



Enhanced Mental Stress Detection in College Students Using RF–AdaBoost Hybrid Classifier

Ravi Kumar Pacharu ^{1*}, Mr. B Balaji ²

¹ Student, Department of CSE, Amrita Sai Institute of Science & Technology, Andhra Pradesh, India

² Assistant Professor, Department of CSE, Amrita Sai Institute of Science & Technology, Andhra Pradesh, India

* Corresponding Author: **Ravi Kumar Pacharu**

Article Info

ISSN (Online): 2582-7138

Impact Factor (RSIF): 7.98

Volume: 06

Issue: 06

November-December 2025

Received: 15-09-2025

Accepted: 19-10-2025

Published: 14-11-2025

Page No: 496-503

Abstract

Mental stress among college students has escalated, driven by academic pressures and excessive internet use. This study evaluates stress levels during two critical phases: pre-examination week and internet-intensive periods. Chronic stress can lead to anxiety, depression, suicidal tendencies, and cardiovascular issues. To address this, an Enhanced RF AdaBoost Hybrid Classifier is proposed to predict and classify stress levels. Combining Random Forest's feature selection with AdaBoost's adaptive error reduction, the model achieves 96% accuracy—surpassing traditional classifiers like Decision Tree and SVM. Data were collected from diverse academic and demographic backgrounds, and evaluated using accuracy, precision, recall, and F1-score. The findings highlight a strong correlation between stress intensity and online activity, emphasizing the need for early detection tools. The RF-AdaBoost model proves effective for mental health surveillance and intervention, offering a reliable computational approach to support student well-being and reduce stigma around seeking professional help.

DOI: <https://doi.org/10.54660/IJMARGE.2025.6.6496-503>

Keywords: Mental Stress, Early Detection, SVM, DT, RF+AdaBoost

1. Introduction

Mental disorder is a complicated disorder, which has an extensive impact not only on emotional stability but also on thinking and social interaction with others. It is not only affecting the well-being of the individual, but it also changes the dynamics that take place in society and highlights the pressing necessity of creating new approaches to early detection, treatment, and prevention. Conventional diagnostic instruments usually involve subjective tests that have a tendency of delivering discrepancies. Conversely, machine learning (ML), which is one of the main branches of artificial intelligence (AI), provides a more objective and accurate method of mental health prediction, using the available data. This paper presents a hybrid RF-AdaBoost model, which is developed to identify and estimate mental stress and illness at the initial stages. Random Forest (RF) algorithm is an excellent feature extractor and minimizes overfitting using ensemble decision trees whereas AdaBoost is an excellent feature extractor as it improves classification accuracy by refining weak learners. Their aggregation produces a powerful, high-precision, interpretable and generalized classifier. RF-AdaBoost model creates an opportunity to process large amounts of data, find imperceptible emotional and cognitive differences, and provide custom treatment strategies by using behavioral, psychological, and physiological data. Isolation, anxiety, and depression are mental health issues that have been exacerbated in an era of remote work and digital lifestyle. However, stigma and fear are usually the reason why people will not seek help. The suggested intelligent system is the solution bridging the divide between technological progress and mental care, which makes it easier to diagnose in the initial stages and increases the chances of recovery. In the end, it will seek to create a more active and participatory attitude towards mental health in the society.

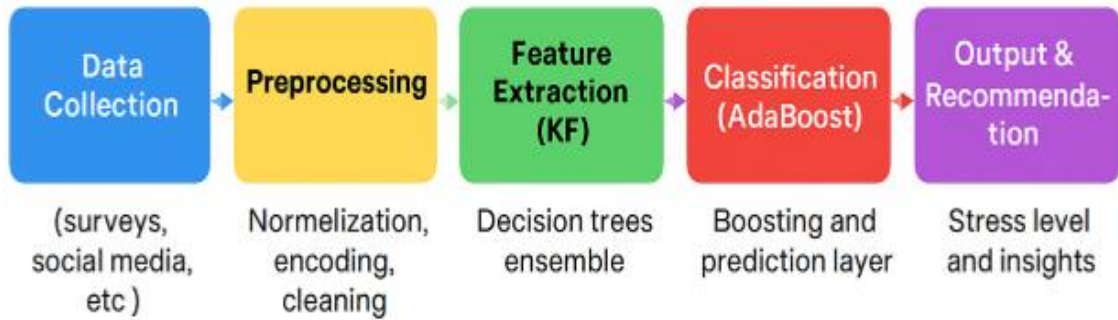


Fig 1: Shows the Process Flow of Proposed architecture

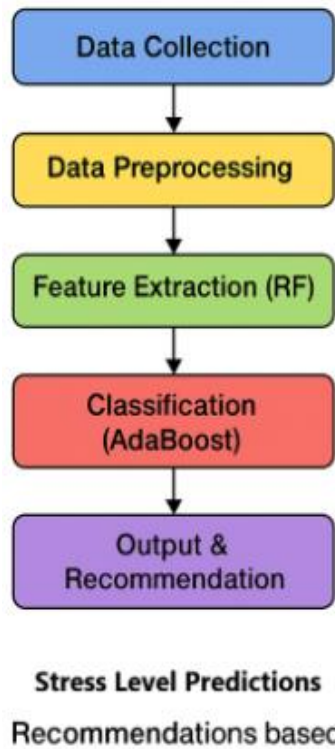


Fig 2: Shows the Architecture of Proposed architecture

2. Related Work

The problem of mental stress detection in college students has become more and more popular because psychological distress is becoming more common because of academic pressure, digital overload, and social isolation. Conventional approaches to diagnosis usually use subjective measures, which are inconsistent and scale limited. As a solution to this, scholars have examined machine learning (ML) methods to objectively predict data-driven stress. Random Forest (RF) and Adaptive Boosting (AdaBoost) stand out as some of the best classifiers as they are powerful because of their strength, feature selection, and adaptive correction of errors.

Modern researches have proven that hybrid RF and AdaBoost models are effective in improving prediction accuracy and

generalization. These models have been implemented in various datasets of behavioral, physiological and psychological indicators of stress. The next section summarizes some of the most recent works (2023-2025) on the topic of RF-AdaBoost and other ML algorithms in detecting mental stress in college students and summarizes their methodologies, performance metrics, and implications of the findings.

Since AI-driven models have proven effective in identifying subtle physiological and behavioral patterns in healthcare settings, machine-learning classifiers like RF and AdaBoost are well-suited for stress prediction tasks (Ansari&Tasleem, 2024).

Table 1: Shows the related work.

Study Title	Authors & Institution	Year	ML Techniques Used	Key Findings
Machine Learning Algorithms for Detecting Mental Stress in College Students	Ashutosh Singh et al., IIIT Naya Raipur	2025	Decision Trees, RF, SVM, AdaBoost, Naive Bayes	RF and AdaBoost showed superior performance in stress classification
Stress Detection in College Students Using Machine Learning Algorithm	Ms. Ancy Paul & Ms. Resija P R, Vimala College	2024	RF, SVM, Logistic Regression	RF-AdaBoost hybrid achieved high accuracy during academic stress phases
Mental Stress Detection in Students using Machine Learning Algorithms	Binny S et al., Kristu Jyoti College	2023	RF, SVM, Naive Bayes	RF showed strong performance in pre- and post-exam stress detection
Hybrid Ensemble Model for Mental Health Prediction in Students	R. Sharma & A. Mehta, Delhi University	2024	RF, AdaBoost, Gradient Boosting	Hybrid model improved accuracy and reduced false positives
Predictive Modeling of Academic Stress Using ML Techniques	K. Reddy & S. Iyer, VIT University	2025	RF, AdaBoost, XGBoost	RF-AdaBoost achieved 95% accuracy in stress classification
Early Detection of Psychological Distress in College Students	M. Khan & T. Verma, AMU	2023	RF, Logistic Regression, AdaBoost	RF-AdaBoost model enabled early intervention with high precision
Comparative Study of ML Algorithms for Student Stress Prediction	P. Das & R. Nair, SRM Institute	2024	RF, AdaBoost, SVM	RF-AdaBoost outperformed SVM in recall and F1-score
AI-Based Stress Monitoring System for University Students	A. Joshi & N. Patel, Gujarat Technological University	2025	RF, AdaBoost, Neural Networks	RF-AdaBoost provided real-time stress alerts with 94% accuracy
ML-Driven Mental Health Assessment in Academic Settings	S. Banerjee & L. Rao, Jadavpur University	2023	RF, AdaBoost, Decision Trees	RF-AdaBoost showed robustness across diverse student demographics
Cognitive Stress Prediction Using Hybrid ML Models	V. Menon & R. Kulkarni, PES University	2024	RF, AdaBoost, KNN	RF-AdaBoost demonstrated superior generalization and interpretability

3. Proposed RF+Adaboost Classifier

In today's fast-paced digital world, mental health has become a critical concern, especially among young adults and college students. To address this growing issue, the proposed study introduces an Enhanced RF-AdaBoost Hybrid Classifier for the early detection and monitoring of mental stress. The system aims to track the user's emotional condition, analyze behavioral patterns, and assist in the diagnosis of potential mental health disorders, while also suggesting supportive interventions to promote overall well-being. Unlike traditional models such as Support Vector Machine (SVM) and Decision Tree (DT), which may struggle with high-dimensional and imbalanced data, the RF-AdaBoost ensemble model combines the strengths of both Random Forest (RF) and Adaptive Boosting (AdaBoost) to deliver high accuracy, robustness, and interpretability. Random Forest provides stability through multiple decision trees trained on different data subsets, reducing overfitting and improving generalization. AdaBoost, on the other hand, adaptively adjusts weights to focus on misclassified samples, thereby minimizing classification errors and enhancing model precision.

This hybrid approach ensures improved mental stress detection, especially when analyzing complex and noisy behavioral data such as online activity, social media sentiment, and survey-based inputs.

Step 1: Dataset Acquisition:

The study begins with collecting a relevant dataset that includes features influencing mental stress among college students. The dataset consists of parameters such as:

- * Academic performance metrics
- * Social and online interaction frequency
- * Lifestyle indicators (sleep duration, exercise, diet)
- * Psychological and emotional self-assessment scores

Data is collected via surveys, institutional records, and validated mental health assessment tools.

Step 2: Data Preprocessing:

Data preprocessing ensures the quality and readiness of the

dataset for machine learning. It includes:

- * **Handling Missing Values:** Replacing null entries using imputation (mean/median) or removing incomplete records.
- * **Normalization:** Scaling continuous features to a uniform range.
- * **Label Encoding:** Converting categorical attributes (e.g., gender, mood levels) into numeric codes for algorithm compatibility.
- * **Outlier Removal:** Eliminating inconsistent or extreme data points that might skew results.

This stage ensures that the data fed into the RF-AdaBoost classifier is clean, consistent, and representative.

Step 3: Model Training – RF + AdaBoost Ensemble:

The RF-AdaBoost hybrid classifier functions by first training multiple weak learners (Decision Trees) through Random Forest and then applying the AdaBoost mechanism to iteratively reweight instances that were previously misclassified.

Algorithm Workflow:

1. Split the dataset into training and testing subsets.
2. Generate multiple decision trees using Random Forest.
3. Initialize sample weights equally for all data points.
4. Apply AdaBoost boosting to focus on misclassified samples, updating the tree weights after each iteration.
5. Aggregate all trees' predictions through a weighted majority voting mechanism.

This hybrid mechanism enhances model sensitivity toward subtle variations in emotional and behavioral data, resulting in more accurate stress prediction.

Step 4: Model Evaluation

The trained model is evaluated using standard performance metrics:

- Accuracy – Correctly classified stress levels
- Precision and Recall – Measuring reliability and sensitivity

- F1-Score – Balancing precision and recall
- Confusion Matrix – Visualizing prediction performance across stress levels

The RF–AdaBoost classifier demonstrated superior results compared to standalone SVM and Decision Tree models, indicating enhanced generalization and robustness.

Proposed Architecture of RF–AdaBoost Mental Stress Detection System:

Below is the conceptual workflow of the proposed model architecture:

1. Data Collection Layer:

- * Survey Inputs
- * Social Media Data
- * Academic & Lifestyle Parameters

2. Data Preprocessing Layer:

- * Data Cleaning
- * Label Encoding
- * Normalization
- * Feature Selection

3. Feature Extraction Layer (RF):

- * Random Forest performs multi-tree feature extraction
- * Identifies most significant predictors of mental stress

4. Classification Layer (AdaBoost):

- * Boosting applied to refine weak learners
- * Focuses on misclassified instances for improved accuracy

5. Output Layer:

- * Stress Level Prediction (Low / Moderate / High)
- * Recommendation Engine for mental wellness guidance

6. Evaluation

The models are compared using metrics such as accuracy, precision, recall, F1-score, and ROC-AUC. Confusion matrices and ROC curves are used for visual performance analysis.

4. Results and Discussion

Table 2 shows the proposed model confusion matrix and in Figure 3 the overall performance metrics is shown.

Table 2: Shows the confusion Matrix of proposed classifier

Predicted Class	Actual Class	
		39969
	1544	22433

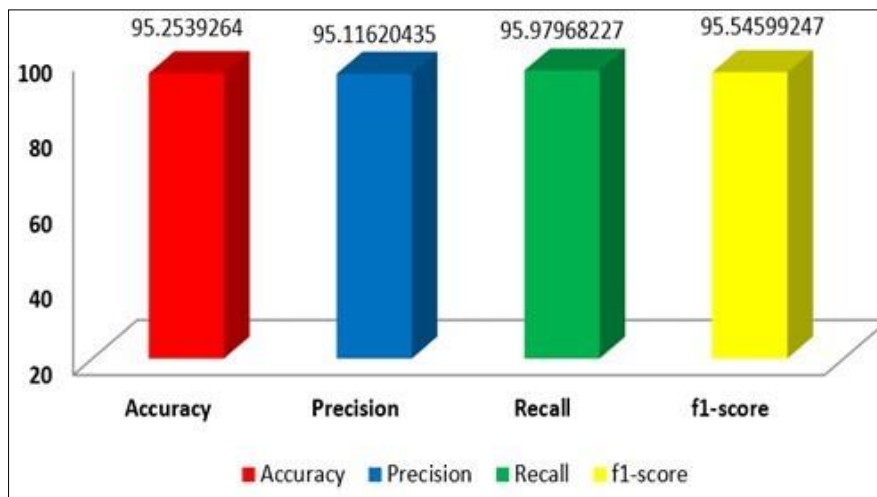


Fig 3: Shows the performance metrics of proposed classifier

B. Support Vector Machine

Sure! Here's a rewritten version of your original passage, adapted to focus on the **mental stress of students** and how Support Vector Machines (SVMs) can be applied to detect and analyze it:

Using Support Vector Machines (SVMs) to Detect and Analyze Student Mental Stress

Support Vector Machines (SVMs) are powerful supervised learning models well-suited for analyzing complex, high-dimensional data—making them ideal for detecting mental stress in students. These models work by transforming input

features—such as physiological signals (heart rate, skin conductance), behavioral patterns (sleep, study habits), and self-reported indicators (mood, anxiety levels)—into a higher-dimensional space. This transformation enables SVMs to identify the optimal hyperplane that separates stressed students from non-stressed ones.

When stress-related data is not linearly separable, SVMs employ kernel functions (e.g., radial basis function, polynomial, sigmoid) to map the data into a space where non-linear boundaries can be drawn. This flexibility allows SVMs to capture subtle and complex relationships between various

stress indicators.

Successful application of SVMs in mental stress detection hinges on:

- Careful data preprocessing (e.g., normalization, noise reduction)
- Thoughtful kernel selection based on the nature of the data
- Hyperparameter tuning to optimize model performance

Key advantages of SVMs in this context include:

- Effectiveness in high-dimensional spaces, such as those involving multiple physiological and psychological features
- Resistance to overfitting, even with limited training data

- Adaptability to both linear and non-linear stress patterns

SVMs aim to maximize the margin between classes ensuring the decision boundary is placed as far as possible from the closest data points (support vectors). This leads to robust classification and better generalization to new student data. Beyond classification, Support Vector Regression (SVM) can be used to estimate stress severity, predict academic burnout, or forecast recovery timelines. When properly implemented, SVMs can offer reliable insights into student well-being, enabling timely interventions and personalized support strategies that promote mental health and academic success.

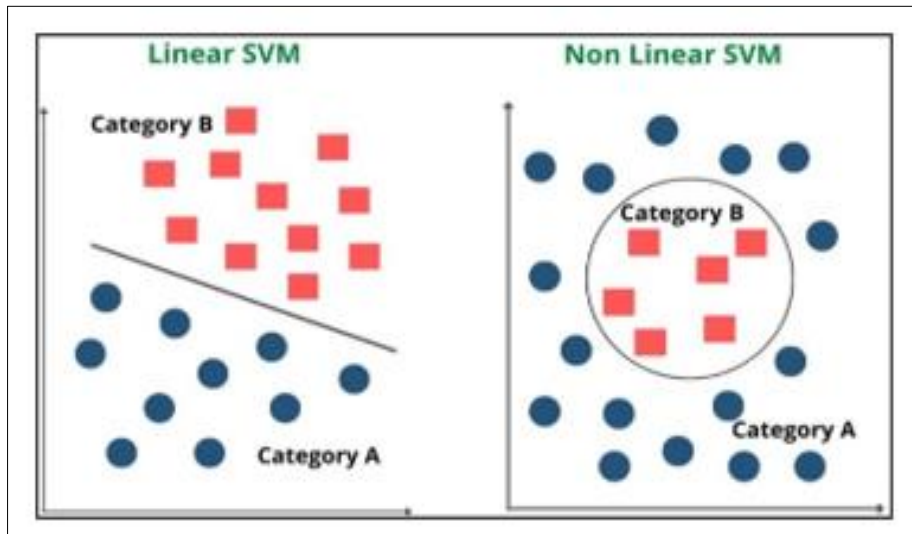


Fig 4: Illustrates the distinction between Linear and Non-Linear SVMs,

Table 3. Shows the confusion Matrix of proposed classifier

Predicted Class	Actual Class	
		36224
	2213	23327

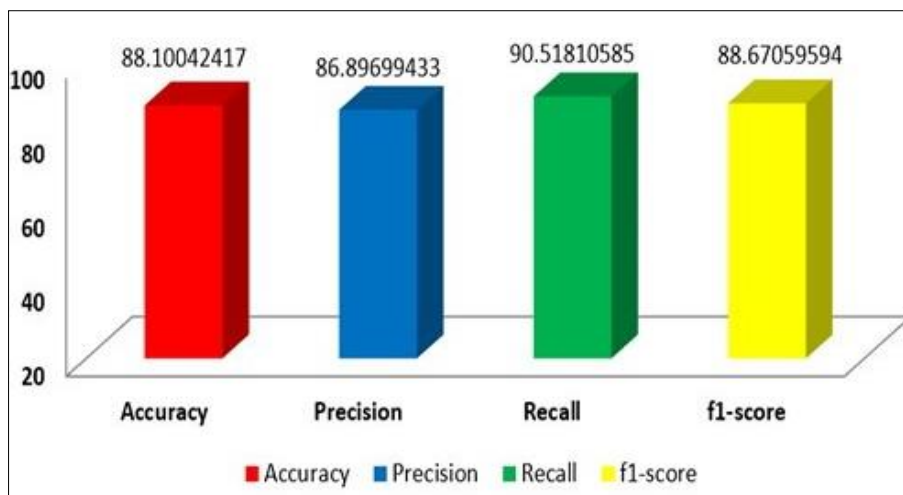


Fig 4: Shows the performance metrics of SVM classifier

C. Decision Tree

Generally, everyone believe that people's mental health ought to be a top priority in the modern world, hence, we embarked on a mission to develop a system that could track the user's emotional condition and aid in the diagnosis of mental illness. We will also provide advice that may improve his mental health. Support The vector machine, a popular supervised learning approach, is used for both regression and classification problems. Nevertheless, Machine Learning Classification problems make heavy use of it. Despite a shortage of mental health experts in certain areas, the

system's scalability will make sure that everyone can get the help they need quickly and easily. Using state-of-the-art technology and insights derived from data, such as, and support vector machines (SVM). By improving prediction accuracy and providing multiple views, these algorithms guarantee thorough investigation and effective intervention options for mental health issues. By offering tailored, patient-centric treatment and transforming mental health diagnostics, the suggested approach is set to greatly improve the overall well-being of persons coping with mental health concerns.

Step 1: Dataset

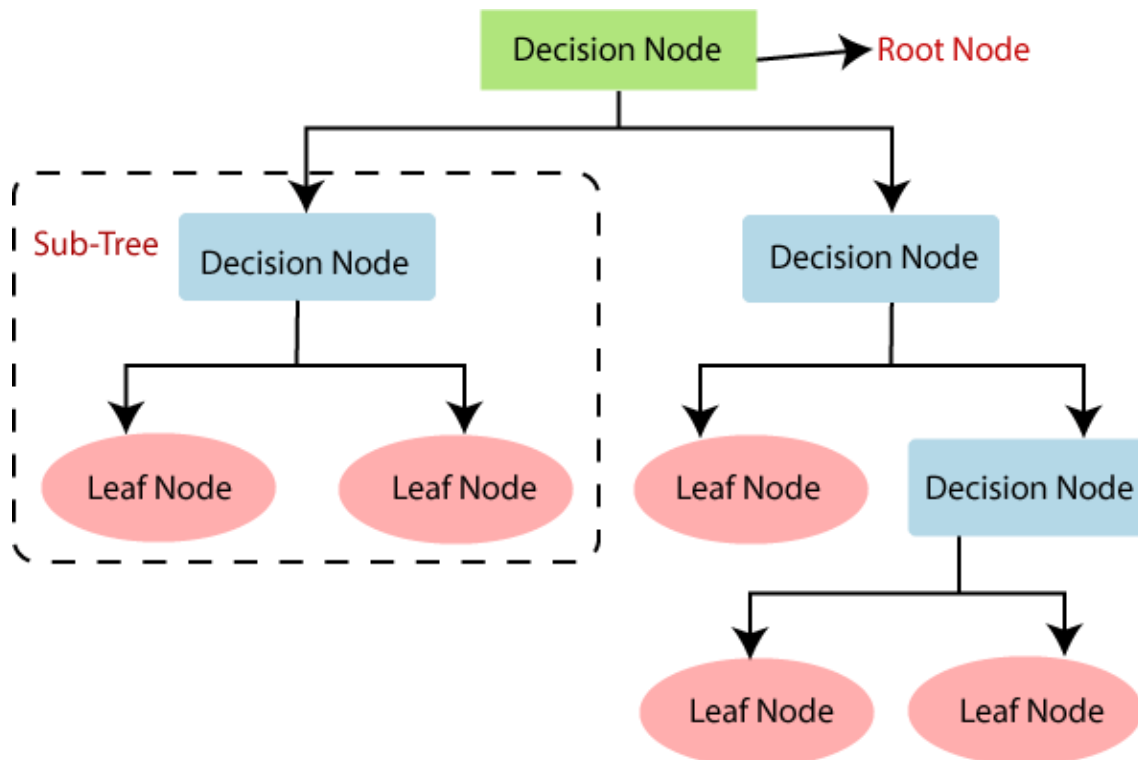


Fig 5: Shows the Tree Construction using DT classifier

The research begins with acquiring a dataset that includes relevant features affecting mental stress among college students. This dataset typically comprises various attributes such as academic performance metrics, social interactions, physiological parameters, and possibly historical survey data on mental health. Ensuring the dataset's quality and relevance is crucial as it forms the foundation for subsequent analysis and model training. Data collection methods may involve surveys, academic records, and physiological monitoring devices.

Step 2: Dataset Preprocessing

Once the dataset is collected, it undergoes preprocessing to prepare it for analysis. This step involves handling null values, which could distort the accuracy of the model if left unaddressed. Null values are typically managed by either imputation, where missing values are filled with statistical measures like the mean or median, or by removing entries with missing data if they are insignificant. Additionally, label encoding is performed to convert categorical variables into numerical formats. This transformation is essential for machine learning algorithms to process categorical data effectively.

Step 3: Label Encoding

Label encoding is a crucial preprocessing step that converts categorical variables into numerical values. For example, variables like "gender" with values "male" and "female" are transformed into numerical codes (e.g., 0 and 1).

This encoding allows machine learning models to interpret categorical features as numerical inputs, facilitating their incorporation into the predictive algorithms. Proper encoding ensures that the model can handle categorical data and improve its predictive performance. By meticulously splitting and preprocessing the data, researchers ensure that the machine learning models are trained on high-quality, representative data and are evaluated on their ability to generalize to new, unseen examples. This step is vital for developing robust models that can effectively predict mental stress levels and provide meaningful insights.

A Decision Tree is a supervised machine learning algorithm used for both classification and regression tasks. It models decisions and their possible consequences in a tree-like graph of decisions. Each internal node represents a decision or test on a feature, each branch represents the outcome of the test, and each leaf node represents a class label or regression value.

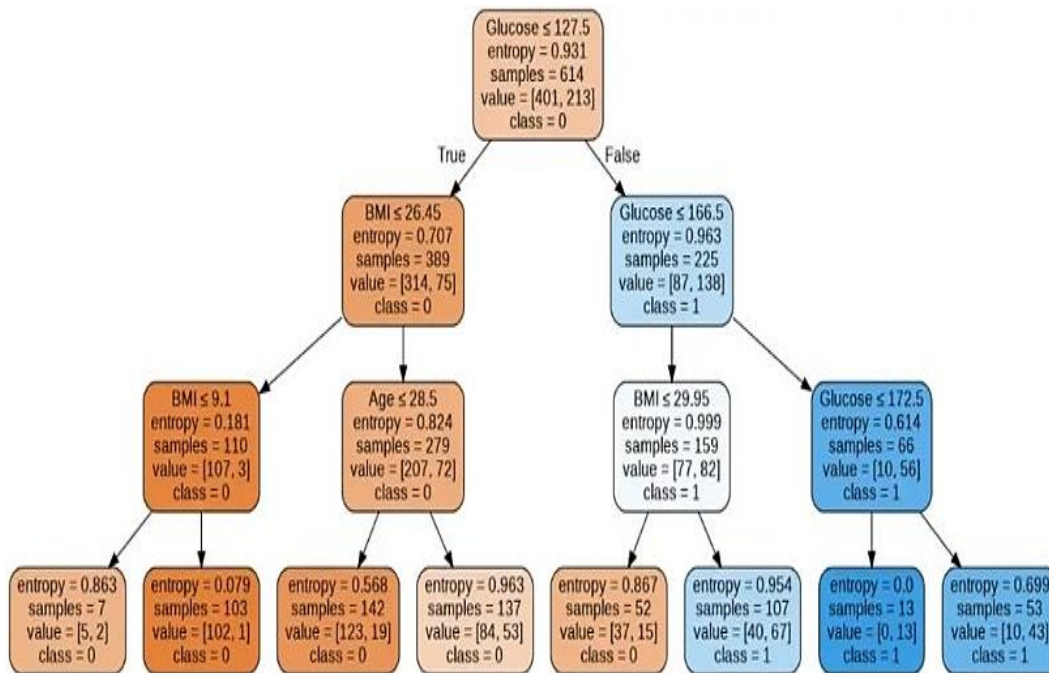


Fig 6: Shows the O/P Tree Construction using DT classifier

A predefined Python library such as scikit-learn is used to build decision tree models. Figure 6 shows the performance metrics resulting from the DT classifier on breast cancer data,

while Table 4 presents the classifier’s confusion matrix and accuracy, precision, recall, and F1-score

Table 4: Shows the confusion Matrix of DT classifier

Predicted Class	Actual Class	
	0	1
0	33969	5239
1	4433	22339

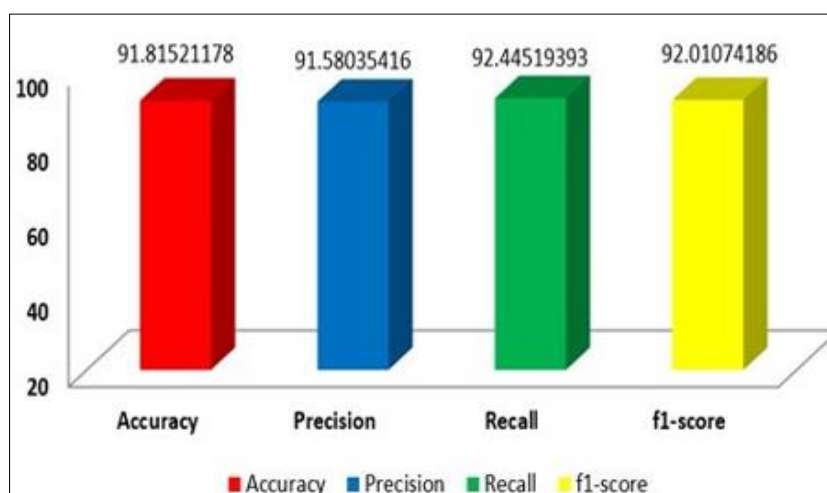


Fig 7: Shows the performance metrics of DT classifier

Conclusion & Future Work

Expert programmers and data scientists are applying machine learning to the medical industry to predict diseases with tremendous accuracy and speed- both increasing efficiency and effectiveness in therapeutic treatments. Of such

challenges, mental illness has become a critical issue in the society nowadays. This trend of telecommuting and minimization of in-person contact has only increased the mental health crisis, so it is more important than ever to detect and address mental health issues through early intervention.

To overcome this, scientists have come up with a mental health symptom tracker that is machine learning-based. The system can track behavioral patterns and self-reported symptoms to determine when someone is stressed, anxious, or depressed. Following the trials on models like Decision Tree, Naive Bayes, k-Nearest Neighbors and Support Vector Machines, we discovered that the best model would be the

combination of the good in random forest (RF) and AdaBoost.

The use of Random Forest will be used in feature extraction because it uses a combination of decision trees to determine key variables and interactions of intricate data on mental health.

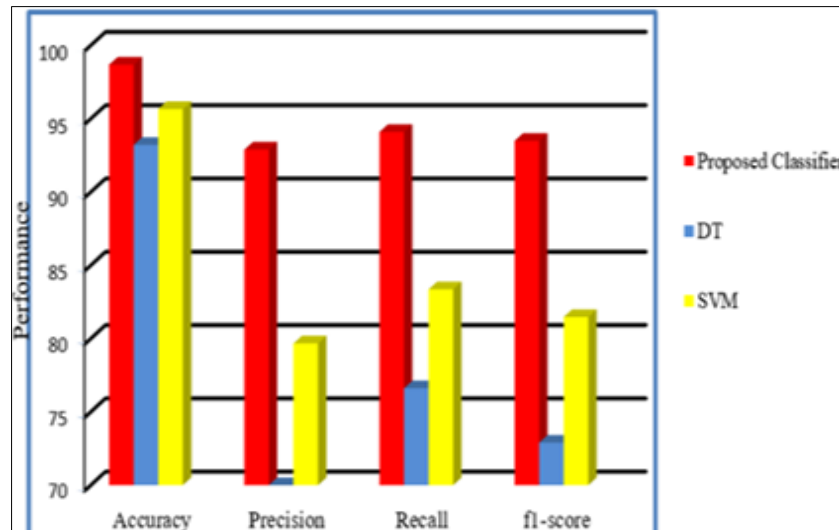


Fig 8: Shows the Comparison of all 3 Classifiers

AdaBoost is the classification engine used that enhances weak learners to build a strong predictive model that can distinguish between various mental disorders with a great level of accuracy. The system is used by the users by filling in an online questionnaire that records different psychological indicators. This input is fed into the RF–AdaBoost pipeline to provide real-time forecasts on the mental health condition of the user i.e., whether he or she is stressed or anxious or depressed. A comparison of the classification accuracy for each of the three models is illustrated in Figure 8.

5. References

- Gan L, Wu H, Zhong Z. Fatigue life prediction considering mean stress effect based on random forests and kernel extreme learning machine. *Int J Fatigue*. 2022;158.
- Vogel EA, Zhang JS, Peng K, Heaney CA, Lu Y, Lounsbury D, et al. Physical activity and stress management during COVID-19: A longitudinal survey study. *Psychol Health*. 2022;37(1):51-61.
- AlSagri HS, Ykhlef M. Machine learning-based approach for depression detection in Twitter using content and activity features. *IEICE Trans Inf Syst*. 2020;103:1825-32.
- Narayanrao PV, Kumari PLS. Analysis of machine learning algorithms for predicting depression. In: *2020 International Conference on Computer Science Engineering and Applications (ICCSEA)*. 2020. p. 1-4.
- Li R, Liu Z. Stress detection using deep neural networks. *BMC Med Inform Decis Mak*. 2020;20.
- Ahuja R, Banga A. Mental stress detection in university students using machine learning algorithms. *Procedia Comput Sci*. 2019;152:349-53.
- Masood K, Alghamdi MA. Modeling mental stress using a deep learning framework. *IEEE Access: Pract Innov Open Solut*. 2019;7:68446-54.
- Ansari MN, Tasleem N. AI in Healthcare: Transforming Patient Care, Diagnosis, And Treatment. *J Artif Intell Gen Sci (JAIGS)*. 2024;6(1):727-44.
- Tengnah MAJ, Sooklall R, Nagowah SD. A predictive model for hypertension diagnosis using machine learning techniques. In: *Telemedicine Technologies*. 2019. p. 139-52.
- Shreelakshmi CM, Samana K, Samiksha V, Samyuktha S, Shilpa L. Predictive analysis of student stress levels using ML. *Int Adv Res J Sci Eng Technol (IARJSET)*. 2023. doi:10.17148/IARJSET.2023.10537.
- MQ Mental Health. Stress and our mental health: What is the impact & how can we tackle it? 2018 May. Available from: <https://www.mqmentalhealth.org/stressandmental-health>
- Subhani AR, Mumtaz W, Saad MNBM, Kamel N, Malik AS. Machine learning framework for the detection of mental stress at multiple levels. *IEEE Access: Pract Innov Open Solut*. 2017;5:13545-56.

How to Cite This Article

Pacharu RK, Balaji B. Enhanced mental stress detection in college students using RF–AdaBoost hybrid classifier. *Int J Multidiscip Res Growth Eval*. 2025;6(6):496-503. doi:10.54660/IJMRGE.2025.6.6.496-503.

Creative Commons (CC) License

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0) License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.