Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 3, July 2024
DOI : https://doi.org/10.33395/sinkron.v8i3.13914
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Machine Learning to Predict Student Satisfaction Level Using KNN Method and Naive Bayes Method

Dinda Julia Arfah[1)*], Masrizal[2)], Irmayanti[3)]
[1,2,3)]Universitas Labuhanbatu, Indonesia
[1)] dindajuliaarvahharahap@gmail.com, [2)] masrizal120405@gmail.com,
[3)]irmayantiritonga2@gmail.com,

Submitted : Jul 19, 2024 | **Accepted** : Jul 29, 2024 | **Published** : Jul 30, 2024

Abstract: This research aims to apply machine learning techniques in classifying student satisfaction levels at the Faculty of Science and Technology, using the K-Nearest Neighbors (KNN) and Naive Bayes methods. This method was chosen because of its ability to manage classification data and provide accurate predictions regarding student satisfaction with the faculty. It is hoped that this research will provide a deeper understanding of the factors that influence student satisfaction as well as the potential for developing a better evaluation system in the future. This research was carried out through structured stages, starting from selecting the right data to collect relevant information, designing the model by applying the KNN and Naive Bayes methods, to evaluating the performance of the model being built. The data used consisted of 110 student data, where the classification results showed that 104 students were satisfied and 6 other students were dissatisfied with the faculty. The evaluation process produced excellent accuracy, with the Test and Score results and confusion matrix showing an accuracy level exceeding 90%. In conclusion, this research succeeded in showing that the KNN and Naive Bayes methods were effective in classifying the level of student satisfaction at the Faculty of Science and Technology. The results obtained confirm that both methods are reliable in managing and analyzing student satisfaction data efficiently, and provide valuable insights for educational institutions to improve student services and experiences in the future.

Keywords: KNN method; Naïve Bayes Method; Machine Learning; Classification Model; Satisfaction Level

INTRODUCTION

Labuhanbatu University, which is located on Jalan SisingSafegaraja No.126 A KM 3.5 Aek Tapa, Bakaran Batu, Kec. Rantau Sel., Kab. Labuhanbatu, North Sumatra, is a higher education institution that is developing in the area. The university boasts four faculties that focus on providing diverse study programs, covering key areas of science and scholarship. The faculties at Labuhanbatu University include the Faculty of Economics and Business, the Faculty of Science and Technology, the Faculty of Teacher Training and Education, and the Faculty of Law. With this diversity, the university strives to meet the needs and interests of students in various fields of study. One of the prominent faculties is the Faculty of Science and Technology, which has four superior study programs, namely Information Systems, Information Technology, Agrotechnology and Information Management. The success of this study program is reflected in the B accreditation they obtained, indicating a high standard of educational quality. However, there is one study program, namely Information Management, which has achieved B

*name of corresponding author

accreditation but is still at Diploma III (D3) level. This shows the university's focus on providing education that meets the needs of industry and society, while continuing to improve the quality and quality of graduates. Even though a faculty at a university has obtained B accreditation, this is not automatically the only benchmark that the faculty has good quality . It is important to realize that the quality of education and student satisfaction are complex aspects and require more in-depth assessment. In order to understand the extent to which the faculty meets the expectations and needs of its students, a research on the level of student satisfaction with the Faculty of Science and Technology is a very relevant step.

Labuhanbatu University has indeed achieved good educational quality standards, reflected in the accreditation it has obtained. This good accreditation is an important factor in attracting the interest of many students to choose Labuhanbatu University as a place to pursue higher education. In particular, the Faculty of Science and Technology at this university receives special attention due to its proven reputation. The satisfaction of students and female students at the Faculty of Science and Technology is certainly an aspect that needs to be considered. In a dynamic academic environment, every student has their own level of satisfaction regarding their experience while studying. This can be influenced by various factors, such as the quality of teaching, facilities, curriculum and academic support provided. In facing various challenges and competition in the world of higher education, Labuhanbatu University continues to strive to improve the quality of educational services so that they can meet the expectations and needs of students, especially in the Faculty of Science and Technology. By paying attention to student satisfaction levels, universities can better understand their needs and continue to innovate to create better learning environments and motivate students to achieve success in science and technology. Research on the level of student satisfaction at the Faculty of Science and Technology is a very relevant and significant step in understanding the dynamics of the academic environment at Labuhanbatu University. Therefore, the author wants to investigate and analyze the level of satisfaction of these students in facing the reality of conditions in their faculty. Involving the participation of students and female students who already have college experience can provide in-depth insight into various aspects that influence their satisfaction. This research will detail factors that can influence satisfaction levels, such as teaching quality, facilities, curriculum, academic support, and interaction between students and lecturers. Inclusive research methods will be used to obtain comprehensive and accurate data. Surveys, interviews and documentation analysis will be effective research instruments for evaluating the level of satisfaction of students and students.

This research will implement a machine learning approach to analyze the level of satisfaction of students at the Faculty of Science and Technology (Wahyuni, 2022). The two machine learning methods that will be applied are the Naive Bayes Method and the K-Nearest Neighbor (KNN) Method (Anggoro & Novitaningrum, 2021). This approach is expected to provide an effective solution in predicting satisfaction levels based on data collected from respondents. The Naive Bayes method will be used to classify the level of student satisfaction based on a number of identified attributes. By utilizing probability theory, the Naive Bayes algorithm can determine the probability of satisfaction levels based on related information contained in the dataset. This method is renowned for its speed and efficiency in handling categorical data. Meanwhile, the K-Nearest Neighbor (KNN) method will provide a more contextual approach by considering the proximity between instances in the attribute space (Ula, Ulva, Mauliza, Ali, & Said, 2022). By identifying patterns that may be hidden in the data and determining the level of similarity between the data, KNN can help predict the level of student satisfaction more accurately. The combination of the two methods is expected to provide stronger and more reliable results. This machine learning analysis can not only provide a deeper understanding of student satisfaction levels, but can also be the basis for developing a prediction system that can be applied more widely in the academic environment. By utilizing machine learning technology, it is hoped that this research can make a significant contribution to developing strategies for improving and increasing the quality of education at the Faculty of Science and Technology.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 3, July 2024 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v8i3.13914 p-ISSN : 2541-044X

LITERATURE REVIEW

Machine learning is a branch of artificial intelligence that focuses on developing algorithms and models that enable computers to learn and make decisions or predictions based on data (Assegie, 2021) (Nugraha & Romadhony, 2023). This process involves training a model using an existing dataset, which is then used to recognize patterns, classify data, or make predictions without being explicitly programmed (Pulgar-Sánchez et al., 2021) (Santoso, Santoso, & Sandjaya, 2023). Machine learning has a wide range of applications, including speech recognition, image analysis, medical diagnosis, and market prediction, and is significantly changing the way technology interacts with the world (Wibawa, Sutikno, & Sasongko, 2023) (Aldi, Nozomi, Sentosa, & Junaidi, 2023).

The Naive Bayes method is a classification technique based on Bayes' theorem with the assumption of independence between features (Pratama, Yanris, Nirmala, & Hasibuan, 2023) (Madjid, Ratnawati, & Rahayudi, 2023). Although this assumption is often unrealistic in the real world, this method remains effective and efficient in many situations, especially with large and complex data (Hasibuan, Sihombing, & Nasution, 2023) (Anam, Rahmiati, Paradila, Mardainis, & Machdalena, 2023). Naive Bayes works by calculating the probability of each possible class for an instance and then selecting the class with the highest probability (Lubis & Chandra, 2023) (Saleh, Dharshinni, Perangin-Angin, Azmi, & Sarif, 2023). This method is often used in text recognition, such as spam email classification, sentiment analysis, and recommendation systems.

The K-Nearest Neighbors (KNN) method is a non-parametric algorithm used for classification and regression. In essence, KNN works by identifying the closest 'k' data points in the feature space of the data to be classified or predicted (Arifuddin, Pinastawa, Anugraha, & Pradana, 2023) (Adjani, Fauzia, & Juliane, 2023). Then, the classification or prediction results are determined based on the majority of labels or values from the nearest neighbors (Istiadi, Rahman, & Wisnu, 2023). KNN is very simple and intuitive, but can be very effective with the right dataset. The main challenges in using KNN include choosing the right value for 'k' and the influence of different feature sizes and scales, which often requires data normalization (Violita, Yanris, & Hasibuan, 2023).

METHOD

The Naive Bayes and K-Nearest Neighbors (KNN) methods are two approaches that are often used in analyzing student satisfaction levels in science and technology faculties (Pratama et al., 2023). The Naive Bayes method is based on Bayes' theorem which assumes independence between features, which makes it possible to calculate the probability that a student will feel satisfied or not based on factors such as teaching quality, facilities, academic support, and social interaction (Nasution, Dar, & Nasution, 2023). Using survey data collected from students, Naive Bayes can quickly classify satisfaction responses based on patterns detected in the dataset. This method is effective due to its speed and ability to handle large amounts of data, although its assumption of independence is sometimes not completely realistic. On the other hand, the K-Nearest Neighbors (KNN) method works by identifying the 'k' nearest neighbors of each student based on the same features, such as teaching quality and facilities. By analyzing data from nearest neighbors, KNN can predict a student's level of satisfaction by comparing him with a group of students who have similar characteristics. The advantage of KNN is its ability to handle data with a non-linear structure and does not require certain data distribution assumptions. However, KNN can be inefficient with very large datasets and requires data normalization to avoid bias due to differences in feature scales. These two methods, when implemented correctly, can provide valuable insights for science and technology faculty to improve the quality of education and services provided to students. The stages in this research are as follows. The K-Nearest Neighbors (KNN) and Naive Bayes methods are employed to classify student satisfaction levels at the Faculty of Science and Technology, Universitas Labuhanbatu. KNN works by analyzing the closest data points to determine the classification of student satisfaction, leveraging its simplicity and effectiveness in handling non-linear data. On the other hand, Naive Bayes, a probabilistic classifier based on Bayes' Theorem, uses features of student feedback to predict satisfaction levels. Both methods provide a comprehensive analysis of students' satisfaction, allowing for a detailed understanding of factors contributing to whether students feel satisfied or dissatisfied with their educational experience.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 3, July 2024 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v8i3.13914 p-ISSN : 2541-044X

Data Selection : Stages carried out to collect data and select data that is suitable for use for research.
Classification Model Design : The steps taken to create or design a model to classify data.
Classification Result : stages that will provide the results of the classification in this research.
Designing Evaluation Model: The steps taken to create a model design for evaluating the methods used in this research.
Evaluation Result : Stages that will explain the results of the evaluation method used in this research.

RESULT

Data Selection

The data selection stage is the stage carried out to collect data that will be used in this research. So next the author will select the data that will be used and is suitable for use so that he can carry out research well. This step is very important to ensure that the data used for analysis meets the research objectives. By selecting relevant data, researchers can avoid noise and outliers that can interfere with the analysis results.

Data Training

Training data is data that will be used to help the data classification process. Training data is also often referred to as training data which is tasked with assisting the data classification process in order to obtain good and correct classification data.

Table 1. Data Training

Full name	Lecturer's Attitude	Bureau Services	Security	Comfort	KeCleanan Ruang Kelas	Category
Angga prayoga	Indifferent and Late Entering the Room	Friendly	Not safe	Comfortable	Clean	Satisfied
Fandini	Indifferent and Late Entering the Room	Indifferent	Not safe	Comfortable	Clean	Not satisfied
Farhan zamzami nasution	Friendly and Kind	Friendly	Safe	Uncomfortable	Clean	Satisfied
Febrita Damena Sinaga	Indifferent and Late Entering the Room	Friendly	Not safe	Uncomfortable	Not clean	Not satisfied
Holong Marluha Sihite	Indifferent and Late Entering the Room	Indifferent	Safe	Uncomfortable	Not clean	Not satisfied
Jilfia Jannah	Friendly and Kind	Indifferent	Safe	Comfortable	Clean	Satisfied
Rahmi Azizi	Friendly and Kind	Indifferent	Not safe	Comfortable	Not clean	Not satisfied
Rizki Safesyah	Friendly and Kind	Friendly	Not safe	Comfortable	Clean	Satisfied
Rudi Ansor hasibuan	Friendly and Kind	Friendly	Not safe	Uncomfortable	Not clean	Not satisfied
Sabrina Erika Kristina	Friendly and Kind	Friendly	Safe	Uncomfortable	Not clean	Satisfied

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 3, July 2024 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v8i3.13914 p-ISSN : 2541-044X

In the table above is training data that will be used to help the data classification process in data mining. With the training data above, the classification process can be carried out well. The training data that will be used is 10 training data.

Data Testing

Testing data is research sample data that is used for data classification, so testing data will later be classified with the help of training data.

Table 2. Data Testing

Full name	Lecturer's Attitude	Bureau Services	Security	Comfort	Comfort
Wahyu Hidayat	Friendly and Kind	Friendly	Safe	Comfortable	Clean
Agus Santoso	Friendly and Kind	Friendly	Safe	Comfortable	Clean
Agus Supriyadi	Friendly and Kind	Friendly	Safe	Comfortable	Clean
Ahmad Fadli	Friendly and Kind	Friendly	Safe	Comfortable	Clean
Ali Maulana	Friendly and Kind	Friendly	Safe	Comfortable	Clean
Andi Anggraeni	Friendly and Kind	Friendly	Safe	Comfortable	Clean
Andi Pratama	Friendly and Kind	Friendly	Safe	Comfortable	Clean
Andi Wijaya	Friendly and Kind	Friendly	Safe	Comfortable	Clean
Angga prayoga	Indifferent and Late Entering the Room	Indifferent	Safe	Comfortable	Clean
Anisa Puspita	Friendly and Kind	Friendly	Safe	Comfortable	Not clean
Anisa Putri	Friendly and Kind	Friendly	Safe	Comfortable	Not clean
Arief Wibowo	Friendly and Kind	Friendly	Safe	Comfortable	Clean
Arvida	Friendly and Kind	Friendly	Safe	Comfortable	Clean
Ayu Lestari	Friendly and Kind	Friendly	Safe	Comfortable	Clean
Bambang Susilo	Friendly and Kind	Friendly	Safe	Comfortable	Clean
Beni Rahman	Friendly and Kind	Friendly	Safe	Comfortable	Clean
Beni Setiawan	Friendly and Kind	Friendly	Safe	Comfortable	Clean
Budi Setiawan	Friendly and Kind	Friendly	Safe	Comfortable	Clean
Dedi Kusuma	Friendly and Kind	Friendly	Safe	Comfortable	Clean

In the table above is the testing data that will be used and classified in machine learning. For the testing data used, there were 110 student data. However, what the author shows above is only 19 data, the data will all be displayed later in the classification results.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 3, July 2024 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v8i3.13914 p-ISSN : 2541-044X

Classification Model Design

Classification model design is a process carried out to design a classification model that will be used to classify data.

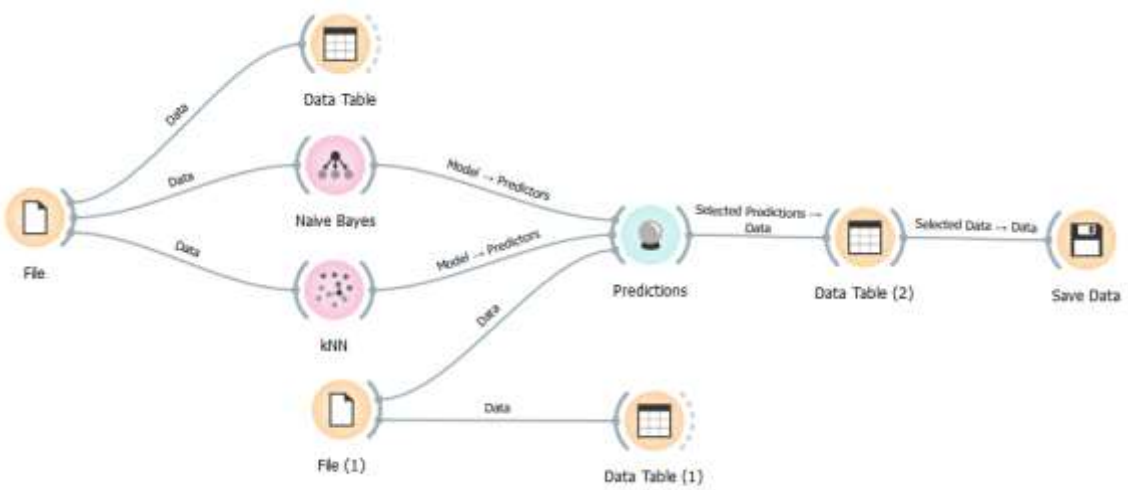

Fig 1. Classification Models

The image above is a model designed so that classification can be carried out correctly. The model above was designed using the orange application. The methods used are the Naïve Bayes method and the KNN method.

Classification Result

The author describes the classification results obtained in tabular form which can be seen in the table below.

Table 3. Classification Results Data

Full name	Lecturer's Attitude	Bureau Services	Security	Comfort	Comfort	Category
Agus Santoso	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Agus Supriyadi	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Ahmad Fadli	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Ali Maulana	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Andi Anggraeni	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Andi Pratama	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Andi Wijaya	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Angga prayoga	Indifferent and Late Entering the Room	Indifferent	Safe	Comfortable	Clean	Not satisfied
Anisa Puspita	Friendly and Kind	Friendly	Safe	Comfortable	Not clean	Satisfied

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 3, July 2024 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v8i3.13914 p-ISSN : 2541-044X

Anisa Putri	Friendly and Kind	Friendly	Safe	Comfortable	Not clean	Satisfied
Arief Wibowo	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Arvida	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Ayu Lestari	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Bambang Susilo	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Beni Rahman	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Beni Setiawan	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Budi Setiawan	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Dedi Kusuma	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Desi Kartika	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Desi Puspita	Indifferent and Late Entering the Room	Friendly	Safe	Uncomfortable	Not clean	Satisfied
Dewi Anggraini	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Dian Permana	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Dinda Julia Arfah	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Dini Andriani	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Dini Pratiwi	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Doni Pratama	Indifferent and Late Entering the Room	Indifferent	Safe	Comfortable	Clean	Not satisfied
Doni Saputra	Friendly and Kind	Friendly	Safe	Comfortable	Not clean	Satisfied
Eka Anggraini	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Eka Saputra	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Fadli Akbar	Indifferent and Late Entering the Room	Indifferent	Safe	Comfortable	Clean	Not satisfied
Fajar Nugraha	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Fajar Setiawan	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Fandini	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 3, July 2024
DOI : https://doi.org/10.33395/sinkron.v8i3.13914
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Farhan zamzami	Friendly and Kind	Friendly	Safe	Comfortable	Not clean	Satisfied
Febrita Damena Sinaga	Friendly and Kind	Friendly	Safe	Comfortable	Not clean	Satisfied
Fika Nurhaliza	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Fitri Permata	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Hana Tasya	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Hasan Basri	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Hasanudin Pratama	Friendly and Kind	Friendly	Safe	Comfortable	Not clean	Satisfied
Hendra Gunawan	Friendly and Kind	Friendly	Safe	Comfortable	Not clean	Satisfied
Hendri Susanto	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Heri Gunawan	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Hesti Putri	Friendly and Kind	Friendly	Safe	Comfortable	Not clean	Satisfied
Imam Maulana	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Indra Gunawan	Friendly and Kind	Friendly	Safe	Comfortable	Not clean	Satisfied
Intan Pratama	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Ira Kartika	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Irfan Maulana	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Irfan Santoso	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Jilfia Jannah	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Joko Susanto	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Lia Anggraini	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Linda Pratiwi	Friendly and Kind	Friendly	Safe	Comfortable	Not clean	Satisfied
Lisa Anggraeni	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Lusi Apriani	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Lusi Kartika	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Maya Pratiwi	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 3, July 2024
DOI : https://doi.org/10.33395/sinkron.v8i3.13914
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Maya Sari	Indifferent and Late Entering the Room	Friendly	Safe	Uncomfortable	Not clean	Satisfied
Mira Santoso	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Mita Anggraini	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Nani Permata	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Nia Aulia Sari Lubis	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Nia Lestari	Friendly and Kind	Friendly	Safe	Comfortable	Not clean	Satisfied
Nina Kartika	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Nisa Kartika	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Putra Andika Ritonga	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Putri Permata	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Rahmadhani Rtg	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Rahmat Hidayat	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Rahmi Azizi	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Ratna Dewi	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Reza Pratama	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Rian Nugraha	Friendly and Kind	Friendly	Safe	Comfortable	Not clean	Satisfied
Rika Dewi	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Rina Anggraeni	Indifferent and Late Entering the Room	Indifferent	Safe	Uncomfortable	Not clean	Not satisfied
Rina Kartika	Friendly and Kind	Friendly	Safe	Comfortable	Not clean	Satisfied
Rina Santoso	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Rio Saputra	Indifferent and Late Entering the Room	Indifferent	Not safe	Comfortable	Clean	Not satisfied
Risky Amaldi Harahap	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Rizki Safesyah	Friendly and Kind	Friendly	Safe	Comfortable	Not clean	Satisfied
Rizky Putra	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 3, July 2024 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v8i3.13914 p-ISSN : 2541-044X

Rudi Ansor hasibuan	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Rudi Pratama	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Rudi Setiawan	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Sabrina Erika	Indifferent and Late Entering the Room	Friendly	Safe	Uncomfortable	Not clean	Satisfied
Salman Al farisi	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Santi Lestari	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Sari Anggraini	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Sari Pratama	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Sinta Melati	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Siska Puspita	Indifferent and Late Entering the Room	Friendly	Safe	Uncomfortable	Not clean	Satisfied
Siti Nurhaliza	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Taufik Hidayat	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Taufik Pratama	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Tommy Firdaus	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Tommy Saputra	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Tongku Hamonangan	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Vivi Lestari	Friendly and Kind	Friendly	Safe	Comfortable	Not clean	Satisfied
Wahyu Hidayat	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Wahyu Nugroho	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Widianto Pratama	Friendly and Kind	Friendly	Safe	Comfortable	Not clean	Satisfied
Widianto Rahman	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Wulan Permata	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Wulan Santoso	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Yudi Pratama	Indifferent and Late Entering the Room	Indifferent	Safe	Comfortable	Clean	Not satisfied

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 3, July 2024 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v8i3.13914 p-ISSN : 2541-044X

Yudi Santoso	Indifferent and Late Entering the Room	Friendly	Safe	Uncomfortable	Not clean	Satisfied
Yuni Franata Sinurat	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Yuni Saputri	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied
Yusuf Kurniawan	Friendly and Kind	Friendly	Safe	Comfortable	Clean	Satisfied

The classification results obtained showed that of the 110 student data used, 104 were satisfied student data at the Faculty of Science and Technology and the remaining 6 student data were dissatisfied at the Faculty of Science and Technology.

Designing Evaluation Model

Evaluation model design is a model designed to obtain evaluation results from the method used. The evaluation model will also be designed in the Orange application.

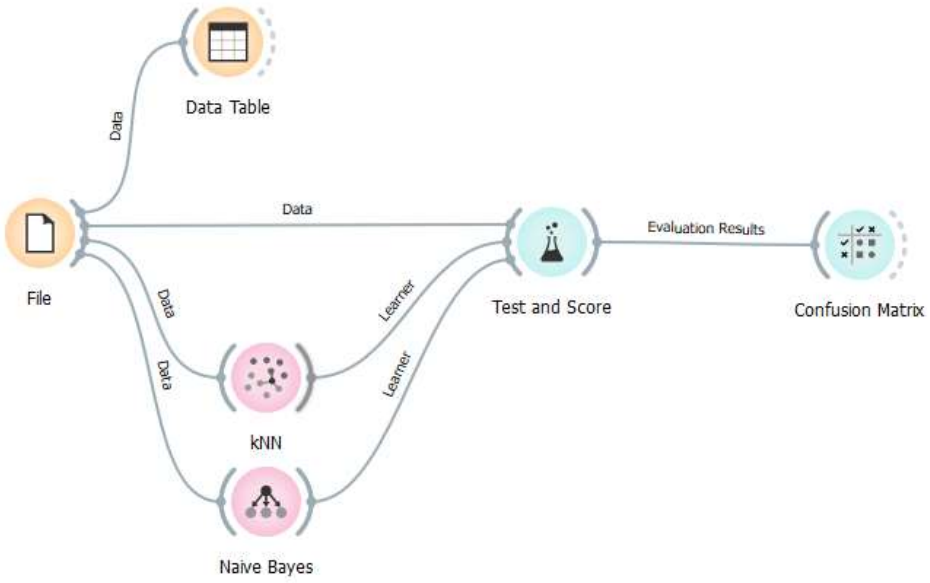

Fig 2. Evaluation Model

The image above is an image designed in the orange application. The purpose of designing the model above is to find out the results of method evaluation. For the evaluation results, the author will use 2 widgets, namely Test and Score and the Confusion Matrix widget.

Evaluation Result

Test and Score

Table 4. Test and Score

Model	AUC	CA	F1	Precision	Recall	MCC
KNN	0.913	0.991	0.991	0.991	0.991	0.909
Naïve Bayes	0.984	0.955	0.961	0.975	0.955	0.721

The evaluation results provided by the test and score widget are very large, although the results do not reach 100%. But the results are enormous. The accuracy results from the KNN method are 91% and

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 3, July 2024
DOI : https://doi.org/10.33395/sinkron.v8i3.13914
e-ISSN : 2541-2019
p-ISSN : 2541-044X

the accuracy results provided by the Naïve Bayes method are 98%. The two results provided state that this method is very suitable for use in this research.

Confusion Matrix

Table 5. Confusion Matrix Results in the K-Nearest Neighbor (KNN) Method

		Predicted		
		Satisfied	Not Satisfied	∑
Actual	Satisfied	104	0	104
	Not Satisfied	1	5	6
	∑	105	5	110

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 104 data, for True Negative (TN) results which are 5 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.

$$\boldsymbol{Accuracy} = \frac{\mathbf{104+5}}{\mathbf{104+5+0+1}} \mathbf{+ 100\%}$$ Then the Accuracy value = 99%

$$\boldsymbol{Precision} = \frac{\mathbf{104}}{\mathbf{104+0}} \mathbf{+ 100\%}$$ Then the Precision value = 100%

$$\boldsymbol{Recall} = \frac{\mathbf{104}}{\mathbf{104=1}} \mathbf{+ 100\%}$$ Then the Recall value = 99%

For the accuracy results of the confusion matrix for the KNN method, the results obtained are very large. This is because the results obtained almost reached 100%, namely the accuracy obtained was 99%.

Table 6. Confusion Matrix Results in the Naïve Bayes Method

		Predicted		
		Satisfied	Not Satisfied	∑
Actual	Satisfied	99	5	104
	Not Satisfied	0	6	6
	∑	99	11	110

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 99 data, for True Negative (TN) results which are 6 data, for False Positive (FP) results which are 5 and for False Negative (FN) results which are 0 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.

$$\boldsymbol{Accuracy} = \frac{\mathbf{99+6}}{\mathbf{99+6+5+0}} \mathbf{+ 100\%}$$ Then the Accuracy value = 95%

$$\boldsymbol{Precision} = \frac{\mathbf{99}}{\mathbf{99+5}} \mathbf{+ 100\%}$$ Then the Precision value = 95%

$$\boldsymbol{Recall} = \frac{\mathbf{99}}{\mathbf{99+0}} \mathbf{+ 100\%}$$ Then the Recall value = 100%

The accuracy results obtained by the confusion matrix for the Naïve Bayes method are very large. This is because the results obtained exceed the value of 90%. The results obtained were 95%.

*name of corresponding author

DISCUSSIONS

This research aims to analyze the level of student satisfaction at the Faculty of Science and Technology using the K-Nearest Neighbors (KNN) and Naive Bayes methods. Of the 110 student data analyzed, the classification results showed that 104 students were satisfied with the faculty, while 6 other students were dissatisfied. The KNN and Naive Bayes methods were chosen because of their ability to handle classification data and provide accurate results in determining the level of satisfaction. The analysis process begins with data pre-processing to ensure that the data used is clean and ready for analysis. The data is then divided into two groups: training data and testing data. KNN and Naive Bayes algorithms are applied to training data to build a classification model, which is then tested with test data to evaluate model performance. The evaluation results show that these two methods provide a very high level of accuracy, with Test and Score results reaching more than 90%.

Further evaluation using the confusion matrix also showed consistent results with an accuracy rate of more than 90%. The confusion matrix provides an overview of the number of correct and incorrect predictions made by the model, which shows that the KNN and Naive Bayes methods are able to predict student satisfaction levels well. These high results confirm that both methods are very effective in this research, able to capture patterns and trends in the student data used. Overall, the results of this research show that the KNN and Naive Bayes methods are very suitable for analyzing student satisfaction levels at the Faculty of Science and Technology. The high level of accuracy provides confidence that the model built is reliable and valid in this context. These findings provide valuable insight for faculty to understand student satisfaction and can be used as a basis for improving the quality of services and student experiences in the future.

CONCLUSION

This research aims to analyze the level of student satisfaction at the Faculty of Science and Technology using the K-Nearest Neighbors (KNN) and Naive Bayes methods. Of the 110 student data analyzed, the classification results showed that 104 students were satisfied with the faculty, while 6 other students were dissatisfied. The KNN and Naive Bayes methods were chosen because of their ability to handle classification data and provide accurate results in determining student satisfaction levels. The evaluation results show that these two methods provide a very high level of accuracy, with Test and Score results reaching more than 90%. The confusion matrix also shows consistent results with an accuracy rate of more than 90%. These high results confirm that the KNN and Naive Bayes methods are very effective and suitable for use in this research, able to predict the level of student satisfaction at the Faculty of Science and Technology.

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

Aldi, F., Nozomi, I., Sentosa, R. B., & Junaidi, A. (2023). Machine Learning to Identify Monkey Pox Disease. *Sinkron*, *8*(3), 1335–1347. https://doi.org/10.33395/sinkron.v8i3.12524

Anam, M. K., Rahmiati, R., Paradila, D., Mardainis, M., & Machdalena, M. (2023). Application of Naïve Bayes Algorithm for Non-Cash Food Assistance Recipients in Kampar Regency. *Sinkron*, *8*(1), 433–441. https://doi.org/10.33395/sinkron.v8i1.12032

Anggoro, D. A., & Novitaningrum, D. (2021). Comparison of accuracy level of support vector machine (SVM) and artificial neural network (ANN) algorithms in predicting diabetes mellitus disease. *ICIC Express Letters*, *15*(1), 9–18. https://doi.org/10.24507/icicel.15.01.9

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

*name of corresponding author

Assegie, T. A. (2021). An optimized K-Nearest neighbor based breast cancer detection. *Journal of Robotics and Control (JRC)*, *2*(3), 115–118. https://doi.org/10.18196/jrc.2363

Hasibuan, S. A., Sihombing, V., & Nasution, F. A. (2023). Analysis of Community Satisfaction Levels using the Neural Network Method in Data Mining. *Sinkron*, *8*(3), 1724–1735. https://doi.org/10.33395/sinkron.v8i3.12634

Istiadi, I., Rahman, A. Y., & Wisnu, A. D. R. (2023). Identification of Tempe Fermentation Maturity Using Principal Component Analysis and K-Nearest Neighbor. *Sinkron*, *8*(1), 286–294. https://doi.org/10.33395/sinkron.v8i1.12006

Lubis, A. I., & Chandra, R. (2023). Forward Selection Attribute Reduction Technique for Optimizing Naïve Bayes Performance in Sperm Fertility Prediction. *Sinkron*, *8*(1), 275–285. https://doi.org/10.33395/sinkron.v8i1.11967

Madjid, F. M., Ratnawati, D. E., & Rahayudi, B. (2023). Sentiment Analysis on App Reviews Using Support Vector Machine and Naïve Bayes Classification. *Jurnal Dan Penelitian Teknik Informatika*, *8*(1), 556–562. Retrieved from https://doi.org/10.33395/sinkron.v8i1.12161

Nasution, R. F., Dar, M. H., & Nasution, F. A. (2023). Implementation of the Naïve Bayes Method to Determine Student Interest in Gaming Laptops. *Sinkron*, *8*(3), 1709–1723. https://doi.org/10.33395/sinkron.v8i3.12562

Nugraha, A. B., & Romadhony, A. (2023). Identification of 10 Regional Indonesian Languages Using Machine Learning. *Sinkron*, *8*(4), 2203–2214. https://doi.org/10.33395/sinkron.v8i4.12989

Pratama, H. A., Yanris, G. J., Nirmala, M., & Hasibuan, S. (2023). *Implementation of Data Mining for Data Classification of Visitor Satisfaction Levels*. *8*(3), 1832–1851.

Pulgar-Sánchez, M., Chamorro, K., Fors, M., Mora, F. X., Ramírez, H., Fernandez-Moreira, E., & Ballaz, S. J. (2021). Biomarkers of severe COVID-19 pneumonia on admission using data-mining powered by common laboratory blood tests-datasets. *Computers in Biology and Medicine*, *136*(July). https://doi.org/10.1016/j.compbiomed.2021.104738

Saleh, A., Dharshinni, N., Perangin-Angin, D., Azmi, F., & Sarif, M. I. (2023). Implementation of Recommendation Systems in Determining Learning Strategies Using the Naïve Bayes Classifier Algorithm. *Sinkron*, *8*(1), 256–267. https://doi.org/10.33395/sinkron.v8i1.11954

Santoso, B. C., Santoso, H., & Sandjaya, J. (2023). Development of Independent Taekwondo Training Machine Learning With 3D Pose Model Mediapipe. *Sinkron*, *8*(3), 1427–1434. https://doi.org/10.33395/sinkron.v8i3.12571

Ula, M., Ulva, A. F., Mauliza, M., Ali, M. A., & Said, Y. R. (2022). Application of Machine Learning in Determining the Classification of Children'S Nutrition With Decision Tree. *Jurnal Teknik Informatika (Jutif)*, *3*(5), 1457–1465. https://doi.org/10.20884/1.jutif.2022.3.5.599

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

Wahyuni, W. (2022). Analisis Sentimen terhadap Opini Feminisme Menggunakan Metode Naive Bayes. *Jurnal Informatika Ekonomi Bisnis*, *4*, 148–153. https://doi.org/10.37034/infeb.v4i4.162

Wibawa, H. A., Sutikno, S., & Sasongko, P. S. (2023). Optic Disc Detection on Retina Image using Extreme Learning Machine. *SinkrOn*, *8*(2), 1064–1073. https://doi.org/10.33395/sinkron.v8i2.12123

*name of corresponding author