Journal of Information Systems and Informatics
Vol. 7, No. 2, June 2025 e-ISSN: 2656-4882 p-ISSN: 2656-5935
DOI: 10.51519/journalisi.v7i2.1087 Published By DRPM-UBD

Development of a Student Depression Prediction Model Based on Machine Learning with Algorithm Performance Evaluation

Penni Wintasari Simarmata[1], Putri Taqwa Prasetyaningrum[2]

[1,2]Information System, Mercu Buana University, Yogyakarta, Indonesia
Email: [1]penniwitasari1122@gmail.com, [2]putri@mercubuana-yogya.ac.id

Abstract

This research explores the implementation of machine learning to predict depression among university students using a dataset of 2.028 responses containing PHQ-9 scores and academic-demographic attributes. The research implements a structured modeling process involving feature selection, normalization, the model's efficacy was gauged through a suite of evaluate measures, encompassing accuracy, precision, recall, F1-score, The support vector machine (SVM) model's accuracy improved from 58.8% to 99.5% after hyperparameter tuning. This investigation lends itself to the advancement of a proactive identification framework, which hold potential for incorporation within collegiate mental well-being surveillance infrastructures. Future implementations may consider real-time models and expand data sources through digital counseling systems and behavioral analytics

Keywords: Classification Algorithm, Depression Prediction, Machine Learning, Model Development, Model Evaluation

1. INTRODUCTION

Psychological conditions, particularly those affecting students, have become a growing concern, gaining significant attention in recent years due to the pressures they face—academic demands, social expectations, and uncertainty about the future [1]. Among these, depression stands out as one of the most prevalent psychological disorders. This condition is often marked by persistent low mood, prolonged feelings of sadness, hopelessness, guilt, and worthlessness, which severely impact one's emotional well-being [2]. According to the World Health Organization (WHO), approximately 280 million people worldwide were living with depressive disorders in 2019, with a disturbing rise in prevalence among adolescents and young adults. A study revealed that 29% of university students suffer from anxiety disorders, and 25% experience depression, with varying degrees of severity, from mild to severe [3]. These mental health challenges significantly influence various aspects of cognition, emotion, and behaviour,

ultimately diminishing motivation to engage in daily activities or maintain a social life [4].

Despite the growing mental health crisis, many students hesitate to seek professional help, largely due to the stigma surrounding mental health issues. The fear of being labelled as "weak" or incapable often prevents them from accessing the psychological services they need [5]. In addition, the lack of mental health awareness and the shortage of professional resources further complicate efforts to provide timely and effective support, hindering early intervention.

In light of these challenges, recent advancements in Artificial Intelligence (AI), particularly in Machine Learning (ML), have presented new opportunities for detecting and diagnosing mental health issues. ML algorithms can analyse diverse data types—such as surveys, behavioural patterns, and psychological symptoms—offering a more objective, efficient, and scalable approach to mental health detection [6][7]. Prior research has demonstrated the potential of ML-based depression prediction models as effective tools for early intervention and detection [8], [9].

However, existing studies often suffer from key limitations, including a lack of interpretability, a failure to consider localized academic contexts, and an over-reliance on single algorithms without comparative analysis using balanced datasets [10][11]. These shortcomings diminish the practical applicability of the findings, especially in varied academic environments where student populations differ significantly in their psychological and academic backgrounds. These gaps highlight the need for a more holistic and systematic approach that incorporates transparency in model development, fairness in data representation, and robust performance evaluation across multiple algorithms.

In response to these challenges, this research proposes a comprehensive and comparative framework designed to address these limitations. The study compares the performance of six prominent ML algorithms Support Vector Machine (SVM), Logistic Regression, K-Nearest Neighbors (KNN), Random Forest, Decision Tree, and Naive Bayes using a public student dataset [12]. Additionally, the research employs Synthetic Minority Over-sampling Technique (SMOTE) to address class imbalance and uses Shapley Additive Explanations (SHAP) to improve model interpretability [13].

This research is motivated by the increasing need for early mental health detection among university students, particularly within the context of data-driven interventions. Accordingly, the study aims to answer two critical questions: First, which machine learning algorithm provides the most accurate prediction of depression in university students? Second, how do hyperparameter tuning and data

balancing techniques impact the overall performance of these predictive models? These questions serve as the foundation for evaluating the technical and practical effectiveness of different machine learning approaches in addressing mental health challenges within academic settings.

Ultimately, the goal of this research is to develop and evaluate predictive models that utilize various machine learning techniques to detect depression levels among university students. By doing so, the study also aims to identify the most effective algorithms in terms of accuracy and robustness, as well as the key features that contribute to predictive outcomes. In the broader context, this research aspires to support the integration of AI-driven systems into university-level mental health monitoring and intervention strategies, offering a new tool for promoting mental well-being in academic environments.

2. METHODS

This research employs a quantitative approach to address the challenges associated with depression detection in college students by leveraging machine learning (ML) algorithms. The primary goal of this study is to develop a prediction model that can effectively assess depression levels based on survey and questionnaire data [14]. The study is both descriptive and comparative, as it not only outlines the process of developing the prediction model but also compares the performance of several ML algorithms to identify the most optimal model [15][16].

2.1. Research Stages

The research follows a systematic series of steps, starting with a literature review and progressing through model evaluation and implementation. Each step contributes to constructing a reliable depression prediction system using machine learning techniques. The overall flow of the research is summarized in Figure 1, illustrating the methodology used in this study. The stages of the research process are as follows:

1) Literature Review

The study begins with an extensive literature review, which aims to explore existing works in the domain of depression detection using machine learning. This step provides foundational knowledge on previous methodologies, models, evaluation techniques, and the relevance of psychological instruments such as the Patient Health Questionnaire-9 (PHQ-9) [17]. Understanding previous approaches helps establish a theoretical framework for the current research and guides the choice of algorithms and data sources.

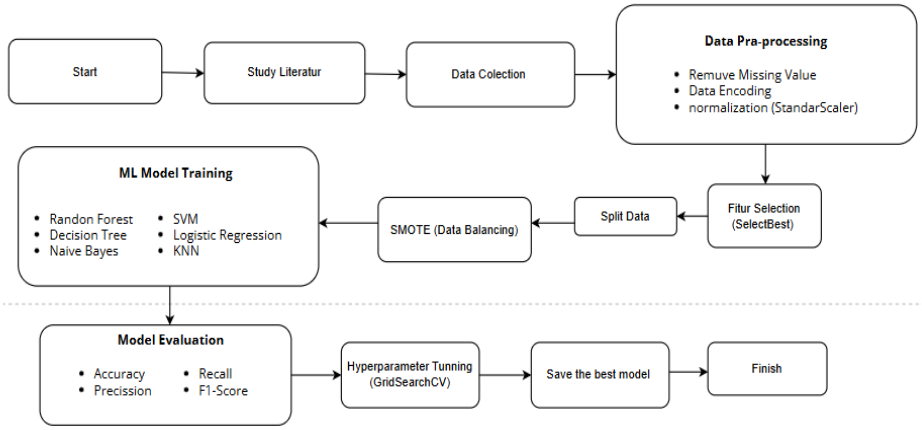

Figure 1. Flowchart of the Research

2) Data Collection

The dataset used in this research is sourced from Figshare under the title "MHP Anxiety Stress Depression Dataset of University Students." It contains 2,028 responses from university students across 15 institutions in Bangladesh, both public and private. The dataset includes a range of demographic, academic, and psychological features, which are crucial for building an effective depression prediction model [18]. Each entry in the dataset consists of a row representing a student and columns indicating different variables or features. These features include general information such as age, gender, university, department, academic year, CGPA, and whether the student receives a scholarship. Additionally, the dataset includes responses to the nine questions of the Patient Health Questionnaire-9 (PHQ-9), a widely used tool to assess depression [17]. The survey was developed by a team of experts and distributed through faculty representatives using an online form. The dataset has been validated for sample size and reliability, making it suitable for building predictive models in student mental health research. Although the data is specific to Bangladesh, the PHQ-9 instrument used is globally validated and aligns with DSM-IV diagnostic criteria, ensuring its relevance and applicability in broader academic contexts. This dataset serves as the foundation for training and developing machine learning algorithms aimed at predicting depression levels among university students, contributing to the field of mental health research.

3) Data Pre-processing

Data preprocessing is a crucial stage in ensuring the quality and consistency of the dataset before analysis. The preprocessing steps include:

a) Handling Missing Values: Missing data points are imputed using the mean for numerical features and the mode for categorical features to avoid biases in the model training process.
b) Encoding Categorical Data: Categorical variables are converted into numerical form using Label Encoding or One-Hot Encoding, enabling the machine learning algorithms to process them effectively.
c) Normalization: Numerical features are normalized using the StandardScaler method. This ensures that all features are on a comparable scale, preventing larger-valued features from dominating the learning process. The formula used for normalization as shown in Equestion 1.

d) $$z = \frac{x-\mu}{\sigma} \quad (1)$$

Where x is the feature value, μ is the average value, whileσ denotes the standard deviation

This step is essential for maintaining the stability and convergence of scale-sensitive algorithms.

4) Feature Selection

To improve model performance and reduce dimensionality, feature selection is performed using the SelectKBest method with an ANOVA F-test as the scoring function [20]. This technique ranks the features based on their statistical significance relative to the target variable (depression levels) and selects the top-k features that contribute most to predicting depression. Feature selection helps reduce overfitting, enhance training efficiency, and improve the interpretability of the model by focusing only on the most relevant features.

5) SMOTE (Synthetic Minority Over-sampling Technique)

The dataset exhibits an imbalance in the target class, with some depression levels being underrepresented. To address this issue and prevent model bias, SMOTE (Synthetic Minority Over-sampling Technique) is applied. SMOTE generates synthetic data points for the minority class by averaging the values of the nearest neighbors in feature space. The key parameter used in SMOTE is k_neighbors=5, meaning each new synthetic sample is generated based on the average of the five nearest neighbors. This parameter was chosen after conducting a trial-and-error process combined with sensitivity analysis to ensure optimal class balancing without overfitting [21].

6) Data Split

The dataset is split into two portions: 80% of the data is allocated for training, and 20% is reserved for testing. The split is performed in a stratified manner, ensuring that the class distribution remains consistent in both sets. This method is crucial for ensuring that the model evaluation is not biased by skewed class distributions, providing a fair and accurate assessment of model performance [22].

7) Machine Learning Model Training

In this study, six machine learning models are compared: Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Naïve Bayes, Logistic Regression, Random Forest, and Decision Tree. The selection of these algorithms is based on their established effectiveness in various classification tasks, as demonstrated in previous studies [25]. Each model has distinct characteristics that make them suitable for different types of data and classification problems.

a) Random Forest and Decision Tree: These tree-based models are chosen for their ability to handle complex data and produce models that are easily interpretable through visualization. Both algorithms are known for their high classification accuracy, making them ideal for tasks that require clear decision boundaries and explainable results [26].
b) Support Vector Machine (SVM): SVM is particularly effective in classifying high-dimensional data. Its strength lies in its ability to maximize the margin between classes, which helps improve classification accuracy, especially when dealing with complex datasets [27].
c) Logistic Regression: This model is selected for its simplicity and efficiency in modeling the linear relationship between features and the target variable. Despite its straightforward nature, logistic regression can provide reliable and interpretable results, making it a strong choice for problems with a linear decision boundary [28].
d) Naive Bayes: The Naive Bayes algorithm is chosen due to its assumption of independence between features, which makes it computationally efficient and fast. It is particularly well-suited for classification tasks with simpler structures or when working with text-based data [29].
e) K-Nearest Neighbors (KNN): KNN is a distance-based algorithm that classifies data points based on the majority vote of their nearest neighbors. This model is particularly effective when dealing with balanced datasets and has been shown to produce good prediction results when the data has a clear, well-defined structure [30].

In this research, these six algorithms are trained on the prepared dataset to determine the most effective approach for predicting depression levels among university students. Each model's strengths and characteristics make it suitable for

different aspects of the prediction task, and comparing their performance will help identify the best algorithm for this particular application.

8) Model Evaluation

Once the models are trained, they are evaluated using the test dataset, which was not involved in the training process to ensure unbiased performance measurement. The evaluation metrics used include:

a) Accuracy: The proportion of correct predictions out of all predictions made. It is useful for balanced datasets but less reliable when class distribution is skewed.

$$\text{Akurasi} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \quad (2)$$

b) Precision: The proportion of true positives among all positive predictions. It indicates the accuracy of the model in predicting positive cases.

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \quad (3)$$

c) Recall (Sensitivity): The proportion of true positives correctly identified by the model. It measures how well the model detects actual positive cases.

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \quad (4)$$

d) F1-Score: The harmonic mean of precision and recall, providing a single metric that balances the two, especially useful when dealing with imbalanced datasets.

$$\text{F1} - \text{Score} = 2.\frac{\text{Precision}\,.\,\text{Recall}}{\text{Precision+Recall}} \quad (5)$$

These metrics collectively provide a comprehensive overview of each model's ability to handle the depression prediction task, particularly in the context of imbalanced datasets [23].

9) Hyperparameter Tuning (GridSearchCV)

To optimize the performance of the models, **GridSearchCV** is used for hyperparameter tuning, which performs an exhaustive search over a specified parameter grid. The process is combined with 5-fold cross-validation to ensure robust validation and avoid overfitting. This dual validation strategy strengthens the model's generalization capability and enhances its practical applicability [24].

10) Saving the Best Model

After evaluating the models, the one with the best performance—based on validation data—is saved for future use. This step ensures that the final model used for deployment is not just based on the training data but also reflects its performance in real-world scenarios. By saving the best model, we avoid retraining in the future, thus saving time and resources while ensuring consistent results.

11) Conclusion

The research concludes once the optimal model is identified, validated, and prepared for digital deployment. The final model will support early detection of depression among university students, facilitating timely intervention and mental health monitoring within academic environments.

3. RESULTS AND DISCUSSION

3.1. Initial Data Exploration

The first step in model development is to analyze the data. The dataset used is student mental healt data sourced from Fighshare with a total dataset of 2028 before, and after pre-processing it becomes 1989 data and there are 19 features including demographic, academic, and depression symptoms collected through reflective questions. The target label used is Depression Label, with the category of student depression classification. The initial visualization shows an unbalanced class distribution, with a predominance in the moderate depression category, and a much smaller amount of data in the severe or mild depression categories.

3.2. Data Pre-processing

The goal of data pre-processing is to assess the precision of the data before it is applied to machine learning model training. The first step is to handle empty values, where numerical features such as age and curren CGPA are imputed with mean values, while categorical features such as gendre and university are imputed with mode. Next, categorical data encoding is performed such as Label Encoding for binary features and One-Hot Encodding for features with multiple categories. Numerical features are then normalized using StandardScaler to have a balanced scale. The target label (Depression Label) is categorized from the total depression value into several classes, such as not depressed to severely depressed. Finally, the dataset is split into two segments: 80% for training and 20% for testing by startification to maintain a balanced label distribution.

3.3. Feature Selection

The features or patterns used include demographic and academic data, as well as 9 depression symptom questions that match the indicators in PHQ-9 (Patient Healt Questionnaire-9). In the figure below, you can see the graph of the feature selection results.

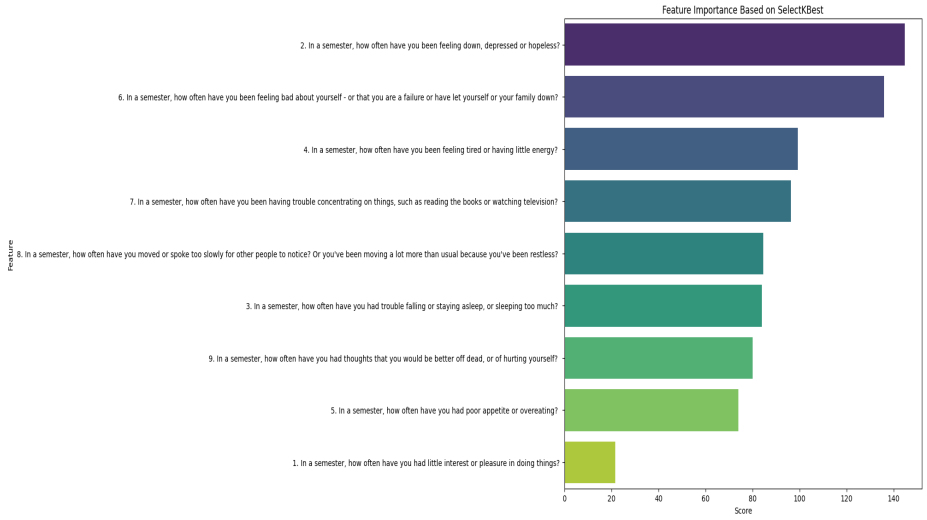

Figure 2. Feature Selection Output

The graph shows the importance of the features based on selectbest. From the correlation results and visual exploration, the psychological symptom features showed strong relationships with depression scores and labels, and were therefore retained in the model.

3.4. Machine Learning Model Evaluation

During the preliminary phase, six different machine learning algorithms were evaluated using standard preprocessed data and the data was split into training and testing sets in an 80:20 ratio. This evaluation aims to determine the baseline performance of the tested models Including algorithms like Support Vector Machine (SVM), Logistic Regression, Random Forest, Decision Tree, Naive Bayes, and K-Nearest Neighbors (KNN). Evaluation is carried out using a stratified train-test split technique so that the class distribution remains balanced in training and testing data. The assessment of each model was done using four main metrics, Specifically, accuracy, precision, recall, and F1-Score. A summary of the preliminary outcomes for each model is shown in Table 1.

Table 1. ML Model Comparison Evaluation Results

Model	Accuracy	Precision	Recall	F1-Score
SVM	0.587940	0.550473	0.587940	0.562694
Random Forest	0.482412	0.478805	0.482412	0.474088
Decision Tree	0.462312	0.454648	0.462312	0.452599
Logistic Regression	0.404523	0.363206	0.404523	0.367422
KNN	0.366834	0.364505	0.366834	0.360025
Naive Bayes	0.314070	0.295676	0.314070	0.292111

Based on the evaluation results, it is evident that the Support SVM model delivers the highest performance compared to all other models. there is an F1-Score value of 56.27%, precision of 55.04%, and relatively balanced accuracy and recall of 58.79%. Random Forest, Decision Tree, Logistic regression show Accuracy, Precision, Recall, and F1-Score around 40% and above and still below SVM. These results indicate that tree models tend to be affected by class imbalance in the data. Similarly, the low logistic regression model is caused by the linear assumptions of this model, making it less able to capture complex relationships between features in the data. Meanwhile, models such as Naive Bayes and KNN showed the lowest performance, with F1-Score of only 30% each, which is most likely due to the assumption of independence between features. This suggests that SVM has a more stable classification ability in recognizing different classes of depression than other models

3.5. Model Tunning

After the initial model evaluation, a hyperparameter tuning process was carried out to enhance the performance of each classification model. This tuning was implemented using the GridSearchCV method from the scikit-learn library, which performs an exhaustive search over a specified parameter grid to identify the most optimal combination. The tuning process aimed to maximize model accuracy, and was conducted using 5-fold cross-validation to ensure the results were reliable and generalizable. This cross-validation technique divides the data into five subsets, using four for training and one for validation in each iteration, thus minimizing bias and variance in performance estimation. The tuning outcomes, along with the corresponding accuracy scores for each model configuration, are presented in the Figure 3.

After the tunning process, SVM showed the best performance with a cross-validation score of 59.78% and a test accuracy of 99.50%. This result indicates that Support vector machine can identify complex patterns in the data, especially after adjusting the C, kernel, and gamma parameters. Logistic regression also experienced significant improvement with test accuracy reaching 64.57%. Meanwhile, Random Forest and Decision Tree did not show significant

improvement, even tended to stagnate despite tuning. Naive Bayes remains the lowest performing model, in line with the assumption of strict data distribution and simple model structure, while KNN shows a fairly good performance improvement, especially when using Manhattan metric and distance-based weighting.

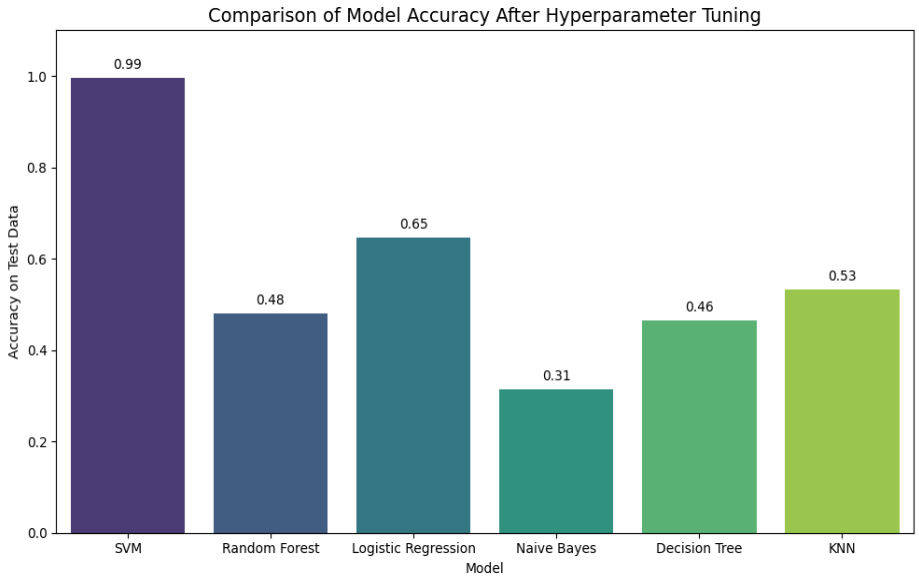

Figure 3. Comparison Chart of Model Accuracy after Hyperparameter Tunning

3.6. Handling Class Imbalance with SMOTE

To address the issue of imbalanced data distribution, the SMOTE (Synthetic Minority Over-sampling Technique) method was employed. Imbalanced datasets, where one class significantly outnumbers the others, can lead to biased model performance, particularly in classification tasks where the model tends to favor the majority class. SMOTE tackles this issue by generating synthetic samples of the minority class based on feature-space similarities between existing instances. This approach not only helps to balance the class distribution but also preserves the underlying structure of the data. Figure 4 and Figure 5 illustrates the class distribution before and after applying the SMOTE technique, demonstrating a more balanced dataset that allows for fairer and more accurate model training.

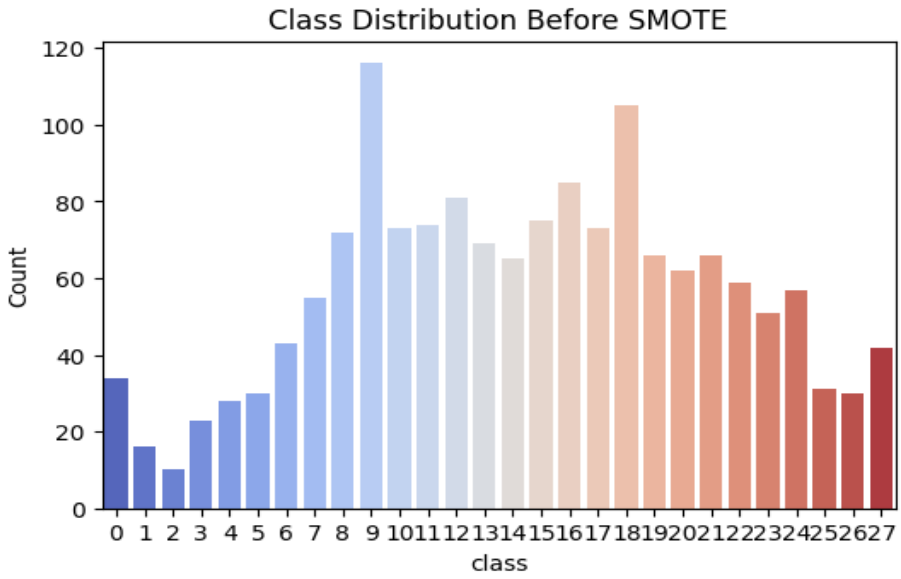

Figure 4. Class Distribution Before SMOTE

Figure 4 illustrates the class distribution before applying SMOTE, highlighting the imbalance in the dataset. It is evident that certain classes, such as classes 9 and 18, contain significantly more data than other classes. This imbalance poses a risk to the machine learning model, as it may become biased toward the majority classes, resulting in poor prediction performance for the minority classes.

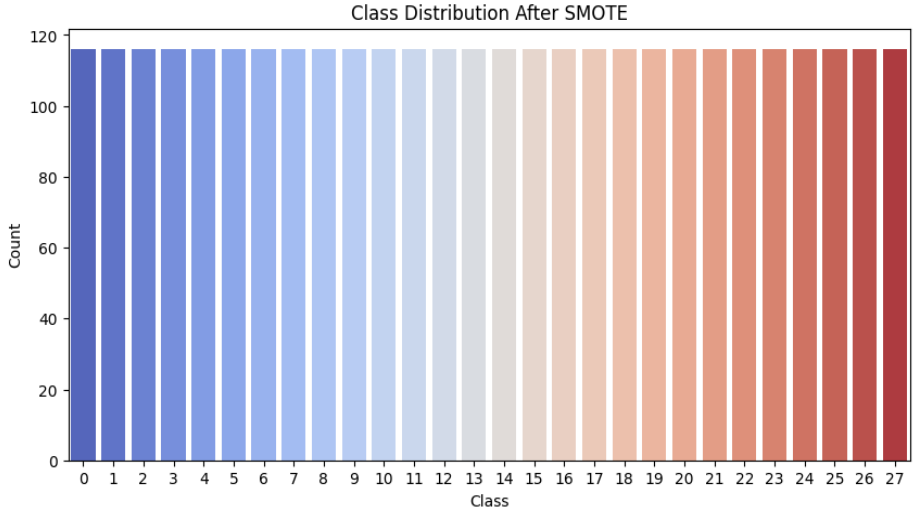

Figure 5. Class Distribution After SMOTE

After applying SMOTE, the class distribution is much more balanced. Each class now contains approximately 115 data points, ensuring equal representation across all classes. This demonstrates that SMOTE effectively generated synthetic data for the minority classes, addressing the imbalance and creating a more even distribution of data.

3.7. Final Evaluation and Interpretation of SVM Model

The final evaluation was conducted on the Support Vector Machine (SVM) model that had gone through the stages of preprocessing, feature selection, data balancing using SMOTE, and hyperparameter tuning with GridSearchCV. The model performed very well with 99% accuracy on the test data. The overall average and weighted average metrics for precision, recall, and F1-score reached 99% to 100% respectively, indicating that the model was able to classify all classes consistently and accurately. The graph is shown in the figure below.

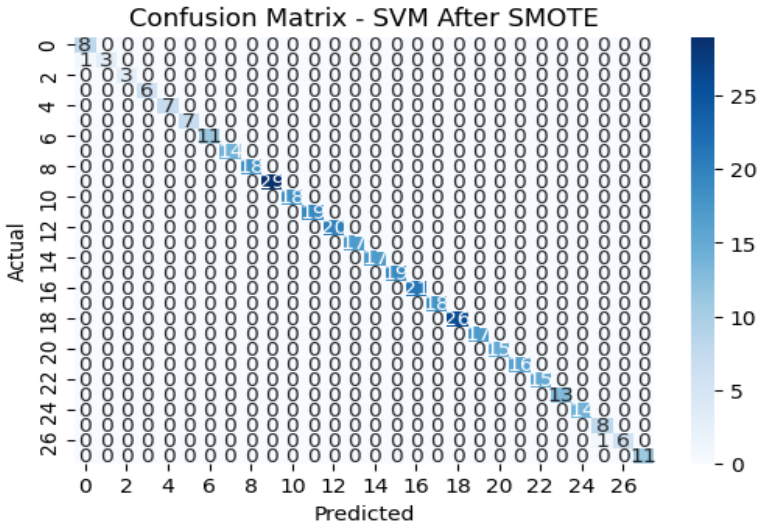

Figure 6. Confusion Matrix After SMOTE

The class-wise evaluation showed that almost all classes had perfect metric scores. However, there was a slight decrease in recall in class 1 (mild depression) by 75%, indicating challenges in distinguishing mild symptoms from other adjacent classes. Meanwhile, precision remained high, indicating that the predictions were very precise. To support the interpretation of the model, SHAP analysis was used, which showed that the most influential features in prediction were feelings of discouragement, sleep disturbance, fatigue, and thoughts of self-harm. These features correspond to common indicators of depression, corroborating that the model has learned from clinically relevant patterns.

3.8. System Interface Overview

The system interface is designed to be user-friendly and intuitive, ensuring that users can navigate through the process of depression assessment with ease. The login page serves as the initial step in the user journey. Upon visiting the system, users are prompted to either register for a new account or log in using their existing credentials. The registration process ensures that only authorized users can access the depression prediction tool, safeguarding sensitive mental health data. Secure authentication is paramount for protecting personal information and ensuring that the system is used responsibly. After registration, users can log in and gain access to the full features of the depression prediction system. Figure 7 illustrates the Login/Registration page, which provides the user with this authentication interface.

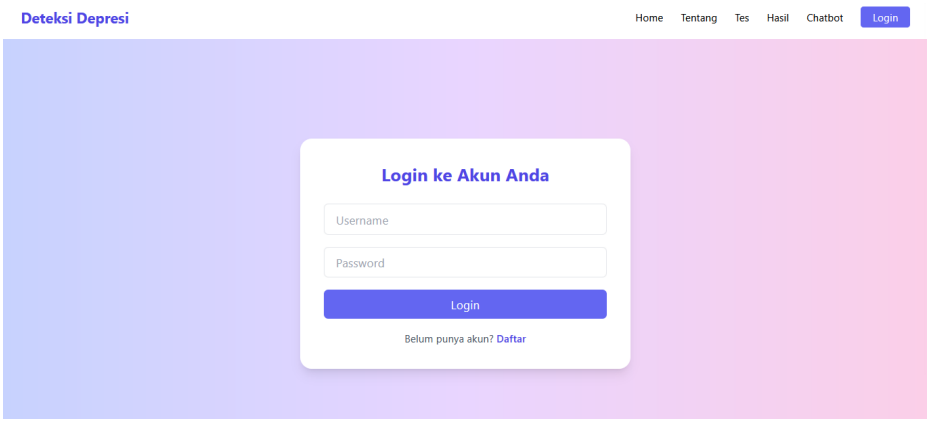

Figure 7. Login/Register

Figure 8. Home Page

Once users successfully log in, they are directed to the home page, which functions as the system's main dashboard. This page features a prominently displayed button that users can click to begin the depression assessment test. The home page is designed for easy navigation, guiding users smoothly to the next step of the process. Its clean, simple layout encourages engagement and provides a seamless transition into the testing phase. Figure 8 shows the Home Page, where users can start the test with a simple click. The About page provides users with valuable information about the system, including its purpose, objectives, and the machine learning model driving the depression prediction. This transparency is essential for building trust with users, helping them understand the underlying technology and methodology behind the tool. By explaining the system's goals and how it works, the About page ensures that users are well-informed about what to expect and the reliability of the system they are using. Figure 9 displays the About Page, which educates users on the background and functionality of the system.

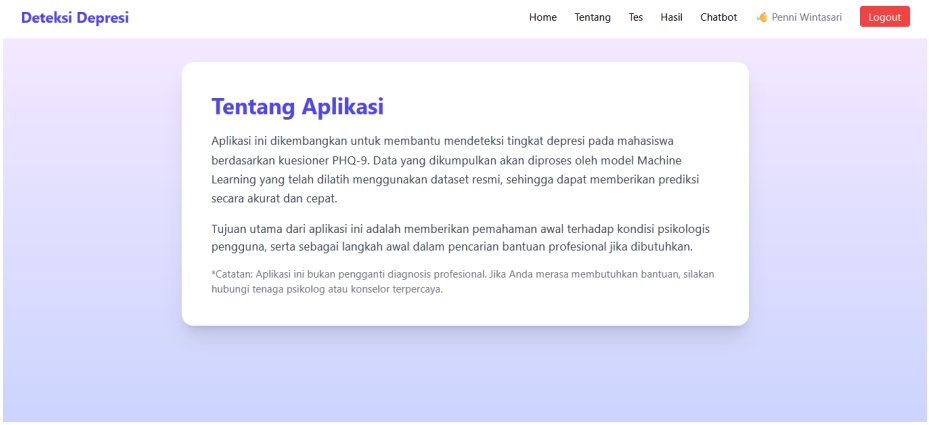

Figure 9. About page

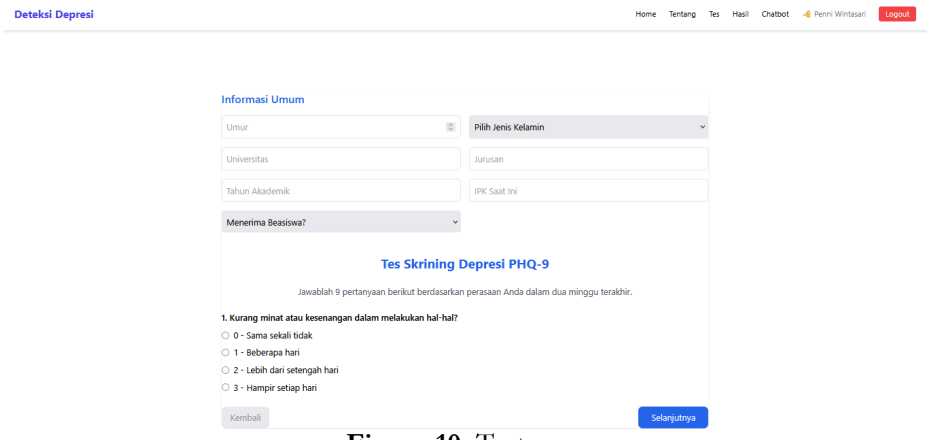

Figure 10. Test page

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

On the Test page, users are presented with a series of questions, which include both demographic details and the PHQ-9 depression assessment items. To generate accurate prediction results, users must complete all the questions. The page is designed for clarity and ease of use, with a straightforward layout that minimizes any potential for confusion. This ensures that users can provide accurate responses, which are crucial for generating reliable predictions. The design of the Test page is focused on reducing errors and making the experience as user-friendly as possible. Figure 10 shows the Test Page, where users can fill in their responses.

After completing the questionnaire, users are directed to the Result page, where the depression prediction results are displayed. The results are generated by the trained SVM model, and the page includes an interpretation of the results. This interpretation helps users understand the meaning of their score and offers guidance on potential next steps. The Result page is an essential part of the system as it provides real-time feedback, demonstrating the practical application of the depression prediction model. Users are given actionable insights that they can use to take the next steps in managing their mental health. Figure 11 depicts the Result Page, where users can view their depression prediction and corresponding interpretation.

Figure 11. Result page

The system interface is designed to guide users through each stage of the depression prediction process, from registration to result interpretation. With its clear, intuitive layout, the system ensures that users can easily access the tool, complete the assessment, and receive meaningful feedback on their mental health status. The design is user-cantered, aiming to foster a supportive environment where individuals can learn about and monitor their mental health with ease.

3.9. Black Box Testing

Black box testing is a crucial phase in the software testing process, where the internal workings of the system are not known to the tester. The focus is entirely on testing the system's functionality based on its inputs and expected outputs, without any knowledge of the code or internal logic. This type of testing helps ensure that the system functions as intended from the user's perspective and that all features work correctly in various scenarios. In the context of the depression prediction system, black box testing was performed to evaluate how well the system responds to various inputs and whether it produces the expected results. This testing involved simulating real-world usage by inputting different sets of data into the system and checking the corresponding outputs. Key areas of focus during black box testing included the accuracy of the predictions, the functionality of the user interface, and the overall user experience, including the interaction between the system's components. Table 2 summarizes the key aspects of black box testing performed on the system, along with the results for each tested feature.

Table 2. black box testing performed

Test Case	Test Description	Expected Outcome	Actual Outcome	Pass /Fail
Login Functionality	Test the login process with valid and invalid credentials	System should allow login with valid credentials and reject invalid ones	Login with valid credentials was successful, invalid login was rejected	**Pass**
Home Page Navigation	Test the navigation from the login page to the home page	Clicking the "Start Test" button should take the user to the test page	Navigation to the test page was seamless and functional	**Pass**
Test Page Completion	Test the completion of all fields in the PHQ-9 questionnaire	All fields should be filled in, and a prompt should appear if any field is left empty	The system correctly prompted when required fields were missing	**Pass**
Prediction Result Generation	Test the result page after completing the test	The system should generate a depression prediction based on the user's responses	Prediction results were accurately displayed and aligned with input data	**Pass**
System Response to Edge Cases	Test how the system responds to extreme or non-standard inputs	System should either handle edge cases or display an appropriate error message	System correctly handled edge cases or	**Pass**

Test Case	Test Description	Expected Outcome	Actual Outcome	Pass /Fail
			displayed error messages	
Data Privacy and Security	Test the system's handling of sensitive data, ensuring no leakage	Sensitive information such as personal data and responses should be securely stored	The system maintained data privacy and security at all stages	**Pass**

3.10. Discussion

The results of this study demonstrate that the Support Vector Machine (SVM) consistently outperforms the other machine learning models evaluated, achieving an impressive accuracy of 99.50%. This superior performance can be attributed to SVM's ability to establish optimal decision boundaries in high-dimensional feature spaces, which is especially beneficial when distinguishing between closely related classes such as "mild" and "moderate" depression. Given the nature of the PHQ-9 questionnaire, which tends to produce overlapping response patterns, SVM's strength lies in its capacity to create wide margins between these classes, allowing it to classify the data more effectively. This aligns with previous studies that have highlighted SVM as a strong performer in psychological and sentiment analysis tasks, where nuanced distinctions are necessary [20][9].

In contrast, tree-based algorithms, such as Random Forest and Decision Tree, showed weaker results in this study. Despite hyperparameter tuning, their performance only slightly improved, suggesting that these algorithms struggle to generalize when faced with complex, imbalanced datasets. The tendency of these models to overfit on minority classes likely contributed to this limitation. Similarly, Logistic Regression struggled with the task due to its reliance on linear relationships, which failed to capture the non-linear associations between psychological symptoms commonly found in depression. The poor performance of K-Nearest Neighbors (KNN) and Naive Bayes can be attributed to their sensitivity to feature scaling and the unrealistic independence assumptions made by Naive Bayes, which are not well suited to the interdependent nature of the features in psychological datasets.

A major contribution of this study is the use of SMOTE (Synthetic Minority Over-sampling Technique) to address the issue of class imbalance, a common challenge in mental health datasets. By generating synthetic examples for the underrepresented classes, SMOTE helped balance the dataset, improving the models' ability to make accurate predictions across all categories. This approach improves upon previous studies that did not adequately address class imbalance,

which often leads to biased model performance, favoring the majority class [13][29].

Another critical component of this study is the interpretability of the machine learning model. By utilizing SHAP analysis (Shapley Additive Explanations), the study identified key features such as "sleep disturbance," "low energy," and "thoughts of self-harm" as the most influential factors in determining the depression prediction results. These features are consistent with well-established clinical symptoms of depression, lending credibility to the model's reasoning process and ensuring that the results are grounded in psychological theory. These insights can be valuable in practical applications, allowing universities and mental health professionals to design early intervention programs targeting students who exhibit these specific symptoms.

The optimized SVM model, with its exceptional accuracy of 99.5%, demonstrates strong potential for real-world applications in digital mental health tools. This model could be integrated into university mental health services or mobile applications, enabling rapid depression screening through a simple self-assessment like the PHQ-9. Such integration could streamline response times and support proactive mental health initiatives on campuses, helping institutions address mental health concerns before they become more severe.

However, while the model shows promise, any real-world implementation must be accompanied by safeguards to address potential ethical concerns. For example, there is a risk of misclassification, and data privacy must be rigorously protected. Incorporating human oversight into the process and ensuring secure data handling protocols are in place will be crucial for responsible use. As such, the deployment of this tool should be accompanied by clear guidelines to protect users' privacy and ensure its ethical application.

In summary, this study highlights the effectiveness of combining machine learning techniques, SMOTE, and SHAP analysis to create a powerful and transparent depression detection tool. By addressing both the technical challenges of class imbalance and the need for model interpretability, the system provides a comprehensive solution for early depression detection. The high accuracy and transparency of this model make it well-suited for integration into institutional mental health monitoring systems, offering a promising tool for proactive mental health interventions.

4. CONCLUSION

This study successfully developed a machine learning-based model for predicting depression among university students, employing a comprehensive workflow that

included data cleaning, feature selection, addressing class imbalance using SMOTE, training six machine learning algorithms, and fine-tuning hyperparameters through GridSearchCV. The Support Vector Machine (SVM) algorithm emerged as the most effective, achieving an impressive 99.5% accuracy rate while maintaining balanced precision and recall after optimization. The use of SHAP analysis further enhanced model transparency, revealing key features that align with clinically recognized depression symptoms. These findings contribute to the growing body of technology-driven tools for early mental health detection, enabling educational institutions to identify vulnerable students and provide targeted support strategies. However, the study has its limitations. The dataset, sourced from universities in Bangladesh, may limit the model's generalizability to other cultural or educational contexts. Additionally, the model has not yet been tested in real-world applications or with diverse student populations, which may affect its practical efficacy and reliability across various environments.

Future research should focus on integrating the model into live digital counseling platforms, ensuring it complements existing mental health services. Collaboration with mental health practitioners and student support departments will be crucial to ensure the tool's meaningful use and proper follow-up. Ethical considerations, such as safeguarding data privacy, reducing the risk of misclassification, and obtaining informed user consent, must also be addressed to prevent negative consequences and avoid stigmatization of individuals seeking support.

REFERENCES

[1] E. S. Gisela, E. A. Kinkie, A. Sabbilla, and U. Subroto, "Pengaruh Stres Akademik terhadap Kesejahteraan Psikologis Mahasiswa Semester Akhir yang Terlambat Lulus," *Jurnal Mahasiswa Humanis*, vol. 5, no. 1, Jan. 2025, doi: 10.47065/josh.v6i1.5930.

[2] E. Junilia and A. K. Dharmawan, "Menelusuri Jejak Depresi di Kalangan Mahasiswa: Sebuah Eksplorasi Gejala Khas," *MagnaSalus: Jurnal Keunggulan Kesehatan*, vol. 6, no. 3, 2024.

[3] D. Muriyatmoko, Dihin, A. Musthafa, and M. Fa-Idzaa, "Perbandingan Metode Support Vector Machine dan Random Forest dalam Menganalisis Pengaruh Musik Terhadap Penurunan Tingkat Stress Mahasiswi Semester 7 saat Skripsi (Studi Kasus: Universitas Darussalam Gontor)," in *Prosiding Seminar Nasional Amikom Surakarta*, vol. 2, pp. 128-135, 2024.

[4] M. Naufal, D. W. Utomo, and R. P. Tresyani, "Early Detection of Mental Health Disorders based on Sentiment using Stacking Method," *Sistemasi: Jurnal Sistem Informasi*, vol. 14, no. 1, pp. 271-280, 2025.

[5] S. Juwariyah, A. Hulvi, N. Riduan, and K. Kusrini, "Mengukur Faktor Demografi Psikologis: Memprediksi Depresi, Kecemasan, dan Stres dengan menggunakan Machine Learning," *Komputika : Jurnal Sistem*

Komputer, vol. 13, no. 2, pp. 149–156, Sep. 2024, doi: 10.34010/komputika.v13i2.11793.

[6] J. Homepage *et al.*, "Research in the Mathematical and Natural Sciences Klasifikasi Tingkat Depresi Mahasiswa Menggunakan Image Recognition dengan Support Vector Machine," *Res. Math. Nat. Sci*, vol. 4, no. 1, pp. 30–36, 2025, doi: 10.55657/rmns.v4i1.193.

[7] P. T. Prasetyaningrum, P. Purwanto, and A. F. Rochim, "Consumer Behavior Analysis in Gamified Mobile Banking: Clustering and Classifier Evaluation," *Online) Journal of System and Management Sciences*, vol. 15, no. 2, pp. 290–308, 2025, doi: 10.33168/JSMS.2025.0218.

[8] M. R. Sudrajat and M. Zakariyah, "Penerapan Natural Language Processing dan Machine Learning untuk Prediksi Stres Siswa SMA Berdasarkan Analisis Teks," *Building of Informatics, Technology and Science (BITS)*, vol. 6, no. 3, Dec. 2024, doi: 10.47065/bits.v6i3.6180.

[9] H. Syahputra, S. I. Naibaho, M. A. Maulana, I. Zulfahmi, and E. P. Sinaga, "Perbandingan Algoritma Support Vector Machine (SVM) dan Decision Tree Untuk Deteksi Tingkat Depresi Mahasiswa," *BINA INSANI ICT JOURNAL*, vol. 10, no. 1, pp. 52–61, 2023, doi: https://doi.org/10.51211/biict.v10i1.2304.

[10] H. Syahputra, S. I. Naibaho, M. A. Maulana, I. Zulfahmi, and E. P. Sinaga, "Perbandingan Algoritma Support Vector Machine (SVM) dan Decision Tree Untuk Deteksi Tingkat Depresi Mahasiswa," *BINA INSANI ICT JOURNAL*, vol. 10, no. 1, pp. 52–61, 2023, doi: https://doi.org/10.51211/biict.v10i1.2304.

[11] A. U. Haspriyanti and P. T. Prasetyaningrum, "Penerapan Data Mining Untuk Prediksi Layanan Produk Indihome Menggunakan Metode K-Nearst Neighbor," *Journal of Information System and Artificial Intelligence*, vol. 1, no. 2, pp. 100-107, 2021.

[12] I. Setiawan, I. F. Yasin, and Y. T. Desianti, "Komparasi Kinerja Algoritma Random Forest, Decision Tree, Naïve Bayes, dan KNN dalam Prediksi Tingkat Depresi Mahasiswa menggunakan Student Depression Dataset," *Jurnal Ilmu Komputer dan Teknologi*, vol. 6, no. 1, pp. 47-58, 2025.

[13] R. Fahlapi, H. Hermanto, A. Y. Kuntoro, L. Effendi, R. O. Nitra, and S. Nurlela, "Prediction of Employee Attendance Factors Using C4.5 Algorithm, Random Tree, Random Forest," *Semesta Teknika*, vol. 23, no. 1, 2020, doi: 10.18196/st.231254.

[14] P. T. Prasetyaningrum, P. Purwanto, and A. F. Rochim, "Enhancing Element Game Classification: Effective Techniques for Handling Imbalanced Classes," *International Journal of Intelligent Engineering and Systems*, vol. 17, no. 1, pp. 555–571, 2024, doi: 10.22266/ijies2024.0229.47.

[15] M. Fadhilla, R. Wandri, A. Hanafiah, P. R. Setiawan, Y. Arta, and S. Daulay, "Analisis Performa Algoritma Machine Learning Untuk Identifikasi

Depresi Pada Mahasiswa," *Journal of Informatics Management and Information Technology*, vol. 5, no. 1, pp. 40–47, Jan. 2025, doi: 10.47065/jimat.v5i1.473.

[16] W. Wahyuningsih and P. T. Prasetyaningrum, "Enhancing Sales Determination for Coffee Shop Packages through Associated Data Mining: Leveraging the FP-Growth Algorithm," *Journal of Information Systems and Informatics*, vol. 5, no. 2, pp. 758–770, May 2023, doi: 10.51519/journalisi.v5i2.500.

[17] U. E. Chigbu, S. O. Atiku, and C. C. Du Plessis, "The Science of Literature Reviews: Searching, Identifying, Selecting, and Synthesising," *Publications*, vol. 11, no. 1, Mar. 2023, doi: 10.3390/publications11010002.

[18] P. J. Yu, "A Machine Learning Method for Detecting Depression Among College Students," *Int J Comput Appl*, vol. 185, no. 24, pp. 44–51, Jul. 2023, doi: 10.5120/ijca2023923003.

[19] A. E. Karrar, "The Effect of Using Data Pre-Processing by Imputations in Handling Missing Values," *Indonesian Journal of Electrical Engineering and Informatics*, vol. 10, no. 2, pp. 375–384, Jun. 2022, doi: 10.52549/ijeei.v10i2.3730.

[20] A. Chatterjee *et al.*, "Statistical Analysis of Online Public Survey Lifestyle Datasets: A Machine Learning and Semantic Approach," Nov. 28, 2023. doi: 10.21203/rs.3.rs-2864069/v1.

[21] I. Komang Dharmendra, I. Made, A. W. Putra, and Y. P. Atmojo, "Evaluasi Efektivitas SMOTE dan Random Under Sampling pada Klasifikasi Emosi Tweet," *Informatics for Educators and Professionals : Journal of Informatics*, vol. 9, no. 2, pp. 192–193, 2024.

[22] I. Muraina, "Ideal dataset splitting ratios in machine learning algorithms: general concerns for data scientists and data analysts," in *7th International Mardin Artuklu Scientific Research Conference*, pp. 496-504, 2022.

[23] D. Martin Ward Powers, "Evaluation: From Precision, Recall And F-Measure To Roc, Informedness, Markedness & Correlation," *Article in Journal of Machine Learning Technologies*, vol. 2, no. 1, pp. 37–63, 2011, doi: 10.9735/2229-3981.

[24] M. R. R. Saelan and A. Subekti, "K-Best Selection Untuk Meningkatkan Kinerja Artificial Neural Network Dalam Memprediksi Range Harga Ponsel," *INTI Nusa Mandiri*, vol. 19, no. 1, pp. 10–16, Jul. 2024, doi: 10.33480/inti.v19i1.5554.

[25] A. Supoyo and P. T. Prasetyaningrum, "Analisis Data Mining Untuk Memprediksi Lama Perawatan Pasien Covid-19 Di DIY," *Bianglala Inform*, vol. 10, no. 1, pp. 21-29, 2022.

[26] A. Fauzi, N. Maulidah, R. Supriyadi, H. Nalatissifa, and S. Diantika, "Prediksi Harga Properti Di Indonesia Menggunakan Algoritma Random Forest," *RIGGS: Journal of Artificial Intelligence and Digital Business*, vol. 4, no. 1, pp. 43-49, 2025, doi: 10.31004/riggs.v4i1.367.

[27] H. Eldo, A. Ayuliana, D. Suryadi, G. Chrisnawati, and L. Judijanto, "Penggunaan Algoritma Support Vector Machine (SVM) Untuk Deteksi Penipuan pada Transaksi Online," *Jurnal Minfo Polgan*, vol. 13, no. 2, pp. 1627–1632, Oct. 2024, doi: 10.33395/jmp.v13i2.14186.

[28] F. R. U. Hasanah and M. Yollanda, "Penerapan Model Regresi Logistik Terhadap Indeks Pembangunan Manusia (IPM) di Provinsi Sumatera Barat Tahun 2019–2021," *JOSTECH Journal of Science and Technology*, vol. 2, no. 2, pp. 199-208, 2022.

[29] J. Zeniarja, A. Salam, and F. A. Ma'ruf, "Seleksi Fitur dan Perbandingan Algoritma Klasifikasi untuk Prediksi Kelulusan Mahasiswa," *Jurnal Rekayasa Elektrika*, vol. 18, no. 2, Jul. 2022, doi: 10.17529/jre.v18i2.24047.

[30] N. T. Ujianto, Gunawan, H. Fadillah, A. P. Fanti, A. D. Saputra, and I. G. Ramadhan, "Penerapan algoritma K-Nearest Neighbors (KNN) untuk klasifikasi citra medis," *IT-Explore: Jurnal Penerapan Teknologi Informasi dan Komunikasi*, vol. 4, no. 1, pp. 33–43, Feb. 2025, doi: 10.24246/itexplore.v4i1.2025.pp33-43.