ISBN: 978-628-96613-2-3. ISSN: 2414-6390. Digital Object Identifier: https://dx.doi.org/10.18687/LEIRD2025.1.1.880

Implementation of an Ensemble Stacking Model for Early Prediction of Depression in University Students

Huayllas Tirado, Sergio Alexander[1], Arellano Vilchez, Franklin Antoni[1], Briones Zúñiga, José Luis[1], and Huamán Aguirre, Arnold Anthony[1]
[1]Universidad Tecnológica del Perú, Perú, *U19221872@utp.edu.pe, U17102807@utp.edu.pe, C19980@utp.edu.pe, C19532@utp.edu.pe*

Abstract*– The increasing number of mental disorders among university students, particularly depression, highlights the urgency of early and effective interventions to prevent serious consequences on their academic and social performance. This study proposes an Ensemble Stacking-based model that combines Logistic Regression, Random Forest, and Decision Tree to predict depression, using variables such as gender, age, academic performance, and lifestyle factors. The featured model (Case 3) achieved 97.67% accuracy and 97.00% cross-validation, optimized by specific hyperparameter settings. Advanced preprocessing techniques, such as SMOTE for class balancing and PCA for dimensionality reduction, ensured the robustness and generalizability of the model. The methodology included an XGBoost metamodel, leveraging the diversity of the base models to improve accuracy and mitigate overfitting on complex data. These results demonstrate the effectiveness of the proposed approach for early detection of depression in educational contexts, providing practical tools that can facilitate timely and personalized interventions.

Keywords*-- Stacking assembly; detection of depression; university; prediction model; machine learning.

I. INTRODUCTION

In recent years, there has been a notable rise in the number of university students experiencing mental health disorders, particularly depression. This issue is especially concerning due to its strong association with anxiety and suicidal ideation [1], [2]. The increasing prevalence of these disorders underscores the urgent need for early detection strategies that enable timely interventions and adequate support for affected students [3], [4], [5], [6], [7]. Failure to identify symptoms promptly can lead to more severe conditions, exacerbating academic stress and hindering social interactions, ultimately impacting students' overall well-being.

The use of machine learning (ML) techniques in the field of mental health has grown significantly, especially in the early detection of disorders such as depression. Recent studies have demonstrated the effectiveness of ML in various contexts, such as the analysis of subjective well-being [8], [9] and suicide risk [10], [11]. Models such as SVM and Random Forest have achieved accuracies of over 90% in identifying indicators of anxiety and depression, using specialized patient health questionnaires [12], [13] and EEG recordings [14], [15]. These findings underscore the potential of ML in providing reliable, data-driven tools for mental health assessment, enabling earlier interventions and more personalized support for individuals at risk.

Recent research also highlights the relationship between higher education and increased depressive symptoms, especially during the COVID-19 pandemic, which intensified psychological stress and the prevalence of mental disorders [16], [17], [18]. These investigations demonstrate ML's ability to integrate multiple data sources, such as emotional interactions and physiological characteristics, with models such as Gradient Boosting and CNN achieving accuracies close to 98% [19], [20].

Despite significant advancements in the application of ML for mental health, much of the existing research relies on individual algorithms, often yielding inconsistent accuracy levels. This limitation underscores the need for more comprehensive approaches that combine the strengths of multiple models to enhance predictive performance. This article addresses this gap by implementing a hybrid modeling strategy to predict depression among university students. By integrating Logistic Regression, Random Forest, and Decision Tree models, the study aims to achieve improved accuracy and generalization. The findings are expected to advance early detection strategies for depression in educational settings, enabling more effective and timely interventions to support students in need.

II. METHODOLOGY

In this section, we detail the base models used, the experimental hyperparameters, the metamodel implemented in the Ensemble Stacking and the metrics selected for performance validation.

A. Ensemble Stacking

Figure 1, shows the Ensemble Stacking technique that integrates predictions from base models such as Logistic Regression, Random Forest and Decision Tree, using an XGBoost metamodel to generate the final prediction [21]. This approach takes advantage of the diversity of the base models to improve robustness and reduce the risk of overfitting, which is crucial in mental health applications with complex and highly nonlinear data.

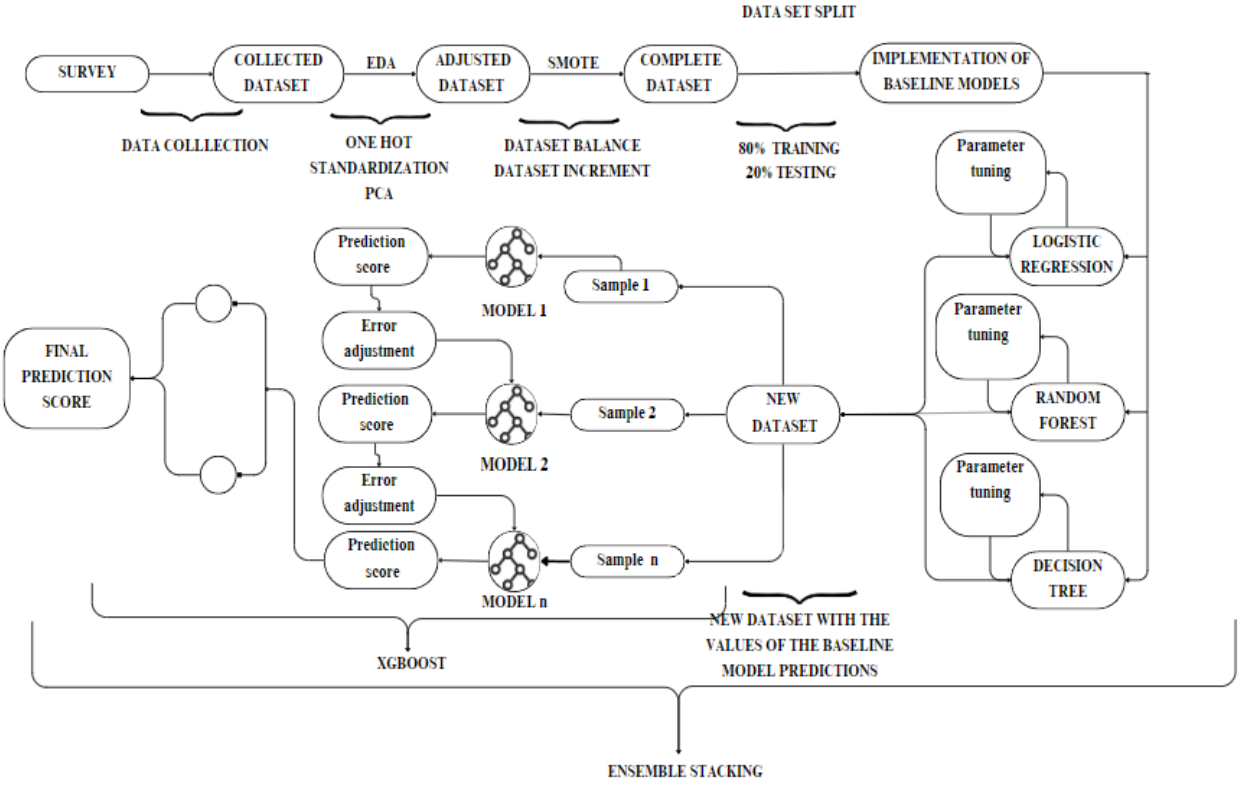

Fig. 1. Diagram illustrating the application process of the Ensemble Stacking method with Logistic Regression, Random Forest, and Decision Tree as base models, and XGBoost as the meta-model.

B. Base Models

1) Logistic Regression

This statistical model is used to predict the probability of an event by categorizing the outcome into two possible states. It calculates probabilities using a logistic function of the predictor variables, allowing for an easy interpretation of how each feature contributes to the predicted outcome [22]. This not only enables predictions about the students' depression status but also helps to understand the relative influence of predictor variables on the probability of depression. The statistical model for Logistic Regression is presented in equation (1).

$$p = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n)}} \qquad (1)$$

Where:
p is the probability of the event occurring.
β_0 is the intercept term.
$\beta_1, \beta_2, \ldots, \beta_n$ re the coefficients of the independent variables $x_1, x_2, \ldots, x_n$.

2) Decision Tree

This predictive model uses a tree structure to represent decisions and their possible consequences, including classification outcomes or regression values. At each node of the tree, a decision is made based on a single feature, making the model easy and intuitive to understand. This model is particularly useful for identifying critical variables that influence an outcome [23]. It can handle a large number of inputs regardless of variable dependencies and provide a ranking of variable importance, which is crucial for identifying the main predictors of depression.

For a decision tree, a common metric for choosing the best split is information gain. Information gain (ΔE) is presented in eq. (2).

$$\Delta E = E(S) - \sum_{i=1}^{k} \frac{|S_i|}{|S|} E(S_i) \qquad (2)$$

Where:

$E(S)$ is the entropy of the original dataset S.
S_i are the subsets after the split.

|S| is the size of the original dataset.
K is the number of resulting subsets from the partition.

3) Random Forest

This ensemble learning algorithm constructs multiple decision trees during training and produces the output based on the mode (classification) or average (regression) of the predictions from all trees. This method is highly effective for managing large datasets with multiple predictor variables, providing robustness against overfitting and improving overall prediction accuracy. Randomness is introduced by selecting a random subset of features at each tree split [24]. This method helps us visualize and clearly understand how decisions are made based on student characteristics.

Random Forest consists of many decision trees. For classification, the prediction ($\hat{y}$) is presented in equation (3), and for regression, the prediction ($\hat{y}$) is presented in equation (4).

$$\hat{y} = \text{mode}\{T_b(x)\}_{b=1}^{B} \quad (3)$$

$$\hat{y} = \frac{1}{B}\sum_{b=1}^{B} T_b(x) \quad (4)$$

Where:

$T_b(x)$ is the prediction of the b-th tree for input x.
B is the total number of trees in the forest.
.

C. Data Preprocessing

A rigorous data preprocessing was conducted on survey data obtained from a private university in Lima, Peru. The variables included Gender (X1), Age (X2), University Cycle (X3), Major (X4), Overall GPA (X5), Housing Situation (X6), Relationship Status (X7), Social Activity (X8), Economic Situation (X9), Employment Status (X10), Exercise (X11), Extracurricular Activities (X12), Mental Health Care (X13), Stress Level (X14), Sleep Hours (X15), Nutrition (X16), Height (X17), Weight (X18), Consumption of other substances (X19), Academic Social Situation (X20), Significant Loss (X21), Negative Thoughts (X22), and Depression (X23). Variable X3 is a quantitative numerical variable (int64), while the rest are qualitative variables (object).

This meticulous data preparation process ensures the integrity and applicability of the developed machine learning models. Table I details the Python libraries used in various stages of preprocessing and segmentation of the dataset.

D. Hyperparameters

The configuration of hyperparameters is a crucial aspect of deep learning as it enhances the results during the training of the model. Table II provides a description of the hyperparameters considered in this research.

TABLE I
PYTHON LIBRARIES USED IN DATA PREPROCESSING BEFORE IMPLEMENTING BASE MODELS

Library	Description
pandas	Data manipulation and analysis
sklearn.preprocessing	Data preprocessing for standardization and other transformations
sklearn.decomposition	Used for PCA (Principal Component Analysis)
imblearn.over_sampling	SMOTE (Synthetic Minority Over-sampling Technique) for oversampling
sklearn.model_selection	Tools for data splitting, cross-validation, and more
matplotlib.pyplot	Data visualization
seaborn	Statistical data visualization

TABLE II
MODELS AND HYPERPARAMETERS

Model	Hyperparameters
Logistic Regression	**solver**: Algorithm to address optimization problems.
	penalty: Specifies the norm used in the penalty.
	C: Inverse of regularization strength; smaller values specify stronger regularization.
Decision Tree	**criterion:** Function to measure the quality of a split in a decision tree, including 'gini' to ensure resulting groups are as uniform as possible, and 'entropy' to maximize information gain.
	max_depth: Maximum depth of the tree. Deeper trees learn finer details of the data and may overfit.
	min_sample_split: Minimum number of samples required to be a leaf node.
Random Forest	**n_estimators:** Number of trees in the forest. A larger number of trees can improve performance but increases computational cost.
	max_depth: Maximum depth of the trees.
	min_sample_leaf: Minimum number of samples required to split an internal node.

E. Classification Evaluation Metrics

The effectiveness of the integrated model was evaluated using the metrics described in Table III. Accuracy, which measures the proportion of true positives (TP) relative to all predicted positive outcomes (TP + FP), was used to determine the model's ability to correctly classify cases, complemented by other metrics to address class imbalance [25]. In addition to these metrics, learning curves were employed to visualize how training and validation accuracies evolve as the size of the training dataset increases or as training epochs progress. This helps to identify if the model suffers from overfitting (where performance on the training set is significantly superior to the validation set) or underfitting (where performance is inadequate on both sets), providing a clear visual perspective on the balance between learning and generalization [26].

TABLE III
DESCRIPTION OF METRICS AND THE LIBRARIES USED FOR MODEL VALIDATION

Metric	Formula
Accuracy	$\frac{TP + TN}{TP + TN + FP + FN}$
Precision	$\frac{TP}{TP + FP}$
Recall	$\frac{TP}{TP + FN}$
F1-Score	$2 * \frac{Precision * Recall}{Precision + Recall}$
Specificity	$\frac{TN}{TN + FP}$
Accuracy	$\frac{TP + TN}{TP + TN + FP + FN}$

A 5-fold cross-validation was implemented to ensure robust and generalizable model evaluation [27]. Finally, a confusion matrix and classification report provided a detailed performance analysis by class, showing true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN). Metrics such as precision, recall (which measures the proportion of true positives over the sum of true positives and false negatives, TP / (TP + FN)), and F1-score (the harmonic mean of precision and recall), were crucial for evaluating how the model handled imbalanced classes [28]. These values, the closer to 1, better indicated the model's performance in terms of the mentioned metrics.

III. RESULTS

A. Data preprocessing

The dataset was obtained through an anonymized online survey conducted at a private university in Lima, Peru. It was divided into feature (X) and label (Y) components, where Y represents the target variable indicating the presence or absence of depression. Likewise, during the preprocessing stage of the dataset, a cleaning technique was applied consisting of the elimination of records containing missing values. The data were then divided into 80% for training and 20% for testing, using stratified random sampling to maintain class proportionality in the training and test sets [29].

To correct the imbalance between the classes in the dataset, we applied the synthetic minority oversampling (SMOTE) technique. This strategy generated synthetic data for the minority class based on its nearest neighbors, thus balancing the classes and improving the generalization of the model [22]. Figure 2 illustrates the homogenization achieved by this process, increasing the total number of examples to 2000. Categorical variables were transformed into binary vectors by One-Hot coding, which allowed the models to detect patterns without assuming a hierarchy between categories [23]. After normalization of the numerical variables, Principal Component Analysis (PCA) was applied, selecting the 40 most important principal components from a total of 81variables maximizing the variance captured by 83.63%, as shown in Figure 3 and 4. This step simplified the representation of our data without compromising its essential integrity [24].

B. Experimental Setup

To assess the performance and stability of the model, different hyperparameter configurations for Logistic Regression, Random Forest, and Decision Tree models were explored through three case studies. Table IV details the experimental cases used to optimize the models.

C. Learning Curves

Figure 5, shows three learning curves corresponding to the analysis of the performance of the ensemble stacking method applied to predict the detection of depression in university students. In each case a different experiment is represented, in which different hyperparameters of the model are fitted, and allows observing the behavior both in the training data and in the cross-validation. The red lines reflect the score obtained in the training data, while the green lines show the performance of the model in the cross-validation, which evaluates the generalization capability of the model on unseen data.

In the first case, the training score remains consistently high, with values above 0.97 reaching a metric of 0.975, as can be seen in Table 5, but the accuracy metric in the cross-validation starts lower, around 0.78, and gradually increases as the iterations increase until reaching a value of 0.967. However, the distance between the two scores suggests that the model may be overfitted, i.e., it is learning the patterns in the training data too well, but is not generalizing optimally.

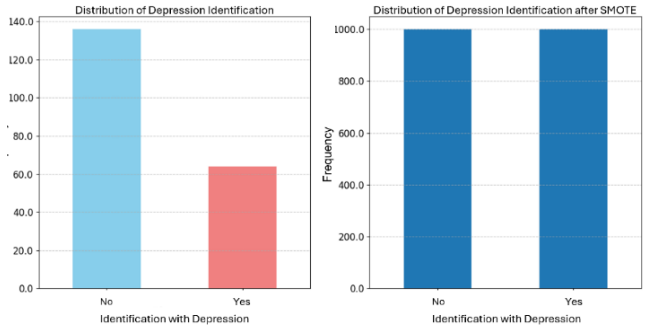

Fig. 2. Distribution of Responses on Depression Identification among University Students before and after SMOTE

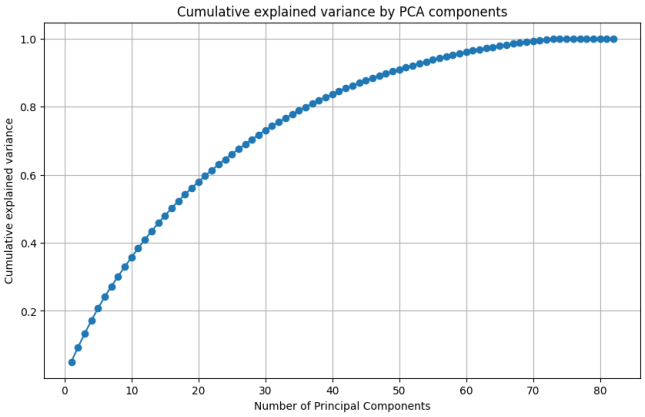

Fig. 3. the cumulative explained variance by PCA

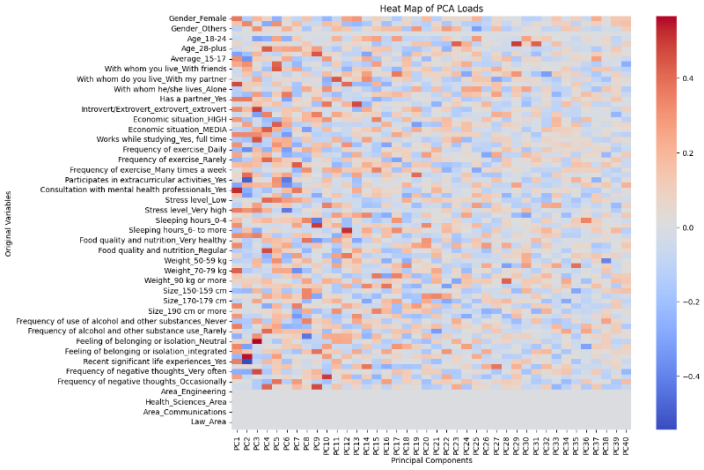

Fig. 4. the cumulative explained variance by PCA

The second experiment follows a similar pattern, with a high and stable training score of approximately 0.971, but the accuracy in cross-validation improves markedly compared to the first case with an approximate metric of 0.97. Although there is still a difference between the training and validation scores, the validation curve reaches 0.97 and shows less variability, indicating an improvement in the model's ability to generalize to unseen data.

TABLE IV
EXPERIMENTAL CASES

Case	Models	Hyperparameters	Value
1	Logistic Regression	Solver	'liblinear'
		Penalty	'l1'
		C	0.5
	Random Forest	n_estimators	150
		max_depth	20
		min_samples_split	5
	Decision Tree	criterion	'entropy'
		max_depth	20
		min_samples_leaf	2
2	Logistic Regression	solver	'liblinear'
		penalty	'l1'
		C	0.5
	Random Forest	n_estimators	100
		max_depth	10
		min_samples_split	2
	Decision Tree	criterion	'gini'
		max_depth	10
		min_samples_leaf	1
3	Logistic Regression	solver	'saga'
		penalty	'none'
		C	1.5
	Random Forest	n_estimators	200
		max_depth	None
		min_samples_split	10
	Decision Tree	criterion	'gini'
		max_depth	None
		min_samples_leaf	4

Finally, in the third case, the model shows the best performance of the three experiments. The training score remains high, as in the previous cases, but the accuracy in cross-validation grows faster and stabilizes around 0.976 with lower variability. This scenario indicates a better balance between the model fit and its generalization ability, suggesting that the hyperparameters selected in this experiment allow for superior performance..

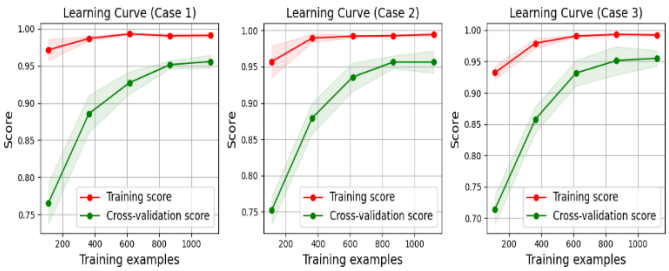

Fig. 5. Learning Curves for cases 1, 2 and 3

D. *Confusion Matrix*

Figure 6 shows the performance of the model in terms of correct predictions and errors for the “No Depression” and “Depression” classes. In all three cases, the model shows a high

level of accuracy, with a higher number of correct predictions in both classes. However, slight differences between the cases are observed in the classification errors. Likewise, case 3 is confirmed as the best performing model, standing out for having the lowest number of errors, with only 4 false negatives and 10 false positives, which reinforces the results of the learning curve, where this case showed the best generalization.

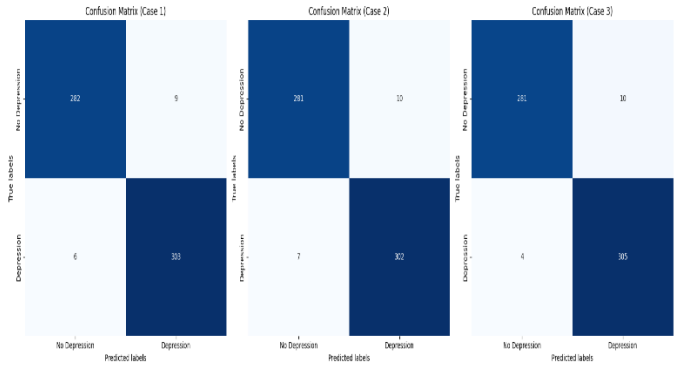

Fig. 6. Confusion Matrix for cases 1, 2 and 3

E. Accuracy, Precision, and Cross-Validation

According to the report in Table V, all three cases exhibit high performance with accuracy, precision, recall, and F1-Score values being quite similar between the "No" and "Yes" classes. Specifically, Cases 1 and 2 report an accuracy of 97.50% and 97.17%, showing remarkable consistency with slight variations in metrics between classes, indicating a balance in predictions for both classes. While case 3, with a slightly higher accuracy of 97.67%, maintains high and balanced metrics for both classes.

TABLE V
CLASSIFICATION REPORT OF CASES 1,2 AND 3

Case	Class	Precision	Recall	F1-Score	Cross-Validation	Support
1	No	0.9792	0.969	0.974	0.9635	291
	Yes	0.9712	0.980	0.975		309
	Accuracy	0.9750				600
2	No	0.9757	0.965	0.970	0.9700	291
	Yes	0.9679	0.977	0.972		309
	Accuracy	0.9717				600
3	No	0.9860	0.965	0.975	0.9700	291
	Yes	0.9683	0.987	0.977		309
	Accuracy	0.9767				600

Additionally, cross-validation metrics have been incorporated, revealing further precision regarding the model's stability and reliability:

- **Case 1:** Cross-validation for the "No" class shows an average of 0.9635, which is consistent with the other metrics.
- **Case 2:** The "No" class has a cross-validation of 0.9700, indicating robust performance despite data variations.
- **Case 3:** Cross-validation for the "No" class reaches 0.9700, reflecting effective generalization capability.

These results suggest not only consistency in accuracy and precision metrics but also in the model's ability to generalize, reinforced by cross-validation. With balanced supports of 291 for the "No" class and 309 for the "Yes" class in each case, the results highlight the model's strong capability to classify with high precision under different data subsets.

F. CONCLUSION

This study demonstrated the efficacy of the Ensemble Stacking method for predicting the presence of depression among Peruvian university students. Case 1 showed a good performance with an accuracy of 97.50% and a cross-validated accuracy of 96.35%, using Logistic Regression, Random Forest and Decision Tree models. Case 2 obtained an accuracy of 97.17% and a cross-validated accuracy of 97.00%. Case 3, with a slightly better performance than the previous cases, achieved an accuracy of 97.67% and a cross-validated accuracy of 97.00%. The model of case 3 stood out for its high accuracy and generalization capacity, which makes it suitable for practical application in educational environments. Therefore, the overall objective of generating evidence that the ensemble stacking method significantly improves the Early Prediction of Depression in College Students was achieved.

Evidence supports that the integration of machine learning techniques such as logistic regression, Random Forest, Decision Tree and XGBoost significantly improves the accuracy and adaptability of models to different data samples. This approach is crucial for the early detection of depression in university settings, enabling more timely and accurate interventions. The results contribute significantly to the scientific literature, providing a solid foundation for future research that aims to incorporate more complex models and additional variables.

REFERENCES

[1] Y. Zhai et al., "Machine learning predictive models to guide prevention and intervention allocation for anxiety and depressive disorders among college students," Journal of Counseling & Development, Oct. 2024, doi: 10.1002/jcad.12543.

[2] 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, p. 122, 2022, doi: 10.1504/IJDATS.2022.10049553.

[3] Y. Shen et al., "Detecting risk of suicide attempts among Chinese medical college students using a machine learning algorithm," J Affect Disord, vol. 273, pp. 18–23, Aug. 2020, doi: 10.1016/j.jad.2020.04.057.

[4] S. Ware et al., "Automatic depression screening using social interaction data on smartphones," Smart Health, vol. 26, p. 100356, Dec. 2022, doi: 10.1016/j.smhl.2022.100356.

[5] M. Zhang, K. Yan, Y. Chen, and R. Yu, "Anticipating interpersonal sensitivity: A predictive model for early intervention in psychological disorders in college students," Comput Biol Med, vol. 172, p. 108134, Apr. 2024, doi: 10.1016/j.compbiomed.2024.108134.

[6] M. M. Qirtas, E. Zafeiridi, E. B. White, and D. Pesch, "The relationship between loneliness and depression among college students: Mining data derived from passive sensing," Digit Health, vol. 9, Jan. 2023, doi: 10.1177/20552076231211104.

[7] H. G. B. dos Santos, S. R. Marcon, M. M. Espinosa, M. N. Baptista, and P. M. C. de Paulo, "Factors associated with suicidal ideation among university students," Rev Lat Am Enfermagem, vol. 25, no. 0, 2017, doi: 10.1590/1518-8345.1592.2878.

[8] R. B. King, Y. Wang, L. Fu, and S. O. Leung, "Identifying the top predictors of student well-being across cultures using machine learning and conventional statistics," Sci Rep, vol. 14, no. 1, p. 8376, Apr. 2024, doi: 10.1038/s41598-024-55461-3.

[9] Y. Wang, R. King, and S. O. Leung, "Understanding Chinese Students' Well-Being: A Machine Learning Study," Child Indic Res, vol. 16, no. 2, pp. 581–616, Apr. 2023, doi: 10.1007/s12187-022-09997-3.

[10] J. Cohen et al., "A Feasibility Study Using a Machine Learning Suicide Risk Prediction Model Based on Open-Ended Interview Language in Adolescent Therapy Sessions," Int J Environ Res Public Health, vol. 17, no. 21, p. 8187, Nov. 2020, doi: 10.3390/ijerph17218187.

[11] P. Marti-Puig, C. Capra, D. Vega, L. Llunas, and J. Solé-Casals, "A Machine Learning Approach for Predicting Non-Suicidal Self-Injury in Young Adults," Sensors, vol. 22, no. 13, p. 4790, Jun. 2022, doi: 10.3390/s22134790.

[12] A. Priya, S. Garg, and N. P. Tigga, "Predicting Anxiety, Depression and Stress in Modern Life using Machine Learning Algorithms," Procedia Comput Sci, vol. 167, pp. 1258–1267, 2020, doi: 10.1016/j.procs.2020.03.442.

[13] M. Bader, M. Abdelwanis, M. Maalouf, and H. F. Jelinek, "Detecting depression severity using weighted random forest and oxidative stress biomarkers," Sci Rep, vol. 14, no. 1, p. 16328, Jul. 2024, doi: 10.1038/s41598-024-67251-y.

[14] A. Al-Ezzi, A. A. Al-Shargabi, F. Al-Shargie, and A. T. Zahary, "Complexity Analysis of EEG in Patients With Social Anxiety Disorder Using Fuzzy Entropy and Machine Learning Techniques," IEEE Access, vol. 10, pp. 39926–39938, 2022, doi: 10.1109/ACCESS.2022.3165199.

[15] A. Arsalan and M. Majid, "A study on multi-class anxiety detection using wearable EEG headband," J Ambient Intell Humaniz Comput, vol. 13, no. 12, pp. 5739–5749, Dec. 2022, doi: 10.1007/s12652-021-03249-y.

[16] B. Gavurova, S. Khouri, V. Ivankova, M. Rigelsky, and T. Mudarri, "Internet Addiction, Symptoms of Anxiety, Depressive Symptoms, Stress Among Higher Education Students During the COVID-19 Pandemic," Front Public Health, vol. 10, Jun. 2022, doi: 10.3389/fpubh.2022.893845.

[17] S. Van de Velde et al., "Depressive symptoms in higher education students during the first wave of the COVID-19 pandemic. An examination of the association with various social risk factors across multiple high- and middle-income countries," SSM Popul Health, vol. 16, p. 100936, Dec. 2021, doi: 10.1016/j.ssmph.2021.100936.

[18] J. Lee, H. J. Jeong, and S. Kim, "Stress, Anxiety, and Depression Among Undergraduate Students during the COVID-19 Pandemic and their Use of Mental Health Services," Innov High Educ, vol. 46, no. 5, pp. 519–538, Oct. 2021, doi: 10.1007/s10755-021-09552-y.

[19] Md. S. Khan, N. Salsabil, Md. G. R. Alam, M. A. A. Dewan, and Md. Z. Uddin, "CNN-XGBoost fusion-based affective state recognition using EEG spectrogram image analysis," Sci Rep, vol. 12, no. 1, p. 14122, Aug. 2022, doi: 10.1038/s41598-022-18257-x.

[20] E. H. Houssein, S. Mohsen, M. M. Emam, N. Abdel Samee, R. I. Alkanhel, and E. M. G. Younis, "Leveraging explainable artificial intelligence for emotional label prediction through health sensor monitoring," Cluster Comput, vol. 28, no. 2, p. 86, Apr. 2025, doi: 10.1007/s10586-024-04804-w.

[21] K. Ayushi, K. N. Babu, N. Ayyappan, J. R. Nair, A. Kakkara, and C. S. Reddy, "A comparative analysis of machine learning techniques for aboveground biomass estimation: A case study of the Western Ghats, India," Ecol Inform, vol. 80, p. 102479, May 2024, doi: 10.1016/j.ecoinf.2024.102479.

[22] M. Meysami et al., "Utilizing logistic regression to compare risk factors in disease modeling with imbalanced data: a case study in vitamin D and cancer incidence," Front Oncol, vol. 13, Sep. 2023, doi: 10.3389/fonc.2023.1227842.

[23] H. Blockeel, L. Devos, B. Frénay, G. Nanfack, and S. Nijssen, "Decision trees: from efficient prediction to responsible AI," Front Artif Intell, vol. 6, Jul. 2023, doi: 10.3389/frai.2023.1124553.

[24] Z. Shao, M. N. Ahmad, and A. Javed, "Comparison of Random Forest and XGBoost Classifiers Using Integrated Optical and SAR Features for Mapping Urban Impervious Surface," Remote Sens (Basel), vol. 16, no. 4, p. 665, Feb. 2024, doi: 10.3390/rs16040665.

[25] S. Lin et al., "Using machine learning to develop a five-item short form of the children's depression inventory," BMC Public Health, vol. 24, no. 1, p. 1118, Apr. 2024, doi: 10.1186/s12889-024-18657-w.

[26] M. Meysami et al., "Utilizing logistic regression to compare risk factors in disease modeling with imbalanced data: a case study in vitamin D and cancer incidence," Front Oncol, vol. 13, Sep. 2023, doi: 10.3389/fonc.2023.1227842.

[27] E. Kee, J. J. Chong, Z. J. Choong, and M. Lau, "A Comparative Analysis of Cross-Validation Techniques for a Smart and Lean Pick-and-Place Solution with Deep Learning," Electronics (Basel), vol. 12, no. 11, p. 2371, May 2023, doi: 10.3390/electronics12112371.

[28] S. Orozco-Arias, J. S. Piña, R. Tabares-Soto, L. F. Castillo-Ossa, R. Guyot, and G. Isaza, "Measuring Performance Metrics of Machine Learning Algorithms for Detecting and Classifying Transposable Elements," Processes, vol. 8, no. 6, p. 638, May 2020, doi: 10.3390/pr8060638.

[29] C. Tang, D. Wang, A.-H. Tan, and C. Miao, "EEG-Based Emotion Recognition via Fast and Robust Feature Smoothing," 2017, pp. 83–92. doi: 10.1007/978-3-319-70772-3_8.