

Predicting Student Dependency on ChatGPT for Academic Tasks Using Naive Bayes Classification

Risha Febrianti[✉]

Universitas Negeri Makassar, Makassar, Indonesia

Sul Fitriana

Universitas Negeri Makassar, Makassar, Indonesia

Asrafah

Universitas Negeri Makassar, Makassar, Indonesia

Stephen Amukune

MATE Institute of Education, Hungary

ABSTRACT

Purpose – This study aims to predict and classify the level of student dependency on ChatGPT in completing academic tasks using the Naive Bayes algorithm to support data-driven decision making in higher education.

Methods – A quantitative survey approach was employed involving 254 active undergraduate students from the Department of Informatics and Computer Engineering at a public university in Indonesia. Data were collected through a Likert-scale questionnaire measuring five behavioral indicators: purpose of ChatGPT use, interaction frequency and duration, understanding of generated outputs, trust in AI responses, and learning independence. The collected data were cleaned, numerically encoded, and labeled into three dependency categories (low, medium, high). A Naive Bayes classification model was implemented using Orange Data Mining and evaluated under three data split scenarios: 90:10, 80:20, and 70:30.

Findings – The results indicate that the 70:30 data split achieved the highest classification performance, with an AUC value of 0.973, accuracy of 85.3%, F1-score of 0.866, and precision of 0.909. These results demonstrate that the Naive Bayes algorithm is effective in identifying distinct patterns of student dependency on ChatGPT based on multidimensional behavioral data.

Research limitations – This study is limited to a single academic program and relies on self-reported questionnaire data, which may constrain the generalizability of the findings across different educational contexts.

Originality – This study provides empirical evidence on the application of probabilistic classification models to assess student dependency on generative AI, contributing to educational decision sciences by informing institutional policies on balanced and responsible AI use in higher education.

 OPEN ACCESS

ARTICLE HISTORY

Received: 05-05-2026

Revised: 21-07-2026

Accepted: 02-08-2026

KEYWORDS

ChatGPT dependency;
Educational data mining;
higher education decision making;
Naive Bayes classification;
Student learning independence

Correspondence Author: ✉ rshfbrnti01@gmail.com

To cite this article : Risha Febrianti, Sul Fitriana, Asrafah & Stephen Amakune. (2026). Predicting Student Dependency on ChatGPT for Academic Tasks Using Naive Bayes Classification. Artificial Intelligence in Educational Decision Sciences, 1(2), 10–21.

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



INTRODUCTION

The development of artificial intelligence (AI) technology has brought about a major transformation in higher education, particularly in the way students complete academic assignments (Search, 2023). One of the most prominent technologies is ChatGPT, a generative language model capable of automatically generating text, answering questions, and constructing arguments (Maula et al., 2024). This phenomenon has encouraged students to utilize ChatGPT as a primary tool for understanding lecture materials and completing text-based assignments (Kurniasari et al., 2025). Research by Ningrum et al. (2024) shows that the use of ChatGPT facilitates the learning process, but also poses new challenges related to academic ethics and the validity of student work. Dependence on ChatGPT has become a serious concern in the world of education, as it can affect students' critical thinking and independent learning abilities (Fathony et al., 2024). Research by Endraswari et al. (2025) reveals that excessive use of ChatGPT has the potential to weaken students' learning initiative and analytical abilities. Additionally, research by Ridwan et al. (2024) shows that ease of use and compatibility with lifestyle are the main drivers for students to adopt ChatGPT, even though they do not fully understand its impact on academic quality (April-september et al., 2025). This phenomenon indicates a shift in learning behavior that needs to be studied systematically, especially in the context of vocational education that emphasizes independence and analytical thinking skills (Cahyanto et al., 2024).

Research by Siregar et al. (2024) highlights that the integration of ChatGPT in education needs to be supported by a behavior classification system to prevent misuse of the technology (Andrea et al., 2025). This study suggests the use of classification algorithms such as Naïve Bayes to monitor patterns of student interaction with AI systems. Meanwhile, research by Rieuwpassa et al. (2024) shows that Naïve Bayes has advantages in computational efficiency and probabilistic interpretation, making it suitable for analyzing user behavior in the context of digital education (Triharningsari et al., 2025). Based on the previous discussion, it is important to understand how classification methods such as Naïve Bayes have been utilized in various studies related to user behavior and dependence on artificial intelligence-based learning technologies (Munggaran et al., 2023). The following section discusses previous studies that form the theoretical and empirical basis for the development of a model for predicting student dependence on ChatGPT (Setiabudi & Sampurno, n.d.).

Research by Rizki & Hidayat (2023) applied the Naïve Bayes Classifier to classify the level of student dependence on online learning technology. Using student survey data, their model achieved an accuracy of 85%, indicating that Naïve Bayes is effective in detecting user behavior patterns in the context of e-learning. This study confirms that probabilistic-based classification is capable of predicting students' tendencies toward digital learning technology (Purbayanto & Suharsono, n.d.). Meanwhile, in the study by Wijaya & Kurniawan (2023), the Naïve Bayes algorithm was applied to analyze student engagement on e-learning platforms and achieved an accuracy of 81%. These results reinforce the evidence that this method can recognize learning interaction patterns well, especially in the context of online student behavior analytics (Nakhipova et al., 2024). In the context of ChatGPT usage, Sari & Nugroho (2024) examined student dependence on ChatGPT using the Naïve Bayes model. With survey data from 254 students, this study showed an accuracy of 88%, where ease of access and time efficiency were the main causes of high dependence (J. Sistem et al., 2023). Similar results were obtained by Rahmawati & Fahmi (2024), who found that the Naïve Bayes model was superior to Logistic Regression with an accuracy of 89% (Mehta et al., 2023). Research by Utami & Pratama (2024) used Multinomial Naïve Bayes to analyze student sentiment towards the use of ChatGPT. From 500 social media and survey data, 76% were classified as positive, 14% as neutral, and 10% as negative, with a total accuracy of 82% (Iskandar et al., 2025). These results show a positive trend among students towards the use of ChatGPT, although there are still concerns about the originality of their work. On the other hand, Putri & Santoso (2023) utilized Naïve Bayes to classify the behavior of AI-based chatbot usage in education, resulting in an accuracy of 82%. This model proved effective in recognizing students' habits in using learning chatbots, including ChatGPT, as digital academic assistants (Ijaz & Yadegaridehkordi, 2024).

Research by Sun & Liu (2025) applying the Naïve Bayes algorithm to analyze social regulation in conversation-based collaborative learning showed an accuracy rate of 87% (Latief et al., 2024),

indicating the model's ability to effectively capture patterns of interaction and student learning behavior. Similar results were also shown by Abdelhamid & Aly (2025), who used Naïve Bayes in the design and evaluation of artificial intelligence-based STEM projects with an accuracy of 84%, confirming that this probabilistic approach is relevant to support AI-based learning processes. Additionally, English & Jagha (2025) utilized the same model to predict learning depth through online dialogue and obtained an accuracy of 83%, demonstrating the consistency of Naïve Bayes performance in the context of text-based interaction analysis. Overall, these three studies show that the Naïve Bayes algorithm has high flexibility and effectiveness in various learning contexts involving AI systems, including conversation-based platforms such as ChatGPT (Nhu et al., 2024).

Based on all of these studies, it can be concluded that the Naïve Bayes method consistently provides accuracy above 80% in various digital education contexts. However, most studies still focus on analyzing general behavior and sentiment, without specifically measuring the level of student dependence on ChatGPT quantitatively (Amijoyo & Mumtaza, 2023). Therefore, this study attempts to fill this gap by developing a Naïve Bayes-based predictive model to measure the level of student dependence on ChatGPT in completing academic assignments (Klimova et al., 2025). Most previous studies have only discussed students' perceptions of ChatGPT usage, without examining the level of dependence on it in completing academic assignments. In fact, excessive use can reduce students' independent thinking and analytical skills. This gap is important to study in order to determine the extent to which dependence on ChatGPT affects the learning process (Atieku-boateng et al., 2024) [21]. This situation creates an urgent need to identify patterns of student dependence more objectively. Without proper understanding, ChatGPT can encourage instant gratification and reduce the quality of learning. Therefore, the Naïve Bayes method was chosen because it can efficiently predict and classify the level of student dependence based on behavioral data patterns. This approach supports the creation of more balanced learning and encourages students to continue thinking independently (Hasanein & Sobaih, 2023).

This study aims to analyze and predict the level of student dependence on ChatGPT in completing academic assignments. Through a data-based approach, this study seeks to identify patterns of student behavior in utilizing ChatGPT, as well as factors that influence the level of dependence. In addition, this study also aims to provide an empirical description of how the use of ChatGPT affects students' learning independence and critical thinking processes. To achieve these objectives, this study proposes the application of the Naïve Bayes method as the main classification technique due to its ability to process data efficiently and produce accurate predictions. Through this method, it is hoped that this research can make a real contribution in the form of a prediction model that can be used to monitor and control the level of student dependence on ChatGPT. The results of this study are expected to form the basis for educational institutions in formulating learning policies that balance the use of AI technology and the strengthening of students' independent thinking skills.

METHOD

This study uses a quantitative survey to examine student dependence on ChatGPT for academic tasks, with the research stages outlined in Figure 1 from problem formulation to result interpretation.

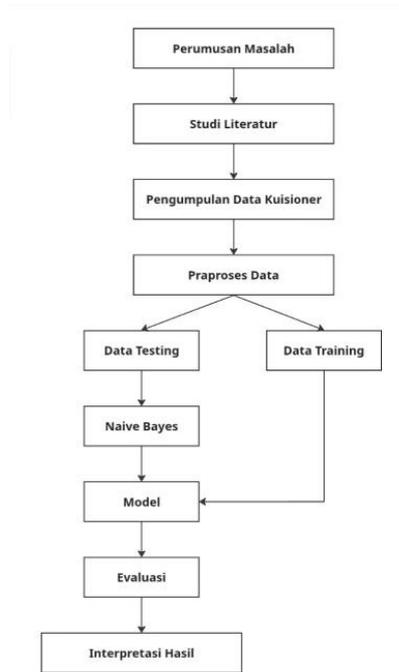


Figure 1. Research Methodology

To ensure that the research was conducted systematically, it was carried out in several stages as follows:

Data Collection

Data were collected via an online Likert-scale questionnaire targeting active students of the Informatics and Computer Education Study Programme, selected through purposive sampling (cohorts 2022–2024, prior ChatGPT use, complete responses). The instrument assessed five components purpose, frequency/duration, understanding, trust, and learning independence to capture motivation, usage patterns, and impact on academic autonomy. Responses were numerically coded for analysis using Naïve Bayes to classify students’ dependence on ChatGPT.

Data Pre-processing

After data collection, pre-processing ensured a clean, structured dataset for Naïve Bayes. Incomplete responses and irrelevant attributes were removed, Likert-scale answers were numerically coded (1–5), and total scores were labelled into low (6–14), medium (15–23), and high (24–31) dependence using a simple spreadsheet formula (Sta et al., 2020).

$$\text{Label} = \begin{cases} \text{Low} & \text{If } S < 14 \\ \text{High} & \text{If } S > 24 \\ \text{Medium} & \text{If } 14 \leq S \leq 24 \end{cases}$$

With $S = \sum_{i=1}^n x_i$, which is the total score of all relevant questions. An example in Excel is written as follows:

```

    =IF(SUM(C3:G3)+B3<14;"Low";IF(SUM(C3:G3)+B3>24;"High";"Medium")),
    
```

This method automatically groups respondents based on their total scores using an equal-width binning approach, in which continuous value ranges are divided into equal-width intervals to form discrete categories. This technique facilitates classification and improves interpretability, thereby providing a strong methodological basis for categorising the level of student dependence on ChatGPT (Xia et al., 2015).

Data Testing

Test data was taken from the part of the dataset that was not used in training to measure the model's ability to predict new data. This data consisted of student questionnaire scores that were processed using Naïve Bayes to generate dependency label predictions, while also calculating accuracy, precision, and recall (Infotech et al., 2025). The evaluation is conducted by comparing the prediction results with the actual labelling results (Rajagukguk et al., 2024), thereby ensuring that the model is not only optimal on the training data but also has good generalisation capabilities in predicting the level of student dependency on ChatGPT.

Data Training

The data training stage uses most of the preprocessed and labeled questionnaire data to train the Naïve Bayes model to recognize patterns of student dependence on ChatGPT (Ramadhani et al., 2024). The dataset is separated into features in the form of ChatGPT usage indicator scores (frequency, duration, understanding, trust, and learning independence) and target labels (low, medium, high). Through this process, the model learns the relationship between features and labels, forming a classification pattern that forms the basis for accurate and reliable predictions (Rifaldi et al., 2025).

Naïve Bayes

The Naïve Bayes algorithm is a simple yet effective probabilistic classification method. Its main principle is to calculate the probability of data falling into a certain class based on the distribution of its features (Tohir et al., 2024). The basic Naïve Bayes formula is derived from Bayes' theorem with the assumption of independence between features, which is expressed as:

$$P(C | X) = \frac{P(C) \cdot \prod_{i=1}^n P(x_i | C)}{P(X)}$$

$P(C|X)$ represents the probability of class C given data X , with $P(C)$ as the prior and $P(x_i|C)$ as the feature likelihood. Naïve Bayes assumes feature independence, simplifying calculations for large datasets. Student questionnaire data were converted into numerical features purpose, frequency, understanding, confidence, and learning independence and mapped to dependency labels (low, medium, high). The dataset was split into training and testing sets to build and evaluate model accuracy and generalization (Mohi, 2020). Naïve Bayes provides interpretable probabilistic predictions, and prior studies confirm its effectiveness in classifying student behavior, including ChatGPT dependence.

Model

Test and Score evaluates classification model performance on unseen test data to assess generalization by comparing predictions with actual labels (Mijwil & Mijwil, 2024). Typically applied via holdout validation, the dataset is split into training and test sets. In this study, Orange Data Mining implemented Test and Score, producing Confusion Matrix, AUC, and key metrics Classification Accuracy, Precision, and F1-Score following standard machine learning definitions.

$$a. \text{ Accuracy (CA)} = \frac{TP+TN}{TP+TN+FP+FN}$$

$$b. \text{ Precision} = \frac{TP}{TP+FP}$$

$$c. \text{ Recall} = \frac{TP}{TP+FN}$$

$$d. \text{ F1 Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

e. AUC is calculated from the area under the ROC curve that compares the True Positive Rate and False Positive Rate.

A high AUC reflects strong class distinction, while high F1 Score and Precision indicate accurate and consistent predictions. Using Test and Score ensures the Naïve Bayes model effectively classifies both training and test data, providing a robust, objective basis for evaluating student dependence on ChatGPT.

Evaluation

The Confusion Matrix is used to describe the performance of a classification model by comparing the prediction results with the actual labels. This matrix consists of four main components, namely True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). Through these components, researchers can determine the number of correct and incorrect predictions in each class, allowing for a more detailed evaluation of the model.

From the Confusion Matrix, the two main metrics used in this study are Accuracy and Precision, with the following formulas:

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

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

Accuracy indicates the overall correctness of predictions, while Precision assesses how well positive predictions match actual labels, providing a more comprehensive evaluation of model performance.

Interpretation of Results

Accuracy across data splits was analyzed to identify the most effective model, while Orange's Rank feature assessed each feature's contribution, revealing key factors and learning sources that influence student learning category predictions.

RESULTS AND DISCUSSION

Data set

This study began with data collection through a closed questionnaire distributed to active students of the Information Technology and Computer Science Study Program at Makassar State University. This questionnaire was designed to measure various aspects of ChatGPT usage in an academic context, including the purpose of use, frequency and duration of interaction, understanding of the results provided, level of trust, and learning independence. From this collection stage, complete responses were obtained from 240 students, which then became the initial dataset for the study. This initial data was summarized and presented in Table 1 below, as the basis for the preprocessing, analysis, and development of the Naïve Bayes classification model. This data provides an initial overview of ChatGPT usage patterns among students and serves as the starting point for all subsequent stages of analysis.

Table 1. Pre-processed Dataset

Initials	Major	Generation	...	I feel that the use of ChatGPT affects ...	I feel more dependent on ChatGPT....	I find it difficult to complete the task...
MIS	PTIK	2022	...	Agree	Agree	Agree
FN	PTIK	2022	...	Agree	Neutral	Agree
AS	PTIK	2023	...	Agree	Strongly agree	Agree
NR	PTIK	2022	...	Disagree	Neutral	Neutral
MM	PTIK	2024	...	Agree...	Agree...	Agree...
...
R	PTIK	2023	...	Agree	Agree	Strongly agree

IB	TEKOM	2023	...	Agree	Strongly agree	Strongly agree
F	TEKOM	2022	...	Disagree	Neutral	Disagree
AG	PTIK	2022	...	Disagree	Disagree	Disagree
TAFW	TEKOM	2022	...	Disagree	Disagree	Disagree

The pre-processed dataset includes comprehensive demographic information on participating students, including their initials, study programs, and enrollment batches, along with their responses to a series of statements specifically designed to capture various aspects of ChatGPT usage in an academic context. These responses were recorded using a Likert scale ranging from "Strongly Agree," "Agree," "Neutral," to "Disagree," which served to measure students' perceptions, experiences, and behavioral tendencies regarding the influence of ChatGPT on their learning process. These initial responses provided a deeper understanding of how students interact with AI tools in an academic environment and formed the basis for further analytical procedures.

After data collection, the next important stage involved manual data pre-processing, which was carried out carefully to ensure that the dataset was clean, well-structured, and suitable for advanced classification modeling. This process includes several key steps: first, irrelevant or non-contributing variables such as student initials, program identifiers, or other identifiers are removed to prevent noise and ensure that the dataset focuses only on features relevant to measuring student dependence on ChatGPT. Second, all responses obtained through the Likert scale are systematically converted into numerical values to facilitate computational processing by the Naïve Bayes algorithm. Third, each student's aggregate responses are used to assign categorical labels representing their level of dependence on ChatGPT, classified into low, medium, and high categories based on predetermined assessment thresholds.

This careful pre-processing stage is crucial to ensure the validity, consistency, and reliability of the data, thereby providing a solid foundation for the predictive modeling process. By preparing the dataset in this way, the Naïve Bayes algorithm is able to learn the underlying patterns in the data, accurately map feature combinations to dependency levels, and generate predictions that reflect students' actual behavioral tendencies. Ultimately, this stage ensures that the model's output is trustworthy and that the analytical conclusions drawn from the research are scientifically valid and practically relevant for understanding student interactions with AI-assisted learning tools.

Tabel 2. Post-Processing Dataset

Final GPA	Number of needs	Aspect 1	Aspect 2	Aspect 3	Aspect 4	Aspect 5	Label
3.94	4	3,25	3,5	3,5	3,5	4,2	Moderate
3.87	3	4,25	3,5	2	2,5	4,2	Moderate
3.72	1	4,75	4,5	3,5	3,5	4,6	Moderate
4	1	3,5	2,75	4	3,25	3,8	Moderate
3.00	2	1	3	3	3	3	Moderate...
...
3.50	1	1	1	3	3,75	3	Low
3.42	1	2,5	1,75	3	5	2,6	Moderate
3.90	1	4	2,25	2,5	2	2,4	Moderate
3.88	2	2	3	1	2,75	1,6	Low
3.76	1	4	3,25	2	4	3,2	Moderate

At this stage, Table 2 shows that the data has undergone a careful cleaning process to ensure its relevance and consistency with the research objectives. Attributes that do not contribute to the classification analysis, such as respondent initials, class year, and study program, were removed to prevent potential bias and reduce the complexity of the dataset. Furthermore, all respondent answers on the Likert scale were converted into numerical form with a value range of 1 to 5, enabling the Naïve Bayes algorithm to process the data effectively. This transformation covers all indicators of ChatGPT usage, including purpose of use, frequency and duration of interaction, understanding of

results, level of trust, and learning independence. The results of this preprocessing stage are then compiled into a table displaying complete numerical variables, such as final GPA, number of academic needs, five aspects of ChatGPT usage behavior, and final classification labels (low, medium, high). These labels are determined based on the combined scores of all indicators, so that each respondent is categorized according to their level of dependence on ChatGPT usage. Thus, the preprocessing stage not only prepares the dataset for model training and testing, but also provides an accurate and consistent numerical representation of student behavior in the context of using ChatGPT for academic purposes.

Result

This study utilised three dataset division scenarios for model training and testing, namely ratios of 90:10, 80:20, and 70:30, with the aim of evaluating the stability and consistency of the Naïve Bayes classification model's performance in predicting the level of student dependence on ChatGPT in completing academic assignments. This division was carried out to understand the influence of the proportion of training data to test data on the prediction accuracy and generalisation ability of the model. Thus, the 90:10 scenario provides most of the data for training and a small amount for testing, the 80:20 scenario offers a balance between training data and test data, and the 70:30 scenario increases the proportion of test data to evaluate generalisation more strictly. By comparing these three scenarios, the study can assess the model's effectiveness in recognising patterns of student dependency, determine the optimal proportion of training data to produce accurate predictions, and gain insights into the model's sensitivity to dataset variations in the context of classifying student dependency on ChatGPT in academic activities.

1. Model evaluation was conducted using several key performance metrics, including Area Under Curve (AUC), which indicates the model's ability to distinguish between classes; Classification Accuracy (CA), which assesses the overall proportion of correct predictions; F1-Score, which balances precision and recall; and Precision, which measures the accuracy of the model's positive predictions. This combination of metrics allows researchers to evaluate the model comprehensively, not only in terms of overall accuracy, but also the quality of predictions for each dependency category.
2. The evaluation results of the three data partitioning scenarios are presented in Table 3, which shows a comparison of the AUC, CA, F1-Score, and Precision values for each scenario. With this analysis, it is possible to determine the data division scenario that provides the most optimal performance while also providing insight into the model's sensitivity to the proportion of training and test data. This comprehensive assessment is important to ensure that the developed model is not only capable of recognising patterns in the training data, but can also be relied upon to predict student dependence on ChatGPT in new data.

Tabel 3. Results of Naïve Bayes Model Evaluation Based on Data Division

Data Division (Test data: Training data)	AUC	Accuracy (CA)	F1-Score	Precision
90:10	0.955	0.760	0.778	0.862
80:20	0.954	0.820	0.837	0.890
70:30	0.973	0.853	0.866	0.909

Based on the results of the Naïve Bayes model evaluation in three data distribution scenarios, namely 90:10, 80:20, and 70:30, it can be seen that the best performance was shown in the 70:30 data distribution. In this scenario, the model obtained an AUC value of 0.973, an accuracy of 85.3%, an F1-Score of 0.866, and a Precision of 0.909, which are the highest values compared to the other two scenarios. The high AUC value indicates that the model is able to distinguish classes very well, while the high accuracy, F1-Score, and Precision indicate that the model is not only able to recognise classes correctly but also consistent in producing relevant predictions. These findings suggest that a proportion of 70% training data and 30% test data provides an optimal balance for the model in learning data patterns while maintaining good generalisation capabilities for new data.

In the 80:20 split scenario, the model's performance is still relatively good with an AUC value of 0.954, accuracy of 0.820, F1-Score of 0.836, and Precision of 0.890. However, all of these metrics are still below the results in the 70:30 data split. This indicates that even though the proportion of training data is quite large, adding a larger amount of test data, as in the 70:30 scenario, provides a more stable evaluation and allows the model to capture data pattern variations better. Conversely, the 90:10 scenario produced the lowest performance, with an accuracy of 0.760 and an F1-Score of 0.778. The very small proportion of test data in this scenario resulted in less representative evaluation and reduced the model's ability to measure overall prediction performance.

Overall, the evaluation results show that a 70:30 data split is the most ideal ratio for this study. This ratio provides a large amount of training data for the model to learn the characteristics of the data, while also providing sufficient test data to produce stable and accurate performance evaluations. Thus, the Naïve Bayes model in this study has proven to have good ability in predicting the level of student dependence on ChatGPT, especially when using a balanced data distribution ratio such as 70:30. In addition to quantitative evaluation through AUC, accuracy, F1-Score, and Precision metrics, the performance analysis of the Naïve Bayes model is also reinforced by the interpretation of the Confusion Matrix, which can be seen in Figure 2. This matrix presents the distribution of model predictions for three target classes, namely Low, Medium, and High, by comparing the actual labels and predicted labels:

		Predicted			Σ
		Rendah	Sedang	Tinggi	
Actual	Rendah	100.0 %	0.0 %	0.0 %	6
	Sedang	25.8 %	74.2 %	0.0 %	31
	Tinggi	0.0 %	8.6 %	91.4 %	35
Σ		14	26	32	72

Figure 2. Interpretation of the Confusion Matrix

Interpretation of Figure 2 indicates that the Naïve Bayes model demonstrates a highly precise classification tendency, particularly in the Low and High classes, with minimal or no misclassifications observed for these groups. This outcome suggests that the underlying data characteristics for students classified as having low or high dependence on ChatGPT are well-defined, with distinct patterns in features such as usage frequency, purpose, understanding, trust, and learning independence, which the model can readily identify. In contrast, the Medium class exhibits a higher degree of classification ambiguity, as evidenced by a 25.8% misclassification rate into the Low class. This ambiguity may arise from overlapping feature values between Medium and Low categories, as well as insufficient representation in the dataset to distinctly separate students with moderate dependence from those with minimal dependence. Such overlap highlights inherent challenges in modeling behavioral data where gradations of dependence are subtle and multidimensional.

The overall distribution of predictions further reinforces the model's robustness, showing a reasonably balanced spread with 14 predictions for Low, 26 for Medium, and 32 for High out of 72 test instances. This balance suggests that the Naïve Bayes model does not exhibit systematic bias towards any single category and is capable of maintaining proportionality in predictions that aligns with the observed distribution of actual data. It also reflects the effectiveness of the feature selection and preprocessing steps in capturing the relevant aspects of student behavior and ChatGPT usage patterns.

When considering the data split ratio, the 70:30 division emerges as the most effective for this study. This ratio provides a sufficiently large training set of 240 instances, allowing the model to learn the nuanced relationships between features and dependence levels while preserving an adequately sized test set to conduct stable and meaningful performance evaluations. The results indicate that with this split, the model achieves both high predictive accuracy and reliability in reflecting real-world student

behaviors. Overall, the findings confirm that the Naïve Bayes model is well-suited for predicting the level of student dependence on ChatGPT, particularly under conditions where training and test datasets are carefully balanced. The study demonstrates that probabilistic models like Naïve Bayes can capture subtle variations in usage patterns and provide actionable insights for understanding students' reliance on AI-based learning tools, which can inform educational policy and instructional design.

CONCLUSION

This study confirms that the Naïve Bayes algorithm can be used effectively to predict the level of student dependence on ChatGPT in completing academic assignments. The research process was carried out systematically, starting with data collection through questionnaires, followed by a pre-processing stage involving data cleaning and coding into numerical form, and finally evaluation using Test and Score in the Orange Data Mining application. The analysis results show that the 70:30 data split scenario provides the most optimal performance with an AUC value of 0.973, accuracy of 85.3%, F1 Score of 0.866, and Precision of 0.909. These findings prove that this ratio is able to maintain a balance between the model's ability to learn data patterns and its ability to generalize to new data. Thus, this study makes a real contribution to the development of a student behavior classification system and can be used as a reference for educational institutions in formulating learning policies that balance the use of AI technology and the strengthening of students' independent thinking. Further research is recommended to expand the number and diversity of samples by involving students from various study programs, faculties, or even different higher education institutions. Increasing the heterogeneity of the dataset will enhance the generalization capability of predictive models and provide insights that are more representative of the broader student population. In addition, employing comparison algorithms such as Random Forest, Support Vector Machine, or Gradient Boosting can offer a more comprehensive evaluation of model performance, enabling researchers to benchmark Naïve Bayes against other advanced machine learning approaches. Incorporating additional variables related to psychological factors, digital literacy, intensity of technology usage, and pedagogical contexts is also crucial for deepening the understanding of students' dependence on AI-based tools like ChatGPT. The adoption of more robust validation strategies, such as stratified k-fold cross-validation, is recommended to increase the reliability, stability, and reproducibility of the prediction model. Finally, designing and implementing a prototype of a classification-based system would allow for practical assessment of the model's effectiveness, enabling real-time monitoring and personalized recommendations within an educational environment. Such steps would not only strengthen the scientific rigor of future studies but also provide actionable insights for institutions aiming to balance AI utilization with the promotion of independent learning skills.

ACKNOWLEDGMENT

The author would like to express sincere gratitude to all students of the Informatics and Computer Engineering Study Program at Makassar State University who participated in this study, dedicating their time and effort to provide valuable responses through the questionnaire. Their active involvement and honest feedback formed the foundation of this research, enabling a comprehensive analysis of student dependence on ChatGPT. Special appreciation is extended to the faculty members, lecturers, and administrative staff of the Department, whose continuous support facilitated the smooth implementation of data collection, provision of necessary resources, and coordination with students. The author also gratefully acknowledges the constructive guidance, insightful suggestions, and critical feedback from colleagues, mentors, and academic peers throughout the process of research conceptualization, design, and methodological refinement. Their expertise and encouragement were instrumental in overcoming challenges encountered during data analysis and in ensuring the rigor and validity of the study. Furthermore, the author recognizes the support from the broader academic community, whose literature and prior research provided invaluable references for contextualizing and interpreting the findings. This study would not have been possible without the collaborative effort, cooperation, and goodwill of all those involved, and the author deeply appreciates their contributions to the successful completion of this research.

AUTHOR CONTRIBUTION STATEMENT

The author was solely responsible for all aspects of this research, taking full responsibility for the conceptualization and formulation of the research problem, the design and structuring of the study, as well as the development of the research methodology. This included designing the data collection instruments, determining the appropriate sampling techniques, and overseeing the administration of the questionnaire to ensure data integrity and relevance. The author also independently conducted the data processing and pre-processing, including cleaning, coding, and labelling the dataset, followed by the implementation of the Naïve Bayes classification model and evaluation of its performance through appropriate metrics. Furthermore, the author carried out the interpretation of the results, critically analyzed the findings in the context of existing literature, and derived meaningful conclusions and implications for practice and further research. All decisions regarding methodological choices, analytical procedures, and the presentation of results were made autonomously by the author. Additionally, the author was solely responsible for the preparation of the manuscript, including the writing, editing, and formatting of the text to meet academic standards, ensuring that all aspects of the study were conducted with scientific rigor, transparency, and integrity.

AI DISCLOSURE STATEMENT

Artificial intelligence tools were used exclusively for linguistic editing and readability improvement. These tools had no role in data analysis, modeling, result interpretation, or decision-making processes. The authors retain full responsibility for the manuscript content.

REFERENCES

- Amijoyo, T., & Mumtaza, A. (2023). *Implementation Naïve Bayes Method To Prediction Of Student Achievement at Bimbel Sonai West Jakarta*. 6(2), 139–144.
- Andrea, D., Putri, R., Putri, E., & Winarsih, S. (2025). *Hubungan Ketergantungan ChatGPT dengan Kemandirian Akademik Siswa SMPN 198 Jakarta*. 6(2), 100–106.
- April-september, N. N. E., Indria, L., & Hakim, M. F. (2025). *Analisis Ketergantungan Mahasiswa Universitas Merangin terhadap ChatGPT dengan Machine Learning*. 7924(September), 35–41.
- Atieku-boateng, H., Addo-tara, A., Osei, R. D., Atieku-boateng, B., & Ameko, S. K. (2024). *Evaluating ChatGPT's Effect on Students' Independent Learning Experiences: The Case of Pentecost University, a Ghanaian Higher Educational Institution*. 11, 20–40. <https://doi.org/10.63467/red.11.2024.art2>
- Cahyanto, H. N., Pamungkas, P., & Zulkarnain, O. (2024). *Pengaruh Penggunaan ChatGPT Terhadap Kemandirian Mahasiswa Dalam Menyelesaikan Tugas Akademik*. PREPOTIF J. Kesehat. Masy., 8(1), 930–935.
- Fathony, M. H., Hizraini, A. A., Aulia, R., & Penggunaan, E. (2024). *Jurnal Review Pendidikan dan Pengajaran, Volume 7 Nomor 4, 2024 | 16601*. 7, 16601–16607.
- Hasanein, A. M., & Sobaih, A. E. E. (2023). *Drivers and Consequences of ChatGPT Use in Higher Education: Key Stakeholder Perspectives*. 2599–2614.
- Ijaz, M., & Yadegaridehkordi, E. (2024). *ChatGPT in the higher education: A systematic literature review and research challenges*. *International Journal of Educational Research*, 127, 102411. <https://doi.org/10.1016/j.ijer.2024.102411>
- Infortech, J., Wahyudi, N., Putri, D. A., Bina, U., & Informatika, S. (2025). *Implementasi algoritma naïve bayes terhadap*. 2(16).
- Iskandar, A., Retnawati, H., Hadi, S., Utami, L., & Amiruddin, E. G. (2025). *Ingénierie des Systèmes d'Information Naïve Bayes Classifier-Based Intelligent System for Student Academic Performance Assessment*. 30(6), 1609–1620.
- J. Sistem et al. (2023). *Analisis Sentimen Chat GPT sebagai Masa Depan Pekerja pada Media Sosial Youtube menggunakan Algoritma Klasifikasi Naïve Bayes*. 14(2), 193–200.
- Klimova, B., Bachmann, P., & Frutos-bencze, D. (2025). *The use of ChatGPT in academia: Perspectives of higher education students*. *Cogent Education*, 12(1). <https://doi.org/10.1080/2331186X.2025.2508216>
- Kurniasari, P., Mardikaningsih, A., & Sari, R. S. (2025). *Dependensi Penggunaan Kecerdasan Buatan AI (Artificial Intelligence) Terhadap Tugas Akademik Mahasiswa*. 4(3).

- Latief, S. A., Muhsin, M. A., & Hasanah, N. I. (2024). *Detection of AI Utilization in Language Learning: A Study on the Implementation of Chat Gpt In Academic Writing Among Indonesian Higher Education Students*. 13(4), 1–10.
- Maula, S. R., Aprillian, S. D., & Rachman, A. W. (2024). *Ketergantungan Mahasiswa Universitas Jember Terhadap Artificial Intelligence (AI)*. 2(1).
- Mehta, S., Info, A., Modeling, P., Graduation, S., Algorithm, N. B., & Data, E. (2023). *Playing Smart with Numbers: Predicting Student Graduation Using the Magic of Naïve Bayes*. 2(1), 60–75.
- Mijwil, M. M., & Mijwil, M. M. (2024). *A Comparative Analysis of Machine Learning Algorithms for Classification of Diabetes Utilizing Confusion Matrix Analysis*. 21(5).
- Mohi, Z. R. (2020). *Orange Data Mining as a tool to compare Classification Algorithms*. 3, 13–23.
- Munggaran, J. P., Alhafidz, A. A., Taqy, M., Aprianti, D., Agustini, R., & Munawir, M. (2023). Sentiment Analysis of Twitter User Opinions Data Regarding the Use of ChatGPT in Education. *Journal of Computer Engineering, Electronics and Information Technology (COELITE)*, 2(2), 75–88.
- Nakhipova, V., Kerimbekov, Y., Umarova, Z., Suleimenova, L., & Botayeva, S. (2024). *Use of the Naïve Bayes Classifier Algorithm in Machine Learning for Student Performance Prediction*. 14(1). <https://doi.org/10.18178/ijiet.2024.14.1.2028>
- Nhu, T., Nguyen, T., Lai, N. V., & Nguyen, Q. T. (2024). *Artificial Intelligence (AI) in Education: A Case Study on ChatGPT's Influence on Student Learning Behaviors*. 0901(2), 105–121.
- Purbayanto, B., & Suharsono, T. N. (n.d.). *Analisis Sentimen Pengguna X terhadap ChatGPT dengan Algoritme Naïve Bayes*. 18(2), 63–71.
- Rajagukguk, N., Eka, I. P., & Kencana, N. (2024). Application of Term Frequency—Inverse Document Frequency in The Naïve Bayes Algorithm For ChatGPT User Sentiment Analysis. *ICAMSAC 2023*. <https://doi.org/10.2991/978-94-6463-413-6>
- Ramadhani, T., Hermawan, P., & Dzikrillah, A. R. (2024). *Penerapan Metode Naïve Bayes untuk Analisis Sentimen pada Ulasan Pengguna Aplikasi ChatGPT di Google Play Store*. 6(1), 430–439. <https://doi.org/10.47065/bits.v6i1.5400>
- Rifaldi, D., Famuji, T. S., Okta, B., Miranda, S., & Purma, F. (2025). *Evaluasi Sentimen Pengguna ChatGPT Menggunakan Naïve Bayes: Tinjauan dari Confusion Matrix dan Classification Report*. 3(2), 81–89.
- Search, I. (2023). *Penggunaan ChatGPT dalam Pembelajaran: Persepsi Mahasiswa*. 22(1), 254–259.
- Setiabudi, D. H., & Sampurno, A. R. (n.d.). *Chatbot for Complex Questions in University Admission using Bidirectional Long-Short Term Memory and Convolutional Neural Network*.
- Sta, U., Id, B. Z., & Id, G. B. (2020). *Discretisation of conditions in decision rules induced for continuous data*. <https://doi.org/10.1371/journal.pone.0231788>
- Tohir, F. A., Irawan, B., & Bahtiar, A. (2024). *Analisis Sentimen Aplikasi ChatGPT Mobile Menggunakan Algoritma Naïve Bayes*. 10(2), 179–192.
- Triharningsari, D., Widayauri, A., Putri, M. A., & Fatihin, A. (2025). *Sentiment Analysis of ChatGPT Exploration Based on Opinions on Platform X Using Naïve Bayes Algorithm*. 4(July 2024), 94–101.
- Xia, Y., Hou, C., Xu, Z., Sun, J., Space, G., & Place, P. C. (2015). *Feature Reduction Using Ensemble Approach*. 309–318.