Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Number 3, July 2025
DOI : https://doi.org/10.33395/sinkron.v9i3.15110
e-ISSN : 2541-2019
p-ISSN : 2541-044X

# Customer Loyalty Classification Using KNN and Decision Tree for Sales Strategy Development

**[1)]Mukhlisin, [2)]Handoyo Widi Nugroho**
[1,2)]Magister Teknik Informatika IIB Darmajaya Lampung
Email: [1] muhklisin.2421210014p@mail.darmajaya.ac.id, [2] handoyo.wn@darmajaya.ac.id

**Submitted** :Jul 27, 2025 | **Accepted** : Aug 2, 2025 | **Published** : Aug 3, 2025

**Abstract:** Customer loyalty is a crucial element in maintaining business continuity in today's competitive digital era. This study aims to classify customer loyalty levels based on sales and transaction behavior data using two supervised machine learning algorithms: *K-Nearest Neighbor* (KNN) and *Decision Tree*. The models were developed and evaluated using Python in the Google Colaboratory environment, utilizing a dataset of 250 customer records. The research process included data preprocessing, feature selection, normalization, data splitting, model building, and evaluation using accuracy, precision, recall, and F1-score metrics. Evaluation results showed that the Decision Tree algorithm delivered the best performance with 99.20% accuracy, 99.50% precision, 99.50% recall, and a 99.50% F1-score. Meanwhile, the KNN algorithm achieved 91.60% accuracy, 91.63% precision, 98.50% recall, and a 94.91% F1-score. These findings indicate that the Decision Tree model is more effective for classifying customer loyalty and can be implemented as a decision support tool for data-driven Customer Relationship Management (CRM) strategies.

**Keywords:** Customer Loyalty; Classification; Decision Tree; K-Nearest Neighbor; Python; Data Mining

## INTRODUCTION

In the increasingly competitive digital business landscape, customer loyalty has become a vital factor in maintaining long-term business growth. Loyal customers not only make repeat purchases but also act as brand advocates, helping companies reduce marketing costs and strengthen market presence. Consequently, identifying and managing customer loyalty effectively is a key aspect of Customer Relationship Management (CRM) strategies(Utomo et al., 2025)(Takalapeta, 2018)

However, many businesses still rely on manual methods or managerial intuition to identify loyal customers. This subjective approach often lacks consistency and may overlook valuable behavioral patterns hidden in transaction data. At the same time, companies are generating large volumes of customer data that remain underutilized due to the absence of systematic analysis methods.(Nosiel et al., 2021)(Wahyudi et al. 2022).

Recent developments in machine learning (ML) offer promising solutions for automating customer loyalty classification. Among the various classification algorithms, K-Nearest Neighbor (KNN) and Decision Tree stand out due to their simplicity, interpretability, and effectiveness in handling small to medium-sized datasets. These algorithms allow businesses to analyze transaction and behavioral data more objectively and generate reliable loyalty predictions(Bounie 2025) (Isyriyah et al., 2024)

While several previous studies have used GUI-based tools like RapidMiner or Orange, such platforms often conceal the full modeling pipeline and limit transparency and reproducibility. This research addresses that gap by implementing a fully code-based approach using Python and Scikit-learn. The novelty of this study lies in presenting a complete and transparent modeling process—from data preprocessing to final evaluation—that can be directly integrated into CRM systems and business operations.(Wijaya & Girsang, 2015)(Naldy & Andri, 2021).

The objective of this study is to build and compare the performance of KNN and Decision Tree algorithms in classifying customer loyalty levels using behavioral and transaction data. The findings are expected to support businesses in creating more targeted marketing strategies and making informed decisions through data-driven customer segmentation (Nosiel et al., 2021)

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Number 3, July 2025
DOI : https://doi.org/10.33395/sinkron.v9i3.15110
e-ISSN : 2541-2019
p-ISSN : 2541-044X

## LITERATURE REVIEW

The use of classification algorithms such as Decision Tree and K-Nearest Neighbor (KNN) has been widely applied in local studies related to customer loyalty(NOVIA RAHMADANA1, ABDUL RAHIM*2, 2024) implemented the C4.5 algorithm to predict loyalty based on customer attributes such as age, payment method, and transaction frequency, with classification results indicating strong model reliability. (Nurzahputra et al., 2016) applied C4.5 to predict customer loyalty at PT. SNS Bekasi and reported an accuracy rate of 94.00%. While studies using KNN for loyalty classification are limited, (Sreevalsan-Nair, 2020)) and a more recent study (Nugroho et al., 2024) show that KNN can be applied to analyze customer satisfaction or loyalty potential, with accuracy ranging from 85% to 87%.

In the context of machine learning implementation using code-based tools, Python is gaining attention in local academic research, although its adoption remains limited compared to GUI-based platforms. (Tritularsih & Prasetyo, 2025) demonstrated the use of Python, particularly the Pandas and Seaborn libraries, for exploratory analysis of customer behavior and the identification of attributes contributing to loyalty patterns. Their study highlights Python's flexibility and transparency compared to GUI tools like RapidMiner, while also emphasizing its advantages in reproducibility and integration with business intelligence systems.

Regarding model performance evaluation, most local studies still focus primarily on accuracy and precision. For example, a study at PT. Kopi Kenangan using the C4.5 algorithm reported an accuracy of 86.96% and a precision of 90% in classifying customer loyalty (Fihir et al., 2010)yet omitted recall and F1-score from the evaluation. Similarly, (Ardani et al., 2022) compared C4.5, Naive Bayes, and SVM for classifying customer satisfaction at Telkomsel and found that C4.5 achieved the highest accuracy (96.50%), but they also failed to report more comprehensive metrics. This highlights a gap in local studies where recall and F1-score—especially important for imbalanced datasets—are often overlooked, despite their value in ensuring more reliable model interpretation.

Although other algorithms such as Random Forest and Support Vector Machine (SVM) are known to deliver strong predictive performance, they were not selected in this study due to several practical considerations. Random Forest, while accurate, tends to generate complex ensemble models that are harder to interpret—an important limitation when results need to be communicated to business stakeholders. Similarly, SVM is computationally intensive and often less intuitive in how it classifies data, especially with small to medium-sized datasets like the one used in this research. By contrast, Decision Tree and KNN offer simpler, more interpretable models that can be readily visualized and explained, making them better suited for business applications that require transparency and direct implementation

## METHOD

Figure 1 shows the research flow that includes the collection of sales and transaction performance data, pre-processing of data (such as normalization, grouping, and removal of irrelevant data), extraction of critical features, division of data into training data and test data, and application of K-Nearest Neighbor (KNN) and Decision Tree algorithms for classification. The next step is to evaluate the model to measure performance and effectiveness in developing a career development strategy for sales(Artana et al., 2025)(Gunia et al., 2024)

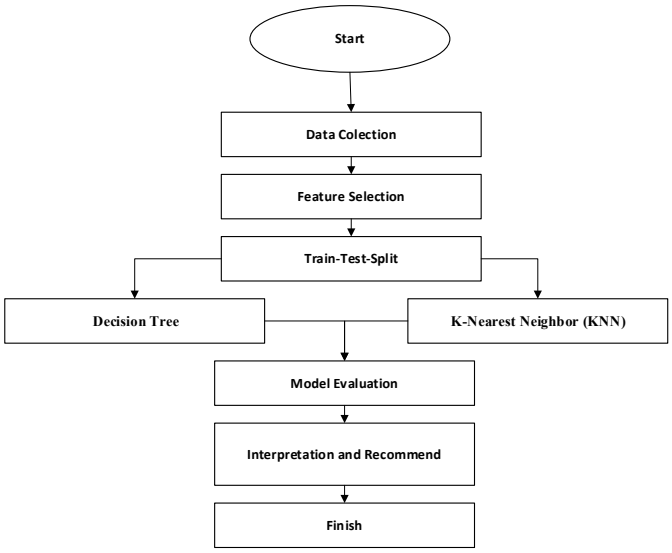


Fig. 1 Research Flow

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Number 3, July 2025
DOI : https://doi.org/10.33395/sinkron.v9i3.15110
e-ISSN : 2541-2019
p-ISSN : 2541-044X

The process begins with Start, where the objectives of the analysis are defined and the working environment is prepared to support a data-driven approach. In the Data Collection phase, relevant customer data such as transaction records, purchase frequency, total spending, and other behavioral indicators are gathered, as these factors may influence loyalty levels. Next, during Feature Selection, important variables are selected through statistical techniques like Pearson correlation to ensure that only relevant features are included, thereby reducing noise and improving the overall performance of the model.

After that, the data is split into training and testing subsets in the Train-Test Split stage, typically using an 80:20 ratio. This ensures that the model can be trained on one portion of the data and evaluated on unseen data to test its generalizability. In the Model Building phase, two classification algorithms—Decision Tree and K-Nearest Neighbor (KNN)—are applied in parallel. The Decision Tree builds a rule-based model by recursively splitting the data based on feature values, while KNN classifies each data point based on the majority class among its *k* closest neighbors in the feature space.

Once the models are constructed, Model Evaluation is carried out using standard performance metrics: accuracy, precision, recall, and F1-score. These metrics provide a comprehensive understanding of how well each algorithm performs in classifying customer loyalty levels. In the Interpretation and Recommendation step, the evaluation results are analyzed to gain insights into customer behavior. Based on these insights, recommendations are made for customer segmentation and strategy development focused on enhancing customer loyalty. Finally, the process ends with the Finish stage, where results are documented, business implications are discussed, and potential future steps such as model deployment or refinement with new data are considered.

This research began with the collection of customer transaction data which included attributes such as purchase frequency, total transaction value, subscription duration, and loyalty labels that have been determined by the management. The data used consists of 250 customer data that has been anonymized to maintain identity confidentiality.

In the pre-processing stage, data cleanup is performed to remove duplication and handle empty values. Feature normalization is done using *the StandardScaler* from the *Scikit-learn* library to avoid bias due to scale differences. Loyalty labels are converted to numeric formats (e.g.: 0 = Disloyal, 1 = Fairly Loyal, 2 = Loyal).

Feature selection is carried out using the Pearson correlation method to determine the strength of the relationship between the feature and the target variable. Features with a correlation below 0.1 are eliminated. Furthermore, visual exploration and initial testing were carried out to validate the importance of the features used.

The dataset is divided into two parts, namely 80% training data and 20% test data, with a random partitioning process using *random_state = 42* to ensure consistent replication of results.

The classification model was built using two algorithms, namely *K-Nearest Neighbor* (KNN) and *Decision Tree*. The optimal parameter value for KNN is determined through *a cross-validation* process, and the best k value is obtained, which is k = 5. For Decision Tree, *a Gini* separation criterion with *maximum depth settings* is used to prevent *overfitting*. All modeling processes are performed in a Python environment using the *Scikit-learn library*.

The dataset was initially divided into training and testing sets using an 80:20 ratio through the train_test_split() function from Scikit-learn. This ensures that the model is trained on one portion of the data and validated on a separate subset to assess its generalization ability. A fixed random_state = 42 was used to maintain reproducibility of results.

In addition to this static split, K-Fold Cross-Validation was performed—using K = 5—to further evaluate the robustness of the models. This technique partitions the training data into 5 subsets (folds), where each fold is used once as validation while the remaining folds are used for training. The results from all folds are averaged to produce a more reliable performance estimate. This cross-validation step helps detect and reduce the risk of overfitting, ensuring that the model performance is not overly dependent on one specific data partition.

The evaluation of model performance was carried out with four main metrics, namely accuracy, precision, recall, and F1-score. The results of the classification obtained are then analyzed to produce recommendations for data-driven sales career development strategies.

## RESULT

This research was conducted using Google Collaboratory which utilizes a CPU and RAM computing capacity of 12.7 GB. The optimal parameters for the K-Nearest Neighbor (KNN) algorithm, such as the number of closest neighbors (k), as well as the cumulative value of variance for PCA and 2DPCA, are determined through a grid search process using cross-validation techniques. The grid search results show the best combination of parameters that result in the highest accuracy in validation data and are computationally efficient. These parameters are then used to build a final classification model. In addition, the Decision Tree algorithm is also optimized with similar approaches, such as depth value adjustment and splitting criteria, to produce accurate and reliable models in supporting the development of sales career development strategies

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Number 3, July 2025
DOI : https://doi.org/10.33395/sinkron.v9i3.15110
e-ISSN : 2541-2019
p-ISSN : 2541-044X

**Value k = 5**

The following Figure 2 presents a visualization of the results of the performance evaluation of the classification model using the K-Nearest Neighbor (KNN) algorithm with the parameter of k = 5. This graph illustrates four key evaluation metrics, namely accuracy, precision, **recall**, and **F1-score**, which are used to assess how well the model is at grouping sales data and transaction performance as a basis for developing sales career development strategies.

The confusion matrix below illustrates the classification results of the KNN algorithm, where true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) can be visualized for each class of customer loyalty. This helps assess how well the model distinguishes between loyal, fairly loyal, and disloyal customers.

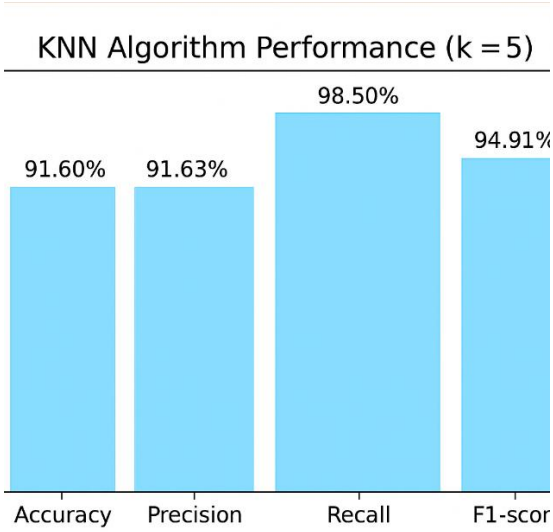


Fig. 2 **k = 5**

**K-Nearest Neighbor (KNN) Algorithm Performance Analysis (k = 5)**

The figure above shows the performance evaluation results of the K-Nearest Neighbor (KNN) algorithm with the parameter of the number of closest neighbors (**k = 5**) in classifying sales data and transaction performance to support sales career development strategies.

From the graph it can be seen that:

1. The accuracy of the KNN model reached 91.60%, which indicates that the model is able to classify data with a fairly high level of accuracy.
2. A precision of 91.63% indicates that of all positive predictions made by the model, more than 91% are correct predictions.
3. Recall has the highest value of 98.50**%,** which means that the model is very good at finding actual positive data, or in other words has a high *level of sensitivity*.
4. The F1-score, which is the harmonic average of precision and recall, is recorded at 94.91%, indicating a good balance between precision and recall.

These results show that the KNN algorithm is quite reliable in classifying data with good accuracy and completeness, although there is still room for improvement in precision. The high recall advantage suggests that this model tends to be better at avoiding the mistake of ignoring important data, which is crucial in the context of sales career development decision-making

**Decision Tree**

After evaluating the Decision Tree algorithm, results were obtained that showed excellent performance in the classification of sales data and transaction performance. To visually illustrate the results of the evaluation, the following graph is displayed that presents the accuracy, precision, recall, and F1-score values of the Decision Tree model:

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Number 3, July 2025
DOI : https://doi.org/10.33395/sinkron.v9i3.15110
e-ISSN : 2541-2019
p-ISSN : 2541-044X

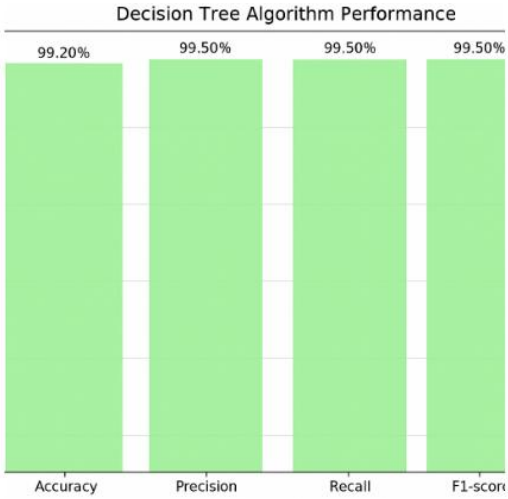


Fig. 3 Value of the Decesion Tree

Figure 3 shows the results of the evaluation of the performance of the Decision Tree algorithm based on four main metrics, namely accuracy, precision, recall, and F1-score. From the graph, it can be seen that the model has excellent performance, with all evaluation metrics being above 99%.

1. The accuracy of 99.20% indicates that the model is able to classify the data very accurately overall.
2. The accuracy and recall reached 99.50%, respectively, which indicates that the model is not only accurate in predicting positive classes, but also very minimal in making type I and type II errors.
3. The **F1-score** value of 99.50% strengthens the model's consistency in handling imbalances between classes, combining precision and recall in one harmonious measure.

Overall, the high performance of Decision Tree's algorithm on sales data and transaction performance shows that this model is very effective to use in developing sales career development strategies based on available historical data.

**Comparison of the two algorithms**

To provide a clearer visual picture of the performance of the two algorithms used in this study, namely K-Nearest Neighbor (KNN) with a value *of k = 5* and Decision Tree, the following is a comparison graph based on four main evaluation metrics: Accuracy, Precision, Recall, and F1-score. This graph aims to show the extent of the difference in the level of accuracy and effectiveness of classification from each algorithm in developing a sales career development strategy based on sales data and transaction performance

Table 1, Decision Tree consistently achieves near-perfect results across all evaluation metrics

| Algorithm | Accuracy | Precision | Recall | F1-Score |
|---|---|---|---|---|
| KNN (k=5) | 91.60% | 91.63% | 98.50% | 94.91% |
| Decision Tree | 99.20% | 99.50% | 99.50% | 99.50% |

Figure 4 presents a comparative bar chart illustrating the performance of the K-Nearest Neighbor (KNN) and Decision Tree algorithms based on four evaluation metrics: Accuracy, Precision, Recall, and F1-score. Each bar represents the percentage value achieved by each algorithm for the respective metric.

As shown in the figure, the Decision Tree algorithm consistently outperforms KNN across all metrics. Decision Tree achieved 99.20% in accuracy, 99.50% in both precision and recall, and a 99.50% F1-score, indicating a balanced and highly accurate classification model. Meanwhile, KNN recorded lower but still respectable scores, with 91.60% accuracy, 91.63% precision, 98.50% recall, and a 94.91% F1-score.

These visual results highlight the superiority of Decision Tree in classifying customer loyalty, especially in maintaining both precision and recall, which are critical in minimizing false predictions and ensuring customer targeting is accurate and effective

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Number 3, July 2025
DOI : https://doi.org/10.33395/sinkron.v9i3.15110
e-ISSN : 2541-2019
p-ISSN : 2541-044X

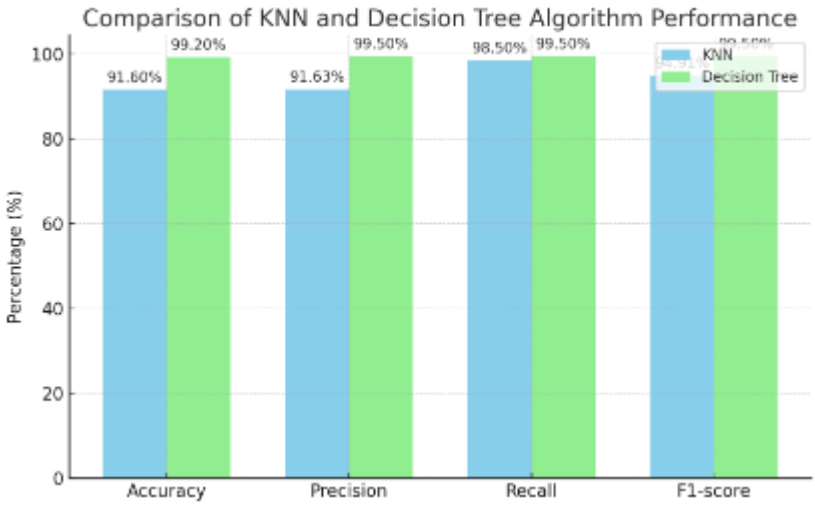


Fig. 4 of the second Algorithm

The graph above shows a comparison of the performance between the KNN and Decision Tree algorithms on sales data and transaction performance in developing a sales career development strategy.

In terms of **accuracy**, the Decision Tree algorithm shows much higher performance than KNN. Decision Tree managed to achieve an accuracy level of 99.20%, while KNN only reached 91.60%. This difference reflects Decision Tree's ability to classify data more accurately and consistently, providing more reliable predictive results.

In terms of **precision**, Decision Tree also excels with a score of 99.50%, compared to KNN which recorded a precision of 91.63%. The high precision of Decision Tree suggests that this model is more effective in reducing false positive misclassification. This means that most of the customers classified as loyal by this model are indeed truly loyal, making them useful for precise decision-making in marketing strategies or sales career development.

In the **recall metric**, both algorithms showed high values, namely 98.50% for KNN and 99.50% for Decision Tree. Although the difference is small, these results confirm that Decision Tree is more efficient at recognizing all the data that actually falls into the loyal category, so the risk of losing important customers can be minimized.

Finally, in the **F1-score** metric, which is a harmonious combination of precision and recall, KNN recorded a score of 94.91%, while Decision Tree achieved 99.50%. The high F1-score on Decision Tree shows that the model is not only accurate in classification, but also balanced in recognizing loyal customers thoroughly and precisely. Thus, Decision Tree proves to be the most stable and optimal algorithm in the context of customer loyalty classification based on sales data and transaction performance

## DISCUSSIONS

The results of this study show that the Decision Tree algorithm consistently outperforms KNN in classifying customer loyalty based on sales and transaction data. While KNN demonstrates strong recall (98.50%), its slightly lower precision (91.63%) indicates a tendency toward false positive classifications, which may lead to inefficient targeting in loyalty programs.

In contrast, the Decision Tree model achieved very high and balanced scores across all evaluation metrics, including 99.50% in precision, recall, and F1-score. These results suggest that the model can more accurately identify truly loyal customers, which is crucial for designing effective Customer Relationship Management (CRM) and personalized marketing strategies. Businesses can use these classification outputs to segment customers, tailor promotional campaigns, and allocate resources more efficiently—ultimately improving retention and lifetime customer value.

To ensure model robustness and reduce overfitting, K-Fold Cross-Validation was applied during the training phase. This method provided more reliable performance estimates compared to a single static data split. However, despite the promising results, the model was still evaluated on the same dataset used for training and validation purposes.

Therefore, to truly assess the generalizability of the model, it is essential to test it on an entirely different or external dataset. This would help determine whether the model performs consistently across various customer behavior patterns and data distributions beyond the original sample.

Additionally, this study is limited by the use of a relatively small dataset (250 records), which may not fully represent the complexity of real-world customer behavior. The dataset also consisted of anonymized and pre-labeled loyalty categories, which may not capture the full dynamics of loyalty over time.

*name of corresponding author

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

Future research should consider incorporating time-based features, evaluating model performance on external datasets, and exploring ensemble methods (e.g., Random Forest or Gradient Boosting) to reduce overfitting while maintaining interpretability.

## CONCLUSION

Based on the evaluation results, the Decision Tree algorithm outperformed the K-Nearest Neighbor (KNN) algorithm in classifying customer loyalty levels based on transaction and sales performance data. Decision Tree achieved superior metrics with 99.20% accuracy, 99.50% precision, 99.50% recall, and a 99.50% F1-score. Meanwhile, KNN with k = 5 yielded 91.60% accuracy, 91.63% precision, 98.50% recall, and a 94.91% F1-score. These findings indicate that Decision Tree provides more accurate and consistent performance, making it a more suitable choice for supporting CRM and sales career development strategies in data-driven environments

## REFERENCES

Ardani, R. A. T., Jupriyadi, Styawati, Saputra, A. W., & Basroni, A. (2022). Implementasi Data Mining Menggunakan Algoritma Apriori untuk Memprediksi Merk Parfum yang Terjual (Studi Kasus: Queen Parfum). *Jurnal Ilmiah Infrastruktur Teknologi Informasi (JIITI)*, *3*(1), 9–15. http://jim.teknokrat.ac.id/index.php/teknologiinformasi/article/view/2324%0Ahttp://jim.teknokrat.ac.id/index.php/teknologiinformasi/article/download/2324/769

Artana, I. P. Y., Dwi, I. M., Asana, P., Nyoman, N., Sastaparamitha, A. J., & Jaya, K. (2025). *Combining U-NET Segmentation and Dimensionality Reduction Methods for K-NN Fish Freshness Classification*. *8*(1), 81–94.

Bounie, D. (2025). *Pengaruh Ulasan Pelanggan Online terhadap Keputusan Pembelian : Kasus Pengaruh Ulasan Pelanggan Online terhadap Pembelian Keputusan : Kasus Video Game*. *1*.

Fihir, M., Martanto, & Hayati, U. (2010). Menggunakan Metode Decision Tree Pada. *(Jurnal Mahasiswa Teknik Informatika*, *7*(6), 3830–3833.

Gunia, E., Irma Purnamasari, A., & Ali, I. (2024). Penerapan Datamining Dalam Menentukan Pola Penjualan Produk Menggunakan Algoritma Fp-Growth. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *8*(2), 2417–2422. https://doi.org/10.36040/jati.v8i2.9506

Isyriyah, L., Baihaqi, I., & Purwiantono, F. E. (2024). Prediksi Loyalitas Pelanggan Pada Fast Moving Consumer Goods Menggunakan Klasifikasi Metode C4.5. *Smatika Jurnal*, *13*(02), 369–380. https://doi.org/10.32664/smatika.v13i02.1115

Naldy, E. T., & Andri, A. (2021). Penerapan Data Mining Untuk Analisis Daftar Pembelian Konsumen Dengan Menggunakan Algoritma Apriori Pada Transaksi Penjualan Toko Bangunan MDN. *Jurnal Nasional Ilmu Komputer*, *2*(2), 89–101. https://doi.org/10.47747/jurnalnik.v2i2.525

Nosiel, N., Sriyanto, S., & Maylani, F. (2021). Perbandingan Teknik Data Mining Untuk Prediksi Penjualan Pada UMKM Gerabah. *Prosiding Seminar Nasional Darmajaya*, *1*, 72–86.

NOVIA RAHMADANA1, ABDUL RAHIM*2, F. Y. (2024). *Analisis Kepuasan Pelanggan Menggunakan Algoritma K-Nearest Neighbors Pada*. *9*, 183–192.

Nugroho, R., Setiawan, I., Akmal, R. N., & Azka, N. (2024). *Evaluasi Keamanan Sistem Informasi Pada SMKN 1 Banyumas Berdasarkan Indeks Keamanan Informasi ( KAMI ) ISO 27001 : 2013*. *6*.

Nurzahputra, A., Ratna Safitri, A., & Aziz Muslim, M. (2016). Klasifikasi Pelanggan pada Customer Churn Prediction Menggunakan Decision Tree. *PRISMA : Prosiding Seminar Nasional Matematika*, 717–722. https://journal.unnes.ac.id/sju/prisma/article/view/21528

Sreevalsan-Nair, J. (2020). K-nearest Neighbors. *Encyclopedia of Earth Sciences Series*, *2020*(2), 100–103. https://doi.org/10.1007/978-3-030-26050-7_170-1

Takalapeta, S. (2018). Penerapan Data Mining Untuk Menganalisis Kepuasan Konsumen Menggunakan Metode Algoritma C4.5. *J I M P - Jurnal Informatika Merdeka Pasuruan*, *3*(3), 34–38. https://doi.org/10.37438/jimp.v3i3.186

Tritularsih, Y., & Prasetyo, H. (2025). Penerapan Machine Learning untuk Pencarian Pelanggan Loyal Berpotensi Menggunakan Metode Python Pandas Seaborn. *Integrasi: Jurnal Ilmiah Teknik Industri*, *10*(1), 12–24. https://doi.org/10.32502/integrasi.v10i1.292

Utomo, D. N. S. S., Bhakti, H. D., & Devi, P. A. R. (2025). Penerapan Algoritma Naïve Bayes Untuk Klasifikasi Penilaian Kinerja Pegawai Di Kedai Xyz. *Kohesi: Jurnal Multidisiplin Saintek*, *7*(1), 61–70. https://ejournal.warunayama.org/index.php/kohesi/article/view/11029

*name of corresponding author

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

Wahyudi, T., Informasi, S., Tinggi, S., Komputer, I., Karya, C., Sawit, D., & Timur, K. J. (2022). Penerapan Data Mining Pada Transaksi Penjualan Pakaian Dengan Menggunakan Algoritma Apriori. *Jupiter*, *14*(2), 473–482.

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

*name of corresponding author