SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 2, April 2024
DOI : https://doi.org/10.33395/sinkron.v8i2.13637
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Analysis of Public Interest in Automatic Motorcycles Using KNN and Neural Network Methods

Nurul Fatma[1)*], Syaiful Zuhri Harahap[2)], Masrizal[3)]
[1)2)3)] Universitas Labuhanbatu, Indonesia
[1]nurul2018fatma@gmail.com, [2)] syaifulzuhriharahap@gmail.com, [3)] masrizal120405@gmail.com,

Submitted :April 4, 2024 | **Accepted** : April 24, 2024 | **Published** : April 25, 2024

Abstract: Research on public interest in automatic motorbikes was carried out with the aim of understanding the factors that influence the decision to purchase an automatic motorbike. Using data mining methods, this research applies the K-Nearest Neighbor (KNN) and Neural Network techniques to identify and analyze people's interest patterns. The data used amounted to 139 samples, of which 127 showed interest in automatic motorbikes, while 12 others showed no interest. The research process begins with data analysis, the next stage is preprocessing, which includes data cleaning, in the model design stage in data mining, two models are built: one using KNN and the other using Neural Network. These two models are designed to classify sample data based on interest in automatic motorbikes. The next stage is model testing. Test results show that both models can classify interests accurately, with most of the sample data being classified correctly. Model evaluation was carried out to measure the effectiveness and accuracy of the two methods. The evaluation results show that both models provide very good performance, with results that almost reach a perfect score. This shows that both methods, KNN and Neural Network, are very effective in classifying and predicting people's interest in automatic motorbikes based on available data. In conclusion, this research not only shows the effectiveness of KNN and Neural Network in data mining for analyzing people's interests, but also provides valuable insights for automatic motorbike manufacturers and sellers about consumer preferences.

Keywords: Classification; Confusion Matrix; Data Mining; K-Nearest Neighbor (KNN); Neural Network;

INTRODUCTION

Motorbikes have become one of the most common and efficient means of land transportation used by many people throughout the world. Ease of operation, compact size, and ability to bypass traffic jams make motorbikes a very popular choice for daily commuting or going somewhere. In addition, the relatively low operational and maintenance costs compared to four-wheeled vehicles make motorbikes an economical alternative for urban and rural mobility. The use of motorbikes is not only limited to daily activities such as going to work or school, but has also become an integral part of various business sectors. Goods and service delivery services, for example, often rely on motorbikes because of their ability to deliver goods quickly and efficiently, especially in areas with heavy traffic. The tourism industry is also utilizing motorbikes as a way to offer unique and personalized travel experiences to tourists, allowing them to explore destinations more freely and in depth.

The number of motorbikes worldwide has increased significantly in recent decades, in line with economic growth and the need for efficient transportation in both urban and rural areas. Factors such as ease of access, ability to overcome traffic jams, and affordable operating costs have strengthened the position of motorbikes as one of the main choices for individual mobility. This increase is also driven by the diversification of products offered by manufacturers, meeting various consumer needs and preferences, ranging from motorbikes for daily activities to special models for hobbies and sports. In the midst of this growth trend, the motorbike industry has done a lot innovation, especially in the development of automatic (automatic) motorbikes which have revolutionized the way of riding. Automatic motorbikes offer ease of operation with an automatic transmission system that eliminates the need to change gears manually, making them very popular among new riders and those looking for maximum comfort when riding (Hidup, Produk, Citra, & Terhadap, 2023). Additionally, motorcycle manufacturers continue to integrate advanced technologies, such as ABS braking systems, traction control, and even smartphone

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 2, April 2024
DOI : https://doi.org/10.33395/sinkron.v8i2.13637
e-ISSN : 2541-2019
p-ISSN : 2541-044X

connectivity, improving safety, comfort, and the overall riding experience. This innovation not only addresses increasing market demands for safer and more comfortable vehicles but also demonstrates the industry's commitment to continuous improvements in vehicle technology. Automatic motorbikes are also very popular among teenagers, especially women, this is because automatic motorbikes have a design that suits their character (Amdhani et al., 2022).

Even though automatic (automatic) motorbikes have gained wide popularity among riders in various parts of the world due to their ease of use, there are still some people who question their effectiveness, especially for use in rural areas (Maulana, Mentary, & Muhibuddin, 2023). This skepticism is often rooted in the perception that automatic motorbikes are less resilient than manual motorbikes in dealing with rough and poorly maintained road conditions, which are generally found in rural areas. Apart from that, concerns regarding the durability and possibly higher maintenance costs of automatic motorbikes compared to manual motorbikes are also reasons why some people prefer to avoid using automatic motorbikes in areas with limited access to maintenance services and spare parts. On the other hand, most people consider automatic motorbikes to be ideal for daily use, especially in urban environments. Automatic motorbikes offer significant ease of operation because the rider does not have to worry about manually adjusting the transmission; they only need to accelerate to move and use the brakes to stop. This convenience makes automatic motorbikes a very popular choice among riders of all ages and levels of riding experience. In addition, the latest innovations in automatic motorbike design and technology, such as increased fuel efficiency, better braking systems and additional safety features, further strengthen the position of automatic motorbikes as comfortable, economical and safe vehicles for daily use in urban areas.

In order to explore more deeply about transportation preferences among the community, especially related to the use of automatic motorbikes, the author plans to conduct research that focuses on people's interest in automatic motorbikes in Sei Rumbiah Village. This village was chosen as a research location because of its unique representation as a rural area that has the potential to experience changes in the dynamics of vehicle use, from the dominance of manual motorbikes to the possibility of increasing interest in automatic motorbikes. This research aims to identify how much interest people have in automatic motorbikes, understand the factors that influence their preferences, and explore the potential influence of the rural environment on the decision to use these vehicles. To obtain accurate and in-depth results, this research will be carried out using data mining techniques, an approach that allows researchers to process and analyze large data sets to discover patterns and trends that are not visible at first glance. Two methods with classification models will be applied: the Nearest Neighbor (KNN) method and the Neural Network. KNN was chosen because of its extraordinary ability to classify data based on the closeness of its characteristics to known data samples, while Neural Network, with a structure and information processing process that resembles the human brain.

Through the combination of these two methods, it is hoped that we can reveal strong insight into the interest of the people in Sei Rumbiah Village towards automatic motorbikes, including the factors that influence their interest. The results of this research will not only provide an overview of current trends in the village but can also provide useful information for motorbike manufacturers, the government, and other stakeholders in developing appropriate strategies to meet land transportation needs in Sei Rumbiah Village in particular, and in rural areas in general. Apart from that, this research is also expected to contribute to academic literature related to vehicle use in rural areas and the application of data mining in transportation studies.

LITERATURE REVIEW

K-Nearest Neighbor (KNN) is a machine learning method that is included in the supervised learning category. This method works based on the principle of proximity between data. In KNN, data will be classified based on how often the labels of its k-nearest neighbors appear (Violita, Yanris, & Hasibuan, 2023) (Istiadi, Rahman, & Wisnu, 2023). K here is a parameter that determines the number of nearest neighbors to be considered (Adjani, Fauzia, & Juliane, 2023) (Arifuddin, Pinastawa, Anugraha, & Pradana, 2023). KNN is very effective for datasets that have a small number of dimensions (features), but can become less efficient as the number of dimensions increases due to the complexity of calculating the distance between data points (Hermawan & Prianggono, 2023) (Triani, Dar, & Yanris, 2023). KNN does not require a learning model in the initial phase, but requires more time in the testing phase because it needs to calculate the distance from new data points to all other data points that already exist in the dataset (A. W. Sari, Hermanto, & Defriani, 2023). This method is widely used in various practical applications such as product recommendations, document classification, and pattern recognition (Zai, Sirait, Nainggolan, Sihombing, & Banjarnahor, 2023).

Neural Network, or neural network, is a paradigm in machine learning inspired by human biological neural networks (M. Sari, Yanris, & Hasibuan, 2023) (Suherman, Hindarto, Makmur, & Santoso, 2023). This method consists of neurons that are connected to each other and can process information in a way that resembles the human brain. Neural networks are very effective in identifying complex and abstract patterns that may not be recognized by other machine learning methods (Lestari, Mawengkang, & Situmorang, 2023). The basic architecture of a neural network includes three main layers: input layer, hidden layer, and output layer. Each neuron in one layer is connected to all the neurons in the next layer, and the strength of these connections (called weights) can be resolved

*name of corresponding author

during the learning or training process (Anwar, Jalinus, & Abdullah, 2023). This method requires a lot of data for training to be effective, but once trained, the neural network is able to make predictions and classifications quickly and accurately (Hindarto, 2023).

METHOD

In this research, there are several stages carried out so that the data can be classified properly and correctly. The stages that will be carried out are as follows.

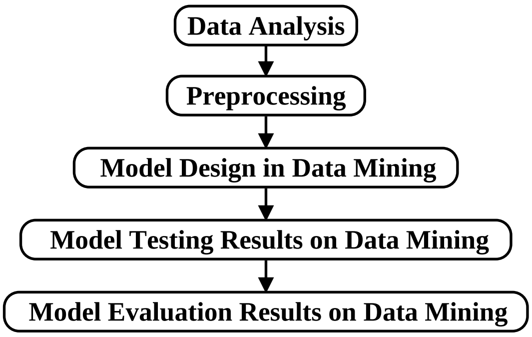

Fig 1. Research Workflow

In the picture above is the workflow for the stages of conducting research on public interest in automatic motorbikes. An explanation of the flow of the stages is as follows.

Data Analysis	:	Data analysis is a data collection process carried out to obtain data sets that will be used to conduct this research.
Preprocessing	:	This process is a process carried out to clean data from data that is not suitable for use and in this process the data will also be arranged in a good form and format.
Model Design in Data Mining	:	This process is the process of designing a model that will be used to classify data.
Model Testing Result on Data Mining	:	At this stage, it is the result of data classification that has been carried out previously with a model that has been designed.
Model Evaluation Result on Data Mining	:	This process is the result of a method evaluation carried out to see the capabilities of the method used.

Confusion Matrix

Table 1. Confusion Matrix

	Prediction Class		
Attribute Class	**Class**	**True**	**False**
	True	True Positive (TP)	False Positive (FP)
	False	False Negative (FN)	True Negative (TN)

From the table above for the explanation below.

1. TP (True Positive), namely the amount of positive data that has a true value.
2. TN (True Negative), namely the amount of negative data that has a true value.
3. FN (False Negative), namely the amount of negative data but which has the wrong value.
4. FP (False Positive), namely the amount of data that is positive but has the wrong value.

$$\boldsymbol{Acuracy = \frac{TP+TN}{TP+TN+FN+FP} \times 100\%}$$ (Siregar, Irmayani, & Sari, 2023)

$$\boldsymbol{Precision = \frac{TP}{TP+FP} \times 100\%}$$ (Mawaddah, Dar, & Yanris, 2023)

$$\boldsymbol{Recall = \frac{TP}{TP+FN} \times 100\%}$$ (Hasibuan, Dar, & Yanris, 2023)

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 2, April 2024
DOI : https://doi.org/10.33395/sinkron.v8i2.13637
e-ISSN : 2541-2019
p-ISSN : 2541-044X

RESULT

Data Analysis

Data analysis is a stage carried out by the author to determine the data that will be used as training data and as testing data. For the data that has been obtained by the author, there are 139 community sample data. So the author will divide the data into 2 data sets. This is done so that the classification process can be carried out in data mining. The attributes that the author uses in the data set are full name, fuel efficiency, easy of use, design and aesthetics, engine performance, selling price, storage space, driving comfort. The features used are in the orange application. For features, there are files, data tables, KNN and Neural Network methods, predictions. This is a feature or widget that will be used to classify data. To evaluate the method, several features are used, namely files, KNN and neural network methods, test and score, confusion matrix and ROC analysis.

Data Training

Training data is data that is used to help the data classification process. For training data which will be taken from a total of 178 data, the author will take 39 data which will be used in the training data. But in the table below, the author did not include all the data. The table above is only a sample of training data to find out the form of training data that will be used.

Table 2
Data Training

Full name	Fuel Efficiency	Ease of Use	Design and Aesthetics	Engine Performance	Selling price	Storage Space	Driving Comfort	Category
Andriani	Fuel Efficient	Easy to use	Good	Good	Expensive	Wide	Comfortable	Interest
Anzali Mukti	Fuel Efficient	Easy to use	Good	Good	Cheap	Wide	Comfortable	Interest
Ari Syahputra	Fuel Efficient	Easy to use	Good	Good	Affordable	Wide	Comfortable	Interest
Bayu Aji Anggara	Fuel Efficient	Easy to use	Good	Good	Affordable	Wide	Comfortable	Interest
Deni Suryanti	Wastes fuel	Easy to use	Just the same	Not Good	Expensive	Wide	Uncomfortable	Not interested
Edi Syahputra	Fuel Efficient	Easy to use	Just the same	Good	Affordable	Wide	Comfortable	Interest
Hendra	Fuel Efficient	Easy to use	Good	Good	Affordable	Wide	Comfortable	Interest
Nur Hasanah	Fuel Efficient	Easy to use	Not Good	Not Good	Expensive	Wide	Uncomfortable	Not interested
Pra Purna Sinta	Fuel Efficient	Easy to use	Good	Good	Cheap	Wide	Comfortable	Interest
Rini Antika Harahap	Fuel Efficient	Easy to use	Good	Good	Cheap	Small	Comfortable	Interest

Data Testing

Testing data is research sample data or can be interpreted as data that will be classified in data mining using a predetermined method. For testing data, 139 data were used. Of the 178 data obtained, 39 data have been included in the training data, then the remaining 139 data will be included in the testing data. The testing data is also almost the same as the training data. The data that the author presents is not all existing data, but only a sample of data.

Table 3
Data Testing

Full name	Fuel Efficiency	Ease of Use	Design and Aesthetics	Engine Performance	Selling price	Storage Space	Driving Comfort
Abdul Rahmad	Fuel Efficient	Easy to use	Just the same	Good	Affordable	Wide	Comfortable
Ade Irma Rambe	Wastes fuel	Easy to use	Just the same	Not Good	Expensive	Wide	Uncomfortable
Ade Pertama	Fuel Efficient	Easy to use	Good	Good	Cheap	Small	Just the same
Agustini Ritonga	Fuel Efficient	Easy to use	Good	Good	Affordable	Wide	Just the same
Aisyah Hasibuan	Fuel Efficient	Easy to use	Good	Good	Affordable	Small	Comfortable
Aisyah Siregar	Fuel Efficient	Easy to use	Good	Good	Cheap	Wide	Comfortable
Aisyah Zahrani	Fuel Efficient	Easy to use	Good	Good	Cheap	Wide	Comfortable
Alfi Khairi Siregar	Fuel Efficient	Easy to use	Good	Good	Affordable	Small	Comfortable
Ali Sofyan Rambe	Fuel Efficient	Easy to use	Good	Good	Affordable	Wide	Uncomfortable
Alifahri Mulyalza	Fuel Efficient	Easy to use	Good	Good	Cheap	Wide	Comfortable
Almuniroh Siregar	Fuel Efficient	Easy to use	Good	Good	Expensive	Wide	Comfortable
Andika Pratama	Fuel Efficient	Easy to use	Good	Good	Cheap	Wide	Comfortable
Andini Fitria	Wastes fuel	Difficult to Use	Not Good	Not Good	Affordable	There isn't any	Just the same
Anggun Khairani	Fuel Efficient	Easy to use	Just the same	Good	Affordable	Wide	Comfortable

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 2, April 2024
DOI : https://doi.org/10.33395/sinkron.v8i2.13637
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Anhar Harahap	Fuel Efficient	Easy to use	Just the same	Good	Cheap	Small	Comfortable
Anita Wati	Fuel Efficient	Easy to use	Good	Good	Cheap	Wide	Comfortable
Anjeli Hasibuan	Fuel Efficient	Easy to use	Just the same	Good	Cheap	Small	Comfortable
Annisa Ritonga	Fuel Efficient	Easy to use	Good	Good	Affordable	Small	Comfortable
Ardiansyah	Wastes fuel	Difficult to Use	Not Good	Not Good	Affordable	Small	Uncomfortable
Ardiansyah	Wastes fuel	Difficult to Use	Not Good	Not Good	Affordable	Small	Uncomfortable
Aria Swindo	Fuel Efficient	Easy to use	Good	Good	Cheap	There isn't any	Uncomfortable
Arini	Fuel Efficient	Easy to use	Good	Good	Cheap	Wide	Comfortable
Arrafkha Mulyalza	Fuel Efficient	Easy to use	Good	Good	Cheap	Wide	Comfortable
Aswet Faizi	Wastes fuel	Difficult to Use	Not Good	Not Good	Expensive	Small	Uncomfortable

In the table above is the data that will be used to classify data in data mining using the KNN and Neural Network methods. The above data was obtained by distributing questionnaires to the Rumbiah village community.

Preprocessing

At the data preprocessing stage, the data that has been obtained, whether training data or testing data, will be selected for suitability, then the data will be cleaned and arranged in an appropriate format. For this research, the data that researchers will use is data in file format.xlsx. The cleaning process means, for example, if there is data that has attributes that are not filled in or there is data that does not have a name, this data will be deleted from the data set that will be used. So the data that will be taken is data that is in accordance with research needs.

Model Design in Data Mining

For data mining model design, it will be designed and formed using the Orange application. The methods used are the KNN and Neural Network methods.

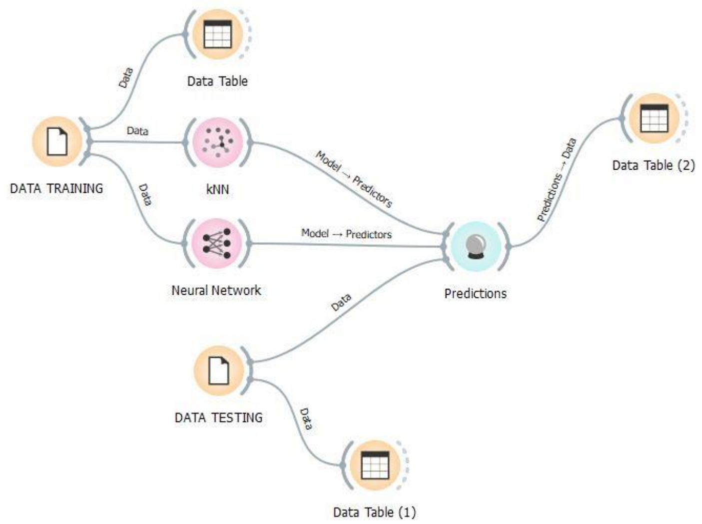

Fig 2. Classification Design Model in Data Mining

In the image above is the model used by the author to classify data in data mining. There are 2 methods used in the model above, namely the K-Nearest Neighbor (KNN) Method and the Neural Network Method. So in this research, the training data will be entered into the knn and neural network methods, then with the model, the testing data will be connected to the training data with the help of a prediction widget. This is done to train testing data so that it can be classified.

Model Testing Results in Data Mining

The model testing results are the results obtained from a previously designed model which is used to classify data in data mining using the KNN and neural network methods. The test results can be seen in the table below.

Table 4
Data Classification Results in Data Mining

Full name	Fuel Efficiency	Ease of Use	Design and Aesthetics	Engine Performance	Selling price	Storage Space	Driving Comfort	Category
Abdul Rahmad	Fuel Efficient	Easy to use	Just the same	Good	Affordable	Wide	Comfortable	Interest
Ade Irma Rambe	Wastes fuel	Easy to use	Just the same	Not Good	Expensive	Wide	Uncomfortable	Not interested
Ade Pertama	Fuel Efficient	Easy to use	Good	Good	Cheap	Small	Just the same	Interest
Agustini Ritonga	Fuel Efficient	Easy to use	Good	Good	Affordable	Wide	Just the same	Interest

*name of corresponding author

Aisyah Hasibuan	Fuel Efficient	Easy to use	Good	Good	Affordable	Small	Comfortable	Interest
Aisyah Siregar	Fuel Efficient	Easy to use	Good	Good	Cheap	Wide	Comfortable	Interest
Aisyah Zahrani	Fuel Efficient	Easy to use	Good	Good	Cheap	Wide	Comfortable	Interest
Alfi Khairi Siregar	Fuel Efficient	Easy to use	Good	Good	Affordable	Small	Comfortable	Interest
Ali Sofyan Rambe	Fuel Efficient	Easy to use	Good	Good	Affordable	Wide	Uncomfortable	Interest
Alifahri Mulyalza	Fuel Efficient	Easy to use	Good	Good	Cheap	Wide	Comfortable	Interest
Almuniroh Siregar	Fuel Efficient	Easy to use	Good	Good	Expensive	Wide	Comfortable	Interest
Andika Pratama	Fuel Efficient	Easy to use	Good	Good	Cheap	Wide	Comfortable	Interest
Andini Fitria	Wastes fuel	Difficult to Use	Not Good	Not Good	Affordable	There isn't any	Just the same	Interest
Anggun Khairani	Fuel Efficient	Easy to use	Just the same	Good	Affordable	Wide	Comfortable	Interest
Anhar Harahap	Fuel Efficient	Easy to use	Just the same	Good	Cheap	Small	Comfortable	Interest
Anita Wati	Fuel Efficient	Easy to use	Good	Good	Cheap	Wide	Comfortable	Interest
Anjeli Hasibuan	Fuel Efficient	Easy to use	Just the same	Good	Cheap	Small	Comfortable	Interest
Annisa Ritonga	Fuel Efficient	Easy to use	Good	Good	Affordable	Small	Comfortable	Interest
Ardiansyah	Wastes fuel	Difficult to Use	Not Good	Not Good	Affordable	Small	Uncomfortable	Interest
Ardiansyah	Wastes fuel	Difficult to Use	Not Good	Not Good	Affordable	Small	Uncomfortable	Interest
Aria Swindo	Fuel Efficient	Easy to use	Good	Good	Cheap	There isn't any	Uncomfortable	Interest
Arini	Fuel Efficient	Easy to use	Good	Good	Cheap	Wide	Comfortable	Interest
Arrafkha Mulyalza	Fuel Efficient	Easy to use	Good	Good	Cheap	Wide	Comfortable	Interest
Aswet Faizi	Wastes fuel	Difficult to Use	Not Good	Not Good	Expensive	Small	Uncomfortable	Not interested

In the table above are the predicted results of the classification of public interest in automatic motorbikes. Of the 139 community data samples, 127 people's data are interested in automatic motorbikes and 12 people's data are not interested in automatic motorbikes. The data above is only a sample, this was done so that the results of the research carried out can be seen.

Model Evaluation Results in Data Mining

Evaluation in the context of data mining is a vital process that aims to test the effectiveness of methods in carrying out classification. Through evaluation, the ability of classification methods such as k-Nearest Neighbors (kNN) and Neural Network can be measured more objectively. This evaluation process not only helps in identifying the strengths and weaknesses of each method in dealing with various types of data and classification situations, but also allows developers and researchers to make adjustments to strategies or algorithms to improve the accuracy and efficiency of the classification models used. Especially in data mining, evaluations like this are critical in guiding data-based decision making, thereby maximizing the value and insights that can be drawn from large data sets. There is a model that will be used to carry out this evaluation, the model will be designed in the orange application. The model is as follows. The K value used in this research is not numerical, but the K value used in this research is categorical. So from the previous classification results the K value is Interest. However, for this evaluation, the k value used by the author is 90. So the 3 evaluation values carried out by the author, the accuracy results obtained must be close to or more than ilia K, which means that the KNN method can be properly used to classify data.

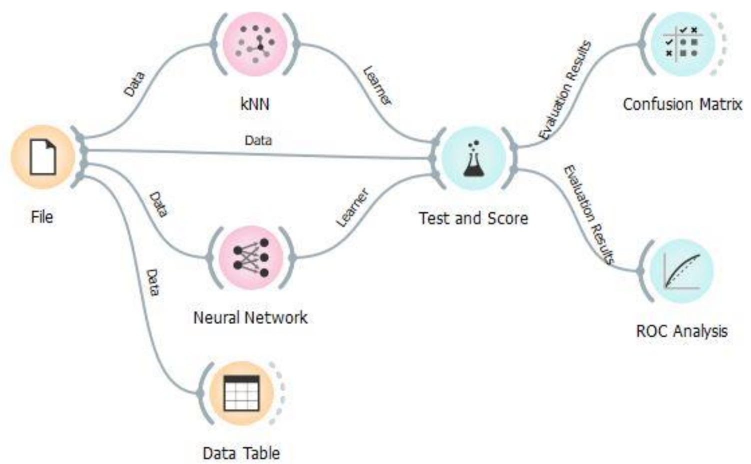

Fig 3. Evaluation Design Model in Data Mining

In the image above is the model design used to evaluate the effectiveness of the KNN and neural network methods. For this evaluation, the author used the test and score widget and the confusion matrix widget.

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 2, April 2024
DOI : https://doi.org/10.33395/sinkron.v8i2.13637
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Test and Score

Table 5. Result of Test and Score

Model	AUC	CA	F1	Precision	Recall	MCC
kNN	0.987	0.971	0.970	0.970	0.971	0.807
Neural Network	1.000	0.993	0.993	0.993	0.993	0.954

In the table above are the accuracy results obtained from the test and score widget. The results above show that the accuracy obtained between the KNN method and the neural network is that the accuracy value obtained by the neural network method is greater than the accuracy value obtained by the KNN method. But the results of both methods are equally good.

Confusion Matrix

The confusion matrix results obtained contained 2 results from different methods, the results are as follows.

Confusion Matrix Results of the K-Nearest Neighbor (KNN) Method

Table 6. Confusion Matrix Results in the K-Nearest Neighbor (KNN) Method

		Predicted		
		Interest	Not Interested	∑
Actual	Interest	126	1	42
	Not Interested	3	9	50
	∑	42	50	92

In the table above is the composition of the confusion matrix results obtained from evaluating the KNN method in data mining. For the results, there are True Positive (TP) results which are 126 data, for True Negative (TN) results which are 9 data, for False Positive (FP) results which are 1 and for False Negative (FN) results which are 3 data. For these results, it is not possible to directly measure the accuracy value, the data above must be calculated first using the formula in the confusion matrix, which is as follows.

$$\boldsymbol{Accuracy} = \frac{\mathbf{126+9}}{\mathbf{126+9+1+3}} + \mathbf{100\%}$$ Then the Accuracy value = 97%

$$\boldsymbol{Precision} = \frac{\mathbf{126}}{\mathbf{126+1}} + \mathbf{100\%}$$ Then the Precision value = 99%

$$\boldsymbol{Recall} = \frac{\mathbf{126}}{\mathbf{126+3}} + \mathbf{100\%}$$ Then the Recall value = 97%

Confusion Matrix Results Neural Network Method

Table 7. Confusion Matrix Results in the Neural Network Method

		Predicted		
		Interest	Not Interested	∑
Actual	Interest	127	0	42
	Not Interested	1	11	50
	∑	42	50	92

In the table above is the composition of the confusion matrix results obtained from evaluating the KNN method in data mining. For the results, there are True Positive (TP) results which are 127 data, for True Negative (TN) results which are 11 data, for False Positive (FP) results which are 0 and for False Negative (FN) results which are 1 data. For these results, it is not possible to directly measure the accuracy value, the data above must be calculated first using the formula in the confusion matrix, which is as follows.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 2, April 2024
DOI : https://doi.org/10.33395/sinkron.v8i2.13637
e-ISSN : 2541-2019
p-ISSN : 2541-044X

$\boldsymbol{Accuracy = \frac{127+11}{127+11+0+1} + 100\%}$ Then the Accuracy value = 99%

$\boldsymbol{Precision = \frac{127}{127+0} + 100\%}$ Then the Precision value = 100%

$\boldsymbol{Recall = \frac{127}{127+1} + 100\%}$ Then the Recall value = 99%

ROC Analysis

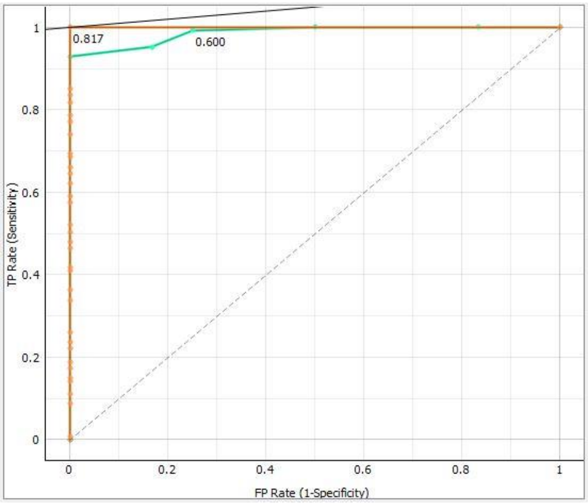

Fig 4. ROC Results Analysis of Community Interest in Automatic Motorbikes

In Figure 4 above are the results of people who are interested in automatic motorbikes using the KNN and neural network methods. The graphic results above show that there are still many people who are interested in using automatic motorbikes.

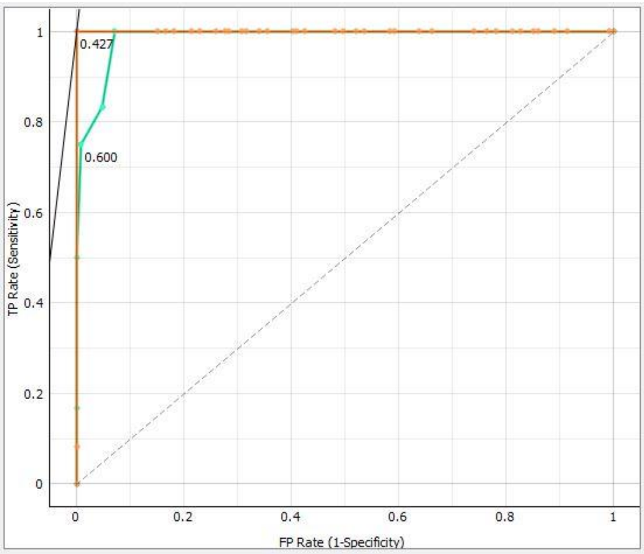

Fig 5. ROC Analysis Results: People Are Not Interested in Automatic Motorbikes

In Figure 5 above are the results of people who are not interested in automatic motorbikes using the KNN and neural network methods. The graphic results above show that not many people are interested in automatic motorbikes. It can be seen in the orange graph line that the neural network method results in more people being interested in automatic motorbikes. For the knn method, the results obtained from this graphic analysis are balanced and can be seen on the green graphic line.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 2, April 2024 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v8i2.13637 p-ISSN : 2541-044X

The two images above are the results of ROC Analysis from the results of the evaluation method used. The results for the knn method are shown in a green graph and for the neural network method they are shown in an orange graph.

DISCUSSIONS

Research on public interest in automatic motorbikes using a data mining approach shows very impressive results. This research applies two main methods in data mining, namely K-Nearest Neighbor (KNN) and Neural Network, to analyze data and predict community interests. These two methods are known for their high ability to handle large and complex data. To assess and compare the effectiveness of the two methods, researchers used several tools in the analysis process. The Test and Score and Confusion Matrix widgets are the main tools for measuring the accuracy of the model being built. From the results obtained through the Test and Score widget, the KNN method shows an accuracy of 98%, while the Neural Network method achieves perfect accuracy, namely 100%. Meanwhile, based on the Confusion Matrix, the accuracy for KNN is 97% and Neural Network is 99%. A comparison of these two methods shows that they both give almost the same results, it's just that the difference is not very significant, namely only a few percent. This shows that both methods are very effective and reliable in classifying data to understand people's preferences for automatic motorbikes. Even though there is a small difference in accuracy values, these two methods are equally suitable for use in similar research or other practical applications in the field of data mining. However, the Neural Network method is superior to the KNN method, this is because the accuracy value obtained by the Neural Network method is higher than the accuracy value obtained by the KNN method, even though it is only 2% different. This can be seen in the accuracy results of the confusion matrix widget.

CONCLUSION

Research on public interest in automatic motorbikes using data mining methods is very relevant in the context of product development and marketing strategies in the automotive industry. In this research, the K-Nearest Neighbor (KNN) and Neural Network techniques are used to classify people's interest in automatic motorbikes. This study collected data from 139 respondents to assess their interest in automatic motorbikes. By using the designed model, the results show that the majority, namely 127 people's data, have an interest in automatic motorbikes, while 12 other people's data have no interest in automatic motorbikes. The KNN and Neural Network models applied in this research succeeded in showing very high classification performance. The KNN method succeeded in achieving an accuracy of 97%, while the Neural Network method achieved a higher accuracy, namely 99%. This almost perfect accuracy shows the effectiveness of these two methods in analyzing and predicting consumer preferences for automatic motorbikes. These results are important for automatic motorbike manufacturers and marketers because they provide insight into potential market proportions and consumer preferences. Apart from that, the success of the KNN and Neural Network methods in this research also shows that these two techniques can be relied on for similar applications in market research and product development in other industries. The use of these advanced data mining techniques not only enriches our understanding of consumer preferences but also paves the way for more targeted innovation and more effective marketing strategies.

REFERENCES

Adjani, K., Fauzia, F. A., & Juliane, C. (2023). Comparison of K-N Earest Neighbor and Naïve Bayes Algorithms for Prediction of Aptikom Membership Activity Extension in 2023. *SinkrOn*, *8*(2), 700–707. https://doi.org/10.33395/sinkron.v8i2.12081

Amdhani, R., Wati, E. E., Purwanto, K., Manajemen, P. S., Hukum, F., Bisnis, E., & Indonesia, D. (2022). *Jurnal manajemen dewantara*. *6*(3), 281–291.

Anwar, B., Jalinus, N., & Abdullah, R. (2023). Weather Forecast In Medan City With Hopfield Artificial Neural Network Algorithm. *Sinkron*, *8*(1), 398–404. https://doi.org/10.33395/sinkron.v8i1.12048

Arifuddin, N. A., Pinastawa, I. W. R., Anugraha, N., & Pradana, M. G. (2023). Classification of Stroke Opportunities with Neural Network and K-Nearest Neighbor Approaches. *SinkrOn*, *8*(2), 688–693. https://doi.org/10.33395/sinkron.v8i2.12228

Hasibuan, F. F., Dar, M. H., & Yanris, G. J. (2023). Implementation of the Naïve Bayes Method to determine the Level of Consumer Satisfaction. *SinkrOn*, *8*(2), 1000–1011. https://doi.org/10.33395/sinkron.v8i2.12349

Hermawan, F., & Prianggono, J. (2023). Crime of theft prediction using Machine Learning K-Nearest Neighbour Algorithm at Polresta Bandar Lampung. *Sinkron*, *8*(3), 1515–1527. https://doi.org/10.33395/sinkron.v8i3.12422

Hidup, P. G., Produk, D., Citra, D. A. N., & Terhadap, M. (2023). *MINAT BELI MOTOR MATIC HONDA SCOOPY PRESTIGE PADA MAHASISWA*. *11*(2).

Hindarto, D. (2023). Enhancing Road Safety with Convolutional Neural Network Traffic Sign Classification. *Sinkron*, *8*(4), 2810–2818. https://doi.org/10.33395/sinkron.v8i4.13124

Istiadi, I., Rahman, A. Y., & Wisnu, A. D. R. (2023). Identification of Tempe Fermentation Maturity Using

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 2, April 2024
DOI : https://doi.org/10.33395/sinkron.v8i2.13637
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Principal Component Analysis and K-Nearest Neighbor. *Sinkron*, *8*(1), 286–294. https://doi.org/10.33395/sinkron.v8i1.12006

Lestari, V., Mawengkang, H., & Situmorang, Z. (2023). Artificial Neural Network Backpropagation Method to Predict Tuberculosis Cases. *Sinkron*, *8*(1), 35–47. https://doi.org/10.33395/sinkron.v8i1.11998

Maulana, Z., Mentary, N., & Muhibuddin, A. (2023). *KARAKTERISTIK KERAGAMAN MORFOLOGI CABAI KATOKKON Capsicum chinense Jacq . Morphological Diversity Characteristics of Chili Katokkon Capsicum cinenses Jacq*. *5*(2), 84–88. https://doi.org/10.35965/jae.v5i2.3458

Mawaddah, A., Dar, M. H., & Yanris, G. J. (2023). Analysis of the SVM Method to Determine the Level of Online Shopping Satisfaction in the Community. *SinkrOn*, *8*(2), 838–855. https://doi.org/10.33395/sinkron.v8i2.12261

Sari, A. W., Hermanto, T. I., & Defriani, M. (2023). Sentiment Analysis Of Tourist Reviews Using K-Nearest Neighbors Algorithm And Support Vector Machine. *Sinkron*, *8*(3), 1366–1378. https://doi.org/10.33395/sinkron.v8i3.12447

Sari, M., Yanris, G. J., & Hasibuan, M. N. S. (2023). Analysis of the Neural Network Method to Determine Interest in Buying Pertamax Fuel. *SinkrOn*, *8*(2), 1031–1039. https://doi.org/10.33395/sinkron.v8i2.12292

Siregar, A. P., Irmayani, D., & Sari, M. N. (2023). Analysis of the Naïve Bayes Method for Determining Social Assistance Eligibility Public. *SinkrOn*, *8*(2), 805–817. https://doi.org/10.33395/sinkron.v8i2.12259

Suherman, E., Hindarto, D., Makmur, A., & Santoso, H. (2023). Comparison of Convolutional Neural Network and Artificial Neural Network for Rice Detection. *Sinkron*, *8*(1), 247–255. https://doi.org/10.33395/sinkron.v8i1.11944

Triani, D. J., Dar, M. H., & Yanris, G. J. (2023). Analysis of Public Purchase Interest in Yamaha Motorcycles Using the K-Nearest Neighbor Method. *Sinkron*, *8*(3), 1238–1254. https://doi.org/10.33395/sinkron.v8i3.12433

Violita, P., Yanris, G. J., & Hasibuan, M. N. S. (2023). Analysis of Visitor Satisfaction Levels Using the K-Nearest Neighbor Method. *SinkrOn*, *8*(2), 898–914. https://doi.org/10.33395/sinkron.v8i2.12257

Zai, F., Sirait, J., Nainggolan, D. W., Sihombing, N. G. D., & Banjarnahor, J. (2023). Comparison Analysis of C4.5 Algorithm and KNN Algorithm for Predicting Data of Non-Active Students at Prima Indonesia University. *Sinkron*, *8*(4), 2027–2035. https://doi.org/10.33395/sinkron.v8i4.12879

*name of corresponding author