flytraps. The WWPA is an iterative algorithm that aims to find the optimal solution in an ample, complex search space. A potential solution is to have every member of the population serve as a waterwheel. The population matrix L may be written as:

$$L = \begin{bmatrix} L_{1,1} & \cdots & L_{1,j} & \cdots & L_{1,m} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ L_{i,1} & \cdots & L_{i,j} & \cdots & L_{i,m} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ L_{N,1} & \cdots & L_{N,j} & \cdots & L_{N,m} \end{bmatrix}$$

where N denotes the number of waterwheels and m the number of decision variables. The initialization of each element is defined as:

$$L_{i,j} = lb_j + r_{i,j}(ub_j - lb_j), \quad i = 1,2,\ldots,N, \; j = 1,2,\ldots,m$$

Here, $r_{i,j}$ is a uniformly distributed random number between 0 and 1, while lb_j and ub_j represent the lower and upper bounds of the jth variable, respectively.

The corresponding objective function values for the population are represented as:

$$F = \begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_N \end{bmatrix} = \begin{bmatrix} F(X_1) \\ F(X_2) \\ \vdots \\ F(X_N) \end{bmatrix}$$

The value of each component of F is the fitness value of a waterwheel. The best solution corresponds to the highest value of the objective function, whereas the worst solution corresponds to the lowest. When the algorithm repeats, the waterwheels will search the search space with different paces. The dynamic nature enables the algorithm to update and find better solutions continually. The algorithm recreates the biological hunting technique of waterwheel plants, which involves pursuing and capturing prey as an exploratory stage. WWPA improves its navigation capabilities by mimicking a waterwheel's approach to an insect, while avoiding local minima. The following equations calculate the new position of a waterwheel:

$$W = r_1 \cdot (P(t) + 2K)$$
$$P(t+1) = P(t) + W \cdot (2K + r_2)$$

If improvement is not observed after three consecutive iterations, the position update follows a Gaussian perturbation rule:

$$P(t+1) = \text{Gaussian}(\mu_P, \sigma) + r_1 \frac{(P(t) + 2K)}{W}$$

Here, r_1 and r_2 are random variables within the ranges $[0,2]$ and $[0,1]$, respectively, and K is an exponentially decreasing control parameter within $[0,1]$. The waterwheel explores the search area in a circular region of diameter W, facilitating broad coverage of the solution space.

3.3.1 Exploitation Phase

During exploitation, the WWPA refines its search around promising areas by mimicking the process of transferring captured prey to feeding tubes. This local search improves convergence toward optimal solutions. The updated positions are computed as:

$$W = r_3 \cdot (KP_{\text{best}}(t) + r_3 P(t))$$
$$P(t+1) = P(t) + KW$$

where r_3 is a random variable in the range $[0,2]$, $P(t)$ is the current position, and P_{best} is the best-known position. If no improvement occurs over three iterations, a new position is generated as:

$$P(t+1) = (r_1 + K)\sin\left(\frac{F}{C}\theta\right)$$

where F and C are independent random variables with ranges $[-5,5]$. The exponential decrease of K is expressed by:

$$K = 1 + \frac{2t^2}{T_{\max}} + F$$

The WWPA process is iterative, continuously refining all positions until the termination condition (maximum iterations $T_{\max}$) is met. Upon completion, WWPA provides the best candidate solution found. Algorithm 1 presents the procedural implementation of WWPA, outlining initialization, exploration, and exploitation phases.

Algorithm 1 Waterwheel Plant Optimization Algorithm (WWPA)

Require: Population size n, maximum iterations $T_{\max}$, control parameters $r, \vec{r}_1, \vec{r}_2, \vec{r}_3, f, c, K$
Ensure: Optimal solution P_{best}

Initialize population of waterwheels $\{P_i\}_{i=1}^{n}$ randomly
Evaluate fitness $f_i = f(P_i)$ for all i
Determine initial best position P_{best}
Set iteration counter $t = 1$
while $t \leq T_{\max}$ **do**
  for $i = 1$ to n **do**
    if $r < 0.5$ **then** ▷ Exploration phase
      Compute $W = \vec{r}_1 \times (P_i(t) + 2K)$
      Update position: $P_i(t+1) = P_i(t) + W \times (2K + \vec{r}_2)$
      if no improvement for 3 iterations **then**
        $P_i(t+1) = \text{Gaussian}(\mu_P, \sigma) + \vec{r}_1 \frac{P_i(t)+2K}{W}$
      end if
    else ▷ Exploitation phase
      $W = \vec{r}_3 \times (K \cdot P_{\text{best}}(t) + \vec{r}_3 \cdot P_i(t))$
      $P_i(t+1) = P_i(t) + K \cdot W$
      if no improvement for 3 iterations **then**
        $P_i(t+1) = (\vec{r}_1 + K) \times \sin\left(\frac{F}{C\theta}\right)$
      end if
    end if
  end for
  Update control parameter: $K = 1 + \frac{2t^2}{T_{\max}^3} + f$
  Randomly refresh $r, \vec{r}_1, \vec{r}_2, \vec{r}_3, f, c$
  Re-evaluate fitness for all P_i
  Update global best P_{best}
  Increment $t = t + 1$
end while
return P_{best} as the optimal solution

Such an algorithmic description could serve as a valuable model for scholars and practitioners seeking to use or interpret the WWPA. It recounts the automatic stimulation and formalized method, which have been the driving forces behind this optimization plan, demonstrating how biological behavior and computational intelligence have worked closely together.

3.4 Evaluation Metrics

The predictive performance of the classification models was evaluated using six conventional evaluation measures, including Accuracy, Sensitivity (True Positive Rate, TPR), Specificity (True Negative Rate, TNR), Positive Predictive Value (PPV), Negative Predictive Value (NPV), and F-Score. All these measurements give the overall impression that the model is correct, reliable, and can balance false positives and false negatives. Table 3 summarizes their definitions and mathematical expressions. A combination of these metrics provides balanced model performance. Accuracy is general correctness; Sensitivity and Specificity are the model's discrimination capabilities; and PPV, NPV, and F-Score are reliability, precision, and balance.

4. EMPIRICAL RESULTS

The effectiveness of the advanced machine learning (ML) models and optimization algorithms that were used in this study is confirmed by their predictive performance. The random forest algorithm was found to be the most precise, and the balance was nearly perfect across all evaluation metrics for all models assessed. This observation indicates that it is highly effective in accurately distinguishing between positive and negative cases. Besides, the overall performance of the model was considerably enhanced when the feature-selection strategy was predetermined using the Random Forest and the newly introduced optimization algorithm, i.e., the Waterwheel Plant Algorithm (WWPA). It was found that F-score values and accuracy increased most, and the balance between precision and recall improved. These results show that combining ML models with metaheuristic optimization can improve the accuracy, reliability, and repeatability of classification results. The WWPA with the Random Forest indicates that this method can be successfully applied to complex data, and the model's results are more detailed in predicting depression.

4.1 Baseline Deep Learning Performance

In the process of thoroughly assessing the quality of the machine learning models in this study, specific standard evaluation metrics were identified, including Accuracy, Sensitivity (True Positive Rate, TPR), Specificity (True Negative Rate, TNR), Positive Predictive Value (PPV), Negative Predictive Value (NPV), and F-score. These metrics were preferred to bring a compromise between the predictive ability of each model and the generalization ability. Accuracy is the percentage of correctly classified samples, and Sensitivity is the model's ability to identify actual positive samples (students with depression) correctly. Specificity, on the other hand, is a metric that measures the percentage of actual negative cases the model correctly identifies, ensuring the model is not overtrained on the symptoms of depression. The model's predictive accuracy in identifying actual positive and negative cases can also be assessed using the PPV and NPV measures. The F-score is the harmonic mean of precision and recall, and it gives a single, balanced score of the performance of the model in the face of class imbalances. All the algorithms were tested and as shown in Table 4, the performance of the Random Forest model was the highest in overall performance. It was also found to classify depression status more accurately than any other type of accuracy was at 90.8 percent. It has high sensitivity (89.36%), and specificity (92.16%), which means that the model is not only consistent in its ability to identify depressed students but also non-depressed students to ensure that false positives and false negatives are reduced. Precision (PPV) value of 91.30% and NPV of 90.38% once again indicate the predictive reliability of the model, and F-Score of 90.32% can affirm the good balance between the precision and recall. The Support Vector Machine (SVM) model ranked just below with the accuracy of 87.37% which is competitive in terms of sensitivity and specificity. Its good performance in terms of metrics portrays its good generalization ability especially when addressing complex and nonlinear relationship within the dataset. The high F-Score (86.49%) of SVM indicates that it is efficient in the correct classification and misclassification trade-offs and thus a good alternative model in the depression detection activities. Naive Bayes (NB) classifier had an accuracy of 85.15% and performed relatively well with a slight margin of difference with SVM and random forest. Its simplicity and probability characteristic render the NB model appropriate towards comprehending the underlying statistical associations among the psychosocial variables, but it is inclined to assume independence of features that may not entirely understand the complex associations in the behavioral and emotional data. Decision Tree (DT) model had the lowest accuracy of 83.33 and moderate values of sensitivity and specificity. This notwithstanding, it is a useful model in explaining the effects of particular features on depression status due to its interpretability. It offers a clear decision-making framework where researchers can image which psychosocial or academic variables can make the most significant contribution to the depressive tendencies. In general, the comparison shows that the Random Forest algorithm is the most efficient baseline model as it is better than other methods in all significant evaluation metrics. This can be explained by the fact that it is an ensemble model and it is a combination of several decision trees that minimize variance and bias resulting in better generalization. The high results of the Random Forest give a good basis on which further enhancement of the performance through incorporation of the proposed optimization algorithm the Waterwheel Plant Algorithm (WWPA) to improve feature selection and model parameter optimization can be done. This equality is likely to give more accurate and consistent prediction as discussed in the other paragraphs. Performance assessment of models using various metrics shows a complete picture of the predictive power of the model and its consistency. Figure 4 is a radar chart that summarizes the performance of the Random Forest model as based on important measures of evaluation, such as Accuracy, Sensitivity (True Positive Rate), Specificity (True Negative Rate), Positive Predictive Value (PPV), Negative Predictive Value (NPV), and F-Score. The figure shows that the Random Forest classifier has steady values throughout the metrics and this indicates a balanced performance and good generalization capacity of the classifier in the dataset.

Histograms were plotted to show the standard deviation and mean of each of the evaluation metrics to better see the consistency and variability of model performance across multiple evaluation metrics. The distribution of key performance measures as Accuracy, Sensitivity (True Positive Rate), Specificity (True Negative Rate), Positive Predictive Value (PPV)

Table 3. Evaluation Metrics and Their Mathematical Definitions

Metric	Mathematical Formula	Interpretation
Accuracy	$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$	Measures the overall proportion of correctly classified students.
Sensitivity (TPR)	$\text{Sensitivity} = \frac{TP}{TP+FN}$	Measures how effectively students with depression are correctly identified.
Specificity (TNR)	$\text{Specificity} = \frac{TN}{TN+FP}$	Measures how effectively students without depression are correctly identified.
Positive Predictive Value (PPV)	$\text{PPV} = \frac{TP}{TP+FP}$	Indicates the reliability of positive predictions.
Negative Predictive Value (NPV)	$\text{NPV} = \frac{TN}{TN+FN}$	Indicates the reliability of negative predictions.
F-Score	$\text{F-Score} = 2 \times \frac{\text{PPV} \times \text{TPR}}{\text{PPV} + \text{TPR}}$	Represents the harmonic mean of precision and recall for balanced classification assessment.

Table 4. Baseline Performance of Machine Learning Models

Model	Accuracy	Sensitivity	Specificity	PPV	NPV	F-Score
Random Forest	0.9082	0.8936	0.9216	0.9130	0.9038	0.9032
Support Vector Machine	0.8737	0.8602	0.8857	0.8696	0.8774	0.8649
Naive Bayes	0.8515	0.8298	0.8704	0.8478	0.8545	0.8387
Decision Tree	0.8333	0.8085	0.8545	0.8261	0.8393	0.8172

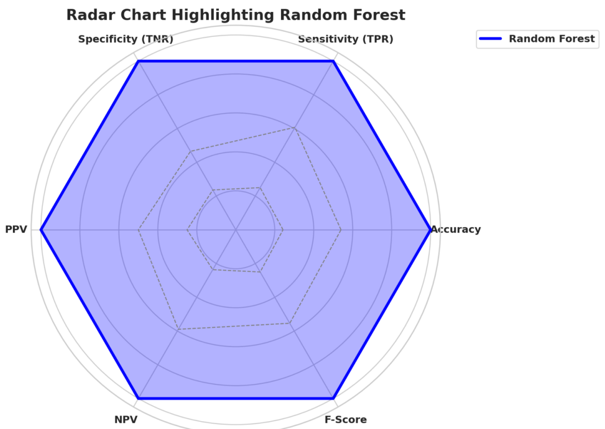

Figure 4. Random Forest Performance Radar Chart

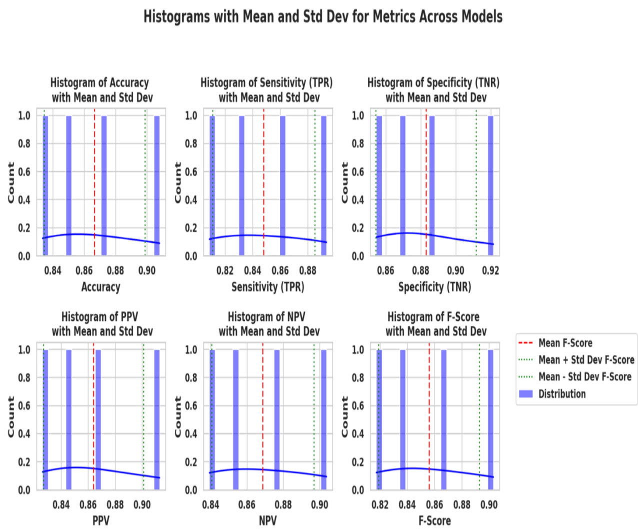

Figure 5. Histograms with Mean and Standard Deviation for Metrics Across Models

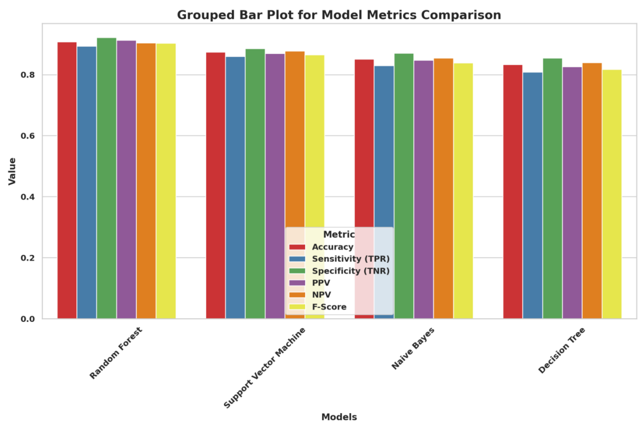

Figure 6. Grouped Bar Plot for Model Metrics Comparison

and Negative Predictive Value (NPV) and F-Score is presented in Figure 5. The addition of the indicators of mean (red dashed line) and standard deviation (green dotted lines) makes it possible to visually determine the central tendency and the dispersion of each of the indicators, which makes it easier to compare the model stability and reliability.

The analytical comparison of various machine learning models gives a better insight into the performance of these models in respect to various measures. Figure 6 depicts a grouped bar plot that demonstrates the performance of four classification algorithms, namely: Random Forest, Support Vector Machine, Naive Bayes and Decision Tree, based on six evaluation metrics, that include: Accuracy, Sensitivity (True Positive Rate), Specificity (True Negative Rate), Positive Predictive Value (PPV), Negative Predictive Value (NPV) and F-Score. Such visualization enables the efficient comparison of the strengths and weaknesses of each model, the Random Forest model can be shown to be the one that has better scores in most of the metrics, though.

4.2 Optimized Model Analysis

In this section, a comparative analysis of the Random Forest model optimized using different metaheuristic algorithms is provided. The main aim of this analysis is to determine the performance of the proposed Waterwheel Plant Algorithm (WWPA) against other popular optimization methods, such as Genetic Algorithm (GA), Particle Swarm Optimization

(PSO), and Grey Wolf Optimizer (GWO). Random Forest hyperparameters and feature selection were optimized by each of the methods of optimization with the purpose of improving the predictive power of the model and reducing its computational cost. Some evaluation metrics were used in order to strictly compare the models performance optimization: Accuracy, Sensitivity (True Positive Rate, TPR), Specificity (True Negative Rate, TNR), Positive Predictive Value (PPV), Negative Predictive Value (NPV), and F-Score. These metrics are an overall measurement of the effectiveness of the classification, which shows the ability of the models to discriminate both between depressive and non-depressive students with high accuracy and reliability.

Table 5. Performance Comparison of Random Forest Optimized with Different Metaheuristic Algorithms

Model	Accuracy	Sensitivity	Specificity	PPV	NPV	F-Score
WWPA + Random Forest	0.9577	0.9502	0.9644	0.9604	0.9553	0.9553
GA + Random Forest	0.9399	0.9331	0.9457	0.9375	0.9419	0.9353
PSO + Random Forest	0.9287	0.9169	0.9387	0.9271	0.9301	0.9220
GWO + Random Forest	0.9186	0.9048	0.9301	0.9149	0.9216	0.9098

As it can be seen in Table 5, the combination of the proposed WWPA and the Random Forest model demonstrated the best performance in all measurements. The model recorded 95.77% accuracy, which is significantly higher than the traditional optimization methods. The Sensitivity value of 95.02%is a confirmation that the model was excellent in identifying the students who are suffering depressive symptoms, and the Specificity of 96.44% shows how precise the model was in identifying the non-depressed ones. In addition, PPV of 96.04% and NPV of 95.53 indicate that the model is robust in both the positive and negative predictive reliability. The value of F-score 95.53% shows that there is a good trade-off in terms of the precision and the recall, and this is what makes the WWPA-optimized Random Forest the best configuration. Comparatively, the GA + Random Forest model had the highest accuracy at 93.99%, which is strong and stable, though the precision and sensitivity of the model are a little less than that of WWPA. The GA evolutionary strategy is also effective at search space exploration, however, the convergence rate was slower and it sometimes experienced premature convergence to local optima. However, the findings indicate that GA is still an effective and reliable optimization technique to make fine-tuning of model parameters. The PSO + Random Forest model had a 92.87% accuracy. Its particle based approach allowed it to search the space in a flexible, adaptive manner but was at a disadvantage in that it also tended to miss optimal regions as it entered a later iteration. The PSO model was not the worst in general (F-Score: 92.20%), but was not as balanced in terms of the sensitivity and Specificity as WWPA and GA. The lowest accuracy (91.86%) was registered with the GWO + Random Forest model compared to optimization techniques. Even though GWO successfully uses the hierarchical hunting behaviour of grey wolves to balance the exploration and exploitation, it approached global optima more sluggish and sometimes did not reach global optima, rather it was stagnant in suboptimal areas. Nevertheless, it had reasonable predictive values regardless of these constraints, indicating the inherent power of the nature-inspired algorithms in improving the work of machine learning. The comparative results show that there are a number of important aspects. First of all, the integration of metaheuristic optimization boosted the predictive performance of the Random Forest model by a significant level, as compared to its original performance. Second, WWPA optimization was the most stable one and yielded the best results on all measures of all tested approaches. It is biologically adapted with the feeding mechanism of waterwheel plant that allows changing between exploration and exploitation. This increases the effectiveness of the search in avoiding local minima to produce better quality solutions. The issue that F-scores and equal Sensitivity-Specificity ratios augmented under the WWPA demonstrates that the algorithm has the ability to regulate across various patterns of student behavior, psychological predictors and academic results. WWPA also optimizes the random forest parameter space to determine the most valuable features and in the process, the reliability and interpretability of the predictive model is enhanced. This dual advantage is advantageous to predicting mental health outcomes since the ability to be clear is critical to the ethical and transparent deployment of AI-based systems. On a larger scale, the higher level of predictive accuracy and interpretability of uniting WWPA and Random Forest could play an important role in educational and mental health contexts. Depression diagnosis of students at an early age can be kept on an accurate diagnosis, followed by intervention and individual approach to psychology. In addition, improved model transparency will enable educational workers, counselors, and researchers to identify the significant psychosocial and behavioral predictors of depression, which will facilitate the application of data in the prophylactic model of mental health. Beyond the academic setting, the present study points to the increasing possibilities of combining ensemble learning with bio-inspired optimization and indicating an implementation of more complex classification tasks in healthcare, behavioral science, and education. This entails the synergy of the interpretability of the Random Forest model, and the adaptive optimization mechanism of the WWPA, which indicates that a powerhouse can be realized, with good performance and believability. These enhancements are the foundation of future studies concerning the designing of intelligent, explainable, and ethically aware predictive systems to be applied in the activities of the human mind. Both the histogram and the kernel density estimation (KDE) were used to acquire further information on the distribution and concentration of the model performance metrics. The results of Figure 7 show the distribution of six critical evaluation metrics, namely Accuracy, Sensitivity (True Positive Rate), Specificity (True Negative Rate), Positive Predictive Value (PPV), Negative Predictive Value (NPV), and F-Score, over the models that were tested. These visualizations indicate that the distributions of all metrics are fairly compact, which suggests that there is high consistency in the models and small variation in the performance of different evaluation runs.

In order to determine the impact of different optimization methods on the success of the Random Forest classifier, a number of hybrid models were made and compared. Figure 8 optimal RF provides the results of four hybrid configurations of WWPA+Random Forest, GA+Random Forest, PSO+Random Forest, and GWO+Random Forest when evaluated through six significant evaluation measures, including Accuracy, Sensitivity (True Positive Rate), Specificity

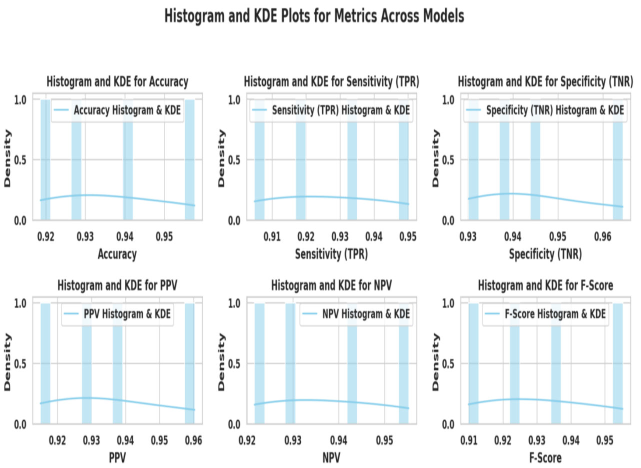

Figure 7. Histogram and KDE Plots for Metrics Across Models

(True Negative Rate), Positive Predictive Value (PPV), Negative Predictive Value (NPV), and F-Score. The visualization also shows the steady performance of all the models, where WWPA+Random Forest has slightly better results in most measures, indicating that the developed model can be effectively used to increase model robustness and predictive balance.

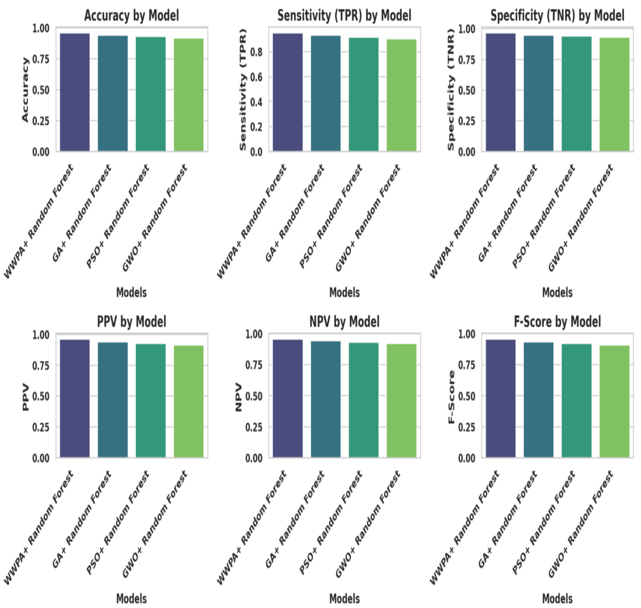

Figure 8. Comparison of Optimized Random Forest Models Across Multiple Metrics

In order to explore the distributional properties and consistency of the model performance further, the model performance in terms of each of the evaluation metrics was discussed using kernel density estimation (KDE) plot. Figure 9 shows the KDE curves of six key performance measures of Accuracy, Sensitivity (True Positive Rate), Specificity (True Negative Rate), Positive Predictive Value (PPV), Negative Predictive Value (NPV), and F-Score, generalized over all models. The fact that the density profiles in the plots are smooth show that most of the models obtained well-clustered performance values, which implies strong reliability and low error in the predictive capability of different algorithms used.

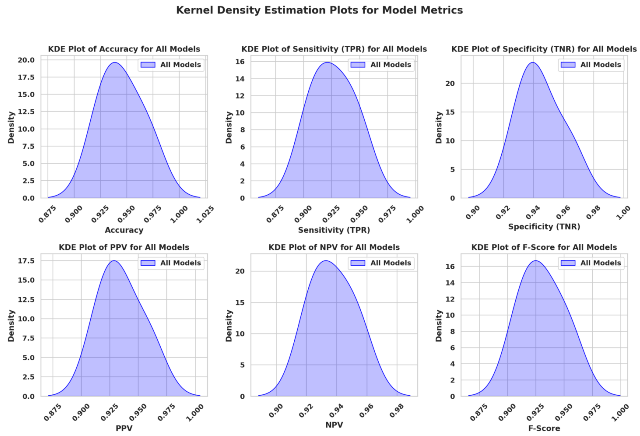

Figure 9. Kernel Density Estimation (KDE) Plots for Model Metrics

5. DISCUSSION

The results of the paper at hand indicate the importance of Machine Learning (ML) and metaheuristic optimization algorithms in improving the predictability and understanding of depression among students. The findings indicated that the Random Forest model was more precise, sensitive, and specific than the basic classifiers, including Support Vector Machine, Naive Bayes, and Decision Tree, and the F-score. This was a high-performance ensemble in the sense that it could sustain nonlinear associations, handle noisy data, and exhibit high-dimensional behavior. Besides, psychosocial and academic variables were included to provide a multidimensional approach, thereby enhancing the predictive ability of the models. This is due to the enhanced performance of the Waterwheel Plant Algorithms (WWPA) optimization (greater than 95%), which justifies its suitability for parameter tuning and the majority of informative feature selection, leading to more balanced and trustworthy classification. This boost implies that the biologically inspired algorithms to which advanced ensemble techniques can be applied can be used in practice and to more complex problems, such as mental health testing. In addition to predictive accuracy, the research will also broaden understanding of the contribution of psychosocial and behavioral determinants to mental health among institutions of learning. It is possible to determine the most significant elements of depression—such as sleep patterns, academic performance, and social interaction—by interpreting the Random Forest model, which produces dynamic, significant results in the psychological domain. It is also possible to refine the WWPA by increasing interpretability, thereby maximizing feature selection to target the most relevant predictors and eliminate redundancy and overfitting. The Particle Swarm Optimization (PSO), when compared with other optimization algorithms, including the Genetic Algorithm (GA), the Particle Swarm Optimization (PSO) and the Grey Wolf Optimizer (GWO), was more convergent and more stable to be evaluated. This fact supports the hypothesis that the dynamism of WWPA exploration versus exploitation cores yields better solutions even in complex, mutually reinforcing datasets. The adaptive search style of the algorithm is another possible application of computational intelligence in the study of behavioral data, as it is based on natural processes. In a broader context, the findings emphasize the misrepresentations that occur when combining Machine Learning and metaheuristic optimization to build educational and mental health analytics. Maximizing accuracy about WWPA is RFA, and it is more accurate in predicting depression. Nevertheless, it is easier to interpret, more dependable, and more ethically transparent, which is a prerequisite for applying AI-based mental health solutions in practice. This capability to pinpoint the at-risk students will allow institutions to devise timely, precise, and therefore offset the impact of the undiagnosed cases of depression on the student in terms of the academic and emotional condition. Besides, the methodological paradigm introduced in this work can be applied to other psychological and educational processes, as the tendency to develop innovative people-centred systems persists. Further studies can include larger data, multimodal behavioural data (such as physiological or social media feedback), and longitudinal tracking to make the models more generalizable to contexts and more relevant in real time. Finally, and most importantly, the implementation of AI, psychology, and optimization in the study is a mandatory measure towards developing predictive mechanisms that are technically efficient, ethically responsible, and human-welfare-oriented.

6. CONCLUSION AND FUTURE WORK

This paper demonstrates that an ensemble learning methodology coupled with a biologically inspired metaheuristic can improve the prediction of depression among university students. The experiment involved comparing psychosocial, demographic, and academic data for 100 Computer Science students using four baseline classifiers —Random Forest, Support Vector Machine, Naive Bayes, and Decision Tree — to determine their predictive accuracy. The Random Forest model was the most stable and accurate across all evaluation measures, indicating that it can be used with multidimensional data. The model with the best predictive accuracy, sensitivity, specificity and F-Score was the Random Forest model with the addition of the proposed Waterwheel Plant Algorithm (WWPA) to select features and hyperparameter optimization to the algorithm rather than the Genetic Algorithms (GA), Particle Swarm Optimization (PSO) or the Gray Wolf Optimizer to the algorithm (GWO). In addition to enhancing predictive modeling, this integration did not compromise interpretability, as it made the psychosocial factors in the case of depression more explicit. The findings therefore support the view that WWPA is an acceptable approach for optimizing a high-performing mental health prediction model with strong generalization, balance, and reliability. There are several avenues for extending and improving this study in the future. Future studies should be designed to use the proposed methodology with larger, multi-institutional, and cross-cultural samples to improve generalizability and further control for sampling bias. The longitudinal data would also enable changes in the mental health of the students to be monitored over a specific period of time and improve the ability to identify trends in depression early. Also, it can be complemented with multimodal data, such as wearable sensor data, smartphone usage patterns and learning management system logs, to feed the predictive model with real-time behavioral and physiological data. Other improvements could enhance the robustness and reliability of the model, including probabilistic calibration, uncertainty quantification, and the use of explainable AI methods such as SHAP and counterfactual reasoning to provide transparent, interventionistic explanations. In practice, issues of equity, privacy, and ethical conduct in education and clinical work will be given priority in future work. However, sensitive student data can be protected in learning applications while maintaining analytical value through privacy-preserving methods such as federated learning or differential privacy. Further, the potential of adaptive optimization techniques such as WWPA, with either Bayesian or local search, could be extended to yield a faster-converging, more scalable optimization scheme. Lastly, the optimized model can be deployed in a real-world academic environment through counselor-in-the-loop systems to provide ongoing model evaluation, model drift detection and proof of impact in the real world. By following these future directions, the proposed RF-WWPA framework can become a trustworthy, ethical, human-centered decision-support tool with the potential to make a difference in early mental health intervention to improve students' well-being in educational institutions.

DATA AVAILABILITY STATEMENT

The smart home energy consumption and weather data used in this study are publicly available at: https://www.kaggle.com/datasets/mdismielhossenabir/psychosocial-dimensions-of-student-life/data.

REFERENCES

[1] P. Varma, M. Junge, H. Meaklim, and M. L. Jackson, "Younger people are more vulnerable to stress, anxiety and depression during covid-19 pandemic: A global cross-sectional survey," *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, vol. 109, p. 110236, 2021.

[2] M. Ciharova, P. Cuijpers, Y. Amanvermez, H. Riper, A. M. Klein, F. Bolinski, L. M. de Wit, C. M. van der Heijde, R. Bruffaerts, S. Struijs, R. W. Wiers, and E. Karyotaki, "Use of tailoring features and reasons for dropout in a guided internet-based transdiagnostic individually-tailored cognitive behavioral therapy for symptoms of depression and/or anxiety in college students," *Internet Interventions*, vol. 34, p. 100646, 2023.

[3] S. Yasin, S. A. Hussain, S. Aslan, I. Raza, M. Muzammel, and A. Othmani, "EEG based major depressive disorder and bipolar disorder detection using neural networks: A review," *Computer Methods and Programs in Biomedicine*, vol. 202, p. 106007, 2021.

[4] A. A. Mentis, D. Lee, and P. Roussos, "Applications of artificial intelligence-machine learning for detection of stress: A critical overview," *Molecular Psychiatry*, vol. 29, no. 6, pp. 1882–1894, 2024.

[5] S. J. Pinto and M. Parente, "Comprehensive review of depression detection techniques based on machine learning approach," *Soft Computing*, vol. 28, no. 17, pp. 10 701–10 725, 2024.

[6] M. Squires, X. Tao, S. Elangovan, R. Gururajan, X. Zhou, U. R. Acharya, and Y. Li, "Deep learning and machine learning in psychiatry: A survey of current

progress in depression detection, diagnosis and treatment," *Brain Informatics*, vol. 10, no. 1, p. 10, 2023.

[7] A. M. Chekroud, J. Bondar, J. Delgadillo, G. Doherty, A. Wasil, M. Fokkema, Z. Cohen, D. Belgrave, R. DeRubeis, R. Iniesta, D. Dwyer, and K. Choi, "The promise of machine learning in predicting treatment outcomes in psychiatry," *World Psychiatry*, vol. 20, no. 2, pp. 154–170, 2021.

[8] P. Gooding and T. Kariotis, "Ethics and law in research on algorithmic and data-driven technology in mental health care: Scoping review," *JMIR Mental Health*, vol. 8, no. 6, p. e24668, 2021.

[9] N. Rost, E. B. Binder, and T. M. Brückl, "Predicting treatment outcome in depression: An introduction into current concepts and challenges," *European Archives of Psychiatry and Clinical Neuroscience*, vol. 273, no. 1, pp. 113–127, 2023.

[10] B. Akay, D. Karaboga, and R. Akay, "A comprehensive survey on optimizing deep learning models by metaheuristics," *Artificial Intelligence Review*, vol. 55, no. 2, pp. 829–894, 2022.

[11] R. Qasrawi, S. P. V. Polo, D. A. Al-Halawa, S. Hallaq, and Z. Abdeen, "Assessment and prediction of depression and anxiety risk factors in schoolchildren: Machine learning techniques performance analysis," *JMIR Formative Research*, vol. 6, no. 8, p. e32736, 2022.

[12] X. Zong, M. Lipowski, T. Liu, M. Qiao, and Q. Bo, "The sustainable development of psychological education in students' learning concept in physical education based on machine learning and the internet of things," *Sustainability*, vol. 14, no. 23, p. 15947, 2022.

[13] I. J. Ratul, M. M. Nishat, F. Faisal, S. Sultana, A. Ahmed, and M. A. A. Mamun, "Analyzing perceived psychological and social stress of university students: A machine learning approach," *Heliyon*, vol. 9, no. 6, p. e17307, 2023.

[14] Z. Ren, Y. Xin, J. Ge, Z. Zhao, D. Liu, R. C. M. Ho, and C. S. H. Ho, "Psychological impact of covid-19 on college students after school reopening: A cross-sectional study based on machine learning," *Frontiers in Psychology*, vol. 12, p. 641806, 2021.

[15] D. Opazo, S. Moreno, E. Álvarez-Miranda, and J. Pereira, "Analysis of first-year university student dropout through machine learning models: A comparison between universities," *Mathematics*, vol. 9, no. 20, p. 2599, 2021.

[16] Y. A. Alsariera, Y. Baashar, G. Alkawsi, A. Mustafa, A. A. Alkahtani, and N. Ali, "Assessment and evaluation of different machine learning algorithms for predicting student performance," *Computational Intelligence and Neuroscience*, vol. 2022, no. 1, p. 4151487, 2022.

[17] X. He, P. Chen, J. Wu, and Z. Dong, "Deep learning-based teaching strategies of ideological and political courses under the background of educational psychology," *Frontiers in Psychology*, vol. 12, p. 731166, 2021.

[18] Z. Xu and Q. Xu, "Psychological factors influencing successful music learning using deep learning technologies," *Interactive Learning Environments*, vol. 32, no. 8, pp. 4652–4662, 2024.

[19] C. Du, C. Liu, P. Balamurugan, and P. Selvaraj, "Deep learning-based mental health monitoring scheme for college students using convolutional neural network," *International Journal on Artificial Intelligence Tools*, vol. 30, no. 06n08, p. 2140014, 2021.

[20] J. Zhao and Y. Qin, "Perceived teacher autonomy support and students' deep learning: The mediating role of self-efficacy and the moderating role of perceived peer support," *Frontiers in Psychology*, vol. 12, p. 652796, 2021.

[21] J.-L. Hung, K. Rice, J. Kepka, and J. Yang, "Improving predictive power through deep learning analysis of k-12 online student behaviors and discussion board content," *Information Discovery and Delivery*, vol. 48, no. 4, pp. 199–212, 2020.

[22] Y. Ding, X. Chen, Q. Fu, and S. Zhong, "A depression recognition method for college students using deep integrated support vector algorithm," *IEEE Access*, vol. 8, pp. 75 616–75 629, 2020.

[23] A. Baba and K. Bunji, "Prediction of mental health problem using annual student health survey: Machine learning approach," *JMIR Mental Health*, vol. 10, no. 1, p. e42420, 2023.

[24] G. Tyulepberdinova, M. Mansurova, T. Sarsembayeva, S. Issabayeva, and D. Issabayeva, "The physical, social, and mental conditions of machine learning in student health evaluation," *Journal of Computer Assisted Learning*, vol. 40, no. 5, pp. 2020–2030, 2024.

[25] F. Ge, D. Zhang, L. Wu, and H. Mu, "Predicting psychological state among chinese undergraduate students in the covid-19 epidemic: A longitudinal study using a machine learning approach," *Neuropsychiatric Disease and Treatment*, 2020, online ahead of print.

[26] P. Chikersal, A. Doryab, M. Tumminia, D. K. Villalba, J. M. Dutcher, X. Liu, S. Cohen, K. G. Creswell, J. Mankoff, J. D. Creswell, M. Goel, and A. K. Dey, "Detecting depression and predicting its onset using longitudinal symptoms captured by passive sensing: A machine learning approach with robust feature selection," *ACM Transactions on Computer-Human Interaction*, vol. 28, no. 1, pp. 3:1–3:41, 2021.

[27] N. Mumenin, A. B. M. K. Hossain, M. A. Hossain, P. P. Debnath, M. N. Della, M. M. H. Rashed, A. Hossen, M. R. Basar, and M. S. Hossain, "Screening depression among university students utilizing GHQ-12 and machine learning," *Heliyon*, vol. 10, no. 17, p. e37182, 2024.

[28] M. I. Siraji, A. A. Rahman, M. M. Nishat, M. A. A. Mamun, F. Faisal, L. I. Khalid, and A. Ahmed, "Impact of mobile connectivity on students' wellbeing: Detecting learners' depression using machine learning algorithms," *PLOS ONE*, vol. 18, no. 11, p. e0294803, 2023.

[29] F. Al-Wesabi, H. Alsolai, A. Hilal, M. Hamza, M. Duhayyim, and N. Negm, “Machine learning based depression, anxiety, and stress predictive model during covid-19 crisis,” *Computers, Materials & Continua*, vol. 70, no. 3, pp. 5803–5820, 2021.

[30] A. B. Siddique and M. Chowdhury, “A machine learning-based approach to predict university students’ depression pattern and mental healthcare assistance scheme using android application,” *International Journal of Data Analysis Techniques and Strategies*, vol. 14, no. 2, pp. 122–139, 2022.