Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 3, July 2023
DOI : https://doi.org/10.33395/sinkron.v8i3.12674
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Implementation of Data Mining for Data Classification of Visitor Satisfaction Levels

Hubban Arfi Pratama[1)*], Gomal Juni Yanris[2)], Mila Nirmala Sari Hasibuan[3)]
[1,2,3)]Universitas Labuhanbatu, Indonesia
[1)] hubbanarfi@gmail.com, [2)]gomaljuniyanris@gmail.com, [3)]milanirmalasari7@gmail.com

Submitted : Jul 10, 2023 | **Accepted** : Jul 25, 2023 | **Published** : Jul 25, 2023

Abstract: An amusement park is a location or place that can provide a special attraction to the public. This is because in amusement parks there is lots of entertainment provided. But not all amusement parks are liked by visitors, usually because the location is still not good enough. Therefore the authors make a study of the level of visitor satisfaction. This research was made so that the writer can determine whether or not the number of visitors is satisfied at the amusement park. To conduct this research, the authors used 2 methods with a classification model in data mining. The methods used are the K-Nearest Neighbor (kNN) method and the Naïve Bayes method. Study this is done using 100 visitor data. The classification results obtained from both methods give the same results. The results obtained were 77 satisfied visitor data at amusement parks and 23 dissatisfied visitors at amusement parks. The result of the two methods used is that many visitors are satisfied with the amusement park. The accuracy results obtained are also very good. This means that these two methods are very suitable to be used as a method with a classification model. The conclusion is that the amusement park has beauty and a great location that can give attraction to visitors. With this research it can be a reference that the K-Nearest Neighbor (kNN) method and the Naïve Bayes method are very suitable for carrying out a data classification.

Keywords: Classification, Confusion Matrix, K-Nearest Neighbor (kNN), Naïve Bayes, Satisfaction Level

INTRODUCTION

An amusement park is a place or location that has attractions such as rides, swimming pools, sleds. Amusement parks are one of the places that are often visited by visitors. No wonder the amusement park has a lot of visitors. An example of an amusement park in Labuhanbatu Regency is Boombara Waterpark Rantauprapat. Boombara Waterpark Rantauprapat is one of the amusement parks that is frequently visited by visitors. Especially on holidays, lots of people come, ranging from children to adults. That's because Boombara Waterpark Rantauprapat has a good attraction, so it can bring in a lot of visitors. From the things that have been described above, the author will make a research on the level of visitor satisfaction at the Boombara Waterpark Rantauprapat amusement park. This research was conducted because of the large number of visitors who came to Boombara Waterpark, so the authors wanted to know the visitor's assessment of Boombara Waterpark Rantauprapat, such as the cleanliness of the location, access to the location, and the service.

Visitor satisfaction is a very important thing for an amusement park, this is because it has an impact on the development or not of an amusement park. This is because if the visitor rating at an amusement park is negative, it will reduce the number of visitors who come and assess visitors to an amusement park is positive, it will increase the number of visitors. Therefore the authors made a study of the level

*name of corresponding author

of visitor satisfaction at Boombara Waterpark Rantauprapat, with the aim that the authors could find out how visitors rated Boombara Waterpark Rantauprapat. This research can later be used as a reference for developing Boombara Waterpark Rantauprapat. If many visitors are satisfied with the amusement park, then Boombara Waterpark Rantauprapat is in great demand by visitors who come to that place. But if many visitors are dissatisfied with Boombara Waterpark Rantauprapat, the dissatisfied visitors can be used as a reference for them to be able to change and develop the place so that visitors can be satisfied with the atmosphere of the amusement park provided. To conduct this research, the authors will carry out a data classification in data mining. So the sample data (visitor data) that has been obtained from the questionnaire will be classified on data mining.

Data mining is a process carried out to extract data into information using techniques statistics and mathematics (Hussain, Dahan, Ba-Alwib, & Ribata, 2018) (Baharuddin, Azis, & Hasanuddin, 2019) (Indrayuni, 2019) (Watratan, B, Moeis, Informasi, & Makassar, 2020). Data mining is also a technique that is used to extract knowledge from data sets using techniques of statistics and mathematics (Uçar & Karahoca, 2021). The data mining will later be stored in a repository using a reasoning pattern technology with mathematical and statistical techniques (Wijaya & Girsang, 2015). he data mining that will be carried out will later use a classification model. The data obtained will be processed on data mining to classify the level of visitor satisfaction at Boombara Waterpark Rantauprapat (Saputra, Widiyaningtyas, & Wibawa, 2018). Data classification is a model that exists in data mining by grouping data with provision predetermined categories (Pour, Esmaeili, & Romoozi, 2022) (Kumar, Chatterjee, & Díaz, 2020). To perform data mining, we need a method or algorithm that can be used to classify data. There are 2 methods that will be used in the data mining process, namely the K-Nearest Neighbor (kNN) method and the Naïve Bayes method. Both of these methods are methods that can be used to classify data.

METHOD

Metode K-Nearest Neighbor (kNN)

The K-Nearest Neighbor (kNN) method is a model that can be used to classify data based on data that is close to the desired object (Liantoni, 2016) (Lubis, Lubis, & Al-Khowarizmi, 2020). The K-Nearest Neighbor (kNN) method can be a suitable method used to classify data in the data mining process. Using the K-Nearest Neighbor (kNN) method can help writers solve a classification problem and has high accuracy (Farhad Khorshid & Mohsin Abdulazeez, 2021). This method can withstand data that has noise and this method is also quite effective if used to perform training on quite large data (Rosso, 2019). This method can also do training quickly and the way to understand it is also quite simple and easy to understand and study. For the general form of the formula for the K-Nearest Neighbor (kNN) method, the following formula can be used:

$$\boldsymbol{Dxy} = \sqrt{\sum_{i=1}^{n}(\boldsymbol{x}_i - \boldsymbol{y}_i)^2}$$

Information
D : Proximity Distance
x : data *training*
y : data *testing*
n : the number of individual attributes between 1s/d n
f : attribute *similarity* function I between cases x and y
I : individual attributes between 1s/d n

Metode Naïve Bayes

The Naïve Bayes method is a method that can be used for simple classification by calculating the probability of a combination of certain data (Azzahra & Wibowo, 2020) (Murwantara, Yugopuspito, & Hermawan, 2020). This method is used to classify certain data with simple probabilities that are arranged

*name of corresponding author

e-ISSN : 2541-2019
p-ISSN : 2541-044X

and designed so they can be used with inter assumptions variables by using techniques of statistics and mathematics (Santoso et al., 2020). This method will later be used to classify visitor satisfaction level data at Boombara Waterpark Rantauprapat. With this method, visitor data will be classified based on certain conditions and categories that have been determined by the author. To use this method, please note that the pattern of each widget used must match, so that later the system that has been designed for data mining can be used properly (Damuri, Riyanto, Rusdianto, & Aminudin, 2021). For the general formula of the method naive bayes which is used as follows:

$$\mathbf{P\ (A\ |\ B):}\ \frac{P\ (B\ |A)\ P\ (A)}{P\ (B)} \qquad \text{(Di \& Duan, 2014)}$$

Information:
A : hypothesis of data A (specific class)
B : data with unknown classes
P (A | B) : Probability of hypothesis based on condition B
P (A) : Probability of hypothesis A
P (B | A) : Probability B when condition A
P (B) : Probability (Azzahra & Wibowo, 2020)

The naïve Bayes method is a method by calculating other related probabilities. After being applied to the Naive Bayes algorithm, this formula produces a basic assumption. In looking at a feature, this algorithm always assumes that the feature is independent, equal, and has a contribution to the result. The way it works is that we try to find the probability of event A, if event B is true. Event B is also referred to as evidence. For P(A) is the a priori of A (the prior probability, i.e. the probability of the event before the proof is seen) and the proof is the attribute value of the unknown instance (event B). For P(A|B) is the posteriori probability of B, that is, the probability that it will occur after the evidence is shown.

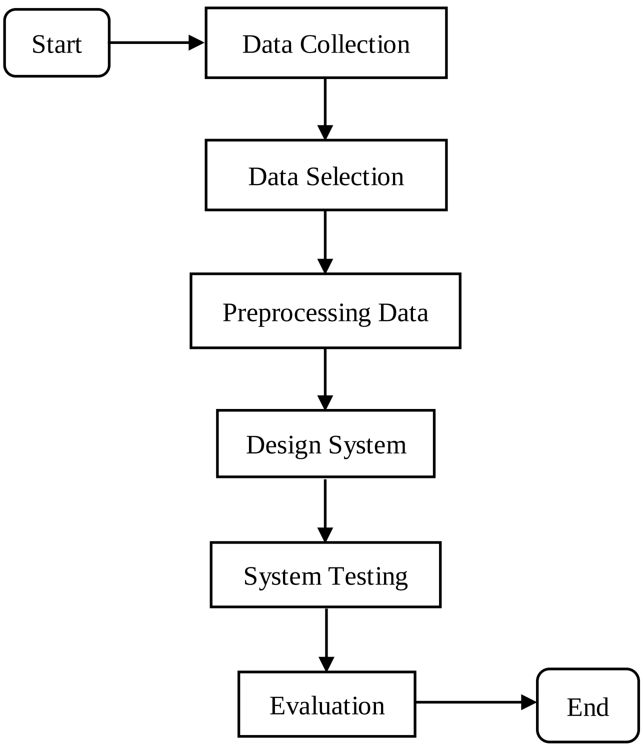

Fig 1. *Classification Process Design*

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 3, July 2023
DOI : https://doi.org/10.33395/sinkron.v8i3.12674
e-ISSN : 2541-2019
p-ISSN : 2541-044X

To classify data in data mining using the K-Nearest Neighbor (kNN) method and the Naïve Bayes method, there are a number of the steps that must be carried out, namely the first stage is to collect visitor data which will be used as research sample data, the second stage is choose and selecting data to be used as research sample data, the third stage is data preprocessing which is the process of selecting appropriate data to be used as research sample data and at this stage too, the data is arranged in a form that can be used in data mining. The selected data is arranged in a file with excel format, namely file.xlsx, the fourth stage is design system in data mining using the K-Nearest Neighbor (kNN) and Naïve Bayes methods so that data can be classified properly, the fifth stage is the classification process using a system that was created previously in data mining and the last stage is evaluation which is the process carried out to determine the accuracy of the method used. Accuracy aims to see eligibility for the method used, can the method do a good classification.

Confusion Matrix

Confusion matrix is an easy to use and effective tool to perform a data classification and easy to be able to determine the results of classification (Yun, 2021). he confusion matrix can be used to carry out a work evaluation of a model and can be used to determine the results of the data mining process.

Table 1
Confusion Matrix

Classification	Predicated Class	
	True	False
Actual: True	True Positive (TP)	False Positive (FP)
Actual: False	False Negative (FN)	True Negative (TN)

To determine the calculation of the confusion matrix, we can do it by calculating accuracy, precision and recall.

Accuracy is calculated with the condition that the prediction number (TP + TN) is divided by the number of samples available, to calculate accuracy, the following formula can be seen (Patil & Tamane, 2018):

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (1)$$

Precision is used to identify positive cases with a high false positive rate, which can be calculated as follows (Agustina, Adrian, & Hermawati, 2021):

$$\text{Precision} = \frac{TP}{TP+FP} \times 100\% \quad (2)$$

In contrast to precision, recall serves to identify positive cases with high false negative values. Recall can be calculated in this way (Normawati & Prayogi, 2021):

$$\text{Recall} = \frac{TP}{TP+FN} \times 100\% \quad (3)$$

RESULT

Data Analysis

The table below is visitor data which is the research sample data. The data was obtained from a questionnaire distributed to people who had been visitors to Boombara Waterpark Rantauprapat.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 3, July 2023
DOI : https://doi.org/10.33395/sinkron.v8i3.12674
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Table 2
Visitor Data

Full Name	Gender	Swimming Pool Cleanliness	Swimming Pool Entrance Fee	Service Quality	Access To Swimming Pool
Adelina	Woman	Not Clean	Cheap	Not Care	Difficult
Ahmad Aldi Amin	Man	Clean	Cheap	Not Care	Easy
Ali Pahang	Man	Clean	Cheap	Not Care	Easy
Ammon Siagian	Woman	Clean	Affordable	Not Care	Easy
Arifin Munthe	Man	Clean	Expensive	Friendly	Easy
Aripinsah Ritonga	Man	Clean	Affordable	Not Care	Easy
Aris Saputra	Man	Not Clean	Cheap	Not Care	Difficult
Asas Pardamean Ritonga	Man	Clean	Cheap	Not Care	Easy
Azwar Anas	Man	Clean	Expensive	Not Care	Difficult
Bida Sari Hasibuan	Woman	Clean	Expensive	Friendly	Easy
Darmawati	Woman	Clean	Cheap	Not Care	Easy
Destifika Andriani	Woman	Clean	Affordable	Friendly	Easy
Dewi Kusuma	Woman	Clean	Cheap	Friendly	Easy
Dian Praja	Man	Not Clean	Affordable	Friendly	Easy
Dicky Afrian	Man	Clean	Affordable	Friendly	Difficult
Dirma	Woman	Clean	Cheap	Friendly	Easy
Doly Andriani Syadid	Man	Clean	Affordable	Not Care	Easy
Elida Yanti Tanjung	Woman	Clean	Expensive	Friendly	Easy
Evi Rianti	Woman	Not Clean	Cheap	Not Care	Difficult
Fatimah Zahara	Woman	Not Clean	Expensive	Friendly	Difficult
Fitrawati	Woman	Clean	Cheap	Not Care	Easy
Fitri Arhoma	Woman	Not Clean	Cheap	Not Care	Difficult
Hariatik	Woman	Clean	Affordable	Friendly	Easy
Hartini Dwitri	Woman	Clean	Expensive	Not Care	Difficult
Harziansyah	Man	Clean	Affordable	Friendly	Easy
Hidayat Amin	Man	Clean	Cheap	Friendly	Difficult
Hotmarida	Woman	Clean	Affordable	Friendly	Difficult
Ika Ayuli	Woman	Clean	Affordable	Friendly	Easy
Imam Syaputra Munthe	Man	Clean	Cheap	Not Care	Easy
Indah Ratu Aulia Ritonga	Woman	Clean	Cheap	Not Care	Easy
Irsad Amsali	Man	Clean	Cheap	Friendly	Difficult
Irwan Efendi Hasibuan	Man	Clean	Cheap	Indifferent	Easy
Ismail Nasution	Man	Clean	Affordable	Not Care	Easy
Ismail Nasution	Man	Clean	Affordable	Friendly	Easy
Jamilatul Adab	Woman	Clean	Expensive	Friendly	Easy
Julia Asmara	Woman	Not Clean	Expensive	Friendly	Difficult
Junaidi	Man	Clean	Affordable	Not Care	Easy
Juni Haryani	Woman	Not Clean	Cheap	Not Care	Difficult
Leli Wifda	Woman	Not Clean	Cheap	Not Care	Difficult
Lenggayani Siregar	Woman	Clean	Cheap	Indifferent	Easy
Leo Sunarta	Man	Clean	Cheap	Friendly	Easy
Liana	Woman	Clean	Affordable	Not Care	Easy

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 3, July 2023 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v8i3.12674 p-ISSN : 2541-044X

Linda Amanda	Woman	Not Clean	Affordable	Friendly	Easy
Maimunah Tanjung	Woman	Not Clean	Cheap	Friendly	Easy
Mariani	Woman	Clean	Cheap	Not Care	Easy
Mhd Nur Nasution	Man	Not Clean	Cheap	Not Care	Difficult
Misni	Woman	Clean	Affordable	Friendly	Difficult
Muhammad Al Akbar	Man	Clean	Affordable	Friendly	Difficult
Muhammad Fratama	Man	Clean	Cheap	Not Care	Easy
Mul Mujiono	Man	Clean	Expensive	Friendly	Easy
Musri Zal Romadhani	Man	Clean	Cheap	Not Care	Easy
Mustika Suhardi	Man	Clean	Expensive	Not Care	Difficult
Nando Dermawan	Man	Not Clean	Expensive	Not Care	Difficult
Nasli Nasution	Man	Clean	Affordable	Not Care	Easy
Nissi Hartat Ritonga	Woman	Clean	Affordable	Not Care	Easy
Nita Hariani	Woman	Clean	Expensive	Friendly	Easy
Nur Azman	Man	Clean	Cheap	Not Care	Easy
Nur Hafni Nasution	Woman	Clean	Expensive	Friendly	Easy
Nurhayati Caniago	Woman	Clean	Cheap	Not Care	Easy
Nurul Mawaddah	Woman	Clean	Cheap	Friendly	Easy
Pidah	Woman	Not Clean	Affordable	Friendly	Easy
Quito Rian Zori	Man	Clean	Cheap	Indifferent	Easy
Ratna Dewi Tanjung	Woman	Clean	Affordable	Not Care	Easy
Reynalda Safira	Woman	Clean	Affordable	Friendly	Easy
Reynaldo Deswara	Man	Clean	Affordable	Friendly	Easy
Ridanti Mahrani	Woman	Clean	Affordable	Friendly	Easy
Rifandi Hasibuan	Man	Clean	Affordable	Not Care	Easy
Rinaldi Afriza	Man	Clean	Cheap	Friendly	Easy
Rini Misriyani	Woman	Clean	Affordable	Not Care	Easy
Riswan	Man	Clean	Affordable	Not Care	Easy
Rodiah Dalimunthe	Woman	Not Clean	Cheap	Not Care	Difficult
Roma	Man	Clean	Expensive	Not Care	Difficult
Rosmawati Hazda	Woman	Clean	Affordable	Friendly	Easy
Rosnima Lubis	Woman	Clean	Affordable	Not Care	Easy
Rostika Simanjuntak	Woman	Clean	Cheap	Not Care	Easy
Rustam Hasibuan	Man	Clean	Affordable	Not Care	Easy
Safrianto	Man	Clean	Cheap	Not Care	Easy
Saifuddin Zuhri	Man	Not Clean	Cheap	Not Care	Difficult
Sarah	Woman	Clean	Cheap	Not Care	Easy
Satria Amanda Negara	Man	Not Clean	Cheap	Not Care	Difficult
Septiandi	Man	Clean	Cheap	Not Care	Easy
Siti Jubaidah	Woman	Clean	Affordable	Friendly	Easy
Sri Handayani Fitri	Woman	Not Clean	Cheap	Friendly	Easy
Sri Hera Wati Dalimunthe	Woman	Clean	Expensive	Not Care	Difficult
Sri Wahyuni Siregar	Woman	Clean	Affordable	Not Care	Easy
Susanto	Man	Clean	Expensive	Not Care	Difficult
Suwastati Sagala	Woman	Clean	Cheap	Not Care	Easy
Syafril	Man	Not Clean	Cheap	Not Care	Easy
Syiful Idek	Man	Clean	Expensive	Friendly	Easy
Tan Hartono	Man	Clean	Affordable	Not Care	Easy
Tumpal Dalimunthe	Man	Clean	Cheap	Not Care	Easy

Tutin Harmita	Woman	Clean	Affordable	Not Care	Easy
Wiwik	Woman	Clean	Cheap	Friendly	Difficult
Yogi Kharisma Dalimunthe	Man	Not Clean	Cheap	Not Care	Difficult
Yuli Wulandari	Woman	Not Clean	Expensive	Friendly	Difficult
Yurnawilis Koto	Woman	Clean	Affordable	Not Care	Easy
Yus Negayenti	Woman	Clean	Cheap	Friendly	Easy
Zirli Zaidan	Man	Clean	Affordable	Friendly	Easy
Zuanda Lubis	Man	Not Clean	Cheap	Not Care	Difficult
Zulkifli Koto	Man	Clean	Affordable	Not Care	Easy

Table 2 is visitor data for Boombara Waterpark Rantauprapat obtained from a questionnaire of 100 visitor data. This data will be used as research sample data which will be classified in data mining.

Data Training

Training data is data that will be used to assist the classification process in data mining. The training data was selected and arranged in the same format as the sample data, namely the excel file.xlsx file format so that the data can be used for the classification process in data mining.

Table 3
Data Training

Full Name	Gender	Swimming Pool Cleanliness	Swimming Pool Entrance Fee	Service quality	Access To Swimming Pool	Category
Amin Hidayat	Man	Clean	Cheap	Friendly	Difficult	Satisfied
Arif Widianto	Man	Clean	Affordable	Friendly	Difficult	Satisfied
Arini Mawaddah	Man	Not Clean	Cheap	Not Care	Easy	Not Satisfied
Arsyad Tholib Pohan	Man	Clean	Cheap	Friendly	Difficult	Satisfied
Azizah	Woman	Not Clean	Cheap	Not Care	Difficult	Not Satisfied
Azman Dalimunthe	Man	Clean	Cheap	Not Care	Easy	Satisfied
Efendi Hasibuan	Man	Clean	Affordable	Not Care	Easy	Satisfied
Elfina Harahap	Woman	Clean	Expensive	Friendly	Easy	Satisfied
Eli Kurnia	Woman	Not Clean	Cheap	Not Care	Difficult	Not Satisfied
Fitri Aini	Woman	Not Clean	Cheap	Friendly	Easy	Satisfied
Fitri Febriani	Woman	Not Clean	Cheap	Not Care	Difficult	Not Satisfied
Haris Tonang	Man	Not Clean	Cheap	Not Care	Difficult	Not Satisfied
Hartono	Man	Clean	Expensive	Not Care	Difficult	Not Satisfied
Hotma Sarida	Woman	Clean	Affordable	Friendly	Difficult	Satisfied
Imam Prasetyo	Man	Clean	Cheap	Friendly	Difficult	Satisfied
Indah Purnama	Woman	Not Clean	Cheap	Not Care	Difficult	Not Satisfied

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 3, July 2023
DOI : https://doi.org/10.33395/sinkron.v8i3.12674
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Intan Pratiwi	Woman	Clean	Expensive	Not Care	Difficult	Not Satisfied
Ira Rahmawati	Woman	Clean	Cheap	Not Care	Easy	Satisfied
Joko	Man	Clean	Expensive	Friendly	Easy	Satisfied
Julpan Efendi	Man	Clean	Expensive	Not Care	Difficult	Not Satisfied
Juni Pane	Woman	Not Clean	Expensive	Friendly	Difficult	Not Satisfied
Juwanda	Man	Clean	Affordable	Not Care	Easy	Satisfied
Karisma	Man	Clean	Affordable	Not Care	Easy	Satisfied
Kurnia Putri'	Woman	Clean	Expensive	Friendly	Easy	Satisfied
Linda Putri Hasibuan	Woman	Not Clean	Cheap	Not Care	Difficult	Not Satisfied
Mayang Sari	Woman	Not Clean	Cheap	Friendly	Easy	Satisfied
Muhammad Efendi	Man	Clean	Cheap	Not Care	Easy	Satisfied
Murni Adinda	Woman	Clean	Affordable	Friendly	Difficult	Satisfied
Nasli Nasution	Man	Clean	Affordable	Not Care	Easy	Satisfied
Nono Rambe	Man	Not Clean	Expensive	Not Care	Difficult	Not Satisfied
Nur Izmi	Woman	Clean	Cheap	Not Care	Easy	Satisfied
Nurokhim	Man	Not Clean	Cheap	Not Care	Difficult	Not Satisfied
Rezi Syahputra	Man	Clean	Affordable	Not Care	Easy	Satisfied
Rini Antika	Woman	Clean	Affordable	Not Care	Easy	Satisfied
Rohami	Woman	Clean	Affordable	Not Care	Easy	Satisfied
Roman Sitorus	Man	Clean	Expensive	Not Care	Difficult	Not Satisfied
Sari Handayani	Woman	Clean	Affordable	Not Care	Easy	Satisfied
Saskia Harahap	Woman	Not Clean	Cheap	Not Care	Difficult	Not Satisfied
Satria Hamdani	Man	Not Clean	Cheap	Not Care	Difficult	Not Satisfied
Siska Pratiwi	Woman	Clean	Affordable	Not Care	Easy	Satisfied
Sri Wahdini	Woman	Clean	Expensive	Not Care	Difficult	Not Satisfied
Suhardi	Man	Clean	Expensive	Not Care	Difficult	Not Satisfied
Sulistiawati	Woman	Clean	Cheap	Not Care	Easy	Satisfied
Surasto	Man	Clean	Cheap	Not Care	Easy	Satisfied
Syaiful Harahap	Man	Not Clean	Cheap	Not Care	Difficult	Not Satisfied
Syarah Tanjung	Woman	Clean	Cheap	Not Care	Easy	Satisfied
Tati Irawati	Woman	Clean	Affordable	Not Care	Easy	Satisfied
Tiara Putri	Woman	Clean	Affordable	Not Care	Easy	Satisfied
Yogi Dalimunthe	Man	Not Clean	Cheap	Not Care	Difficult	Not Satisfied
Yulia Ningsih	Woman	Not Clean	Expensive	Friendly	Difficult	Not Satisfied

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 3, July 2023
DOI : https://doi.org/10.33395/sinkron.v8i3.12674
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Yusnanda Riani	Woman	Clean	Cheap	Friendly	Easy	Satisfied
Zahra Hasibuan	Woman	Not Clean	Expensive	Friendly	Difficult	Not Satisfied
Zaki Harahap	Man	Clean	Cheap	Not Care	Easy	Satisfied
Zuhri Rambe	Man	Not Clean	Cheap	Not Care	Difficult	Not Satisfied

Table 4 is a table that contains visitor data for Boombara Waterpark Rantauprapat. Data training is used to assist the classification process that will be carried out in data mining.

Table 4
Visitor Column Data

No	Attribute	Type	Role	Values
1	Full Name	Text	Meta	
2	Gender	Categorical	Feature	Man, Woman
3	Swimming Pool Cleanliness	Categorical	Feature	Clean, Not Clean
4	Swimming Pool Entrance Fee	Categorical	Feature	Affordable, Cheap, Expensive
5	Service Quality	Categorical	Feature	Friendly, Not Care
6	Access to Swimming Pool	Categorical	Feature	Difficult, Easy
7	Category	Categorical	Target	Satisfied, Not Satisfied

Table 4 above is an attribute of visitor data that has been obtained from the questionnaire. These attributes are used as research parameters to determine the level of visitor satisfaction at Boombara Waterpark Rantauprapat. In each attribute there are several values that become research parameters. To be able to get the results of the classification of satisfied and dissatisfied visitors at Boombara Waterpark Rantauprapat, the category attribute in the role column needs to be changed from feature to target. This is done so that later when the classification process is executed, the category attributes can provide answers to the classification results.

Data Selection Process (Preprocessing)

Data preprocessing is a process of selecting data that is suitable for use in data mining (Negara, Muhardi, & Putri, 2020). The process is carried out so that the data used can provide good results. Therefore the need for a process for data selection. After the data has been selected, the data will be compiled into a form that suits the needs of data mining (Al-Rasheed, 2021). Data is compiled in excel format, file with file.xlsx format.

Data Mining Process

The data mining process will be carried out using a classification model using 2 methods capable of carrying out a data classification, namely the K-Nearest Neighbor (kNN) method and the Naïve Bayes method. In this classification model, the data will be grouped based on the same data and the grouping of data is also adjusted according to needs. Data grouping will be divided into 2, namely visitors who are satisfied and visitors who are not satisfied at Boombara Waterpark Rantauprapat.

*name of corresponding author

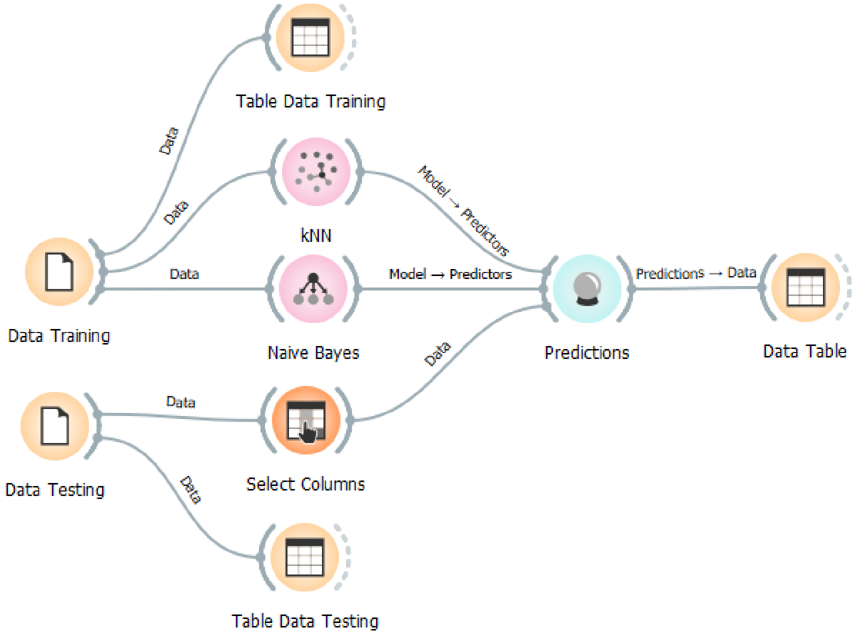

Fig 2. *System Design Process in Data Mining*

Figure 2 is the process of designing a widget pattern in data mining so that it can be used in the data classification process using the K-Nearest Neighbor (kNN) method and the Naïve Bayes method. The design of this pattern is done so that the sample data can be classified perfectly.

Classification Model Testing Process

In this process a testing process system with a classification model is done so that the data can be categorized according to each class. This process will be carried out in data mining using the K-Nearest Neighbor (kNN) method and the Naïve Bayes method.

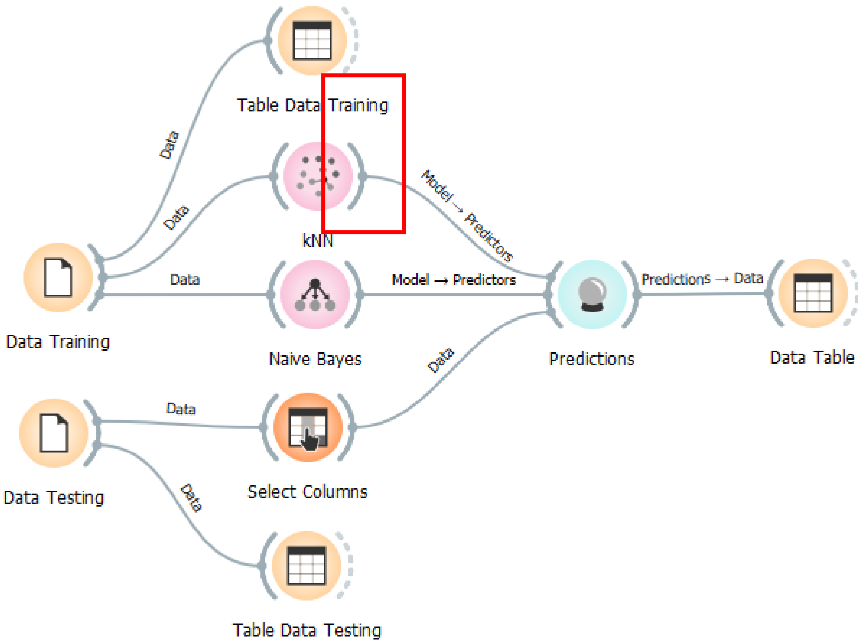

Fig 3. *Classification Process in Data Mining*

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 3, July 2023
DOI : https://doi.org/10.33395/sinkron.v8i3.12674
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Figure 3 is a process testing system which was designed previously in data mining to be able to classify visitor satisfaction level data at Boombara Waterpark Rantauprapat. The method used can be seen in the widget located inside the red box. The methods used are the K-Nearest Neighbor (kNN) method and the Naïve Bayes method.

Classification Model Predictions Process

This stage is the prediction result of the classification model that has been carried out using the K-Nearest Neighbor (kNN) method and the Naïve Bayes method. Result Classification can be seen on table under.

Table 5
Classification Model Prediction Results

Full Name	Gender	Swimming Pool Cleanliness	Swimming Pool Entrance Fee	Service quality	Access To Swimming Pool	kNN	Naive Bayes
Adelina	Woman	Not Clean	Cheap	Not Care	Difficult	Not Satisfied	Not Satisfied
Ahmad Aldi Amin	Man	Clean	Cheap	Not Care	Easy	Satisfied	Satisfied
Ali Pahang	Man	Clean	Cheap	Not Care	Easy	Satisfied	Satisfied
Ammon Siagian	Woman	Clean	Affordable	Not Care	Easy	Satisfied	Satisfied
Arifin Munthe	Man	Clean	Expensive	Friendly	Easy	Satisfied	Satisfied
Aripinsah Ritonga	Man	Clean	Affordable	Not Care	Easy	Satisfied	Satisfied
Aris Saputra	Man	Not Clean	Cheap	Not Care	Difficult	Not Satisfied	Not Satisfied
Asas Pardamean Ritonga	Man	Clean	Cheap	Not Care	Easy	Satisfied	Satisfied
Azwar Anas	Man	Clean	Expensive	Not Care	Difficult	Not Satisfied	Not Satisfied
Bida Sari Hasibuan	Woman	Clean	Expensive	Friendly	Easy	Satisfied	Satisfied
Darmawati	Woman	Clean	Cheap	Not Care	Easy	Satisfied	Satisfied
Destifika Andriani	Woman	Clean	Affordable	Friendly	Easy	Satisfied	Satisfied
Dewi Kusuma	Woman	Clean	Cheap	Friendly	Easy	Satisfied	Satisfied
Dian Praja	Man	Not Clean	Affordable	Friendly	Easy	Satisfied	Satisfied
Dicky Afrian	Man	Clean	Affordable	Friendly	Difficult	Satisfied	Satisfied
Dirma	Woman	Clean	Cheap	Friendly	Easy	Satisfied	Satisfied
Doly Andriani Syadid	Man	Clean	Affordable	Not Care	Easy	Satisfied	Satisfied
Elida Yanti Tanjung	Woman	Clean	Expensive	Friendly	Easy	Satisfied	Satisfied
Evi Rianti	Woman	Not Clean	Cheap	Not Care	Difficult	Not Satisfied	Not Satisfied
Fatimah Zahara	Woman	Not Clean	Expensive	Friendly	Difficult	Not Satisfied	Not Satisfied
Fitrawati	Woman	Clean	Cheap	Not Care	Easy	Satisfied	Satisfied
Fitri Arhoma	Woman	Not Clean	Cheap	Not Care	Difficult	Not Satisfied	Not Satisfied
Hariatik	Woman	Clean	Affordable	Friendly	Easy	Satisfied	Satisfied
Hartini Dwitri	Woman	Clean	Expensive	Not Care	Difficult	Not Satisfied	Not Satisfied
Harziansyah	Man	Clean	Affordable	Friendly	Easy	Satisfied	Satisfied
Hidayat Amin	Man	Clean	Cheap	Friendly	Difficult	Satisfied	Satisfied
Hotmarida	Woman	Clean	Affordable	Friendly	Difficult	Satisfied	Satisfied

*name of corresponding author

e-ISSN : 2541-2019
p-ISSN : 2541-044X

Ika Ayuli	Woman	Clean	Affordable	Friendly	Easy	Satisfied	Satisfied
Imam Syaputra Munthe	Man	Clean	Cheap	Not Care	Easy	Satisfied	Satisfied
Indah Ratu Aulia Ritonga	Woman	Clean	Cheap	Not Care	Easy	Satisfied	Satisfied
Irsad Amsali	Man	Clean	Cheap	Friendly	Difficult	Satisfied	Satisfied
Irwan Efendi Hasibuan	Man	Clean	Cheap	Indifferent	Easy	Satisfied	Satisfied
Ismail Nasution	Man	Clean	Affordable	Not Care	Easy	Satisfied	Satisfied
Ismail Nasution	Man	Clean	Affordable	Friendly	Easy	Satisfied	Satisfied
Jamilatul Adab	Woman	Clean	Expensive	Friendly	Easy	Satisfied	Satisfied
Julia Asmara	Woman	Not Clean	Expensive	Friendly	Difficult	Not Satisfied	Not Satisfied
Junaidi	Man	Clean	Affordable	Not Care	Easy	Satisfied	Satisfied
Juni Haryani	Woman	Not Clean	Cheap	Not Care	Difficult	Not Satisfied	Not Satisfied
Leli Wifda	Woman	Not Clean	Cheap	Not Care	Difficult	Not Satisfied	Not Satisfied
Lenggayani Siregar	Woman	Clean	Cheap	Indifferent	Easy	Satisfied	Satisfied
Leo Sunarta	Man	Clean	Cheap	Friendly	Easy	Satisfied	Satisfied
Liana	Woman	Clean	Affordable	Not Care	Easy	Satisfied	Satisfied
Linda Amanda	Woman	Not Clean	Affordable	Friendly	Easy	Satisfied	Satisfied
Maimunah Tanjung	Woman	Not Clean	Cheap	Friendly	Easy	Satisfied	Satisfied
Mariani	Woman	Clean	Cheap	Not Care	Easy	Satisfied	Satisfied
Mhd Nur Nasution	Man	Not Clean	Cheap	Not Care	Difficult	Not Satisfied	Not Satisfied
Misni	Woman	Clean	Affordable	Friendly	Difficult	Satisfied	Satisfied
Muhammad Al Akbar	Man	Clean	Affordable	Friendly	Difficult	Satisfied	Satisfied
Muhammad Fratama	Man	Clean	Cheap	Not Care	Easy	Satisfied	Satisfied
Mul Mujiono	Man	Clean	Expensive	Friendly	Easy	Satisfied	Satisfied
Musri Zal Romadhani	Man	Clean	Cheap	Not Care	Easy	Satisfied	Satisfied
Mustika Suhardi	Man	Clean	Expensive	Not Care	Difficult	Not Satisfied	Not Satisfied
Nando Dermawan	Man	Not Clean	Expensive	Not Care	Difficult	Not Satisfied	Not Satisfied
Nasli Nasution	Man	Clean	Affordable	Not Care	Easy	Satisfied	Satisfied
Nissi Hartat Ritonga	Woman	Clean	Affordable	Not Care	Easy	Satisfied	Satisfied
Nita Hariani	Woman	Clean	Expensive	Friendly	Easy	Satisfied	Satisfied
Nur Azman	Man	Clean	Cheap	Not Care	Easy	Satisfied	Satisfied
Nur Hafni Nasution	Woman	Clean	Expensive	Friendly	Easy	Satisfied	Satisfied
Nurhayati Caniago	Woman	Clean	Cheap	Not Care	Easy	Satisfied	Satisfied
Nurul Mawaddah	Woman	Clean	Cheap	Friendly	Easy	Satisfied	Satisfied
Pidah	Woman	Not Clean	Affordable	Friendly	Easy	Satisfied	Satisfied
Quito Rian Zori	Man	Clean	Cheap	Indifferent	Easy	Satisfied	Satisfied
Ratna Dewi Tanjung	Woman	Clean	Affordable	Not Care	Easy	Satisfied	Satisfied
Reynalda Safira	Woman	Clean	Affordable	Friendly	Easy	Satisfied	Satisfied
Reynaldo Deswara	Man	Clean	Affordable	Friendly	Easy	Satisfied	Satisfied
Ridanti Mahrani	Woman	Clean	Affordable	Friendly	Easy	Satisfied	Satisfied
Rifandi Hasibuan	Man	Clean	Affordable	Not Care	Easy	Satisfied	Satisfied

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 3, July 2023
DOI : https://doi.org/10.33395/sinkron.v8i3.12674
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Rinaldi Afriza	Man	Clean	Cheap	Friendly	Easy	Satisfied	Satisfied
Rini Misriyani	Woman	Clean	Affordable	Not Care	Easy	Satisfied	Satisfied
Riswan	Man	Clean	Affordable	Not Care	Easy	Satisfied	Satisfied
Rodiah Dalimunthe	Woman	Not Clean	Cheap	Not Care	Difficult	Not Satisfied	Not Satisfied
Roma	Man	Clean	Expensive	Not Care	Difficult	Not Satisfied	Not Satisfied
Rosmawati Hazda	Woman	Clean	Affordable	Friendly	Easy	Satisfied	Satisfied
Rosnima Lubis	Woman	Clean	Affordable	Not Care	Easy	Satisfied	Satisfied
Rostika Simanjuntak	Woman	Clean	Cheap	Not Care	Easy	Satisfied	Satisfied
Rustam Hasibuan	Man	Clean	Affordable	Not Care	Easy	Satisfied	Satisfied
Safrianto	Man	Clean	Cheap	Not Care	Easy	Satisfied	Satisfied
Saifuddin Zuhri	Man	Not Clean	Cheap	Not Care	Difficult	Not Satisfied	Not Satisfied
Sarah	Woman	Clean	Cheap	Not Care	Easy	Satisfied	Satisfied
Satria Amanda Negara	Man	Not Clean	Cheap	Not Care	Difficult	Not Satisfied	Not Satisfied
Septiandi	Man	Clean	Cheap	Not Care	Easy	Satisfied	Satisfied
Siti Jubaidah	Woman	Clean	Affordable	Friendly	Easy	Satisfied	Satisfied
Sri Handayani Fitri	Woman	Not Clean	Cheap	Friendly	Easy	Satisfied	Satisfied
Sri Hera Wati Dalimunthe	Woman	Clean	Expensive	Not Care	Difficult	Not Satisfied	Not Satisfied
Sri Wahyuni Siregar	Woman	Clean	Affordable	Not Care	Easy	Satisfied	Satisfied
Susanto	Man	Clean	Expensive	Not Care	Difficult	Not Satisfied	Not Satisfied
Suwastati Sagala	Woman	Clean	Cheap	Not Care	Easy	Satisfied	Satisfied
Syafril	Man	Not Clean	Cheap	Not Care	Easy	Not Satisfied	Not Satisfied
Syiful Idek	Man	Clean	Expensive	Friendly	Easy	Satisfied	Satisfied
Tan Hartono	Man	Clean	Affordable	Not Care	Easy	Satisfied	Satisfied
Tumpal Dalimunthe	Man	Clean	Cheap	Not Care	Easy	Satisfied	Satisfied
Tutin Harmita	Woman	Clean	Affordable	Not Care	Easy	Satisfied	Satisfied
Wiwik	Woman	Clean	Cheap	Friendly	Difficult	Satisfied	Satisfied
Yogi Kharisma Dalimunthe	Man	Not Clean	Cheap	Not Care	Difficult	Not Satisfied	Not Satisfied
Yuli Wulandari	Woman	Not Clean	Expensive	Friendly	Difficult	Not Satisfied	Not Satisfied
Yurnawilis Koto	Woman	Clean	Affordable	Not Care	Easy	Satisfied	Satisfied
Yus Negayenti	Woman	Clean	Cheap	Friendly	Easy	Satisfied	Satisfied
Zirli Zaidan	Man	Clean	Affordable	Friendly	Easy	Satisfied	Satisfied
Zuanda Lubis	Man	Not Clean	Cheap	Not Care	Difficult	Not Satisfied	Not Satisfied
Zulkifli Koto	Man	Clean	Affordable	Not Care	Easy	Satisfied	Satisfied

Table 5 is the prediction results obtained from the mining process using the K-Nearest Neighbor (kNN) method and the Naïve Bayes method with a classification model. The classification results obtained from 100 visitor data using the K-Nearest Neighbor (kNN) method and the Naïve Bayes method stated that 77 visitor data were satisfied with Boombara Waterpark Rantauprapat and 23 visitor data were dissatisfied with Boombara Waterpark Rantauprapat. The classification results obtained from the two methods are the same.

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 3, July 2023
DOI : https://doi.org/10.33395/sinkron.v8i3.12674
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Classification Model Evaluation Results

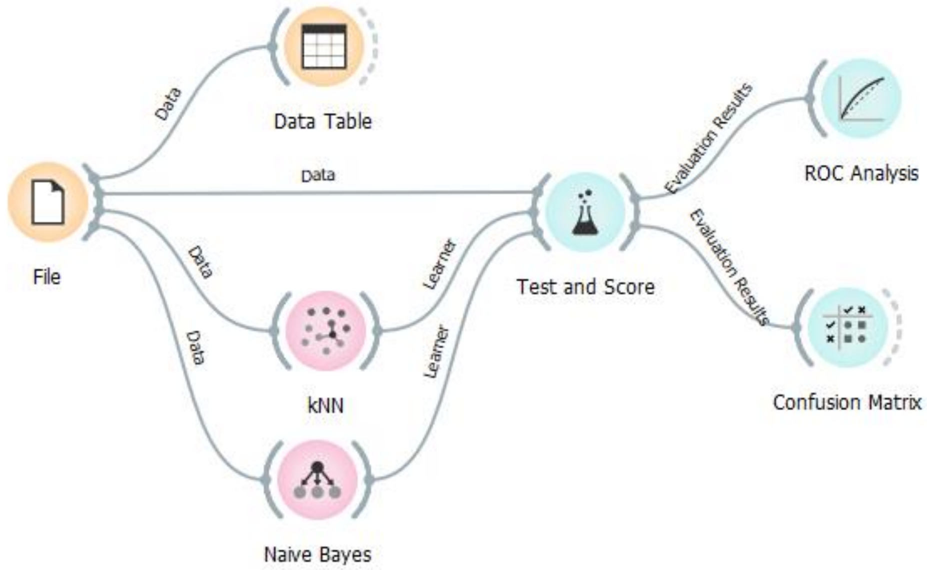

Fig 4. *Design Widget Evaluation Process*

Figure 4 is an evaluation process of the results of the classification that has been done before. The evaluation process aims to determine the accuracy of the method used. Also in this process, the author will determine the results of tests and scores and confusion matrix. Both widgets are used to determine the accuracy of the method that has been used. processed Also, the author will display the graphical results of each method used.

Evaluation Results with Test and Score

Table 6
Results with Test and Score

Model	AUC	CA	F1	Precision	Recall
kNN	0.978	0.990	0.990	0.990	0.990
Naïve Bayes	0.999	0.990	0.990	0.990	0.990

After the authors evaluate the classification model using the K-Nearest Neighbor (kNN) method, the Test and Score results for AUC are 0.978, results for CA are 0.990, results for F1 are 0.990, results for Precision are 0.990 and results for Recall are 0.990. While the evaluation results with the classification model obtained with using the Naïve Bayes method, the test and score results for AUC were 0.999, results for CA were 0.990, results for F1 were 0.990, results for Precision were 0.990 and results for Recall were 0.990.

Evaluation Result with Confusion Matrix

The confusion matrix is a widget that is used as a measuring tool for classification techniques by calculating the correctness of data that has been classified using the K-Nearest Neighbor (kNN) method and the Naïve Bayes method.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 3, July 2023
DOI : https://doi.org/10.33395/sinkron.v8i3.12674
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Table 7
Confusion Matrix results using the K-Nearest Neighbor (kNN) method

Actual \ Predicted	Satisfied	Not Satisfied	∑
Satisfied	77	0	55
Not Satisfied	1	22	45
∑	60	40	100

Table 8 is the result of the confusion matrix obtained from the evaluation of the Classification model. The results of the confusion matrix are True Positive (TP) is 77. True Negative (TN) is 22, False Positive (FP) is 0 and False Negative (FN) is 1. Then the values for accuracy, precision and recall are as follows:

$$\boldsymbol{Accuracy = \frac{77+22}{77+22+0+1} \times 100\%}$$ Then the Accuracy value = 99%

$$\boldsymbol{Presisi = \frac{77}{77+0} \times 100\%}$$ Then the Precision value = 100%

$$\boldsymbol{Recall = \frac{77}{77+1} \times 100\%}$$ Then the Recall value = 98,7%

Table 8
Confusion Matrix results using the Naïve Bayes method

Actual \ Predicted	Satisfied	Not Satisfied	∑
Satisfied	77	0	55
Not Satisfied	1	22	45
∑	60	40	100

Table 8 is the result of the confusion matrix obtained from the evaluation of the Classification model. The results of the confusion matrix are True Positive (TP) is 77. True Negative (TN) is 22, False Positive (FP) is 0 and False Negative (FN) is 1. Then the values for accuracy, precision and recall are as follows:

$$\boldsymbol{Accuracy = \frac{77+22}{77+22+0+1} \times 100\%}$$ Then the Accuracy value = 99%

$$\boldsymbol{Presisi = \frac{77}{77+0} \times 100\%}$$ Then the Precision value = 100%

$$\boldsymbol{Recall = \frac{77}{77+1} \times 100\%}$$ Then the Recall value = 98,7%

Evaluation Result with ROC Curve

Roc Analysis is obtained from the evaluation results of the Classification model with the addition of the ROC Analysis widget. The ROC Analysis results will be displayed in the form of graphic images obtained from data mining processing.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 3, July 2023
DOI : https://doi.org/10.33395/sinkron.v8i3.12674
e-ISSN : 2541-2019
p-ISSN : 2541-044X

ROC Evaluation Results Analysis using the K-Nearest Neighbor (kNN) Method

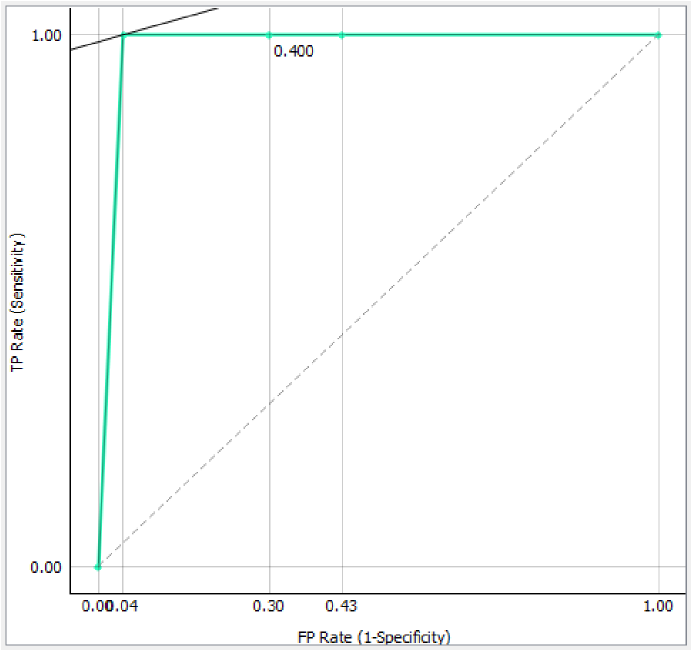

Fig 5. *ROC Analysis of Satisfied Visitors to Boombara Waterpark Rantauprapat*

Figure 5 is the ROC result Analysis Satisfied visitors at Boombara Waterpark Rantauprapat using the K-Nearest Neighbor (kNN) method. The results obtained were 0.400.

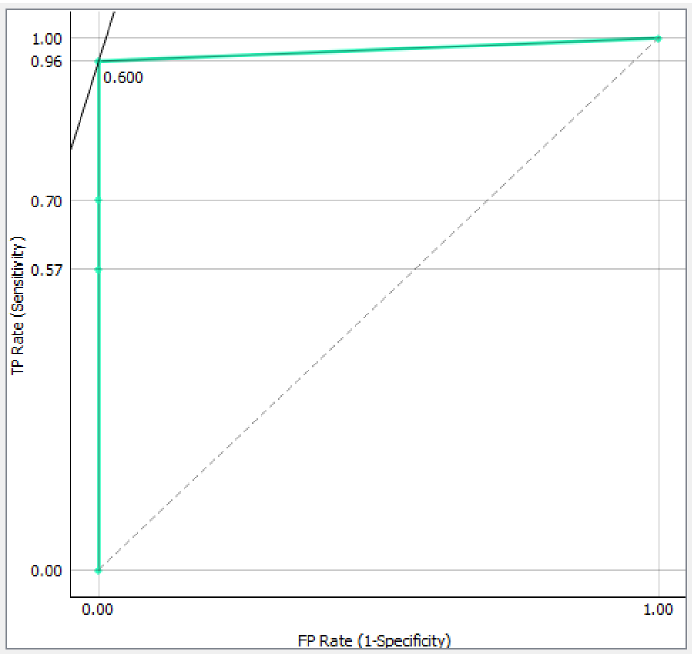

Fig 6. *ROC Analysis of Dissatisfied Visitors to Boombara Waterpark Rantauprapat*

Figure 6 is the result of the ROC analysis of dissatisfied visitors to Boombara Waterpark Rantauprapat using the K-Nearest Neighbor (KNN) method. The results obtained were 0.600.

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 3, July 2023
DOI : https://doi.org/10.33395/sinkron.v8i3.12674
e-ISSN : 2541-2019
p-ISSN : 2541-044X

ROC Evaluation Results Analysis using the Naïve Bayes Method

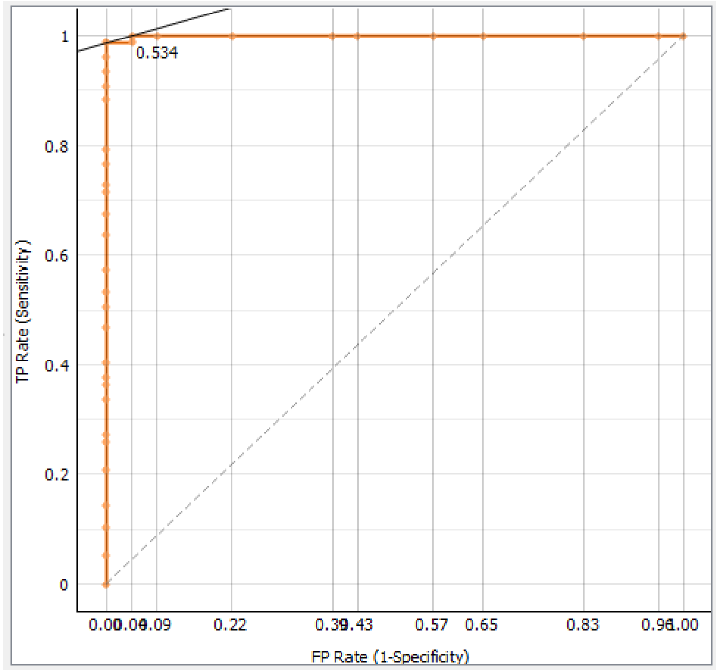

Fig 7. *ROC Analysis of Satisfied Visitors to Boombara Waterpark Rantauprapat*

Figure 7 is the ROC result Analysis Satisfied visitors at Boombara Waterpark Rantauprapat using the Naïve Bayes method. The results obtained were 0.534.

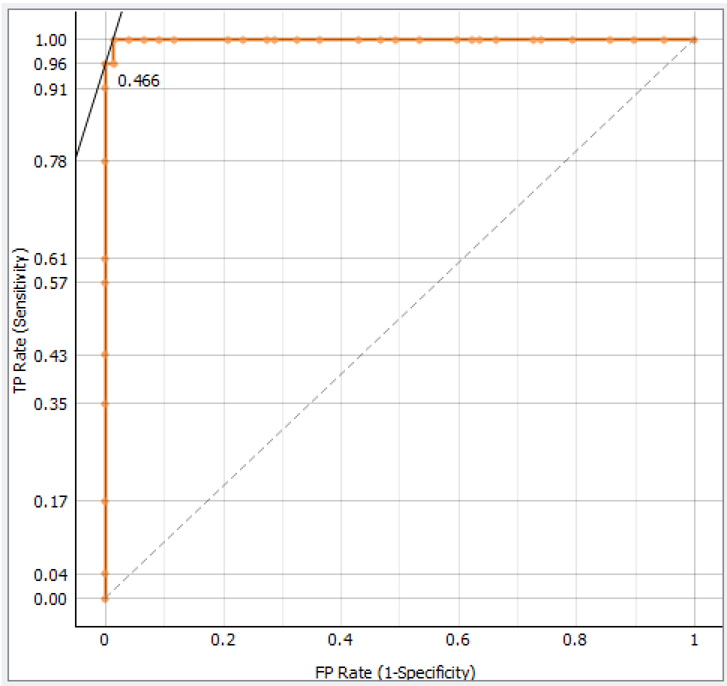

Fig 8. *ROC Analysis of Dissatisfied Visitors to Boombara Waterpark Rantauprapat*

Figure 5 is the ROC result Analysis of Dissatisfied visitors to Boombara Waterpark Rantauprapat using the Naïve Bayes method. The results obtained were 0.466.

*name of corresponding author

DISCUSSIONS

To determine whether an amusement park is good or not, it is necessary to conduct a study that can provide results about whether the place is good or bad. So in this study the author will determine the level of visitor satisfaction at Boombara Waterpark Rantauprapat. This research was conducted to determine how many or nope Satisfied visitors at Boombara Waterpark Rantauprapat. This research was conducted on data mining using 2 methods with a classification model. The methods used are the K-Nearest Neighbor (kNN) method and the Naïve Bayes method. Second method This is a method that can perform a data classification. So this method is very suitable for classifying the level of visitor satisfaction at Boombara Waterpark Rantauprapat.

The classification results obtained from 100 visitor data using the K-Nearest Neighbor (KNN) method show that 77 visitors are satisfied with Boombara Waterpark Rantauprapat and 23 visitors are dissatisfied with Boombara Waterpark Rantauprapat. The classification results obtained from 100 visitor data using the Naive Bayes method show that 77 visitors are satisfied with Boombara Waterpark Rantauprapat and 23 visitors are dissatisfied with Boombara Waterpark Rantauprapat. The results of the two methods used give good results. This is because there are still many visitors who are satisfied with Boombara Waterpark Rantauprapat. For comparison of the results of the classification of the two methods used is 1:1, this is because the results obtained have the same value.

To determine the suitability and feasibility of the method used as a method with the classification model, the authors have created a design that can be used to determine the accuracy of the method used in this study. The accuracy value will later become a reference for the feasibility and suitability of the method used to carry out a data classification. To determine accuracy results, the author uses 2 widgets that can provide accurate results from the method used, namely the Test and Score widget and the Confusion Matrix widget. The accuracy of the Test and Score results obtained using the K-Nearest Neighbor (kNN) method was 0.798 (representation obtained was 79.8%) and for the Test and Score results obtained using the Naïve Bayes method was 0.999 (representation obtained was 99 ,9%). Test and Score results obtained from both methods own the difference is quite far, namely about 20%. However, both methods are suitable for use as a method with a classification model.

For Confusion Matrix widget accuracy results using the K-Nearest Neighbor (kNN) method is 99% and for Confusion Matrix widget accuracy results using the Naïve Bayes method is 99%. The results of the accuracy of the two methods used using the Confusion Matrix widget have a very good value. This means that with the Confusion Matrix widget, the accuracy results obtained are almost perfect and for comparison the accuracy of the two methods used with the Confusion Matrix widget is 1:1. This is because the results are given the same value. So the K-Nearest Neighbor (kNN) method and the Naïve Bayes method are methods that are very suitable to be used as methods for processing data with a classification model.

CONCLUSION

The classification results obtained give good results, this is because the accuracy results obtained are almost perfect. This means that the two methods used in this study are very suitable if used as a method with a classification model. The classification results obtained from the two methods used give good results. The results obtained stated that many visitors were satisfied with Boombara Waterpark Rantauprapat. The conclusion is that Boombara Waterpark Rantauprapat has a very good appeal, this is because many visitors who can and like the space provided. With this research, it can be used as a reference to determine whether it is good or not A location can be done on data mining.

REFERENCES

Agustina, N., Adrian, A., & Hermawati, M. (2021). Implementasi Algoritma Naïve Bayes Classifier untuk Mendeteksi Berita Palsu pada Sosial Media. *Faktor Exacta*, *14*(4), 1979–276. https://doi.org/10.30998/faktorexacta.v14i4.11259

Al-Rasheed, A. (2021). Identification of important features and data mining classification techniques in predicting employee absenteeism at work. *International Journal of Electrical and Computer Engineering*, *11*(5), 4587–4596. https://doi.org/10.11591/ijece.v11i5.pp4587-4596

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 3, July 2023
DOI : https://doi.org/10.33395/sinkron.v8i3.12674
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Azzahra, S. A., & Wibowo, A. (2020). Analisis Sentimen Multi-Aspek Berbasis Konversi Ikon Emosi dengan Algoritme Naïve Bayes untuk Ulasan Wisata Kuliner Pada Web Tripadvisor. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, *7*(4), 737. https://doi.org/10.25126/jtiik.2020731907

Baharuddin, M. M., Azis, H., & Hasanuddin, T. (2019). Analisis Performa Metode K-Nearest Neighbor Untuk Identifikasi Jenis Kaca. *ILKOM Jurnal Ilmiah*, *11*(3), 269–274. https://doi.org/10.33096/ilkom.v11i3.489.269-274

Damuri, A., Riyanto, U., Rusdianto, H., & Aminudin, M. (2021). Implementasi Data Mining dengan Algoritma Naïve Bayes Untuk Klasifikasi Kelayakan Penerima Bantuan Sembako. *JURIKOM (Jurnal Riset Komputer)*, *8*(6), 219. https://doi.org/10.30865/jurikom.v8i6.3655

Di, P., & Duan, L. (2014). New naive Bayes text classification algorithm. *Shuju Caiji Yu Chuli/Journal of Data Acquisition and Processing*, *29*(1), 71–75. https://doi.org/10.11591/telkomnika.v12i2.4180

Farhad Khorshid, S., & Mohsin Abdulazeez, A. (2021). Breast Cancer Diagnosis Based on K-Nearest Neighbors: a Review. *Journal Of Archaeology Of Egypt/Egyptology*, *18*(4), 1927–1951.

Hussain, S., Dahan, N. A., Ba-Alwib, F. M., & Ribata, N. (2018). Educational data mining and analysis of students' academic performance using WEKA. *Indonesian Journal of Electrical Engineering and Computer Science*, *9*(2), 447–459. https://doi.org/10.11591/ijeecs.v9.i2.pp447-459

Indrayuni, E. (2019). Klasifikasi Text Mining Review Produk Kosmetik Untuk Teks Bahasa Indonesia Menggunakan Algoritma Naive Bayes. *Jurnal Khatulistiwa Informatika*, *7*(1), 29–36. https://doi.org/10.31294/jki.v7i1.1

Kumar, A., Chatterjee, J. M., & Díaz, V. G. (2020). A novel hybrid approach of SVM combined with NLP and probabilistic neural network for email phishing. *International Journal of Electrical and Computer Engineering*, *10*(1), 486–493. https://doi.org/10.11591/ijece.v10i1.pp486-493

Liantoni, F. (2016). Klasifikasi Daun Dengan Perbaikan Fitur Citra Menggunakan Metode K-Nearest Neighbor. *Jurnal ULTIMATICS*, *7*(2), 98–104. https://doi.org/10.31937/ti.v7i2.356

Lubis, A. R., Lubis, M., & Al-Khowarizmi. (2020). Optimization of distance formula in k-nearest neighbor method. *Bulletin of Electrical Engineering and Informatics*, *9*(1), 326–338. https://doi.org/10.11591/eei.v9i1.1464

Murwantara, I. M., Yugopuspito, P., & Hermawan, R. (2020). Comparison of machine learning performance for earthquake prediction in Indonesia using 30 years historical data. *Telkomnika (Telecommunication Computing Electronics and Control)*, *18*(3), 1331–1342. https://doi.org/10.12928/TELKOMNIKA.v18i3.14756

Negara, A. B. P., Muhardi, H., & Putri, I. M. (2020). Analisis Sentimen Maskapai Penerbangan Menggunakan Metode Naive Bayes dan Seleksi Fitur Information Gain. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, *7*(3), 599. https://doi.org/10.25126/jtiik.2020711947

Normawati, D., & Prayogi, S. A. (2021). Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter. *Jurnal Sains Komputer & Informatika (J-SAKTI*, *5*(2), 697–711. Retrieved from http://ejurnal.tunasbangsa.ac.id/index.php/jsakti/article/view/369

Patil, R., & Tamane, S. (2018). A comparative analysis on the evaluation of classification algorithms in the prediction of diabetes. *International Journal of Electrical and Computer Engineering*, *8*(5), 3966–3975. https://doi.org/10.11591/ijece.v8i5.pp3966-3975

Pour, E. S., Esmaeili, M., & Romoozi, M. (2022). Breast cancer diagnosis: a survey of pre-processing, segmentation, feature extraction and classification. *International Journal of Electrical and Computer Engineering*, *12*(6), 6397–6409. https://doi.org/10.11591/ijece.v12i6.pp6397-6409

Rosso, G. A. (2019). Milton. *William Blake in Context*, (September), 184–191. https://doi.org/10.1017/9781316534946.021

Santoso, H. A., Rachmawanto, E. H., Nugraha, A., Nugroho, A. A., Setiadi, D. R. I. M., & Basuki, R. S. (2020). Hoax classification and sentiment analysis of Indonesian news using Naive Bayes optimization. *Telkomnika (Telecommunication Computing Electronics and Control)*, *18*(2), 799–806. https://doi.org/10.12928/TELKOMNIKA.V18I2.14744

Saputra, M. F. A., Widiyaningtyas, T., & Wibawa, A. P. (2018). Illiteracy classification using K means-

*name of corresponding author

naïve bayes algorithm. *International Journal on Informatics Visualization*, *2*(3), 153–158. https://doi.org/10.30630/joiv.2.3.129

Uçar, T., & Karahoca, A. (2021). Benchmarking data mining approaches for traveler segmentation. *International Journal of Electrical and Computer Engineering*, *11*(1), 409–415. https://doi.org/10.11591/ijece.v11i1.pp409-415

Watratan, A. F., B, A. P., Moeis, D., Informasi, S., & Makassar, S. P. (2020). Implementation of the Naive Bayes Algorithm to Predict the Spread of Covid-19 in Indonesia. *Journal of Applied Computer Science and Technology*, *1*(1), 7–14.

Wijaya, A., & Girsang, A. S. (2015). Use of Data Mining for Prediction of Customer Loyalty. *CommIT (Communication and Information Technology) Journal*, *10*(1), 41. https://doi.org/10.21512/commit.v10i1.1660

Yun, H. (2021). Prediction model of algal blooms using logistic regression and confusion matrix. *International Journal of Electrical and Computer Engineering*, *11*(3), 2407–2413. https://doi.org/10.11591/ijece.v11i3.pp2407-2413

*name of corresponding author