

Application of the KNN Algorithm to Assess Customer Satisfaction at A2 Collection Sei Silau Timur

Lutfi Anniswa Sitorus ^{1*}, Jeperson Hutahaean ², Cecep Maulana ³

^{1*,2,3} Information Systems Study Program, Faculty of Computer Science, Universitas Royal, Asahan Regency, North Sumatra, Indonesia

Email: lutfianniswastr@gmail.com ^{1*}, jepersonhutahean@gmail.com ², maulanac36@gmail.com ³

Article history:
Received March 14, 2026
Revised April 6, 2026
Accepted April 8, 2026

Abstract

The development of information technology and data mining in recent years has changed the way retail businesses, including small and medium-scale fashion businesses, collect, analyze, and utilize customer data to improve services and marketing strategies. In addition, through the main discussion carried out in this study, it aims to be able to analyze the factors that influence the level of customer satisfaction at Amel Fashion Prapat Janji based on the attributes of product quality, price, comfort of use, and service. In addition, this study develops and applies the K-Nearest Neighbor (K-NN) algorithm to classify the level of customer satisfaction more objectively, measurably and data-based. And in addition, for the Research Method section used in this study, a qualitative approach was chosen because the focus of this study is to explore the meaning, perception, and direct experience of business actors in the marketing and distribution process. So based on that, this study shows the results that the application of the K-Nearest Neighbor (KNN) algorithm in the customer satisfaction classification system at A2 Collection Sei Silau Timur is able to provide an effective solution in managing and analyzing customer evaluation data. This website-based system has succeeded in changing the assessment process that was previously carried out manually to be more structured, systematic, and easily accessible. Based on the system's calculations, the resulting distance values, such as 2.354, categorized as "Satisfied" and 2.325, categorized as "Dissatisfied," indicate that the proximity of attribute values significantly influences the classification results. Although the difference in distance values is relatively small, the system is still able to determine the class based on the dominance of the nearest neighbor data.

Keywords:

Utilization of the K-Nearest Neighbor (KNN) Algorithm; Data Mining Concept; Customer Satisfaction Classification; A2 Collection Store.

1. INTRODUCTION

The development of information technology and data mining in recent years has changed the way retail businesses, including small and medium-sized fashion businesses, collect, analyze, and utilize customer data to improve services and marketing strategies. In the fashion sector, customer satisfaction is not only a temporary metric of consumer satisfaction, but also a determinant of loyalty, repeat purchase frequency, and word-of-mouth recommendations that directly affect business continuity (Amalia, 2024). Therefore, fashion businesses need to have a customer satisfaction assessment mechanism that is structured, objective, and adaptive to changes in consumer preferences. The use of analytical methods such as classification algorithms in data mining provides opportunities for business owners to transform customer data into operational information, for example, identifying customers at risk of switching, products that frequently generate complaints, or service attributes that need improvement (Arini et al., 2025).

Amel Fashion Prapat Janji, a local fashion retailer, faces challenges typical of Indonesian MSMEs: limited resources for in-depth customer research, manual or unstructured feedback collection methods, and difficulty prioritizing data-driven improvement actions. Frequent complaints in fashion retail businesses,

such as size discrepancies, limited stock variety, discrepancies between product images and physical product expectations, and speed of service, indicate that intuitive satisfaction analysis alone is no longer sufficient (Banjarnahor et al., 2023). To develop an effective improvement strategy, Amel Fashion requires an analytical tool capable of processing both unstructured and structured data, classifying customer satisfaction levels, and generating operationalizable recommendations. Studies on the application of classification techniques to customer satisfaction show that an algorithm-based approach can provide a more objective and replicable picture of satisfaction patterns (Dhewayani et al., 2022).

Based on customer satisfaction data for 15 products sold at Amel Fashion Store, it can be seen that there is a quite striking difference in assessment between products that satisfy and do not satisfy customers. Several products such as Patterned Hijabs, Oversized T-shirts, Casual Women's Blouses, and Premium Cotton Dresses received high scores on aspects of quality, price, comfort, and service, resulting in a stable level of satisfaction. Meanwhile, products such as Knitted Cardigans, Unisex Flannel Shirts, and Women's Sneakers received lower scores on several attributes, indicating that these products do not meet customer expectations. This imbalance between products with high and low levels of satisfaction indicates that the quality of products and services provided is still uneven (Diansyah, 2022).

These conditions can affect customer perceptions of overall service quality. Products that are deemed unsatisfactory have the potential to reduce customer trust, especially if issues such as low quality, inappropriate pricing, or user inconvenience are not immediately resolved. Therefore, Amel Fashion Store needs to conduct a more comprehensive evaluation of products that receive low scores, as well as improve service standards to meet customer expectations. Utilizing analytical methods such as the K-Nearest Neighbor (KNN) algorithm can also help identify customer satisfaction patterns so that the store can make more informed decisions in improving product and service quality (Fahmi et al., 2025).

One of the simplest yet proven effective classification algorithms in analyzing customer data is K-Nearest Neighbor (K-NN). This algorithm works with a non-parametric principle, where each new data will be classified based on a number of nearest neighbors (k) in the feature space, then the class is determined through the majority of labels from those neighbors. With its characteristics as lazy learning, K-NN does not require a complex model training process so it is very suitable for use by business actors who have limited computing resources but need fast and accurate analysis results. In addition, the ease of interpretation offered by K-NN allows business owners to review nearest neighbors as a basis for decision making, so that the process of evaluating and improving service quality can be carried out in a more focused and data-based manner (Fansyuri, 2020).

Previous research has shown that the K-Nearest Neighbor method can help classify customer satisfaction levels more accurately by comparing new customer data to previous customer patterns or characteristics. Through this approach, companies can determine whether customer satisfaction is in the satisfied, less satisfied, or dissatisfied category, so that the relationship between employee performance and customer satisfaction levels can be analyzed more clearly and measurably (Heti Aprilianti et al., 2025).

The application of K-NN to customer satisfaction studies offers several practical advantages for Amel Fashion. First, K-NN can utilize various relevant input variables such as product quality, size accuracy, stock variety, price, shopping experience (cashier service, store layout), convenience (cleanliness, facilities), and transaction processing speed. Second, because K-NN does not assume a specific distribution of data, this algorithm is flexible in dealing with heterogeneous data, for example, the combination of numerical and categorical variables commonly found in customer satisfaction surveys [9]. Third, the classification results (e.g., categories: very satisfied, satisfied, sufficient, dissatisfied) facilitate customer segmentation mapping so that improvement strategies can be directed specifically, for example, prioritizing stock improvements for segments that complain most frequently, or retention programs for customers at risk of leaving. Applied research in Indonesia confirms that the K-NN algorithm is able to classify customer satisfaction levels with adequate performance in various service and retail domains (Indraputra & Fitriana, 2020).

However, there are several challenges and limitations that need to be considered when implementing K-NN to assess customer satisfaction at Amel Fashion. The first challenge is data quality and completeness: K-NN performance is highly dependent on the representativeness of historical data, feature scale, and handling of missing values or outliers. If survey data is scarce or biased (e.g., a predominance of highly satisfied respondents), classification results can be misleading (Iranti & Nuryana, 2024). The second challenge concerns the choice of the k parameter and the distance metric. Choosing too small a k makes the model sensitive to noise, while choosing too large a k can obscure class distinctions; the choice of distance metric (Euclidean, Manhattan, or a metric specific to categorical data) also affects the results. The third challenge is interpretability in business practice: while K-NN is relatively easy to explain, converting classification results into operational policies requires cross-functional understanding (stock, service, marketing). Literature on K-NN applications in similar cases often includes data preprocessing, normalization, and cross-validation to ensure reliable results (Lillah et al., 2023).

The application of K-NN at Amel Fashion is not only to determine whether customers are satisfied or not, but also helps store owners make more informed decisions. The results of the K-NN classification can be used to see which products need to be improved, manage stock so that there is no excess or shortage, improve service quality, and create promotions that are more in line with customer needs. K-NN analysis can be

followed up immediately so that stores can improve services and increase customer satisfaction more quickly and effectively (Muhamad Soleh, Nurul Hidayati, 2021).

Previous research applying K-NN to customer satisfaction often indicates the need for cross-validation to obtain robust k-parameters and avoid overfitting or underfitting. With a careful research design, K-NN can provide a fairly reliable estimate of satisfaction levels for operational improvement measures at Amel Fashion (Nasrullah, 2021). Based on these problems, the author conducted research and presented it in the form of a thesis with the title: "Application of the KNN Algorithm to Assess Customer Satisfaction at A2 Collection Sei Silau Timur" is relevant and urgent. This research is expected to not only present an academic evaluation of the accuracy of K-NN in the context of local fashion retail, but also provide practical guidance for Amel Fashion to improve the quality of products and services measurably. The results of the research are expected to help business owners formulate evidence-based improvement priority policies so as to increase satisfaction, loyalty, and business competitiveness in the increasingly competitive local market.

2. RESEARCH METHOD

Through the research method section is an important stage used as a guideline in conducting research so that the results obtained can be scientifically accounted for. In this study, the method used is designed to support the process of implementing the K-Nearest Neighbor (KNN) algorithm in assessing the level of customer satisfaction at A2 Collection Sei Silau Timur. The approach taken includes the process of data collection, data processing, to the analysis stage using the KNN algorithm to produce accurate information related to customer satisfaction. With a systematic and structured research method, it is hoped that this study will be able to provide a clear picture of the level of customer satisfaction and become a basis for decision making to improve service quality (Pratama, 2023).

2.1. Identification of Problems

Problem identification is the initial stage in conducting research that aims to obtain the information needed to achieve research objectives. The problems at Amel Fashion Prapat Janji are less than optimal promotion, manual distribution recording, and lack of sales data integration (Pratama, 2023).

2.2. Data Collection

Data collection is the process of gathering and measuring information from various sources to obtain a complete and accurate picture. Data were collected through in-depth interviews, field observations, and documentation of business activities. Primary data were obtained directly from the distributor owner of Toko Pada Amel Fashion Prapat Janji through direct interaction in the field, while secondary data as supporting data were obtained from documents or business records in the form of the organizational structure of Toko Pada Amel Fashion Prapat Janji, sales data in the last few months, product distribution and delivery records, promotional media (Rahmadinia, Enjel Erika LorencisLubis, Aji Priansyah, Yolanda R.W.N., 2023).

2.3. Research Data Set

This research dataset is compiled based on customer assessment data for products available at A2 Collection Sei Silau Timur. The data consists of 15 product entries with several main attributes, namely quality, price, comfort of use, and service, each of which is assessed using a scale of 1 to 5. In addition, there are customer satisfaction labels classified into "satisfied" and "dissatisfied" as targets in the application of the K-Nearest Neighbor (KNN) algorithm. Each row of data represents one type of product, such as blouses, pants, robes, and shoes, which are assessed directly by customers based on their experiences. Variations in values for each criterion indicate differences in customer perceptions of the products offered. This dataset plays an important role as training and test data in the KNN classification process, so that the system can predict the level of customer satisfaction with new products based on the similarity of attribute values they have (Ratih Yulia Hayuningtyas, 2022).

Table 1. Sales Product Data

No	Product name	Quality (1-5)	Price (1-5)	Wearing Comfort (1-5)	Service (1-5)	Satisfaction (Satisfied/Dissatisfied)
1	Casual Women's Blouse	5	4	5	4	Satisfied
2	Women's Jeans	4	3	4	4	Satisfied
3	Premium Cotton Gamis	5	4	5	5	Satisfied
4	Unisex Hoodie	4	3	4	3	Satisfied
5	Oversized T-shirt	5	5	5	4	Satisfied
6	Pleated Skirt	4	4	4	4	Satisfied
7	Knitted Cardigan	3	3	3	4	Not satisfied
8	Unisex Flannel Shirt	3	2	3	3	Not satisfied
9	Girls' Outfits	4	4	4	5	Satisfied

10	Patterned Hijab	5	5	5	5	Satisfied
11	Inner Cuff	4	4	5	4	Satisfied
12	Teenage Dress	4	3	4	4	Satisfied
13	Women's Sneakers	3	3	3	3	Not satisfied
14	Fashion Sling Bag	4	4	4	3	Satisfied
15	Women's Wallet	4	3	4	4	Satisfied

2.4. Research Variables

The variables used in this study consist of independent variables and dependent variables. Independent variables include product quality, price, comfort of use, and service, each of which is measured using a rating scale of 1 to 5. These four variables represent the main factors that influence customer perceptions of a product. Meanwhile, the dependent variable in this study is the level of customer satisfaction which is classified into two categories, namely "satisfied" and "dissatisfied". The relationship between independent and dependent variables is analyzed using the K-Nearest Neighbor (KNN) algorithm to find proximity patterns between data. Thus, the system can classify the level of customer satisfaction based on the similarity of attribute values. The selection of these variables is based on their relevance in reflecting the overall customer experience with the product offered (Suprianto et al., 2024).

2.5. Analysis Stages

The analysis stage in this study begins with the data collection process obtained through customer assessments of products based on several criteria, namely quality, price, comfort of use, and service. Next, a data pre-processing stage is carried out to ensure the data is clean, consistent, and ready to be used in the calculation process. After that, the data is divided into training data and test data as a basis for implementing the K-Nearest Neighbor (KNN) algorithm. In the main analysis stage, distance calculations between data are carried out using certain methods, such as Euclidean Distance, to determine the closeness between the test data and the training data. The closeness value is used to determine the majority class from a number of nearest neighbors (K value). The classification results are then evaluated to determine the level of accuracy of the system in predicting customer satisfaction accurately (Tamami et al., 2024).

2.6. Utilization of the K-Nearest Neighbor Algorithm

K-Nearest Neighbor (K-NN) is a classification method that determines the category of new data based on its proximity to the nearest training data, using a distance measure such as Euclidean distance, without building a prior model (lazy learning). The sequence of steps in the process of finding missing values with KNN is as follows:

- a. Determine K
Determine amount centroids (K) in a way random, Where k is amount The closest observation will be used. There is no specific method for determining the k value in the K-NN method. If the k value is too small, there will be a lot of noise which will reduce the level of accuracy in the classification, but if it is too large, it will also can cause errors in limiting the values taken and indirectly affect accuracy.
- b. Calculate the Euclidian distance between examples of data with missing values and complete data
Calculate the distance between the Euclidean that has a missing value in the jth observation and other observations that do not have a missing value in that variable. according to the Euclidean distance calculation in formula 1.

$$d_{(x,y)} = \sqrt{\sum_{j=1}^s (x_j - y_j)^2} \dots\dots\dots(1)$$

Where:

$d_{(x,y)}$ = Distance Euclid

j = data testing

s = number of attributes

x_{aj} = mark from attribute jth Which contain data Which is lost

y_{bj} = mark from attribute jth other Which contains complete data

- c. Look for k based on distance Euclid minimum
Mark j on k observation shortest will used on process imputation For observations that have missing values.
- d. Calculating the value of k
The closest observation will get the highest score.
- e. Calculate the average value at k
Calculates the mean value of the k shortest observations that do not have missing values.

$$x_j = \frac{\sum_{k=1}^k w_k v_k}{\sum_{k=1}^k w_k} \dots \dots \dots (2)$$

Where:

x_j = Weight Mean Estimation

K = number of parameters k Which used

w_k = mark observation neighbor closest K

v_k =value in data complete on attribute Which contain data is lost based on parameter k

f. Do process imputation mark missing value

Missing values are used using the average value obtained in step 5 (Wahyu Sudrajat, 2023).

3. RESULTS AND DISCUSSION

The results and discussion section is the core of the research that presents findings based on data processing and interpretation of the results obtained. In this study, the analysis results were obtained through the application of the K-Nearest Neighbor (KNN) algorithm in assessing the level of customer satisfaction at A2 Collection Sei Silau Timur. The collected data was then processed to produce a classification of customer satisfaction levels based on predetermined parameters. Furthermore, the results were analyzed in depth to determine patterns, trends, and factors that influence the level of customer satisfaction. Through this discussion, it is hoped that it can provide a clearer understanding of the effectiveness of using the KNN algorithm and its contribution in helping businesses improve the quality of service to customers (Wahyu Sudrajat, 2023).

3.1. Data Processing

The data processing in this study was carried out through several systematic stages to produce data ready for analysis. The initial data obtained from customer assessments was first selected and checked for completeness to ensure there were no missing or inconsistent values. Next, each attribute, such as quality, price, usability, and service, was converted into a numeric form on a scale of 1–5 so that it could be calculated using the KNN method. After that, the data was normalized to equalize the range of values so that no attribute dominated the distance calculation. The processed data was then used to calculate the distance between data using Euclidean Distance. The result of this process was a proximity value that was used to determine the classification of customer satisfaction based on the majority of nearest neighbors. This process ensured more accurate and objective analysis results.

3.2. Utilization of the K-Nearest Neighbor Algorithm

The use of the K-Nearest Neighbor (KNN) algorithm in the study "Application of the KNN Algorithm to Assess Customer Satisfaction at A2 Collection Sei Silau Timur" is used as a classification method to determine the level of customer satisfaction with the product. This algorithm works by comparing new data with previous data that already has a satisfaction label. Each product data is represented in the form of numeric attributes, namely quality, price, comfort of use, and service. Next, the distance between the test data and all training data is calculated using methods such as Euclidean Distance. After the distance is obtained, a number of K nearest data are selected as a reference. The satisfaction class is determined based on the majority of the nearest neighbors. This approach allows the system to provide objective predictions based on data proximity patterns. And here is a description of its use:

- a. The first step in implementing the K-Nearest Neighbor (K-NN) algorithm is to determine the value of the parameter K (the number of nearest neighbors) and prepare the training data set. At this stage, it is determined how many nearest neighbors will be involved in the calculation to predict the class or label of the new assessment data. The K value generally uses an odd number (for example, $K = 3$ or $K = 5$) to avoid a tie when determining the majority class (majority voting). The data used as a knowledge base contains attributes of Product Quality, Price, Comfort of Use, and Store Service with a scale range of 1 to 5, along with the target class, namely "Satisfied" or "Dissatisfied".
- b. The second step is to determine the Test Data (New Data). The second step is to prepare assessment data from new customers (test data) whose satisfaction labels are not yet known. This data is entered by the cashier after the customer has completed their purchase. As an example, calculation, there is one new customer with the following assessment details:

Table 1. Test Data (New Customer Assessment)

No	Customer Name	Quality	Price	Comfort	Service	Satisfaction
1	Budi Santoso	4	3	4	3	?
2	Siti Aminah	5	5	4	5	?
3	Joko Pranoto	2	3	2	2	?
4	Rina Marlina	4	4	5	4	?
5	Ahmad Fauzi	3	2	3	3	?
6	Princess Diana	5	4	5	5	?
7	Hendra Gunawan	3	4	3	3	?
8	Maya Sari	4	5	4	4	?
9	Rizky Pratama	2	2	3	2	?
10	Ayu Lestari	5	5	5	4	?

- c. The third step is calculating the distance. The system then calculates the distance between the test data (Budi Santoso) and all training data in the database using the Euclidean Distance formula. Here's an example of the distance calculation for the first five training data sets:
- 1). d_1 (Alfino) = $\sqrt{((4-5)^2 + (3-4)^2 + (4-5)^2 + (3-4)^2)} = \sqrt{(1+1+1+1)} = \sqrt{4} = 2,000$
 - 2). d_2 (Aditia) = $\sqrt{((4-3)^2 + (3-3)^2 + (4-2)^2 + (3-3)^2)} = \sqrt{(1+0+4+0)} = \sqrt{5} = 2,236$
 - 3). d_3 (Adjie) = $\sqrt{((4-4)^2 + (3-5)^2 + (4-5)^2 + (3-4)^2)} = \sqrt{(0+4+1+1)} = \sqrt{6} = 2,449$
 - 4). d_4 (Aga) = $\sqrt{((4-2)^2 + (3-3)^2 + (4-2)^2 + (3-2)^2)} = \sqrt{(4+0+4+1)} = \sqrt{9} = 3,000$
 - 5). d_5 (Agung) = $\sqrt{((4-5)^2 + (3-4)^2 + (4-4)^2 + (3-5)^2)} = \sqrt{(1+1+0+4)} = \sqrt{6} = 2,449$
- d. After all distances are calculated, the system sorts the distance values from smallest (closest/most similar) to largest. From this sequence, the top K data points are selected. If K is set to 3, the system will select the three data points with the smallest distances.

Table 2. (K=3)

Ranking	Training Data Name	Smallest Distance	Satisfaction Label
1	Alfino Siregar	2,000	Satisfied
2	Aditia Saputra	2,236	Not satisfied
3	Adjie al fikri	2,449	Satisfied

- e. Step 5: Determining Classification Results (Majority Voting). The final step is to determine the predicted results for the test data (new customers) based on the majority of labels from the K nearest neighbors that were sorted in the previous step. Based on Table 3, from the 3 nearest neighbors, the comparison of the number of labels is as follows:

Table 3. Results of Classification Determination Based on Majority Voting

New Customer Name	K value	Number of "Satisfied" Labels	Number of "Dissatisfied" Labels	Final Classification Results
Budi Santoso	3	2	1	Satisfied

3.3. System Implementation

The system implementation section explains the stages of applying the design results into a usable system. In this study, the implementation was carried out by building a system capable of applying the K-Nearest Neighbor (KNN) algorithm to assess customer satisfaction levels at A2 Collection Sei Silau Timur. This system is designed to process collected customer data, then process it through the KNN calculation stages to produce a satisfaction level classification. In addition, this stage also explains the system interface, workflow, and how users can utilize the system to support the decision-making process.

1. Login Page View

To access the system, a valid username and password combination is required on the login page to access the main page. Users with access rights to the Customer Satisfaction Classification System include Owners, Admins, and Cashiers.



Figure 1. Login Page Display

2. Classification Results Report Page View

This classification results report page displays a summary of customer assessment data that has been processed using the K-Nearest Neighbor (K-NN) algorithm. On this page, users can view detailed classification history information, including the date, customer name, assessed product name, score details for each attribute (Quality, Price, Comfort, and Service), and the final predicted satisfaction label. Furthermore, this page also features a special function for printing the report document.

No	Tanggal	Nama	Produk	Detail	Status
1	27/03/24	Arini	Essen's Good	2.354	Pu
2	28/03/24	-	Essen's Good	2.325	Dissat

Figure 2. Classification Results Report Page Display

4. CONCLUSION

Based on the research results, it can be concluded that the application of the K-Nearest Neighbor (KNN) algorithm in the customer satisfaction classification system at A2 Collection Sei Silau Timur is able to provide an effective solution in managing and analyzing customer evaluation data. This website-based system has successfully transformed the assessment process that was previously carried out manually into a more structured, systematic, and easily accessible one. By utilizing assessment data covering aspects of quality, price, comfort of use, and service, the system can process this information into a basis for more objective decision-making. This shows that digitization through the KNN method can increase efficiency in data management and minimize recording errors that often occur in manual processes. In addition, the use of Euclidean distance calculations and the majority voting method in the KNN algorithm has proven to be able to classify customer satisfaction levels well. Based on the results of the system calculations, it can be seen that the resulting distance values, such as 2.354 which is categorized as "Satisfied" and 2.325 which is categorized as "Dissatisfied", indicate that the proximity of attribute values significantly influences the classification results. Although the difference in distance values is relatively small, the system is still able to determine the class based on the dominance of the nearest neighbor data.

REFERENCES

- Amalia, N., & Nurdin, N. (2024). Performance Analysis of the K-Nearest Neighbor (K-NN) Method in Classifying Hajj Pilgrims Service Satisfaction. *Jurnal Media Informatika*, 6(1), 337-343.
- Banjarnahor, J., Siregar, R., Lumbantobing, C. F., Ridho, M. A., & Zuhdi, M. F. A. (2023). K-Nearest Neighbor (Knn) Analysis for Clothing Sales Classification Based on Materials Used. *Jurnal Sistem Informasi Dan Ilmu Komputer Prima (JUSIKOM PRIMA)*, 7(1), 42-49. <https://doi.org/10.34012/jurnalsisteminformasidanilmukomputer.v7i1.3903>
- Dhewayani, F. N., Amelia, D., Alifah, D. N., Sari, B. N., & Jajuli, M. (2022). Implementasi K-Means Clustering untuk Pengelompokan Daerah Rawan Bencana Kebakaran Menggunakan Model CRISP-DM. *Jurnal Teknologi Dan Informasi*, 12(1), 64-77. <https://doi.org/10.34010/jati.v12i1.6674>

- Diansyah, S. (2022). Klasifikasi Tingkat Kepuasan Pengguna dengan Menggunakan Metode K-Nearest Neighbour (KNN). *Jurnal Sistim Informasi Dan Teknologi*, 4, 7–12. <https://doi.org/10.37034/jsisfotek.v4i1.114>
- Fahmi, M. F., Helilantar, R., & Farida, I. N. (2025). Implementasi Metode K-Nearest Neighbor untuk Klasifikasi Tingkat Kepuasan Pelanggan pada Toko Buku Anugerah Ilmu. (2025). *Prosiding SEMNAS INOTEK (Seminar Nasional Inovasi Teknologi)*, 9(2), 1471-1480. <https://doi.org/10.29407/x4mg2w61>
- Fansyuri, M. (2020). Analisa Algoritma Klasifikasi K-Nearest Neighbor Dalam (Study Kasus Pt. Trigatra Komunikatama) Analysis of K-Nearest Neighbor Classification Algorithm in Determining Accuracy Value on Customer Pleasure (Case Study on Pt. Trigatra Komunikatama). *Humanika: Jurnal Ilmu Sosial, Pendidikan, dan Humaniora*, 3(1), 29-33.
- Heti Aprilianti, Khothibul Umam, & Maya Rini Handayani. (2025). Comparative Study of SVM, KNN, and Naïve Bayes for Sentiment Analysis of Religious Application Reviews. *Journal of Applied Informatics and Computing*, 9(3), 920–927. <https://doi.org/10.30871/jaic.v9i3.9482>
- Indraputra, R. A., & Fitriana, R. (2020). K-Means clustering data COVID-19. *Jurnal Teknik Industri*, 10(3), 275-282.
- Iranti, A. M., & Nuryana, I. K. D. (2024). Implementasi Algoritma K-Nearest Neighbors dalam Mengetahui Kepuasan Pengguna Aplikasi CapCut. *Journal of Emerging Information Systems and Business Intelligence (JEISBI)*, 5(3), 266–277. <https://doi.org/10.26740/jeisbi.v5i3.63772>
- Lillah, M. R. R. L., Maylawati, D. S. ., Zulfikar, W. B. ., Uriawan, W. ., & Wahana, A. . (2023). Implementasi Algoritma K-Nearest Neighbor (KNN) untuk Analisis Sentimen Pengguna Aplikasi Tokopedia . *Intellect : Indonesian Journal of Learning and Technological Innovation*, 2(2), 171–184.
- Murtiwiayati, M., Arini, R. F. ., Safitri, L. ., Widiyanti, S. ., & Setiadi, I. . (2025). Implementasi Algoritma K-Nearest Neighbors untuk Klasifikasi Sentimen pada Ulasan Pakaian Anak di Platform TikTok Shop. *Jurnal Minfo Polgan*, 14(2), 2921-2928. <https://doi.org/10.33395/jmp.v14i2.15538>
- Nasrullah, A. H. (2021). Implementasi algoritma Decision Tree untuk klasifikasi produk laris. *Jurnal Ilmiah Ilmu Komputer Fakultas Ilmu Komputer Universitas Al Asyariah Mandar*, 7(2), 45-51. <https://doi.org/10.35329/jiik.v7i2.203>
- Pratama, A. R. (2023). Futsal Berbasis Web Menggunakan Metode Rapid Application Development (RAD). *VI(03)*, 63–69.
- Rahmadinia, Enjel Erika LorencisLubis, Aji Priansyah, Yolanda R.W.N., T. M. (2023). PENERAPAN DATA MINING UNTUK MEMREDIKSI HARGA BAHAN PANGAN DI INDONESIA MENGGUNAKAN ALGORITMA K-NEAREST NEIGHBOR. *Jurnal Mahasiswa Akuntansi Samudra*, 4(4), 223–235. <https://doi.org/10.33059/jmas.v4i4.7074>
- Ratih Yulia Hayuningtyas, R. S. (2022). Implementasi Data Mining Dengan Algoritma Multiple Linear Regression Untuk Memprediksi Penyakit Diabetes. *Jurnal Teknik Komputer AMIK BSI*, 8(1), 40–44. <https://doi.org/10.31294/jtk.v4i2>
- Rifa'i, H., Ryan Hamonangan, Dian Ade Kurnia, Kaslani, & Mulyawan. (2022). Implementasi Algoritma Decision Tree Dalam Klasifikasi Kompetensi Siswa. *KOPERTIP: Jurnal Ilmiah Manajemen Informatika Dan Komputer*, 6(1), 15–20. <https://doi.org/10.32485/kopertip.v6i1.131>
- Soleh, M., Hidayati, N., & Krisdian, F. D. M. (2021). PENERAPAN METODE ASSOCIATION RULE - MARKET BASKET ANALYSIS UNTUK MENINGKATKAN DAYA SAING TOKO SWALAYAN KECIL. *Jurnal Aplikasi Ilmu Teknik Industri (JAPTI)*, 2(1), 1–9. <https://doi.org/10.32585/japti.v2i1.1469>
- Sudrajat, W., & Cholid, I. (2023). K-NEAREST NEIGHBOR (K-NN) UNTUK PENANGANAN MISSING VALUE PADA DATA UMKM. *Jurnal ReKayasa Sistem Informasi Dan Teknologi*, 1(2), 54–63. <https://doi.org/10.59407/jrsit.v1i2.77>

- Suprianto, A., Surapati, U., Akbar, Y., & Hidayat, A. Z. (2024). Classification of Customer Satisfaction with the K-Nearest Neighbor Algorithm in Relation to Employee Performance at PT. Airkon Pratama. *International Journal Software Engineering and Computer Science (IJSECS)*, 4(3), 882–895. <https://doi.org/10.35870/ijsecs.v4i3.2948>
- Tamami, R., Yanti, D., & Anugrah, R. (2024). Analisis Kepuasan Pelanggan Menggunakan Algoritma C4.5 Pada Aplikasi Betukang.Id. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(6), 12507–12511. <https://doi.org/10.36040/jati.v8i6.11937>