Indonesian Journal of Electrical Engineering and Computer Science
Vol. 44, No. 1, October 2026, pp. 186~196
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v44.i1.pp186-196

Data-driven psychological state classification using artificial intelligence

Yudi Ramdhani[1], Win Ce[2], Ghanim Kanugrahan[3], Yulita Anggelia[4], Febriyanti Panjaitan[5], Rizki Tri Prasetio[1]

[1]Department of Information System, Faculty of Creative Technology, Satu University, Bandung, Indonesia
[2]Information of Systems Department, School of Information Systems, Binus University, Jakarta, Indonesia
[3]Department of Computer Science, Faculty of Creative Technology, Satu University, Bandung, Indonesia
[4]Department of Psychology, Faculty of Business and Humanities, Satu University, Bandung, Indonesia
[5]Department of Computer Science, Faculty of Creative Technology, Satu University, Palembang, Indonesia

Article Info

Article history:

Received Feb 10, 2026
Revised Aug 26, 2026
Accepted Sep 18, 2026

Keywords:

Anxiety
DASS-21
Depression
Machine learning
Neural network
Stress
Support vector machine

ABSTRACT

This study uses data collected from university students through the depression anxiety stress scales-21 (DASS-21) questionnaire to assess how effectively various machine learning algorithms predict degrees of depression, anxiety, and stress. The algorithms studied were K-nearest neighbors (KNN), Naive Bayes, decision tree (DT), random forest (RF), support vector machine (SVM), and neural network (NN). Among them, the NN achieved the highest overall performance, demonstrating superior accuracy and consistency across all categories. The SVM also demonstrated competitive performance, achieving high accuracy and precision, although slightly below that of the NN. RF and DT produced decent results; however, KNN and Naive Bayes were less reliable, particularly in predicting anxiety. Overall, the findings suggest that NN represents a promising approach for predicting mental health conditions using DASS-21 data. Still, more testing with different datasets would be beneficial in determining how well it performs in general. DT or RF may be suitable alternatives in situations requiring minimal processing capacity.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Win Ce
Information of Systems Department
School of Information Systems, Binus University
Jakarta City, Indonesia
Email: wn@binus.edu

1. INTRODUCTION

With the rapid growth of the internet and technology, data is now being generated in massive volumes and various formats. Advances in computing have also made storing this data cheaper and more efficient. However, raw data holds little value unless it's processed with a clear purpose. To effectively collect, store, and analyze large-scale data, the right tools and techniques are essential. Data mining offers a solution by combining statistics, machine learning, and computing methods to uncover hidden patterns and relationships that traditional methods may miss. As data continues to grow both structured and unstructured managing and making sense of it has become a major challenge that requires smart strategies [1], [2]. Each data mining algorithm has its own strengths and weaknesses, as they are designed to solve specific types of problems. That's why it's important to clearly understand and define the problem before choosing the right algorithm. Knowing what each algorithm is best suited for helps in selecting the most effective one for the task [2]. Machine learning has advanced rapidly in recent years, transforming key sectors like healthcare,

finance, engineering, and education. As a branch of artificial intelligence, it allows systems to analyze data and make decisions autonomously with minimal human input. For classification tasks, neural network (NN) and support vector machine (SVM) have been extensively applied in areas dealing with image-based and language-related data. These approaches belong to supervised learning, where models are trained on labeled datasets. These developments are enabled by advancements in data collection and storage technology [3], [4].

People today are more ambitious and continuously looking for opportunities to advance in their careers. As a result, problems such as anxiety, despair, stress, and dissatisfaction have grown so widespread that many people regard them as typical aspects of professional life. According to the World Health Organization (WHO), depression is the most frequent mental illness, impacting over 300 million people globally. This increased worry has prompted many health professionals to concentrate on mental health studies. However, machines continue to struggle to distinguish between anxiety, depression, and stress, emphasizing the need for effective learning algorithms to enable accurate diagnoses. The WHO emphasizes that true health comprises both physical and mental well-being [5]. Depression is often diagnosed using the patient health questionnaire (PHQ), whereas the DASS-21 is commonly used to screen for symptoms of depression, anxiety, and stress [6]. Common symptoms of depression include memory issues, trouble concentrating, loss of interest, appetite changes, feelings of worthlessness, and suicidal thoughts, all of which can affect school, work, and social life. Anxiety often shows up as nervousness, irritability, fatigue, insomnia, digestive problems, panic attacks, and difficulty focusing. Stress symptoms may include restlessness, trouble relaxing, low energy, frequent headaches, overreacting, and a weakened immune system. Many symptoms of mental health conditions—like insomnia, chest pain, fatigue, and trouble focusing—often appear in more than one disorder. Because of this overlap, it can be difficult for machines to correctly classify these conditions [5]-[7]. Mental health problems are often related to social and emotional factors such as problems with friends, parents getting divorced, or pressure from school. In more serious cases, these issues can lead to drug use, self-harm, dropping out of school, or even thoughts of suicide. That's why it's very important to detect the symptoms early and accurately, so people can get the right treatment through therapy or medication. Psychological assessments are usually done using two types of tests: self-report questionnaires and projective tests [8]-[11]. Projective tests, like the Rorschach test and the house-tree-person (HTP), ask people to respond in ways that reveal their personality and mental health. On the other hand, self-report tests let people answer questions about their feelings and thoughts. Each method has its own strengths and weaknesses [11], [12].

Machine learning can help by making fast and accurate predictions or diagnoses, and it also helps reduce the time needed for psychological analysis. As intelligent systems, they improve their classification and prediction abilities through continuous learning from feedback. In psychological assessment, machine learning is increasingly used to evaluate aspects such as subjective well-being, personality traits, and mental health conditions like stress, anxiety, and depression. Several systematic reviews have highlighted its strong potential in measuring psychological constructs and detecting mental disorders [12]-[14]. Machine learning can be an effective method for detecting psychopathological signs in university students. To date, limited research has systematically compared multiple machine learning algorithms for classifying psychological severity levels among university students. While machine learning has been widely applied to detect stress, anxiety, and depression in the general population, studies specifically focusing on undergraduate students remain scarce. Therefore, this study evaluates and compares the performance of several machine learning models in predicting stress, anxiety, and depression levels among university students. The objective is to identify the most appropriate algorithm for mental health classification within this population [15]-[17].

Using DASS-21 data from diverse participants, the study tested five machine learning algorithms to classify mental health conditions by severity. Due to class imbalance in the confusion matrix, F1-score was added as an evaluation metric, with random forest (RF) showing the highest overall accuracy [5]. A previous study applied data mining techniques to analyze DASS-42 responses from 680 education students at Necmettin Erbakan University. The researchers applied several algorithms, including Naive Bayes (NB), neural network (NN), logistic regression (LR), SVM, and RF. The results showed that the artificial neural network (ANN) had the best accuracy for detecting depression (99.26%), anxiety (98.67%), and stress (97.35%). These results show that ANN can recognize complex patterns in psychological data [2]. Other recent studies have also examined stress related to work by using feature selection methods like Chi-Square, Information Gain, and PCA. These features were combined with classifiers such as Naive Bayes, RF, SVM, ANN, and gradient boosted trees (GBT), mostly in the insurance sector. The best performance came from combining correlation-based feature selection with SVM, showing that filter-based techniques are useful for classifying stress levels [18].

This study focused on detecting cyberbullying content on Twitter using labeled data validated by linguists. It compared classifiers like NB, decision tree (DT), LR, and SVM, applying Bag of Words with unigram, bigram, and trigram weighting. SVM achieved the highest accuracy (76%) and F-measure (82%), while DT had the best recall. The results highlight SVM's effectiveness in identifying cyberbullying on social media [19]. Earlier studies used simple metrics and limited linguistic cues like pronouns and sentiment

words to assess psychological states. Later research expanded this by including emotion, sentiment polarity, and feelings to improve detection. Using a large forum dataset on anxiety, depression, suicidal thoughts, and normal posts, the study showed that combining these features with machine learning, especially SVM and deep learning, improved accuracy in identifying psychological conditions through text [20]. A previous study predicted psychological well-being (PWB) using machine learning on a large U.S. dataset. PWB was measured through self-reported happiness, financial satisfaction, and job satisfaction. K-nearest neighbors (KNN) was used to rank feature importance, revealing that marital status had a stronger impact on happiness than wealth, while job prestige was key to job satisfaction. Forward Selection helped identify the best variable combination, allowing the model to classify low PWB individuals with 70–80% accuracy. The study also applied ordered probit regression to assess each factor's influence [21]. Stress is a global health issue with serious psychological impacts, making early detection essential. In one study, stress data were collected using dry-electrode EEG during the stroop color word task (SCWT). EEG signals were analyzed using fast fourier transform (FFT) to extract average power across four frequency bands. Algorithms like DA, DT, KNN, NB, and SVM were tested for accuracy, sensitivity, and specificity. SVM outperformed the others, achieving the highest accuracy up to 99.98%, particularly in the alpha and beta bands, making it the most effective method for stress detection [22].

Breast cancer remains a major global health threat for women, highlighting the importance of early diagnosis. In prior research, machine learning was applied to improve breast cancer prediction, testing ten algorithms such as DT, LR, NB, SVM, and gradient boosting. Performance was assessed using six key measures: specificity, sensitivity, precision, accuracy, F1-score, and ROC curve analysis. The results showed Gradient Boosting outperforming other models, reaching an impressive F1-score of 96.77% [23]. Landslides are natural disasters that can threaten lives and damage infrastructure. A previous study compared three models for mapping landslide vulnerability in Van Yen District, Yen Bai Province: Analytic Hierarchy Process, Frequency Ratio, and RF. Using 70% of the data for training and 30% for testing, the study considered 13 landslide-triggering factors. While AHP and FR showed good results with AUC scores of 0.842 and 0.852, RF outperformed both with an AUC of 0.949. The findings highlight the strong potential of machine learning in landslide vulnerability analysis [24]. Modern predictive systems are increasingly adopting machine learning and deep learning techniques to forecast household food security, enabling policymakers to deliver timely aid during crises. A recent analysis of Ethiopia's Household Income, Consumption, and Expenditure survey compared multiple predictive approaches, including traditional methods (KNN, LR, SVM, NB) and advanced neural networks (MLP, ANN, CNN). The ANN demonstrated superior performance across all evaluation metrics, achieving 99.15% accuracy along with exceptional scores in precision, recall, F1 measure, mean absolute error, and root mean squared error [25].

This study examined mental health disorders, specifically depression, anxiety, and stress, using responses from the DASS-21 questionnaire. The primary purpose is to investigate how machine learning algorithms can categorize psychological states into five severity levels: normal, mild, moderate, severe, and extremely severe. This study compares the predictive performance of different categorization models, including KNN, DT, RF, SVM, NB, and NN. Machine learning is seen as an appropriate method because of its capacity to manage enormous datasets and create correct classifications, whereas traditional statistical approaches frequently struggle with sophisticated data processing. As a result, this study uses machine learning approaches to aid in the evaluation of mental health disorders utilizing DASS-21 data.

2. METHOD

This research evaluates how effectively various machine learning models can assess and predict anxiety, depression, and stress levels using DASS-21 survey data. The responses are categorized into five severity levels: Normal, Mild, Moderate, Severe, and Extremely Severe. Figure 1 illustrates the research workflow from data preparation through final classification. We tested multiple algorithms, including DTs, SVM, Naive Bayes, RF, KNN, and NNs, to determine which performs best at predicting mental health states from questionnaire responses. The primary objective is to identify the most accurate model for detecting different intensities of these psychological conditions.

2.1. Dataset

The dataset used in this research was collected from the DASS-21 questionnaire. This questionnaire was available from 2017 to 2019 and could be filled out by anyone. University students completed it voluntarily to get personal feedback about their mental health. Only the responses from participants who gave permission for their data to be used in research were included. The study collected 39,775 responses using the DASS-21 questionnaire. This assessment consists of 21 items divided equally into three subscales:

depression, anxiety, and stress. Participants rated each statement on a 4-point frequency scale indicating how often they experienced each symptom during the previous week, with these response options:

a) 1 = Never (did not apply to me at all)
b) 2 = Sometimes (applied to me occasionally)
c) 3 = Often (applied to me frequently)
d) 4 = Almost always (applied to me most of the time)

The data in this study were collected using the DASS-21 questionnaire, which consists of 21 questions. Each group of symptoms stress, anxiety, and depression includes seven questions. Responses to each question could be provided either in text or numeric form [5]. The details of the DASS-21 questionnaire items are presented in Table 1, which lists the questions related to anxiety, depression, and stress.

After calculating the final score for each respondent based on their DASS-21 questionnaire responses, the data were then classified into five levels of severity: Normal, Mild, Moderate, Severe, and Extremely Severe. The classification criteria for each severity level are shown in Table 2 [5].

The distribution of the dataset by severity level shows that most respondents fall into the Extremely Severe category, with 16,008 individuals for anxiety, 13,738 for depression, and 6,240 for stress. This distribution indicates that severe anxiety symptoms are more dominant in the dataset, while stress symptoms are more commonly found in the Normal category. In total, the dataset analyzed in this study includes 39,775 respondents. The detailed distribution can be seen in Table 3.

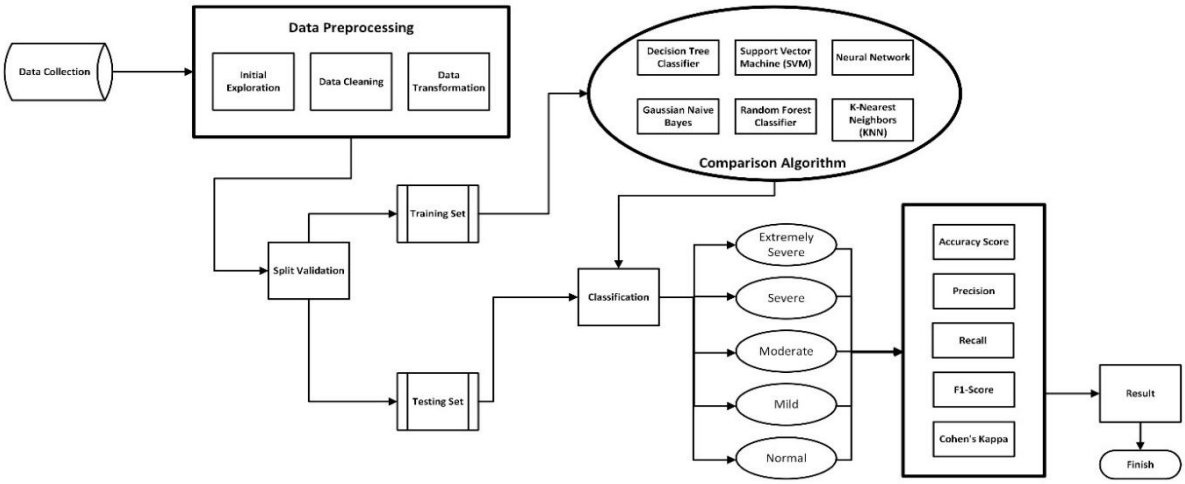

Figure 1. Research method

Table 1. Questionnaires on anxiety, depression and stress

No	Anxiety	Depression	Stress
1	I was aware of dryness of my mouth	I couldn't seem to experience any positive feeling at all	I tended to over-react to situations
2	I experienced breathing difficulty	I felt that I had nothing to look forward to	I found it difficult to relax
3	I felt scared without any good reason	I felt I wasn't worth much as a person	I felt that I was using a lot of nervous energy
4	I was aware of the action of my heart in the absence of physical exertion	I felt down-hearted and blue	I felt that I was rather touchy
5	I felt I was close to panic	I was unable to become enthusiastic about anything	I found it hard to wind down
6	I was worried about situations in which I might panic and make a fool of myself	I felt that life was meaningless	I found that I was very irritable
7	I experienced trembling	I found it difficult to work up the initiative to do things	I found myself getting agitated

Table 2. Severity levels

Symptom level	Anxiety score	Depression score	Stress score
Normal	0-7	0-9	0-14
Mild	8-9	10-13	15-18
Moderate	10-14	14-20	19-25
Severe	15-19	21-27	26-33
Extremely Severe	20+	28+	33+

Table 3. Dataset distribution

Severity level	Anxiety cases	Depression cases	Stress case
Normal	8902	8480	13300
Mild	2473	3771	4959
Moderate	7687	8106	7166
Severe	4705	5680	8110
Extremely Severe	16008	13738	6240
Total		39775	

In this study, bar charts are used to demonstrate how respondents are classified according to the intensity of anxiety, depression, and stress. The statistics reveal that the majority of respondents fell into the Extremely Severe group, particularly for anxiety and depression. In terms of stress, however, the majority of people fall under the Normal category. This is obvious from the height of the bars, with the bars for acute anxiety and sadness being the tallest. The charts demonstrate that many participants in this sample exhibit strong symptoms of anxiety and depression, but stress levels are more evenly distributed between Normal and Severe. These visual results correspond to the data presented earlier in Figure 2.

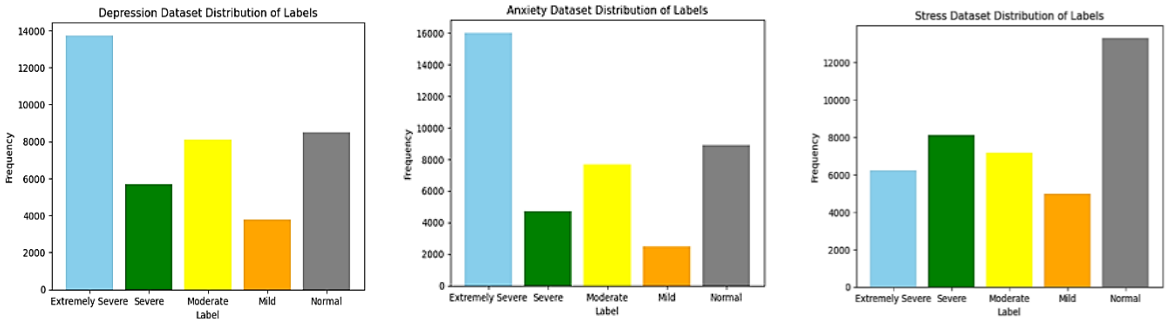

Figure 2. Dataset distribution graph

2.2. Data preprocessing

In the data exploration stage, we started by looking at the dataset to understand its structure, the number of records, the types of variables, and how the labels were distributed. We checked how many participants were in each category of anxiety, depression, and stress, and also reviewed the severity levels. We also made some visualizations to help see the data patterns more clearly before starting the modeling process. During the data cleaning step, we worked to improve the quality of the data by removing duplicate entries, fixing missing values, and making sure the important columns were complete. We also removed any data that was not related to the study, so it would not affect the results. After that, in the data transformation stage, we changed categorical values into numbers to make the data suitable for machine learning. We also applied normalization and label encoding so that the data would be in the right format. Finally, we split the dataset into features (X) and labels (y) and then divided them into training and testing sets. These steps are important to help the model work more accurately and reliably in the next phases.

2.3. Split data

Following data preprocessing, we split the dataset into training and testing subsets using a 70-30 ratio. This standard practice enables models to learn patterns from the majority of available data while reserving a portion for unbiased performance evaluation. The training set (70%) served to develop the predictive models, while the testing set (30%) provided unseen data for validation. Table 4 presents the detailed distribution of cases across all categories in both training and testing partitions.

Table 4. Split data distribution

	Data training			Data testing		
	Depression	Anxiety	Stress	Depression	Anxiety	Stress
Extremely Severe	9659	11261	4393	4079	4747	1847
Severe	2622	1726	3498	1149	747	1461
Moderate	5668	5357	5008	2438	2330	2158
Mild	5922	6199	9248	2558	2703	4052
Normal	3971	3299	5695	1709	1406	2415
Total		27842			11933	

2.4. Modeling

2.4.1. Naive Bayes

The Naive Bayes classifier is a straightforward probabilistic model based on fundamental statistical principles. Originally developed from the work of 18th-century mathematician Thomas Bayes, this approach has become widely adopted in machine learning due to its computational efficiency and intuitive probabilistic foundation. Its simplicity and speed make it particularly useful for classification tasks where quick, interpretable results are needed. The algorithm assumes that all input features are independent and have the same importance in predicting the result. Although this assumption is not always true, Naive Bayes still works well in many situations. The key idea of this method is based on three main concepts: prior probability, likelihood, and posterior probability [26].

2.4.2. Random forest

RF is a machine learning technique that generates several DTs from random subsets of a dataset. It then aggregates the data from all of the trees to get a final prediction. Classification tasks require a majority vote. If it is regression, it computes the average. This strategy improves the model's accuracy and stability. It also minimizes the likelihood of overfitting, which frequently occurs with single DTs. RF works effectively with huge datasets, especially ones with many attributes and complex interactions among the data [27]. The model diagram of the RF algorithm is shown in Figure 3.

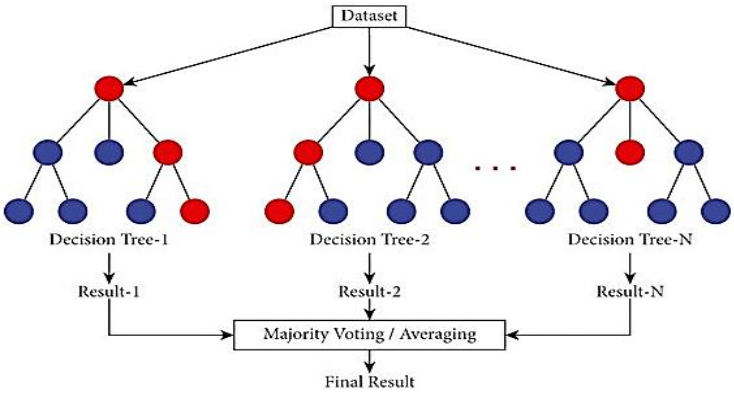

Figure 3. RF diagram [28]

2.4.3. KNN

KNN is a simple but useful machine learning algorithm. It can be used for both classification and regression. The main idea of KNN is to classify a new data point by looking at the "k" closest points in the dataset. It checks which class is the most common among those nearby points and then assigns that class to the new point. To find the closest points, KNN usually calculates the distance using a method called Euclidean distance. After the distances are calculated, the algorithm picks the k nearest points and checks which class appears the most often. That class becomes the prediction for the new data point [29], [30]. The basic formula for KNN using Euclidean distance is shown in (1).

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

In the KNN algorithm, a new instance (**x**) is classified by measuring its distance to every existing data point (**y**) in the feature space, where **n** indicates the total number of features. The system identifies the **k** closest neighbors based on these distance calculations and assigns the new point to the class that appears most frequently among them.

2.4.4. Decision tree

The DT algorithm employs a hierarchical branching structure to classify data through sequential feature-based splits as shown in Figure 4. Beginning at the root node, the model recursively partitions the dataset by selecting optimal features that maximize class separation at each decision point. The splitting process continues until terminal nodes (leaves) are reached, which contain the final classification predictions. This intuitive structure, where branches represent decision rules and leaves denote outcomes, makes DTs both interpretable and versatile for diverse datasets [31].

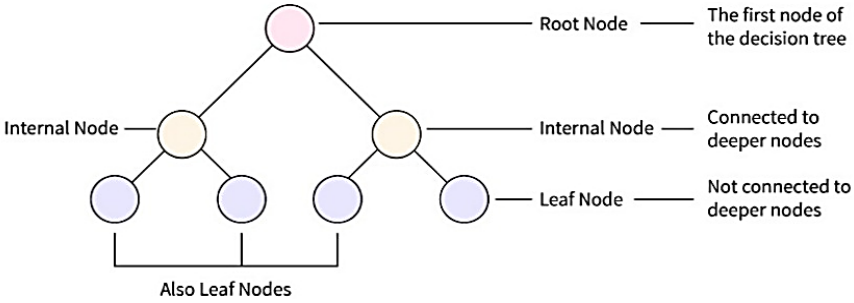

Figure 4. DT example [32]

2.4.5. Support vector machine

SVM is a machine learning method used for both classification and regression tasks as shown in Figure 5. Instead of just drawing any dividing line, SVM finds the *best possible* boundary (called a hyperplane) to separate different classes while maximizing the margin between them. What makes SVM powerful is its ability to classify data that isn't linearly separable. By using the kernel trick, it maps data into a higher-dimensional space where a clear separation becomes possible. This makes SVM highly effective for complex or high-dimensional datasets [33].

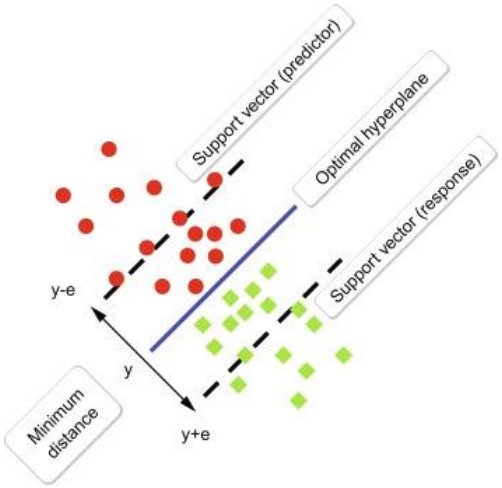

Figure 5. Illustration of the SVM concept

2.4.6. Neural network

A NN is a model inspired by how the human brain works. It has layers of "neurons" that are connected to each other. In classification problems, the data passes through an input layer, one or more hidden layers, and an output layer that gives the result. NNs are good at finding complex patterns in data. They learn by adjusting the strength of the connections between neurons using a method called backpropagation. One popular type of NN for image and visual data is the convolutional neural network (CNN) [34].

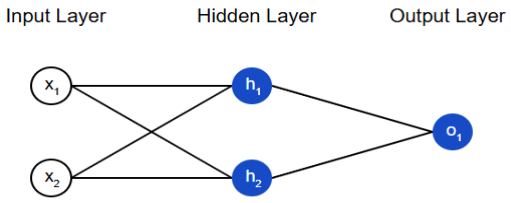

Figure 6. Illustration of NN algorithm [35]

2.5. Validation

To evaluate model performance, we employed several key metrics that collectively provide a comprehensive assessment. Accuracy measures the overall correctness of predictions by calculating the percentage of accurate classifications. Precision evaluates the model's ability to correctly identify positive cases by determining the ratio of true positives to all predicted positives. Recall assesses how effectively the model captures all actual positive instances from the dataset. The F1-score offers a balanced evaluation by combining precision and recall, which proves particularly valuable when working with imbalanced data distributions. Additionally, Cohen's Kappa provides insight into the agreement between predicted and actual classifications while accounting for chance agreement. Together, these metrics enable a robust and multi-faceted evaluation of our models' capabilities in classifying depression, anxiety, and stress levels with high reliability.

3. RESULTS AND DISCUSSION

3.1. Results

The classification results indicate clear differences in performance among the six machine learning models. The NN achieved the highest performance, with all predictions correctly classified across depression, stress, and anxiety severity levels. SVM demonstrated comparable results, with only minimal misclassifications. RF and DT also performed well overall, although some classification errors were observed, particularly in mild and moderate stress categories. In contrast, Naïve Bayes showed the highest misclassification rates, especially in anxiety and depression severity levels, where several extremely severe cases were assigned to incorrect classes. Overall, the findings consistently indicate that the NN outperformed the other models, followed by SVM, RF, DT, and Naïve Bayes. This ranking aligns with the comparative results obtained from accuracy, precision, recall, and F1-score evaluations.

The performance results of six machine learning algorithms are summarized in Table 5, highlighting their effectiveness in predicting depression, stress, and anxiety. For the KNN algorithm, accuracy ranged from 92.36% to 94.1%, with the highest error rate of 7.64% occurring in stress prediction. Both precision and recall were above 88%, and Cohen's Kappa was around 0.9, indicating a high level of agreement. RF also performed well, with accuracy over 96%, low error rates, and Cohen's Kappa above 0.96. The DT algorithm achieved accuracy between 95.57% and 96.66%, with similarly strong Cohen's Kappa values ranging from 0.94 to 0.96. In contrast, Naive Bayes showed the weakest performance, particularly for anxiety prediction, with the lowest accuracy at 82.45%, an error rate of 17.55%, and a Cohen's Kappa of 0.76, indicating moderate agreement. SVM delivered near-perfect results, achieving accuracy between 99.98% and 99.99%. However, among all models tested, the NN demonstrated the most optimal and consistent performance, reaching 100% accuracy, precision, recall, and F1-score across all categories, with a perfect Cohen's Kappa of 1, confirming that the model made no classification errors.

Table 5. Classification results

Model	Condition	Accuracy	Error rate	Positive accuracy	Recall	F1-score	Cohen's Kappa
KNN	Depression	94,1	5,9	92,69	92,27	92,41	0,92
	Stress	92,36	7,64	90,84	91,59	91,14	0,9
	Anxiety	92,74	7,26	89,17	88,89	88,86	0,9
SVM	Depression	99,98	0,02	99,97	99,99	99,98	1
	Stress	99,99	0,01	99,99	99,99	99,99	1
	Anxiety	99,98	0,02	99,99	99,97	99,98	1
Random forest	Depression	97,34	2,66	96,67	96,66	96,66	0,97
	Stress	96,76	3,24	96,07	96,21	96,14	0,96
	Anxiety	97,34	2,66	96,3	96,8	96,55	0,96
Decision tree	Depression	96,57	3,43	95,85	95,8	95,82	0,96
	Stress	95,57	4,43	94,63	94,9	94,75	0,94
	Anxiety	96,66	3,34	95,45	96,19	95,81	0,95
Naïve Bayes	Depression	89,21	10,79	86,25	87,79	86,7	0,86
	Stress	89,63	10,37	88,24	88,37	87,98	0,87
	Anxiety	82,45	17,55	71,57	68,09	66,62	0,76
Neural network	Depression	100	0	100	100	100	1
	Stress	100	0	100	100	100	1
	Anxiety	100	0	100	100	100	1

3.2. Limitations

Our assessment of various algorithms for predicting depression, anxiety, and stress levels revealed that NNs and SVM outperformed others, achieving nearly flawless accuracy (100%) across all mental health categories. This exceptional performance indicates their superior ability to detect intricate patterns in DASS-

21 responses. RF and DT also demonstrated strong results, with accuracy exceeding 95%, though they had marginally higher error rates than SVM and NNs. In contrast, KNN and Naive Bayes showed comparatively weaker performance, particularly in anxiety prediction, where accuracy dropped significantly. Naive Bayes' limitations likely stem from its inherent assumption of independent features an assumption that doesn't hold for DASS-21 data, where symptoms often correlate. These results highlight that models adept at modeling non-linear relationships (e.g., NNs, SVM) are better suited for mental health prediction tasks. However, the NNs perfect scores raise concerns about potential overfitting, especially since the study relied on a single dataset without rigorous cross-validation. To enhance reliability, future research should incorporate cross-validation and diverse datasets to test model generalizability. For real-world implementation, RF and DT remain viable options, offering a balance of high accuracy and computational efficiency making them practical alternatives to more complex NNs.

4. CONCLUSION

Our study evaluated different machine learning methods for predicting mental health conditions (depression, stress, and anxiety) using DASS-21 survey responses. The NN model achieved perfect results - identifying all cases correctly without any mistakes. The SVM method came very close to perfect accuracy, making only occasional minor errors. Both RF and DT approaches performed well, correctly identifying over 95% of cases, though not quite as precisely as the top two methods. On the other hand, Naive Bayes and KNN techniques were less effective, especially for anxiety detection, with more incorrect classifications compared to the other models. Based on these findings, the NN is the most recommended model for predicting levels of depression, stress, and anxiety due to its outstanding accuracy. However, further testing on different datasets is necessary to ensure its generalizability. For scenarios with limited computational resources, DT and RF remain efficient alternative options.

REFERENCES

[1] M. Chen, S. Mao, and Y. Liu, "Big data: a survey," *Mobile Networks and Applications*, vol. 19, no. 2, pp. 171–209, Jan. 2014, doi: 10.1007/s11036-013-0489-0.

[2] S. A. Sulak and N. Koklu, "Analysis of depression, anxiety, stress scale (DASS-42) with methods of data mining," *European Journal of Education*, vol. 59, no. 4, 2024, doi: 10.1111/ejed.12778.

[3] J.-Y. Yoon, G. Gweon, and Y. J. Yoo, "Supervised diagnostic classification of cognitive attributes using data augmentation," *PLOS ONE*, vol. 19, no. 1, p. e0296464, Jan. 2024, doi: 10.1371/journal.pone.0296464.

[4] N. K. Hikmawati, Y. Ramdhani, and D. P. Alamsyah, "Optimizing forest fire detection using PSO, neural networks, and k-Fold cross-validation," in *2024 Ninth International Conference on Informatics and Computing (ICIC)*, Oct. 2024, pp. 1–6, doi: 10.1109/icic64337.2024.10956756.

[5] A. Priya, S. Garg, and N. P. Tigga, "Predicting anxiety, depression and stress in modern life using machine learning algorithms," *Procedia Computer Science*, vol. 167, pp. 1258–1267, 2020, doi: 10.1016/j.procs.2020.03.442.

[6] C. Zanon *et al.*, "Examining the dimensionality, reliability, and invariance of the depression, anxiety, and stress scale–21 (DASS-21) across eight countries," *Assessment*, vol. 28, no. 6, pp. 1531–1544, Jan. 2020, doi: 10.1177/1073191119887449.

[7] A. Thapar, O. Eyre, V. Patel, and D. Brent, "Depression in young people," *The Lancet*, vol. 400, no. 10352, pp. 617–631, Aug. 2022, doi: 10.1016/s0140-6736(22)01012-1.

[8] J. Bantjes *et al.*, "Mental health and academic failure among first-year university students in South Africa," *South African Journal of Psychology*, vol. 51, no. 3, pp. 396–408, Oct. 2020, doi: 10.1177/0081246320963204.

[9] E. Kohls, S. Baldofski, R. Moeller, S.-L. Klemm, and C. Rummel-Kluge, "Mental health, social and emotional well-being, and perceived burdens of university students during COVID-19 pandemic lockdown in Germany," *Frontiers in Psychiatry*, vol. 12, Apr. 2021, doi: 10.3389/fpsyt.2021.643957.

[10] E. Sheldon *et al.*, "Prevalence and risk factors for mental health problems in university undergraduate students: A systematic review with meta-analysis," *Journal of Affective Disorders*, vol. 287, pp. 282–292, May 2021, doi: 10.1016/j.jad.2021.03.054.

[11] B. L. Schaab *et al.*, "How do machine learning models perform in the detection of depression, anxiety, and stress among undergraduate students? A systematic review," *Cadernos de Saúde Pública*, vol. 40, no. 11, 2024, doi: 10.1590/0102-311xen029323.

[12] J. McLeod, *Doing research in counselling and psychotherapy*. SAGE Publications Ltd, 2022.

[13] P. Kumar, S. Garg, and A. Garg, "Assessment of anxiety, depression and stress using machine learning models," *Procedia Computer Science*, vol. 171, pp. 1989–1998, 2020, doi: 10.1016/j.procs.2020.04.213.

[14] J. Chung and J. Teo, "Mental health prediction using machine learning: taxonomy, applications, and challenges," *Applied Computational Intelligence and Soft Computing*, vol. 2022, pp. 1–19, Jan. 2022, doi: 10.1155/2022/9970363.

[15] L. F. Ramos-Lima, V. Waikamp, T. Antonelli-Salgado, I. C. Passos, and L. H. M. Freitas, "The use of machine learning techniques in trauma-related disorders: a systematic review," *Journal of Psychiatric Research*, vol. 121, pp. 159–172, Feb. 2020, doi: 10.1016/j.jpsychires.2019.12.001.

[16] Y. Wu, K. Mao, L. Dennett, Y. Zhang, and J. Chen, "Systematic review of machine learning in PTSD studies for automated diagnosis evaluation," *npj Mental Health Research*, vol. 2, no. 1, 2023, doi: 10.1038/s44184-023-00035-w.

[17] R. S. McIntyre *et al.*, "Synthesizing the evidence for ketamine and esketamine in treatment-resistant depression: an international expert opinion on the available evidence and implementation," *American Journal of Psychiatry*, vol. 178, no. 5, pp. 383–399, May 2021, doi: 10.1176/appi.ajp.2020.20081251.

[18] A. Hashmi, "A comparative study of classification of occupational stress in the insurance sector using machine learning and filter feature selection techniques," *Revista Gestão Inovação e Tecnologias*, pp. 5781–5801, 2021, doi: 10.47059/revistageintec.v11i4.2623.

[19] A. Muzakir, H. Syaputra, and F. Panjaitan, "A comparative analysis of classification algorithms for cyberbullying crime detection: an experimental study of Twitter social media in Indonesia," *Scientific Journal of Informatics*, vol. 9, no. 2, pp. 133–138, Oct. 2022, doi: 10.15294/sji.v9i2.35149.
[20] X. Du and Y. Sun, "Linguistic features and psychological states: A machine-learning based approach," *Frontiers in Psychology*, vol. 13, 2022, doi: 10.3389/fpsyg.2022.955850.
[21] Y. Yang, "The economics of psychological well-being evidence from the united states using machine learning method," 2022, doi: 10.2139/ssrn.4105068.
[22] Y. Badr, F. Al-Shargie, U. Tariq, F. Babiloni, F. Al Mughairbi, and H. Al-Nashash, "Classification of mental stress using dry EEG electrodes and machine learning," in *2023 Advances in Science and Engineering Technology International Conferences (ASET)*, Feb. 2023, pp. 1–5, doi: 10.1109/aset56582.2023.10180884.
[23] R. R. Kadhim and M. Y. Kamil, "Comparison of machine learning models for breast cancer diagnosis," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 12, no. 1, p. 415, Mar. 2023, doi: 10.11591/ijai.v12.i1.pp415-421.
[24] T. D. Khuc *et al.*, "Comparison of multi-criteria decision making, statistics, and machine learning models for landslide susceptibility mapping in Van Yen District, Yen Bai Province, Vietnam," *International Journal of Geoinformatics*, vol. 19, no. 7, pp. 33–45, 2023.
[25] M. Nigus and H. L. Shashirekha, "A comparison of machine learning and deep learning models for predicting household food security status," *International Journal of Electrical and Electronics Research*, vol. 10, no. 2, pp. 308–311, 2022, doi: 10.37391/ijeer.100241.
[26] K. L. Kohsasih and Z. Situmorang, "Comparative analysis of C4.5 and Naïve Bayes algorithms in predicting cerebrovascular disease," *Jurnal Informatika*, vol. 9, no. 1, pp. 13–17, Apr. 2022, doi: 10.31294/inf.v9i1.11931.
[27] G. L. Postiglione *et al.*, "Race-performance parameters differentiating world-best from national-level swimmers: a race video analysis and machine-learning approach," *International Journal of Sports Physiology and Performance*, vol. 20, no. 8, pp. 1034–1042, Aug. 2025, doi: 10.1123/ijspp.2024-0486.
[28] M. Y. Khan, A. Qayoom, M. S. Nizami, M. S. Siddiqui, S. Wasi, and S. M. K.-R. Raazi, "Automated prediction of good dictionary EXamples (GDEX): A comprehensive experiment with distant supervision, machine learning, and word embedding-based deep learning techniques," *Complexity*, vol. 2021, no. 1, Jan. 2021, doi: 10.1155/2021/2553199.
[29] R. Ehsani and F. Drabløs, "Robust distance measures for k NN classification of cancer data," *Cancer Informatics*, vol. 19, Jan. 2020, doi: 10.1177/1176935120965542.
[30] H. Saadatfar, S. Khosravi, J. H. Joloudari, A. Mosavi, and S. Shamshirband, "A new K-Nearest neighbors classifier for big data based on efficient data pruning," *Mathematics*, vol. 8, no. 2, p. 286, Feb. 2020, doi: 10.3390/math8020286.
[31] A. Naswin and A. P. Wibowo, "Performance analysis of the decision tree classification algorithm on the pneumonia dataset," *International Journal of Artificial Intelligence in Medical Issues*, vol. 1, no. 1, pp. 1–9, May 2023, doi: 10.56705/ijaimi.v1i1.83.
[32] A. Singh, "Decision tree in machine learning." https://www.appliedaicourse.com/blog/decision-tree-in-machine-learning/ (accessed Jun. 24, 2025).
[33] Y. Ramdhani, R. T. Prasetio, and R. Hidayat, "Comparation of SVM algorithm and neural network with feature optimization based on genetic algorithm in determining immunotherapy success in cancer disease," *Proceedings of the International Conference on Industrial Engineering and Operations Management Monterrey*, pp. 3142–3149, 2021.
[34] S. Susanti, "Searching Sahih Hadiths based on queries using neural models and fasttext," *Journal of Applied Data Sciences*, vol. 6, no. 1, pp. 272–285, Jan. 2024, doi: 10.47738/jads.v6i1.467.
[35] V. Zhou, "Machine learning for beginners: an introduction to neural networks a simple explanation of how they work and how to implement one from scratch in Python." https://victorzhou.com/blog/intro-to-neural-networks/ (accessed Jun. 24, 2025).

BIOGRAPHIES OF AUTHORS



Yudi Ramdhani is a lecturer and researcher at Satu University in the Information Systems Study Program. He holds a Master's degree in Computer Science and has focused his research on Data Mining, Data Science, and Decision Support Systems. With several years of experience in both teaching and research, he has been actively involved in developing data-driven academic projects and innovative curriculum design. He has published research papers in reputable journals and regularly contributes to academic forums and seminars. He can be contacted at email: yudi.ramdhani@univ.satu.ac.id.



Win Ce is an academic and professional that have been active as a lecturer at BINUS University and SATU University, specifically within the Information Systems study program. His experience is not limited to academia; he frequently serves as a speaker and expert for various companies, fulfilling an expert role in the industry. Win Ce, S.Kom., M.M.'s research interests span several key areas in information technology, including software development, machine learning, decision support systems (DSS), and information systems. He can be contacted at email: wn@binus.edu.



Ghanim Kanugrahan received his Bachelor's degree in Computer Science from Universitas Dian Nuswantoro in 2019, and his Master's degree in Information Technology from Universitas Indonesia in 2021. His main areas of expertise include data science, sentiment analysis, and machine learning. He is currently a lecturer in the Informatics Program at Satu University Bandung, where he is actively involved in teaching, research, and community service. He can be contacted at email: ghanim.kanugrahan@univ.satu.ac.id.



Yulita Anggelia is an academic professional currently serving as a Subject Content Coordinator in the Psychology study program at SATU University. Her dedication to education is reflected in her active involvement in curriculum development and academic coordination within the program. With a background rooted in psychology and education, Yulita frequently contributes to academic discussions and institutional initiatives aimed at enhancing the quality of teaching and learning. She can be contacted at email: yulita.anggelia@univ.satu.ac.id.



Febriyanti Panjaitan earned her Doctoral degree from the Faculty of Engineering, Universitas Sriwijaya, with a concentration in Informatics Engineering, in 2023. Her main areas of expertise include data science, machine learning, and deep learning, with a research focus on developing artificial intelligence (AI) models for complex data analysis. She is a lecturer in the Informatics Engineering Program Faculty of Creative Technology at Satu University, and she is actively involved in teaching, research, and community service. She can be contacted at email: pebriyanti.panjaitan@univ.satu.ac.id.



Rizki Tri Prasetio is a lecturer and researcher at SATU University in the Information Systems Study Program. He holds a Master’s degree in Computer Science and focuses his research on Data Mining, Data Science, and Decision Support Systems. With several years of experience in both teaching and research, he has actively contributed to the development of data-driven academic projects and innovative curriculum design. He has published research papers in reputable journals and regularly participates in academic forums and seminars. He can be contacted at email: rizki.prasetio@univ.satu.ac.id.