KLIK: Kajian Ilmiah Informatika dan Komputer
ISSN 2723-3898 (Media Online)
Vol 4, No 1, Agustus 2023, Hal 592-600
DOI 10.30865/klik.v4i1.963
https://djournals.com/klik

Naïve Bayes Classifier and Decision Tree Algorithms for Classifying Payment Data

Berti Sari Br Sembiring, Mahdianta Pandia, Harianta Sembiring, Desi Margaretta*

Program Studi Sistem Informasi, STMIK Kristen Neumann Indonesia, Medan, Indonesia
Email: [1]bertisari0@gmail.com, [2]mahdianta03@gmail.com, [3]hariantasembiring12345@gmail.com, [4,*]desibirink30@gmail.com
Email Penulis Korespondensi: bertisari0@gmail.com

Abstract−In this study the authors will analyze the comparison of the naïve Bayes classifier and decision tree methods in the classification of transaction data types of payments that are often made by customers where the method will analyze which model has the best percentage. The author uses the Kaggle deals payment data set. The data mining methods used to classify data are naïve Bayes classifier, decision tree, and rule based. For this study the Naïve Bayes Classifier method will be used. The results of the research on the accuracy of data classification using a decision tree have an accuracy value of 95.60%. where the predicted data yes with yes answers totaled 232 and answer no 17 with a class precision value of 93.17%. While the predictions for no with yes answers totaled 5 and for answers no totaled 246 with a class precision value of 93.17%. Based on the results of research using the naïve Bayes classifier and decision tree, it is possible to classify data on types of deals payment based on age ranges with different accuracies. From the percentage results, the decision tree method has the highest or best percentage with a value of 95.60%, while the Naïve Bayes classifier has a value of 92.20%.

Keywords: Naïve Bayes Classifier; Decision Trees; Classification

1. INTRODUCTION

Data mining or data mining in the digital era has become a strategic step for use in various fields. One area that uses data mining is sales and payment transactions [1].

The evolution of data cannot be separated from the current development of information technology which is capable of collecting very large amounts of data [2]. In addition to the ever-increasing data mining needs, various classification algorithms have emerged so that large amounts of data can be processed [3].

Classification is the process of creating a model or function that describes and characterizes the class of information. Classification is a form of analysis of information extraction models that classify information classes [1]. Classification is important in determining service class, because service class affects service quality [4]. The results of the classification process in service classes are used as a source of information for business needs and national policies in the field of data and communication technology [5].

In today's digital era, the convenience of data storage and data access is very important. The data received by the company is processed into information so that it becomes knowledge that can be used for planning, decision making, monitoring, evaluation [6]. To extract knowledge from data that has been processed, certain methods are needed so that this knowledge can be useful for users. The method used is data mining. Data mining is used to extract knowledge automatically from large data by looking for interesting patterns contained in the data [7].

In this study the authors will analyze the comparison of the naïve Bayes classifier and decision tree methods in the classification of transaction data types of payments that are often made by customers where the method will analyze which model has the best percentage [8]. The author uses the Kaggle deals payment data set.

The data mining methods used to classify data are naïve Bayes classifier, decision tree, and rule based. For this study the Naïve Bayes Classifier method will be used [9]. The advantage of using the Naïve Bayes classifier method compared to other methods is that it is a statistical classification method that can be used to predict the probability of membership of a class, besides that it is proven to have high accuracy and speed when applied to large databases [10].

Comparison of algorithms is a comparison of 2 or more algorithms to find out which algorithm is the best of these algorithms. Comparison of these classification algorithms to determine which classification produces the highest accuracy value based on the type of payment data [11].

Another algorithm that can be used to perform sentiment analysis is the Decision Tree. Decision Tree is a simple representation for classification, the process in the Decision Tree is to convert data in the form of tables into a tree or tree and then change the shape of the tree into rules or rules [12].

The Naive Bayes Algorithm is a method that can handle predictive factors in the next month and trends directly. The advantage of this algorithm is that it has a very good ability to estimate data that has trend patterns and predictions for the next month [13].

The Naïve Bayes classifier is a fairly simple probabilistic classification method. The calculation carried out in this method is on a set of opportunities by adding up the frequency and combination of values from the dataset owned. The Naïve Bayes Classifier method assumes that each variable in each category is independent of the other [14].

A tree is a data structure consisting of nodes and edges. The nodes in a tree are divided into three, namely root nodes (root/node), branch/internal nodes (branch/internal nodes) and leaf nodes (leaf nodes) [15]. A decision tree is a simple representation of a classification technique for a number of classes. thus, where both internal and root nodes are marked with attribute names, the edges are labeled with possible attribute values and leaf nodes are marked with different classes [10].

With the growth of n, the smoothed minimum depth of decision trees solving the problem of recognition deterministically is either bounded from above by a constant, or grows as a logarithm, or linearly [16]. For other cases (decision trees solving the problem of recognition nondeterministically, and decision trees solving the membership problem deterministically and nondeterministically), with the growth of n, the smoothed minimum depth of decision trees is either bounded from above by a constant or grows linearly [17].

The decision tree is a method with the basic concept of turning data into a decision tree and its rules [6]. The selected variable will result in a constraint with the same data, and a simple decision tree can be generated with fewer iterations. This decision tree is based on rules that aim to divide the number of populations with heterogeneous properties into more detailed and homogeneous characteristics [14].

In research (Eka Miranda, et al 2018) classifies customers into two classes, namely potential customer classes and non-potential customer classes using classification prediction attributes consisting of Occupation, Type of Payment, Tenor and Age [9]. The results of the study show that the Naïve Bayes Classifier has been able to classify customers into two classes, namely potential customers and non-potential customers with their respective accuracy values as follows: Sensitivity 97%, Specificity 99.8%, Precision 99.8%, Recall 97 %, Accuracy 97%, Error Rate 3% [2].

In the preprocessing stage before carrying out the data mining process to find out the amount of data per attribute then carry out the data mining process using the naïve Bayes algorithm, with a combined percentage split of 60 to 90 percent showing the results of each accuracy value of 80.137%, 78.0822%, 81, 5060%, and 83.5616%. the best in the percentage split combination of 90% with an accuracy percentage of 83.5616%. And the results of the evaluation using the confusion matrix show that the status of consumers who have the most and are superior to the employee class and student class is the Student Class [18].

Based on previous research where the naïve Bayes classifier and decision tree methods can be used in data classification, the authors will compare the two methods in classifying payment type transaction data.

2. RESEARCH METHODOLOGY

2.1 Research Stages

The implementation of this research is illustrated by using a fishbone flowchart as shown in the following figure.

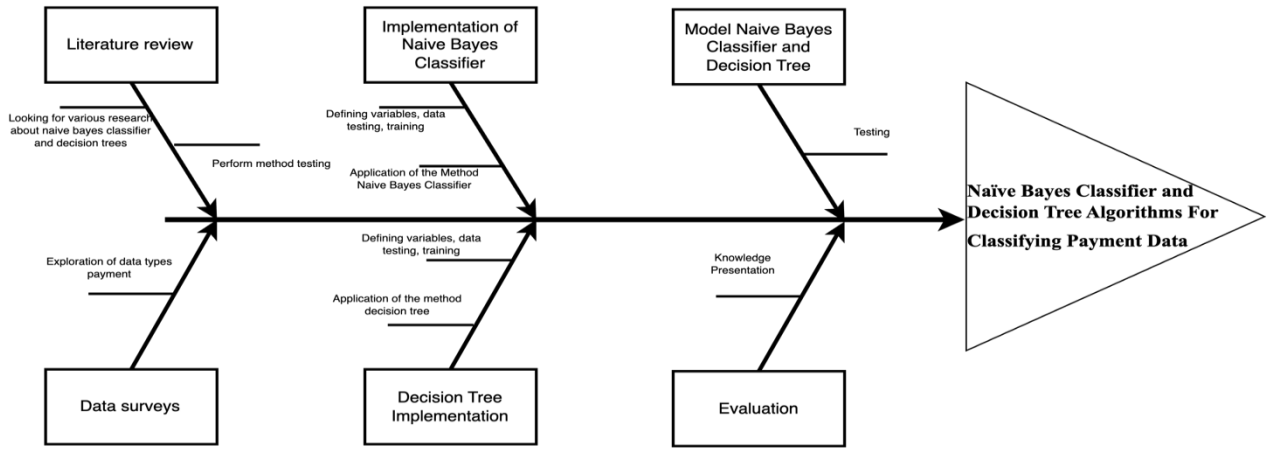

Figure 1. Research Fishbone Model

This framework is the steps that will be taken in order to solve the problem to be discussed. Figure 2. below is the framework (framework) used in this study.

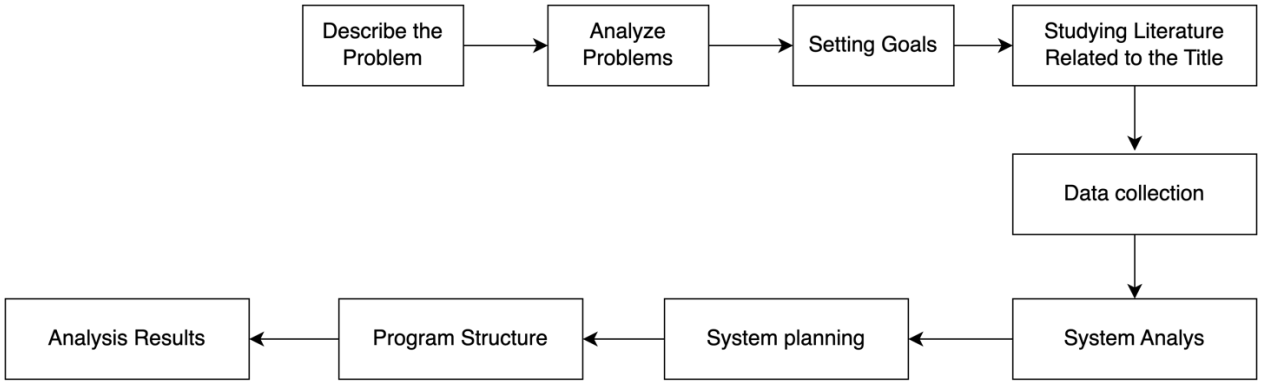

Figure 2. Research Procedure Framework

Based on the framework in Figure 2, each step can be described as follows:

1. Describe the Problem
   Describe the problem clearly to get the results of a comparative study of the naïve Bayes classifier and decision tree methods in the classification of payment type data.
2. Problem Analysis
   The problem analysis step is a step to understand the problem whose scope or boundaries have been determined. By analyzing the predetermined problem, it is hoped that the problem can be well understood.
3. Setting Goals
   Based on the understanding of the problems of the problems, the objectives to be achieved in this study were determined. This goal determines the targets to be achieved, especially those that can overcome existing problems.
4. Studying Literature Related to the Title
   To achieve the goal, then studied some of the literature that is expected to be used. Then the literature studied is selected which will be used in this study. Literature sources were obtained from books and journals.
5. Data Collection
   The data needed is data that will be used as material for research, namely data on types of payment deals from Kaggle.
6. System Analysis
   System analysis is quite important to do, because here the author must know the weaknesses of the system, obstacles, constraints and opportunities that are not able to be achieved by the current system in order to find alternative solutions to the problem.
7. System Design
   The system is designed using rapid miner tools to manage payment type data by looking at the accuracy and percentage of calculations for each method.
8. Program Structure
   Program Structure Design is a design that describes the relationship between a communication system with other communication systems.
9. Analysis Results
   At this stage it will provide the results of the research analysis of the naïve Bayes classifier and decision tree models which produce the percentage model of each method.

2.2 Naïve Bayes Classifier

The Naïve Bayes classifier is a fairly simple probabilistic classification method. The calculation carried out in this method is on a set of opportunities by adding up the frequency and combination of values from the dataset owned. The Naïve Bayes Classifier method assumes that each variable in each category is independent of the other [14].

Naive Bayes is one classic machine learning algorithm based on a Bayesian network, which is usually applied to classification problems and has excellent performance. When using naive Bayes for classification, from the perspective of probability, for the item to be classified, it needs to calculate the probability of the item appearing in each target category, then choose the most significant probability and find the corresponding category, as the classification result [19]

The Naïve Bayes Classifier algorithm has no relationship between one attribute and another, or in other words one attribute has no effect on other attributes, even though these attributes may be related [1].

Classification capabilities in Naïve Bayes are based on the Bayes theorem. In addition, the Naïve Bayes Classifier can also be said to be a method for making predictions in the future based on previous experience [14]. Based the Bayes theorem formula can be written as follows:

$P(H|X) = P(X|H).P(H)$
$P(X)$ ..... (1)

where:

X : unknown class data
H : hypothesis from data X, whose class specifications are known
$P(H|X)$: the probability of conjecture H based on state X (posterior probability)
$P(H)$: the probability of conjecture H (prior probability)
$P(X|H)$: the probability of X based on the state of the conjecture H
$P(X)$: the probability of X

2.3 Decision Tree

A tree is a data structure consisting of nodes and edges. The nodes in a tree are divided into three, namely the root node (root/node), branch/internal node (branch/internal node) and leaf node (leaf node) [20].

The decision tree is a simple representation of a classification technique for a finite number of classes, where internal nodes and root nodes are marked with attribute names, the edges are labeled with possible attribute values and leaf nodes are marked with different classes [20].

Decision Tree is an algorithm commonly used for decision making. The Decision Tree will look for solutions to problems by making criteria as nodes that are interconnected to form a tree-like structure. Decision tree is a predictive model for a decision using a hierarchical or tree structure. Each tree has branches, branches represent an attribute that must be fulfilled to go to the next branch until it ends in a leaf. The concept of data in the Decision Tree is data expressed in the form of a table consisting of attributes and records [21].

The decision tree is a method with the basic concept of turning data into a decision tree and its rules. The selected variable will result in a constraint with the same data, and a simple decision tree can be generated with fewer iterations. This decision tree is based on rules that aim to divide the number of populations with heterogeneous properties into more detailed and homogeneous characteristics [14].

3. RESULT AND DISCUSSION

The data mining process is carried out using naïve Bayes classification and decision trees because these algorithms can be used in classifying payment type data.

In this study two main research approaches were carried out, namely the qualitative and the qualitative approaches quantitative approach. A qualitative approach is used to analyze the literature review regarding variables that affect the selling price. While the quantitative approach is research the method used to examine a particular population or sample from a data set used for classification of sales prices [21]. The research methodology carried out consisted of 4 stages, which can be seen as follows:

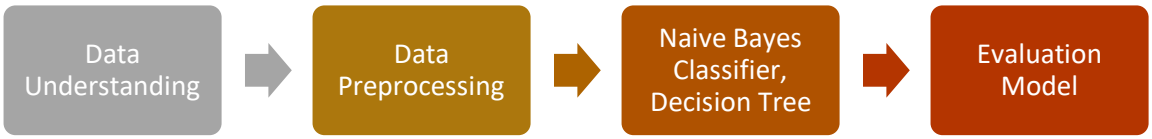

Figure 3. Research Methodology

3.1 Data Understanding

At this stage is data collection, analyzing data to understand the data to be used, identifying problems by understanding the substance in the data and looking for interesting things in the data. Broadly speaking, data understanding is used to check data so that we can identify problems in the data we get. This stage provides an analytical foundation for a study by making a summary and identifying potential problems in the data. This stage must also be carried out carefully and not in a hurry by data practitioners. Examples of data visualization performed by data practitioners. Usually, if you are not careful, insight or conclusions cannot be found immediately. The data in this study were obtained from the Kaggle deals dataset. Data deals for payment types with a total of 500 with 1 special attribute and 3 regular attributes.

Row No.	Future Cus...	Age	Gender	Payment M...
1	yes	44	male	credit card
2	no	86	female	credit card
3	yes	17	female	credit card
4	no	57	female	credit card
5	no	48	female	credit card
6	no	39	female	credit card
7	no	78	male	credit card
8	yes	38	male	credit card
9	no	80	male	cheque
10	yes	33	male	credit card
11	yes	54	male	credit card
12	no	53	male	cash
13	no	63	male	cash
14	no	91	male	cash
15	yes	17	female	credit card
16	yes	22	male	credit card

Figure 4. Deals Payment Type Data

3.2 Data Preprocessing

The next process before the algorithm model is made is data preprocessing. In this research, preprocessing techniques were used, namely: cleansing, data aggregation, checking for missing values.

The resulting data from Kaggle still contains redundancies which will interfere with the classification process, so preprocessing is needed to filter and clean. The following is the process of naïve bayes classifier and decision tree.

3.3 Naïve Bayes Classifier

When classifying in the naïve Bayes classification, training data is needed which will measure the data to be classified. Retrieving 500 Kaggle deals payment type data and labeling them manually to become training data. The training data will be classified into naïve Bayes classification and a model will be created and this model will automatically become test data and be given an automatic classification.

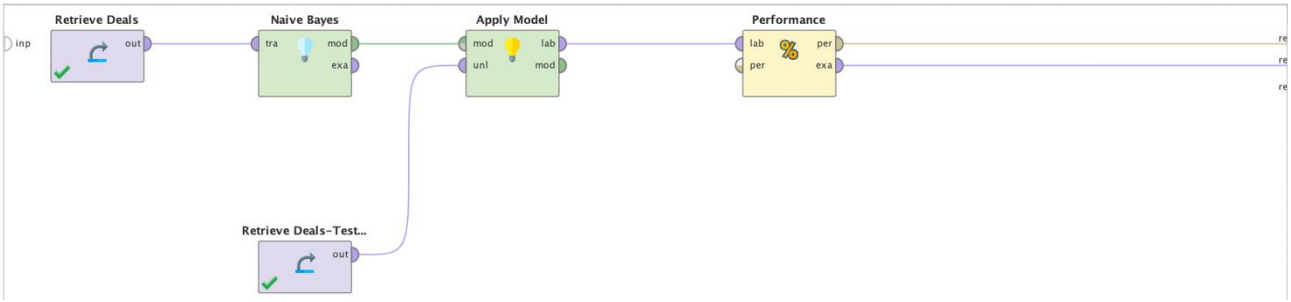

Figure 5. Naïve Bayes Classifier Process

In processing the classification using rapid miner tools by using the payment type deals dataset on Kaggle. The dataset will be processed naïve Bayes by applying the model and testing the performance of the results of the classification model. From the results of the training data and test data with the naïve Bayes classifier, a predictive data table and confidence value will be formed. So that a classification of payment types will be formed as shown in the following data.

Row No.	Future Cus...	prediction(...	confidence(...	confidence(...	Age	Gender	Payment M...
1	yes	yes	0.836	0.164	44	male	credit card
2	no	no	0.006	0.994	86	female	credit card
3	yes	yes	0.758	0.242	17	female	credit card
4	no	no	0.161	0.839	57	female	credit card
5	no	no	0.307	0.693	48	female	credit card
6	no	no	0.476	0.524	39	female	credit card
7	no	no	0.123	0.877	78	male	credit card
8	yes	yes	0.889	0.111	38	male	credit card
9	no	no	0.012	0.988	80	male	cheque
10	yes	yes	0.918	0.082	33	male	credit card
11	yes	yes	0.679	0.321	54	male	credit card
12	no	no	0.238	0.762	53	male	cash
13	no	no	0.102	0.898	63	male	cash
14	no	no	0.003	0.997	91	male	cash
15	yes	yes	0.758	0.242	17	female	credit card
16	yes	yes	0.954	0.046	22	male	credit card
17	yes	yes	0.758	0.242	17	female	credit card
18	no	no	0.161	0.839	57	female	credit card
19	yes	yes	0.947	0.053	25	male	credit card
20	no	no	0.271	0.729	50	female	credit card
21	no	no	0.368	0.632	46	male	cash
22	yes	yes	0.938	0.062	28	male	credit card

Figure 6. Results of Naïve Bayes Classifier Process Test and Training Data

The results of the test and training data where the prediction is true if the data is correct then the prediction is yes, as in the first row where the data is yes and the prediction results are yes with a confidence value of 0.836, then if the data is no then the prediction results are no which indicates conformity data, all results from the classification model are in accordance with the modeling process.

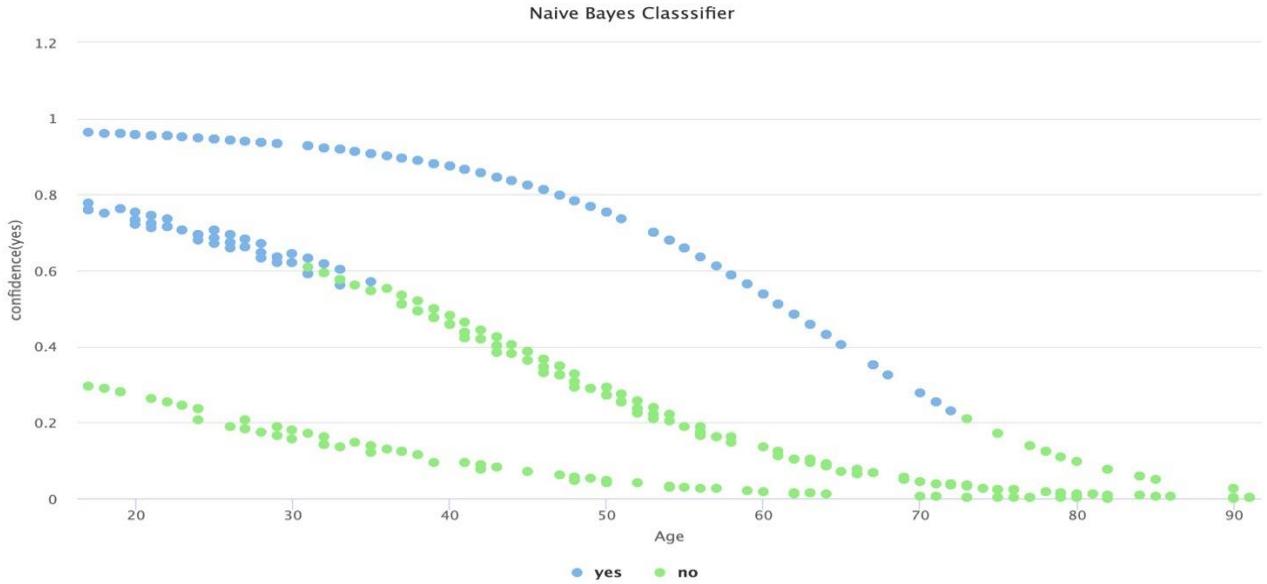

Figure 7. Graph of Classification Results of the Naïve Bayes Classifier Method

Furthermore, after the classification results with the naïve Bayes classifier, the following is a graph where the graphic results display data with the variable age, personal data on payments. In the graph, it can be seen that the age range of 20 – 40 makes payments while those over 50 have not made payment transactions so that the pattern has decreased on the chart.

accuracy: 92.20%

	true yes	true no	class precision
pred. yes	217	19	91.95%
pred. no	20	244	92.42%
class recall	91.56%	92.78%	

Figure 8. Naïve Bayes Classifier Accuracy Percentage

The results of the training and testing data will be seen for accuracy which has 92.20% accuracy obtained based on true and false values in the data after being processed for classification using the Naïve Bayes classification model. Where from the amount of data 500 with the predicted value of yes 217 and value no 19 with an accuracy class of 91.95%. While the prediction no with the answer yes 20 and no 244 with an accuracy of 92.42%.

3.4 Decision Tree

In the next process of classification with a decision tree where the training data process will be the measurement of the data to be classified. Retrieving 500 Kaggle deals payment type data and labeling them manually to become training data. The training data will be classified into a decision tree and a model will be created and this model will automatically become test data and be given an automatic classification.

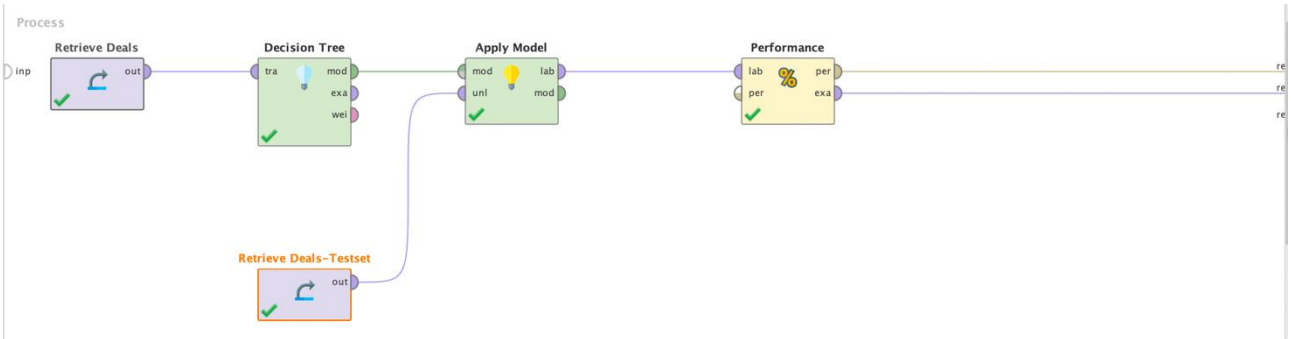

Figure 9. Picture of the Decision Tree Process

In the process using rapid miner tools which use data deals for the type of payment, then apply a decision tree using a model, then determine the performance of the classification.

Row No.	Future Cus...	prediction(...	confidence(...	confidence(...	Age	Gender	Payment M...
1	yes	yes	0.847	0.153	44	male	credit card
2	no	no	0	1	86	female	credit card
3	yes	yes	0.995	0.005	17	female	credit card
4	no	no	0	1	57	female	credit card
5	no	no	0	1	48	female	credit card
6	no	no	0	1	39	female	credit card
7	no	yes	0.847	0.153	78	male	credit card
8	yes	yes	0.847	0.153	38	male	credit card
9	no	no	0	1	80	male	cheque
10	yes	yes	0.847	0.153	33	male	credit card
11	yes	yes	0.847	0.153	54	male	credit card
12	no	no	0.068	0.932	53	male	cash
13	no	no	0.068	0.932	63	male	cash
14	no	no	0.068	0.932	91	male	cash
15	yes	yes	0.995	0.005	17	female	credit card
16	yes	yes	0.995	0.005	22	male	credit card

Figure 10. Results of Test Data and Decision Tree Process Training

The results of the data from the process of classifying the types of deals payment using 500 data, the results of the prediction data are yes if the data is yes and the prediction results are worth no if the data is no, from the first data it has a confidence value of 0.847, and 0.153.

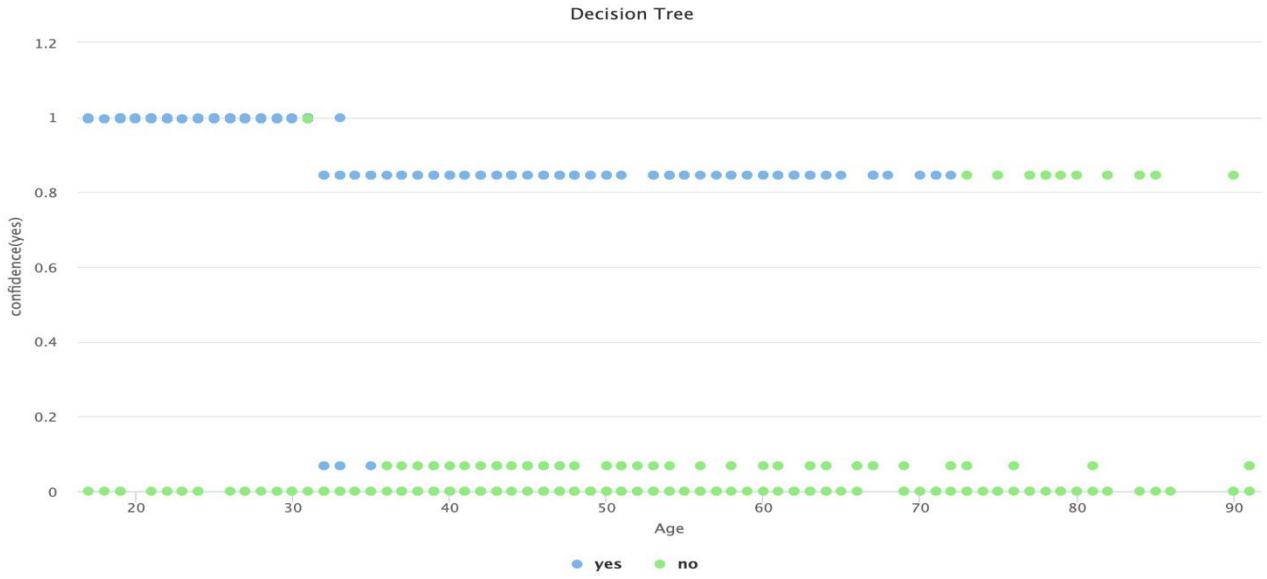

Figure 11. Graph of Classification Results of the Decision Tree Method

The results of the classification chart of the decision tree method are relatively balanced values based on the age variable data, on the graph the confidence data is with an interval of 0 to 0.2 with no answers, while yes answers are with intervals of 0.8 to 1.

accuracy: 95.60%

	true yes	true no	class precision
pred. yes	232	17	93.17%
pred. n pred. yes	5	246	98.01%
class recall	97.89%	93.54%	

Figure 11. Percentage Accuracy of the Decision Tree

The results of the accuracy of data classification using a decision tree has an accuracy value of 95.60%. where the predicted data yes with yes answers totaled 232 and answer no 17 with a class precision value of 93.17%. While the predictions for no with yes answers totaled 5 and for answers no totaled 246 with a class precision value of 93.17%.
Based on the results of research using the naïve Bayes classifier and decision tree, it is possible to classify data on types of deals payment based on age ranges with different accuracies. From the percentage results, the decision tree method has the highest or best percentage with a value of 95.60%, while the Naïve Bayes classifier has a value of 92.20%.

4. CONCLUSION

In this study the authors will analyze the comparison of the naïve Bayes classifier and decision tree methods in the classification of transaction data types of payments that are often made by customers where the method will analyze which model has the best percentage. The author uses the Kaggle deals payment data set. The data mining methods used to classify data are naïve Bayes classifier, decision tree, and rule based. For this study the Naïve Bayes Classifier method will be used. The advantage of using the Naïve Bayes classifier method compared to other methods is that it is a statistical classification method that can be used to predict the probability of membership of a class, besides that it is proven to have high accuracy and speed when applied to large databases. The results of the accuracy of data classification using a decision tree has an accuracy value of 95.60%. where the predicted data yes with yes answers totaled 232 and answer no 17 with a class precision value of 93.17%. While the predictions for no with yes answers totaled 5 and for answers no totaled 246 with a class precision value of 93.17%. Based on the results of research using the naïve Bayes classifier and decision tree, it is possible to classify data on types of deals payment based on age ranges with different accuracies. From the percentage results, the decision tree method has the highest or best percentage with a value of 95.60%, while the Naïve Bayes classifier has a value of 92.20%. This decision tree method has better accuracy than the naïve Bayes classifier. In the future it will be interesting for future researchers to be more specific in determining the imbalanced data set.

REFERENCES

[1] A. Zainal Macfud, A. Pandu Kusuma, and W. Dwi Puspitasari, "Analisis Algoritma Naive Bayes Classifier (NBC) Pada Klasifikasi Tingkat Minat Barang Di Toko Violet Cell," *J. Mhs. Tek. Inform.*, vol. 7, no. 1, pp. 1–8, 2023.

[2] E. Miranda and Julisar, "DATA MINING DENGAN METODE KLASIFIKASI NAÏVE BAYES UNTUK MENGKLASIFIKASIKAN PELANGGAN Eka Miranda , Julisar Program Sistem Informasi , Program Studi Sistem Informasi , Universitas Bina Nusantara," *Infotech*, vol. 4, no. 9, pp. 6–12, 2018.

[3] Yunita and I. H. Ikasari, "Perbandingan Metode Klasifikasi C4.5 dan Naïve Bayes untuk Mengukur Kepuasan Pelanggan," *Inform. Univ. Pamulang*, vol. 6, no. 3, pp. 2622–4615, 2021, [Online]. Available: http://openjournal.unpam.ac.id/index.php/informatika456.

[4] Y. Alkhalifi, A. Zumarniansyah, R. Ardianto, N. Hardi, and A. E. Augustia, "Comparison of Naive Bayes Algorithm and C.45 Algorithm in Classification of Poor Communities Receiving Non Cash Food Assistance in Wanasari Village Karawang Regency," *J. Techno Nusa Mandiri*, vol. 17, no. 1, pp. 37–42, 2020, doi: 10.33480/techno.v17i1.1191.

[5] N. Nurajijah and D. Riana, "Algoritma Naïve Bayes, Decision Tree, dan SVM untuk Klasifikasi Persetujuan Pembiayaan Nasabah Koperasi Syariah," *J. Teknol. dan Sist. Komput.*, vol. 7, no. 2, pp. 77–82, 2019, doi: 10.14710/jtsiskom.7.2.2019.77-82.

[6] T. Olsson, M. Ericsson, and A. Wingkvist, "To automatically map source code entities to architectural modules with Naive Bayes," *J. Syst. Softw.*, vol. 183, p. 111095, 2022, doi: 10.1016/j.jss.2021.111095.

[7] C. Aulia Sari, A. Sukmawati, R. Putri Aprilli, P. Sarah Kayaningtias, and N. Yudistira, "Perbandingan Metode Naïve Bayes, Support Vector Machine Dan Decision Tree Dalam Klasifikasi Konsumsi Obat," *J. Litbang Edusaintech*, vol. 3, no. 1, pp. 33–41, 2022, [Online]. Available: http://dx.doi.org/10.51402/jle.v3i1.47.

[8] T. A. Q. Putri, A. Triayudi, and R. T. Aldisa, "Implementasi Algoritma Decision Tree dan Naïve Bayes Untuk Klasifikasi Sentimen Terhadap Kepuasan Pelanggan Starbucks," *J. Inf. Syst. Res.*, vol. 4, no. 2, pp. 641–649, 2023, doi: 10.47065/josh.v4i2.2949.

[9] N. Deepa, J. Sathya Priya, and T. Devi, "Towards applying internet of things and machine learning for the risk prediction of COVID-19 in pandemic situation using Naive Bayes classifier for improving accuracy," *Mater. Today Proc.*, vol. 62, pp. 4795–4799, 2022, doi: 10.1016/j.matpr.2022.03.345.

[10] P. Khansong, J. Karnjana, S. Laitrakun, and S. Usanavasin, "Customer Service Improvement based on Electricity Payment Behaviors Analysis using Data Mining Approaches," *2020 Jt. Int. Conf. Digit. Arts, Media Technol. with ECTI North. Sect. Conf. Electr. Electron. Comput. Telecommun. Eng. ECTI DAMT NCON 2020*, pp. 114–117, 2020, doi: 10.1109/ECTIDAMTNCON48261.2020.9090699.

[11] I. B. A. Peling, I. N. Arnawan, I. P. A. Arthawan, and I. G. N. Janardana, "Implementation of Data Mining To Predict Period of Students Study Using Naive Bayes Algorithm," *Int. J. Eng. Emerg. Technol.*, vol. 2, no. 1, p. 53, 2017, doi: 10.24843/ijeet.2017.v02.i01.p11.

[12] C. M. S. Ramdani, A. N. Rachman, and R. Setiawan, "Comparison of the Multinomial Naive Bayes Algorithm and Decision Tree with the Application of AdaBoost in Sentiment Analysis Reviews PeduliLindungi Application," *Int. J. Inf. Syst. Technol. Akreditasi*, vol. 6, no. 4, pp. 419–430, 2022.

[13] H. D. Wijaya and S. Dwiasnati, "Implementasi Data Mining dengan Algoritma Naïve Bayes pada Penjualan Obat," *J. Inform.*, vol. 7, no. 1, pp. 1–7, 2020, doi: 10.31311/ji.v7i1.6203.

[14] A. Tanza and D. T. Utari, "Comparison of the Naïve Bayes Classifier and Decision Tree J48 for Credit Classification of Bank Customers," *EKSAKTA J. Sci. Data Anal.*, vol. 3, no. 2, pp. 70–77, 2022, doi: 10.20885/eksakta.vol3.iss2.art2.

[15] R. N. Whidhiasih, "Model Naïve Bayes dan Decision Tree untuk Prediksi Penerimaan Spesifikasi Mobil," *JREC (Journal Electr. ...*, vol. 10, no. 2, pp. 87–98, 2022, [Online]. Available:

https://jurnal.unismabekasi.ac.id/index.php/jrec/article/view/5600%0Ahttps://jurnal.unismabekasi.ac.id/index.php/jrec/article/download/5600/2394.

[16] H. Sastypratiwi, Yulianti, and H. Muhardi, "Uji Komparasi Algoritma Naïve Bayes dan Decision Tree Classification Menggunakan Covid-19 Dataset," *J. Edukasi dan Penelit. Inform.*, vol. 8, no. 1, pp. 1–6, 2022, doi: 10.26418/jp.v8i1.49841.

[17] M. Moshkov, "Decision trees for regular factorial languages," *Array*, vol. 15, no. January, p. 100203, 2022, doi: 10.1016/j.array.2022.100203.

[18] D. Syamsudin, Y. C. D. Halundaka, and A. Nugroho, "Prediksi Status Konsumen Produk Celana Menggunakan Naïve Bayes," *JOINTECS (Journal Inf. Technol. Comput. Sci.*, vol. 5, no. 3, p. 177, 2020, doi: 10.31328/jointecs.v5i3.1435.

[19] W. Guo, G. Wang, C. Wang, and Y. Wang, "Distribution network topology identification based on gradient boosting decision tree and attribute weighted naive Bayes," *Energy Reports*, vol. 9, pp. 727–736, 2023, doi: 10.1016/j.egyr.2023.04.256.

[20] A. H. Nasrullah, "Implementasi Algoritma Decision Tree Untuk Klasifikasi Produk Laris," *J. Ilm. Ilmu Komput.*, vol. 7, no. 2, pp. 45–51, 2021, doi: 10.35329/jiik.v7i2.203.

[21] P. H. Putra, B. Purba, and Y. A. Dalimunthe, "Random forest and decision tree algorithms for car price prediction," vol. 1, no. 2, pp. 81–89, 2023.