Jurnal Teknik Informatika (JUTIF) Vol. 7, No. 1, February 2026, Page. 169-184
P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2026.7.1.5259

Predicting Anxiety of STMIK Palangkaraya Students Using K-Means Clustering and Gaussian Naïve Bayes

Maura Widyaningsih*[1], Rosmiati[2], Paholo Iman Prakoso[3]

[1]Program Studi Teknik Informatika , STMIK Palangka Raya, Palangka Raya, Indonesia
[2]Program Studi Sistem Informasi , STMIK Palangka Raya, Palangka Raya, Indonesia
[3]DIKE, Fakultas MIPA, Universitas Gadjah Mada, Yogyakarta, Indonesia

Email: [1]maurawidya@gmail.com

Received : Aug 11, 2025; Revised : Sep 9, 2025; Accepted : Sep 22, 2025; Published : Feb 15, 2026

Abstract

Academic anxiety is a common psychological problem experienced by students, especially before final exams, which impacts learning performance and mental well-being. This study aims to identify and predict students' anxiety levels using a Machine Learning approach, specifically the web framework Gradio, through a combination of the K-Means Clustering and Gaussian Naïve Bayes (GNB) methods. The research instrument used a Google Form-based questionnaire modified from the Zung Self-Rating Anxiety Scale (ZSAS) with 20 items (K1–K20) on a Likert scale (0–3). Data were obtained from 110 students of the Information Systems and Informatics Engineering Study Program at STMIK Palangkaraya. The research process consisted of five main stages: pre-processing, clustering using the K-Means algorithm, training the GNB classification model, evaluation, and prediction of new data. The clustering results categorized the data into three levels of anxiety: Low, Median, and High. The GNB model showed 95% accuracy with a balanced distribution of evaluation metrics (precision, recall, and F1 score). Comparison with other algorithms shows that while SVM achieved the highest accuracy (100%), GNB was more balanced in handling uneven class distributions and more practical for implementation in web-based systems. This prediction system has the potential to be used as an early detection tool for student anxiety, while also supporting educational institutions in designing more targeted psychological interventions. Further improvements can be made by expanding the scope of respondents, balancing the data distribution, and testing other machine learning methods to improve model generalization. The program and data are available at: https://github.com/maurawidya75/StudentAnxiety2025.

Keywords : Gaussian Naïve Bayes, K-Means Clustering, Machine Learning, Student Anxiety, Student Prediction

1. INTRODUCTION

Academic anxiety is a common psychological problem experienced by students, especially when facing final exams. This condition not only affects academic performance but also has a direct impact on students' psychological well-being. Academic pressure, environmental expectations, and limited stress management skills are the main factors that trigger anxiety. If not addressed appropriately, this condition can reduce concentration, decrease academic performance, and increase the risk of mental health disorders [1][2].

Factors causing anxiety can be grouped into internal and external factors. Internal factors include self-confidence, mental health, and time management skills [3][4], while external factors include family pressure, social support, and the difficulty of exam material [5][6][7]. Understanding these factors is crucial for educational institutions to develop appropriate intervention strategies to help students overcome academic anxiety [8].

Several studies have demonstrated the effectiveness of various ML algorithms, such as Naïve Bayes, SVM, KNN, Decision Tree, and Random Forest, in predicting student health status, including physical, mental, and social aspects [9][10][11][12]. For example, Random Forest achieved an accuracy of up to 99.4% in predicting college students' health status [11] , while KNN excelled in predicting the severity of anxiety, stress, and depression with 95% accuracy [12]. Most of these studies still focus on large-scale data and have not integrated many psychometric instruments.

Data-driven approaches and machine learning (ML) techniques are increasingly being applied to support mental health prediction, as they are capable of producing systematic and accurate results [13]. To address this research gap, this study proposes a hybrid approach that combines unsupervised learning (K-Means clustering) and supervised learning (Gaussian Naïve Bayes). K-Means is used to classify survey data into Low, Median, and High categories. Gaussian Naïve Bayes (GNB) itself is known to be efficient for data with normal distributions [14][15] and independent features, and is often used in survey-based prediction [16][17]. Similar studies have shown that the hybrid approach can improve prediction accuracy to 89.6%, higher than single methods [17][18][19].

The survey-based measurement instrument in this study used a modified Zung Self-Rating Anxiety Scale (ZSAS), focusing on psychological aspects with indicators such as symptom frequency, tension level, and academic readiness [20]. The instrument's responses were based on a Likert scale, which can provide a quantitative overview of students' psychological state [17][19].

Data were collected through a Google Form-based questionnaire, developed using a modification of the ZSAS instrument. It consisted of 20 items (K1–K20) focusing on symptoms of academic anxiety leading up to final exams. A total of 110 students from the Information Systems and Informatics Engineering Study Programs at STMIK Palangkaraya participated in completing the questionnaire.

The system was built using the Gradio web framework, with research stages that included pre-processing, clustering with K-Means, model training using the GNB algorithm, performance evaluation using a Confusion Matrix, and prediction of new data.

The primary objectives of this study were to identify and classify student anxiety levels based on survey data and to develop a web-based prediction system for early detection of student anxiety and support psychological interventions in higher education settings. Unlike previous research, this study not only evaluates the performance of GNB but also compares it with five other ML algorithms (SVM, KNN, Logistic Regression, Decision Tree, Random Forest) using accuracy, precision, recall, and F1-score metrics.

The main contribution of this study is the integration of the ZSAS psychometric instrument with ML methods to predict student anxiety levels and the development of a web-based Gradio prediction system that combines K-Means Clustering and GNB. This study also develops a robust classification model that addresses data imbalance, leading to balanced evaluation performance. Furthermore, this research offers practical implications for educational institutions, enabling them to detect student anxiety levels more effectively and design targeted psychological interventions.

2. METHOD

Figure 1 illustrates the research process for a case study on classifying student anxiety before final exams.

This study began with the development of a Google Form-based questionnaire comprising 20 Likert-scale questions (1–4). The questions were designed to measure academic stress, study preparation, social support, and mental health among students at STMIK Palangkaraya. The questionnaire data were coded using Likert scaling, with the following conversion rules: “Tidak sama sekali” (Not at all) = 0, “Sedikit” (Slightly) = 1, “Agak” (Somewhat) = 2, “Sangat” (Very) =3.

The pre-processing stage began with extracting data from Google Sheets. Irrelevant columns such as timestamp, name, and academic ID were removed. Question labels were converted to K1–K20 for better structure. Categorical data were encoded as numeric values (0–3). Missing values were handled using mean imputation normalization, which is suitable for normally distributed data [21].

Vol. 7, No. 1, February 2026, Page. 169-184
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.1.5259

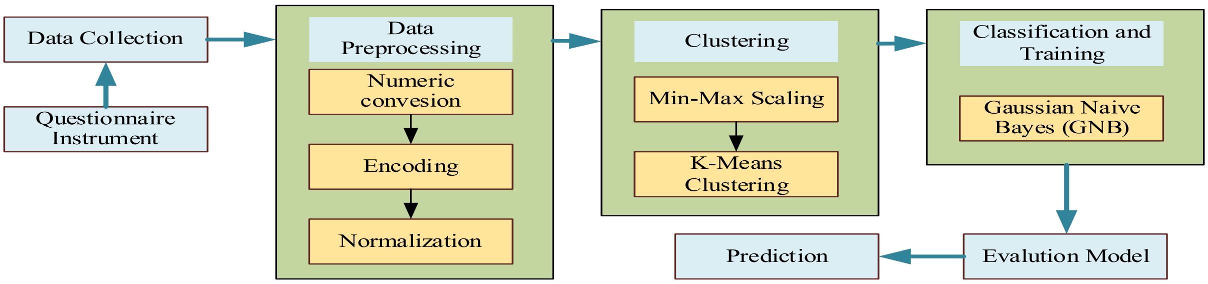


Figure 1. Research Stages

After pre-processing, the next step was clustering using the K-Means Clustering algorithm. This step identified natural patterns in student anxiety. Data were grouped into three clusters: ("Rendah" = Low, "Sedang" = Median, and "Tinggi" = High. Min-Max Scaling standardized features to reduce bias from imbalanced class distributions [22]. The resulting labeled dataset was used in classification.

After clustering, the process proceeded to the training and classification stages, which utilized GNB by dividing the dataset into 80% training data and 20% testing data. In these stages, the model calculated the prior probability of each class, the likelihood of each feature, and the posterior probability to determine the most likely category of student anxiety.

The model was evaluated using a Confusion Matrix. Metrics such as accuracy, precision, recall, and F1-score were calculated. This evaluation assessed model performance and considered data distribution in the results.

The final stage is prediction, where the model is applied to new data containing the same 20 questions. This new data undergoes identical preprocessing steps. The model then predicts anxiety categories for the new data.

2.1. Zung Anxiety Scale

The Zung Self-Rating Anxiety Scale (ZSAS) is a widely used tool to quantify anxiety. Developed by Dr. William W. K. Zung in 1971, it has become a benchmark instrument for assessing anxiety.

The scale comprises statements assessing prevalent anxiety symptoms. These statements address affective (emotional/worry/anxiety), somatic (physical), psycho-cognitive (thoughts and focus), and motor (behavior and activity) manifestations [22]. The 20 items reflect the core components of the Zung scale.

2.2. Likert Scale

The Likert scale is widely used in research to describe attitudes, opinions, or perceptions towards a topic or statement [23]. It consists of questions or statements with answer choices on an ordinal scale, usually ranging from 1 (Not at all) to 4 (Very) [24][25]. Respondents pick the option that best reflects their feelings [24] .

2.3. Imputation

Imputation is a technique in data preprocessing used to fill in missing values in a dataset, allowing for accurate and consistent analysis or training of machine learning models [26]. The purpose of implementing this method is to minimize bias and information loss due to incomplete data, thereby making the results of the analysis or model more representative of the overall data. Simple Imputation applies the Mean Imputation technique by filling in missing values with the average of the related column, and is used for normally distributed data [27].

Vol. 7, No. 1, February 2026, Page. 169-184
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.1.5259

2.4. K-Means Clustering

K-Means Clustering is a widely used unsupervised algorithm that groups data into K clusters based on the similarity of their features. Its main goal is to minimize the distance between each data point and its cluster centre (centroid) [28][29]. K-Means works well for high-dimensional datasets and is suitable for preprocessing in data mining or machine learning tasks, such as classification [30]. Formula 1 shows the main function of K-Means Clustering, and Formula 2 shows Min-Max scale normalization.

$$J = \sum_{i-1} \sum_{x \epsilon C_i} ||x - \mu_i|| \quad (1)$$

Formula 2 defines the main function of K-Means Clustering, where C_i is the i-th cluster, μ_i is the centroid of cluster C_i, and $\| x - \mu_i \|$ is the distance between data x and the centroid.

$$x_{scaled} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (2)$$

Formula 3 illustrates the Min-Max scale normalization, where x represents the original value of a feature, x_{min} denotes the minimum value of the feature, and x_{max} denotes the maximum value of the feature [22] [29].

2.5. Gaussian Naïve Bayes

Naïve Bayes classification is a simple yet effective approach. This algorithm utilizes probabilistic theory to model the relationships between variables and produce robust predictions, even in the presence of imbalanced data. Naïve Bayes has several advantages, such as computational efficiency, the ability to handle independent variables directly, and sufficient accuracy on small datasets [31][32].

Gaussian Naïve Bayes is one of the Naïve Bayes methods, where the primary assumption is that each feature (input variable) in the data follows a normal (Gaussian) distribution within each class, described by its mean and variance [33]. Mean and variance are determined as in formulas (3) to (5):

1. Building a Distribution Model
   By calculating μ and σ^2 from the training data for each feature in each class, GNB constructs a Gaussian probability distribution:

$$P(x_i \mid y = c) = \frac{1}{\sqrt{2\pi\sigma_{c,i}^2}} exp(-\frac{(x_i - \mu_{c,i})^2}{2\sigma_{c,i}^2} \quad (3)$$

   Where x_i = the value of the i-th , $y = c$ = Class c, $\mu_{c,i}$ = the mean of the i-th feature in class c, $\sigma_{c,i}^2$ = the variance of the i-th feature in class c.

2. Measuring Likelihood
   The probability values above are the likelihood that a particular feature value occurs in that class. The mean (μ) centers the distribution around the mean value of the feature for that class. The variance (σ^2) measures how wide the probability distribution is (the greater the variance, the wider the distribution).
3. Combining Features (Naive Bayes Assumption)
   Because Naive Bayes assumes independence between features, the likelihoods of all features are combined (multiplied) to calculate the total probability of a sample belonging to a class:

$$P(x \mid y = c) = i = \prod_{i=1}^{n} P(x_i \mid y = c) \quad (4)$$

4. Class Prediction
   The posterior probability is calculated using Bayes' Theorem.

Vol. 7, No. 1, February 2026, Page. 169-184
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.1.5259

$$P(y = c \mid x) \propto P(y = c) \cdot P(x \mid y = c) \quad (5)$$

The class with the highest probability will be selected as the prediction

2.6. Confusion Matrix

A confusion matrix (also known as an error matrix) is a commonly used evaluation tool in various classification tasks, including semantic segmentation. For semantic segmentation, the confusion matrix can be adapted to consider the individual pixels in an image. A confusion matrix is a matrix-like table that evaluates performance, measuring the extent to which a classification model successfully predicts the correct class across the entire dataset.

There are four classification terms in the confusion matrix: True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). Overall Accuracy (OA) measures the percentage of cases in which the model correctly predicts a value, calculated as (TP + TN) / (TP + FP + TN + FN). Precision measures the number of correct predictions out of all positive predictions, defined as TP / (TP + FP). Recall (also known as Sensitivity) measures the percentage of positive cases successfully identified by the model, defined as TP / (TP + FN). The F1-score, the harmonic mean of precision and recall, is calculated as 2*Precision*Recall / (Precision + Recall). The following are the metric formulas (Formulas 6 and 7) used to evaluate model performance [34][35]:

$$Accuracy = \frac{True\ Positives + True\ Negatives}{Total\ Samples} \quad (6)$$

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives} \quad (7)$$

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives} \quad (8)$$

$$F1 - Score = 2 \,.\, \frac{Precision\ .Recall}{Precision + Recall} \quad (9)$$

3. RESULT

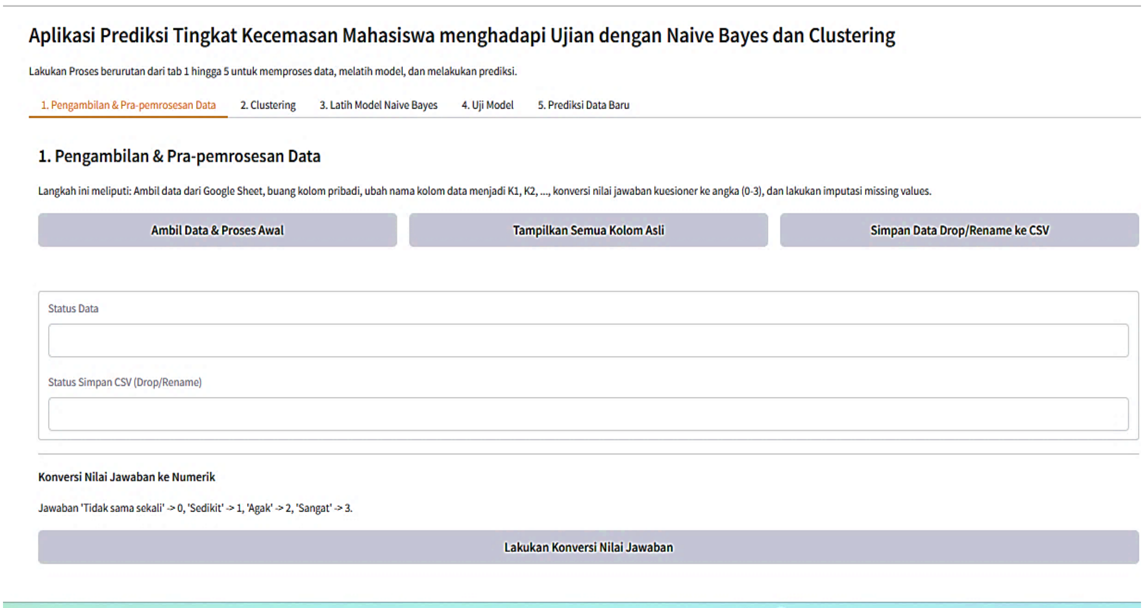


Figure 2. Anxiety Prediction System Interface

Jurnal Teknik Informatika (JUTIF) Vol. 7, No. 1, February 2026, Page. 169-184
P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2026.7.1.5259

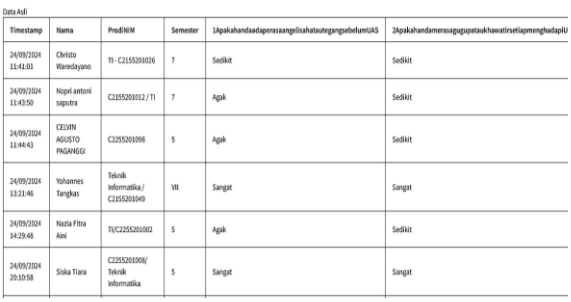
The developed system is web-based using Python in Google Colab with Gradio interface integration. The implementation system consists of five main stages: Pre-processing, Clustering, Model Training, Evaluation, and Prediction of New Data. Figure 2 illustrates the initial display of the system designed to detect student anxiety levels

3.1. Pre-processing

Data were obtained from a Google Form survey, which included a total of 110 respondents, comprising 38 Information Systems students and 72 Informatics Engineering students. The pre-processing stages included:

1. Data collection and initial processing for loading the raw data.
2. Column drop or rename: the process of cleaning columns 1–3 and changing the question labels to K1–K20. This process eliminates unnecessary columns for clustering.
3. Response value conversion: the process of transforming the answers into numeric values ("Tidak sama sekali" = Not at all (0), "Sedikit" = Slightly (1), "Agak" = Somewhat (2), "Sangat" = Very (3)).
4. Data imputation: the process of filling in values using the simple mean imputation normalization technique. This stage ensures the data is ready for further analysis.

Figure 3 shows the results from process 1 to process 2: the original data is then transferred to the data drop results and column names are renamed.



Data Asli

Timestamp	Nama	ProdiNIM	Semester	1ApakahandaadaperasaangelisahatautegangsebelumUAS	2ApakahandamerasagugupataukhawatirsetiapmenghadapiU
24/09/2024 11:41:01	Christo Waredayano	TI - C2155201026	7	Sedikit	Sedikit
24/09/2024 11:43:50	Nopri antoni saputra	C2155201012 / TI	7	Agak	Sedikit
24/09/2024 11:44:43	CELVIN AGUSTO PAGANGGI	C2255201098	5	Agak	Sedikit
24/09/2024 13:21:46	Yohannes Tangkas	Teknik Informatika / C2155201049	VII	Sangat	Sangat
24/09/2024 14:29:48	Nazia Fitra Aini	TI/C2255201002	5	Agak	Sedikit
24/09/2024 20:10:58	Siska Tiara	C2255201008/ Teknik Informatika	5	Sangat	Sangat

(a) Original data

Data setelah Drop Kolom & Rename

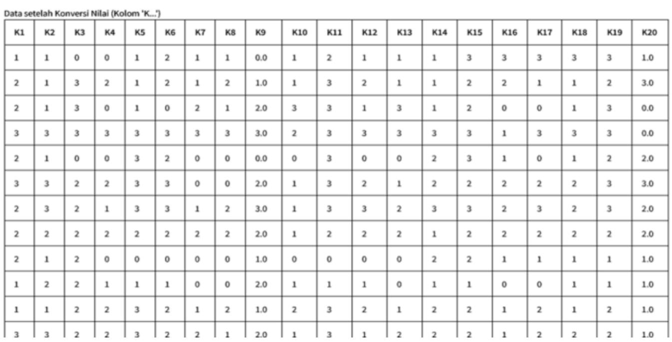
K1	K2	K3	K4	K5	K6	K7	K8	K9	K10	K11	K12	K13
Sedikit	Sedikit	Tidak sama sekali	Tidak sama sekali	Sedikit	Agak	Sedikit	Sedikit	Tidak sama sekali	Sedikit	Agak	Sedikit	Sedikit
Agak	Sedikit	Sangat	Agak	Sedikit	Agak	Sedikit	Agak	Sedikit	Sedikit	Sangat	Agak	Sedikit
Agak	Sedikit	Sangat	Tidak sama sekali	Sedikit	Tidak sama sekali	Agak	Sedikit	Agak	Sangat	Sangat	Sedikit	Sangat
Sangat	Sangat	Sangat	Sangat	Sangat	Sangat	Sangat	Sangat	Sangat	Agak	Sangat	Sangat	Sangat
Agak	Sedikit	Tidak sama sekali	Tidak sama sekali	Sangat	Agak	Tidak sama sekali	Tidak sama sekali	Tidak sama sekali	Tidak sama sekali	Sangat	Tidak sama sekali	Tidak sama sekali

(b) Data drop and column rename

Figure 3. Results of retrieving original data, dropping data, and renaming columns

The results shown in Figure 2 show that unnecessary columns were eliminated. The clustering process only requires data with the required columns. Therefore, columns 1, 2, and 3 were excluded. The original data, after eliminating columns, had their column headings changed to simplify the process.

Figure 4 illustrates the subsequent results of process 3, which converts text response codes to numeric values.

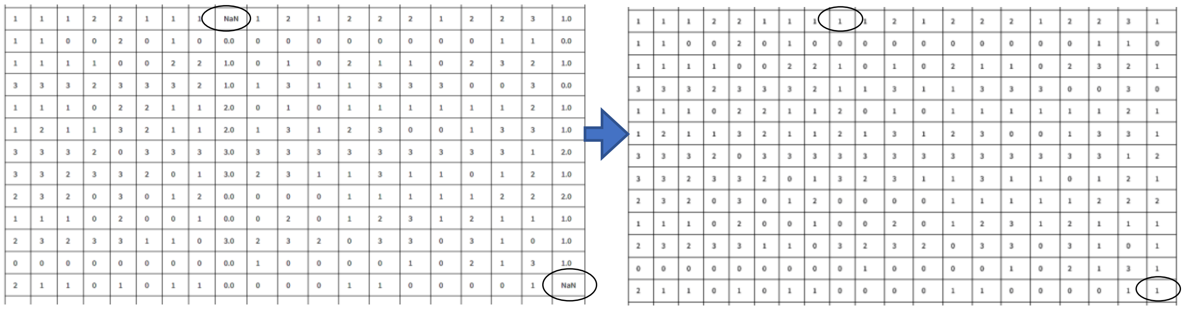


Data setelah Konversi Nilai (Kolom 'K...')

K1	K2	K3	K4	K5	K6	K7	K8	K9	K10	K11	K12	K13	K14	K15	K16	K17	K18	K19	K20
1	1	0	0	1	2	1	1	0.0	1	2	1	1	1	3	3	3	3	3	1.0
2	1	3	2	1	2	1	2	1.0	1	3	2	1	1	2	2	1	1	2	3.0
2	1	3	0	1	0	2	1	2.0	3	3	1	3	1	2	0	0	1	3	0.0
3	3	3	3	3	3	3	3	3.0	2	3	3	3	3	3	1	3	3	3	0.0
2	1	0	0	3	2	0	0	0.0	0	3	0	0	2	3	1	0	1	2	2.0
3	3	2	2	3	3	0	0	2.0	1	3	2	1	2	2	2	2	2	3	3.0
2	3	2	1	3	3	1	2	3.0	1	3	3	2	3	3	2	3	2	3	2.0
2	2	2	2	2	2	2	2	2.0	1	2	2	2	1	2	2	2	2	2	2.0
2	1	2	0	0	0	0	0	1.0	0	0	0	0	2	2	1	1	1	1	1.0
1	2	2	1	1	1	0	0	2.0	1	1	1	0	1	1	0	0	1	1	1.0
1	1	2	2	3	2	1	2	1.0	2	3	2	1	2	2	1	2	1	2	1.0
3	3	2	2	3	2	2	1	2.0	1	3	1	2	2	2	1	2	2	2	1.0

(a) Convert text to numeric

P-ISSN: 2723-3863
E-ISSN: 2723-3871
Vol. 7, No. 1, February 2026, Page. 169-184
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.1.5259



(b) Imputation Process

Figure 4. Results of the data conversion and imputation process

Figure 4 shows the conversion of all text data to numeric values as described in step 3, using a predetermined scale. Next, the imputation process fills in empty values (NaN) to prevent bias in subsequent steps. Specifically, a "NaN" was found in the K9 feature for the 85-th data point and the K20 feature for the 97-th. After imputation, these empty values are filled with the value 1, ensuring the data is ready for the clustering stage.

3.2. Clustering Process

K-Means Clustering begins with Min-Max scaling normalization, ensuring that each feature is in the range [0, 1], thereby preventing any single feature from dominating. The clustering results produce three classifications: Low, Median, and High for all 110 data sets. Figure 5 displays the average cluster score for each feature (K1–K20) for the three clusters resulting from the Low, Median, and High clustering results.

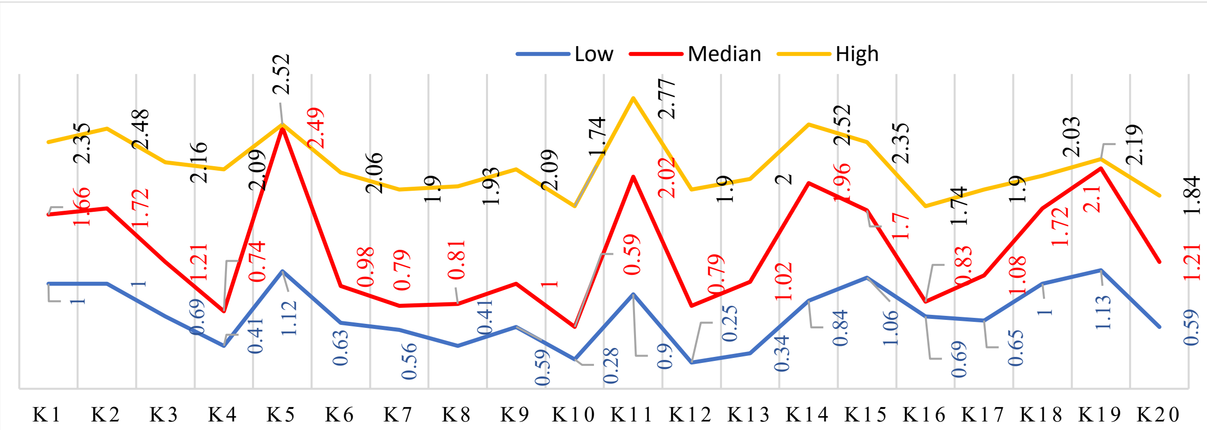


Figure 5. Average features per cluster

Figure 5 shows that the High cluster consistently has an average score above 2, the Median cluster is in the 1–2 range, while the Low cluster tends to be <1 for most features. The most striking difference is seen in K5 and K11, where the Low cluster has a very low score (<0.5), while the High cluster is above 2.5, making it the strongest indicator in distinguishing anxiety categories. Furthermore, features K10 and K14 also show a consistent increasing pattern from low to high, supporting the classification. Thus, the clustering results confirm that K5, K11, K10, and K14 are the dominant features distinguishing student anxiety levels.

The result of the data grouping process using the K-Means Clustering method is that all data is labeled ("Rendah" (Low)= 0, "Sedang" (Median)= 1, "Tinggi" (High) =2), allowing the data to proceed to the training data process, some sample results are shown in Figure 6.

Vol. 7, No. 1, February 2026, Page. 169-184
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.1.5259

K1	K2	K3	K4	K5	K6	K7	K8	K9	K10	K11	K12	K13	K14	K15	K16	K17	K18	K19	K20	Cluster	Cluster_Label
1	1	0	0	1	2	1	1	0	1	2	1	1	1	3	3	3	3	3	1	1	Sedang
2	1	3	2	1	2	1	2	1	1	3	2	1	1	2	2	1	1	2	3	2	Tinggi
2	1	3	0	1	0	2	1	2	3	3	1	3	1	2	0	0	1	3	0	1	Sedang
3	3	3	3	3	3	3	3	3	2	3	3	3	3	3	1	3	3	3	0	2	Tinggi
2	1	0	0	3	2	0	0	0	0	3	0	0	2	3	1	0	1	2	2	1	Sedang
3	3	2	2	3	3	0	0	2	1	3	2	1	2	2	2	2	2	3	3	2	Tinggi
2	3	2	1	3	3	1	2	3	1	3	3	2	3	3	2	3	2	3	2	2	Tinggi
2	2	2	2	2	2	2	2	2	1	2	2	2	1	2	2	2	2	2	2	2	Tinggi
2	1	2	0	0	0	0	0	1	0	0	0	0	2	2	1	1	1	1	1	0	Rendah
1	2	2	1	1	1	0	0	2	1	1	1	0	1	1	0	0	1	1	1	0	Rendah
1	1	2	2	3	2	1	2	1	2	3	2	1	2	2	1	2	1	2	1	2	Tinggi
3	3	2	2	3	2	2	1	2	1	3	1	2	2	2	1	2	2	2	1	2	Tinggi
2	3	1	0	3	0	0	1	2	0	3	3	1	3	0	0	0	0	3	2	1	Sedang
3	3	2	3	3	1	1	2	3	3	3	2	2	3	0	2	2	2	2	2	2	Tinggi
3	3	3	2	2	1	2	1	2	1	3	1	2	2	3	1	1	2	1	1	2	Tinggi

Figure 6. K-Means Clustering classification results

From the results of the Clustering process on a total of 110 data points, the classification results obtained were: Low = 32 data points; Median = 47 data points; and High = 31 data points. This suggests that normalization using the Min-Max scale can help mitigate bias in classification results, thereby reducing the dominance of the most prevalent class. As seen, the previous class was dominated by the Median class, followed by the Low class, while the High class was less, with the Min-Max technique being able to distribute the classes more evenly.

3.3. Training with Gaussian Naïve Bayes

The training process was performed using Gaussian Naive Bayes (GNB) to learn the relationship between the K1–K20 features and the clustering labels. The dataset was divided into 88 (80%) train and 22 (20%) test datasets.

Table 1 shows the results of the data learning process, specifically for obtaining prior probabilities of training dataset.

Table 1. Prior Probability

Class Label	Encoded Class	Number of Training Samples	Prior Probability
Low	0	26	0.295455
Median	1	37	0.420455
High	2	25	0.284091

Prior Probability indicates the likelihood of each class occurring before considering the feature factors (K1–K20). The Prior Probability values show that the Median class (0.42) has the highest baseline probability, followed by Low (0.29) and High (0.28). Although the difference is not statistically significant, it suggests that the Median class is slightly more dominant in the training dataset.

Figure 7, Figure 8, and Figure 9 show the likelihood (mean and variance) obtained for Low, Median, and High classes, where the mean value indicates the average feature value for a particular class, and the variance indicates how spread out the feature data is in a specific class.

Analysis of the mean and variance distribution of features K1–K20 shows that the mean values increase from the Low to the Median and High classes, reflecting a linear classification pattern. The most discriminatory features between classes are K1, K2, K3, K11, and K12. The High class has a high mean and low variance, allowing for sharp but potentially biased predictions. The Median class exhibits greater variance and overlaps with the High class, which can cause confusion for the model.

P-ISSN: 2723-3863
E-ISSN: 2723-3871
Vol. 7, No. 1, February 2026, Page. 169-184
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.1.5259

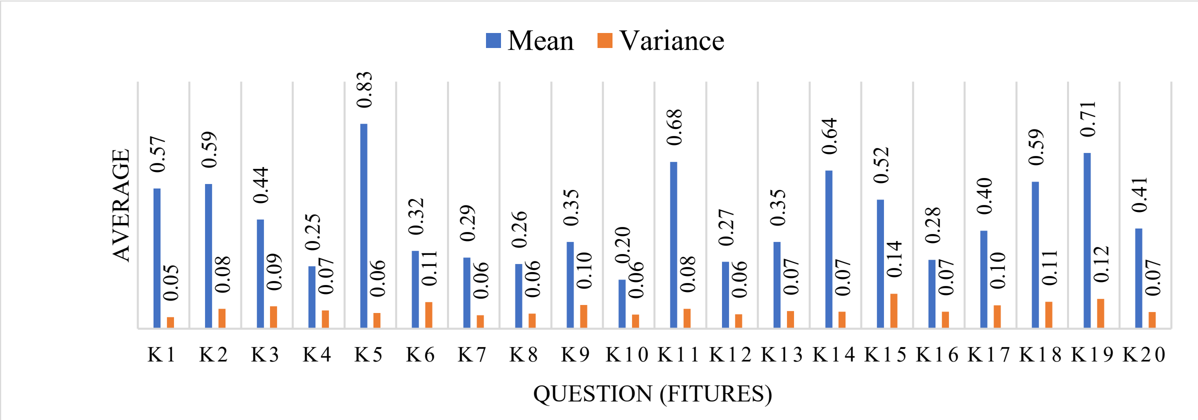


Figure 7. Likelihood Gain (Mean and Variance) in Low-Class

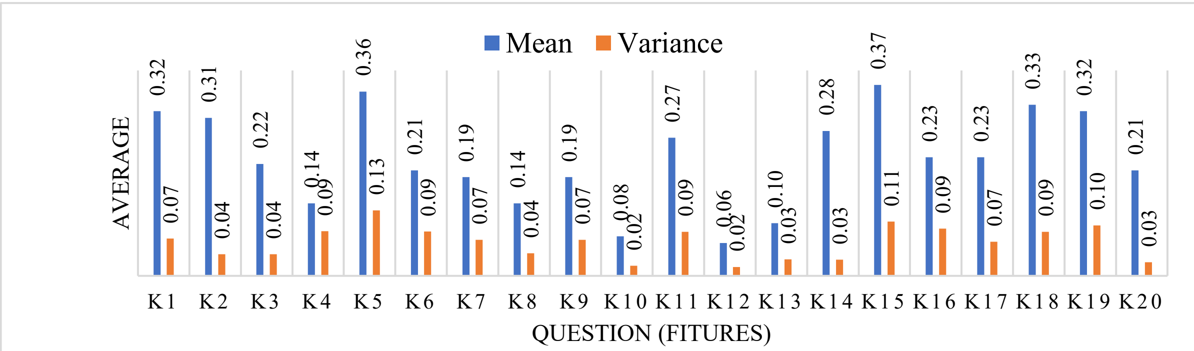


Figure 8. Likelihood Gain (Mean and Variance) in the Median-Class

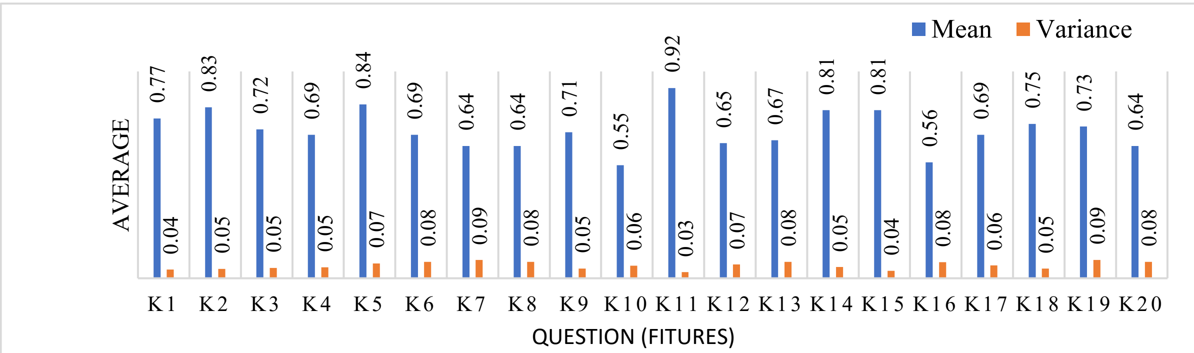


Figure 9. Likelihood Gain (Mean and Variance) in High Class

3.4. Performance Evaluation

Model evaluation used a Confusion Matrix on the test data. Table 2 summarizes the Gaussian Naive Bayes model's performance on predictions.

Table 2. Confusion Matrix evaluation results

	Precision	Recall	F1-score	Support
Low	1.00	0.83	0.91	6
Median	0.91	1.00	0.95	10
High	1.00	1.00	1.00	6
accuracy			0.95	22
macro avg	0.97	0.94	0.95	22
weighted avg	0.96	0.95	0.95	22

Jurnal Teknik Informatika (JUTIF)
P-ISSN: 2723-3863
E-ISSN: 2723-3871

Vol. 7, No. 1, February 2026, Page. 169-184
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.1.5259

Table 2 shows that the model achieved 95% accuracy, indicating strong classification performance. The High class achieved perfect precision, recall, and F1-score (1.00), indicating the model was able to recognize all data in this class without error. The Median class had a recall of 1.00 and a precision of 0.91, indicating all Median class data was correctly classified, although there were small misclassifications from other classes into this class. Meanwhile, the Low class had perfect precision (1.00) but a recall of 0.83, indicating that the model did not recognize some Low class data.

The Macro Average is obtained by averaging the precision, recall, and F1-score values for each class, regardless of the data size. This is suitable for measuring the performance of imbalanced datasets. The results aim to measure performance equally for all classes, including both frequent and rare ones.

The Weighted Average is the average of each class, considering the proportion of data in each class. It is obtained by multiplying each metric (precision, recall, F1) by the number of data points per class, then calculating the weighted average. The results aim to provide a more realistic assessment of model performance in imbalanced datasets.

3.5. New Data Prediction

Figure 10 illustrates the new data prediction interface, which enables users to input answers to 20 questions. The test example demonstrates that the system can classify new responses into the High category with a posterior probability of the High class equal to 1. This shows the system's practical use in early identification of students' anxiety levels.

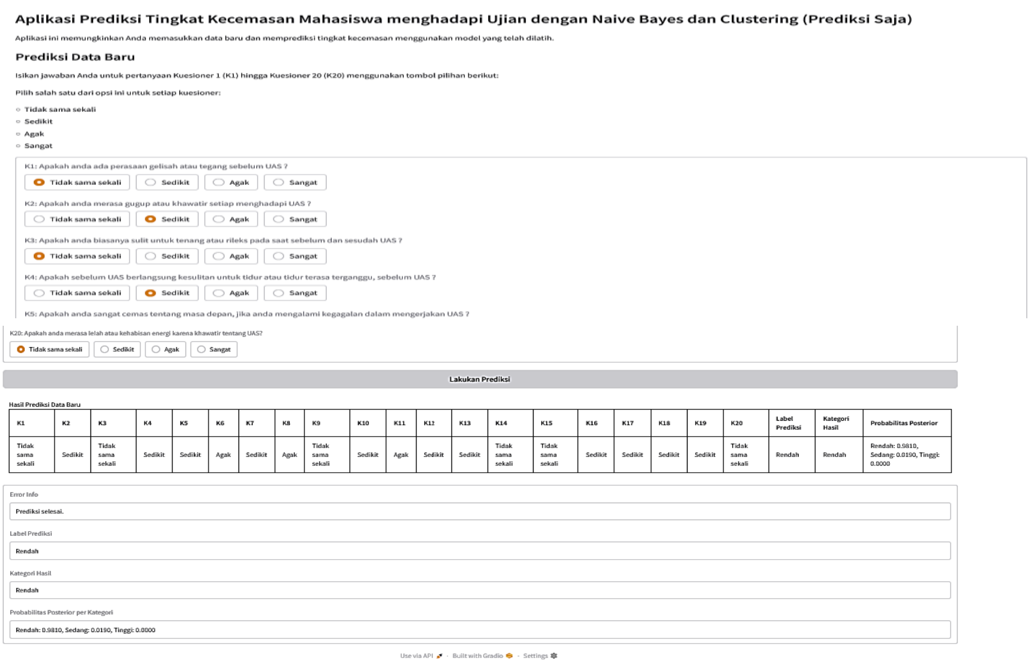


Figure 10. Process view menu "Predict new data"

Figure 9 shows that the system successfully classifies new data into its categories and provides measurements that instill confidence in its predictions, making it practical for identifying early levels of student anxiety.

3.6. Comparison of Machine Learning Models

Validate performance by comparing it with five other ML models: Logistic Regression, Support Vector Machine, K-Nearest Neighbor, Decision Tree, and Random Forest.

Following this validation, Table 3 compares ML training approaches that could aid early detection and intervention in student anxiety issues. The dataset comes from the clustering process.

Vol. 7, No. 1, February 2026, Page. 169-184
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.1.5259

Table 3. Confusion Matrix evaluation results

Model	Accuracy %	Low (0)			Median (1)			High (2)		
		Precision	Recall	F1-score	Precision	Recall	F1-score	Precision	Recall	F1-score
Logistic Regression	0.86	1	0.67	0.8	0.82	0.9	0.86	0.86	1	0.92
SVM	1	1	1	1	1	1	1	1	1	1
KNN	0.91	0.83	0.83	0.83	0.9	0.9	0.9	1	1	1
Decision Tree	0.86	1	0.67	0.8	0.82	0.9	0.86	0.86	1	0.92
Random Forest	0.91	1	0.67	0.8	0.83	1	0.91	1	1	1
GNB	0.95	1	0.83	0.91	0.91	1	0.95	1	1	1

SVM achieved perfect accuracy (100%), with precision, recall, and F1-score of 1.00 for all classes. Naïve Bayes achieved 95% accuracy, noted for its simplicity and probabilistic interpretation. KNN and Random Forest each reached 91% accuracy. Logistic Regression and Decision Tree both showed 86% accuracy, which was lower compared to the other models.

Building on these results, Figure 11 and Figure 12 display performance evaluation graphs that compare the model outcomes on the testing dataset, as presented in Table 3, with the Confusion Matrix measurement. The measurement results are described in terms of Accuracy, Precision, Recall, and F1-score.

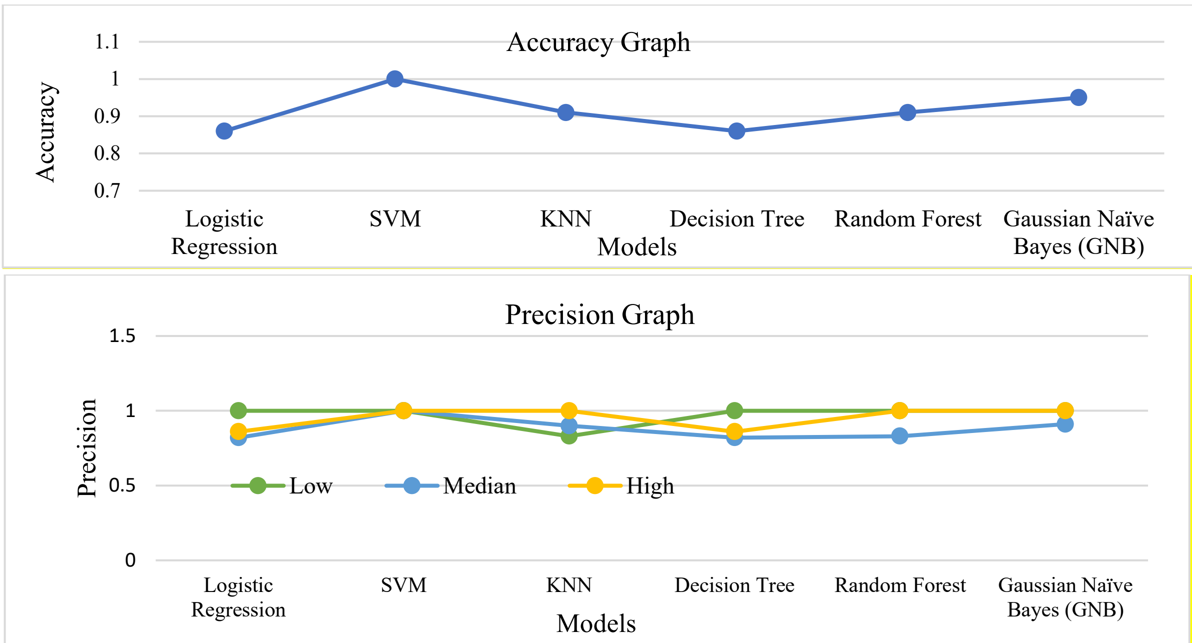
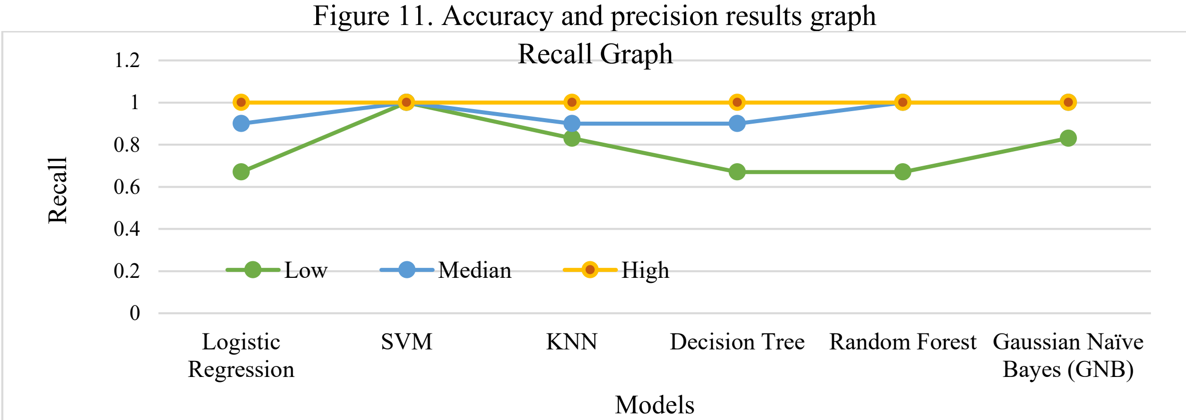


Figure 11. Accuracy and precision results graph



Vol. 7, No. 1, February 2026, Page. 169-184
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.1.5259

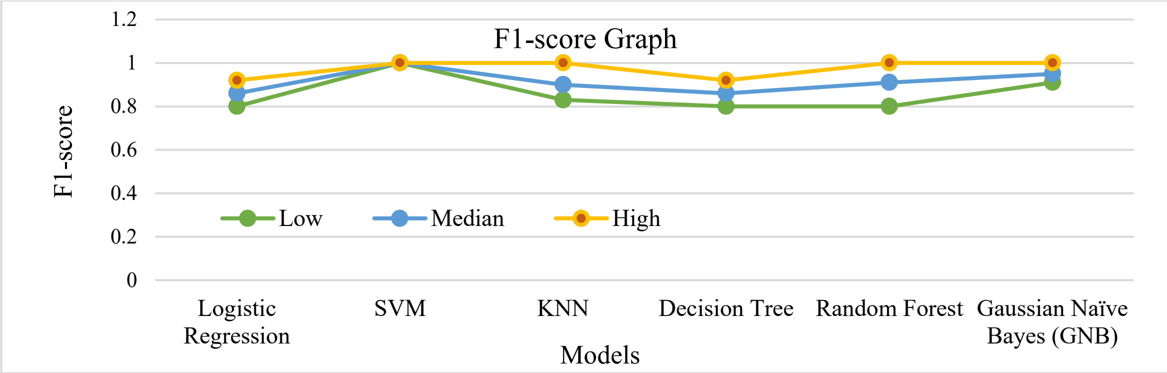


Figure 12. Graph of comparison results of recall and F1-score

Figure 11 shows that SVM achieved the highest accuracy (100%), followed by GNB (95%) and Random Forest (91%). Notably, GNB remained competitive with high accuracy despite imbalanced data. Regarding precision, the High class was perfectly consistent (1.00) across almost all algorithms. The Median class maintained high and stable precision, while the Low class fluctuated more, especially in Logistic Regression and Decision Tree (0.67). Transitioning to recall, Figure 12 shows that the High class's recall was perfectly stable (1.00), meaning all data were detected. Conversely, the Low class was the most difficult to recognize, with varying recall values (0.67–0.83), except for SVM, which achieved a perfect recall of 1.00. The High-class F1-score was near perfect across all models. Importantly, GNB provided the best balance across all classes (0.91, 0.95, 1.00), demonstrating stability in both precision and sensitivity despite imbalanced data.

This comparison demonstrates that while SVM excels in accuracy, GNB remains a suitable choice due to its stability, efficiency on limited datasets, and the ability to provide probabilistic interpretations. This is particularly relevant in the context of early detection of student anxiety, as it directly links GNB's strengths to practical application.

To further illustrate these findings, Table 4 presents the classification results of different machine learning algorithms. The class labels are coded as follows: L = Low, M = Median, H = High. Each row represents the actual class, while each column represents the predicted class for each algorithm.

Table 4. Confusion Matrix evaluation results

Class	Logistic Regression			SVM			KNN			Decision Tree			Random Forest			GNB			Support
	L	M	H	L	M	H	L	M	H	L	M	H	L	M	H	L	M	H	
L	4	2	0	6	0	0	5	1	0	4	2	0	4	2	0	5	1	0	6
M	0	9	1	0	10	0	1	9	0	0	9	1	0	10	0	0	10	0	10
H	0	0	6	0	0	6	0	0	6	0	0	6	0	0	6	0	0	6	6

Table 4 shows that, in general, the three classes Low, Median, and High can be predicted with varying degrees of accuracy by each algorithm. All algorithms performed very well for the High class, as indicated by a correct prediction score of 6 with no misclassifications, showing that the High class had the most easily separated patterns by the model. For the Median class, the best performance was demonstrated by SVM, Random Forest, and GNB, which successfully predicted 10 samples correctly without error. Other algorithms, such as Logistic Regression, KNN, and Decision Tree, still experienced slight misclassifications, mapping some samples to the Low or High class.

In contrast, the Low class appeared to be the most challenging to predict consistently. Logistic Regression and Decision Tree tended to misclassify some samples to the Median (2 cases). Meanwhile, KNN and GNB were more stable, with only one error, and SVM excelled, achieving perfect predictions for this class.

Jurnal Teknik Informatika (JUTIF)
P-ISSN: 2723-3863
E-ISSN: 2723-3871

Vol. 7, No. 1, February 2026, Page. 169-184
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.1.5259

4. DISCUSSIONS

The results of this study demonstrate a close relationship between the clustering and classification stages in detecting student anxiety levels. The clustering process, using the K-Means method, successfully grouped the data into three distinct categories: Low, Median, and High. The feature distribution pattern indicates that K5, K10, K11, and K14 were the dominant variables distinguishing the three groups. The low cluster consistently had scores below 1, the Median cluster ranged between 1 and 2, and the High cluster exceeded 2.5 for several features. These results confirm that the unsupervised learning approach is capable of generating valid labels to support further classification processes. Thus, clustering plays a crucial role in the initial stage of building a labeled dataset, which underpins the predictive model's performance.

Further, building on these findings, the classification results demonstrate that the GNB (National Learning) model effectively utilizes the clustering results. This model achieved an overall accuracy of 95% with balanced performance in precision, recall, and F1-score across all classes. Although the low class is relatively more challenging to recognize than the Median and High classes, GNB maintained stable predictions. Furthermore, comparison results with five other machine learning models showed that SVM achieved the highest accuracy (100%), while GNB performed more evenly across all evaluation metrics, including those with imbalanced class distributions. This indicates that while SVM excels in accuracy, GNB is more robust to data variability [36] and more efficient in practical implementations in web-based detection systems.

Furthermore, inter-class predictive comparison analysis revealed that the SVM model provided the most consistent results in classifying all three anxiety categories [9]. Random Forest and GNB also demonstrated competitive performance, particularly in recognizing the Median and high anxiety classes with minimal error rates. In contrast, Logistic Regression, KNN, and Decision Tree had relatively higher error rates, particularly for the more challenging to separate low and Median anxiety classes. This pattern confirms that algorithms with probabilistic (GNB) and margin-based (SVM) approaches are superior in handling data with imbalanced distributions compared to other algorithms [12].

K-Means Clustering and classification using GNB are appropriate approaches for early detection of student anxiety [18][36]. SVM can indeed be considered an alternative with perfect accuracy [37], but its limitations in terms of interpretability and higher computational requirements make it less practical to integrate into web-based systems. GNB itself provides more stable results, is simple, and can maintain the balance of evaluation metrics, so it is seen as a suitable model in the context of real implementation in educational environments.

5. CONCLUSION

This research successfully developed a web-based Gradio student anxiety prediction system using five main stages: preprocessing, clustering, model training, evaluation, and prediction. The results showed that the preprocessing stage successfully converted 20 categorical features (K1–K20) into numerical form, allowing for more accurate normalization and statistical analysis, without the discovery of significant missing values.

Building upon this, K-Means Clustering successfully identified natural grouping patterns (Low, Median, High). However, the data distribution between clusters was unbalanced, necessitating normalization scaling to reduce bias. Min-Max scaling provided the clustering solution to minimize bias in the classification output.

Following clustering, predictions using Gaussian Naive Bayes (GNB) achieved a high accuracy of 95%, with stable precision (the proportion of true positive predictions among all positive predictions), recall (the proportion of true positive predictions among all actual positives), and F1-score (the harmonic mean of precision and recall) values, especially for the perfectly identified High-class.

Furthermore, comparisons with other algorithms showed that Support Vector Machine (SVM) achieved the highest accuracy (100%), but Gaussian Naive Bayes (GNB) was more consistent in handling imbalanced class distributions. This confirms that the GNB algorithm is an appropriate and practical decision-making model, implemented in a web-based system, to support the early detection of student anxiety. As a result, this system can contribute to preventative efforts in psychological services in higher education settings by efficiently and measurably mapping student conditions early.

Given that the imbalance was a problem in this study, addressing it in future work is crucial. Therefore, if questionnaires are included, complete responses to each question are required to eliminate any missing data. Further feature analysis is also essential to simplify the model and improve interpretation. Once the model is optimized, it can be integrated into a simple application as a tool to classify anxiety levels.

CONFLICT OF INTEREST

The authors declare there is no conflict of interest in this work.

ACKNOWLEDGEMENT

The authors would like to thank STMIK Palangkaraya for supporting publication funds and using laboratories for experiments and testing.

REFERENCES

[1] M. Bieleke, T. Goetz, T. Yanagida, E. Botes, A. C. Frenzel, and R. Pekrun, "Measuring emotions in mathematics: the Achievement Emotions Questionnaire—Mathematics (AEQ-M)," *ZDM - Math. Educ.*, vol. 55, no. 2, 2023, doi: 10.1007/s11858-022-01425-8.

[2] R. P. George, P. M. Donald, H. H. K. Soe, S. C. Tee, J. Toh, and M. J. Q. Cheah, "Prevalence of Symptoms of Depression, Anxiety, and Stress among Undergraduate Dental Students in Malaysia.," *J. Contemp. Dent. Pract.*, vol. 23, no. 5, pp. 532–538, May 2022, [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/35986462

[3] A. S. Nugroho, R. T. Sari, H. Cahyono, P. Siswanto, and O. Jambari, "Investigating Causing Factors of Speaking Anxiety," *EDUKASIA J. Pendidik. dan Pembelajaran*, vol. 4, no. 2, pp. 1289–1294, 2023, [Online]. Available: http://jurnaledukasia.org

[4] Y. Zhenlei, M. Boyuan, S. Lin, G. Chunxia, and H. Qiang, "Identification of knowledge anxiety factors among researchers based on grounded theory," *Heliyon*, vol. 10, no. 4, 2024, doi: 10.1016/j.heliyon.2024.e25752.

[5] M. Zineldin, "Neurological and psychological determinants of depression, anxiety, and life quality," *Int. J. Prev. Med.*, vol. 12, no. 1, 2021, doi: 10.4103/ijpvm.ijpvm_237_19.

[6] E. Zhou *et al.*, "Psychosocial factors associated with anxious depression," *J. Affect. Disord.*, vol. 322, pp. 39–45, 2023, doi: 10.1016/j.jad.2022.11.028.

[7] K. M. Keyes and J. M. Platt, "Annual Research Review: Sex, gender, and internalizing conditions among adolescents in the 21st century – trends, causes, consequences," *Journal of Child Psychology and Psychiatry and Allied Disciplines*, vol. 65, no. 4. 2024. doi: 10.1111/jcpp.13864.

[8] Q. Chen, "Causes and Treatment of Anxiety Disorder," *Lect. Notes Educ. Psychol. Public Media*, vol. 9, no. 1, 2023, doi: 10.54254/2753-7048/9/20230230.

[9] L. Luo *et al.*, "Predictors of depression among Chinese college students: a machine learning approach," *BMC Public Health*, vol. 25, no. 1, p. 470, Feb. 2025, doi: 10.1186/s12889-025-21632-8.

[10] I. J. Ratul, M. M. Nishat, F. Faisal, S. Sultana, A. Ahmed, and M. A. Al Mamun, "Analyzing Perceived Psychological and Social Stress of University Students: A Machine Learning Approach," *Heliyon*, vol. 9, no. 6, 2023, doi: 10.1016/j.heliyon.2023.e17307.

[11] G. Tyulepberdinova, M. Mansurova, T. Sarsembayeva, S. Issabayeva, and D. Issabayeva, "The

physical, social, and mental conditions of machine learning in student health evaluation," *J. Comput. Assist. Learn.*, vol. 40, no. 5, pp. 2020–2030, Oct. 2024, doi: 10.1111/jcal.12999.

[12] S. S. Malik and A. Khan, "Anxiety, Depression and Stress prediction among College Students using Machine Learning Algorithms," in *2023 2nd International Conference on Electrical, Electronics, Information and Communication Technologies, ICEEICT 2023*, 2023, pp. 1–5. doi: 10.1109/ICEEICT56924.2023.10157693.

[13] S. Tribedi, A. Biswas, S. K. Ghosh, and A. Ghosh, "Machine Learning Based Anxiety Prediction of General Public from Tweets During COVID-19," in *Studies in Computational Intelligence*, vol. 963, 2022, pp. 291–312. doi: 10.1007/978-3-030-74761-9_13.

[14] F. T. Cruz, E. E. C. Flores, and S. J. C. Quispe, "Prediction of depression status in college students using a Naive Bayes classifier based machine learning model," *Psychol. Comput. Sci.*, Jul. 2023, [Online]. Available: http://arxiv.org/abs/2307.14371

[15] R. Sahoo, B. K. Mishra, and B. R. Das, "Odia Text Classification Using Naïve Bayes Algorithm : An Empirical Study," *ECS Trans.*, vol. 107, no. 1, 2022, doi: 10.1149/10701.8175ecst.

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

[17] I. Kaur, Kamini, J. Kaur, Gagandeep, S. P. Singh, and U. Gupta, "Enhancing explainability in predicting mental health disorders using human–machine interaction," *Multimed. Tools Appl.*, 2024, doi: 10.1007/s11042-024-18346-1.

[18] D. Goutam, V. Rani, and H. S. i Sain, "Mental Health illness Prediction with Hybrid Machine Learning Approach," *Int. Res. J. Mod. Eng. Technol. Sci.*, vol. 06, no. 01, pp. 3893–3898, 2024, doi: 10.56726/irjmets49045.

[19] K. Vaishnavi, U. N. Kamath, B. A. Rao, and N. V. S. Reddy, "Predicting Mental Health Illness using Machine Learning Algorithms," in *Journal of Physics: Conference Series*, 2022. doi: 10.1088/1742-6596/2161/1/012021.

[20] D. A. Dunstan and N. Scott, "Norms for Zung's Self-rating Anxiety Scale," *BMC Psychiatry*, vol. 20, no. 1, 2020, doi: 10.1186/s12888-019-2427-6.

[21] J. H. Li *et al.*, "Comparison of the effects of imputation methods for missing data in predictive modelling of cohort study datasets," *BMC Med. Res. Methodol.*, vol. 24, no. 1, 2024, doi: 10.1186/s12874-024-02173-x.

[22] C. Wongoutong, "The impact of neglecting feature scaling in k-means clustering," *PLoS One*, vol. 19, no. 12, p. e0310839, Dec. 2024, doi: 10.1371/journal.pone.0310839.

[23] M. Koo and S.-W. Yang, "Likert-Type Scale," *Encyclopedia*, vol. 5, no. 1, p. 18, Feb. 2025, doi: 10.3390/encyclopedia5010018.

[24] J. Mumu, B. Tanujaya, R. Charitas, and I. Prahmana, "Likert Scale in Social Sciences Research: Problems and Difficulties," *FWU J. Soc. Sci.*, vol. 16, no. 4, 2022, doi: 10.51709/19951272/Winter2022/7.

[25] A. F. Kiliç, I. Uysal, and B. Kalkan, "An alternative to likert scale: Emoji," *Journal of Measurement and Evaluation in Education and Psychology*, vol. 12, no. 2. 2021. doi: 10.21031/epod.864336.

[26] M. Liu *et al.*, "Handling missing values in healthcare data: A systematic review of deep learning-based imputation techniques," *Artificial Intelligence in Medicine*, vol. 142. pp. 178–210, 2023. doi: 10.1016/j.artmed.2023.102587.

[27] K. Kotan and S. Kırışoğlu, "Cyclical hybrid imputation technique for missing values in data sets," *Sci. Rep.*, vol. 15, no. 1, p. 6543, Feb. 2025, doi: 10.1038/s41598-025-90964-7.

[28] A. M. Ikotun, A. E. Ezugwu, L. Abualigah, B. Abuhaija, and J. Heming, "K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data," *Inf. Sci. (Ny).*, vol. 622, pp. 178–210, 2023, doi: 10.1016/j.ins.2022.11.139.

[29] R. Mussabayev, N. Mladenovic, B. Jarboui, and R. Mussabayev, "How to Use K-means for Big Data Clustering?," *Pattern Recognit.*, vol. 137, p. 109269., 2023, doi: 10.1016/j.patcog.2022.109269.

[30] X. Hu, X. Chen, W. Liu, and G. Dai, "Road Traffic Status Prediction Approach Based on Kmeans-Decision Tree Model," *J. Eng. Proj. Prod. Manag.*, vol. 12, no. 2, 2022, doi:

Vol. 7, No. 1, February 2026, Page. 169-184
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.1.5259

10.32738/JEPPM-2022-0010.
[31] Nurul Rismayanti and Aulia Putri Utami, "Improving Multi-Class Classification on 5-Celebrity-Faces Dataset using Ensemble Classification Methods," *Indones. J. Data Sci.*, vol. 4, no. 2, 2023, doi: 10.56705/ijodas.v4i2.78.
[32] K. Maswadi, N. A. Ghani, S. Hamid, and M. B. Rasheed, "Human activity classification using Decision Tree and Naïve Bayes classifiers," *Multimed. Tools Appl.*, vol. 80, no. 14, 2021, doi: 10.1007/s11042-020-10447-x.
[33] M. V. Anand, B. Kiranbala, S. R. Srividhya, K. C., M. Younus, and M. H. Rahman, "Gaussian Naïve Bayes Algorithm: A Reliable Technique Involved in the Assortment of the Segregation in Cancer," *Mob. Inf. Syst.*, vol. 2022, no. 436946, pp. 1–7, 2022, doi: 10.1155/2022/2436946.
[34] S. M. Piryonesi and T. E. El-Diraby, "Data Analytics in Asset Management: Cost-Effective Prediction of the Pavement Condition Index," *J. Infrastruct. Syst.*, vol. 26, no. 1, 2020, doi: 10.1061/(asce)is.1943-555x.0000512.
[35] M. Fahmy Amin, "Confusion Matrix in Binary Classification Problems: A Step-by-Step Tutorial," *J. Eng. Res.*, vol. 6, no. 5, 2022, doi: 10.21608/erjeng.2022.274526.
[36] Herman, H. Darwis, Nurfauziyah, R. Puspitasari, D. Widyawati, and A. Faradibah, "Comparative Analysis of Anxiety Disorder Classification Using Algorithm Naïve Bayes, Decision Tree and K-NN," in *2025 19th International Conference on Ubiquitous Information Management and Communication (IMCOM)*, IEEE, Jan. 2025, pp. 1–6. doi: 10.1109/IMCOM64595.2025.10857485.
[37] T. Wang, C. Xue, Z. Zhang, T. Cheng, and G. Yang, "Unraveling the distinction between depression and anxiety: A machine learning exploration of causal relationships," *Comput. Biol. Med.*, vol. 174, p. 108446, May 2024, doi: 10.1016/j.compbiomed.2024.108446.