International Journal of Scientific Research in Computer Science, Engineering and Information Technology

ISSN : 2456-3307 OPEN ACCESS

Available Online at : www.ijsrcseit.com

doi : https://doi.org/10.32628/CSEIT241011

The Comparative Early Prediction Model for Cardiovascular Disease Using Machine Learning

Sri Sumarlinda*[1,2], Azizah binti Rahmat[1], Zalizah binti Awang Long[1], Wiji Lestari[2]

*[1] Malaysian Institute Information Technology, Kuala Lumpur, Malaysian

[2] Information System, Universitas Duta Bangsa Surakarta, Surakarta, Indonesian

ARTICLEINFO

Article History:

Accepted: 15 Dec 2023

Published: 07 Jan 2024

Publication Issue

Volume 10, Issue 1

January-February-2024

Page Number

24-33

ABSTRACT

Cardiovascular disease (CVD) is a leading cause of death and a major contributor to disability. Early detection of cardiovascular disease using ANFIS has the potential to reduce costs and simplify treatment. This study aims to develop a prediction model using ANFIS (Adaptive Neuro-Fuzzy Inference System) for early detection of cardiovascular disease. The dataset used consists of 500 data with 12 features, including various risk factors such as blood sugar levels, cholesterol, uric acid, systolic blood pressure, diastolic blood pressure, body mass index (BMI), age, smoking habits, lifestyle, genetic factors, and gender, and one label feature. This study compares cardiovascular disease prediction models using machine learning methods, namely Support Vector Machine (SVM), K-Nearest Neighbor (K-NN), and ANFIS. The development of the KNN algorithm involves the value of K=5 with the Euclidian distance measure. The SVM algorithm used a kernel cache of 200 and a convergence epsilon of 0.001. The ANFIS model was built using 500 data sets divided into training (70%) and testing (30%) data, with learning rate variations of 0.01, 0.05, 0.1, 0.2, and 0.5. The results of testing the early detection model show for SVM, the accuracy value is 0.760, the precision value is 0.839, and the recall value is 0.671. For the KNN model, the accuracy value is 0.758, the precision value is 0.768, and the recall value is 0.771. As for the ANFIS model, the accuracy value reaches 0.989, precision value 0.996, and recall value 0.988. The model using ANFIS has the highest performance. Further study of the model using ANFIS with learning rate variations shows that a learning rate of 0.1 provides the most optimal performance.

Keywords: Prediction Model, Cardiovascular Disease, Support Vector Machine (SVM), K-Nearest Neighbor (K-NN), Adaptive Neuro-Fuzzy Inference System (ANFIS).

I. INTRODUCTION

Cardiovascular disease (CVD) is a leading cause of death and a major contributor to disability. China and India globally have the highest rates of CVD cases[1]. Risk assessment plays a crucial role in reducing the global burden of CVD [2]. According to the WHO, heart disease ranks highest among the ten leading causes of death [3]. Global statistics show that cardiovascular disease dominates as the leading cause of death[4]. CVD is not only the leading cause of death worldwide but also contributes to health system expenditure [5]. In 2014, stroke became the fourth leading cause of death in Japan, with cerebral infarction as the leading cause [6]. Early detection of cardiovascular disease using ANFIS has the potential to reduce costs and simplify treatment [7].

Model comparison has been done by applying several machine learning methods or algorithms for early detection of cardiovascular disease. KNN and SVM algorithms, for example, can be compared for early detection of cardiovascular disease [8]. Extra trees, random florets, adaBoost and gradients boosting algorithms are used for early prediction of cardiovascular disease [9]. Comparative model with hyper parameter optimization used for early detection of cardiovascular disease using KNN, SVM, SGD, Random florests, neural network, naïve bayes, logistic regression, and adaBoost [10].

Adaptive Neuro-Fuzzy Inference System (ANFIS) is an innovative method that combines the intelligence of artificial neural networks (ANN) and the flexibility of fuzzy inference systems (FIS). ANFIS was developed to address challenges in predictive modelling. ANFIS combines the advantages of ANNs in handling nonlinear complexity with the advantages of fuzzy inference systems in dealing with uncertainty and decision complexity. The ANFIS development process involves data separation, classification, fuzzy rule formation, inference, defuzzification, model training, validation, testing, and optimization. ANFIS effectively solves complex nonlinear situations by integrating neural networks and fuzzy inference [11].

Support Vector Machine (SVM) is a machine learning method that can be used to create early prediction models in cardiovascular disease. SVM operates by finding the optimal hyperplane to separate two classes of data, where in the context of cardiovascular disease early prediction models, the classes include high-risk patients and low-risk (healthy) patients. The success of the SVM model in early cardiovascular disease prediction is strongly influenced by the data representation as well as the setting of SVM parameters [8] [9].

The K-Nearest Neighbours (KNN) algorithm is a machine learning method that can be used for early detection of cardiovascular disease. KNN works by classifying new data based on the majority of classes in KNN, where the data will be attributed to the most dominant class among its K nearest neighbours. KNN can be a good choice, especially if the dataset is relatively small and there is no strong linear assumption between feature relationships. As with other models, data representation plays an important role in improving the performance of the KNN algorithm [8], [10] .

This research focuses on building comparative models using SVM, KNN and ANFIS. The purpose of this research is to compare and find parameters that produce optimal values in early detection. Performance testing uses RMSE, accuracy and F1-Score. From the performance test, the three models will be compared to find a better model.

II. METHODS AND MATERIAL

This session will discuss the materials and algorithms that will be used and research methods.

A. Data Set

The 500 data obtained from X Surakarta Hospital involved information on patients with heart disease and stroke. This data included patient ID, blood sugar, cholesterol, uric acid, systolic and diastolic BMI, age, smoking habits, physical activity, lifestyle, genetic factors, and gender.

B. Support Vector Machine (SVM)

Support Vector Machine (SVM) is a group of supervised learning methods used for classification, regression, and outlier detection. The main advantages of SVMs lie in their effectiveness in high-dimensional spaces and the utilization of a subset of training points in the decision function, known as support vectors. SVMs form a hyperplane in an infinite dimensional space, which can be used for both regression and classification purposes [10] , [12] [13], [14].

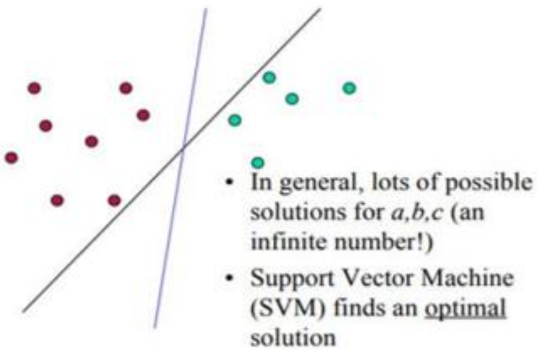

Figure 1. Hyperplane in SVM

C. K-Nearest Neighbor (K-NN)

The K-Nearest Neighbor (KNN) algorithm is a supervised learning algorithm that is often used for prediction and classification . The main advantages of the KNN algorithm are its high accuracy rate, its good ability to deal with outliers, and the absence of specific assumptions regarding the data. Determining the K value is a crucial factor in the implementation of this algorithm. To measure the similarity of data with labels, this algorithm uses the Euclidean distance with a certain formula [8], [15].

$$y(x, y) = \sqrt{\sum_{i}^{n} (x_1 - y_1)^2} \quad (1)$$

Where:
d(x,y) distance of x and y data
xi is the first training data
yi is the first testing data

The stages of the KNN algorithm are:
- Determine the value of k
- Calculate the distance with Euclidian distance
- Sort the training data
- Determine the class

D. Adaptive Neuro Fuzzy Inference System (ANFIS)

Sugeno rule-based fuzzy inference model and Adaptive Neuro Fuzzy Inference System (ANFIS) model can be compared in terms of functionality. Functionally, ANFIS has an identical design to the radial function artificial neural network, except for a few exceptions. In addition, the rules in ANFIS are flexible. For a radial function network to be comparable to a first-order Sugeno model based on fuzzy rules, several requirements must be met. To generate all outputs, the rules must apply a uniform aggregation technique, such as weighted average or weighted sum [15][16], [17]. The number of fuzzy rules and activation functions must be equal. Each activation function requires a membership function for each input, especially when the rule base has more than one entry. Consistency between the fuzzy rules and the activation function needs to be maintained with the rules and neurons on the output side.

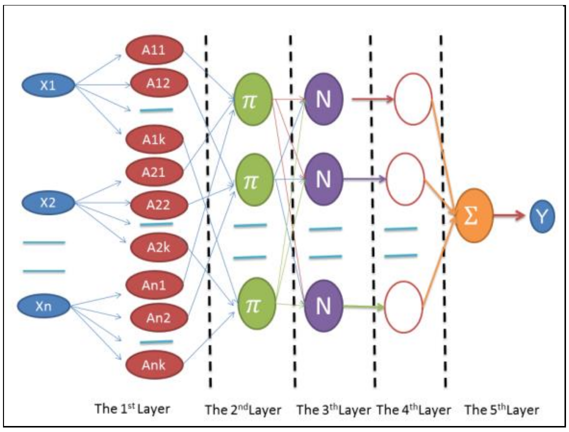

Figure 2. Architecture ANFIS

Suppose there are 2 inputs x1, x2 and output y. There are 2 rules in the rule base of Sugeno model:

If x1 is A1 and x2 is B1 Then y1 = C11x1 + C12x1 + C10
If x1 is A1 and x2 is B2 Then y2 = C21x2 + C22x2 + C20

If the rule predicates are w1 and w2, then the weighted average can be calculated as:

$$y = \frac{w_1 y_1 + w_2 y_2}{w_1 + w_2} = \overline{w_1} y_1 + \overline{w_2} y_2 \qquad (2)$$

ANFIS network consists of layers [16], [17]. The layers are:

1. The 1st layers

Each neuron in the first layer is adaptive to the parameters of the activation function. The output of each neuron is the degree of membership given by the input membership function, namely:

$$\mu(x) = \frac{1}{1 + \left|\frac{x-c}{a}\right|^{2b}} \qquad (3)$$

Where a, b and c are parameters called premise parameters.

2. The 2nd Layer

Each neuron in the second layer is a fixed neuron whose output is derived from the input. The AND operator is usually used. Each node represents a rule predicate in the system.

$$w_k = \prod_1^n \mu_{nk} = \mu_{A1k} . \mu_{A2k} . \mu_{3k} \dots \mu_{nk} \qquad (4)$$

3. The 3th Layer

Each neuron in the third layer is a fixed node that is the result of calculating the ratio of predicates (w) from the rule to the total number of predicates.

$$\overline{w}_k = \frac{w_k}{\sum_1^k w_k} = \frac{w_k}{w_1 + w_2 + w_3 + \dots + w_k} \qquad (5)$$

4. The 4th Layer

Each neuron in the fourth layer is adaptive to an output.

$$\overline{w}_i . y_i = \overline{w}_i (c_{i1} . x_1 + c_{i2} . x_2 + c_{i3} . x_3 + \dots + c_{in} . x_n + c_{i0} \qquad (6)$$

Where $\overline{w}_i$ is the normalized firing power at the third layer and cij is the parameter of the neuron. The parameters on the neuron are called consequence parameters

5. The 5th Layer

Each neuron in the fifth layer is a fixed node that is the sum of all inputs.

E. Research Methods

The purpose of this study is to compare models and find parameters that produce optimal values in early detection. stages in this study, which are described in Figure 3:

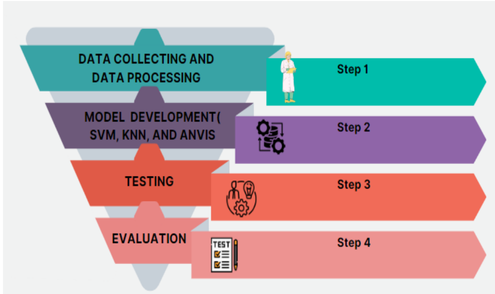

Figure 3. Research Method Diagram

This research goes through four stages, namely data collecting and data processing, SVM, K-NN, and ANFIS model development, testing and evaluation, testing and evaluation.

1. Data Colleting and Data Processing

The study involved collecting data from a number of clinics, hospitals and respondents. After obtaining the data, a tracking and cleaning process was conducted to address missing or incomplete data. Then, the data was adjusted and improved during the data processing stage, including the percentage split between training and testing data. In total, 500 data points were collected for this study.

2. Model Development

Model development in this study uses three algorithms namely SVM, KNN and ANFIS. KNN algorithm development uses the value of K = 5 with the size of the Euclidian distance. SVM algorithm used with kernel cache 200 and epsilon convergence 0.001. ANFIS built the model using 500 data sets divided into training (70%) and testing (30%) data. A Gaussian membership function was applied, and ANFIS development was performed using Python software. The training process involved 1000 epochs, with learning rate variations of 0.01, 0.05, 0.1, 0.2, and 0.5.

3. Testing & Evaluation Method

This research uses a combination of appropriate evaluation metrics to assess the performance of the ANFIS model, which include prediction error with Root Mean Square Error (RMSE), accuracy, and F1-Score.

a. Root Mean Square Error

Root Mean Square Error (RMSE) is an evaluation metric that measures the extent of the difference between the predicted value and the actual value, as described earlier [15]. The RMSE formula:

$$RMSE = \sqrt{\frac{\sum_1^N (y_i - y_{out})^2}{N}} \quad (7)$$

N is number of datasets; Yi is the real value or label form dataset and Y out the prediction value form ANFIS process.

b. Accuracy

Accuracy is an evaluation metric that measures the extent to which a model can predict correctly [18]. The accuracy formula:

$$Accuracy = \left(\frac{TP+TN)}{N}\right) \quad (8)$$

c. Precision

Precision is one of the model performance evaluation metrics in the context of classification. Precision measures the extent to which the positive predictions made by the model are correct or relevant [19].

$$Precision = \left(\frac{TP)}{TP+FP}\right) \quad (9)$$

d. Recall

Recall is a classification model performance evaluation metric that measures the extent to which the model is able to capture or detect all true positive cases [20].

$$Recall = \left(\frac{TP)}{TP+FN}\right) \quad (10)$$

e. F1-Score

F1-Score is a metric that combines precision and recall

$$F1 - score = 2\ x \left(\frac{Precision\ x\ Recall)}{Precesion+Recall}\right) \quad (11)$$

Where:

TP : True Positive
TN : True Negative
N : Number of data
FP : False Positive
FN : False Negative

III.RESULTS AND DISCUSSION

Model implementation was carried out using three algorithms, namely SVM, KNN, and ANFIS. Model development using ANFIS involves more parameter variations. The ANFIS model was built by dividing the data into training and testing data in a 70%:30% ratio. The training dataset consists of 70% of the total 500 datasets, which is 350 data, while the testing dataset covers 30% of the total 500 datasets, with a total of 150 data.

A. Result 1

The results of the model with 1000 epochs and learning rate variations can be found in Table 1 to Table 5.

Table 1. The result of model with epochs 1000 and

MODEL (Epochs 1000, LR = 0.01)					
DATA SET	RMSE	ACCURACY	F1-SCORE	PRECISION	RECALL
Training Data	0.055	0.980	0.986	0.985	0.988
Testing Data	0.288	0.913	0.940	0.953	0.927

The table displays the results of calculating the accuracy, F1-score, precision and recall values on the training data and testing dataset, which shows that the accuracy, F1-score, precision and recall values on the testing data are smaller than the values on the training data. For the RMSE value on the training data which is smaller than the RMSE value on the testing data.

Table 2. The result of model with epochs 1000 and learning rate 0.05

MODEL (Epochs 1000, LR = 0.05)					
DATA SET	RMSE	ACCURACY	F1-SCORE	PRECISION	RECALL
Training Data	0.053	0.989	0.992	0.996	0.988
Testing Data	0.330	0.893	0.925	0.961	0.891

The table shows the results of model evaluation on training and testing data. If the accuracy, F1-Score, precision, and recall values on the training data are higher than those on the testing data, this may indicate overfitting.

Table 3. The result of model with epochs 1000 and learning rate 0.1

MODEL (Epochs 1000, LR = 0.1)					
DATA SET	RMSE	ACCURACY	F1-SCORE	PRECISION	RECALL
Training Data	0.051	0.986	0.990	0.996	0.984
Testing Data	0.537	0.900	0.929	0/970297	0.891

Table 3 shows that the values of RMSE, accuracy, F1-Score, precision and recall with a learning rate of 0.1 show that the training data is higher than the value of the testing data.

Table 4. The result of model with epochs 1000 and learning rate 0.2

MODEL (Epochs 1000, LR = 0.2)					
DATA SET	RMSE	ACCURACY	F1-SCORE	PRECISION	RECALL
Training Data	0.115	0.960	0.973	0.977	0.969
Testing Data	0.332	0.880	0.917	0.934	0.900

From Table 4, it can be seen that the values of RMSE, accuracy, F1-Score, precision, and recall with a learning rate of 0.2 show that the performance of the model on the training data is higher than the values found on the testing data.

Table 5. The result of model with epochs 1000 and learning rate 0.2

	MODEL (Epochs 1000, LR = 0.5)				
DATA SET	RMSE	ACCURACY	F1-SCORE	PRECISION	RECALL
Training Data	0.149	0.943	0.961	0.961	0.961
Testing Data	0.255	0.893	0.927	0.935	0.918

From Table 5, it can be seen that the values of RMSE, accuracy, F1-Score, precision, and recall with a learning rate of 0.5 show that the performance of the model on the training data is higher than the values found on the test data.

The results of the prediction model implementation with 1000 epochs and various learning rates from table 1 to 5 are compared in table 6.

Table 6. Comparison performance of ANFIS using learning rate with value 0.01, 0.05, 0.1, 0.2 and 0.5

	MODEL					
	TRAINING DATA			TESTING DATA		
LEARNING RATE	RMSE	ACCURACY	F1-SCORE	RMSE	ACCURACY	F1-SCORE
0.01	0.055	0.980	0.986	0.288	0.913	0.940
0.05	0.053	0.989	0.925	0.330	0.893	0.992
0.1	0.051	0.986	0.990	0.537	0.900	0.929
0.2	0.115	0.960	0.973	0.332	0.880	0.917
0.5	0.149	0.943	0.961	0.255	0.893	0.927

To prove that our model is better, we conducted a comparison test with other machine learning algorithms. It can be seen in the table and graph below:

Table 7. Comparison performance and accuracy

Machine Learning	Accuracy	Precision	Recall
ANFIS	0.989	0.996	0.988
SVM	0.760	0.839	0.671
KNN	0.758	0.768	0.771

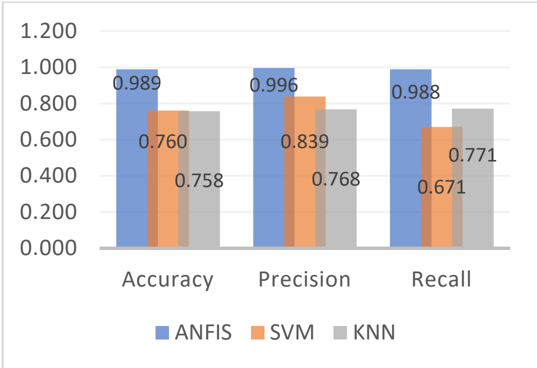

Figure 3: Comparison performance and accuracy

Based on Table 9 and Figure 2, it can be seen that ANFIS performs better than SVM. The accuracy of ANFIS is higher, with a difference of 23%, and the precision and recall improve by 16% and 32%, respectively. Overall, ANFIS excels in this evaluation. In addition, ANFIS also outperformed KNN, with an accuracy difference of 23% and an increase in precision and recall of 23% and 22%, respectively. In conclusion, ANFIS still performed better than SVM and KNN in this evaluation.

B. Result 2

The comparison results of leaning rate learning rate variations for RMSE, accuracy and F1-Score can be explained in Figures 1 to 3.

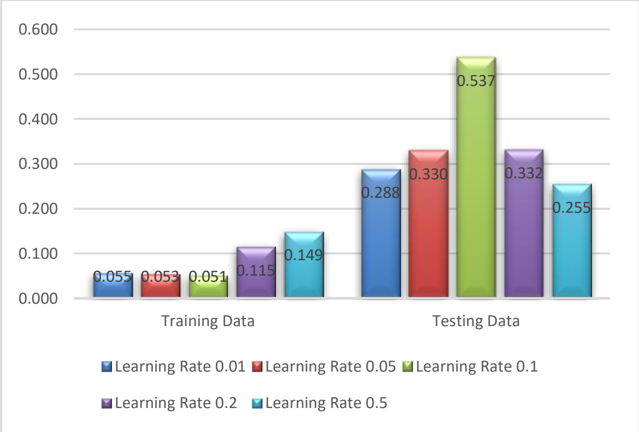

Figure 4: Comparison RMSE value for learning rate 0.01, 0.05, 0.1, 0.2, and 0.5

Figure 4 shows the RMSE comparison between training and testing data with various learning rates. It can be seen that at a learning rate of 0.01, the RMSE of the training data is 23% better than that of the testing data. The same is true for learning rates 0.05 and 0.1, with differences of 28% and 49% respectively. At a learning rate of 0.2, the RMSE on the training data is still 22% better, while at a learning rate of 0.5, the difference is 11%.

Figure 5: Comparison Accuracy value for learning rate 0.01, 0.05, 0.1, 0.2, and 0.5

Figure 5 shows the comparison of accuracy values between training data and testing data with variations in learning rate. It is found that in the training data, the accuracy value with a learning rate of 0.01 is 7% higher than that of the testing data. Similarly, at a learning rate of 0.05, the accuracy value in the training data is 10% higher than that in the testing data. A similar pattern is also seen at a learning rate of 0.1, where the accuracy value in the training data is 9% higher than the testing data. At a learning rate of 0.2, the accuracy value in the training data is 8% higher than the testing data. Finally, at a learning rate of 0.5, the accuracy value in the training data is 5% higher than the testing data.

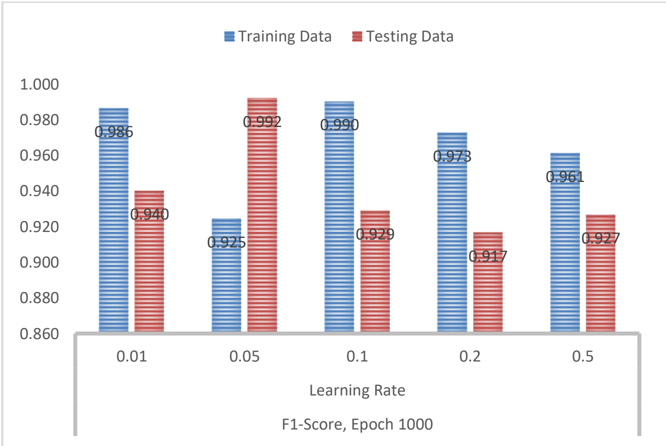

Figure 6: Comparison F1-Score value for learning rate 0.01, 0.05, 0.1, 0.2, and 0.5

Figure 6 shows the comparison of F1-Score values between training data and testing data with various learning rates. The results show that in the training data, F1-Score with a learning rate of 0.01 is higher by 5% compared to the testing data. Conversely, at a learning rate of 0.05, it can be seen that the F1-Score value in the training data is lower by 7% compared to the testing data. A similar pattern is seen at learning rate 0.1, where the F1-Score in the training data is higher by 6% compared to the testing data. At a learning rate of 0.2, the F1-Score in the training data is higher by 6% compared to the testing data. At a learning rate of 0.5, the F1-Score on the training data is higher by 3% compared to the testing data.

C. Discussion

The results of ANFIS implementation with various learning rates show that the RMSE values in Tables 1 to 5 of the model provide more accurate predictions on training data than on testing data. The evaluation of accuracy, F1-Score, precision, and recall also shows consistency, with the values on the training data being higher than on the testing data. This indicates that the model tends to generalize better from the training data, indicating a good fit for the model.

ANFIS showed superiority over SVM and KNN, with accuracy 23% higher than SVM, precision increased by 16%, and recall significantly improved by 32%. The evaluation results confirm the ability of ANFIS to recognize patterns and classify data [21].

When compared to KNN, ANFIS again showed superior performance. The accuracy of ANFIS is higher, with a difference of 23%. The increase in precision reached 23%, and recall also increased by 22% when compared to KNN. Thus, ANFIS remained consistent in performing better than the KNN method, highlighting the effectiveness of this model in the specific evaluation context [22].

ANFIS has performed better than SVM and KNN in past classification tasks. These advantages can provide a better understanding of the model's fit to the data and its potential applicability in real-world situations [23].

Comparative analysis of RMSE, accuracy, and F1-Score values between training and testing data with varying learning rates highlights the performance of the model at various learning rates. First of all, the RMSE values of the training data show significant differences with the testing data at every learning level. Increasing the learning rate from 0.01 to 0.5 provides consistent improvements, with differences of 23%, 28%, 49%, 22%, and 11%, respectively. The RMSE comparison analysis shows that the model has a better ability to minimize errors on the training data, but may face challenges in generalizing to the testing data [24].

Furthermore, the accuracy evaluation shows a similar pattern. At each learning level, the accuracy value on the training data is higher than that on the testing data. The difference ranges from 5% to 10%. The accuracy evaluation shows that the model tends to be superior in classifying previously viewed data compared to data that has never been viewed [14].

Finally, a comparison of the F1-Score values provides additional insight into the performance of the model. Although there is variation at certain learning rates, the general pattern shows that the F1-Score on the training data is higher than that on the testing data. This difference ranges from 3% to 7%, highlighting the model's ability to maintain a balance between precision and recall on the training data. Overall, these results indicate a gap between the model's performance on training and testing data. Further studies are needed to determine the optimal learning rate for the model to achieve a good balance between performance on both data sets [25].

IV.CONCLUSION

The development of the cardiovascular disease early detection model was conducted using a dataset of 500, which included features such as patient ID, blood sugar, cholesterol, uric acid, systolic and diastolic BMI, age, smoking habits, physical activity, lifestyle, genetic factors, and gender. These dataset labels are defined as 1 to indicate susceptible and 0 for non-susceptible (normal). The development of the early detection model involved three algorithms, namely SVM, K-NN, and ANFIS. Specific parameters of each algorithm were used in the model development process. Performance evaluation of the three models was conducted by measuring RMSE, accuracy, and F1-Score. The test results show that ANFIS provides the best performance. Furthermore, the model using ANFIS was varied with learning rates of 0.01, 0.05, 0.1, 0.2, and 0.5. The results show that the model with a learning rate of 0.1 provides optimal performance in early detection of cardiovascular disease.

V. REFERENCES

[1]. G. Zhiting, T. Jiaying, H. Haiying, Z. Yuping, Y. Qunfei, and J. Jingfen, “Cardiovascular disease risk prediction models in the Chinese population- a systematic review and meta-analysis,” BMC

Public Health, vol. 22, no. 1, Dec. 2022, doi: 10.1186/s12889-022-13995-z.

[2]. C. Krittanawong et al., “Machine learning prediction in cardiovascular diseases: a meta-analysis,” Sci Rep, vol. 10, no. 1, Dec. 2020, doi: 10.1038/s41598-020-72685-1.

[3]. N. Louridi, S. Douzi, and B. El Ouahidi, “Machine learning-based identification of patients with a cardiovascular defect,” J Big Data, vol. 8, no. 1, Dec. 2021, doi: 10.1186/s40537-021-00524-9.

[4]. Y. Huang et al., “Using a machine learning-based risk prediction model to analyze the coronary artery calcification score and predict coronary heart disease and risk assessment,” Comput Biol Med, vol. 151, no. PB, p. 106297, 2022, doi: 10.1016/j.compbiomed.2022.106297.

[5]. M. Vaduganathan, G. A. Mensah, J. V. Turco, V. Fuster, and G. A. Roth, “The Global Burden of Cardiovascular Diseases and Risk: A Compass for Future Health,” J Am Coll Cardiol, vol. 80, no. 25, pp. 2361–2371, 2022, doi: 10.1016/j.jacc.2022.11.005.

[6]. D. Teoh, “Towards stroke prediction using electronic health records 08 Information and Computing Sciences 0806 Information Systems,” BMC Med Inform Decis Mak, vol. 18, no. 1, Dec. 2018, doi: 10.1186/s12911-018-0702-y.

[7]. S. Harjai and S. K. Khatri, “An Intelligent Clinical Decision Support System Based on Artificial Neural Network for Early Diagnosis of Cardiovascular Diseases in Rural Areas,” Proceedings - 2019 Amity International Conference on Artificial Intelligence, AICAI 2019, pp. 729–736, 2019, doi: 10.1109/AICAI.2019.8701237.

[8]. W. Lestari and S. Sumarlinda, “IMPLEMENTATION OF K-NEAREST NEIGHBOR (KNN) AND SUPORT VECTOR MACHINE (SVM) FOR CLASIFICATION CARDIOVASCULAR DISEASE.” [Online]. Available: https://archive.ics.uci.edu/ml/datasets/heart+disease.

[9]. N. A. Baghdadi, S. M. Farghaly Abdelaliem, A. Malki, I. Gad, A. Ewis, and E. Atlam, “Advanced machine learning techniques for cardiovascular disease early detection and diagnosis,” J Big Data, vol. 10, no. 1, Dec. 2023, doi: 10.1186/s40537-023-00817-1.

[10]. J. Saputra, C. Lawrencya, J. M. Saini, and S. Suharjito, “Hyperparameter optimization for cardiovascular disease data-driven prognostic system,” Vis Comput Ind Biomed Art, vol. 6, no. 1, Dec. 2023, doi: 10.1186/s42492-023-00143-6.

[11]. A. A. M. Ahmed and S. M. A. Shah, “Application of adaptive neuro-fuzzy inference system (ANFIS) to estimate the biochemical oxygen demand (BOD) of Surma River,” Journal of King Saud University - Engineering Sciences, vol. 29, no. 3, pp. 237–243, Jul. 2017, doi: 10.1016/j.jksues.2015.02.001.

[12]. K. Balasubramanian, “Improved adaptive neuro-fuzzy inference system based on modified glowworm swarm and differential evolution optimization algorithm for medical diagnosis,” Neural Comput Appl, vol. 33, no. 13, pp. 7649–7660, 2021, doi: 10.1007/s00521-020-05507-0.

[13]. O. Taylan, A. S. Alkabaa, H. S. Alqabbaa, E. Pamukçu, and V. Leiva, “svm,” Biology (Basel), vol. 12, no. 1, Jan. 2023, doi: 10.3390/biology12010117.

[14]. M. M. Sherzoy, “Atterberg Limits Prediction Comparing SVM with ANFIS Model,” Journal of Geoscience, Engineering, Environment, and Technology, vol. 2, no. 1, p. 20, 2017, doi: 10.24273/jgeet.2017.2.1.16.

[15]. H. Moayedi et al., “Novel hybrids of adaptive neuro-fuzzy inference system (ANFIS) with several metaheuristic algorithms for spatial susceptibility assessment of seismic-induced landslide,” Geomatics, Natural Hazards and Risk, vol. 10, no. 1, pp. 1879–1911, 2019, doi: 10.1080/19475705.2019.1650126.

[16].S. O. Sada and S. C. Ikpeseni, “Heliyon Evaluation of ANN and ANFIS modeling ability in the prediction of AISI 1050 steel machining performance,” Heliyon, vol. 7, no. December 2020, p. e06136, 2021, doi: 10.1016/j.heliyon.2021.e06136.

[17].A. F. Bendary, A. Y. Abdelaziz, M. M. Ismail, K. Mahmoud, M. Lehtonen, and M. M. F. Darwish, “Proposed ANFIS Based Approach for Fault Tracking , Detection , Clearing and Rearrangement for Photovoltaic System,” 2021.

[18].S. Mariadoss and F. Augustin, “Enhanced sugeno fuzzy inference system with fuzzy AHP and coefficient of variation to diagnose cardiovascular disease during pregnancy,” Journal of King Saud University - Computer and Information Sciences, vol. 35, no. 8, Sep. 2023, doi: 10.1016/j.jksuci.2023.101659.

[19].Z. Lyu et al., “Back-Propagation Neural Network Optimized by K-Fold Cross-Validation for Prediction of Torsional Strength of Reinforced Concrete Beam,” Materials, vol. 15, no. 4, 2022, doi: 10.3390/ma15041477.

[20].X. Li, D. Bian, J. Yu, M. Li, and D. Zhao, “Using machine learning models to improve stroke risk level classification methods of China national stroke screening,” BMC Med Inform Decis Mak, vol. 19, no. 1, Dec. 2019, doi: 10.1186/s12911-019-0998-2.

[21].I. M. El-hasnony, S. I. Barakat, and R. R. Mostafa, “Optimized ANFIS Model Using Hybrid Metaheuristic Algorithms for Parkinson ’ s Disease Prediction in IoT Environment,” vol. 8, 2020.

[22].M. Z. Abbas et al., “An adaptive-neuro fuzzy inference system based-hybrid technique for performing load disaggregation for residential customers,” Sci Rep, vol. 12, no. 1, pp. 1–14, 2022, doi: 10.1038/s41598-022-06381-7.

[23].H. B. Ly, B. T. Pham, D. Van Dao, V. M. Le, L. M. Le, and T. T. Le, “Improvement of ANFIS model for prediction of compressive strength of manufactured sand concrete,” Applied Sciences (Switzerland), vol. 9, no. 18, Sep. 2019, doi: 10.3390/app9183841.

[24].A. Yonar and H. Yonar, “Modeling air pollution by integrating ANFIS and metaheuristic algorithms,” Model Earth Syst Environ, no. 0123456789, 2022, doi: 10.1007/s40808-022-01573-6.

[25].A. A. Ewees and M. A. Elaziz, “Improved Adaptive Neuro-Fuzzy Inference System Using Gray Wolf Optimization: A Case Study in Predicting Biochar Yield,” Journal of Intelligent Systems, vol. 29, no. 1, pp. 924–940, 2020, doi: 10.1515/jisys-2017-0641.

Cite this article as :

Sri Sumarlinda, Azizah binti Rahmat, Zalizah binti Awang Long, Wiji Lestari, "The Comparative Early Prediction Model for Cardiovascular Disease Using Machine Learning", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 1, Issue 1, pp.24-33, January-February-2024. Available at doi : https://doi.org/10.32628/CSEIT241011
Journal URL : https://ijsrcseit.com/CSEIT241011