Vol.7 No.1 Tahun 2026
Jurnal Kecerdasan Buatan, Komputasi dan Teknologi Informasi
https://ejournal.unuja.ac.id/index.php/core
E-ISSN : 2774-7875 and P-ISSN : 2775-0124

Comparative Analysis of Decision Tree, Neural Network, K-NN, GBT, SVR and Random Forest Algorithms for House Price Prediction

Sahal Abdillah[1], Mochamad Taufiqurrohman Abdul Aziz Zein[2], Edy Sulistiyanto[3]
[1,2,3] Nahdlatul Ulama Al Ghazali University of Cilacap, Cilacap, Indonesia

Article Info

Article history:

Received : 11 10, 2025
Revised : 12 21, 2025
Accepted : 03 07, 2026

Keywords:

Machine Learning;
House Price Prediction;
Data Mining;
Altair Studio;
K-Nearest Neighbor;

Abstract

This study aims to analyze and compare the performance of several machine learning algorithms for predicting house prices. The dataset was obtained from Kaggle with variables of house price, building area, land area, number of bedrooms, number of bathrooms, and garage. The algorithms used include Decision Tree, Neural Network, K-Nearest Neighbor, Gradient Boosted Trees, Support Vector Regression, and Random Forest. The study was conducted using Altair Studio through pre-processing, training, testing, and evaluation stages based on performance, computation time, and model complexity. The results showed that K-NN obtained the best RMSE value of 0.608, followed by Neural Network 0.630, Decision Tree 0.649, Random Forest 0.656, SVR 0.662, and GBT 0.747. Based on MAE, Neural Network obtained the best value of 0.464, followed by Random Forest 0.474, Decision Tree 0.478, K-NN 0.485, SVR 0.501, and GBT 0.589. Overall, K-NN was the best model because it had the lowest RMSE, the fastest training time of 0.004 seconds, and a relatively simple model complexity.

Corresponding Author:

Sahal Abdillah,
mlsahal45@gmail.com

INTRODUCTION

A house is one of the primary human needs that serves not only as a place to live but also as a high-value economic investment asset. Accurate house price determination is essential because it is still frequently performed manually based on intuition or personal experience, often resulting in inconsistencies and market imbalances (Rahayuningtyas et al., 2021). In the digital era, the application of Machine Learning technology has become an effective alternative for predicting house prices objectively and efficiently. By utilizing historical data, predictive models can be developed to estimate house prices by considering various factors such as location, building area, land area, number of bedrooms, and other supporting facilities (Ridho et al., 2022).

Several previous studies have applied Machine Learning algorithms for property price prediction. Study Fitri (2023) demonstrated that Linear Regression, Random Forest Regression, and Gradient Boosted Trees Regression were able to generate house price predictions with low prediction errors. Based on the evaluation of these three methods using a house price dataset, Random Forest Regression achieved the best performance with an RMSE value of 0.440 and the highest accuracy of 81.5%. This method outperformed Linear Regression (RMSE = 0.515) and Gradient Boosted Trees Regression (RMSE = 0.508). In addition, the dataset size and preprocessing process significantly affected prediction accuracy, indicating that larger datasets and additional preprocessing techniques could further improve model performance (Fitri, 2023).

Meanwhile, study (Hidayah et al. 2024) showed that the combination of Decision Tree, Support Vector Regression, and Random Forest algorithms improved the accuracy of house price prediction models. Based on the RMSE analysis, Random Forest produced the lowest prediction error with an RMSE value of 3.204, making it the best-performing model among the evaluated methods. However, the Random Forest model still exhibited an R^2 difference of 0.113, indicating a tendency to overfit the training data compared to the Linear Regression model, which showed a smaller R^2 difference of 0.020.

Algorithms such as Decision Tree, Neural Network, K-Nearest Neighbor, Gradient Boosted Trees, Support Vector Regression, and Random Forest possess different characteristics and levels of complexity that influence prediction accuracy and computational efficiency (Cucos & Iantovics, 2024; Alarfaj et al., 2022). These algorithms were selected because of their ability to produce highly accurate predictions while maintaining computational efficiency, with each algorithm offering distinct advantages in learning data patterns, handling complex variables, and prediction speed (Karimi et al., 2025; Özlem & Tan, 2022). Several recent studies have emphasized the importance of selecting algorithms that match the characteristics of the dataset to achieve optimal prediction performance (Lee et al., 2021). Therefore, a comprehensive comparative analysis is required to determine the most effective algorithm for improving model performance, particularly in terms of prediction accuracy and efficiency (Sari & Prabowo, 2025).

Despite these studies, no previous research has simultaneously compared six widely used algorithms using the Altair Studio platform for house price prediction analysis. Altair Studio was selected because of its comprehensive capabilities in supporting data analysis by allowing users to visually construct analytical workflows through interconnected operators, including data preprocessing, predictive and descriptive modeling, model evaluation, visualization, and automated learning features, making it highly efficient for Data Mining tasks (Taranto-Vera et al., 2021). This study is expected to provide a more comprehensive understanding of the effectiveness of various Machine Learning algorithms in predicting house prices. Furthermore, the findings are expected to serve as a valuable reference for system developers and property business practitioners in selecting a prediction model that is more accurate, efficient, and easy to implement using Altair Studio.

METHOD

The method used in this study employs a Data Mining approach to analyze and compare the performance of several Machine Learning algorithms for house price prediction. Data Mining is a systematic approach to discovering hidden information from data so that it can be utilized optimally. It not only focuses on information extraction but also plays an important role in developing predictive models capable of estimating future values or behaviors based on the available data (Warjiyono et al., 2024).

This Data Mining approach utilizes various Machine Learning algorithms designed to identify hidden patterns in historical data and generate accurate prediction models (Manullang et al., 2021). The algorithms used in this study include Decision Tree, Neural Network, K-Nearest Neighbor, Gradient Boosted Trees, Support Vector Regression, and Random Forest. Each algorithm has different characteristics, advantages, and levels of complexity. Decision Tree was selected because of its flexibility in providing clear visualizations, making prediction results easier to interpret (Illahi et al., 2023). Neural Network was chosen for its capability to process large and complex datasets as well as its flexibility in capturing non-linear relationships among variables (Ferdiansyah et al., 2024). K-Nearest Neighbor was employed as a simple yet effective distance-based method for predicting values by identifying the K nearest neighbors of the testing data based on the training data (Aqsha, 2025). Gradient Boosted Trees and Random Forest are ensemble learning algorithms that improve prediction accuracy by combining multiple decision tree models. Meanwhile, Support Vector Regression was selected because of its ability to handle high-dimensional data, its robustness against overfitting, and its flexibility in modeling both linear and non-linear relationships (Handani et al., 2024).

The research process consisted of several sequential stages: problem identification, literature review, data collection, data preprocessing, algorithm implementation, result evaluation, comparative analysis, and conclusion. Each stage was conducted systematically to develop an effective house price prediction model for each algorithm evaluated. The detailed explanation of each research stage is presented as follows:

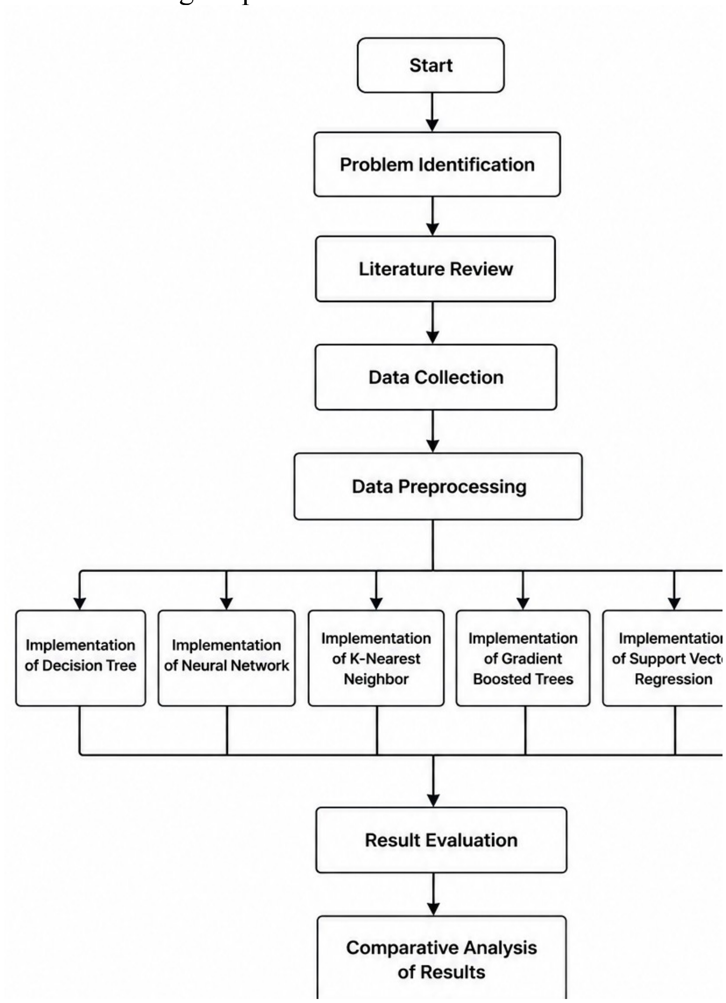

Figure. 1. Research Stages

Problem Identification, Researchers analyzed the challenges that arise when individuals or businesses struggle to accurately estimate house prices due to market fluctuations and interrelated factors such as building area, land area, number of bedrooms, and supporting facilities. Therefore, this study identified and analyzed the application and analysis of machine learning algorithms, specifically Decision Trees, Neural Networks, K-Nearest Neighbors, Gradient Boosted Trees, Support Vector Regression, and Random Forests, in modeling and predicting house prices. The ultimate goal was to determine the most effective and accurate method for predicting house prices using the Altair Studio platform.

Literature Study, A literature review was conducted to gain a deeper understanding of the basic theory and application of machine learning algorithms in various prediction contexts, particularly in the problem of property price prediction. This study included an explanation of the working principles of each algorithm, its advantages and disadvantages, and the relevance of each algorithm for processing data with diverse characteristics. Furthermore, the reviewed literature also included the results of previous research comparing the performance of several predictive algorithms, providing an overview of the accuracy, computational efficiency, and complexity of the resulting models.

Data Collection, This stage includes the process of searching, selecting, and collecting relevant house price datasets as a basis for developing a predictive model. The dataset used in this study was obtained from the Kaggle platform and consists of 1,010 records with 6 attributes: house price, building area, land area, number of bedrooms, number of bathrooms, and garage. The house price variable is used as the prediction target, while the other attributes are used as predictor variables. Before being used in the modeling process, a data quality check was conducted, including checking for missing values, duplicate data, and outliers. The results showed that the dataset had no missing values, so no data imputation process was required. In addition, several extreme values (outliers) were found in several numeric attributes. However, these values were retained because they still represent the real conditions of house characteristics and prices, so it is hoped that the model can learn data patterns more representatively.

Data Preprocessing, At this stage, the collected datasheets will be processed. Data preprocessing itself is performed using Altair Studio in several stages. Altair Studio was chosen because it offers a variety of intelligent and informative methods, such as classification, clustering, and prediction, allowing users to apply techniques tailored to their analysis needs (Sudarsono et al., 2021). The data preprocessing stages include data import, feature selection, data cleaning, data splitting, model visualization, and model testing. Next, comes the Algorithm Implementation process: at this stage, the preprocessed data will be implemented using six machine learning algorithms: Decision Tree, Neural Network, K-Nearest Neighbor, Gradient Boosted Trees, Support Vector Regression, and Random Forest. This implementation process produces predicted values for house prices.

Result Evaluation, Model performance evaluation is conducted by considering three main aspects: performance, computation time, and model complexity. The performance evaluation uses two regression metrics: Root Mean Square Error (RMSE) and Mean Absolute Error (MAE). RMSE measures the root mean square of the difference between predicted and actual values. Meanwhile, MAE calculates the average absolute difference between predicted and actual values. The combination of RMSE and MAE can provide a more comprehensive picture of model accuracy (Sukmaningtyas et al., 2024). Then, computational time is evaluated during model training and testing. Training time indicates how quickly the model can learn patterns from the training data, while testing time measures the model's efficiency in predicting new data. Finally, model complexity is evaluated, which describes the level of complexity of the structure and parameters used in the algorithm development and implementation process. Each algorithm has different complexity determinants, depending on the characteristics of the method used.

Comparative Analysis: This stage is conducted to assess the extent to which each algorithm is able to produce accurate, efficient, and balanced predictions in the context of house price modeling. This analysis is conducted quantitatively by considering the results of the model evaluation, namely model performance, computational time, and the complexity of each prediction model.

In conclusion, this study aims to provide a comprehensive overview of the performance of several machine learning algorithms in predicting house prices using the same dataset and a uniform evaluation method. The evaluation was conducted through three main aspects: model performance based on RMSE and MAE, computational time efficiency during the training and testing stages, and model complexity, which includes the structure and key parameters of each algorithm. This approach is expected to produce objective and measurable comparisons and provide recommendations for the best algorithm that balances accuracy, efficiency, and complexity in house price modeling.

RESULTS AND ANALYSIS

The research results were obtained through the training and testing process of six machine learning algorithms, namely Decision Tree, Neural Network, K-Nearest Neighbor, Gradient Boosted Trees, Support Vector Regression and Random Forest. Each algorithm was tested using the same dataset to make the comparison results more objective and consistent. The evaluation was carried out based on three main indicators, namely the level of prediction performance as measured by Root Mean Square Error (RMSE) and Mean Absolute Error (MAE), computational time efficiency in the training and testing stages, and model complexity as reviewed from the structure and main parameters of each algorithm. The following are the results of the model evaluation conducted:

1. Model Performance Evaluation Results

Based on the model performance evaluation stage, RMSE and MAE values were obtained for each algorithm model. RMSE indicates the average absolute error between the predicted house price and the actual house price. A lower RMSE value indicates that the predicted house price is closer to the actual house price. Meanwhile, MAE indicates the average absolute difference between the actual house price and the predicted house price. The lower the MAE value, the higher the model's accuracy in predicting house prices. This indicates that the difference between the predicted house price and the actual house price is smaller and the model is able to produce predictions closer to reality (Fauzi et al., 2025). The following table compares the RMSE and MAE values obtained by the models.

Table 1. Comparison results of RMSE and MAE values

Algoritma	RMSE	MAE
Decision Tree	0,649	0,478
Neural Network	0,630	0,464
K-Nearest Neighbor	0,608	0,485
Gradient Boosted Trees	0,747	0,589
Support Vector Regression	0,662	0,501
Random Forest	0,656	0,474

Based on the evaluation results shown in the table, KNN obtained the lowest RMSE value of 0.608, making it the model with the smallest mean squared error. This indicates that KNN is better at handling large errors than other algorithms. Neural Networks ranked second with an RMSE of 0.630, followed by Decision Tree (0.649), Random Forest (0.656), and SVR (0.662). Meanwhile, GBT had the highest RMSE value of 0.747, making it the lowest performing model compared to the other models.

In terms of MAE, Neural Networks performed best with the smallest value of 0.464. This indicates that Neural Networks produce predictions with the smallest mean absolute error, resulting in more consistent results. Random Forest (0.474) and Decision Tree (0.478) also performed quite well, slightly ahead of Neural Networks. Meanwhile, KNN (0.485) was slightly higher, although still within the competitive range. The SVR (SVR) of 0.501 performed lower than the previous model, while GBT again performed worse with the highest MAE value of 0.589.

Overall, it can be concluded that KNN excels in minimizing the mean squared error (RMSE), while Neural Network excels in minimizing the mean absolute error (MAE). Random Forest and Decision Tree can be categorized as models with intermediate performance because their results are quite stable, although not as good as KNN and Neural Network. SVR demonstrated relatively good performance, while GBT was the algorithm with

DOI: https://doi.org/10.33650/coreai.v6i2

the lowest performance in this study. Thus, the KNN and Neural Network algorithms can be considered the best models, with KNN being more effective in reducing large errors, while Neural Network is better at providing consistent predictions.

2. Computation Time Evaluation Results

The computation time of the prediction models created for each algorithm was measured. Computation time is divided into two parts: model training and model testing. The training computation time indicates how quickly the model can learn from the training data, while the testing computation time indicates the model's efficiency in predicting new data. The following table compares the computation time achieved by each model.

Table 2. Comparison results of training and testing time

Algoritma	Training	Testing
Decision Tree	0,006 s	0,004 s
Neural Network	0,246 s	0,003 s
K-Nearest Neighbor	0,004 s	0,023 s
Gradient Boosted Trees	0,164 s	0,004 s
Support Vector Regression	0,174 s	0,002 s
Random Forest	0,371 s	0,056 s

Based on the measurement results, the fastest training time was achieved by KNN (0.004 s), followed by Decision Tree (0.006 s). Meanwhile, the algorithms with the longest training times were Random Forest (0.371 s) and Neural Network (0.246 s), indicating that complex models require a longer training process. Gradient Boosted Trees (0.164 s) and SVR (0.174 s) were in the intermediate category.

In the testing phase, the fastest algorithm was SVR (0.002 s), followed by Neural Network (0.003 s). KNN, on the other hand, took the longest time to produce predictions, at 0.023 s. Random Forest (0.056 s) also took longer than the other algorithms, while Decision Tree and GBT had similar testing times, at 0.004 s.

Based on these results, it can be concluded that KNN is highly efficient in the training phase but less efficient in the testing phase, while SVR and GBT are relatively balanced between training and testing. Decision Trees can also be considered efficient because they require a short time for both stages, although their accuracy is not as good as KNN or Neural Networks. Conversely, Neural Networks require a long training time but are more efficient for testing. Random Forests require longer computation times for both training and testing, reflecting the higher level of model complexity compared to other algorithms.

3. Model Complexity Evaluation Results

The complexity evaluation in this study was conducted by considering the parameters generated by each algorithm.

Table 3. Results of model complexity comparison

Algoritma	**Main Parameters**	**Complexity**
Decision Tree	tree depth, leaf nodes	11 depth, 71 leaves
Neural Network	layer, neuron	3 layer, 12 neuron
K-Nearest Neighbor	K value, amount data, dimensi	K = 5, 707 data, 5 dimensi
Gradient Boosted Trees	tree, tree depth, leaf nodes	50 tree, depth 5, 24 leaves

Support Vector Regression	kernel, C value, epsilon	200 kernel, C = 10, ε = 0,5
Random Forest	tree, tree depth	100 tree, 10 depth

In a Decision Tree, complexity is determined by the tree depth and the number of leaf nodes. Tree depth describes the number of levels from the root to the furthest leaf, while leaf nodes are the final nodes containing decisions or predicted values. Test results show that the Decision Tree model has a tree depth of 11 levels with 71 leaf nodes. This indicates a relatively deep tree structure with numerous branches, so its complexity can be categorized as medium.

In Neural Networks, complexity is determined by the number of layers and neurons. Layers consist of input, hidden, and output layers, while neurons function as processing units in the network. Based on the analysis, the Neural Network model uses three layers with a total of 12 neurons. The complexity of this model is not considered high, but it is still greater than that of a Decision Tree because it involves signal propagation between neurons through the hidden layer.

In K-Nearest Neighbor, complexity is determined by the K value, the amount of training data, and the number of dimensions. The K value indicates the number of nearest neighbors used in the prediction, while the number of data and dimensions reflects the dataset size and the number of input variables. Test results show that the KNN model, using a K value of 5 using the weighted neighbor method, was trained using 707 data sets with 5 input variables. KNN complexity is relatively low during the training stage, but increases during the prediction stage because each test data set must calculate its distance from the entire training data set.

In Gradient Boosted Trees, complexity depends on the number of trees, tree depth, and leaf nodes. Based on the analysis, the GBT model uses 50 trees with a tree depth of 5 levels and an average of 24 leaf nodes. This indicates that GBT has a more complex structure than a single Decision Tree, as it involves multiple trees built incrementally (boosting) to increase accuracy.

In Support Vector Regression, complexity is determined by the use of kernels, the C parameter, and the epsilon parameter. The kernel functions to map the data to a higher-dimensional space, the C parameter controls the level of regularization, while the epsilon sets the margin of error tolerance. Test results show that SVR uses 200 kernels, with C = 10 and epsilon = 0.5. This condition indicates that SVR has a relatively high level of complexity, primarily due to the large number of kernel calculations involved.

Finally, in Random Forest, complexity is determined by the number of trees and tree depth. Based on test results, the Random Forest model built 100 trees with a tree depth of 10 levels. This indicates that Random Forest is significantly more complex than a single Decision Tree, as it involves hundreds of trees combined through a bagging mechanism to produce more stable and accurate predictions.

Overall, Decision Trees can be categorized as algorithms with relatively low complexity. Neural Networks and KNN are at an intermediate level of complexity. GBT, SVR, and Random Forest are considered high-complexity algorithms because they involve many trees or kernels. This complexity is directly proportional to the computational time required and affects the trade-off between model accuracy, interpretability, and efficiency.

CONCLUSIONS

Based on the evaluation results of the six machine learning algorithms used in this study Decision Tree, Neural Network, K-Nearest Neighbor, Gradient Boosted Trees, Support Vector Regression, and Random Forest the K-Nearest Neighbor algorithm is the model with the best performance in predicting house prices. Test results show that KNN achieved an RMSE value of 0.608 and an MAE value of 0.485, indicating the lowest prediction error rate compared to other algorithms. In terms of computational time, KNN also demonstrated the highest efficiency, with a training time of 0.004 seconds and a testing time of 0.023 seconds. Furthermore, KNN has a relatively simple model complexity yet still produces high prediction accuracy. These findings demonstrate that a simple model with a distance-based approach can produce accurate results without requiring lengthy computational processes.

Suggestions for further research include developing the model by adding more diverse and relevant variables to the determinants of house prices, such as geographic location, physical condition of the building, accessibility, and local economic factors. Future research is also recommended to implement the KNN algorithm in the form of a web-based prediction system or interactive application, so that the results of this study will not only contribute to the academic field, but also can be directly utilized by the public, property developers, and business actors in estimating house prices automatically and efficiently. With this development, it is hoped that the house price prediction system can be more adaptive, accurate, and applicable in various dynamic property market conditions.

REFERENCES

Rahayuningtyas, E. F., Rahayu, F. N., & Azhar, Y. (2021). Prediksi Harga Rumah Menggunakan General Regression Neural Network. Jurnal Informatika, 8(1), 59-66. https://doi.org/10.31294/ji.v8i1.9036

Ridho, I. I., Mahalisa, G., & Sari, D. R. (2022). Metode Neural Network untuk Penentuan Akurasi Prediksi Harga Rumah. Technologia, 13(1), 56-58. https://doi.org/10.31602/tji.v13i1.6252

Fitri, E. (2023). Analisis Perbandingan Metode Regresi Linier, Random Forest Regression dan Gradient Boosted Trees Regression untuk Prediksi Harga Rumah. Journal of Applied Computer Science and Technology (JACOST), 4(1). https://doi.org/10.52158/jacost.v4i1.491

Hidayah, F., Angesti, S. J., & Widyastuti, Y. P. (2024). Prediksi Harga Rumah di Boston Menggunakan Metode Linear Regression, Support Vector Regression, Decision Tree dan Random Forest Regression. Department of Statistics, 1-9.

Cucos, A. M., & Iantovics, L. B. (2024). Comparative Study of Random Forest, Gradient Boosted Trees, Feedforward Neural Networks and Convolutional Neural Networks Using Fingerprints and Molecular Descriptors for Adverse Drug Reaction Prediction. Procedia Computer Science, 1895-1904. https://doi.org/10.1016/j.procs.2024.09.702

Alarfaj, F. K., Malik, I., Khan, H. U., Almusallam, N., Ramzan, M., & Ahmed, M. (2022). Credit Card Fraud Detection Using State-of-the-Art Machine Learning and Deep Learning Algorithms. IEEE Access, 10, 39700-39715. https://doi.org/10.1109/ACCESS.2022.3166891

Karimi, P., Hosseini Varzaqani, M. Y., & Ghanbari, S. (2025). Machine Learning and Deep Learning Algorithms for Prediction of Different Parameters of a 2D Permanent Magnetic Lattice. Journal of Applied Physics, 137(18). https://doi.org/10.1063/5.0261091

Özlem, Ş., & Tan, O. F. (2022). Predicting Cash Holdings Using Supervised Machine Learning Algorithms. Financial Innovation, 8(1), 1-19. https://doi.org/10.1186/s40854-022-00351-8

Lee, Y. W., Choi, J. W., & Shin, E. H. (2021). Machine Learning Model for Predicting Malaria Using Clinical Information. Computers in Biology and Medicine, 129. https://doi.org/10.1016/j.compbiomed.2020.104151

Sari, B. W., & Prabowo, D. (2025). Analisis Perbandingan Prediksi Harga Rumah dengan Random Forest, Gradient Boosting, dan XGBoost. Indonesian Journal of Learning and Technological Innovation (INTELECT), 4(1), 42-51. https://doi.org/10.57255/intellect.v4i1.1385

Taranto-Vera, G., Galindo-Villardón, P., Merchán-Sánchez-Jara, J., Salazar-Pozo, J., Moreno-Salazar, A., & Salazar-Villalva, V. (2021). Algorithms and Software for Data Mining and Machine Learning: A Critical Comparative View from a Systematic Review of the Literature. Journal of Supercomputing, 77(10), 11481-11513. https://doi.org/10.1007/s11227-021-03708-5

Warjiyono, Rais, A. N., Alfarobi, I., Hadi, S. W., & Kurniawan, W. (2024). Analisa Prediksi Harga Jual Rumah Menggunakan Algoritma Random Forest Machine Learning. JURSISTEKNI (Jurnal Sistem Informasi dan Teknologi Informasi), 6(2), 416-423. https://doi.org/10.52005/jursistekni.v6i2.323

Manullang, N., Sembiring, R. W., Gunawan, I., Parlina, I., & Irawan, I. (2021). Implementasi Teknik Data Mining untuk Prediksi Peminatan Jurusan Siswa Menggunakan Algoritma C4.5. Jurnal Ilmu Komputer dan Teknologi, 2(2), 1-5. https://doi.org/10.35960/ikomti.v2i2.700

Illahi, P. K., Viana, A. R., Fitria, N., Permata, M., & Pratama, M. Y. (2023). Application of Decision Tree Algorithm and Linear Regression for Breast Cancer Classification. Institut Riset dan Publikasi Indonesia (IRPI), 86-92.

Ferdiansyah, M. R., et al. (2024). Prediksi Produksi Tanaman Padi di Sumatera dengan Menggunakan Algoritma Neural Network. Journal of Multidisciplinary Inquiry in Science Technology and Educational Research, 2(1B), 1454-1467.

Aqsha, D. (2025). Perbandingan Kinerja Algoritma Extreme Gradient Boosting dan Random Forest untuk Prediksi Harga Rumah di Jabodetabek. Jurnal Ilmu Komputer dan Sistem Informasi, 13(1), 1-7. https://doi.org/10.24912/jiksi.v13i1.32863

Handani, A. A., Siregar, A. M., & Al Mudzakir, T. (2024). Prediksi Harga Rumah di Karawang Menggunakan Algoritma Multiple Linear Regression dan Support Vector Regression. 5(2), 33-40.

Sudarsono, B. G., Leo, M. I., Santoso, A., & Hendrawan, F. (2021). Analisis Data Mining Data Netflix Menggunakan Aplikasi RapidMiner. JBASE - Journal of Business and Audit Information Systems, 4(1). https://doi.org/10.30813/jbase.v4i1.2729

Sukmaningtyas, Y. N., Akbar, R. M., & Rohma, G. (2024). Penerapan Predictive Analytics untuk Analisis Faktor-faktor yang Mempengaruhi Performa Akademik Siswa. 4(2), 127-145. https://doi.org/10.29240/arcitech.v4i2.12048

Fauzi, A., Maulidah, N., Supriyadi, R., Nalatissifa, H., & Diantika, S. (2025). Prediksi Harga Properti di Indonesia Menggunakan Algoritma Random Forest. 4(1), 43-49. https://doi.org/10.31004/riggs.v4i1.367