Research Article

Prediction Analysis of Package C Student Graduation at the *Bollo DMansel* Community Learning Activity Center (*PKBM*) with the Naïve Bayes Algorithm Method

[1*]Muhammad Yassir, [1]Wanda Cahyani

1 Department of Information Systems and Technology, Institut Teknologi dan Bisnis Nobel Indonesia

* Corresponding Author: myassir669@gmail.com

Abstract: Education plays a crucial role in improving the quality of human resources and is the key to a nation's progress. *Bollo DMansel* Community Learning Activity Center *(PKBM*) in West Papua provides a Paket C Equivalency Education program to help those who are underserved by formal education. The main challenge in this program is to increase student graduation rates. This research aims to analyze and predict the graduation of Paket C students at *PKBM Bollo DMansel* using the Naive Bayes algorithm method. The data used includes historical student data from 2021 to 2023, with a total of 128 students. The research steps include data collection, data pre-processing, Naive Bayes algorithm application, and prediction model evaluation. The results show that the Naive Bayes algorithm can provide graduation prediction with fairly high accuracy. The factors that most influence student graduation were identified, including attendance, test scores, and participation in activities. This research makes a real contribution to improving the quality of education at *PKBM Bollo DMansel* by providing a prediction tool to identify students at risk of not graduating so that timely intervention can be provided.

Keywords: Graduation Prediction; *PKBM Bollo Dmansel*; Naive Bayes; Data Mining; RapidMiner.

Citation: M.Yassir, W.Cahyani, " Prediction Analysis of Package C Student Graduation *at the Bollo DMansel* Community Learning Activity Center (*PKBM*) with the Naive Bayes Algorithm Method", *Iota*, **2024**, ISSN 2774-4353, Vol.04, 03. https://doi.org/10.31763/iota.v4i3.751

Academic Editor : Adi, P.D.P

Received : July, 22 2024

Accepted : July, 29 2024

Published : August, 06 2024

1. Introduction

Education has a crucial role in improving the quality of human resources and is the key to the progress of a nation. Education is a conscious and planned effort to create an active process of self-potential development, where learners are expected to gain skills, noble morals, intelligence, self-control, and spiritual strength that are beneficial to themselves and society [1]. In the Indonesian context, education is regulated by the 1945 Constitution Chapter XIII Article 31 paragraph (1) which states that every citizen has the right to receive teaching, and paragraph (2) which states that the government seeks and organizes a national teaching system regulated by law [2]. Informal education or out-of-school education is one solution to reach people who are not served by formal education. One form of informal education is the Community Learning Activity Center (PKBM). *PKBM* offers various programs, including Equivalency Education Package C, which is equivalent to upper secondary education. The program is designed to help those who have dropped out of school, failed the national exam, or do not have access to formal education, to get an education equivalent to High School or Vocational School [3].

PKBM Bollo DMansel is one of the informal education institutions located on Merdeka Street, Sidomulyo Village, Oransbari Sub-district, South Manokwari District, West Papua Province. This *PKBM* specifically offers the Equivalency Education Package C program. *PKBM Bollo DMansel* plays an essential role in improving the education level of the local community by providing learning opportunities for those who are not reached by the formal education system.

Ensuring student graduation in the Package C program is essential, as this graduation not only determines the students' future but also affects the quality of education at the *PKBM* itself. Predicting student graduation can assist educational institutions in identifying students who are at risk of not graduating so that additional assistance can be provided before the exam. By predicting graduation, *PKBM* can plan more appropriate and effective interventions, increase graduation rates, and ultimately improve the quality of education provided [4].

This research uses the Naive Bayes method [1,2,3,4,5] which is one of the techniques or methods of data mining that can be used for the classification process using probabilities and statistics that can predict future opportunities based on previous experience [5]. This algorithm uses probabilities and statistics to predict outcomes based on historical data. In the context of student graduation prediction, the Naive Bayes algorithm can analyze test score data and other factors to predict whether a student will graduate or not. This method was chosen because of its ability to handle large data and provide accurate prediction results with relatively fast computation time. Based on this background, the problems formulated in this study are how to apply the Naive Bayes algorithm to predict the graduation of Package C students at *PKBM Bollo DMansel*, what factors affect student graduation at *PKBM Bollo DMansel*, and how accurate the prediction of student graduation using the Naive Bayes algorithm at *PKBM Bollo DMansel*. The objectives of this study are to apply the Naive Bayes algorithm to predict the graduation of Package C students at PKBM Bollo DMansel, identify the factors that influence student graduation at *PKBM Bollo DMansel*, and assess the accuracy of predicting student graduation using the Naive Bayes algorithm [6,7,8,9,10] at *PKBM Bollo DMansel*.

2. Theory

2.1 Quality Education

Quality Education Education is the main foundation for improving the quality of human resources that play an important role in the progress of a nation. Effective education not only provides knowledge and skills but also shapes character, ethics, and positive attitudes. According to recent studies, quality education has a significant impact on reducing poverty and improving social welfare [6], [7]. Education also plays an important role in shaping individuals' adaptability to social and technological change.

2.2 Equality Education

Equality Education Equality education through non-formal programs such as those offered by the Community Learning Activity Center (*PKBM*) provides opportunities for individuals who are not served by formal education. The Package C Equivalency Education program, for example, aims to provide education equivalent to upper secondary level for those who have dropped out of school or do not have access to formal education [8]. Challenges in this program include improving student completion rates, which are influenced by various factors such as attendance, participation in activities, and test scores. Identifying and understanding these factors is crucial to developing effective intervention strategies.

2.3 Naïve Bayes Algorithm

The Naive Bayes algorithm is a simple but effective probability-based classification method, which is based on Bayes' Theorem with the assumption of independence between variables. This algorithm has been widely used in various applications, including text classification, medical diagnosis, and prediction of educational outcomes, due to its ability to provide accurate results even with strong independence assumptions [11,12,13,14,15].

2.4 Application of Naive Bayes

This research applies the Naive Bayes algorithm to analyze and predict the graduation of Package C students at *PKBM Bollo DMansel*. The research process includes collecting historical student data from 2021 to 2023, pre-processing the data, applying the Naive Bayes algorithm, and evaluating the prediction model. By

identifying the factors that most influence student graduation, the model is expected to help identify students who are at risk of not graduating, allowing for timely intervention. The results show that the Naive Bayes algorithm [16,17,18,19,20] can provide predictions with fairly high accuracy, which can be used to improve the quality of education at *PKBM Bollo DMansel*.

3. Method

3.1 Analysis Data

This research involves several main stages, starting with data collection from PKBM Bollo DMansel's database containing demographic information, test scores, and student attendance. The next stage is data pre-processing, which includes data cleaning, missing data handling, and data normalization. Missing data was imputed using the mean imputation method for continuous variables and mode imputation for categorical variables. Furthermore, data exploration was conducted to understand the distribution of data and relationships between variables. This analysis used graphical visualization and descriptive statistics to provide a clear picture of the analyzed data.

Figure 1. Flowchart of Research Method Stages

3.2 System Architecture

The architecture of the student graduation prediction system with the Naive Bayes algorithm consists of several main components as follows:

3.2.1 Data Input

Student data is obtained from the *PKBM Bollo DMansel* database, including demographic variables (age, gender, socio-economic background), test scores (daily test scores, midterm exams, final semester exams), and attendance (percentage of attendance, number of unexcused absences). This data was collected for three years (2021-2023), with 30 students, 40 students, and 58 students per year, respectively. Data collection was comprehensive to ensure the accuracy and completeness of the information used in the analysis.

3.2.2 Data Pre-Processing

Data Pre-Processing The collected data is then further processed to ensure its quality and suitability for predictive analysis. The pre-processing stage includes cleaning the data by identifying and removing outliers and dealing with missing data. Missing data was imputed using the mean imputation method for continuous variables such as test scores and attendance, and mode imputation for categorical variables such as gender. Furthermore, the data is normalized to ensure that all variables are on the same scale, which is important for improving the performance of the Naive Bayes algorithm.

3.2.3 Modelling

The modeling process is performed using the Naive Bayes algorithm, which is known to be effective for classification based on conditional probabilities. This algorithm was chosen due to its ability to handle datasets with continuous and categorical variables and its efficiency in processing large data. The model was built using training data covering 80% of the total dataset, while the remaining 20% was used as test data. The process involved dividing the dataset, building the model by calculating the probability of passing based on the attributes, and classifying the students into pass or fail categories.

3.2.4 Model Evaluation

Model performance is evaluated using various metrics such as accuracy, precision, recall, and F1-score. Accuracy measures the proportion of correct predictions out of all predictions made by the model [12]. Precision calculates the proportion of positive predictions that are positive, while recall measures the proportion of positive events that are detected by the model. F1-score is the harmonic mean of precision and recall, which provides a balanced picture of the model's performance, especially in handling imbalanced data.

3.2.5 Prediction

Prediction The evaluated and refined model is then applied to predict the graduation of new students. This process involves using the model to predict graduation status based on new input data that includes demographic variables, test scores, and attendance. This prediction helps in identifying students who are at risk of not graduating so that appropriate interventions can be made early.

3.2.6 Output

The results of prediction and analysis of factors affecting graduation are displayed in the form of a comprehensive report. This report includes visualization of prediction results, analysis of variable distribution, and identification of key factors affecting student graduation. This output is useful for *PKBM Bollo DMansel* managers to formulate more effective education strategies and improve student graduation rates.

3.3 Naive Bayes Method

Naive Bayes is one of the methods in machine learning that uses the principles of probability calculation and statistics proposed by an English scientist named Thomas Bayes. This method is used to estimate the probability of a future event based on experience from previous events [21,22,23,24,25]. In particular, Naive Bayes is often used in document classification and text classification. The main principle is to use Bayes' theorem to calculate the posterior probability of a given class of data, based on the prior probability and likelihood of the features. Naive Bayes assumes independence between features.

$$P\ (C|X) = \frac{P\ (X|C).P(C)}{P(X)} \quad (1)$$

$P\ (C|X)$ is the probability of a class C features X.
$P\ (C|X)$ is the probability of the feature X given class C.
$P\ (C)$ is the a priori probability of the class C.
$P\ (X)$ is the a priori probability of the feature X.

3.4 Implementation with RapidMiner

RapidMiner is open-source software. RapidMiner can be said to be a solution for analyzing data mining, text mining, and predictive analysis. RapidMiner also uses various descriptive and predictive techniques to provide insight to users so that they can make the best decisions [14]. In addition, RapidMiner is also the right choice because it is easy to use. The implementation steps include: 1) Data Input: Entering student data into RapidMiner. 2) Data Pre-Processing: Performing data cleaning, handling missing values, and data normalization using RapidMiner's pre-processing module. 3) Model Building: Using the Naive Bayes module to build a graduation prediction model. 4) Model Evaluation: Using RapidMiner's evaluation module to measure model performance with accuracy, precision, recall, and F1-score metrics. 5) Prediction: Apply the model that has been built to predict the graduation of new students. RapidMiner makes this whole process easy by providing an intuitive user interface and a variety of ready-made modules that can be used without having to write code manually.

4. Result and Discussion

4.1 Research Results

This research uses a dataset of PKBM Bollo DMansel students from 2021 to 2023, which consists of 128 students. The data used includes demographic information, test scores, and student attendance. After going through the data pre-processing stage, an analysis is carried out using the Naive Bayes algorithm to predict student graduation.

4.2 Raw Data Description

Raw data is initial data or original data that has not been processed or analyzed further. This raw data is collected from its source without any additional modification or manipulation. Raw Data Collection is also known and referred to as Primary Data, then the use of raw data that has been collected from trusted sources of PKBM Bollo DMansel administrators is the first step in this research process. The raw data obtained consisted of historical student data that included demographic information, test scores, attendance, and participation in activities. Table 1 shows an example of the raw data display used in this research.

Table 1. Raw data of *PKBM Bollo DMansel*

NIK	Student Name	TTL	Gender	UH score	Mid-term	Final exam	Attendance	Graduation
9202294107xxxxxx	Matarenia Saroi	Masabui, 01 Juli 1998	P	85	80	80	90	L
9211014306xxxxxx	Anti Induwek	Muari, 31 Maret 1998	P	70	75	78	85	L
9202291201xxxxxx	Demianus Mandacan	Inggemboisi, 09 Juli 1988	L	60	65	70	80	L
9202292509xxxxxx	Ronald Ruben Mansumber	Oransbari, 25 September 1982	L	90	75	75	75	L
9211026512xxxxxx	Natalia Kristiani Mansumber	Ransiki, 25 Desember 1994	P	65	70	72	88	L
9202296304xxxxxx	Elsina Yosina Lince Akwan	Oransbari, 23 April 1990	P	70	75	80	90	L
9202290103xxxxxx	Lamek Saroi	Masabui, 01 April 2001	L	60	75	70	80	L
9202291205xxxxxx	Beni Saroi	Masabui, 12 Mei 1964	L	55	60	65	78	TL
9202294206xxxxxx	Supi Saroi	Masabui, 27 Juni 1990	L	58	62	68	82	TL
9202290107xxxxxx	Dominggus Indou	Ikum, 01 Juli 1974	L	75	80	82	90	L
9106012711xxxxxx	Abraham Sayori	Masabui, 01 Juli 1998	L	80	82	85	88	L
…….	……..	……	…	…	…	…	…	……

This data includes several important attributes for analysis which include information such as student's full name, place and date of birth (TTL), student's gender (Jenkel), average daily exam score (UH Score), mid-semester exam score (UTS Score), final semester exam score (UAS Score), percentage of student attendance during the study period (Attendance %), and student graduation status (Graduation Status). These attributes provide a comprehensive picture of the academic and attendance profile of students at *PKBM Bollo DMansel* from 2021 to 2023.

Data Pre-Processing Data pre-processing is an important stage in text analysis where all comments in the raw data are processed to clean the text from elements that are not significant for the classification to be performed [5]. The data pre-processing stage includes several important steps to ensure the quality of the data used in the analysis: 1) Data Cleaning: Identifying and removing inconsistent or irrelevant data 2) Missing Data Handling: Filling in missing values using mean imputation for continuous variables such as test scores and attendance, and mode imputation for categorical

variables such as gender and activity participation. 3) Data Normalization: Aligning the scales of all variables to improve the performance of the Naïve Bayes algorithm.

4.3 Application of Naïve Bayes

Application of Naïve Bayes Algorithm with RapidMiner To apply the Naive Bayes algorithm, this research uses the RapidMiner data mining tool. The stages that can be carried out are as follows: 1) Importing Datasets: - Student datasets are imported into RapidMiner from CSV files [10]. The student dataset import process is shown in Figure 2.

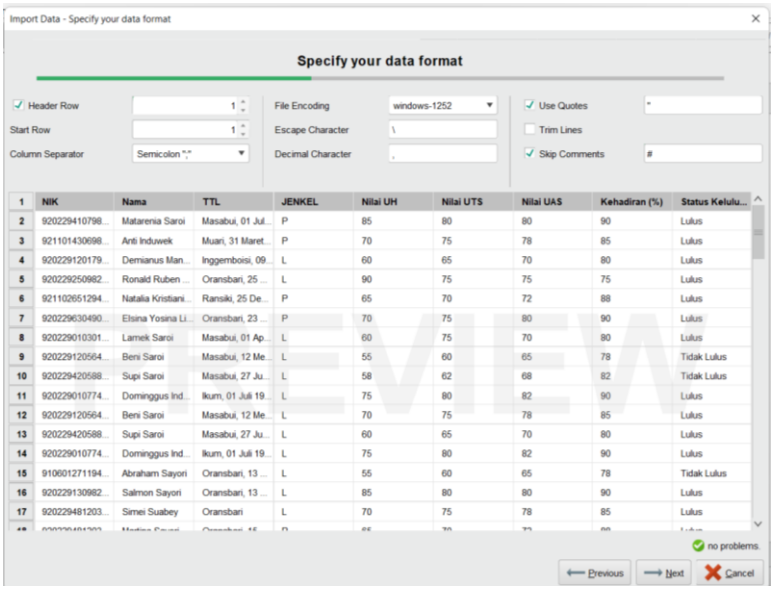

Figure 2. Import student dataset

4.4 Data Pre-processing

Data Pre-processing: - Used the 'Replace Missing Values' operator to address missing data - The 'Normalize' operator was used to normalize test scores and attendance. 3) Dataset Split: - The dataset was split into two parts: 80% for training and 20% for testing using the 'Split Data' operator. Testing using the split data operator is shown in Figure 3.

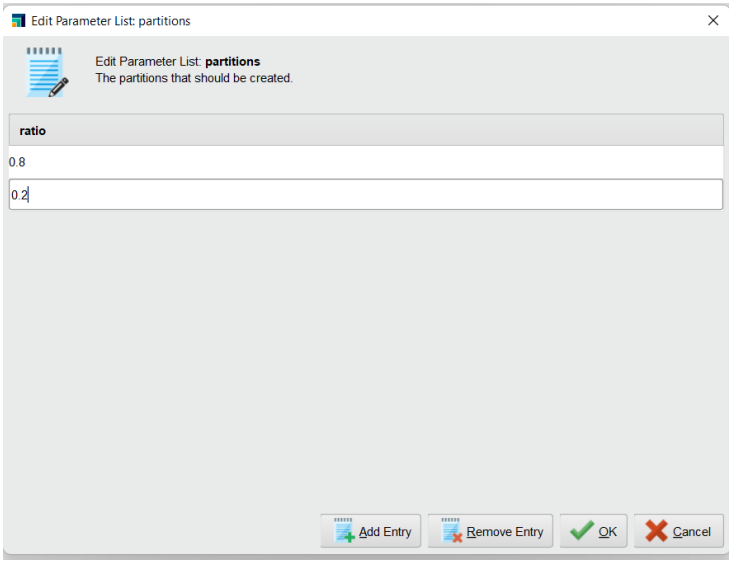

Figure 3. Testing using the split data operator

4.5 Model Training

Model Training: - The 'Naive Bayes' operator is applied to the training dataset to build a prediction model. 5) Model Evaluation: - The trained model is tested using the testing dataset. - The 'Apply model' operator is used to apply the model to the testing data. - The 'Performance (Classification)' operator is used to evaluate the performance of the model with metrics such as accuracy, precision, recall, and F1-score [15]. Model Evaluation is shown in Figure 4.

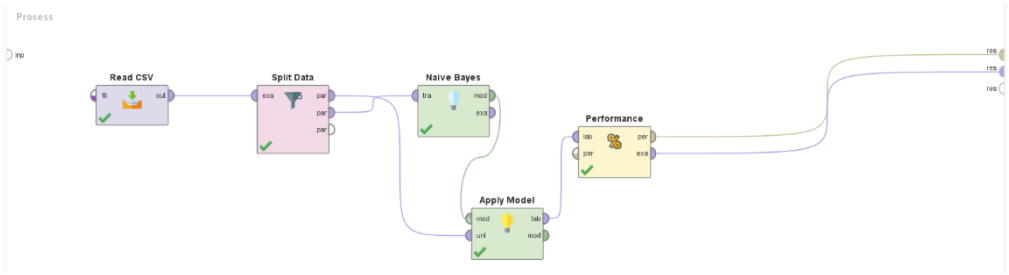

Figure 4. Model Evaluation

Model Evaluation Results After the Naive Bayes model is applied, the model performance evaluation results are shown in Table 2 and Figure 5.

Table 2. Model Performance Evaluation

Student Name	UH score	Mid-term test score	Final exam score	Attendance (%)	Graduation Status	Confidence (Passed)	Confidence (Not Passed)	Prediction (Completion Status)
Matarenia Saroi	85	80	80	90	L	1,0	0,0	L
Anti Induwek	70	75	78	85	L	1,0	0,0	L
Demianus Mandacan	60	65	70	80	L	0,0	1,0	TL
Ronald Ruben Mansumber	90	75	75	75	L	1,0	0,0	L
Natalia Kristiani Mansumber	65	70	72	88	L	1,0	0,0	L
Elsina Yosina Lince Akwan	70	75	80	90	L	1,0	0,0	L
Lamek Saroi	60	75	70	80	L	0,0	1,0	L
Beni Saroi	55	60	65	78	TL	0,0	1,0	TL
Supi Saroi	58	62	68	82	TL	1,0	0,0	TL
Dominggus Indou	75	80	82	90	L	1,0	0,0	L
Abraham Sayori	80	82	85	88	L	1,0	0,0	L
……..	…	…	…	…	……	……..	……	……

accuracy: 86.27%

	true Lulus	true Tidak Lulus	class precision
pred. Lulus	85	4	95.51%
pred. Tidak Lulus	10	3	23.08%
class recall	89.47%	42.86%	

Figure 5. Confusion Matrix Prediction Model

From the evaluation results, the Naive Bayes model performed quite well with an accuracy of 86.37%. This accuracy reflects the percentage of correct predictions out of all predictions made by the model, which includes correct predictions for the "Pass" and "Fail" classes. This figure shows that overall, the model can correctly predict most of the students who passed and did not pass, despite the difference in performance for the two classes. The model showed high precision for the "Pass" class of 95.51%, meaning that out of all the "Pass" predictions, 95.51% were students who passed. The recall for the "Pass" class is also high, at 89.47%, indicating that out of all the students who passed, the model managed to identify 89.47% correctly. This combination of high precision and recall results in an excellent F1 score of 92.23%, indicating that the model is very effective in identifying graduating students with a very small number of errors.

In contrast, the model performance for the "Not Passed" class is much lower. Precision for the "Did Not Pass" class is only 23.08%, indicating that of all the "Did Not Pass" predictions, only 23.08% did not pass. Recall for this class is also low, at 42.86%, meaning the model only identified 42.86% of the students who did not pass. The F1 score for the "Not Passed" class of 30.01% shows that the model often incorrectly predicts students who should not have passed as passed, and also misses many students who did not pass. The imbalance between the "Pass" and "Not Pass" classes likely contributed to this disparate performance, and improving performance on the "Not Pass" class may require balancing the dataset or adding more informative features.

4.6 Analysis of Factors Affecting Graduation

Based on the analysis using the Naive Bayes algorithm, the factors that have the most influence on student graduation are identified, including final semester exam scores, attendance, and participation in activities.

Table 3. Factors Affecting Graduation

Factor	Influence (%)
End of Semester Exam Score	45
Attendance	30
Activity Participation	25

End-of-Semester Exam Grades End-of-semester exam grades have the greatest influence on student graduation. As the main indicator of a student's academic ability, these grades reflect the understanding of the material taught during a semester [16].

Attendance Student attendance is also an important factor affecting graduation [17]. Students with high attendance rates tend to have a greater chance of graduating as they spend more time following the learning process.

Participation in Activities Participation in activities, such as group discussions and projects, also contributes to student passing. Students who actively participate tend to understand the material better and have better study skills.

5. Discussion

The output of this research shows that the Naive Bayes algorithm is effective in predicting student graduation at PKBM Bollo DMansel with an accuracy of 86.37%. From the model evaluation, it was found that the precision and recall for the "Pass" class reached 95.51% and 89.47% respectively, indicating that the model is very good at identifying students who will graduate. However, the performance of the model for the "Not Passed" class is less satisfactory with a precision of 23.08% and a recall of 42.86%, indicating that the model is still often wrong in predicting students who do not pass. The main factors affecting graduation have been identified through further analysis of the

features used in the model. End-of-semester exam scores, attendance, and participation in activities proved to be significant indicators in determining student graduation. Final semester exam scores reflect students' mastery of the material, while attendance and participation in activities show students' commitment and involvement in the learning process.

The implementation of this prediction model can assist PKBM Bollo DMansel in identifying students who are at risk of not graduating early. Thus, interventions can be made proactively to provide additional support for students in need. For example, students with low attendance or minimal participation in activities can be given more attention to increase their motivation and engagement, while students with low end-of-semester exam scores can be given additional tutoring. This allows PKBM Bollo DMansel to improve the overall graduation rate and ensure that all students have a fair chance to succeed.

6. Conclusion

The research conducted provides output or results that the application of the Naive Bayes algorithm in predicting the graduation of Package C students at PKBM Bollo DMansel can produce predictions with a fairly high accuracy. From the model evaluation results, it was found that the prediction accuracy was 86.37% with good precision and recall for the "Pass" class. However, the performance of the model for the "Not Passed" class still needs to be improved. The main factors affecting student graduation were identified to include final semester exam scores, attendance, and participation in activities. Using this prediction model, PKBM Bollo DMansel can identify students who are at risk of not graduating earlier, so that appropriate interventions can be made to increase their chances of graduating.

7. Suggestion

To improve the accuracy of predicting the graduation of Package C students at PKBM Bollo DMansel, it is recommended that the dataset be enlarged and updated regularly by adding variables such as family support and learning motivation. Class balancing also needs to be considered by applying oversampling or undersampling techniques. In addition, compare the performance of the Naive Bayes algorithm with other algorithms such as Decision Tree or Random Forest. The prediction results can be used to design intervention programs such as additional tutoring and counseling. Evaluate the model periodically with new data and update it if necessary. Also consider the use of technology, such as an automated notification system to students and parents about the risk of not graduating, to increase their involvement in the education process. The implementation of these suggestions is expected to improve prediction performance and help increase student pass rates at PKBM Bollo DMansel.

Acknowledgments: Thanks to fellow lecturers and students at the Department of Information Systems and Technology, Nobel Institute of Technology and Business Indonesia, hopefully, research on Naive Bayes continues to be developed and can achieve and produce accurate predictions for various purposes. and hopefully, this research continues to be developed and can be a good citation for scientific development, especially in the same discipline.

Author contributions: All authors are responsible for building Conceptualization, Methodology, analysis, investigation, data curation, writing—original draft preparation, writing—review and editing, visualization, supervision of project administration, funding acquisition, and have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. F. Ramadhani, Al-Khowarizmi and I. P. Sari, "Improving the Performance of Naïve Bayes Algorithm by Reducing the Attributes of Dataset Using Gain Ratio and Adaboost," 2021 International Conference on Computer Science and Engineering (IC2SE), Padang, Indonesia, 2021, pp. 1-5, doi: 10.1109/IC2SE52832.2021.9792027.
2. A. Spada, M. Fiore, and A. Galati, "The Impact of Education and Culture on Poverty Reduction: Evidence from Panel Data of European Countries," Soc Indic Res, 2023, doi: 10.1007/s11205-023-03155-0.
3. S. Reddy P, S. R and V. M, "Enhancing Accuracy in Crop Yield Estimation Through the Integration of Naive Bayes and Convolutional Neural Networks," 2023 Intelligent Computing and Control for Engineering and Business Systems (ICCEBS), Chennai, India, 2023, pp. 1-5, doi: 10.1109/ICCEBS58601.2023.10449096.
4. E. Aditya, Z. Situmorang, B. H. Hayadi, M. Zarlis and Wanayumini, "New Student Prediction Using Algorithm Naive Bayes And Regression Analysis In Universitas Potensi Utama," 2022 4th International Conference on Cybernetics and Intelligent System (ICORIS), Prapat, Indonesia, 2022, pp. 1-6, doi: 10.1109/ICORIS56080.2022.10031391.
5. N. Hafidhoh, A. P. Atmaja, G. N. Syaifuddiin, I. B. Sumafta, S. M. Pratama, and H. N. Khasanah, "Machine Learning untuk Prediksi Kegagalan Mesin dalam Predictive Maintenance System," Jurnal Masyarakat Informatika, vol. 15, no. 1, pp. 56–66, May 2024, doi: 10.14710/jmasif.15.1.63641.
6. H. Gao, "Spam Sorting Based on Naive Bayes Algorithm," 2023 International Conference on Blockchain Technology and Applications (ICBTA), Beijing, China, 2023, pp. 81-84, doi: 10.1109/ICBTA60381.2023.00023.
7. A. Sastrawijaya, E. Pujiastuti, and S. Maemuna, "Penyusunan Kurikulum Operasional Pendidikan Kesetaraan Berbasis Kearifan Lokal pada PKBM," Jurnal Kewarganegaraan, vol. 7, no. 2, 2023.
8. V. N. Reddy and R. Priyanka, "Analysis of Novel Seizure Detection system using Linear Regression Algorithm to Improvise the Performance Efficiency in Comparison with Naive Bayes," 2022 5th International Conference on Contemporary Computing and Informatics (IC3I), Uttar Pradesh, India, 2022, pp. 1910-1914, doi: 10.1109/IC3I56241.2022.10072424.
9. M. Rafi Nauf Akbar, E. Kau Suni, J. Meruya Selatan No, and J. Barat, "Penerapan Data Mining Untuk Memprediksi Penjualan Makanan Pada Kopi Krintji Menggunakan Algoritma Regresi Linear dan C4.5," Jurnal Ilmu Teknik dan Komputer, vol. 6, no. 2, 2022.
10. W. Xiaofang, L. Lan, Z. Qianyin, L. Fengyu, L. Jiawei and H. Di, "Constructing Naive Bayesian Classification Model by Spark for Big Data," 2020 17th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP), Chengdu, China, 2020, pp. 306-309, doi: 10.1109/ICCWAMTIP51612.2020.9317362.
11. F. K. Mega, Y. Yaddarabullah and S. D. H. Permana, "Classification Auction Motorcycle and Car In South Jakarta District Attorney Using Naïve Bayes," 2021 IEEE International Conference on Communication, Networks and Satellite (COMNETSAT), Purwokerto, Indonesia, 2021, pp. 168-171, doi: 10.1109/COMNETSAT53002.2021.9530827.
12. S. Mandala, A. I. Ramadhan, M. Rosalinda, W. M. S. Yafooz and R. H. Khohar, "DDoS Detection by Using Information Gain-Naïve Bayes," 2022 2nd International Conference on Intelligent Cybernetics Technology & Applications (ICICyTA), Bandung, Indonesia, 2022, pp. 283-288, doi: 10.1109/ICICyTA57421.2022.10038054.
13. V. Vijay and P. Verma, "Variants of Naïve Bayes Algorithm for Hate Speech Detection in Text Documents," 2023 International Conference on Artificial Intelligence and Smart Communication (AISC), Greater Noida, India, 2023, pp. 18-21, doi: 10.1109/AISC56616.2023.10085511.
14. T. Liu, "Research on Naive Bayes Integration Method based on Kmeans++ digital teaching clustering," 2023 8th International Conference on Information Systems Engineering (ICISE), Dalian, China, 2023, pp. 302-305, doi: 10.1109/ICISE60366.2023.00069.
15. P. Mai et al., "Implementasi Data Mining Menggunakan Algoritma Apriori Dalam Menentukan Persediaan Barang (studi kasus : toko sinar harahap)," 2022. [Online]. Available: https://jurnal.umj.ac.id/index.php/just-it/index

16. P. C. Kumar and B. T. Geetha, "Efficient Removal of Real Time Rain Streaks from A Image using Novel Naive Bayes (NB) Compare over Linear Regression (LR) with Improved Accuracy," 2023 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI), Chennai, India, 2023, pp. 1-6, doi: 10.1109/ACCAI58221.2023.10199408.
17. M. Yassir, H. Syam, and H. Nur, "Higher Order Thinking Skills (HOTs) based Assessment for Learning: A Model for Computer Networks Learning in Vocational School," 2022. [Online]. Available: www.ajouronline.com
18. M. Yassir, "Pembangunan Sistem Informasi Akademik (Studi Kasus : STIA AL Gazali Barru)." [Online]. Available: https://elektroda.uho.ac.id/
19. I. Fadil, M. A. Helmiawan, F. Supriadi, A. Saeppani, Y. Sofiyan and A. Guntara, "Waste Classifier using Naive Bayes Algorithm," 2022 10th International Conference on Cyber and IT Service Management (CITSM), Yogyakarta, Indonesia, 2022, pp. 1-5, doi: 10.1109/CITSM56380.2022.9935894.
20. S. K. L. Sameer and P. Sriramya, "Improving the Accuracy for Prediction of Heart Disease by Novel Feature Selection Scheme using Decision tree comparing with Naive-Bayes Classifier Algorithms," 2022 International Conference on Business Analytics for Technology and Security (ICBATS), Dubai, United Arab Emirates, 2022, pp. 1-8, doi: 10.1109/ICBATS54253.2022.9758926.
21. Y. K. R. Konduru and K. V. Kanimozhi, "Innovative Method of Predicting Crime Rate by CNN Model Compared with Naive Bayes Algorithm to Increase Accuracy," 2023 Intelligent Computing and Control for Engineering and Business Systems (ICCEBS), Chennai, India, 2023, pp. 1-3, doi: 10.1109/ICCEBS58601.2023.10448816.
22. R. Ismayanti, F. Alameka, D. Mirwansyah, N. W. W. Sari, A. Rahim and Riyayatsyah, "Identification of Pests on Black Orchid Plants Using Naïve Bayes Method Based on Leaf Image Texture," 2022 International Conference of Science and Information Technology in Smart Administration (ICSINTESA), Denpasar, Bali, Indonesia, 2022, pp. 167-172, doi: 10.1109/ICSINTESA56431.2022.10041471.
23. B. Marapelli, S. Kadiyala and C. S. Potluri, "Performance Analysis and Classification of Class Imbalanced Dataset Using Complement Naive Bayes Approach," 2023 Advanced Computing and Communication Technologies for High Performance Applications (ACCTHPA), Ernakulam, India, 2023, pp. 1-7, doi: 10.1109/ACCTHPA57160.2023.10083369.
24. R. Kurniawan, Nurkholis, M. Z. Ahmad Nazri, F. Lestari, R. Salambue and Sukamto, "A Web-Based Zakat Recipient Determination System using the Naïve Bayes Algorithm," 2022 5th International Conference of Computer and Informatics Engineering (IC2IE), Jakarta, Indonesia, 2022, pp. 337-341, doi: 10.1109/IC2IE56416.2022.9970014.
25. S. S and S. Sivanantham, "A Novel Method for Automatic Payment of Car Parking Charges with Random Forest and Comparison over Naïve-Bayes Classification," 2023 Intelligent Computing and Control for Engineering and Business Systems (ICCEBS), Chennai, India, 2023, pp. 1-5, doi: 10.1109/ICCEBS58601.2023.10448938.