Iraqi Journal for Computers and Informatics
University of Information Technology and Communications
Vol. 51, No. 2, 2025, pp. 24-32

DOI: https://doi.org/10.25195/ijci.v49i2.618
Print ISSN: 2313-190X, Online ISSN: 2520-4912

Research Article

A Comparative Analysis of the Effectiveness of Multiple Models for Predicting Heart Failure using Data Mining

[1]Ahmed Sami Jaddoa
Business Informatics College
University of Information Technology and Communications
Baghdad, Iraq
ahmed.sami@uoitc.edu.iq

[2]Juliet Kadum
Department of Computer Science,
College of Science,
University of Diyala,
Baqubah, Iraq
julietkadum@uodiyala.edu.iq

[3]Amaal Kadum
Computer Science Department
College of Education,
Mustansiriyah University
Baghdad, Iraq
amaal_kadum3@uomustansiriyah.edu.iq

ARTICLE INFO
Article History
Received: 03/07/2025
Accepted: 24/07/2025
Published: 17/08/2025

ABSTRACT

One of the most fatal and well-known diseases worldwide, heart disease claims the lives of many people every year. In order to preserve lives, early detection regarding such disease is essential. One of the quickest, practical, and affordable methods of disease detection is Data Mining DM, an artificial intelligence AI technology. Human life is saved by healthcare services through prompt and efficient decision-making. For forecasting, decision-making, and disease prediction, DM technologies are essential. This research predicts heart disease using DM algorithms. There are 14 attributes in Cleveland dataset, including blood fat, blood pressure, gender, and age. The probability regarding patients developing heart disease in the future can be forecasted by analyzing such parameters. For classifying if heart disease is present or absent, two classification algorithms are used: Logistic Regression LR and K-Nearest Neighbor KNN. The precision, accuracy, f-score, and recall of the suggested model are evaluated. The outcomes of suggested model were tested using the heart disease dataset. Without preprocessing the dataset's variation values, the LR and KNN algorithms achieved the highest accuracy (61% and 71%, respectively). The algorithms (LR and KNN) preprocessed the dataset's variation values to get the highest accuracy (90% and 93%). In order to improve data driven medical decision-making, the presented research demonstrates how well DM algorithms work to increase heart disease prediction accuracy.

Keywords: *Heart Failure Prediction, Data Mining, LR, K-NN, Cleveland;*

1. INTRODUCTION

Predicting heart disease can be defined as one of the most complex cases of diseases in the world of medical science. One of the most vital human body organs is the heart. A World Health Organization WHO survey indicates that over 17 million people all over the world die from heart disease annually. There are many different heart disease types, such as arrhythmia, congenital heart disease, and coronary artery disease. Vertigo, chest pain, and profuse sweating are all symptoms of heart disease in patients [1]. A healthy heart and one with heart failure are depicted in Fig. 1 [2].

Iraqi Journal for Computers and Informatics
University of Information Technology and Communications
Vol. 51, No. 2, 2025, pp. 24-32
DOI: https://doi.org/10.25195/ijci.v49i2.618
Print ISSN: 2313-190X, Online ISSN: 2520-4912

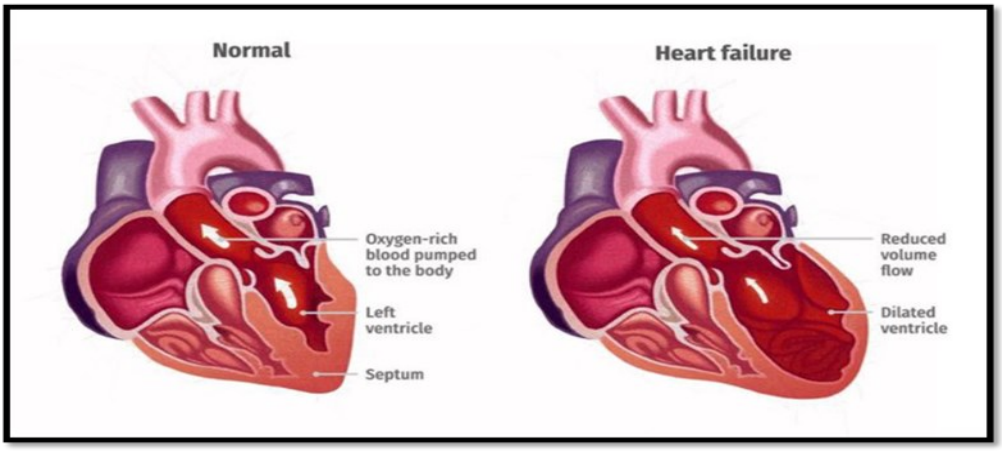

Fig1: Illustration of a healthy heart and one with heart failure [2].

Using DM technology in healthcare systems has become more widely recognized recently. With the use of statistical techniques, ML, and database systems. M is the process of finding trends and patterns in large datasets. DM techniques are widely used for extracting knowledge for a variety of purposes, such as the healthcare services sector [3].

DM is the process of extracting previously undiscovered, implicit, and potentially valuable information that is related to data in a non-trivial way. To put it briefly, it is the process of collecting knowledge from data by studying it from many perspectives. Large volumes of data regarding patients, disease diagnoses, etc. are produced by the healthcare sector these days. In order to find hidden patterns in data, DM offers a variety of tools. Service quality is one of the biggest issues facing the healthcare sector. Accurate disease diagnosis and effective patient treatment are prerequisites for quality of service. It is unacceptable when a poor diagnosis results in catastrophic outcomes [4].

Heart failure is a major global health and economic concern and one of the main causes of mortality. Early detection is a useful approach for lowering mortality rates as well as enhancing patients' quality of life, even in the face of medical advancements in treating this illness. In this regard, DM methods are seen as potential instruments in the prediction regarding heart failure. For classifying patients as either at risk or not at risk, classification algorithms like LR are utilized to evaluate clinical data including gender, age, blood lipid levels, blood pressure, and other critical criteria. Despite that such advancements, there is still a gap in the efficient application regarding such algorithms to actual patient data because of the different environmental and demographic factors. Thus, one crucial step in raising the accuracy regarding early heart failure detection is creating a predictive model depending on several algorithms with parameter optimization as well as pre-processing methods. This will enhance patient outcomes and enable more data-driven medical decision-making.

Several significant advances in the fields of DM applications and medical diagnostics are presented in this work:

1. Comparing classification algorithm performance: Evaluating how well DM algorithms like K-NN and LR predict heart failure.

2. Offering a reliable predictive model: To increase the accuracy regarding early heart failure detection, a predictive model is being developed depending on several algorithms with parameter optimization as well as pre-processing methods.

3. Evaluation without and with preprocessing: To determine how data preprocessing affects model performance, a comparison study is carried out, showing gains in accuracy and other evaluation criteria.
4. High Prediction Accuracy: The accuracy of the suggested models is up to 93% following preprocessing, which shows the usefulness of DM in the early detection of heart failure.

Research gaps

Several research gaps and limitations can be identified in previous studies [5-10] that used the same dataset. Notably, varying training-to-test ratios were used, and some studies omitted data normalization altogether. To address these issues and ensure consistency, the current work adopts a 90% training/10% testing split, which is appropriate given the relatively small dataset size. By addressing these gaps, we achieved better results than the previous work.

Iraqi Journal for Computers and Informatics
University of Information Technology and Communications
Vol. 51, No. 2, 2025, pp. 24-32
DOI: https://doi.org/10.25195/ijci.v49i2.618
Print ISSN: 2313-190X, Online ISSN: 2520-4912

2. LITERATURE REVIEW

A major provocation in healthcare area is heart disease forecasting. Numerous researchers have experience in this specific field. Here are a few explanations of research papers that are currently available in the literature. T.M. **Nithya and T.Sowndhariyaa et al. 2022** [5], researchers used various classification algorithms, which include Naïve Bayes NB, SVM, LR, K-NN, and Decision Trees, to provide a Cleveland heart disease prediction system in the present work. The medical history dataset includes Thalach (maximum heart rate), Chest Pain CP, and other information that can result in heart disease in an individual or patient. The University of California, Cleveland Machine Learning Repository dataset has been used in this research in order to assess a number of ML algorithms for the prediction of heart disease, such as NB, SVMs, Neural Network NN, LR, K-NN, and DT. According to the findings of this study, SVM algorithm has the highest accuracy compared to any algorithm when it comes to heart disease prediction, with an 85% accuracy. **Md. Imam Hossain et al. 2023** [6], and the goal of this study was accurately predicting heart disease by analyzing the many components in patient data. With using a correlation-based feature set selection approach with best search first, the most crucial features with regard to the prediction of heart diseases had been found. Gender, age, obesity, smoking, physical activity, diet, kind of chest pain, stress, diastolic blood pressure, previous chest pain, ECG, diabetes, troponin, and target were discovered to be the most significant factors in heart disease diagnosis. Two heart disease dataset types (all features and chosen features) have been used for testing and comparing several AI algorithms, including NB, LR, NB, K-NN, SVM, DT, RF, and Multi-Layer Perceptron MLP. When put to comparison with other AI methods and the use of all input features, RF's use of selected features yielded the highest accuracy rate (90%). **Bhatt, Chintan M. et al. 2023** [7], and this research provides a Cleveland heart disease dataset and suggests a k-pattern clustering technique with Huang starting in order to increase the accuracy of classification. Models like RF, XGBoost XGB, DT, MLP, and DT are employed. After being trained using 80:20 split data, the models' accuracy was as follows: XGBoost: 87.020% (with no cross-validation) and 86.870% (with cross-validation), RF: 86.920% (with no cross-validation) and 87.050% (with cross-validation), MLP: 86.940% (with no cross-validation) and 87.280% (with cross-validation), DT: 86.530% (with no cross-validation) and 86.370% (with cross-validation). With the maximum accuracy of 87.280%, the multi-layer cross-validation method beat all other algorithms in terms of accuracy, according to the basic research's findings. **AL-Jammali, Karrar. 2023** [8], and the goal of this study is accurately predicting heart diseases by analyzing the many components in patient data. Heart disease has been considered as the leading cause of a large percentage of deaths globally. To examine the effectiveness of predictive DM approaches, a number of experiments were conducted on the data to provide a Cleveland heart disease dataset. The findings indicate that the DT performs badly and SVM outperforms other predictive methods like ANN. According to the findings of this investigation, SVM is the most accurate algorithm for predicting heart disease, with maximum accuracy of 88.89%. **Ouhmad Inass et al. 2024** [9], and this research seeks to overcome such shortcomings through putting forth an improved heart failure prediction method. RF, K-NN, SVM, and DT have been among the ML algorithms used. With 88.04% accuracy, RF outperformed the others. In order to choose the best course of action, the study compares various algorithms and talks about model validation methods. **R. Renugadevi and Nivethitha. A. 2024** [10], and this research employed ML techniques for predicting heart disease incidence. To ascertain whether a person has heart disease, the model employs a number of methods as well as algorithms, such as SVM, LR and RF. Through the integration of such algorithms and techniques, the system seeks to precisely classify people according to whether or not they have heart disease. With an accuracy of 89.39%, LR was the most accurate, according to the data.

3. METHODOLOGY

The output regarding Heart Disease Model, which is suggested depending on a dataset from open-source Kaggle website, is the prediction of whether heart disease will exist or not. There are multiple main stages to the suggested prediction. Each stage consists of multiple phases that cooperate to achieve the goals. Preprocessing as well as classification are the main steps of the prediction process. The main diagram of heart disease is depicted in Fig 2, and Heart disease process is shown in Fig3.

Iraqi Journal for Computers and Informatics
University of Information Technology and Communications
Vol. 51, No. 2, 2025, pp. 24-32
DOI: https://doi.org/10.25195/ijci.v49i2.618
Print ISSN: 2313-190X, Online ISSN: 2520-4912

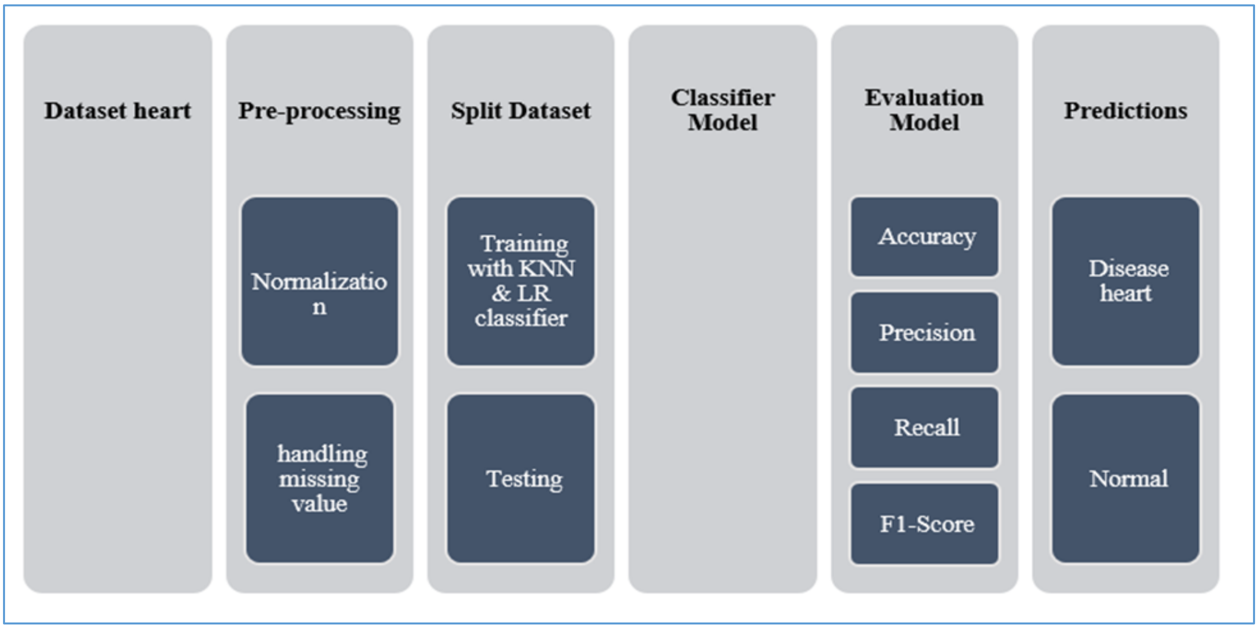

Fig. 2: Illustrates a general block diagram of heart disease.

Fig. 3: Workflow representation of the heart disease

3.1. HEART DISEASE DATASET

Data on 303 patients who had heart disease evaluations can be found in the Cleveland Heart Disease dataset. The data-set can be downloaded from open-access sources such as UCI-ML repository as well as Kaggle web-site https://www.kaggle.com/datasets/ritwikb3/heart-disease-cleveland. A total of 14 attributes, including clinical and demographic data like sex, age, resting blood pressure, kind of chest discomfort, exercise test results, and serum cholesterol level, are used for representing each patient. Each of the 303 records in the dataset represents a distinct patient. Values for each of the 14 attributes are contained in the data in each record, along with the dataset's prediction of the presence of heart disease or its absence. It has 5 numeric values and 8 nominal ones. All those features can be described in detail are as follows:

1. Age: Patients Age in years (Numeric)
2. Sex: Genders (Female: 0; Male: 1) (Nominal)
3. trest_bps: The patient's blood pressure level at the mode of resting in mm/HG (Numerical)
4. cp: Chest pain type experienced by a patient, it is classified to 4 classes. 0 representing typical angina, 1 representing atypical angina, 2 representing the non-angina pain, 3 representing the asymptomatic (Nominal)
5. fbs: Levels of blood sugar on fasting > 120mg/dl represented as 1 for "true" and 0 for "false" (Nominal)
6. chol: Serum cholesterol in mg/dl (Numeric)
7. thalach: Which represents the highest achieved heart rate (Numeric)
8. rest_ecg: Results of the electro-cardiogram while resting are denoted in three distinctive values: 0 : represents the normal 1: having ST-T wave abnormality (T wave inversions and/or ST elevation or depression of > 0.05mV) 2: shows definite or probable left ventricular hypertrophy by Estes' criteria (Nominal)

Iraqi Journal for Computers and Informatics
University of Information Technology and Communications
Vol. 51, No. 2, 2025, pp. 24-32
DOI: https://doi.org/10.25195/ijci.v49i2.618
Print ISSN: 2313-190X, Online ISSN: 2520-4912

9. old_peak: ST-depression that is induced by Exercise in relative with the state of resting (Numeric)
10. ex_ang: Angina that is induced by an exercise, where 0 depicts “NO” and 1 depicts “Yes” (Nominal)
11. ca: Represents the number of the major vessels (which is 0 to 3) (nominal)
12. slope: ST segment that is measured based on the slope throughout the peak exercise 0: represents up sloping; 1: represents flat; 2: represents the down sloping (Nominal)
13. target: Which represents target variable that is to be predicted 1 meaning that the patient suffers from a heart disease and 0 meaning that the patient is normal.
14. thal: The thalassemia blood disorder, where 0 denotes “NULL” 1 represents the normal flow of blood 2: represents the fixed defect (no blood flow in a part of heart) 3: represents a reversible defect (i.e., a blood flow is observed however, it isn’t normal(nominal)

3.2. DATA MINING (DM)

DM represents the process of gaining information and patterns from massive data amounts. Other names for it include knowledge mining from data, data/pattern analysis, extraction of knowledge, and that process of knowledge discovery [11]. KDD is a fundamental step in DM methods. DM, which is also referred to as knowledge discovery in data-bases, finds nontrivial extraction related to implied, previously unknown, and probably helpful information from facts inside sources. Along with KDD, data mining is frequently handled since word alternatives. Knowledge Discovery in a database includes DM. The next steps make up the iterative process for Knowledge Discovery Process Descriptions: Gaining knowledge regarding the application domain as well as the objectives of DM procedure, such as choosing a target data set, Combining and verifying the data set, data transformation, preprocessing, and cleaning, developing models and hypotheses, selecting appropriate DM algorithms, interpretation and display of the results, verification and testing of results [12].

Preprocessing:

The method's first stage is data preprocessing. In DM, it indicates the process of cleaning as well as converting unintelligible raw data into readable data so that DM models can be trained on it. In data preprocessing, data cleaning, data transformation, and data reduction take place. Data transformation includes normalization as well as aggregation, whereas data reduction involves the reduction of the amount of data. Data cleaning includes filling in missing values as well as smoothing noisy data. To avoid any inefficiencies and to achieve greater accuracy, data is deleted during this operation, along with any lost values and dots [13].

Normalization:

The features are scaled within a pre-defined range, usually between 0 and 1, by normalization, also known as Min-Max scaling. This method resolves problems caused by different feature scales that would otherwise affect DM models' convergence and performance. The efficacy of different DM models is increased through implementing suitable scaling, which guarantees equitable treatment of features and mitigates dominance by particular attributes as a result of scale differences [13].

Preprocessing and normalization play a critical role in the improvement of model accuracy, robustness, and generalization. By scaling features to a common range and cleaning the data, these steps help the model learn more effectively, reduce bias caused by dominant features, and lower variance by minimizing sensitivity to noise. This leads to more stable and reliable predictions, especially when working with small or imbalanced datasets.

3.3. MODELING

3.3.1. K-Nearest Neighbors (KNN)

A robust nonparametric technique used for both regression and classification applications is K-NN. In DM, the most popular approach for both regression and classification problems is K-NN. KNN is a supervised learning algorithm. It means to predict the output from input data. This classifier's classification procedure consists of three stages. In step 1, the K-value is determined. Step 2 ranks each test sample by calculating distance between all of the training data. Step 3 will use the majority vote method to provide the class name for the test sample data. Determine the new point's Euclidean distance from the current points in Equation (1): [14].

Iraqi Journal for Computers and Informatics
University of Information Technology and Communications
Vol. 51, No. 2, 2025, pp. 24-32
DOI: https://doi.org/10.25195/ijci.v49i2.618
Print ISSN: 2313-190X, Online ISSN: 2520-4912

$$\text{Distance} = \sqrt{(x2 - x1)^2 + (y2 - y1)^2} \qquad (1)$$

3.3.2. LOGISTIC REGRESSION (LR)

A supervised learning method is logistic regression. It is among the top algorithms for DM. Because of its high accuracy, LR is becoming increasingly popular and utilized for prediction and classification. Regression models come in two varieties: multinomial logistic regression models as well as binary logistic models. The target variable in binary logistic regression model may have a value of 1 or 0. Depending on the independent factors, the algorithm predicts the dependent variables. Depending on the user's symptoms, personal information, and medical test results (independent variables/attributes), the LR model in this project predicts whether or not the patient has heart disease with the particular stage of the disease (target) [15]. For predicting an output variable, an LR model considers one or more independent factors and examines their connection. Equation (2) gives p(x) as the LR function [16].

$$\text{P(x)} = \frac{1}{1+e^{-f(x)}} \qquad (2)$$

It is therefore usually in the range of 0 and 1. The probability for a given x is read as a function of p(x). if p(x) equals 1, it indicates that the output will be 1with certainty. Consequently, (1- p(x)) represents the probability that the output will be 0.

3.4. EVALUATION MEASUREMENTS

Rapid miner tools, such as software applications for DM, ML processing, and predictive analysis, are utilized to assess the prediction's accuracy, F1-score, precision, and recall. It makes sense to measure performance based on projected outcomes [17].

Accuracy: Frequency with which a model can correctly predict the result or class of a particular sample is measured by accuracy, as shown in equation (3).

$$\text{Accuracy} = \frac{TP + TN}{TP + FN + FP + TN} \qquad (3)$$

Recall: The ratio of the accurately detected positive cases, which are known as true positive cases as well, to total of the true positive as well as the false negative cases is measured by using a metric that is called recall. It assesses how well the model can detect actual positive instances, as shown in equation (4).

$$\text{Recall} = \frac{TP}{TP + FN} \qquad (4)$$

Precision: Which is a metric that is used for evaluating ratio of the accurately predicted positive cases (which are referred to as the true positive cases) to total number of the positive prediction cases that have been generated by the model, as shown in equation (5).

$$\text{Precision} = \frac{TP}{TP + FP} \qquad (5)$$

f1-score: for tasks of binary classification, f1-score is a frequently used performance metric combining recall with precision. It is determined through taking the harmonic mean of the recall and precision, which yields a single number that reflects their balanced combination, as shown in equation (6).

$$\text{f1-score} = 2 * \frac{Recall * Precision}{Recall + Precision} \qquad (6)$$

4. RESULTS AND DISCUSSION

When LR and KNN are used, the preprocessing dataset affects how well such algorithms perform. F1-score, precision, recall, and accuracy are used to assess performance. Comparison between LR and KNN that is applied to data-set with no pre-processing is listed in Table (1). Comparison results of LR and KNN applied to the data-set with pre-processing normalization of the dataset are shown in Table (2).

Iraqi Journal for Computers and Informatics
University of Information Technology and Communications
Vol. 51, No. 2, 2025, pp. 24-32
DOI: https://doi.org/10.25195/ijci.v49i2.618
Print ISSN: 2313-190X, Online ISSN: 2520-4912

TABLE 1: Results of the classifier algorithms for the dataset without preprocessing.

Algorithm	*Precision*	*Recall*	*F1-score*	*Accuracy*
KNN	%88	%68	%76	**%70**
LR	%66	%66	%66	%61

TABLE II: Results of the classifier algorithms with preprocessing normalization of dataset.

Algorithm	*Precision*	*Recall*	*F1-score*	*Accuracy*
KNN	%100	%89	%94	**%93**
LR	%94	%88	%91	%90

Tables 1 and 2 are necessary for the results' graphical representations. The accuracy results regarding the classification techniques without preprocessing the dataset are shown in Fig. 4. The accuracy results of classification techniques with preprocessing that uses normalization for the data-set characteristics are depicted in Fig5.

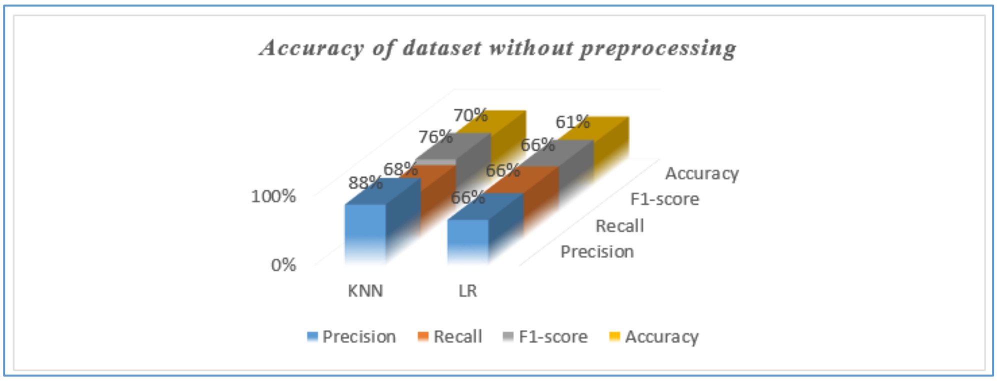

Fig. 4: Results of the classification methods for data-set without preprocessing.

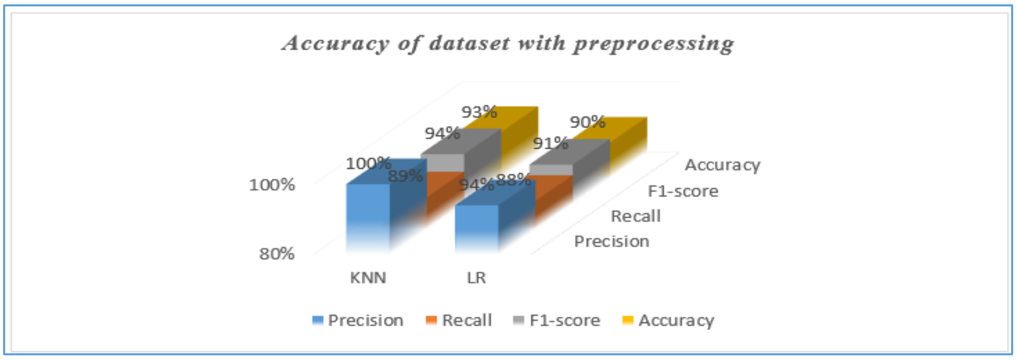

Fig. 5: Results of the classification methods for data-set with preprocessing.

These points could be presented based on a variety of experiments, a set of assessment metrics, and a comparison with earlier related research. Taking use of the preprocessing step to improve the suggested model's performance.

Iraqi Journal for Computers and Informatics
University of Information Technology and Communications
Vol. 51, No. 2, 2025, pp. 24-32
DOI: https://doi.org/10.25195/ijci.v49i2.618
Print ISSN: 2313-190X, Online ISSN: 2520-4912

Yet, as fig. 5 illustrates, this improvement depends on the type of preprocessing used in Min-Max Normalization procedure.

DISCUSSION

Classifier techniques like KNN, which achieved the greatest accuracy regarding 70% without preprocessing the dataset's missing values, and LR, which reached a %61 accuracy, compromise the accuracy of the suggested model. With the dataset's normalization preprocessing, KNN had the best accuracy (%93), whereas LR had the highest accuracy (%90). According to experiments, the performance of the suggested model is much improved by applying a preprocessing phase utilizing Min-Max Normalization, particularly in the case when KNN algorithm is used. These findings emphasize how crucial it is to select suitable preprocessing methods according to the kind of algorithm being employed. To sum up, such findings highlight how crucial the preprocessing stage is to enhancing model performance, particularly when utilizing the KNN algorithm. The effect of various preprocessing methods on model performance in diverse application domains warrants more investigation. The results we obtained are good by using the Min-Max normalization method compared to previous works that worked on the same dataset but did not address the variation in values that affect the performance and accuracy of prediction.

CONCLUSION AND FUTURE WORK

This study confirms that preprocessing especially normalization using the Min-Max method significantly improves the performance of classification models for heart disease prediction. KNN accuracy increased from 70% to 93%, and LR from 61% to 90% after applying normalization. These results highlight the importance of addressing missing values and feature scaling to enhance model accuracy and generalization, particularly with small datasets. Future work should explore other preprocessing techniques and evaluate their effects across different algorithms. Testing on larger and more diverse data-sets, as well as combining preprocessing with ensemble or hybrid models, could further improve prediction performance.

REFERENCES

[1] M. Ahmed and I. Husien, “Heart Disease Prediction Using Hybrid Machine Learning: A Brief Review,” J. Robot. Control, vol. 5, no. 3, pp. 884–892, 2024, doi: 10.18196/jrc. v5i3.21606.

[2] J. Kiran, N. Debbarma, and S. Ganjala, “Heart Disease Prediction Using Machine Learning,” Smart Innov. Syst. Technol., vol. 326 SIST, no. 1, pp. 263–272, 2023, doi: 10.1007/978-981-19-7513-4_24.

[3] R. Fadnavis, K. Dhore, D. Gupta, J. Waghmare, and D. Kosankar, “Heart disease prediction using data mining,” J. Phys. Conf. Ser., vol. 1913, no. 1, 2021, doi:10.1088/1742-6596/1913/1/012099.

[4] C. S.Dangare and S. S. Apte, “Improved Study of Heart Disease Prediction System using Data Mining Classification Techniques,” Int. J. Comput. Appl., vol. 47, no. 10, pp. 44–48, 2012, doi: 10.5120/7228-0076.

[5] T. Sowndhariyaa, T. M. Nithya, and A. Info, “51 International Journal for Modern Trends in Science and Technology Heart Disease Prediction using Machine Learning Algorithms,” Mach. Learn. Algorithms. Int. J. Mod. Trends Sci. Technol., vol. 8, no. 07, pp. 51–57, 2022.

[6] M. I. Hossain et al., “Heart disease prediction using distinct artificial intelligence techniques: performance analysis and comparison,” Iran J. Comput. Sci., vol. 6, no. 4, pp. 397–417, 2023, doi: 10.1007/s42044-023-00148-7.

[7] K. M. Shiwangi, J. K. Sandhu, and R. Sahu, “Effective Heart-Disease Prediction by Using Hybrid Machine Learning Technique,” Proc. Int. Conf. Circuit Power Comput. Technol. ICCPCT 2023, pp. 1670–1675, 2023, doi: 10.1109/ICCPCT58313.2023.10245785.

[8] K. AL-Jammali, “Prediction of Heart Diseases Using Data Mining Algorithms,” Inform., vol. 47, no. 5, pp. 57–62, 2023, doi: 10.31449/inf.v47i5.4467.

[9] I. Ouhmad et al., “Original Article Revolutionizing Heart Failure Prediction with Artificial Intelligence,” vol. 2, pp. 11–18, 2024.

[10] V. R. Gandla, D. V. Mallela, and R. Chaurasiya, “Heart failure prediction using machine learning,” AIP Conf. Proc., vol. 2705, no. 3, pp. 372–378, 2023, doi: 10.1063/5.0133750.

Iraqi Journal for Computers and Informatics
University of Information Technology and Communications
Vol. 51, No. 2, 2025, pp. 24-32

DOI: https://doi.org/10.25195/ijci.v49i2.618
Print ISSN: 2313-190X, Online ISSN: 2520-4912

[11] A. S. Ashour, N. Dey, and D. N. Le, “Biological data mining: Techniques and applications,” Min. Multimed. Doc., vol. 1, no. 4, pp. 161–172, 2017, doi: 10.1201/b21638.

[12] R. B. Diwate and A. Sahu, “Data Mining Techniques in Association Rule: A Review,” Int. J. Computer Science Inf. Technol., vol. 5, no. 1, pp. 227–229, 2014, [Online]. Available: http://connection.ebscohost.com/c/articles/99091355/data-mining-techniques-association-rule-review.

[13] “INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING Prediction of Heart Disease using data mining Techniques Based on Hybrid CNN-light GBM Method,” vol. 12, no. 4, pp. 291–301, 2024.

[14] J. Nayeem, S. Rana, and R. Islam, “Prediction of Heart Disease Using Machine Learning Algorithms,” vol. 1, no. 3, pp. 22–26, 2022.

[15] Kavya S., Prathanya Sree, D. M., N. B., and S. R., “Heart Disease Prediction Using Logistic Regression,” J. Coast. Life Med., vol. 11, no. 7, pp. 573–579, 2023, [Online]. Available: https://www.jclmm.com/index.php/journal/article/view/380.

[16] R. Baxani and M. Edinburgh, “Heart Disease Prediction Using Machine Learning Algorithms Logistic Regression, Support Vector Machine and Random Forest Classification Techniques,” SSRN Electron. J., 2022, doi: 10.2139/ssrn.4151423.

[17] Mulyawan, A. Bahtiar, G. Dwilestari, F. M. Basysyar, and N. Suarna, “Data mining techniques with machine learning algorithm to predict patients of heart disease,” IOP Conf. Ser. Mater. Sci. Eng., vol. 1088, no. 1, p. 012035, 2021, doi:10.1088/1757-899x/1088/1/012035.