Journal of Applied Informatics and Computing (JAIC)
Vol.10, No.4, August 2026, pp. 3311~3317
e-ISSN: 2548-6861

Comparative Analysis of K-Nearest Neighbors and Support Vector Machine for Student Stress Prediction

Luthfi Khoirul Umam[1], Setyoningsih Wibowo[2], Agung Handayanto[3]
[123]Program Studi Informatika, Universitas PGRI Semarang
luthfikhoirulumam2@gmail.com [1], setyoningsihwibowo@upgris.ac.id [2], agunghan@upgris.ac.id [3]

Article Info

Article history:

Received 2026-06-05
Revised 2026-07-24
Accepted 2026-08-08

Keyword:

K-Nearest Neighbors,
Support Vector Machine,
Machine Learning,
Student Stress Prediction.

ABSTRACT

Student stress has emerged as an important issue because it can negatively affect both academic achievement and mental well-being. The growing development of machine learning techniques has enabled the creation of predictive models that can classify stress levels using various student-related factors. This research evaluates and compares the effectiveness of the K-Nearest Neighbors (KNN) and Support Vector Machine (SVM) algorithms in predicting student stress categories. The dataset was sourced from Kaggle and contains questionnaire-based data, including variables such as sleep quality, frequency of headaches, academic achievement, study workload, participation in extracurricular activities, and stress level classifications. The study followed several stages, including data preprocessing, feature normalization, model development, and performance assessment. Model evaluation was conducted using accuracy, precision, recall, F1-score, and confusion matrix metrics. To validate the practical implementation of the models, both algorithms were incorporated into a web-based application built with the Flask framework and supported by a MySQL database. The experimental results revealed that the KNN algorithm delivered superior classification performance compared to SVM. KNN obtained an accuracy score of 88.46% and a weighted F1-score of 0.89, whereas SVM achieved 46.15% accuracy with a weighted F1-score of 0.41. These findings suggest that KNN is more effective in identifying patterns within the student survey data. In addition, the successful deployment of the prediction system demonstrates that conventional machine learning methods can be utilized to provide real-time assessments of student stress levels and support mental health monitoring efforts.

I. INTRODUCTION

Student stress has become one of the major challenges in higher education environments [1]. Academic workload, lack of sleep, pressure to achieve good academic performance, and extracurricular activities are several factors that can affect students' psychological conditions. High stress intensity is associated with negative consequences for academic outcomes, mental well-being, and student's ability to maintain productive activities [2]. Therefore, early identification of stress levels is important to support preventive actions and improve student well-being [3].

Along with the advancement of information technology, machine learning has been widely applied in various prediction and classification tasks, including psychological and educational data analysis [4]. Machine learning enables systems to identify patterns from historical data and generate predictions automatically [5]. Machine learning classification tasks are commonly performed using algorithms such as K-Nearest Neighbors (KNN) and Support Vector Machine (SVM K-Nearest Neighbors (KNN) assigns class membership based on the closest neighbouring samples in the feature space, whereas Support Vector Machine (SVM) separates categories by generating a hyperplane with the maximum margin between classes [6].

Findings from prior studies suggest that machine learning algorithms are capable of accurately predicting psychological conditions, including stress levels in students [7]. However, differences in dataset characteristics can significantly influence algorithm performance [8]. Therefore, comparative analysis between algorithms is necessary to determine the most suitable model for a specific dataset [9].

This study conducts a comparative analysis of KNN and SVM classification models to determine their capability in predicting student stress levels using a Kaggle-based student survey dataset. [10]. Comparative evaluation of the KNN and SVM models is conducted using accuracy, precision, recall, F1-score, and confusion matrix analyses. The superior model is subsequently implemented in a Flask-powered web application to provide instant predictions of student stress levels. [11].

The contribution of this study lies in the comparative analysis of KNN and SVM algorithms on student stress prediction and the implementation of the developed machine learning model into an interactive web-based system [12]. The results of this study are expected to provide insights into the effectiveness of classification algorithms in educational psychological data analysis and support the development of intelligent student support systems [13]. Although numerous studies have investigated the application of machine learning algorithms for stress prediction, most previous works focused on binary stress classification, different health-related datasets, or did not integrate the trained models into practical applications. This study contributes by evaluating the performance of K-Nearest Neighbors and Support Vector Machine algorithms on a multi-class student stress dataset consisting of five stress categories derived from behavioral and academic factors. Furthermore, the best-performing model was implemented into a Flask-based web application integrated with a MySQL database, enabling real-time prediction of students' perceived stress levels. Therefore, this study not only provides a comparative analysis of classification algorithms but also demonstrates their practical implementation in a web-based decision support system.

II. METHOD

A. Dataset

The dataset used in this study was obtained from the Kaggle platform, namely the Student Stress Analysis Dataset. The dataset consists of 520 records representing responses from students regarding academic and behavioral factors associated with stress levels. A total of six attributes are available in the dataset, consisting of five predictor variables and one target variable [14].

The predictor variables include sleep quality, headache frequency, academic performance, study load, and extracurricular activity frequency, while the target variable is stress level, which is categorized into five classes ranging from level 1 to level 5. The distribution of stress categories is relatively balanced, with 110 samples belonging to level 1, 105 samples to level 2, 125 samples to level 3, 100 samples to level 4, and 80 samples to level 5.

The stress level labels contained in the dataset were derived from self-reported responses provided by participants in the original survey. The dataset documentation does not indicate the use of standardized psychological assessment instruments such as the Depression Anxiety Stress Scales (DASS-21) or the Perceived Stress Scale (PSS). Therefore, the stress categories used in this study should be interpreted as subjective stress assessments reported by respondents rather than clinically validated diagnoses. Despite this limitation, the dataset remains suitable for evaluating machine learning classification models and exploring relationships between behavioral factors and perceived stress levels.

To evaluate the classification models objectively, the dataset was divided into training and testing sets using an 80:20 ratio. A total of 416 records (80%) were allocated for model training, while 104 records (20%) were reserved for model evaluation. This partitioning strategy was adopted to assess model performance using previously unseen data and to reduce the possibility of overfitting.

The characteristics and class distribution of the dataset are summarized in Tables 1 and 2.

TABLE 1.
CHARACTERISTICS OF DATASET

Characteristic	Value
Number of Samples	520
Predictor Attributes	5
Target Variable	Stress Level
Number of Classes	5
Training Samples	416
Testing Samples	104
Train-Test Ratio	80:20

TABLE 2.
DISTRIBUTION OF STRESS LEVEL CATEGORIES

Stress Level	Number of Samples	Percentage (%)
Level 1	110	21.15
Level 2	105	20.19
Level 3	125	24.04
Level 4	100	19.23
Level 5	80	15.39
Total	520	100.00

B. Data Preprocessing

An analysis of class distribution indicated that the dataset did not exhibit severe class imbalance. Although stress level 5 contained fewer observations compared to the other categories, the difference was not substantial enough to require additional balancing techniques such as oversampling, under sampling, or synthetic data generation methods. Therefore, the original class distribution was maintained throughout the training and evaluation processes.

Before the classification process was performed, the dataset underwent several preprocessing stages to improve data quality and optimize model performance, After preprocessing, the dataset was divided into training data and testing data using an 80:20 ratio. A total of 80% of the data (416 records) were used for model training, while the remaining 20% (104 records) were used for model testing and evaluation. This data splitting strategy was employed to ensure that the developed models could be evaluated objectively using unseen data [15]. The preprocessing stages included:

1. Data Cleaning
   The dataset was checked to ensure there were no inconsistent or incomplete data values.
2. Feature Encoding
   Categorical data were transformed into numerical representations to enable processing by machine learning algorithms.
3. Feature Normalization
   Data normalization was performed using a scaler to ensure that all features had balanced value ranges.
4. Dataset Splitting
   The dataset was divided into training data and testing data to evaluate model performance objectively.

C. K-Nearest Neighbors (KNN)

K-Nearest Neighbors (KNN) is a classification algorithm that determines the class of data based on the closest neighboring data points. The Euclidean distance was used to measure similarity between data points. In this study, the KNN model used a parameter value of k=5 [16].

The KNN algorithm can be mathematically represented as:

$$d(x,y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)}^2$$

where:

- $d(x,y)$= distance between data points
- x_i= feature value of training data
- y_i= feature value of testing data

D. Support Vector Machine (SVM)

Support Vector Machine (SVM) is a supervised learning method that performs classification by identifying an optimal decision boundary between classes. By maximizing the margin between data groups, SVM aims to improve classification accuracy and generalization capability. In this research, the LinearSVC implementation was utilized to classify student stress levels., the LinearSVC model was used to perform classification [17].

Support Vector Machine (SVM) constructs a hyperplane that maximizes the margin between class boundaries, contributing to improved classification performance and robustness on unseen data [18].

E. Model Evaluation

The models performance was measured through various classification evaluation criteria [19], including the following metrics:

1. Accuracy
2. Precision
3. Recall
4. F1-Score
5. Confusion Matrix

Accuracy was calculated using the following formula:

$$Accuracy = \frac{TP + TN}{TP + TN + FP \ + FN}$$

where:
TP = True Positive
TN = True Negative
FP = False Positive
FN = False Negative

F. System Implementation

The developed machine learning models were implemented into a web-based application using Flask as the backend framework. The system allows users to input survey data and obtain stress level prediction results in real time. The application also integrates MySQL database storage to save prediction history and user data [20].

III. RESULT AND DISCUSSION

A. Data Pre-processing Results

Before the model training process begins, the dataset undergoes preprocessing to improve data quality and optimize classification performance. The preprocessing stages include data cleaning, feature encoding, data normalization, and dividing the dataset into training and testing data.

The data cleaning stage ensures that the dataset contains no inconsistent or incomplete data that could affect the prediction results. Next, categorical features are converted into numerical form using an encoding process so they can be processed by the machine learning algorithm.

Data normalization is performed using a scaler to equalize the range of values between features. This step is crucial because the KNN and SVM algorithms are sensitive to differences in data scale. Without normalization, features with a larger range of values can dominate the classification process and degrade model performance.

After preprocessing, the dataset is divided into two subsets: training data and testing data. The training subset is used for model construction, whereas the testing subset is reserved for performance validation.

Figure 1. Data Distribution

The preprocessing stage makes an important contribution to improving data quality so that the classification process can run more optimally.

B. Evaluation Results of the K-Nearest Neighbors (KNN) Algorithm

The K-Nearest Neighbors (KNN) algorithm was implemented using the parameter k=5. Based on the test results using the testing data, the KNN algorithm achieved excellent classification performance in predicting student stress levels. The following is the confusion matrix generated by the KNN model in the figure 2. The K-Nearest Neighbors (KNN) model was implemented using k = 5. This parameter value was selected based on recommendations from previous studies suggesting that small odd-numbered k values often provide a good balance between model complexity and classification performance. No additional hyperparameter optimization techniques, such as Grid Search or Cross Validation, were performed in this study.

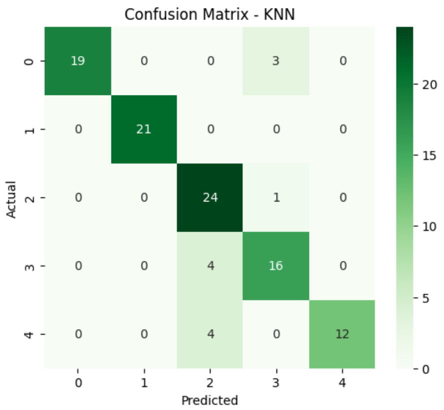

Figure 2. Confusion Matrix KNN

Based on the KNN algorithm's confusion matrix, most of the data was correctly classified in each stress level category. The main diagonal values in the confusion matrix are dominant, indicating that the model has a high rate of correct predictions.

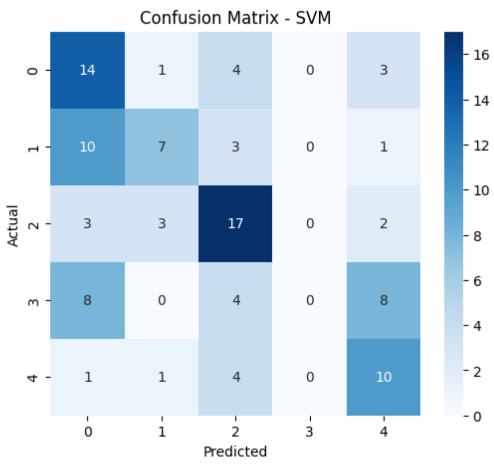

Figure 3. Confusion Matrix KNN

In the low and medium stress categories, the model was able to recognize data patterns very well with a relatively small number of misclassifications. Some mispredictions occurred in stress levels with close characteristics, such as between the medium and high stress categories. However, overall, the number of misclassifications remained relatively low. The confusion matrix results demonstrate that the KNN algorithm has a good ability to recognize similar patterns among student data, resulting in stable and accurate classification performance, as shown in the table 2.

TABLE 3.
KNN RESULTS

Matrix	Value
Accuracy	88,46%
Precision	0,90
Recall	0,88
F1-score	0,89

Based on the evaluation results, the KNN algorithm was able to correctly classify most stress level categories. This is evident in the confusion matrix, where the majority of values fall on the main diagonal, indicating correct classification results. The misclassification rate in the KNN model was relatively small and only occurred in a few stress level categories with close characteristics. This indicates that the KNN algorithm is capable of recognizing similar patterns between data sets well.

The high performance achieved by KNN indicates that the dataset used in this study has a data distribution pattern that is quite suitable for distance-based classification. Because KNN works by measuring the proximity between data points, the previous normalization process also significantly increased model accuracy.

Furthermore, the relatively small size of the dataset also contributed to KNN's performance. The KNN algorithm generally performs best on small to medium-sized datasets because relationships between data points can still be effectively identified. These results also align with several previous studies showing that KNN performs well in survey-based data classification and user behavior.

C. Evaluation Results of the Support Vector Machine (SVM) Algorithm

To assess the classification capability of Support Vector Machine (SVM), the LinearSVC model was trained and tested using the prepared dataset. The obtained results revealed that SVM was less effective than KNN in distinguishing stress level categories. A detailed overview of the classification outcomes is provided through the confusion matrix shown in Figure 3. The Support Vector Machine (SVM) model was implemented using the LinearSVC classifier with default parameter settings. A linear kernel was chosen because of its computational efficiency and suitability for evaluating baseline classification performance. Hyperparameter tuning was not conducted, and the model was trained using the default configuration provided by the scikit-learn library.

Based on the SVM algorithm's confusion matrix, it appears that there are still quite a few classification errors in several stress level categories. The main diagonal values in the confusion matrix are not as strong as those obtained by the KNN algorithm, indicating that the SVM model's classification ability is still suboptimal.

Several stress level categories experienced significant mispredictions to other classes. This indicates that the SVM model has difficulty accurately separating data patterns, especially in data with overlapping distributions.

The poor performance of the SVM algorithm's confusion matrix indicates that the linear approach used was unable to capture the complexity of the data patterns in the student survey dataset used in this study.

TABLE 4.
SVM RESULT

Matrix	Value
Accuracy	46.15%
Precision	0.39
Recall	0.46
F1-score	0.41

The evaluation results showed that the SVM model struggled to accurately separate several stress level categories. This was evident in the confusion matrix, where a significant amount of data was misclassified into other categories.

One of the main factors contributing to the low performance of the SVM was the use of a linear kernel on a dataset with a complex and non-linear data distribution. Under these conditions, the linear hyperplane formed by the SVM was unable to optimally separate the data classes.

Unlike KNN, which operates based on data similarity, the SVM attempts to find the best dividing boundary between classes. When data distribution patterns overlap, the performance of the linear SVM model can experience a significant decline.

Furthermore, the characteristics of student survey-based datasets also tend to have subjective variations in answers between respondents. This condition makes the relationship between features and stress level categories more difficult to separate linearly.

The results of this study indicate that the performance of the SVM algorithm is significantly influenced by the characteristics of the dataset and the kernel selection used. The use of a non-linear kernel such as the Radial Basis Function (RBF) has the potential to improve model performance in future research.

D. Comparative Analysis of KNN and SVM Algorithms

The performance comparison between the KNN and SVM algorithms is shown in Table 3.

TABLE 5.
COMPARISON RESULTS

Metrik	KKN	SVM
Accuracy	88.46%	46.15%
Precision	0.90	0.39
Recall	0.88	0.46
F1-Score	0.89	0.41

Based on the evaluation results, the KNN algorithm demonstrated significantly better performance than SVM across all test metrics. KNN achieved an accuracy of 88.46%, while SVM only achieved an accuracy of 46.15%.

This difference in performance indicates that the student stress level dataset used in the study is more suitable for a distance-based similarity classification approach than a linear separation approach. KNN is able to better recognize proximity patterns between data, resulting in a lower classification error rate. This is supported by the confusion matrix results, which show that KNN has a higher number of correct predictions than SVM. In contrast, the linear SVM model has difficulty determining class boundaries due to overlapping and non-linear data distributions. This leads to a higher classification error rate. From a practical implementation perspective, the high performance of KNN makes this algorithm more suitable for use in predicting student stress levels. The model's ability to produce more accurate predictions can aid in the early identification of student stress.

The results of this study also demonstrate that the choice of machine learning algorithm needs to be tailored to the

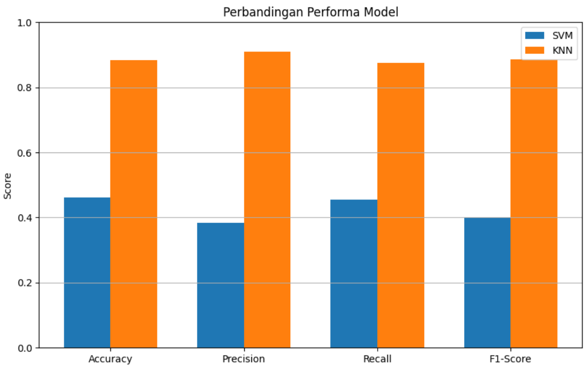

Figure 4. model performance comparison

characteristics of the dataset used. Although SVM is known as a powerful classification algorithm, in this study, KNN proved more effective for the student survey dataset.

E. Discussion

The substantial performance difference observed between KNN and SVM can be attributed to the characteristics of the dataset used in this study. KNN achieved an accuracy of 88.46%, whereas SVM only obtained 46.15%. This indicates that the student stress dataset contains local similarity patterns that are more effectively captured by distance-based classification methods.

KNN determines the class of a sample by considering the labels of neighboring observations. Since the predictor variables in this study, such as sleep quality, headache frequency, academic performance, study load, and extracurricular activities, exhibit gradual changes between stress categories, the similarity-based mechanism employed by KNN can recognize these patterns effectively [21].

In contrast, the SVM model was implemented using the LinearSVC classifier with default parameters and without hyperparameter optimization. The use of a linear decision boundary may not adequately represent complex and overlapping relationships among stress-related variables. Consequently, several stress categories were misclassified, leading to a significant reduction in predictive performance.

The confusion matrix analysis further supports these findings, showing that KNN produced fewer classification errors compared to SVM. Therefore, KNN appears to be more suitable for classifying student stress levels in survey-based datasets characterized by subjective responses and overlapping class distributions.

In addition to the comparative performance analysis, the variables employed in this study are widely recognized as factors associated with students' perceived stress levels. Sleep quality and study load have frequently been identified as important determinants of psychological well-being among university students [22]. Poor sleep quality may reduce concentration, impair emotional regulation, and increase susceptibility to stress [23]. while excessive academic workload can contribute to psychological pressure, particularly when students encounter difficulties in managing assignments, examinations, and other academic responsibilities.

Headache frequency may also represent physical manifestations related to psychological distress, whereas participation in extracurricular activities can provide opportunities for social interaction, recreation, and emotional regulation that potentially help alleviate stress [24]. This study focused on comparative classification performance rather than identifying the relative contribution of individual predictors. Nevertheless, the selected variables were included based on their relevance to student behavioral conditions and psychological well-being reported in previous studies. Therefore, the prediction results should be interpreted as estimates of perceived stress levels influenced by behavioral and academic factors rather than clinical assessments of mental disorders.

From a practical perspective, the developed stress prediction system may serve as an early screening tool for identifying students who potentially experience elevated stress levels. The prediction results can be utilized by students as a form of self-awareness regarding their psychological condition, while academic advisors may use the information to provide timely academic guidance and support. In addition, university counseling services may benefit from the system by identifying students who require further attention and recommending appropriate interventions. Nevertheless, the system is intended to complement existing psychological support services and should not be considered a substitute for professional mental health assessment or diagnosis.

Several limitations of this study should be acknowledged. First, the dataset utilized in this research was obtained from Kaggle and consisted of secondary survey data collected from respondents with unspecified demographic backgrounds. Consequently, the findings may not fully represent the characteristics of students from different universities, regions, or cultural settings. Second, the stress labels were derived from respondents' self-assessments rather than standardized psychological instruments, which may affect the consistency of the reported stress levels. Finally, this study focused on comparing conventional machine learning algorithms using a single train-test split scenario. Future studies are encouraged to evaluate the proposed models under more diverse experimental settings and larger datasets collected through validated psychological assessments.

The findings of this study are generally in agreement with previous research that demonstrated the effectiveness of machine learning approaches for stress prediction among students. Zahrah and Muljono [23]. reported that machine learning models can be utilized as effective tools for early stress detection in educational settings. Similarly, Tariq et al [2]. highlighted the potential of predictive modeling techniques for identifying stress conditions among university students. Daza et al [6]. also indicated through a systematic review that machine learning algorithms have shown promising performance in predicting anxiety and stress in higher education environments. In comparison, the present study employed a multi-class stress dataset consisting of five stress categories and found that KNN achieved an accuracy of 88.46%, outperforming the LinearSVC model. Furthermore, the proposed model was integrated into a Flask-based web application supported by a MySQL database, allowing users to perform stress level predictions through an interactive web interface.

IV. CONCLUSION

The experimental findings demonstrate that the choice of classification algorithm significantly influences the accuracy of student stress prediction. Among the two evaluated models, K-Nearest Neighbors (KNN) consistently produced superior results, achieving an accuracy of 88.46% and a weighted F1-score of 0.89. In contrast, the Support Vector Machine (SVM) model showed considerably lower predictive performance, indicating that a linear decision boundary was less effective in representing the characteristics of the dataset.

The better performance achieved by KNN suggests that the student stress dataset contains patterns that are more effectively captured through similarity-based classification. By utilizing neighboring data instances during prediction, KNN was able to classify stress categories with fewer misclassifications compared to SVM, as reflected in the confusion matrix analysis.

Beyond the comparative evaluation, this research also demonstrates the practical applicability of machine learning in educational environments through the development of a Flask-based web application. The integration of the trained model into an interactive platform enables real-time stress level prediction and provides a foundation for future intelligent systems aimed at supporting student well-being. Future studies may explore larger datasets, additional predictive factors, and alternative machine learning approaches to further improve prediction performance. Future studies are encouraged to utilize datasets collected using validated psychological assessment instruments to improve the reliability and generalizability of stress prediction models.

REFERENCES

[1] R. Ahmed, M. Bibi, and S. Syed, “Improving Heart Disease Prediction Accuracy Using a Hybrid Machine Learning Approach: A Comparative study of SVM and KNN Algorithms,” *Int. J. Comput. Inf. Manuf. IJCIM*, vol. 3, no. 1, pp. 49–54, Jun. 2023, doi: 10.54489/ijcim.v3i1.223.

[2] R. Tariq, M. G. Orozco-del-Castillo, M. T. Zamir, M. S. Ramírez-Montoya, and T. Wilberforce, “Explainable artificial intelligence for predictive modeling of student stress in higher education,” *Sci. Rep.*, vol. 15, no. 1, p. 38375, Nov. 2025, doi: 10.1038/s41598-025-22171-3.

[3] S. Anisa, A. Komarudin, and E. Ramadhan, “Sistem Klasifikasi Untuk Menentukan Tingkat Stress Mahasiswa Secara Umum Menggunakan Metode K-Nearest Neighbors,” *J. Inform. Teknol. Dan Sains Jinteks*, vol. 6, no. 3, pp. 568–578, Aug. 2024, doi: 10.51401/jinteks.v6i3.4317.

[4] Asri Mulyani, Sarah Khoerunisa, and Dede Kurniadi, “Perbandingan Kinerja Algoritma KNN dan SVM Menggunakan SMOTE untuk Klasifikasi Penyakit Diabetes,” *J. Nas. Tek. Elektro Dan Teknol. Inf.*, vol. 14, no. 1, pp. 25–34, Feb. 2025, doi: 10.22146/jnteti.v14i1.15198.

[5] S. Wiyono, D. S. Wibowo, M. F. Hidayatullah, and D. Dairoh, “Comparative Study of KNN, SVM and Decision Tree Algorithm for Student’s Performance Prediction,” *Int. J. Comput. Sci. Appl. Math.*, vol. 6, no. 2, p. 50, Aug. 2020, doi: 10.12962/j24775401.v6i2.4360.

[6] A. Daza, N. Saboya, J. I. Necochea-Chamorro, K. Zavaleta Ramos, and Y. D. R. Vásquez Valencia, “Systematic review of machine learning techniques to predict anxiety and stress in college students,” *Inform. Med. Unlocked*, vol. 43, p. 101391, 2023, doi: 10.1016/j.imu.2023.101391.

[7] M. Marjanović, M. Kovačević, B. Bajat, and V. Voženílek, “Landslide susceptibility assessment using SVM machine learning algorithm,” *Eng. Geol.*, vol. 123, no. 3, pp. 225–234, Nov. 2011, doi: 10.1016/j.enggeo.2011.09.006.

[8] C. Eklund, A. Söderlund, and M. L. Elfström, “Evaluation of a Web-Based Stress Management Program for Persons Experiencing Work-Related Stress in Sweden (My Stress Control): Randomized Controlled Trial,” *JMIR Ment. Health*, vol. 8, no. 12, p. e17314, Dec. 2021, doi: 10.2196/17314.

[9] U. Madububambachu, A. Ukpebor, and U. Ihezue, “Machine Learning Techniques to Predict Mental Health Diagnoses: A Systematic Literature Review,” *Clin. Pract. Epidemiol. Ment. Health*, vol. 20, no. 1, p. e17450179315688, Jul. 2024, doi: 10.2174/0117450179315688240607052117.

[10] M. Fadhilla, R. Wandri, A. Hanafiah, P. R. Setiawan, Y. Arta, and S. Daulay, “Analisis Performa Algoritma Machine Learning Untuk Identifikasi Depresi Pada Mahasiswa,” vol. 5, no. 1, 2025.

[11] S. Gedam, S. Dutta, and R. Jha, “Analyzing mental stress in Indian students through advanced machine learning and wearable technologies,” *Sci. Rep.*, vol. 15, no. 1, p. 20610, Jul. 2025, doi: 10.1038/s41598-025-06918-6.

[12] N. Hafidhoh, A. P. Atmaja, G. N. Syaifuddiin, I. B. Sumafta, S. M. Pratama, and H. N. Khasanah, “Machine Learning untuk Prediksi Kegagalan Mesin dalam Predictive Maintenance System,” *J. Masy. Inform.*, vol. 15, no. 1, pp. 56–66, May 2024, doi: 10.14710/jmasif.15.1.63641.

[13] P. Lialiou and I. Maglogiannis, “Students’ Burnout Symptoms Detection Using Smartwatch Wearable Devices: A Systematic Literature Review,” *AI Sens.*, vol. 1, no. 1, p. 2, May 2025, doi: 10.3390/aisens1010002.

[14] M. Mustaqim, A. Gunawan, Y. B. Pratama, and I. Zaliman, “Pengembangan Chatbot Layanan Publik Menggunakan Machine Learning Dan Natural Languange Processing,” *J. Inf. Technol. Soc.*, vol. 1, no. 1, pp. 1–4, Jun. 2023, doi: 10.35438/jits.v1i1.16.

[15] M. S. I. Ovi, J. Hossain, M. R. A. Rahi, and F. Akter, “Protecting Student Mental Health with a Context-Aware Machine Learning Framework for Stress Monitoring,” Aug. 01, 2025, *arXiv*: arXiv:2508.01105. doi: 10.48550/arXiv.2508.01105.

[16] E. H. Aldian Karim, “Implementasi Machine Learning untuk Prediksi Harga Mobil Bekas dengan Algoritma Regresi Linear berbasis Web,” *J. Ilm. Komputasi*, vol. 21, no. 4, pp. 595–602, Dec. 2022, doi: 10.32409/jikstik.21.4.3327.

[17] P. P. P. A. N. W. F. I. R.H. Zer, B. H. Hayadi, and A. R. Damanik, “Pendekatan Machine Learning Menggunakan Algoritma C4.5 Berbasis Pso Dalam Analisa Pemahaman Pemrograman Website,” *J. Inform. Dan Tek. Elektro Terap.*, vol. 10, no. 3, pp. 150–156, Aug. 2022, doi: 10.23960/jitet.v10i3.2700.

[18] D. I. Sumantiawan, S. Amaliyah, S. Narulita, M. Kholilurrahman, and A. Y. Suharmanto, “Analisis Faktor Status Depresi Kehidupan Mahasiswa Menggunakan Machine Learning,” *Digit. Transform. Technol.*, vol. 4, no. 1, pp. 705–713, Aug. 2024, doi: 10.47709/digitech.v4i1.4577.

[19] A. Ramadhani, N. Iriadi, and R. Hidayat, “Implementasi Teknologi Rest API Dengan Node Js Untuk Aplikasi Rekomendasi Destinasi Wisata,” vol. 4, no. 1, pp. 22–29, 2025.

[20] A. Y. Shdefat, N. Mostafa, Z. Al-Arnaout, Y. Kotb, and S. Alabed, “Optimizing HAR Systems: Comparative Analysis of Enhanced SVM and k-NN Classifiers,” *Int. J. Comput. Intell. Syst.*, vol. 17, no. 1, p. 150, Jun. 2024, doi: 10.1007/s44196-024-00554-0.

[21] M. Zeidan, R. Ballout, N. Al-Akl, and S. A. Kharroubi, “Exploring the relationship between extracurricular activities and stress levels among university students: A cross-sectional study,” *PLOS One*, vol. 20, no. 8, p. e0329888, Aug. 2025, doi: 10.1371/journal.pone.0329888.

[22] E. Wijayanti, A. Eka Yuliani, and S. Bulut, “The Relationship Between Sleep Quality and Stress Level in Biology Education Students,” *Indones. J. Sci. Educ.*, vol. 9, no. 1, pp. 49–57, Apr. 2025, doi: 10.31002/ijose.v9i1.1703.

[23] F. N. Zahrah and M. Muljono, “Machine Learning untuk Deteksi Stres Pelajar: Perceptron sebagai Model Klasifikasi Efektif untuk Intervensi Dini,” *Edumatic J. Pendidik. Inform.*, vol. 8, no. 2, pp. 764–773, Dec. 2024, doi: 10.29408/edumatic.v8i2.28011.

[24] G. Barbayannis, M. Bandari, X. Zheng, H. Baquerizo, K. W. Pecor, and X. Ming, “Academic Stress and Mental Well-Being in College Students: Correlations, Affected Groups, and COVID-19,” *Front. Psychol.*, vol. 13, p. 886344, May 2022, doi: 10.3389/fpsyg.2022.886344.