



Comparing Optimization Algorithms in ANN Models for House Price Prediction in Pekanbaru

Doni Winarso^{1*}, Edo Arribe², Syahril³, Aryanto⁴, Muhandi⁵, Sharulniza Musa⁶

^{1,2,3,4}Department of Information Systems, Faculty of Computer Science, Universitas Muhammadiyah Riau, Riau, Indonesia

⁵Departement of Informatics Engineering, Faculty of Computer Science, Universitas Hangtuah Pekanbaru, Riau, Indonesia

⁶Malaysian Institute of Information Technology, University Kuala Lumpur, Kuala Lumpur, Malaysia

¹doniwinarso@umri.ac.id, ²edoarribe@umri.ac.id, ³syahril@umri.ac.id, ⁴aryanto@umri.ac.id, ⁵muhandi.yudie@gmail.com,

⁶shahrulniza@unikl.edu.my

Abstract

This study evaluates the performance of five optimization algorithms in Artificial Neural Network (ANN) models for predicting house prices in Pekanbaru. The optimizers tested include Adam, AdaDelta, Stochastic Gradient Descent (SGD), Nadam, and Adaptive Sharpness-Aware Minimization (ASAM). A total of 3,149 house sales records were collected from rumah123.com between January and December 2024. After cleaning 148 incomplete entries, 3,001 valid records remained. The dataset included seven features: price, location, number of bedrooms, number of bathrooms, land area, building area, and garage capacity, with the location encoded using one-hot encoding. The research involved a literature review, problem formulation, data acquisition, preprocessing, model development, and evaluation. Model performance was assessed using the Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Root Mean Square Error (RMSE). The results show that SGD consistently achieved the best performance, particularly at a 90:10 train-test split, with the lowest MAPE (1.74%) and MSE (0.3279). Adam and Nadam also performed well, while ASAM had the highest error (MAPE 6.14%). These findings indicate that SGD was the most effective optimizer for this dataset. Future research should explore larger datasets and advanced hyperparameter tuning to improve the generalizability of this model.

Keywords: AdaDelta; stochastic gradient descent (SGD); adaptive moment estimation (Adam); adaptive sharpness-aware minimization (ASAM); artificial neural network (ANN); house price prediction; optimization; nadam

How to Cite: D. Winarso, Edo Arribe, Syahril, Aryanto, Muhandi, and Shahrulniza Musa, "Comparing Optimization Algorithms in ANN Models for House Price Prediction in Pekanbaru", *J. RESTI (Rekayasa Sist. Teknol. Inf.)*, vol. 9, no. 4, pp. 781 - 787, Aug. 2025.

Permalink/DOI: <https://doi.org/10.29207/resti.v9i4.6619>

Received: May 2, 2025

Accepted: July 2, 2025

Available Online: August 17, 2025

This is an open-access article under the CC BY 4.0 License

Published by Ikatan Ahli Informatika Indonesia

1. Introduction

Housing constitutes one of the essential human necessities, positioned alongside food and clothing. According to Maslow's Hierarchy of Needs, housing is classified as a physiological necessity [1], [2]. A house provides shelter and a sense of security, serving as a vital component in human well-being [3]. It is essential to emphasize that housing should not be regarded merely as a luxury, but rather as a basic necessity for human life [4]. In the city of Pekanbaru, Riau Province, Indonesia housing prices have exhibited significant fluctuations, impacting both buyers and sellers in the real estate market.

In recent years, Pekanbaru, Riau, has recorded the highest increase in housing prices compared to other regions in Indonesia. According to the Residential The

Residential Property Price Survey (SHPR) conducted by Bank Indonesia in the third quarter of 2024 [5], the Residential Property Price Index (IHPR) increased by 1.94% year-on-year (yoy), surpassing the 1.66% (yoy) growth observed in the previous quarter. Notably, Pekanbaru recorded the highest IHPR growth at 4.55% (yoy), followed by Yogyakarta at 4.51% (yoy) and Pontianak at 3.12% (yoy). Such abnormal price increases negatively impact both sellers and buyers. For sellers, disproportionate price hikes may hinder property sales and create a negative perception of housing affordability. Conversely, for buyers, rising prices limit purchasing power and force many to consider rental options instead of ownership.

Housing prices are generally influenced by various features, which can be categorized into non-

geographical and geographical factors [6], [7]. Non-geographical factors include attributes such as the number of bedrooms, floor area, and building age, while geographical factors involve proximity to city centers, accessibility to public transportation, and the quality of surrounding educational institutions [8]. Understanding the pricing patterns and determinants is not only relevant for potential buyers and sellers but also crucial for policymakers, real estate analysts, and urban planners [9].

The volatility of housing prices has created a growing demand for sophisticated predictive models that support stakeholders in formulating well-informed decisions. The unpredictability of housing prices has prompted the need for advanced prediction models that can assist stakeholders in making informed decisions [10]. In this regard, Machine Learning (ML) techniques provide effective solutions through their capacity to extract patterns from historical data and generate precise predictions. [11]. Computer-based predictive systems help determine optimal buying and selling times, reducing risks associated with market volatility [9]. Among various ML approaches, Artificial Neural Networks (ANN) have been extensively studied in recent years owing to their powerful ability to handle nonlinear challenges. By efficiently modeling complex dependencies among variables, they provide a versatile framework for solving diverse computational and predictive problems. [12].

Artificial Neural Networks (ANN) have been extensively acknowledged as powerful approaches for addressing regression tasks, primarily due to their ability to model complex and nonlinear associations between input and output variables [13]. By incorporating multiple layers including an input layer, several hidden layers, and an output layer ANNs facilitate the discovery of underlying patterns within large-scale datasets [14], [15]. Prior studies have demonstrated the efficiency of ANN in real estate price

prediction. For instance, Rampini and Re Cecconi [12] implemented ANN to predict housing prices in the Italian real estate market with high accuracy. Similarly, [16] employed deep learning techniques in Sakarya, Turkey, to estimate housing prices effectively. However, despite its advantages, ANN is susceptible to overfitting, a condition where the model memorizes training data instead of generalizing patterns, leading to decreased performance on unseen data [8], [17], [18].

To address these limitations, optimization algorithms play a crucial role in enhancing ANN performance [19]. Optimization refers to refining existing models or designing optimal frameworks to achieve superior results [20], [21]. Various optimization techniques have been explored in prior research. For example, A study has been conducted to compare Stochastic Gradient Descent (SGD), Adaptive Delta (AdaDelta), and Adaptive Moment Estimation (Adam) in image classification tasks and found that Adam provided the highest accuracy [1]. Other optimization methods, such as Nesterov Adam (Nadam)[22], Adaptive Sharpness-aware Minimization (AdaSAM) and SGD [23], [24], have also shown effectiveness in improving ML models.

Given the background of housing price fluctuations and the need for accurate predictions, this study aims to optimize the Artificial Neural Network model using various optimization algorithms. By incorporating different optimization techniques, the proposed model seeks to enhance prediction accuracy and generalization performance. This research aims to contribute to real estate analytics by offering insights for buyers, sellers, and policymakers on housing market dynamics.

2. Methods

This study was conducted systematically, with each stage explained within the framework as shown in Figure 1:

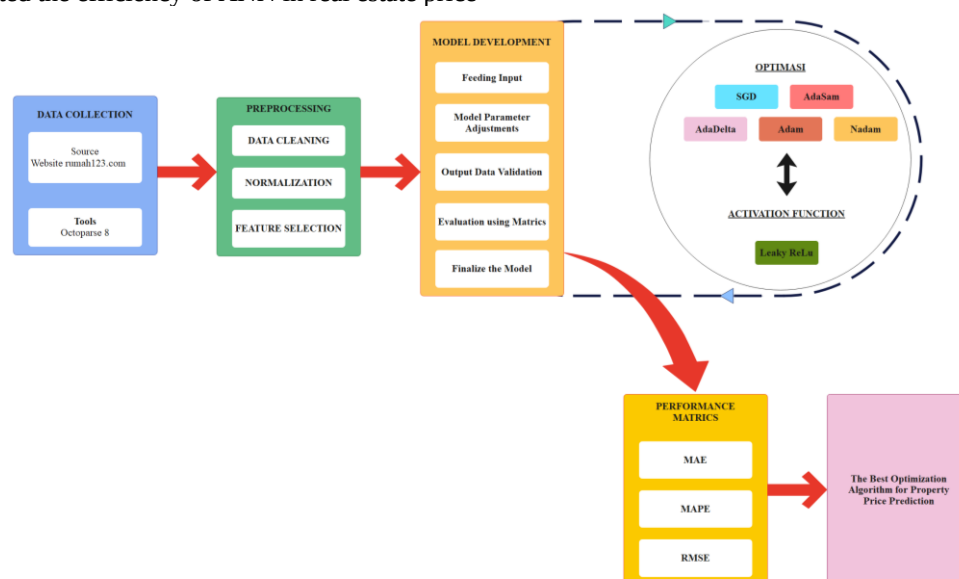


Figure 1. Research Stages for ANN-Based House Price Prediction

2.1 Data Collection

The dataset employed in this study consists of housing sales data from Pekanbaru, systematically collected using web scraping techniques. The data were sourced from Rumah123.com, one of Indonesia's leading real estate listing platforms, and the extraction was performed with the Octoparse 8 application. This method allowed for efficient and structured data collection from numerous property listings available on the website. The collected data represent actual market listings, providing a reliable basis for developing a predictive model of housing prices.

The dataset includes seven key features frequently employed in real estate valuation: price, location, the number of bedrooms, bathrooms, land area, and building area and garage capacity. Among these, the price acts as the target variable for prediction, as it represents the final market value of a property. Location is a critical determinant of price, influenced by factors such as neighborhood accessibility, proximity to public facilities, and ongoing regional development. Properties located in prime or strategically developed areas often command significantly higher prices compared to those in less developed or remote zones.

Other variables also contribute substantially to property valuation. The number of bedrooms and bathrooms serves as an indicator of comfort, privacy, and suitability for family living, thereby influencing buyer interest. Land area and building area are important in estimating the physical size and potential utility of the property, where larger dimensions typically correlate with higher price tags. Garage capacity, though sometimes overlooked, is particularly relevant in urban settings like Pekanbaru, where personal vehicle ownership is common; thus, having sufficient parking space can increase a house's attractiveness and market value. These features collectively provide a multidimensional view of each property, enabling more accurate price prediction through machine learning models.

2.2 Preprocessing

Data preprocessing serves as a fundamental process for improving the overall quality of the data and ensuring optimal integration with machine learning models [25], [26]. The process begins with the removal of duplicate records and irrelevant attributes, such as id, which do not contribute to predictive accuracy and may introduce noise into the model. Eliminating such attributes improves the efficiency of the modeling process.

Handling outliers is another critical step, as their presence can significantly disrupt model performance [27]. Various techniques, including trimming, winsorizing, or data transformation, can be applied to mitigate their impact [28], [29]. Addressing missing values is equally important to prevent biased analysis. Common approaches for handling missing data include

imputation based on mean, median, or model-driven techniques [30], [31].

To ensure consistency across features, data normalization is applied, particularly for numerical variables such as house price, land area, and building size. Normalization scales these values within a defined range, typically between 0 and 1, to prevent disproportionate influence from features with larger numerical scales. Categorical data, such as location, is converted into numerical representations through encoding techniques to enable effective use in predictive modeling.

After completing the data cleaning and transformation processes, the dataset is partitioned into distinct subsets for training and testing purposes. This step ensures that the predictive model can be accurately evaluated before being deployed on new data, ultimately improving its reliability and generalization capability.

2.3 Model Development Process

Ensuring that the developed Artificial Neural Network can generate accurate and reliable house price predictions is a critical aspect of model development. The initial stage involves providing relevant input data, where key features such as price, location, the number of bedrooms, bathrooms, land area, building area, and garage capacity are used as input variables.

Once the data is incorporated, the next step is adjusting model parameters to optimize learning. This includes selecting the appropriate optimization method and activation function, which are essential for improving model performance and prediction accuracy. Proper parameter selection is essential to ensure that the model effectively captures patterns within the data.

The subsequent step is validating the model's output, where predictions are compared against historical data to assess accuracy levels. This validation process ensures that the model is well-calibrated and capable of generating reliable estimates. By analyzing the discrepancies between predicted and actual values, further adjustments can be made to improve model performance.

To assess the predictive capabilities of the model, several performance metrics were employed. The primary metrics include Mean Absolute Error (MAE), Mean Squared Error (MSE), Mean Absolute Percentage Error (MAPE), and Root Mean Squared Error (RMSE). These measures facilitate a quantitative evaluation of the model's performance and assist in pinpointing aspects that may require further improvement. The following mathematical formulas were utilized to calculate the statistical metrics. [13], [32], [33], [34].

Root Mean Square Error (RMSE) is a commonly used metric that measures the square root of the average of the squared differences between the actual values (y_i) and the predicted values ($\hat{y}_i$). RMSE provides an indication of the magnitude of prediction errors, with

larger errors being penalized more due to the squaring process. It is useful for evaluating the overall accuracy of regression models, especially when large deviations are critical to detect [35]. The formula for RMSE is shown in Equation 1.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (1)$$

where y_i denotes the actual (true) value, $\hat{y}_i$ represents the predicted (output) value, and n is the total number of observations.

Mean Absolute Percentage Error (MAPE) measures the average absolute percentage deviation between observed and predicted values. By expressing errors in percentage terms, this metric facilitates easier interpretation across varying scales. However, MAPE is sensitive to zero or near-zero actual values, which can lead to distortion or undefined results [35], [36]. The formula for MAPE is shown in Equation 2.

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left(\left| \frac{y_i - \hat{y}_i}{y_i} \right| \right) \times 100 \quad (2)$$

where y_i denotes the actual (true) value, $\hat{y}_i$ is the predicted (output) value, and n represents the total number of observations.

Mean Absolute Error (MAE) calculates the average magnitude of the absolute errors between actual and predicted values. Unlike RMSE, it treats all errors equally, making it more robust to outliers. MAE is intuitive and represents the average absolute deviation in the original unit of the data [36]. The formula for MAE can be shown in Equation 3.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (3)$$

where y_i represents the actual (true) value, $\hat{y}_i$ denotes the predicted (output) value, and n is the total number of observations.

Mean Squared Error (MSE) measures the average of the squared differences between the actual and predicted values. Like RMSE, it penalizes larger errors more severely, making it useful for identifying models with significant variance in prediction. MSE is also commonly used in optimization because it is differentiable [36]. The formula for MSE can be shown in Equation 4.

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (4)$$

Where y_i denotes the actual (true) value, $\hat{y}_i$ represents the predicted value, and n is the total number of observations.

2.4 Proposed Model

The final stage in the model development process is known as the finalization phase, which plays a pivotal role in ensuring the model's readiness for real-world application. This phase involves selecting the most suitable optimization algorithm based on comprehensive evaluations conducted in the previous

stages. Each algorithm's performance is carefully assessed through various metrics such as Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), and Mean Squared Error (MSE), as defined in Equation (1), Equation (2), Equation (3), and Equation (4), respectively, to determine its ability to produce accurate and stable predictions. By comparing these outcomes, the algorithm that consistently demonstrates the best results across validation datasets is chosen as the final optimizer for the model.

Once the most effective optimization strategy is identified, the model undergoes a refinement process to reach its peak predictive capability. This may involve additional fine-tuning of the hyperparameters and re-training using the full dataset to ensure generalizability. At this stage, the model was tested under conditions that simulated real-world data to verify its robustness and reliability. The goal is to minimize the prediction error while maintaining computational efficiency, particularly when the model is intended to be integrated into larger decision-making systems or user-facing applications.

After achieving optimal performance, the model is prepared for deployment to predict house prices in the city of Pekanbaru. Deployment involves not only making the model accessible via an application interface or software system but also ensuring that it can handle live data input with consistent accuracy. This step is crucial for delivering actionable insights to stakeholders such as homebuyers, real estate agents, property developers, and financial institutions. A well-deployed model enables these parties to make informed decisions based on data-driven price estimates, thereby enhancing transparency and efficiency in the real estate market.

3. Results and Discussions

The results of the conducted experiments are presented in this section, organized into several main subsections.

The research data, obtained through a web scraping process, initially comprised 3,149 rows and 8 attributes. Prior to analysis, the dataset underwent a comprehensive preprocessing phase, which included handling missing values, removing duplicates, and correcting inconsistencies. As a result of this cleaning process, the final dataset was reduced to 3,001 valid entries. A summary of the cleaned dataset is presented, as shown in Table 1.

Table 1. Data Research After Cleaning

Attribute	Non-Null Count	Dtype
Advantages	3001 non-null	object
Price	3001 non-null	int64
Location	3001 non-null	object
Bedrooms	3001 non-null	int64
Bathrooms	3001 non-null	int64
Land_area (m ²)	3001 non-null	int64
Building_area (m ²)	3001 non-null	int64
Additional_info	3001 non-null	object

This stage comprises a series of steps to build, train, and evaluate the Artificial Neural Network (ANN) model. In this study, the ANN model employs optimization algorithms, namely Adam, Nadam, SGD, and AdaDelta. To provide a systematic overview of the implementation and evaluation process, the following pseudocode outlines the key steps involved:

```
BEGIN
  LOAD house price dataset
  REMOVE incomplete records
  ENCODE categorical variables using
  one-hot encoding
  NORMALIZE numerical features using
  MinMaxScaler
  SPLIT dataset into training and
  testing sets
  DEFINE optimizers = [Adam, Nadam,
  SGD, AdaDelta]

  FOR EACH optimizer in optimizers DO:
    INITIALIZE ANN model with:
      - Input layer: number of features
      - Hidden layers: ReLU activation
      - Output layer: linear activation

    COMPILE model with:
      - Loss: Mean Squared Error (MSE)
      - Optimizer: current optimizer
      - Metrics: MAE, MAPE

    TRAIN model on training set
      - Epochs = N
      - Batch size = B
      - Validation split = 0.2

    EVALUATE model on testing set
    COMPUTE RMSE = sqrt (MSE)
    PRINT optimizer and evaluation
    metrics
  END FOR
END
```

As shown in Table 2, the findings in this study are derived from the test results and are calculated based on Equation (1), Equation (2), Equation (3), and Equation (4), which represent MSE, MAE, MAPE, and RMSE, respectively.

Table 2. Result Test

Optimization	Split Data	MSE	MAE	MAPE	RSME
Adam	90 : 10	0.7101	0.3724	1.80%	0.5040
	80 : 20	0.5947	0.3538	1.71%	0.4699
	70 : 30	0.8807	0.4137	1.99%	0.5781
AdaDelta	90 : 10	0.4603	0.3576	1.74%	0.5001
	80 : 20	0.4822	0.3513	1.70%	0.4775
	70 : 30	0.5200	0.3607	1.75%	0.4804
SGD	90 : 10	0.3279	0.3575	1.74%	0.4915
	80 : 20	0.4127	0.3573	1.74%	0.4813
	70 : 30	0.3693	0.3538	1.72%	0.4705
Nadam	90 : 10	0.4461	0.3582	1.74%	0.4887
	80 : 20	0.4103	0.3473	1.68%	0.4604
	70 : 30	0.4395	0.3514	1.70%	0.4697
Asam	90 : 10	2.4834	1.2779	6.14%	1.5759
	80 : 20	3.2414	1.41	6.79%	1.8004
	70 : 30	2.4501	1.2264	5.89%	1.5653

This section discusses the comparative performance of four primary optimization algorithms—SGD, Adam, Nadam, and AdaDelta—used in Artificial Neural Network (ANN) models for house price prediction. The performance is evaluated across three different train-test split configurations: 90:10, 80:20, and 70:30. Evaluation metrics include MSE, Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Root Mean Square Error (RMSE).

Stochastic Gradient Descent (SGD) demonstrated the most consistent and optimal performance across all data splits. On the 90:10 split, it achieved the lowest MSE (0.3279) and MAPE (1.74%), with similarly strong results on the 80:20 and 70:30 splits. These outcomes suggest that SGD is particularly well-suited for structured regression problems such as house price prediction, where a simple and deterministic optimization path yields better generalization.

Adam, known for its adaptive learning rate and momentum, showed variable performance. It performed well on the 80:20 split, achieving a MSE of 0.5947 and MAPE of 1.71%. However, its performance degraded significantly with the 70:30 split, with increased error rates. This instability across different data distributions indicates that Adam may be sensitive to the training data proportion and structure in this context.

Nadam, which integrates Adam with Nesterov momentum, exhibited slightly more stable behavior than Adam. Its best performance was also observed on the 80:20 split, yielding the lowest MAPE among all optimizers (1.68%). Despite this, it did not outperform SGD in terms of consistency or average performance across splits.

AdaDelta showed moderate and stable performance, but its evaluation metrics remained generally higher than those of SGD and Nadam. While it maintained stability across different splits, its relative underperformance suggests that it is less suited for the problem compared to other optimizers.

In contrast, ASAM (Adaptive Sharpness-Aware Minimization) performed the worst across all metrics and data splits. For example, on the 90:10 split, ASAM recorded a significantly high MSE (2.4834) and MAPE (6.14%). These results indicate that ASAM may not be appropriate for this type of structured regression problem, possibly due to its sensitivity to sharp minima in the loss landscape that do not generalize well.

The experimental results revealed several key findings regarding the performance of different optimization algorithms in predicting house prices. Among the optimizers evaluated, Stochastic Gradient Descent (SGD) consistently demonstrated the best overall performance across all data split configurations, indicating its robustness and adaptability to varying training scenarios. In contrast, Nadam achieved the lowest Mean Absolute Percentage Error (MAPE) of

1.68% in the 80:20 train-test split configuration. However, this optimal result was not consistently replicated across other split ratios, suggesting that Nadam's performance may be sensitive to changes in data distribution.

Meanwhile, the Adam optimizer exhibited fluctuating outcomes throughout the experiments. Although it performed well in some settings, its accuracy declined noticeably in wider test splits, indicating reduced generalization ability in scenarios with less training data. The AdaDelta optimizer demonstrated stable performance across all configurations but failed to achieve high accuracy when compared to the other methods, making it a reliable yet less competitive choice. Finally, the Adaptive Sharpness-Aware Minimization (ASAM) algorithm consistently produced the weakest results in this study, suggesting that it may not be suitable for the type of data and prediction task used in this research.

These findings highlight that although adaptive optimizers like Adam and Nadam are popular in deep learning, traditional optimizers like SGD may deliver better generalization and robustness for structured tabular regression tasks such as house price prediction—particularly when the input features are well-normalized and free of noise.

4. Conclusions

This study has evaluated the impact of several optimization algorithms on the performance of ANN models for house price prediction in Pekanbaru. The results show that SGD offers the most accurate and reliable performance across different data splits, achieving the lowest MAPE values and consistent generalization. Adam and AdaDelta also delivered robust results with slight variations in error rates, making them viable alternatives depending on specific modeling needs. Although Nadam demonstrated competitive performance in selected scenarios, its higher loss values limit its consistency. Conversely, ASAM yielded the poorest outcomes, indicating its inadequacy for the dataset and prediction task at hand. Overall, the study confirms the importance of optimizer selection in ANN-based predictive modeling and suggests that conventional optimizers such as SGD, Adam, and AdaDelta continue to offer strong performance in structured regression tasks.

Notwithstanding its contributions, this study has certain limitations. First, The dataset employed for model training and validation was confined to housing price data from a single geographic region, namely Pekanbaru, which may restrict the generalizability of the findings to other regions with different market dynamics. Second, the study employed a fixed ANN architecture across all experiments, without exploring the effects of different network depths, activation functions, or regularization techniques, which may influence the performance of certain optimizers. Third,

hyperparameter tuning was performed manually within constrained bounds; more advanced techniques such as Bayesian optimization or grid search might yield different outcomes.

Future research should focus on expanding the geographical scope of the dataset to include diverse regions, allowing for more generalizable conclusions. Additionally, experimentation with various ANN architectures and the integration of ensemble learning techniques could provide a deeper understanding of optimizer behavior in complex models. It would also be beneficial to investigate the impact of automated hyperparameter tuning methods and hybrid optimization strategies to further enhance model performance. Moreover, exploring the applicability of these optimizers to alternative predictive tasks such as rental price estimation, urban development forecasting, or time-series-based real estate analysis could offer broader insights into their effectiveness across domains.

Acknowledgements

The authors would like to express their sincere gratitude to the Majelis Pendidikan Tinggi Penelitian dan Pengembangan (Dikti Litbang) Muhammadiyah for the financial support provided through the Hibah Riset Nasional Muhammadiyah Batch VIII Tahun 2024 (Grant No. 0258.380/1.3/D/2025). The authors also wish to thank Universitas Muhammadiyah for the continuous support and research facilities that have contributed significantly to the successful completion of this study.

References

- [1] E. F. Rahayuningtyas, F. N. Rahayu, and Y. Azhar, "Prediksi Harga Rumah Menggunakan General Regression Neural Network," *J. Inform.*, vol. 8, no. 1, pp. 59–66, 2021, doi: 10.31294/ji.v8i1.9036.
- [2] A. J. Trivedi and A. Metha, "Maslow's Hierarchy of Needs. A Theory of Human Motivation. Psychological," *Int. J. Res. all Subj. Multi Lang.*, vol. 7, no. 6, pp. 38–41, 2019.
- [3] S. Jagannath, B. Gatersleben, and E. Ratcliffe, "Flexibility of the home and residents' psychological wellbeing," *J. Environ. Psychol.*, vol. 96, no. March, p. 102333, 2024, doi: 10.1016/j.jenvp.2024.102333.
- [4] D. E. Aniobi, C. O. Ochuba, and S. B. Nguideen, "House Price Prediction: Comparative Analysis of Regression-Based Machine Learning Algorithms," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 11, no. 10, pp. 1550–1557, 2023, doi: 10.22214/ijraset.2023.56232.
- [5] Bank Indonesia, "Survei Harga Property Residential Triwulan III 2024," 2024.
- [6] T. Liu, J. Wang, L. Liu, Z. Peng, and H. Wu, "What Are the Pivotal Factors Influencing Housing Prices? A Spatiotemporal Dynamic Analysis Across Market Cycles from Upturn to Downturn in Wuhan," *Land*, vol. 14, no. 2, 2025, doi: 10.3390/land14020356.
- [7] X. Liang, Y. Liu, T. Qiu, Y. Jing, and F. Fang, "The effects of locational factors on the housing prices of residential communities: The case of Ningbo, China," *Habitat Int.*, vol. 81, no. August, pp. 1–11, 2018, doi: 10.1016/j.habitatint.2018.09.004.
- [8] G. Gao and L. Zhang, "Demand Forecasting of Emergency Materials based on Fuzzy Case-based Reasoning," *Front. Econ. Manag.*, 2022.
- [9] A. B. Adetunji, O. N. Akande, F. A. Ajala, O. Oyewo, Y. F.

- Akande, and G. Oluwadara, "House Price Prediction using Random Forest Machine Learning Technique," *Procedia Comput. Sci.*, vol. 199, pp. 806–813, 2021, doi: 10.1016/j.procs.2022.01.100.
- [10] S. B. Jha, R. F. Babiceanu, V. Pandey, and R. K. Jha, "Housing Market Prediction Problem using Different Machine Learning Algorithms: A Case Study," 2020.
- [11] S. Yu, T. Zhang, Y. Zhang, Z. Huang, H. Gao, and ..., "Intelligent setting of process parameters for injection molding based on case-based reasoning of molding features," *J. Intell. ...*, 2020, doi: 10.1007/s10845-020-01658-y.
- [12] L. Rampini and F. Re Cecconi, "Artificial intelligence algorithms to predict Italian real estate market prices," *J. Prop. Invest. Financ.*, vol. 40, no. 6, pp. 588–611, 2021, doi: 10.1108/JPIF-08-2021-0073.
- [13] P. Q. P. Nguyen, D. T. Nguyen, N. H. T. Yen, Q. Le, N. T. Nguyen, and N. D. K. Pham, "Machine Learning-Driven Insights for Optimizing Ship Fuel Consumption: Predictive Modeling and Operational Efficiency," *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 15, no. 1, pp. 27–35, 2025, doi: 10.18517/ijaseit.15.1.12374.
- [14] M. G. M. Abdolrasol et al., "Artificial neural networks based optimization techniques: A review," *Electron.*, vol. 10, no. 21, 2021, doi: 10.3390/electronics10212689.
- [15] S. F. Ahmed et al., *Deep learning modelling techniques: current progress, applications, advantages, and challenges*, vol. 56, no. 11. Springer Netherlands, 2023. doi: 10.1007/s10462-023-10466-8.
- [16] M. Ozdemir, K. Yildiz, and B. Buyuktanur, "Housing Price Estimation with Deep Learning: A Case Study of Sakarya Turkey," *Bilecik Şeyh Edebali Üniversitesi Fen Bilim. Derg.*, vol. 9, no. 1, pp. 138–151, 2022, doi: 10.35193/bseufbd.998331.
- [17] S. N. A. Rahman, N. H. A. Maimun, M. N. Razali, and S. Ismail, "The artificial neural network model (ANN) for Malaysian housing market analysis," *Plan. Malaysia*, vol. 17, no. 1, pp. 1–9, 2019, doi: 10.21837/pmjournal.v17.i9.581.
- [18] A. Yağmur, M. Kayakuş, and M. Terzioğlu, "House price prediction modeling using machine learning techniques: a comparative study," *Aestimum*, vol. 81, pp. 39–51, 2022, doi: 10.36253/aestim-13703.
- [19] İ. Gölcük, F. B. Özsoydan, and E. D. Durmaz, "An improved arithmetic optimization algorithm for training feedforward neural networks under dynamic environments," *Knowledge-Based Syst.*, vol. 263, 2023, doi: 10.1016/j.knosys.2023.110274.
- [20] F. Mehmood, S. Ahmad, and T. K. Whangbo, "An Efficient Optimization Technique for Training Deep Neural Networks," *Mathematics*, vol. 11, no. 6, 2023, doi: 10.3390/math11061360.
- [21] A. O. Fajemisin, D. Maragno, and D. den Hertog, "Optimization with constraint learning: A framework and survey," *Eur. J. Oper. Res.*, vol. 314, no. 1, pp. 1–14, 2024, doi: 10.1016/j.ejor.2023.04.041.
- [22] S. Zahara, Sugianto, and M. B. Ilmiddaviq, "Consumer price index prediction using Long Short Term Memory (LSTM) based cloud computing," *J. Phys. Conf. Ser.*, vol. 1456, no. 1, 2020, doi: 10.1088/1742-6596/1456/1/012022.
- [23] H. Sun et al., "AdaSAM: Boosting sharpness-aware minimization with adaptive learning rate and momentum for training deep neural networks," *Neural Networks*, vol. 169, no. June 2023, pp. 506–519, 2024, doi: 10.1016/j.neunet.2023.10.044.
- [24] S. Nithiyanantha Vasagam, B. Ravikumar, R. Kavibharathi, J. Keerthana, R. Sathya Narayanan, and K. Geetika, "Prediction of leather footwear export using learning algorithms based on ANN model," *Expert Syst. Appl.*, vol. 238, no. PA, p. 121809, 2024, doi: 10.1016/j.eswa.2023.121809.
- [25] A. Bagheri Tofighi, A. Ahmadi, and H. Mosadegh, "A novel case-based reasoning system for explainable lung cancer diagnosis," *Comput. Biol. Med.*, vol. 185, no. December 2024, 2025, doi: 10.1016/j.combiomed.2024.109547.
- [26] S. Z. Golazad, A. Mohammadi, A. Rashidi, and M. Ilbeigi, "From raw to refined: Data preprocessing for construction machine learning (ML), deep learning (DL), and reinforcement learning (RL) models," *Autom. Constr.*, vol. 168, no. PA, p. 105844, 2024, doi: 10.1016/j.autcon.2024.105844.
- [27] S. Huang, J. Dang, C. C. Shekter, H. A. Yenikomshian, and J. Gillenwater, "A systematic review of machine learning and automation in burn wound evaluation: A promising but developing frontier," *Burns*, vol. 47, no. 8, pp. 1691–1704, 2021, doi: 10.1016/j.burns.2021.07.007.
- [28] C. S. K. Dash, A. K. Behera, S. Dehuri, and A. Ghosh, "An outliers detection and elimination framework in classification task of data mining," *Decis. Anal. J.*, vol. 6, no. January, p. 100164, 2023, doi: 10.1016/j.dajour.2023.100164.
- [29] E. Mekelburg and J. Strauss, "Pooling and winsorizing machine learning forecasts to predict stock returns with high-dimensional data," *J. Empir. Financ.*, vol. 79, no. August, p. 101538, 2024, doi: 10.1016/j.jempfin.2024.101538.
- [30] A. Rácz and A. Gere, "Comparison of missing value imputation tools for machine learning models based on product development cases studies," *Lwt*, vol. 221, no. September 2024, 2025, doi: 10.1016/j.lwt.2025.117585.
- [31] S. Alam, M. S. Ayub, S. Arora, and M. A. Khan, "An investigation of the imputation techniques for missing values in ordinal data enhancing clustering and classification analysis validity," *Decis. Anal. J.*, vol. 9, no. October, p. 100341, 2023, doi: 10.1016/j.dajour.2023.100341.
- [32] S. Roy, A. Ghosh, A. K. Das, and R. Banerjee, "Development and validation of a GEP model to predict the performance and exhaust emission parameters of a CRDI assisted single cylinder diesel engine coupled with EGR," *Appl. Energy*, vol. 140, pp. 52–64, 2015, doi: 10.1016/j.apenergy.2014.11.065.
- [33] A. Jierula, S. Wang, T. M. Oh, and P. Wang, "Study on accuracy metrics for evaluating the predictions of damage locations in deep piles using artificial neural networks with acoustic emission data," *Appl. Sci.*, vol. 11, no. 5, pp. 1–21, 2021, doi: 10.3390/app11052314.
- [34] B. D. Ross-Veitia et al., "Machine learning regression algorithms to predict emissions from steam boilers," *Heliyon*, vol. 10, no. 5, 2024, doi: 10.1016/j.heliyon.2024.e26892.
- [35] N. F. Prasetyo, W. Witanti, and A. I. Hadiana, "Forecasting Stock Returns Using Long Short-Term Memory (LSTM) Model Based on Inflation Data and Historical Stock Price Movements," *Resti*, vol. 9, no. 3, pp. 453–464, 2025, doi: https://doi.org/10.29207/resti.v9i3.6422.
- [36] D. Suryani and M. Fadhillah, "Indonesian Crude Oil Price (ICP) Prediction Using Support Vector Regression Algorithm," *J. RESTI*, vol. 8, no. 1, pp. 127–134, 2024, doi: 10.29207/resti.v8i1.5551.