DOI: 10.14421/ijid.2022.3733
IJID (International Journal on Informatics for Development), *e-ISSN: 2549-7448*
Vol. 11, No. 2, 2022, Pp. 252-261

Forecasting: Analyze Online and Offline Learning Mode with Machine Learning Algorithms

Farida Ardiani
Information System
Technology University of Yogyakarta
Yogyakarta, Indonesia
ardianifarida@gmail.com

Rodhiyah Mardhiyyah
Computer Engineering
Technology University of Yogyakarta
Yogyakarta, Indonesia
rodhiyah.office@gmail.com

Izaaz Azaam Syahalam
Informatics
Technology University of Yogyakarta
Yogyakarta, Indonesia
izaazsyahalam@gmail.com

Nasmah Nur Amiroh
Informatics
Technology University of Yogyakarta
Yogyakarta, Indonesia
nasmah466@gmail.com

Article History
Received November 3rd, 2022
Revised December 30th, 2022
Accepted February 1st, 2023
Published February 2023

***Abstract* — Since the pandemic occurred, in March 2020, learning activities have changed from an offline to an online learning mode. This is the first time, such a huge change has occurred, simultaneously in the entire hemisphere. This learning mode opens a new discourse regarding the impact on the learning mode and educational evaluation results. The author aims to compare the results of the educational evaluation of the online learning mode during the pandemic with offline learning mode, so that differences will be known, as well as can be used to predict student learning outcomes, in order to obtain an overview of the effectiveness and efficiency of a learning mode. Data collection is carried out as an initial step in data processing, based on the final results of student learning, in certain courses taken every semester starting in 2017-2022. The data consists of 6 indicators, namely CI1-CI4, grades, and letter grades. The result of this study is the prediction of a more effective learning mode used, as decision support carried out by the forecasting method, comparing the Naïve Bayes and Decision Tree algorithm in getting the best accuracy value, by analyzing the learning mode offline to online.**

Keywords–educational evolution; online learning mode; decision tree; Naïve Bayes classifier; RapidMiner.

1 INTRODUCTION

The global outbreak of the COVID-19 pandemic has spread throughout the world, affecting almost all countries and Regions. According to an official letter and WHO report, on 2 March 2020, Joko Widodo, president of Indonesia announced the first COVID-19 case in Indonesia. Its impact affects learning mode during the new school year. Some schools, colleges, and universities have discontinued face-to-face learning. On March 15, 2020, the president announced measures needed to be taken to encourage work, study, and worship activities at home and suspend any activities involving large numbers of participants. The government is trying to fulfil the right to educational services by conducting distance learning or online using gadgets or laptops to prevent the spread of COVID-19 in universities.

According to KBBI, offline learning mode is an activity that is carried out directly without a platform that requires a network connection or is called a face-to-face meeting. Online learning mode means in a network, connected through a computer network, the internet, and so on which explains that online is an activity carried out online through a platform with an internet network as a support so that it can be interpreted that this is done without a meeting in person. Platform facilities used for online learning modes such as google meet, zoom, classroom, google drive, WhatsApp, and others [1]. Online and offline learning mode certainly has a negative and positive impact [2] so it requires an evaluation of the learning mode and results [3]. The cumulative grade point index is one index that can help predict graduation and the quality of educational institutions [4]. The change in learning mode that occurred very suddenly made us alert, what if in the future the same thing happens, then what steps should we take to anticipate that using online or offline learning modes can still improve the quality of student education? Comparison of the results of online and offline learning modes is intended as the main goal of the authors to predict student learning outcomes so that the lecturer can understand the comparison of student competencies when learning modes are carried out online or offline. By knowing the differences in evaluating learning outcomes through learning models with different methods, it can improve the quality of higher education, which is done by predicting student academic scores so that it helps the process of taking action to improve student performance and reduce failures in human resource management [5]. The results of student learning evaluations are used as a determinant of learning performance [6]. It is hoped that after this research, further research can be carried out to find out what factors influence the results of the learning evaluation. With the results of previous research, the results of student predictions, both online and offline learning modes were carried out to determine the predicted results and the results of offline and online learning mode values that can be used to make considerations in determining effective and efficient learning mode in the education sector in determining student understanding in learning, in accordance with the values obtained.

In previous studies, students of SMAN 1 Singosari showed that the results of the development of learning media in the form of worksheets were used to evaluate student learning outcomes which were divided into two groups alternately, namely online and offline learning groups. Students who get online learning turn to find it difficult to focus on learning. Competency achievement in worksheets shows a higher score in the offline class, which is 93.84, while in the online class, it is 86.10 [7]. In online learning, besides being more difficult for students to focus, it also has the potential to drop out of online learning. A study presents the performance of the Decision Tree, Random-Forest, Support Vector Machine (SVM), and Deep Neural Network (DNN) algorithms to predict the risk of dropping out during online learning by applying learner statistical information, number of system connections, number of lectures, previous semester grade data with Drop out and non-dropout prediction data [8]. As many as 58% of EFL students at Ibnu Khaldun University prefer offline learning to online learning. According to students, learning activities run more interactively, focus more easily, and understand the material more easily. These results were obtained from data analysis on student perceptions of online and offline learning. The data was taken from a questionnaire that students had filled out with questions in the form of a Likert scale [9]. Based on several previous studies that have been described above, reached the conclusion that learning modes affect student grades.

In this research, the steps taken in data processing by processing and analyzing students ' final grades with the decision tree algorithm and naïve Bayes used are student learning outcomes during offline and online learning modes (2018-2021). The tool used by the author to process and analyze data in RapidMiner. The process is carried out using the decision tree algorithm and naïve Bayes classifier. Comparison with the best and highest prediction values will be used as a reference for future learning or forecasting will be carried out by online or offline methods.

2 METHOD

This study uses machine learning algorithms, namely decision tree and naïve Bayes classifier to produce a comparison of online and offline learning model outcomes. In terms of performance, there is a comparison table that compares the several algorithms used. But based on their performance in analyzing the data to be used, decision tree and naive Bayes algorithms are the best algorithms, in terms of performance and time. In addition, based on the data used, the algorithm is in accordance with the requirements for using forecasting. This study resulted in the accuracy of both algorithms so that it can help decision-makers in determining future learning modes both online and offline. This is done by forecasting data mining methods with machine learning algorithms that have been determined. Forecasting is a process of predicting the future as accurately as possible by using all historical information and data as existing knowledge [10]. The forecasting method is used because it is expected to help consider making decisions in the application of effective and efficient learning modes. There are several processes carried out in this study, among others, see Fig. 1.

2.1 Research Data

The object used in this study is the final grade of students used to compare algorithms and predict learning mode. In the previous study, the average writer uses RapidMiner as a data processing tool which is an open-source framework used to assist analysts in preparing data, machine learning mode, text mining, to data analysis [11], [12].

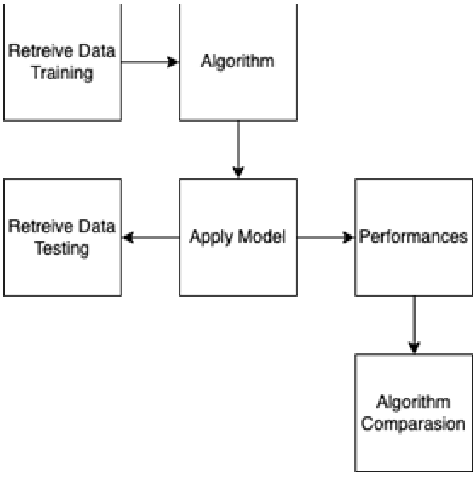

Figure 1. System design functionality

Retrieve data training is the process of inputting data that has been collected and used as a dataset that has gone through the process of normalization and labeling which then part of this dataset is trained to make predictions or perform the functions of another machine learning algorithm according to their respective. The data in this study were taken from a private university in the Special Region of Yogyakarta, Indonesia, with a study program majoring in computer engineering. C1 stands for competency 1. This means that C1 is a student's score in achieving competence, according to the semester's learning plan in each semester there are four kinds of assessment competencies that will eventually produce grades and be converted to letter grades, so letter grades are the final result, the competency achievement score of each student.

Retrieve training data in this study using the data used amounted to 4299 training data, the data is taken from the university's academic information system, which is divided into 2450 training data in the online learning mode period of the academic year 2019-2020, 2020-2021, 2021-2022 and 1849 training data in the offline learning mode period of the academic year 2017-2018, 2018-2019, 2019-2020 in the format .xlsx and .csv purposes, see Table 1.

Indicator x uses the final grades of students in the 2018-2022 academic year, consisting of CI 1, CI 2, CI 3, CI 4, CI stands for Course Indicator, Grades, and Letter Grades. The total data used in this study were 2450 online learning mode data and 1894 offline learning mode data.

Retrieve testing data in the study is the process of collecting test data that has been tested to see the accuracy resulting from the process that has been done in accordance with the purpose of the data to be tested against the training data. Test data can be seen in Table 2.

Table 1. Indicators of Data Training

CI1	CI2	CI3	CI4	Grades	Letter Grades
10	18,75	18,75	16,5	64	B
16	18,75	20	28,5	83,25	A
10	21,25	20	15	66,25	B
10	21,25	18,75	15	65	B
20	21,25	22,5	25,5	89,25	A
12	20	21,25	18	71,25	B
12	20	18,75	12	62,75	B
10	20	18,75	20,1	68,85	B
10	21,25	21,25	30	82,5	A

Table 2. Indicators of Data Testing

CI1	CI2	CI3	CI4	Grades
10	18,75	18,75	16,5	64
16	18,75	20	28,5	83,25
10	21,25	20	15	66,25
10	21,25	18,75	15	65
20	21,25	22,5	25,5	89,25
12	20	21,25	18	71,25
12	20	18,75	12	62,75
10	20	18,75	20,1	68,85
10	21,25	21,25	30	82,5

2.2 Algorithms

Algorithm in mathematics is a procedure or steps taken to obtain a calculation [13]. Decision tree algorithm and naïve Bayes classifier is an algorithm of machine learning used in this study. Machine learning is a technique that helps improve the performance of data by learning through computational processes [14], [15] Decision tree is an algorithm whose attributes use a gain ratio which is usually used for numerical or categorical data [16] and in the process is formed by a result variable Yes (1) and No (0)[17] . The naïve Bayes classifier algorithm is a machine learning algorithm used to calculate probabilities in the process of building a model that is used to calculate the probability of each value of the input variable [16]. This study uses several algorithms including the following:

2.2.1. Decision Tree: A decision tree is a machine learning algorithm whose attributes use a gain ratio which is usually used for numerical or categorical data [16] 11 and in the process is formed by a result variable Yes (1) and No (0) [17]. The decision tree is an algorithm with a supervised approach, which has the main advantages; high classification accuracy and strong robustness [18], [19]. The decision tree algorithm is also quite simple and does not rely on special knowledge of statistics [20], but uses the information reduced by the maximum entropy feature to divide the data in the next step [19]. In addition, this algorithm is good for analyzing and providing rules in forecasting models even though the data used has different periods [21].

2.2.2. *Naïve Bayes Classifier:* Naïve Bayes classifier is one of the many machine learning algorithms used to calculate the probability in the process of building a model that is used to calculate the probability of each value of the input variable [16]. This algorithm has several advantages such as fast calculation, simple algorithm, and high accuracy even though it is used for large data [22], [23] and even this algorithm only requires a small amount of training data, to provide parameter estimates in the classification process [24]. In addition, this naïve bayes classifier algorithm can overcome missing values and still has good performance against irrelevant attributes and noise in the data [22].

2.3 Apply Model

Apply Model is an operator used in RapidMiner to train datasets that have been input using the selected algorithm [11]. In this study, the authors used the decision tree algorithm and naïve Bayes classifier. The process of applying the model in this study is described in a flowchart, see Fig. 2.

2.4 Performances

Performance is an operator RapidMiner operator used for performance evaluation of all types of learning tasks that are executed in an operation [11].

2.5 Algorithm Comparison

Comparison of decision tree algorithm and naïve Bayes classifier in this study was used to compare the effectiveness of online and offline learning mode. This comparison will provide information about the analysis of existing processes so that there will be differences that can be compared between the two algorithms through the resulting performance.

3 RESULT AND DISCUSSION

3.1 Decision Tree Online Result Modeling

Model built using decision tree algorithm based on data train online learning. Fig. 3 shows a knowledge model in the form of a decision tree based on the decision tree algorithm.

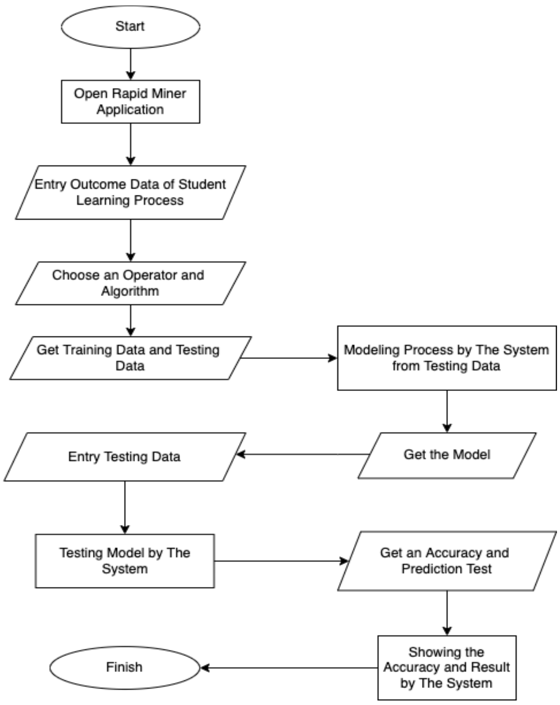

Figure 2. Flowchart

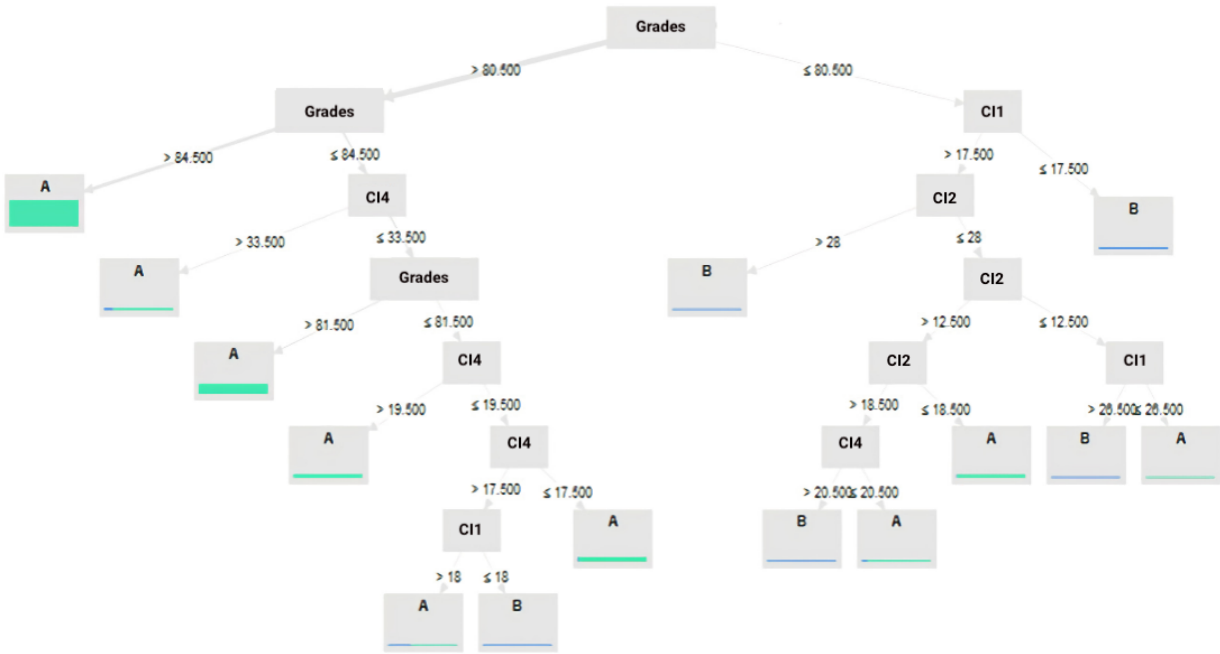

Figure 3. Online model by Decision Tree Algorithm

Table 3. Online Student Final Score Prediction Result Based on Decision Tree Algorithm

Row No.	Prediction (Letter Grades)	Confidence (B)	Confidence (A)	Confidence (E)	Confidence (C)	Confidence (D)	CI1	CI2	CI3	CI4	Grades
1	A	0.133	0.867	0	0	0	11	21	11	38	81
2	A	0.133	0.867	0	0	0	11	21	11	40	83
3	A	0.133	0.867	0	0	0	11	21	12	38	82
4	A	0.133	0.867	0	0	0	11	21	11	41	84
5	A	0.133	0.867	0	0	0	11	21	12	38	82
6	A	0.133	0.867	0	0	0	11	21	11	37	81
7	C	0.014	0	0	0.967	0.019	11	21	11	31	41
8	C	0.014	0	0	0.967	0.019	11	20	9	16	55
9	A	0	1	0	0	0	11	21	11	41	85
10	A	0.133	0.867	0	0	0	11	21	12	38	82
11	A	0.133	0.867	0	0	0	11	21	11	41	84
12	A	0	1	0	0	0	11	21	11	42	85
13	A	0.133	0.867	0	0	0	11	21	11	38	81
14	A	0.133	0.867	0	0	0	11	21	12	38	82
15	A	0	1	0	0	0	12	26	12	37	86

In grouping data to predict student learning outcomes, the decision tree knowledge model uses a number value as the root. Prediction based on each existing weight based on the model formed so that if the weight of the data value at CI 1 is less than 85,000 and less than 17,500. Thus, the data is predicted to get the final grade of the letter in the form of a value of B. The results of the prediction can be seen in Table 3. How to calculate confidence in a decision tree that is by comparing the predicted results with actual results. As long as it is still in the context of prediction then, a trust value cannot be calculated until when the actual value already exists it can be calculated. However, the confidence value can be assumed to be high based on the comparison of the predicted value with the actual value, on the training data used.

Table 3 shows the results of student learning value predictions for the 2021-2022 academic year. The test used 216 test data that predicted the value of student learning mode outcomes. Prediction is based on value indicators contained in data such as CI1, CI2, CI3, CI4, and numeric values with class labels in the form of letter grades, which are used as information on the values to be predicted. Prediction of the letter grades is based on the value of confidence, the greater the value, the prediction results will follow the confidence value. Data with row 1 predicted to get the value of the letter A, based on the value of the greatest confidence in confidence A. The number of students grouped by predicted value can be seen in Table 4.

Based on Table 4, the results of predictions made on the knowledge model were generated from the decision tree algorithm. Data training using online learning mode data. The predicted value of A is obtained by 183 students, the value of B is 21 students, the value of C is 7 students, the value of D is 1 student, and the value of E is 4 students. The resulting accuracy is based on the comparison of the predicted results with the actual situation, the resulting accuracy value can be seen in Table 5.

Table 4. Number of Students Based on Predicted Grades

Row No.	Prediction (Letter Grades)	count (prediction (Letter Grades)
1	A	183
2	B	21
3	C	7
4	D	1
5	E	4

Table 5. The Test Data Accuracy Results Using Decision Tree Algorithm in Online Learning

	true A	true B	true C	true D	true E	class precision
pred.A	183	0	0	0	0	100.00%
pred.B	1	5	0	0	1	71.43%
pred.C	1	0	20	0	0	95.24%
pred.D	0	0	0	4	0	100.00%
pred.E	0	0	0	0	1	100.00%
class recall	98.92%	100.00%	100.00%	100.00%	50.00%	

Table 5 shows the accuracy of the test data performed predicted value of student learning outcomes. Test data accuracy of 98.61%. Process errors contained in the prediction of the value of the letter B as many as 2 errors with a precision class of 71.43% and C as many as 1 error with a precision class of 95.24%.

3.2 *Naïve Bayes Classifier Online Result Modeling*

Fig. 4 is knowledge in the form of rules based on online learning training data. The knowledge model uses a simple distribution where the data are grouped based on the probability value (0.287), the data is included in the Class B group. The predicted result can be seen in Table 6.

Table 6 shows the results of the student's final grade prediction. Testing using test data made predictions of the final value of students, using the algorithm naïve Bayes classifier. Prediction is based on value indicators contained in data such as CI1, CI2, CI3, CI4, and numeric values with class labels in the form of letter grades, which are used as information on the values to be predicted. The higher the confidence value of data, the better the prediction results will follow the confidence value. The data in line I is expected to get the value of the letter A based on the highest confidence value of confidence A. Table 7 shows the number of students grouped by their predicted grades.

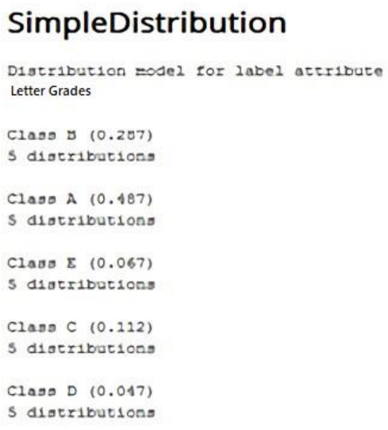

Figure 4. Naïve Bayes classifier model based on online data

Table 6. Online Student Final Score Prediction Result Based on Naïve Bayes Classifier

Row No.	**Prediction (Letter Grades)**	**Confidence (B)**	**Confidence (A)**	**Confidence (E)**	**Confidence (C)**	**Confidence (D)**	**CI1**	**CI2**	**CI3**	**CI4**	**Grades**
1	A	0.109	0.891	0	0.000	0.000	11	21	11	38	81
2	A	0.038	0.962	0	0.000	0.000	11	21	11	40	83
3	A	0.071	0.928	0	0.000	0.000	11	21	12	38	82
4	A	0.022	0.978	0	0.000	0.000	11	21	11	41	84
5	A	0.071	0.928	0	0.000	0.000	11	21	12	38	82
6	A	0.125	0.875	0	0.000	0.000	11	21	11	37	81
7	C	0.000	0.000	0	0.997	0.003	11	21	11	31	41
8	C	0.046	0.000	0	0.952	0.001	11	20	9	16	55
9	A	0.015	0.985	0	0.000	0.000	11	21	11	41	85
10	A	0.071	0.928	0	0.000	0.000	11	21	12	38	82
11	A	0.022	0.978	0	0.000	0.000	11	21	11	41	84
12	A	0.012	0.988	0	0.000	0.000	11	21	11	42	85
13	A	0.109	0.891	0	0.000	0.000	11	21	11	38	81
14	A	0.071	0.928	0	0.000	0.000	11	21	12	38	82
15	A	0.015	0.985	0	0.000	0.000	12	26	12	37	86

Based on Table 7, the results of predictions made on the knowledge model generated from the algorithm naïve Bayes based on training data using online data. The predicted value of A is obtained by 183 students, the value of B is 8 students, the value of C is 16 students, the value of D is 2 students, and the value of E is 4 students. The results' accuracy is based on comparing the predicted values with the actual conditions. The resulting precision values are shown in Table 8.

Table 7. Number of Students Based on Predicted Grades

Row No.	**Prediction (Letter Grades)**	**Count prediction (Letter Grades)**
1	A	186
2	B	8
3	C	16
4	D	2
5	E	4

Table 8 shows the accuracy of the test data performed predicted value of student learning outcomes. Test data accuracy of 93.98%. Process errors occur in the prediction of the value of the letter A with as many as 2 errors with a precision class of 98.92% and B with as many as 11 errors with a precision class of 31.25%.

Table 8. The Test Data Accuracy Results Using Naïve Bayes Classifier Algorithm in Online Learning

accuracy of 93.98%

	true A	true B	true C	true D	true E	class precision
pred.A	184	0	2	0	0	98.92%
pred.B	1	5	10	0	0	31.25%
pred.C	0	0	8	0	0	100.00%
pred.D	0	0	0	4	0	100.00%
pred.E	0	0	0	0	2	100.00%
class recall	99.46 %	100.00 %	40.00 %	100.00 %	100.00 %	

3.3 *Comparison of Online Modeling Results*

The results of both models with the same data can be seen that the decision tree algorithm produces higher accuracy than the algorithm naïve Bayes classifier. The decision tree algorithm predicts that students who get an A score of as much as 183, the same prediction generated by the algorithm naïve Bayes classifier. A prominent difference is found in the prediction of B and C grades, the decision tree algorithm predicted more B grades for 21 students compared to the naïve Bayes classifier which predicted B grades for 8 students. In addition, the value of C in the decision tree algorithm is less with a total of 7 predictions than the naïve Bayes classifier algorithm which produces predictions as many as 27 predictions of student values. In addition, the prediction error generated by the decision tree algorithm is less than that of the naïve Bayes classifier algorithm. Thus, it affects the accuracy of the predictions produced.

3.4 Decision Tree Offline Results Modeling

The Model is built using the decision tree algorithm based on offline learning mode training data and test data. The resulting model based on the decision tree algorithm can be seen in Fig. 5.

Fig. 5 is a knowledge model in the form of a decision tree algorithm based on offline learning mode training data. The knowledge model uses the number value attribute as the root. Data grouping is used to predict student learning outcomes. The prediction is based on the existing weights based on the former model, so for numerical weights above 80,500, the data can receive the final value in the form of a value shown in Table 9.

Table 9 shows the results of the student's final grade prediction. Data testing is done to predict students' final grade using the decision tree algorithm. Prediction is based on value indicators contained in data such as CI1, CI2, CI3, CI4, and numeric values with class labels in the form of letter grades, which are used as information on the values to be predicted. The prediction of the letter grades is based on confidence, where the greater the value, the prediction results will follow the confidence value. The data in line I gets the value of the letter A based on the distribution of the value of confidence A. Table 10 shows the number of students grouped by their predicted grades.

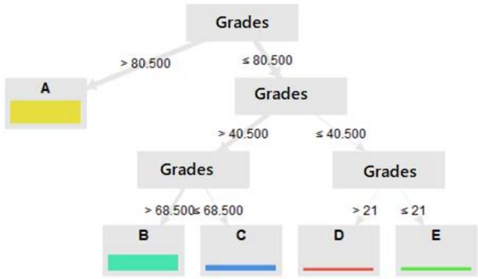

Figure 5. Decision tree model based on offline data

Table 9. Offline Student Final Score Prediction Result Based on Decision Tree Algorithm

Row No.	Prediction (Letter Grades)	Confidence (B)	Confidence (A)	Confidence (E)	Confidence (C)	Confidence (D)	CI1	CI2	CI3	CI4	Grades
1	A	0.001	0.999	0	0	0	11	21	11	38	81
2	A	0.001	0.999	0	0	0	11	21	11	40	83
3	A	0.001	0.999	0	0	0	11	21	12	38	82
4	A	0.001	0.999	0	0	0	11	21	11	41	84
5	A	0.001	0.999	0	0	0	11	21	12	38	82
6	A	0.001	0.999	0	0	0	11	21	11	37	81
7	C	0	0	0	1	0	11	21	11	31	41
8	C	0	0	0	1	0	11	20	9	16	55
9	A	0.001	0.999	0	0	0	11	21	11	41	85
10	A	0.001	0.999	0	0	0	11	21	12	38	82
11	A	0.001	0.999	0	0	0	11	21	11	41	84
12	A	0.001	0.999	0	0	0	11	21	11	42	85
13	A	0.001	0.999	0	0	0	11	21	11	38	81
14	A	0.001	0.999	0	0	0	11	21	12	38	82
15	A	0.001	0.999	0	0	0	12	26	12	37	86

Table 10 shows an overview of the prediction results generated from the decision tree algorithm knowledge model based on training data using offline learning mode data. The results of the prediction for the value of A has a predicted value of 183 values, B has a value of 5, C has a value of 22, D has 2, and E has a value of 4 values. Accuracy is generated by comparing the predicted results with the actual conditions. The accuracy value can be seen in Table 11.

Table 10. Number of Students Based on Predicted Grades

Row No.	Prediction (Letter Grades)	Count (prediction (Letter Grades)
1	A	183
2	B	5
3	C	22
4	D	2
5	E	4

Table 11 shows the accuracy of the test data performed predicted value of student learning outcomes. Test data accuracy of 91.67%. Prediction process errors in the value of B as many as 17 errors with a precision class of 22.73% and C as many as 1 error with a precision class of 80.00%.

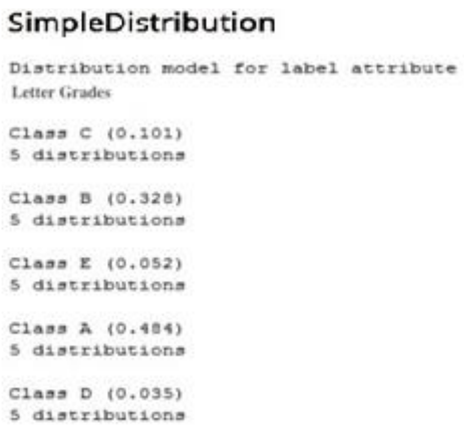

Figure 6. The Knowledge model algorithm Naïve Bayes Classifier

3.5 *Naïve Bayes Classifier Offline Result Modeling*

The Model was created using the algorithm naïve bayes classifier based on data trained online learning. The resulting model based on the naïve Bayes classifier algorithm can be seen in Fig. 6.

Fig. 6 shows a knowledge model in the form of a rule or rule based on online learning training data. The knowledge model uses a simple distribution whose data are grouped by probability value, when the data has a probability value (0.328), the data belongs to the Class B group. The predicted result can be seen in Table 12.

Table 11. The Test Data Accuracy Results Using Decision Tree Algorithm in Offline Learning

accuracy of 91.67%

	true A	true B	true C	true D	true E	class precision
pred.A	183	0	0	0	0	100.00%
pred.B	1	5	16	0	0	22.73%
pred.C	1	0	4	0	0	80.00%
pred.D	0	0	0	4	0	100.00%
pred.E	0	0	0	0	2	100.00%
class recall	98.92%	100.00%	20.00%	100.00%	100.00%	

Table 12 shows the results of students' final score prediction. Testing using test data that has been carried out the prediction process using the naïve Bayes classifier algorithm. Prediction is based on value indicators contained in data such as CI1, CI2, CI3, CI4, and numeric values with class labels in the form of letter grades, which are used as information on the values to be predicted. The prediction of the letter grades is based on confidence, where the greater the value, the prediction results will follow the confidence value. Data with row I predicted to get the value of the letter C, based on the value of confidence spread there on confidence C. The number of students grouped by predicted value can be seen in Table 13.

Table 13. Number of Students Based on Predicted Grades

Row No.	Prediction (Letter Grades)	Count (prediction (Letter Grades)
1	A	39
2	C	172
3	D	1
4	E	4

Based on Table 13, the results of the prediction made by the knowledge model generated from the naïve Bayes classifier algorithm based on offline learning mode data. The A predicted score was obtained by 29 students, the B score was obtained by 0 students, the C score was obtained by 172 students, the D score was obtained by 1 student, and the E score was obtained by 4 students. The resulting accuracy is based on the comparison of the predicted results with the actual situation, the accuracy value can be seen in Table 14.

Table 14. Test Data Accuracy Results Using Naïve Bayes Classifier Algorithm in Online Learning

accuracy of 22.69%

	true A	true B	true C	true D	true E	class precision
pred A	39	0	0	0	0	100.00%
pred.B	148	5	20	0	1	2.91%
pred.C	0	0	0	0	0	0.00%
pred D	0	0	0	4	0	100.00%
pred.E	0	0	0	0	1	100.00%
class recall	21.08 %	100.00 %	0.00 %	100.00 %	50.00 %	

Table 14 shows the accuracy of the test data performed predicted value of student learning mode outcomes. Test data accuracy of 22.69%. Prediction processes errors that occur in the value of B as many as 169 errors with a precision class of 2.91% and C which has no predictive value so as to obtain a value for the precision class of 00.00%.

Table 12. Offline Student Final Score Prediction Result Based on Naïve Bayes Classifier

Row No.	**Prediction (Letter Grades)**	**Confidence (B)**	**Confidence (A)**	**Confidence (E)**	**Confidence (C)**	**Confidence (D)**	**CI1**	**CI2**	**CI3**	**CI4**	**Grades**
1	C	0	0.000	0	1,000	0	11	21	11	38	81
2	C	0	0.000	0	1,000	0	11	21	11	40	83
3	C	0	0.000	0	1,000	0	11	21	12	38	82

4	C	0	0.000	0	1,000	0	11	21	11	41	84
5	C	0	0.000	0	1,000	0	11	21	12	38	82
6	C	0	0.000	0	1,000	0	11	21	11	37	81
7	C	0	0	0	1	0	11	21	11	31	41
8	C	0	0	0	1	0	11	20	9	16	55
9	C	0	0.000	0	1,000	0	11	21	11	41	85
10	C	0	0.000	0	1,000	0	11	21	12	38	82
11	C	0	0.000	0	1,000	0	11	21	11	41	84
12	C	0	0.000	0	1,000	0	11	21	11	42	85
13	C	0	0.000	0	1,000	0	11	21	11	38	81
14	C	0	0.000	0	1,000	0	11	21	12	38	82
15	A	0	0.817	0	0.813	0	12	26	12	37	86

3.6 *Comparison of Offline Modeling Results*

The results of both models with the same data can be seen that the decision tree algorithm produces higher accuracy than the naïve Bayes classifier algorithm. The decision tree algorithm predicts that students who get an A score of as much as 183 while the predictions generated by the algorithm naïve Bayes classifier as many as 29 students. A prominent difference is in the prediction of B and C grades, the decision tree predicts B grades obtained by 5 students, but the naïve Bayes classifier predicts B grades obtained by 0 students. Students who get the value of C in the decision tree algorithm are predicted as many as 22 students, but the prediction algorithm naïve Bayes classifier algorithm produces as many as 172 students. This is an inequality in the results of the offline value modeling process. In addition, the prediction error generated by the decision tree algorithm is less than that of the naïve Bayes classifier algorithm, thus affecting the accuracy of the predictions generated.

3.7 *Comparison of Algorithms*

Based on the process carried out to predict the final grade of students, both online and offline learning modes. Obtained results as in Table 15.

Table 15. Research Result

Algorithm	Method	Prediction "Final Letter Grades of Students"					Accuracy
		A	B	C	D	E	
Decision Tree	Offline	183	21	7	1	4	98.61%
	Online	183	5	22	2	4	91.67%
Naïve Bayes	Offline	186	8	16	2	4	93.98%
	Online	39	0	172	1	4	22.69%

Table 15 is a table of the results of predictions made using the data mining forecasting method with a decision tree and naïve Bayes classifier algorithm. The process is done by testing the data in accordance with both the model methods and algorithms used. The process continues by testing and training with both algorithms to see the difference and the resulting accuracy.

A prediction is strongly influenced by the training data and test data used, the more training data, the higher the accuracy value of the prediction results, there is a possibility that it affects if, why is the predicted value of B lower even though 2 algorithms have been used, the first possibility is, in the Training data B grades are rarely found, which means that student scores are used as training data, rarely get B grades, more students get A, C, D or E grades.

Each algorithm has a learning time, the learning time of the decision tree is relatively fast with a running time of 36.238 (ms), while naïve Bayes has a learning time with a time taken 58.044 (ms). A comparison of algorithms is done to see the best results from the decision tree algorithm and naïve Bayes. After going through the implementation process, the accuracy value obtained from the decision tree algorithm is 95% and the naïve Bayes algorithm is 94%.

4 CONCLUSION

After going through various processes, this study produced information in the form of predicted results of online and offline learning mode grades, with the results of analysis and prediction that a good learning mode is offline, the results of this study support previous research. Based on the algorithm used in the forecasting method, the decision tree algorithm has higher accuracy in predicting the final grade of students with a comparison of the accuracy of the online learning mode of the decision tree of 98.61% and 93.98% of naïve Bayes, while for the offline learning mode in the decision tree algorithm obtained an accuracy of 91.67% and 22.69% of the naïve Byes algorithm. So that from the process of analysis, design, implementation, and discussion, it was concluded that in this study produced information in the form of prediction results and the grades of offline and online learning modes that can be used to make considerations in determining effective and efficient learning modes. Methods used in research is the application of the concept of data mining methods forecasting with decision tree algorithm and naive Bayes classifier in modeling. The decision tree algorithm has a higher predictive accuracy value in online and offline learning modes, compared with the naïve Bayes classifier algorithm in the learning mode, also, especially in online learning mode, the decision tree algorithm potential to get higher A and B grades.

In addition, there are some suggestions that can be done by the next researcher if you want to continue this research such as developing an existing knowledge model through a web-based or mobile system, adding parameters such as the potential for help from people around, the level of understanding of student material or network, and experiments are needed for more learning mode that is applied using training data on the Applied knowledge model so that it can produce accuracy and prediction of a higher percentage of its value.

AUTHOR'S CONTRIBUTION

Farida Ardiani as the first author contributed ideas and data analysis. The second author, Rodhiyah Mardhiyyah, together with the first author, conducted data testing and evaluation, Izaaz Azaam Syahalam, algoritm comparation, and the last author, Nasmah Nur Amiroh, is the main author in this study.

COMPETING INTERESTS

Farida Ardiani, Rodhiyah Mardhiyyah and Nasmah Nur Amiroh as the authors of this article, stated that this article is free from Conflict of Interest (COI) or Competitive Interest (CI).

ACKNOWLEDGMENT

The author would like to thank the ministry of education culture research and technology 2022 for providing funding support in this research.

REFERENCES

[1] A. F. Himmah, 'Implementasi Strategi Metode Ceramah Dalam Pembelajaran Daring Dan Luring Pada Siswa Kelas II Mi Nurul Islam 02 Balung Kulon Di Masa Pandemi', *PESAT: Jurnal Pendidikan, Sosial, dan Agama*, vol. 6, no. 6, pp. 124–133, 2021.

[2] T. A. Nengrum, N. Pettasolong, and M. Nuriman, 'Kelebihan dan Kekurangan Pembelajaran Luring dan Daring dalam Pencapaian Kompetensi Dasar Kurikulum Bahasa Arab di Madrasah Ibtidaiyah 2 Kabupaten Gorontalo', *Jurnal Pendidikan*, vol. 30, no. 1, p. 1, 2021, doi: 10.32585/jp.v30i1.1190.

[3] I. R. Badriyah, A. Akhwani, N. Nafiah, and M. S. Djazilan, 'Analisis Model Pembelajaran Daring dan Luring pada Masa Pandemi Covid-19 di Sekolah Dasar', *Jurnal Basicedu*, vol. 5, no. 5, pp. 3651–3659, 2021, doi: https://doi.org/10.31004/basicedu.v5i5.1373 ISSN.

[4] A. Desiani, S. Yahdin, and D. Rodiah, 'Prediksi Tingkat Indeks Prestasi Kumulatif Akademik Mahasiswa dengan Menggunakan Teknik Data Mining', *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 7, no. 6, p. 1237, 2020, doi: 10.25126/jtiik.2020722493.

[5] R.- Annisa and A.- Sasongko, 'Prediksi Nilai Akademik Mahasiswa Menggunakan Algoritma Naïve Bayes', *JST (Jurnal Sains dan Teknologi)*, vol. 9, no. 1, pp. 1–10, 2020, doi: 10.23887/jstundiksha.v9i1.19488.

[6] A. Qoiriah and Y. Yamasari, 'Prediksi Nilai Akhir Mahasiswa Dengan Metode Regresi (Studi Kasus Mata Kuliah Pemrograman Dasar)', *Journal of Information Engineering and Educational Technology*, vol. 5, no. 1, pp. 40–43, 2021, doi: 10.26740/jieet.v5n1.p40-43.

[7] H. K. Baihaqi, J. Jumadi, A. and Mardiani, and R. Religia, 'Analyzing Student Inquiry In Online And Offline Classes During The Covid-19 Pandemic Through Momentum And Impulse Worksheets', *Momentum: Physics Education Journal*, vol. 5(2), pp. 161–174, 2021.

[8] H. S. Park and S. J. Yoo, 'Early Dropout Prediction In Online Learning Of University Using Machine Learning', *International Journal on Informatics Visualization*, vol. 5(4), pp. 347–353, 2021.

[9] F. L. Damayanti and N. Rachmah, 'Effectiveness Of Online Vs Offline Classes For EFL Classroom: A Study Case In A Higher Education, Journal of English Teaching', *Applied Linguistics and Literatures (JETALL)*, vol. 3(1), p. 19, 2020.

[10] R. J. Hyndman and G. Athanasopoulos, *Forecasting : Principles and Practice*, Second. Texts, 2018.

[11] M. Hofmann and R. Klinkenberg, *Rapid Miner Data Mining Use Cases and Business Analytics Applications*. 2014.

[12] H. Hapsari, M. D. Fauzi, and S. 'Uyun, 'An Application of Mamdani in Selecting Majors in Higher Education', *IJID*, vol. 2, no. 1, 2013.

[13] E. Manurung and P. S. Hasugian, 'Data Mining Tingkat Pesanan Inventaris Kantor Menggunakan Algoritma Apriori pada Kepolisian Daerah Sumatera Utara', *Journal Of Informatic Pelita Nusantara*, vol. 4, no. 2, pp. 8–13, 2019, [Online]. Available: https://e-jurnal.pelitanusantara.ac.id/index.php/JIPN/article/view/608/0

[14] Z.-H. Zhou, *Machine Learning*. Tsinghua University Press, 2021.

[15] G. Chassagnon, M. Vakalopolou, N. Paragios, and M. P. Revel, 'Deep learning: definition and perspectives for thoracic imaging', *Eur Radiol*, vol. 30, no. 4, pp. 2021–2030, 2020, doi: 10.1007/s00330-019-06564-3.

[16] R. Chauhan and H. Kaur, *Predictive Analytics and Data Mining*. 2015. doi: 10.4018/978-1-4666-9562-7.ch019.

[17] Suyatno, *Machine Learning Tingkat Dasar dan Lanjutan*, First Edit. Pasar Buku Planasari No.82, Bandung, 40264: Informatika Bandung, 2018.

[18] E. Pekel, 'Estimation of soil moisture using decision tree regression', *Theor Appl Climatol*, vol. 139, no. 3–4, pp. 1111–1119, 2020, doi: 10.1007/s00704-019-03048-8.

[19] H. F. Zhou, J. W. Zhang, Y. Q. Zhou, X. J. Guo, and Y. M. Ma, 'A feature selection algorithm of decision tree based on feature weight', *Expert Syst Appl*, vol. 164, no. September 2020, p. 113842, 2021, doi: 10.1016/j.eswa.2020.113842.

[20] M. Bansal, A. Goyal, and A. Choudhary, 'A comparative analysis of K-Nearest Neighbor, Genetic, Support Vector Machine, Decision Tree, and Long Short Term Memory algorithms in machine learning', *Decision Analytics Journal*, vol. 3, no. May, p. 100071, 2022, doi: 10.1016/j.dajour.2022.100071.

[21] D. Ramos, P. Faria, L. Gomes, and Z. Vale, 'A Contextual Reinforcement Learning Approach for Electricity Consumption Forecasting in Buildings', *IEEE Access*, vol. 10, pp. 61366–61374, 2022, doi: 10.1109/ACCESS.2022.3180754.

[22] D. A. Putri, D. A. Kristiyanti, E. Indrayuni, A. Nurhadi, and D. R. Hadinata, 'Comparison of Naive Bayes Algorithm and Support Vector Machine using PSO Feature Selection for Sentiment Analysis on E-Wallet Review', *J Phys Conf Ser*, vol. 1641, no. 1, 2020, doi: 10.1088/1742-6596/1641/1/012085.

[23] I. Romli, T. Pardamean, S. Butsianto, T. N. Wiyatno, and E. bin Mohamad, 'Naive Bayes Algorithm Implementation Based on Particle Swarm Optimization in Analyzing the Defect Product', *J Phys Conf Ser*, vol. 1845, no. 1, pp. 0–6, 2021, doi: 10.1088/1742-6596/1845/1/012020.

[24] M. R. Romadhon and F. Kurniawan, 'A Comparison of Naive Bayes Methods, Logistic Regression and KNN for Predicting Healing of Covid-19 Patients in Indonesia', *3rd 2021 East Indonesia Conference on Computer and Information Technology, EIConCIT 2021*, pp. 41–44, 2021, doi: 10.1109/EIConCIT50028.2021.9431845.