Journal of Applied Informatics and Computing (JAIC)
Vol.10, No.1, February 2026, pp. 698~706
e-ISSN: 2548-6861

Comparison of K-Nearest Neighbor, Naïve Bayes, and C4.5 Algorithms for Predicting Academic Stress Risk in Students Based on Psychological Survey Data

Nur Annisa Widya Pertiwi [1]*, Iwan Setiawaan Wibisono [2]**
* Teknik Informatika, Univeritas Ngudi Waluyo
** Teknik Informatika, Universitas Ngudi Waluyo
nurannisawidyap@gmail.com [1], iwansetiawan@unw.ac.id [2]

Article Info

Article history:

Received 2025-12-04
Revised 2025-12-26
Accepted 2026-01-07

Keyword:

Academic Stress,
K-Nearest Neighbor,
Naïve Bayes,
C4.5.

ABSTRACT

Academic stress is a psychological problem experienced by many students and can have an impact on learning achievement, mental health, and quality of life. This study aims to compare the performance of the K-Nearest Neighbor (KNN), Naïve Bayes, and C4.5 (Decision Tree) algorithms in predicting the level of academic stress risk in students based on psychological survey data. Data were obtained from 700 active students at Ngudi Waluyo University through a questionnaire covering physiological, psychological, and behavioral aspects, with a total of 15 indicators using a Likert scale. The data then underwent pre-processing, labeling, standardization, and division into training and test data with a ratio of 80:20. The evaluation was conducted using the Accuracy, Precision, Recall, F1-Score, Confusion Matrix, and AUC-ROC metrics. The results showed that the Naïve Bayes algorithm performed best with an accuracy of 93.26%, precision of 93.35%, recall of 92.26%, and F1-score of 92.80%. The KNN algorithm was in second place with an accuracy of 91.43%, while the C4.5 algorithm had the lowest performance with an accuracy of 80.60%. Based on these results, Naïve Bayes is recommended as the most optimal algorithm for predicting academic stress risk in students using psychological survey data. This study is expected to assist educational institutions in identifying students at risk of stress early on and supporting the development of more effective prevention strategies.

I. INTRODUCTION

Stress is a psychological phenomenon that can affect individuals of all ages. This state can elicit physiological and psychological responses when an individual is subjected to stress. In other words, stress denotes pain resulting from various demands or challenges imposed by external factors [1]. Within the framework of lectures, stress is a factor that influences students' capacity to adjust to academic requirements and learning environments.

Academic stress among students has become one of the most common psychological problems experienced by students and has negative impacts that affect academic achievement, mental health, and the continuity of studies [2]. Academic pressure arises from the demands of achieving good grades, excessive workload, limited time to complete assignments, academic competition, and the need to adapt to the college environment, which are the main factors contributing to academic stress among students. If this condition is not addressed immediately, academic stress in students has the potential to develop into serious psychological disorders, such as excessive anxiety in facing tasks, mental fatigue, and a decrease in motivation or interest in learning.

With the development of information technology today, machine learning approaches are increasingly being used to assist in the process of analyzing and predicting academic stress risk objectively. Several previous studies have applied classification algorithms such as K-Nearest Neighbor (KNN), Naïve Bayes, and Decision Tree C4.5 to predict students' psychological conditions based on survey data or psychological questionnaires [3]. According to Arya's (2024) research, the Naïve Bayes algorithm is capable of producing a relatively high level of accuracy in predicting student academic stress based on Likert questionnaires [2]. Meanwhile, other studies show that the performance of the K-Nearest Neighbor (KNN) and Decision Tree C4.5 algorithms tends to vary and depends on the number of attributes, data characteristics, and class distribution used [4].

Although previous studies have shown the potential of machine learning in predicting academic stress, most of these studies still have a number of shortcomings. Some studies use a relatively limited number of respondents, resulting in poor and suboptimal algorithm generalization capabilities. In addition, many studies apply only one or two algorithms without conducting a comprehensive comparison of the performance between classification methods. Discussions on the interpretability of algorithms in previous studies have not been explored in depth, even though in the fields of psychology and education, the ability to identify the most influential indicators of academic stress is crucial to support the roles of lecturers, counselors, and policymakers in higher education settings. On the other hand, the use of algorithm validation techniques such as K-Fold Cross Validation is still rarely used, which has the potential to cause the model evaluation results to be less stable because they depend on a single data division [1] [5].

Based on the above description, there is still a research gap that needs to be discussed further, namely the need for comparative research that compares the accuracy levels of several classification algorithms and analyzes the differences in algorithm performance and its relationship with the psychological characteristics of students. Therefore, this study aims to compare the performance of the K-Nearest Neighbor (KNN), Naïve Bayes, and C4.5 algorithms in predicting student academic stress risk based on structured psychological survey data.

This study contributes through the use of a dataset with a relatively large and adequate number of respondents, the use of diverse and comprehensive evaluation metrics, and the presentation of algorithm performance analysis that considers the assumptions of the methods and characteristics of the data used. With this approach, this study is expected to contribute to the development of machine learning studies in the fields of psychology and education and provide practical support for universities in their efforts to analyze the risk of academic stress among students.

II. METHOD

This study uses a quantitative method with a comparative experimental approach that aims to test and compare three machine learning algorithms, namely K-Nearest Neighbors (KNN), Naïve Bayes, and C4.5 (Decision Tree). Furthermore, this study aims to determine the most accurate algorithm in predicting the level of academic stress risk in students, based on data obtained from psychological survey results. The researcher used quantitative methods because it involved numerical data obtained from questionnaires filled out by students.

This research was conducted from September 2025 to January 2026 at Ngudi Waluyo University. This location was chosen because it has a diverse student population from various study programs and semesters, so it is considered to represent the general conditions of academic stress on campus. The data used in this study came from the results of a psychological survey created by the researcher and distributed online using Google Forms. This survey was completed by active students from various faculties at Ngudi Waluyo University who were willing to be respondents. The data from the questionnaire was then used as the basis for analysis and comparison of algorithms to predict academic stress levels.

This study used two types of variables. The independent variables were various factors that were thought to influence students' stress levels, such as coursework load, hours of sleep, anxiety levels, social support from friends or family, adaptability, and how students cope with pressure or problems. Meanwhile, the dependent variable was the level of academic stress risk among students, which was the final result to be predicted. This stress level is grouped into three categories, namely low, medium, and high, based on the results of questionnaires filled out by students.

In preparing this study, several stages were carried out systematically. These stages describe the research flow from start to finish, which can be seen in the following figure 1.

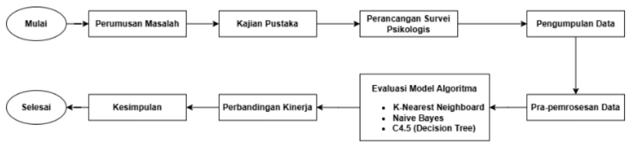

Figure 1. Research Flow

A. *Problem Formulation*

At this stage, the researcher began to identify the main problem, namely the signs of academic stress experienced by students due to pressure from assignments, learning demands, and changes in the learning system. This condition became the reason for predicting the level of student stress using three machine learning algorithms to determine the most accurate and efficient algorithm.

B. *Literarure Review*

At this stage, the aim is to review previous theories and research on the application of the K-Nearest Neighbors (KNN), Naïve Bayes, and C4.5 (Decision Tree) algorithms. This literature review is conducted to understand the performance of each algorithm and to identify the advantages and disadvantages of the algorithms in processing psychological survey data [3].

C. *Pyschological Survey Design*

At this stage , the process of compiling data collection tools was carried out with the aim of measuring and analyzing psychological aspects. This process included compiling questionnaires tailored to the research objectives, identifying the variables and indicators to be measured, selecting the appropriate assessment scales, and testing the validity and reliability of the questionnaires [4]. This psychological survey not only serves as a tool for collecting respondent data but also as a measurement tool to evaluate the level of academic stress risk through measurable indicators and clear rating scales.

D. Data Collection

The data collection stage is an important process in research to obtain information relevant to the topic being studied. At this stage, data is collected through observation, literature study, and online questionnaire distribution using Google Forms to reach a wider and more efficient range of respondents [3] [5].

E. Data Pre-processing

This stage is very important in this research because it affects the quality of data that will be used in the training and testing of algorithms. This stage aims to ensure that the data used is clean, consistent, relevant, and well-structured so that the algorithm can work optimally. Data preprocessing is also the first step in addressing various problems such as missing values, data duplication, and entry errors. Poorly processed data can reduce the performance and accuracy of the algorithm [6].

The dataset in this study was obtained from psychological surveys conducted through online questionnaires using a five-point Likert scale (1-5). These online questionnaires covered 15 indicators of physiological, psychological, and behavioral aspects of academic stress. The total academic stress score for each respondent was calculated by adding up all the scores on the indicators, resulting in a score range of 15 to 75. The following is the Likert scale used in this study.

TABLE I
LIKERT SCALE

Scale	Description
1	Never
2	Rarely
3	Sometimes
4	Often
5	Always

Academic stress levels are grouped into three categories: low, moderate, and high. This grouping is done using an interval score classification approach that is commonly used in the development of psychological scales [7]. This approach assumes that Likert scale data are ordinal numerical data that can be processed through total score calculations and interpreted based on specific value ranges. Based on the accumulated scores, academic stress levels are classified into the following three categories.

TABLE II
STRESS LEVEL CATEGORY

Score	Description
15-30	Low
31-50	Moderate
51-75	High

This method has been widely used in previous studies examining academic stress among students based on questionnaire instruments [2] [8].

Data cleaning was performed to ensure that there was no missing or duplicate data that could potentially cause bias in the analysis process. The verification results showed that all respondent data was complete and valid, allowing it to proceed to the next stage. Next, the academic stress level categories were converted into numerical form through an encoding process so that they could be recognized and processed using machine learning algorithms. In addition, data standardization was performed using the StandardScaler method to equalize the value range of each feature. This stage is important because distance-based algorithms (k), such as K-Nearest Neighbor (KNN), are very sensitive to differences in scale between features, so standardization is necessary to produce a more accurate classification process [9].

F. Algorithm Model Evaluation

This stage is carried out after the survey data has been collected and has gone through the data pre-processing stage. At this stage, researchers begin to apply algorithms to test the performance of each algorithm and determine the algorithm that has the highest accuracy for the data that is ready to be used.

To ensure reliable algorithm evaluation, this study uses a combination of hold-out validation and K-Fold Cross Validation. The dataset is divided into training data and test data with a ratio of 80:20 to measure the algorithm's performance on previously unprocessed data. In addition, 5-Fold Cross Validation is also applied to improve the reliability and generalization ability of the algorithm. The dataset is divided into five equal parts, where four parts are used as training data and one part as test data in each iteration. This process is repeated five times so that each part acts as test data once. The final performance value is obtained from the average of all iterations.

The selection of K=5 is based on previous machine learning research recommendations, which show that five folds provide a balance between evaluation efficiency and stability, especially for medium-sized datasets [10]. The use of cross-validation is important to reduce evaluation bias due to single data division and to ensure that the algorithm's performance reflects its true generalization ability [1].

Algorithm performance evaluation was carried out using the Accuracy, Precision, Recall, F1-Score, Confusion Matrix, and AUC-ROC metrics, which are commonly used in comparative studies of classification algorithms [7] [10-15]. Below is an explanation of the algorithms used in this study.

- *K-Nearest Neighbor (KNN)*

The K-Nearest Neighbor (KNN) algorithm is a method in supervised learning that is used for clustering (classification) or prediction (regression). The way KNN works is quite simple, namely by determining the similarity distance between new data and existing data in the training dataset. The new data is then classified into the group that appears most frequently (majority) among a number of nearby data (usually referred to as the k value) [7] [9]. To calculate the distance between two data points, KNN generally uses Euclidean Distance, which is the straight-line distance between two points in the data space. The formula is as follows:

$$D\ (X_1, X_2) = \sqrt{\sum_{i=1}^{n} (X_{1i} - X_{2i})^2}$$

Explanation:
$D(X_1, X_2)$ = Euclidean distance between two data points,
X_{1i} and X_{2i} = the two attribute values whose distance is to be calculated,
n = the number of attributes used.

After the distance between each point is calculated, a number of k nearest neighbors are selected, then the test data is categorized based on the most dominant class of those neighbors. The advantages of the KNN algorithm lie in its simplicity and adaptability, but its disadvantage is the considerable computational time required when the data volume increases.

- *Naïve Bayes*

The Naïve Bayes algorithm is a classification technique based on Bayes' theorem of probability, with the assumption that each variable or attribute is independent of one another. This approach is used to calculate the probability of a data point belonging to a certain category or class based on the values of its attributes [7] [10] [8]. The equation is expressed as:

$$P(H|X) = \frac{P(X|H)\ \times P(H)}{P(X)}$$

Explanation:
X = data with indetermine classification
H = hypothesis asserting that data X is categorized inside a particular class
P(H|X) = probability of hypothesis H contigent upon condition X
P(X|H) = probability of data X contigent upon the hypothesis H
P(H) = probability of hypothesis H
P(X) = probability of data X

The stages of this process include calculating each class based on training data, then determining the class with the highest probability value for the test data. The Naïve Bayes algorithm is simple but very effective, especially when applied to data with large attributes.

- *C4.5*

The C4.5 algorithm is a data mining algorithm used to classify categories. This algorithm is an extension of the ID3 algorithm. The C4.5 algorithm works by building a decision tree to generate a decision. The decision tree involves the process of forming data (tables) into a tree model, then converting the tree into rules, and simplifying those rules [7]. The formula for calculating information gain is:

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^{n} \frac{|S_i|}{|S|} \times\ Entorpy\ (S_{i)}$$

Explanation:
S = set of cases
A = attribute used for splitting
n = number of partitions of attribute A
$|S_i|$ = number of cases in partition i
|S| = total number of cases
Where:

$$Entropy(S) = -\sum_{i=1}^{n} p_i log_2 p_1$$

Explanation:
S = set of cases
n = number of partitions of the set of cases S
p_1 = proportion of the set of cases in i relative to S

A higher gain rate indicates that the attribute is more effective in separating data. The C4.5 algorithm has the advantage of being able to manage continuous attributes, handle incomplete data, and perform tree pruning to prevent overfitting.

G. Algorithm Comparison

At this stage, algorithm performance is assessed using classification evaluation measures, including Accuracy, Precision, Recall, F1-Score, Confusion Matrix, and AUC-ROC. This stage aims to measure how well each algorithm is able to accurately predict the level of academic stress risk in students. Next, the results of the comparison of the three algorithms are analyzed to determine the algorithm that shows the most optimal performance [1] [10].

H. Conclusion

The conclusion stage is the closing part of the research that contains a summary of the analysis results and main findings. At this stage, an assessment is made of the performance of the K-Nearest Neighbor, Naïve Bayes, and C4.5 algorithms in predicting the risk of academic stress among students. The results are used to determine the most appropriate algorithm based on accuracy and efficiency. In addition, this stage also provides answers to the research questions and serves as a basis for recommendations for future research.

III. Results and Discussion

This chapter presents the results of processing the psychological survey data of students and the performance analysis of each algorithm applied in this study, namely K-Nearest Neighbor (KNN), Naïve Bayes, and C4.5. The results are presented in stages, starting from data collection and pre-processing, implementation of each algorithm, to model performance evaluation based on accuracy, precision, recall, F1-score, confusion matrix, and AUC-ROC value. This chapter is designed to explain the relationship between this study and previous studies, thereby providing a comprehensive overview of the capabilities of each algorithm in predicting the risk of academic stress among students.

A. Data Collection

The data collection process was carried out through a psychological survey distributed online using Google Forms to active students at Ngudi Waluyo University. The online survey method was chosen because it was able to reach a large number of respondents and facilitated the data processing process [4] [5]. During the data collection period, 700 respondents completed the questionnaire completely and validly. All data was then checked for completeness and declared valid to proceed to the next stage, which was pre-processing of the data. The following data was obtained from the questionnaire results.

	Timestamp	Program Studi	Semester (Tulis Angka)	Jenis Kelamin	Usia	q1	q2	q3	q4	q5	q6	q7	q8	q9	q10	q11	q12	q13	q14	q15
0	10/2/2025 16.06.47	S1 Farmasi	7	Perempuan	20 □ 22 Tahun	3	3	4	3	3	3	2	3	3	2	1	3	3	1	3
1	10/2/2025 18.14.34	Teknik Informatika	5	Perempuan	<20 Tahun	5	4	3	3	4	5	3	3	2	4	3	3	3	3	5
2	10/2/2025 18.16.03	Teknik Informatika	5	Laki-Laki	20 □ 22 Tahun	3	4	5	3	5	2	3	3	4	5	5	2	1	1	5
3	10/2/2025 18.18.00	S1 TI	7	Laki-Laki	20 □ 22 Tahun	2	3	3	3	3	4	3	2	2	3	2	1	1	1	2
4	10/2/2025 18.21.39	Teknik informatika	7	Laki-Laki	20 □ 22 Tahun	4	4	3	2	2	3	5	4	3	4	4	4	2	4	4
5	10/2/2025 18.23.07	Teknik informatika	7	Perempuan	20 □ 22 Tahun	3	3	2	2	3	3	3	2	1	3	2	1	1	1	3
6	10/2/2025 18.26.12	S1 Teknik Informatika	7	Laki-Laki	<20 Tahun	2	1	1	1	3	3	2	5	3	3	3	4	1	2	2
7	10/2/2025 18.26.26	S1 Farmasi	7	Perempuan	20 □ 22 Tahun	3	1	4	3	2	3	3	1	1	1	1	1	1	1	1
8	10/2/2025 18.26.53	S1 Gizi	7	Perempuan	20 □ 22 Tahun	4	2	5	3	3	4	3	2	2	2	3	2	1	4	2
9	10/2/2025 18.27.01	s1 teknik informatika	7	Perempuan	20 □ 22 Tahun	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3

Figure 1. Questionnaire Results

B. Data Pre-processing

Data preprocessing is an important step to ensure data quality before it is used in machine learning algorithm training and testing. This step is performed using the Python programming language through the Google Colab platform. The purpose of this step is to prepare the data so that it is clean, structured, and ready to be processed by the algorithm. The initial process involves checking for empty and duplicate data, as well as removing several columns that are not important. This process is called data cleaning. The following are the results of the initial check in the data pre-processing stage.

	Jenis Kelamin	Usia	q1	q2	q3	q4	q5	q6	q7	q8	q9	q10	q11	q12	q13	q14	q15
0	Perempuan	20 □ 22 Tahun	3	3	4	3	3	3	2	3	3	2	1	3	3	1	3
1	Perempuan	<20 Tahun	5	4	3	3	4	5	3	3	2	4	3	3	3	3	5
2	Laki-Laki	20 □ 22 Tahun	3	4	5	3	5	2	3	3	4	5	5	2	1	1	5
3	Laki-Laki	20 □ 22 Tahun	2	3	3	3	3	4	3	2	2	3	2	1	1	1	2
4	Laki-Laki	20 □ 22 Tahun	4	4	3	2	2	3	5	4	3	4	4	4	2	4	4
5	Perempuan	20 □ 22 Tahun	3	3	2	2	3	3	3	2	1	3	2	1	1	1	3
6	Laki-Laki	<20 Tahun	2	1	1	1	3	3	2	5	3	3	3	4	1	2	2
7	Perempuan	20 □ 22 Tahun	3	1	4	3	2	3	3	1	1	1	1	1	1	1	1
8	Perempuan	20 □ 22 Tahun	4	2	5	3	3	4	3	2	2	2	3	2	1	4	2
9	Perempuan	20 □ 22 Tahun	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3

Figure 2. Data Cleaning Results

Next, after the data cleaning process, the labeling process is carried out by summing the total scores for each question on the questionnaire into categories of academic stress risk, namely low, medium, and high. The following are the data labeling results.

	Jenis Kelamin	Usia	q1	q2	q3	q4	q5	q6	q7	q8	q9	q10	q11	q12	q13	q14	q15	Jumlah	Kategori
0	Perempuan	20 □ 22 Tahun	3	3	4	3	3	3	2	3	3	2	1	3	3	1	3	40	Sedang
1	Perempuan	<20 Tahun	5	4	3	3	4	5	3	3	2	4	3	3	3	3	5	53	Tinggi
2	Laki-Laki	20 □ 22 Tahun	3	4	5	3	5	2	3	3	4	5	5	2	1	1	5	51	Tinggi
3	Laki-Laki	20 □ 22 Tahun	2	3	3	3	3	4	3	2	2	3	2	1	1	1	2	35	Sedang
4	Laki-Laki	20 □ 22 Tahun	4	4	3	2	2	3	5	4	3	4	4	4	2	4	4	52	Tinggi
5	Perempuan	20 □ 22 Tahun	3	3	2	2	3	3	3	2	1	3	2	1	1	1	3	33	Sedang
6	Laki-Laki	<20 Tahun	2	1	1	1	3	3	2	5	3	3	3	4	1	2	2	36	Sedang
7	Perempuan	20 □ 22 Tahun	3	1	4	3	2	3	3	1	1	1	1	1	1	1	1	27	Rendah
8	Perempuan	20 □ 22 Tahun	4	2	5	3	3	4	3	2	2	2	3	2	1	4	2	42	Sedang
9	Perempuan	20 □ 22 Tahun	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	45	Sedang

Figure 3. Data Labeling Results

To address differences in score ranges between features, data standardization was performed using the StandardScaler method. Data standardization is particularly important for distance-based algorithms such as K-Nearest Neighbor (KNN) because differences in feature scales can affect distance calculations [6]. The dataset was then divided into two parts, namely training data and test data, with a ratio of 80:20. This division was intended to evaluate the algorithm's generalization ability against previously unrecognized data. The following is the division of training data and test data.

TABLE III
DATA DIVISION

Data Type	Number
Training Data	560
Testing Data	140
	700

C. Algorithm Model Evaluations

After the data has been processed, the next step is to apply three classification algorithms, namely K-Nearest Neighbor (KNN), Naïve Bayes, and C4.5 (Decision Tree). To determine the performance of each algorithm, an assessment was carried out using accuracy, precision, recall, and F1-score metrics, a confusion matrix to see the distribution of predictions, and an AUC-ROC curve to observe the model's ability to distinguish between classes. All of these metrics were used to measure how accurate and consistent the algorithms were in classifying the stress risk levels of students.

1) K-Nearest Neighbor (KNN)

The results of testing the K-Nearest Neighbor (KNN) algorithm with a value or distance of K = 5 show that this algorithm is capable of producing an accuracy rate of 91.43%. The high precision, recall, and F1-score values also indicate that K-Nearest Neighbor (KNN) is quite good and effective in classifying academic stress data in students, especially in the majority class, which is moderate stress. The following are the results of testing the K-Nearest Neighbor (KNN) algorithm.

```
...
--- Evaluasi: K-Nearest Neighbors (K=5) ---
Accuracy : 0.9142857142857143
Precision: 0.9150243110064539
Recall   : 0.9142857142857143
F1-score : 0.9093192648059906

Classification report:
              precision    recall  f1-score   support

      Rendah       0.86      0.75      0.80         8
      Sedang       0.91      0.98      0.95       109
      Tinggi       0.94      0.65      0.77        23

    accuracy                           0.91       140
   macro avg       0.90      0.79      0.84       140
weighted avg       0.92      0.91      0.91       140
```

Figure 4. KNN algorithm evaluation results

However, the confusion matrix results show that there are still some errors in predicting the low stress and high stress categories. This is due to data imbalance, where the moderate stress category has a larger number of samples. The K-Nearest Neighbor (KNN) algorithm performs classification based on the proximity of the distance between classes, so that data in classes with small numbers or minorities tend to be classified into the majority class. The results of this study are in line with Permana's (2021) research, which states that the performance of K-Nearest Neighbor (KNN) is very sensitive to data distribution and data balance. The following are the results of the confusion matrix of the K-Nearest Neighbor (KNN) algorithm.

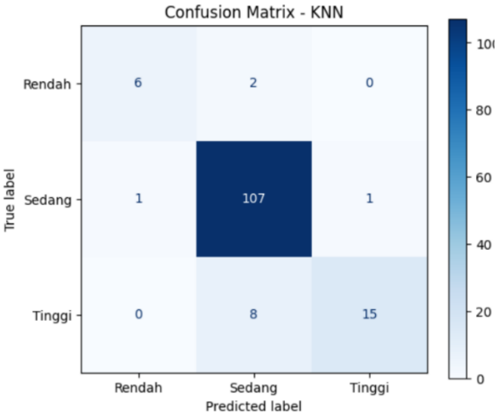

Figure 5. Confusion matrix of the KNN algorithm

The AUC-ROC results show that high values across all classes indicate that the K-Nearest Neighbor (KNN) algorithm has good class discrimination capabilities. However, the performance of K-Nearest Neighbor (KNN) is influenced by data distribution and imbalance. K-Nearest Neighbor (KNN) is less optimal when compared to other probabilistic-based algorithms when applied to psychological survey data. Below are the AUC-ROC results of the K-Nearest Neighbor (KNN) algorithm.

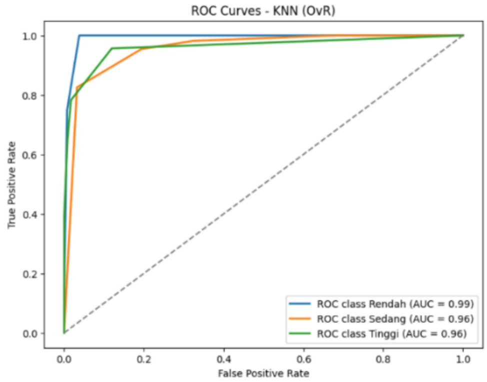

Figure 6. AUC-ROC graph of the KNN algorithm

2) Naïve Bayes

The Naïve Bayes algorithm showed the best performance compared to the other two algorithms tested in this study, with an accuracy rate of 93.26%. The high and balanced precision, recall, and F1-score values in each class indicate that Naïve Bayes has consistent and stable classification capabilities, both in recognizing classes with small numbers and classes with large numbers in the student psychological survey data. The results of the Naïve Bayes algorithm testing are presented visually as follows.

```
... --- Evaluasi: Naive Bayes (GaussianNB) ---
Accuracy : 0.9285714285714286
Precision: 0.9334938077795221
Recall   : 0.9285714285714286
F1-score : 0.9300818858918736

Classification report:
              precision    recall  f1-score   support

      Rendah       0.89      1.00      0.94         8
      Sedang       0.97      0.94      0.95       109
      Tinggi       0.77      0.87      0.82        23

    accuracy                           0.93       140
   macro avg       0.88      0.94      0.90       140
weighted avg       0.93      0.93      0.93       140
```

Figure 7. Evaluation results of the Naive Bayes algorithm

The advantages of the Naïve Bayes algorithm can be explained based on the characteristics of the student psychological survey data used in the study. Likert scale-based data tends to have low correlations between indicators, so the assumption of feature independence in this algorithm is theoretically acceptable [10] [8]. In addition, the Naive Bayes algorithm is known to be effective in handling data with large sample sizes, as shown in Arya's (2024) study, which demonstrated high accuracy in prediction [2].

The confusion matrix results below show that most of the data in the low stress and moderate stress categories were correctly classified by the Naive Bayes algorithm, while the classification error rate in the high stress category was relatively low.

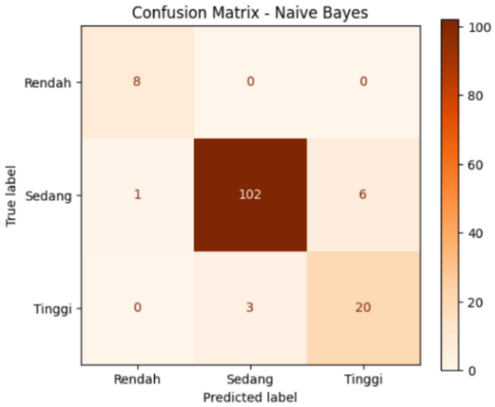

Figure 8. Naive Bayes confusion matrix

The evaluation results using the AUC-ROC curve show a value close to 1, which means that the Naive Bayes algorithm has excellent ability in distinguishing each class in the student psychological survey data. This study confirms that the Naive Bayes algorithm is the most appropriate approach and is very suitable for the characteristics of the student psychological survey data used in this study. Below is a visual representation of the AUC-ROC curve.

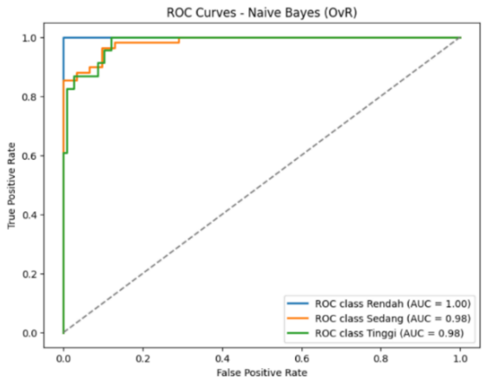

Figure 9. Naive Bayes AUC-ROC graph

3) C4.5

The C4.5 (Decision Tree) algorithm showed the lowest performance compared to the other two algorithms in this study, with an accuracy rate of 80.60%. Based on the evaluation results, the C4.5 (Decision Tree) algorithm has a fairly strong tendency to group data into classes with high numbers, namely the moderate stress category, so that its ability to recognize classes with small numbers becomes less effective and less optimal. The following are the results of testing using the C4.5 (Decision Tree) algorithm.

```
--- Evaluasi: Decision Tree (C4.5 style) ---
Accuracy : 0.8
Precision: 0.791392450785624
Recall   : 0.8
F1-score : 0.7949247253898416

Classification report:
              precision    recall  f1-score   support

      Rendah       0.43      0.38      0.40         8
      Sedang       0.86      0.89      0.87       109
      Tinggi       0.60      0.52      0.56        23

    accuracy                           0.80       140
   macro avg       0.63      0.60      0.61       140
weighted avg       0.79      0.80      0.79       140
```
Figure 10. Evaluation results of the C4.5 algorithm

The confusion matrix results show that the C4.5 (Decision Tree) algorithm still experiences errors in distinguishing between the low stress and high stress categories, as indicated by the presence of data from both classes that are misclassified as moderate stress. The following are the confusion matrix results of the C4.5 (Decision Tree) algorithm.

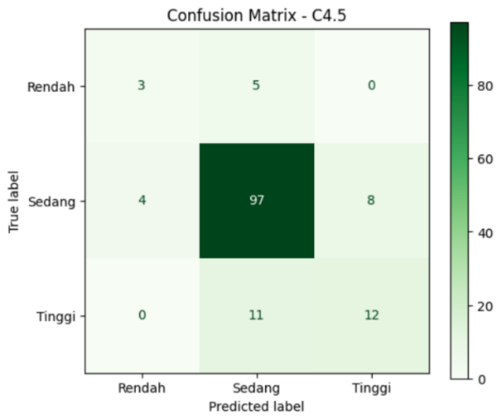

Figure 11. Confusion matrix of the C4.5 algorithm

The performance of the C4.5 (Decision Tree) algorithm is lower than the other two algorithms due to the characteristics of the C4.5 (Decision Tree) algorithm, which is sensitive to data imbalance and tends to form rules based on classes with large amounts of data [10]. In addition, psychological survey data based on the Likert scale is numerical, so the patterns formed are far from optimal when processed using information gain-based separation.

Although the C4.5 (Decision Tree) algorithm has advantages in terms of algorithm interpretability, the results of this study show that the level of accuracy in prediction is a more crucial factor. This study is in line with Renaningtias' (202) research, which shows that the performance of the C4.5 (Decision Tree) algorithm is lower than that of the K-Nearest Neighbor (KNN) algorithm in predicting the risk of academic stress in students [11].

The AUC-ROC values produced by the C4.5 (Decision Tree) algorithm also show suboptimal class discrimination, especially in classes with small amounts of data. The following is the AUC-ROC curve of the C4.5 (Decision Tree) algorithm..

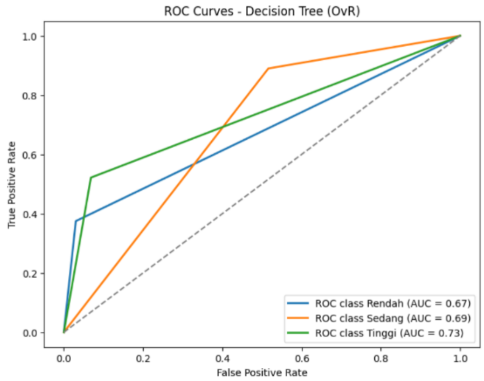

Figure 12. AUC-ROC graph of the C4.5 algorithm

D. *Algorithm Comparison*

At this stage, each algorithm is evaluated for performance using classification metrics that include Accuracy, Precision, Recall, F1-Score, Confusion Matrix, and AUC-ROC, which aim to measure the algorithm's ability to accurately predict the level of academic stress risk in students. Then, the results of this performance comparison are used to determine the algorithm that shows the most optimal performance in predicting the level of academic stress risk in students.

Based on the results of the algorithm evaluation, it can be concluded that the Naïve Bayes algorithm is the most appropriate and optimal algorithm in predicting the level of academic stress risk in students. The advantages of the Naïve Bayes algorithm are not only demonstrated by the highest level of accuracy, but also by its stability of performance between classes and its good and optimal generalization capabilities. The K-Nearest Neighbor (KNN) algorithm is in second place with fairly good performance, but its performance is still affected by the imbalance in data distribution. Meanwhile, the C4.5 (Decision Tree) algorithm shows shortcomings in handling psychological survey data on students, which is numerical and unbalanced.

From the aspect of algorithm interpretability, psychological and behavioral indicators are proven to be the most dominant factors in the process of classifying the level of academic stress risk in students. This study shows that emotional responses, learning behavior patterns, and how students cope with academic pressure contribute to determining stress levels. This information has high practical and strategic value for universities to identify students at risk of academic stress. The following are the results of the algorithm performance comparison in this study.

	Model	Akurasi (%)	Presisi (Weighted) (%)	Recall (Weighted) (%)	F1-Score (Weighted) (%)
0	KNN	91.43	91.50	91.43	90.93
1	Naive Bayes	92.86	93.35	92.86	93.01
2	Decision Tree (C4.5)	80.00	79.14	80.00	79.49

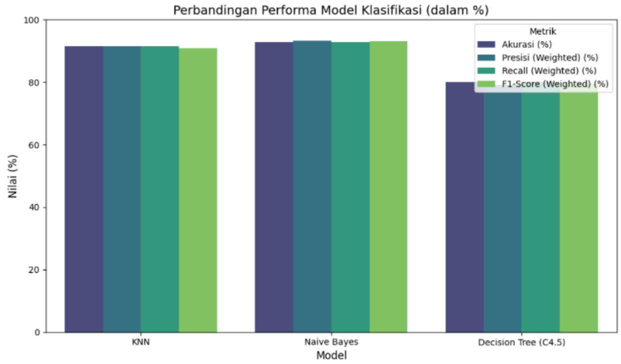

Figure 13. Comparison of algorithm performance

E. *Practical Implications*

The results of this study can be used as a basis for analysis in systematically evaluating the academic stress experienced by students. Information regarding the level of academic stress risk obtained from the analysis of psychological survey data can be used by universities or academic administrators as material for consideration in formulating more responsive academic policies, developing student assistance and guidance programs, and providing counseling services based on actual student data and conditions.

F. *Research Limitations*

This study still has many shortcomings that need to be considered in interpreting the results. First, the research data source only comes from one university, so the level of generalization of this study is only focused on one university with different student characteristics and is still very limited. Second, data collection was conducted using a self-report questionnaire, which allowed for subjective bias among respondents, such as respondents giving answers that did not reflect their actual psychological condition. Third, the classification algorithm used in this study is still very limited to the machine learning approach, so future research is recommended to examine and compare more modern algorithms to produce a more comprehensive, accurate, and highly reliable analysis.

IV. CONCLUSION

Based on the results of the analysis and discussion, it can be concluded that the Naive Bayes algorithm shows the best and most optimal performance in predicting the level of academic stress risk among students compared to the K-Nearest Neighbor (KNN) and C4.5 (Decision Tree) algorithms. Naive Bayes has advantages as shown by its Accuracy (93.26%), Precision (93.35%), Recall (92.26%), and F1-Score (92.80%) values, which are consistently higher in almost all prediction classes.

The superior performance of the Naive Bayes algorithm can be explained by the characteristics of the data used in this study. The data came from a Likert scale-based psychological questionnaire that had relatively low correlations between factors and a stable distribution of values. This is in line with the basic assumption of the Naive Bayes algorithm, which considers each feature to be independent, enabling the Naive Bayes algorithm to effectively calculate class probabilities in psychological survey data.

The K-Nearest Neighbor (KNN) algorithm ranked second with an accuracy of 91.43%. The K-Nearest Neighbor (KNN) algorithm performed well in classes with large amounts of data, especially in the moderate stress category, but its performance declined in the minority classes, namely the low stress and high stress categories. This was due to the sensitivity of the K-Nearest Neighbor (KNN) algorithm to data distribution and class imbalance.

Meanwhile, the C4.5 (Decision Tree) algorithm obtained the lowest performance results with an accuracy of 80.60%. These results indicate that the C4.5 (Decision Tree) algorithm tends to be biased towards the majority class and is less than optimal in distinguishing minority classes. The decision tree structure that is formed is also relatively shallow, making it less capable of capturing complex patterns in multidimensional psychological data.

REFERENCES

[1] D. S. Anastasya Carity, E. Simamora, Z. Indra, and S. Manullang, “Implementasi Algoritma Naïve Bayes Dalam Melakukan Klasifikasi Tingkat Stres Siswa SMA,” 2025.

[2] S. Arya, Anju, and N. A. Ramli, “Predicting the stress level of students using Supervised Machine Learning and Artificial Neural Network (ANN),” *Indian J. Eng.*, vol. 21, no. 56, 2024, doi: 10.54905/disssi.v21i55.e9ije1684.

[3] D. Toresa *et al.*, “Perbandingan Algoritma C4.5 Dan Naïve Bayes Untuk Mengukur Tingkat Kepuasan Mahasiswa Dalam Penggunaan Edlink,” *J. Teknol. Dan Sist. Inf. Bisnis*, vol. 5, no. 3, pp. 250–256, Jul. 2023, doi: 10.47233/jteksis.v5i3.855.

[4] A. R. Sinadia, “Sumikolah: Jurnal Ilmu Pendidikan Aplikasi dan Praktik Konstruksi Skala Psikologi sebagai Instrumen Pengukuran dalam Penelitian Pendidikan,” vol. 2, no. 1, 2024.

[5] I. Novita Sari, L. Fakih Lidimilah, and A. Lutfi, “Proceeding National Conference Of Research And Community Service Sisi Indonesia 2025 Analisis Perbandingan Algoritma Decision Tree, Random Forest, Dan Naïve Bayes Dalam Klasifikasi Gangguan Tidur.” [Online]. Available: https://conference.sinesia.id/ncrcs-sinesia

[6] S. S. Berutu, H. Budiati, J. Jatmika, and F. Gulo, “Data preprocessing approach for machine learning-based sentiment classification,” *J. INFOTEL*, vol. 15, no. 4, pp. 317–325, Nov. 2023, doi: 10.20895/infotel.v15i4.1030.

[7] A. P. Permana, K. Ainiyah, and K. F. H. Holle, “Analisis Perbandingan Algoritma Decision Tree, kNN, dan Naive Bayes untuk Prediksi Kesuksesan Start-up,” *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 6, no. 3, 2021, doi: 10.14421/jiska.2021.6.3.178-188.

[8] R. P. Lestari, N. Surojudin, and E. Budiarto, “Klasifikasi Tingkat Stres pada Mahasiswa Tingkat Akhir menggunakan Metode Naïve Bayes di Universitas Pelita Bangsa Stress Level Classification of Final-Year Students Using Naïve Bayes at Universitas Pelita Bangsa,” 2025. [Online]. Available: https://jurnal.unimed.ac.id/2012/index.php/cess

[9] A. Al Arif, M. Firdaus, Y. Maruhawa, S. AMIK Riau, and J. Purwodadi Panam, “SENTIMAS: Seminar Nasional Penelitian dan Pengabdian Masyarakat Comparison of Data Mining Methods for Predition of Rainfall with C4.5, Naïve Bayes, and KNN Algorithm Perbandingan Metode Data Mining untuk Prediksi Curah Hujan dengan Algoritma C4.5, Naïve B.” [Online]. Available: https://journal.irpi.or.id/index.php/sentimas

[10] M. F. Indriawan, J. P. P. Naibaho, and M. Sanglise, “Analisis Komparasi Algoritma Naïve Bayes Dan C4.5 Dalam Memprediksi Kelulusan Mahasiswa (Studi Kasus: Jurusan S1 Teknik Informatika Universitas Papua),” *JISTECH J. Inf. Sci. Technol.*, vol. 11, no. 2, 2023, doi: 10.30862/jistech.v11i2.107.

[11] N. Renaningtias, T. Eka Putri, E. Putri Purwandari, Y. Stainly Ritonga, and P. Sistem, “Studi Komparasi Algoritma Decision Tree C4.5 dan K-Nearest Neighbor pada Klasifikasi Masa Studi dan Tingkat Stres Mahasiswa”

[12] P. Prasetyawan, I. Ahmad, R. I. Borman, Ardiansyah, Y. A. Pahlevi and D. E. Kurniawan, "Classification of the Period Undergraduate Study Using Back-propagation Neural Network," 2018 International Conference on Applied Engineering (ICAE), Batam, Indonesia, 2018, pp. 1-5, doi: 10.1109/INCAE.2018.8579389.

[13] Y. C. Tapidingan and D. Paseru, “Comparative Analysis of Classification Methods of KNN and Naïve Bayes to Determine Stress Level of Junior High School Students”, Indonesian J. of Inf. Syst., vol. 2, no. 2, pp. 80–89, Feb. 2020.

[14] 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," 2025 19th International Conference on Ubiquitous Information Management and Communication (IMCOM), Bangkok, Thailand, 2025, pp. 1-6, doi: 10.1109/IMCOM64595.2025.10857485.

[15] I. Gunawan and A. P. Widyassari, “Integrating Psychological Stress Indicators with Academic Data for Student Dropout Prediction: A Decision Tree and Expert System Approach ”, ELINVO, vol. 10, no. 2, pp. 131–146, Dec. 2025.