



Analysis of the Impact of Backpropagation Hyperparameter Optimization on Heart Disease Prediction Models

Nita Syahputri^{1*}, Putrama Alkhairi², Enok Tuti Alawiah³

¹Faculty of Engineering and Computer Science, Department of Information Systems, Universitas Potensi Utama, Medan, Indonesia

²Informatics Management Program, STIKOM Tunas Bangsa, Pematangsiantar, Indonesia

³Information Systems, Universitas Bina Sarana Informatika, Karawang, Indonesia

¹nita20d@gmail.com, ²putrama@amiktunasbangsa.ac.id, ³enok.etw@bsi.ac.id

Abstract

Heart disease represents a significant global health concern, underscoring the importance of early and accurate predictive models to minimize complications and improve patient prognosis. The Backpropagation Neural Network (BPNN) has been widely utilized for heart disease prediction; however, its effectiveness is highly dependent on the selection of appropriate hyperparameters, including the number of neurons, activation function, optimizer, and batch size. In this study, the influence of hyperparameter optimization on BPNN performance was investigated. A baseline BPNN model was evaluated alongside an optimized counterpart in which key hyperparameters had been systematically fine-tuned to improve both predictive accuracy and model stability. Both models were trained and validated using an identical dataset, and their performances were assessed based on Accuracy, Precision, Recall, Mean Squared Error (MSE), and Mean Absolute Error (MAE). The optimized model demonstrated marginally higher accuracy (99.11% compared to 99.09%) and slightly lower error rates (MSE and MAE of 0.0089 versus 0.0091). Moreover, it achieved superior precision, indicating enhanced reliability in correctly identifying heart disease cases. While the performance improvement was relatively small, the optimized model exhibited greater consistency and balance. These results emphasize the critical role of hyperparameter tuning in enhancing the predictive capability of neural network models in medical applications. The study contributes to the advancement of more accurate and dependable AI-based tools for early heart disease diagnosis. Future research may benefit from employing advanced optimization strategies such as Bayesian Optimization or Genetic Algorithms and leveraging larger, more diverse datasets to improve generalizability.

Keywords: backpropagation neural network (BPNN); early diagnosis; heart disease prediction; hyperparameter optimization; machine learning

How to Cite: Nita Syahputri, Putrama Alkhairi, and Enok Tuti Alawiah, "Analysis of the Impact of Backpropagation Hyperparameter Optimization on Heart Disease Prediction Models", *J. RESTI (Rekayasa Sist. Teknol. Inf.)*, vol. 9, no. 4, pp. 788 - 795, Aug. 2025.

Permalink/DOI: <https://doi.org/10.29207/resti.v9i4.6473>

Received: March 14, 2025

Accepted: July 23, 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

Heart disease remains a primary contributor to global mortality. As reported by the World Health Organization (WHO), approximately 17.9 million deaths each year are attributed to cardiovascular conditions, including coronary heart disease [1]. In Indonesia, it has been identified by the Ministry of Health as the leading cause of death. The prevalence of heart disease is further intensified by several contributing factors, including poor lifestyle habits, hypertension, diabetes, obesity, and inherited genetic traits [2]-[4]. Therefore, early detection and prediction

of heart disease are essential in reducing complications and improving patient life expectancy [5].

One widely adopted approach for heart disease prediction is Artificial Neural Networks (ANN) [6], which can recognize complex patterns in medical data, thereby enhancing predictive accuracy [7]. In addition to ANN, several other machine learning methods are commonly employed for heart disease detection, including Support Vector Machine (SVM) [8], [9], Random Forest [10], [11], Decision Tree [12], [13], and Logistic Regression [12], [14], which are frequently used in clinical data analysis. Furthermore, deep learning techniques such as Convolutional Neural

Networks (CNN) [15], [16] and Recurrent Neural Networks (RNN) [17], [18] have gained popularity, particularly in processing electrocardiogram (ECG) images and electronic medical records. Other classification-based methods, including k-means clustering and k-Nearest Neighbors (k-NN), can also be utilized in heart disease detection systems. The integration of these techniques enables faster, more accurate, and efficient heart disease prediction, thereby supporting healthcare professionals in making more precise decisions [19]-[21].

A study conducted by [21] demonstrates the effectiveness of Bayesian Optimization in enhancing the performance of the Support Vector Machine (SVM) for early diabetes prediction, achieving a higher accuracy of 95% compared to 94% for the non-optimized SVM model. Additionally, this study utilizes a relevant dataset from Kaggle, providing sufficient data for analysis. However, one of its limitations lies in the lack of a detailed explanation of the Bayesian Optimization process, which may hinder readers unfamiliar with the concept from fully understanding the optimization procedure. Moreover, despite its promising results, the study does not discuss potential limitations or challenges in real-world model implementation.

Similarly, the study conducted by Purwono et al. [22] highlights the advantages of using machine learning algorithms, particularly artificial neural networks, to predict hypertension with a notable accuracy of 85%. Additionally, the research implements hyperparameter optimization using GridSearchCV, demonstrating a more meticulous approach compared to prior studies that primarily focused on accuracy. However, a major limitation of this study is the relatively small dataset, which includes only 219 adult participants, potentially affecting the model's generalizability. While the results are encouraging, the study suggests further optimization of model parameters and the use of larger datasets in future research.

In the medical and technological fields, Artificial Intelligence (AI) has become an effective tool for disease diagnosis and prediction. One widely used approach is Artificial Neural Networks (ANN) [22], [23], particularly the Backpropagation Neural Network (BPNN) [24]. BPNN is a popular machine learning method for medical prediction modeling due to its capability to handle complex and non-linear data. This model operates by adjusting weights through backpropagation, allowing the neural network to enhance its accuracy in predicting medical conditions [25], [26].

Based on the literature review, a key research problem emerges regarding the Backpropagation Neural Network (BPNN). While BPNN is a widely used machine learning method for medical prediction due to its ability to process complex and non-linear data [27],

[28], its performance heavily depends on the selection of optimal hyperparameters, including the number of hidden neurons, learning rate, activation function, and number of epochs. Inappropriate hyperparameter selection may lead to underfitting (inability to recognize patterns within the data) or overfitting (over-adjustment to training data, reducing generalizability to new data) [29], [30]. Consequently, hyperparameter optimization is a crucial step in enhancing the performance of heart disease prediction models.

Several hyperparameter optimization techniques have been developed, such as Grid Search, Random Search, and evolutionary-based optimization algorithms, to determine the optimal parameter combination that improves model accuracy. Through proper hyperparameter optimization, the Backpropagation Neural Network is expected to yield more accurate heart disease predictions, assisting healthcare professionals in making faster and more informed decisions.

Considering these factors, this study aims to analyze the impact of hyperparameter optimization on the performance of the Backpropagation Neural Network in heart disease prediction. Specifically, this research evaluates how variations in certain hyperparameters influence model performance in terms of accuracy, precision, recall, and other evaluation metrics. The findings of this study are expected to contribute to the development of more efficient and accurate AI-based heart disease prediction systems.

2. Methods

This chapter outlines the research methodology used to analyze the impact of hyperparameter optimization on the Backpropagation Neural Network (BPNN) model for heart disease prediction. The methodology covers the dataset description, model design comparisons, and the research stages undertaken.

2.1 Research Dataset

A synthetic dataset was employed in this study to enable the prediction of heart disease risk by integrating various symptoms, lifestyle attributes, and medical background. Each patient is represented by a single row, containing binary indicators (Yes/No) corresponding to relevant symptoms and risk factors, and a designated risk label that classifies the individual as being at either high or low risk for heart disease.

This dataset comprises 70,000 samples, making it well-suited for training machine learning models in classification tasks. The primary objective is to provide a clean and well-structured dataset for researchers, data scientists, and healthcare professionals to explore predictive modeling for cardiovascular health.

The dataset is a side project of EarlyMed, developed by students from Vellore Institute of Technology

(VIT-AP). EarlyMed aims to leverage data science and machine learning for early detection and prevention of chronic diseases. The dataset is publicly available on

Kaggle (Source: Kaggle.com). A sample of the dataset can be seen in Table 1, which provides an overview of the data used in this study.

Table 1. Sample Dataset for Heart Disease Risk Prediction

Chest Pain	Shortness of Breath	Fatigue	Palpitations	Dizziness	etc	Chronic Stress	Gender	Age	Heart Risk
0	0	0	1	0	...	0	0	48	0
0	1	0	1	0	...	0	0	46	0
1	0	0	1	0	...	0	1	66	0
1	1	0	1	0	...	1	1	60	1
0	0	1	0	1	...	0	0	69	0
1	1	0	1	1	...	0	1	55	1
1	1	1	1	1	...	0	1	51	1
1	1	1	1	0	...	1	0	67	1
...	...	...	...	...	...	...	...	...	...
1	1	1	0	0	...	0	0	58	1
1	1	1	1	1	...	1	1	49	1
1	1	1	1	1	...	1	0	51	1
0	0	0	0	0	...	0	1	48	0
1	1	1	1	1	...	0	1	78	1
1	0	1	0	0	...	1	1	48	1
1	0	0	0	1	...	0	0	23	0

The dataset used in this study includes essential features such as age, gender, blood pressure, cholesterol levels, blood sugar levels, heart rate, and other medical examination results. Additionally, it contains a target variable that indicates whether a patient is at risk of developing heart disease. To ensure a fair evaluation of the model's performance, the dataset will be divided into training data and testing data using a specific ratio, such as 80:20 or 70:30.

The dataset consists of two main categories of input features: symptoms and risk factors. The symptom features, represented as binary values (Yes/No), include chest pain, which is a common indicator of heart disease, as well as shortness of breath, which is often linked to cardiac conditions. Other symptoms include unexplained fatigue, palpitations, dizziness or fainting, and swelling in the legs or ankles, which can indicate fluid retention associated with heart failure. Additionally, radiating pain in the arms, jaw, neck, or back may signal angina or a heart attack, while cold sweats and nausea are frequently associated with acute cardiac events.

The risk factors, which can be either binary (Yes/No) or continuous, include age, recorded as a continuous variable, and high blood pressure, high cholesterol levels, and diabetes, which are well-known contributors to cardiovascular diseases. Other significant risk factors include smoking history, obesity, and a sedentary lifestyle, indicating low physical activity levels. Moreover, a family history of heart disease increases the likelihood of developing cardiac conditions, while chronic stress, resulting from prolonged physical, emotional, or psychological strain, may also play a role in increasing cardiovascular risk. Additionally, gender is considered a factor, as heart disease risks may differ between males and females.

The output variable, or heart disease risk label, is a binary classification indicating whether an individual has a low (0) or high (1) risk of heart disease. Before using the dataset for model training, several preprocessing steps will be conducted to improve data

quality and ensure accurate predictions. These steps include handling missing values to maintain data completeness, feature normalization to standardize numerical variables, and feature selection to retain only the most relevant attributes. These preprocessing techniques are essential for optimizing the performance of the predictive model and enhancing its ability to identify individuals at risk of heart disease effectively.

2.2 Proposed Model

This study compares the performance of a standard Backpropagation Neural Network (BPNN) with an optimized BPNN, where specific hyperparameter search techniques are applied. The standard BPNN utilizes manually selected hyperparameters or commonly used default values. In contrast, the optimized BPNN leverages hyperparameter optimization techniques such as Grid Search, Random Search, and Bayesian Optimization to identify the best combination of hyperparameters. Figure 1 presents a comparison between the two Backpropagation Neural Network architectures evaluated in this study.

From Figure 1, it is evident that the Optimized Backpropagation Model uses a smaller number of neurons and the ReLU activation function, which is generally more effective than the Sigmoid function in modern neural networks. Additionally, the Adam Optimizer tends to be more efficient in adjusting the learning rate during training compared to Adagrad. The optimized model also employs a larger batch size (32 vs. 16), which enhances stability and efficiency during the learning process. With the applied optimizations, the Optimized Backpropagation model has a lighter structure with fewer parameters, while still maintaining good accuracy. In contrast, the Standard Backpropagation model has a larger number of parameters, which can lead to higher computational costs and a greater risk of overfitting. The hyperparameters optimized in this study include: (1). Number of neurons in the hidden layer – Determines the model's capacity to learn patterns in the data. (2).

Learning rate – Influences the convergence speed of the model during training. (3). Activation function – Functions like ReLU play a key role in determining the non-linearity of the network. (4). Number of epochs and batch size – Affects the number of iterations and the stability of model training.

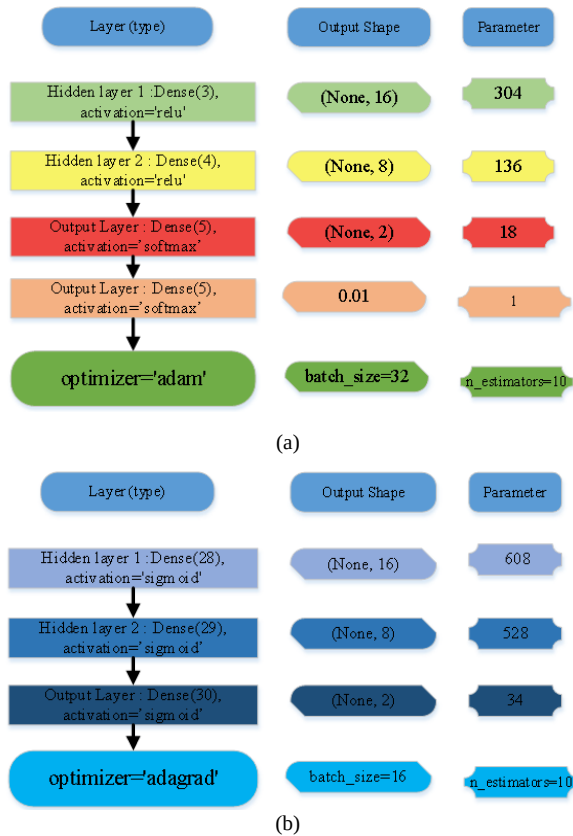


Figure 1. Comparison of Architectures- (a) Optimized Backpropagation; (b) Standard Backpropagation

From the comparison of both models, the Optimized Backpropagation model offers a more efficient approach with fewer parameters and better

performance, thanks to superior hyperparameter selection, including the ReLU activation function, the Adam optimizer, and a larger batch size. On the other hand, the Standard Backpropagation model uses more neurons and parameters, which can increase model complexity without a guaranteed significant improvement in accuracy. Therefore, hyperparameter optimization plays a crucial role in enhancing the performance of heart disease prediction models.

2.3. Research Design

The research framework in this study plays a crucial role in providing a systematic guide for each stage of the research process. The research design is structured to compare the performance of the standard Backpropagation Neural Network (BPNN) model with the hyperparameter-optimized model in heart disease prediction.

The study begins with data collection and preprocessing stages, which include normalization, handling missing values, and selecting relevant features. Following this, the standard BPNN model was developed using default parameters, after which an optimized model was created using techniques such as Grid Search, Random Search, or Bayesian Optimization to identify the best combination of hyperparameters.

Both models are then trained and tested using the same dataset, and the results are compared based on evaluation metrics such as accuracy, precision, recall, F1-score, Mean Squared Error (MSE), and Mean Absolute Error (MAE). Finally, an analysis of the results is conducted to assess how much hyperparameter optimization improves model performance, with the goal of providing recommendations for the development of more accurate and efficient heart disease prediction systems. The research design is illustrated in Figure 2.

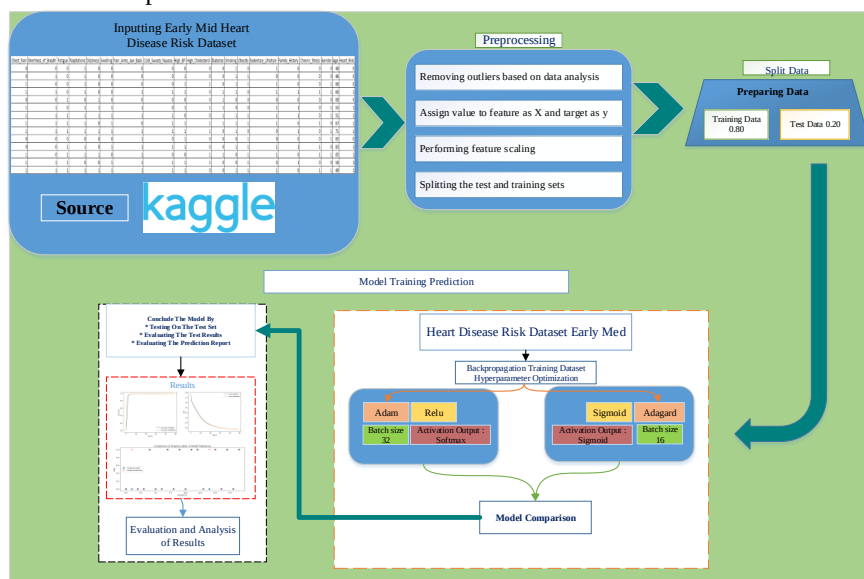


Figure 2. Research Design

3. Results and Discussions

This chapter discusses the results of the study, starting from the data preprocessing process through to the training phase of the heart disease prediction model. The objective of this process is to ensure that the data used is ready for the Backpropagation Neural Network (BPNN) model and to evaluate the model's performance both before and after hyperparameter optimization.

3.1 Data Preprocessing and Data Augmentation

The data preprocessing process is a crucial step in this study to ensure that the dataset used is clean, relevant, and ready for model training. The first step in preprocessing is the missing data check, where missing values are imputed using appropriate methods, such as mean, median, or mode, depending on the type of data involved.

	Chest_Pain	Shortness_of_Breath	Fatigue	Palpitations	Dizziness	\
0	0.0	0.0	0.0	1.0	0.0	
1	0.0	1.0	0.0	1.0	0.0	
2	1.0	0.0	0.0	1.0	0.0	
3	1.0	1.0	0.0	1.0	0.0	
4	0.0	0.0	1.0	0.0	1.0	

	Swelling	Pain_Arms_Jaw_Back	Cold_Sweats_Nausea	High_BP	\
0	0.0	0.0	0.0	0.0	
1	0.0	0.0	0.0	1.0	
2	0.0	0.0	0.0	1.0	
3	0.0	1.0	1.0	1.0	
4	0.0	0.0	0.0	0.0	

	High_Cholesterol	Diabetes	Smoking	Obesity	Sedentary_Lifestyle	\
0	0.0	0.0	1.0	0.0	1.0	
1	0.0	0.0	1.0	1.0	0.0	
2	1.0	0.0	1.0	1.0	1.0	
3	0.0	1.0	1.0	0.0	1.0	
4	0.0	1.0	0.0	0.0	0.0	

	Family_History	Chronic_Stress	Gender	Age	Heart_Risk
0	0.0	0.0	0.0	48.0	0.0
1	0.0	0.0	0.0	46.0	0.0
2	0.0	0.0	1.0	66.0	0.0
3	1.0	1.0	1.0	60.0	1.0
4	0.0	0.0	0.0	69.0	0.0

Figure 3. Sample Data After Data Preprocessing

Figure 3 shows the result after the data normalization process, which standardizes the scale of each feature to enable the model to learn more effectively. This normalization is carried out using methods such as Min-Max Scaling or Standardization. Additionally, feature selection is performed to choose the attributes most influential in predicting heart disease, such as blood pressure, cholesterol levels, and heart rate. With proper preprocessing, the model can operate more optimally, reducing the risk of bias or noise in the data.

3.2 Prediction Model Training

In the model training stage, two Backpropagation Neural Network (BPNN) models are trained: the standard model and the hyperparameter-optimized model. The standard model uses default parameters, such as a higher number of neurons, the Sigmoid activation function, and the Adagrad optimizer, while the optimized model uses fewer neurons, the ReLU activation function, and the Adam optimizer to improve efficiency and accuracy. The training process

is carried out by splitting the dataset into training data and testing data with an 80:20 ratio.

After the training process is complete, evaluation is performed using metrics such as accuracy, precision, recall, F1-score, and Mean Squared Error (MSE) to compare the performance of both models. The experimental results show that the optimized model achieves higher accuracy and lower error compared to the standard model. This indicates that hyperparameter optimization plays a significant role in improving the performance of heart disease prediction models, allowing the model to provide more accurate results that can be used as a tool for early heart disease diagnosis. The comparison of the highest accuracy can be seen in Figure 4.

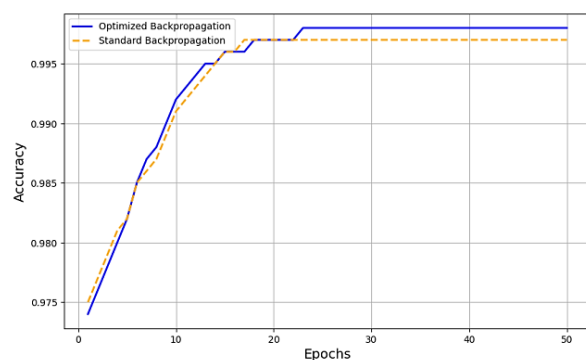


Figure 4. the accuracy comparison between the Optimized Backpropagation and Standard Backpropagation models

The comparison between the Optimized Backpropagation and Standard Backpropagation models (Figure 4) shows that the optimized model achieves faster and higher accuracy. In the early epochs (1 to 10), the optimized model shows a sharp increase in accuracy, quickly reaching around 0.995 and stabilizing thereafter. This faster learning is due to optimized hyperparameters like the ReLU activation function and Adam optimizer, which help the model converge more quickly.

On the other hand, the Standard Backpropagation model improves more slowly, with a steady increase in accuracy, reaching its peak accuracy of about 0.992 after more epochs (around the 30th). This slower progress suggests that the default hyperparameters (Sigmoid activation and Adagrad optimizer) are less efficient in training.

Both models stabilize after around 35 epochs, but the Optimized Backpropagation model consistently achieves higher accuracy, reaching 0.995-0.996 compared to the 0.992 accuracy of the standard model. Although the difference is small, it highlights the effectiveness of hyperparameter optimization in improving model performance.

In conclusion, the Optimized Backpropagation model is not only more accurate but also more efficient, requiring fewer epochs to achieve high performance.

This shows the importance of hyperparameter tuning in improving the efficiency and accuracy of machine learning models.

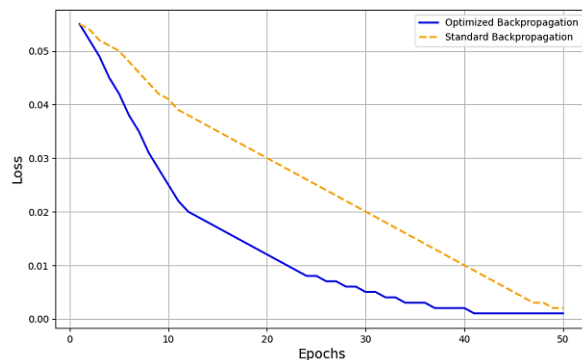


Figure 5. the loss comparison between the Optimized Backpropagation and Standard Backpropagation models

The comparison of loss between the Optimized Backpropagation and Standard Backpropagation models (Figure 5) shows that the optimized model reduces loss much faster. In the early epochs (1 to 10), the optimized model drops its loss from 0.055 to about 0.02, indicating more efficient learning due to optimized hyperparameters like the ReLU activation function and Adam optimizer. This model then stabilizes at a low loss of around 0.005. In contrast, the Standard Backpropagation model decreases its loss more gradually, starting from 0.055 and reaching 0.03 after the first 10 epochs. It continues to decrease slowly, stabilizing at around 0.02. This slower progress suggests that the default hyperparameters, like the Sigmoid activation function and Adagrad optimizer, are less efficient.

By the end of the training, both models stabilize, but the Optimized Backpropagation model achieves a significantly lower final loss (0.005) compared to the Standard model (0.02). This clearly shows that hyperparameter optimization leads to faster and more effective learning, making the optimized model more efficient and accurate.

3.3 Discussion

Here is the visualization of the evaluation results, comparing the performance of the Optimized Backpropagation and Standard Backpropagation models across four evaluation metrics.

The evaluation results comparing the Optimized Backpropagation and Standard Backpropagation

models reveal several key insights across four important metrics: Accuracy, Precision, Recall, and Mean Squared Error (MSE) shown in Figure 6. The Optimized Backpropagation model achieves slightly higher Accuracy (0.9911) compared to the Standard Backpropagation model (0.9909), indicating that the optimized hyperparameters, such as the ReLU activation function and Adam optimizer, help the model make more accurate predictions. Additionally, the optimized model demonstrates a higher Precision (0.9903) than the standard model (0.9879), meaning it is better at correctly identifying patients who have heart disease, reducing the number of false positives.

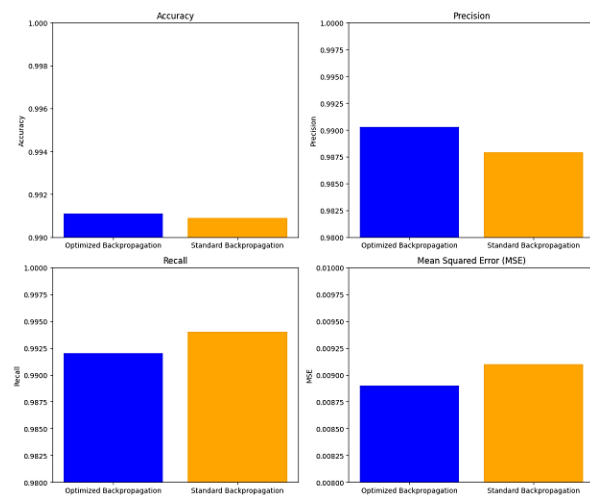


Figure 6. The visualization of the evaluation results Accuracy, Precision, Recall, and Mean Squared Error (MSE)

In terms of Recall, the Standard Backpropagation model performs marginally better, with a value of 0.994, compared to 0.992 for the Optimized Backpropagation model. Recall is crucial for identifying true positive cases, and while the standard model performs slightly better, the difference is minimal. Both models are highly effective at detecting most of the actual heart disease cases.

When considering Mean Squared Error (MSE), the Optimized Backpropagation model outperforms the standard model with a lower MSE of 0.0089, compared to 0.0091 for the standard model. A lower MSE indicates that the optimized model's predictions are closer to the true values, reflecting greater accuracy in its predictions. Here is the visualization of the evaluation results shown in Table 2.

Table 2. Evaluation Results

Model	Accuracy	Precision	Recall	MSE	MAE
Optimized Backpropagation	0.9911	0.9903	0.992	0.0089	0.0089
Standard Backpropagation	0.9909	0.9879	0.994	0.0091	0.0091

The comparison of Mean Absolute Error (MAE) between the Optimized Backpropagation and Standard Backpropagation models shows that the Optimized Backpropagation model performs slightly better (Figure 7- Refer Table 2).

It has a lower MAE of 0.0089 compared to the Standard Backpropagation model, which has an MAE of 0.0091. MAE measures the average magnitude of the errors in the model's predictions, with a lower value indicating more accurate predictions. The results

suggest that the Optimized Backpropagation model is more reliable, as it produces predictions that are closer to the actual values, making it the preferred choice for more accurate heart disease prediction.

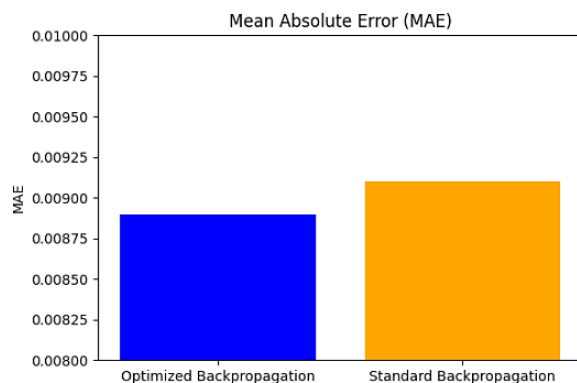


Figure 7. The visualization of the evaluation results Mean Absolute Error (MAE)

4. Conclusions

This study compares the effectiveness of two neural network models—an Optimized Backpropagation Neural Network (BPNN) and a Standard BPNN—for heart disease prediction. The results show that the Optimized BPNN, which uses hyperparameter tuning, performs better across key evaluation metrics such as Accuracy, Precision, Recall, Mean Squared Error (MSE), and Mean Absolute Error (MAE).

The Optimized BPNN achieved higher Accuracy (0.9911) and Precision (0.9903), along with lower MSE (0.0089) and MAE (0.0089), suggesting it produces more reliable and accurate predictions. Although the Standard BPNN recorded a slightly higher Recall (0.994), the difference was marginal. Overall, the Optimized BPNN delivered a more balanced performance by minimizing errors and reducing false positives. Its ability to lower both MSE and MAE further supports its effectiveness for heart disease detection, where accuracy and reliability are essential.

These findings highlight the importance of hyperparameter optimization in enhancing neural network performance, especially in healthcare applications. By adjusting parameters like learning rate, activation function, and optimizer, the optimized model offers a more efficient and robust method for early disease detection. This study contributes to the growing body of research applying artificial intelligence in medical diagnostics and supports the development of more advanced, data-driven healthcare tools.

Future work may involve testing other optimization methods such as Bayesian Optimization or Genetic Algorithms to further improve model accuracy and generalization. Additionally, using a larger and more diverse dataset could increase the model's real-world

relevance, making it more effective for predicting not only heart disease but other medical conditions as well.

References

- [1] A. Jasinska-Piadlo et al., "Data-driven versus a domain-led approach to k-means clustering on an open heart failure dataset," *International Journal of Data Science and Analytics*, vol. 15, no. 1, pp. 49–66, 2023, doi: 10.1007/s41060-022-00346-9.
- [2] P. R. Kumar, "Identification of noteworthy features and data mining techniques for heart disease prediction," *International Journal of Modeling, Simulation, and Scientific Computing*, 2024, doi: 10.1142/S1793962325500102.
- [3] T. R. Ramesh, "PREDICTIVE ANALYSIS OF HEART DISEASES WITH MACHINE LEARNING APPROACHES," *Malaysian Journal of Computer Science*, vol. 2022, pp. 132–148, 2022, doi: 10.22452/mjcs.sp2022no1.10.
- [4] Saiyedfaiazwaris and S. Koteeswaran, "Early Prediction of Heart Conditions by K-Means Consensus Clustering and Convolution Neural Network," vol. 25, no. 3, pp. 6623–6640, 2021.
- [5] N. A. M. Zaini and M. K. Awang, "Performance Comparison between Meta-classifier Algorithms for Heart Disease Classification," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 10, pp. 323–328, 2022, doi: 10.14569/IJACSA.2022.0131039.
- [6] M. Agarwal, "Methodical Implementation of Data Mining Classifiers and ANN for Prediction of Accomplishment of Student Education," *International Journal of Performability Engineering*, vol. 19, no. 9, pp. 587–597, 2023, doi: 10.23940/ijpe.23.09.p4.587597.
- [7] A. P. Windarto, T. Herawan, and P. Alkhairi, "Early Detection of Breast Cancer Based on Patient Symptom Data Using Naive Bayes Algorithm on Genomic Data," in *Artificial Intelligence, Data Science and Applications*, Y. Farhaoui, A. Hussain, T. Saba, H. Taherdoost, and A. Verma, Eds., Cham: Springer Nature Switzerland, 2024, pp. 478–484.
- [8] P. V Naresh, "PSO Optimized CNN-SVM Architecture for Covid - 19 Classification," *International Journal on Recent and Innovation Trends in Computing and Communication*, vol. 11, pp. 205–209, 2023, doi: 10.17762/ijritcc.v11i5s.6646.
- [9] B. S. Reddy, "An optimal superpixel segmentation based transfer learning using AlexNet-SVM model for weed detection," *International Journal of System Assurance Engineering and Management*, 2023, doi: 10.1007/s13198-023-02064-2.
- [10] A. E. K. Gunawan, "Stock Price Movement Classification Using Ensembled Model of Long Short-Term Memory (LSTM) and Random Forest (RF)," *International Journal on Informatics Visualization*, vol. 7, no. 4, pp. 2255–2262, 2023, doi: 10.30630/joiv.7.4.1640.
- [11] T. Kavzoglu, "Predictive Performances of Ensemble Machine Learning Algorithms in Landslide Susceptibility Mapping Using Random Forest, Extreme Gradient Boosting (XGBoost) and Natural Gradient Boosting (NGBoost)," *Arabian Journal for Science and Engineering*, vol. 47, no. 6, pp. 7367–7385, 2022, doi: 10.1007/s13369-022-06560-8.
- [12] E. Dumitrescu, "Machine learning for credit scoring: Improving logistic regression with non-linear decision-tree effects," *European Journal of Operational Research*, vol. 297, no. 3, pp. 1178–1192, 2022, doi: 10.1016/j.ejor.2021.06.053.
- [13] A. H. Wibowo and T. I. Oesman, "The comparative analysis on the accuracy of k-NN, Naive Bayes, and Decision Tree Algorithms in predicting crimes and criminal actions in Sleman Regency," *Journal of Physics: Conference Series*, vol. 1450, no. 1, 2020, doi: 10.1088/1742-6596/1450/1/012076.
- [14] P. Golpour et al., "Comparison of support vector machine, naïve bayes and logistic regression for assessing the necessity

- for coronary angiography,” *International Journal of Environmental Research and Public Health*, vol. 17, no. 18, pp. 1–9, 2020, doi: 10.3390/ijerph17186449.
- [15] Y. Yuhandri, A. P. Windarto, and M. N. H. Siregar, “Improving Brain Tumor Classification Efficacy through the Application of Feature Selection and Ensemble Classifiers,” *Journal of Image and Graphics(United Kingdom)*, vol. 11, no. 4, pp. 397–404, 2023, doi: 10.18178/joig.11.4.397-404.
- [16] A. P. Windarto, Y. Yuhandri, and S. Bukhori, “a Novel Convective Cloud Segmentation for Heavy Rainfall Potential Assessment Using Deep Learning,” *ICIC Express Letters*, vol. 18, no. 4, pp. 313–323, 2024, doi: 10.24507/icicel.18.04.313.
- [17] P. R. Shekar, A. Mathew, and K. V. Sharma, “A hybrid CNN–RNN model for rainfall–runoff modeling in the Potteruvagu watershed of India,” *CLEAN – Soil, Air, Water*, vol. n/a, no. n/a, p. 2300341, 2024, doi: <https://doi.org/10.1002/clen.202300341>.
- [18] Enung, “Hourly Discharge Prediction Using Long Short-Term Memory Recurrent Neural Network (Lstm-Rnn) In The Upper Citarum River,” *International Journal of GEOMATE*, vol. 23, no. 98, pp. 147–154, 2022, doi: 10.21660/2022.98.3462.
- [19] Z. A. Sejuti and M. S. Islam, “A hybrid CNN–KNN approach for identification of COVID-19 with 5-fold cross validation,” *Sensors International*, vol. 4, no. November 2022, p. 100229, 2023, doi: 10.1016/j.sintl.2023.100229.
- [20] R. Kadry and O. Ismael, “A New Hybrid KNN Classification Approach based on Particle Swarm Optimization,” *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 11, pp. 291–296, 2020, doi: 10.14569/IJACSA.2020.0111137.
- [21] W. Sinhashthita, “Improving KNN Algorithm Based on Weighted Attributes by Pearson Correlation Coefficient and PSO Fine Tuning,” 2020.
- [22] C. Karapataki and J. Adamowski, “Comparison of Multivariate Regression and Artificial Neural Networks for Peak Urban Water-Demand Forecasting: Evaluation of Different ANN Learning Algorithms,” *Journal of Hydrologic Engineering*, vol. 17, no. 7, pp. 834–836, 2012, doi: 10.1061/(asce)he.1943-5584.0000472.
- [23] S. Kamada, T. Ichimura, A. Hara, and K. J. Mackin, “Adaptive structure learning method of deep belief network using neuron generation – annihilation and layer generation,” *Neural Computing and Applications*, vol. 0123456789, 2018, doi: 10.1007/s00521-018-3622-y.
- [24] P. Alkhairi, W. Wanayumini, and B. H. Hayadi, “Analysis of the adaptive learning rate and momentum effects on prediction problems in increasing the training time of the backpropagation algorithm,” *AIP Conference Proceedings*, vol. 3048, no. 1, p. 20049, 2024, doi: 10.1063/5.0203374.
- [25] N. A. Setyadi, M. Nasrun, and C. Setianingsih, “Text Analysis for Hate Speech Detection Using Backpropagation Neural Network,” *Proceedings - 2018 International Conference on Control, Electronics, Renewable Energy and Communications, ICCEREC 2018*, pp. 159–165, 2018, doi: 10.1109/ICCEREC.2018.8712109.
- [26] L. Chen, “Intralayer-Connected Spiking Neural Network with Hybrid Training Using Backpropagation and Probabilistic Spike-Timing Dependent Plasticity,” *International Journal of Intelligent Systems*, vol. 2023, 2023, doi: 10.1155/2023/3135668.
- [27] A. P. Windarto, P. Alkhairi, and J. Muslim, “Enhancing Premier League Match Outcome Prediction Using Support Vector Machine with Ensemble Techniques: A Comparative Study on Bagging and Boosting,” *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)(Rekayasa Sistem dan Teknologi Informasi)*, vol. 9, no. 1, pp. 94–103, 2024, doi: <https://doi.org/10.29207/resti.v9i1.6173>.
- [28] N. Razali, A. Mustapha, N. Arbaiy, and P. C. Lin, “Deep Learning for Football Outcomes Prediction based on Football Rating System,” *AIP Conference Proceedings*, vol. 2644, no. November, 2022, doi: 10.1063/5.0104587.
- [29] M. Qu, “An abnormal behaviour recognition of MOOC online learning based on multidimensional data mining,” *International Journal of Autonomous and Adaptive Communications Systems*, vol. 17, no. 4, pp. 369–382, 2024, doi: 10.1504/IJAACS.2024.139403.
- [30] W. N. Ismail, H. A. Alsalamah, M. M. Hassan, and E. Mohamed, “AUTO-HAR: An adaptive human activity recognition framework using an automated CNN architecture design,” *Heliyon*, vol. 9, no. 2, p. e13636, 2023, doi: 10.1016/j.heliyon.2023.e13636.