Comparative Analysis of Machine Learning Algorithms for Investigating Myocardial Infarction Complications

Ali Satty
Department of Mathematics, College of Science, Northern Border University, Arar, Saudi Arabia
ali.hassan@nbu.edu.sa

Mohyaldein M. Y. Salih
Department of Mathematics, College of Science, Northern Border University, Arar, Saudi Arabia
mohyyassin@nbu.edu.sa

Abaker A. Hassaballa
Department of Mathematics, College of Science, Northern Border University, Arar, Saudi Arabia
abaker.abdalla@nbu.edu.sa

Elzain A. E. Gumma
Department of Mathematics, College of Science, Northern Border University, Arar, Saudi Arabia
elzain.Gumma@nbu.edu.sa (corresponding author)

Ahmed Abdallah
Department of Mathematics, College of Science, Northern Border University, Arar, Saudi Arabia
ahmad.adam@nbu.edu.sa

Gamal Saad Mohamed Khamis
Department of Computer Science, College of Science, Northern Border University, Arar, Saudi Arabia
gamal.khamees@nbu.edu.sa p

Received: 30 November 2023 | Revised: 20 December 2023 | Accepted: 22 December 2023

 DOI: https://doi.org/10.48084/etasr.6691

ABSTRACT

Myocardial Infarction (MI) is a condition often leading to death. It arises from inadequate blood flow to the heart, therefore, the classification of MI complications contributing to lethal outcomes is essential to save lives. Machine learning algorithms provide solutions to support the categorization of the MI complication attributes and predict lethal results. This paper compares various machine learning algorithms to classify myocardial infarction complications and to predict fatal consequences. The considered algorithms are Multilayer Perceptron (MLP), Naive Bayes (NB), and Decision Tree (DT). The main objective of this paper is to compare these algorithms in two scenarios: initially using the full dataset once and then using the dataset again, after implementing the WEKA attribute selection algorithm. To accomplish this goal, data from the Krasnoyarsk Interdistrict Clinical Hospital were employed. Results in general revealed that the MLP classifier demonstrated optimal performance regarding the full MI data, whereas the DT classifier emerged as more favorable when the dataset sample size was diminished through an attribute selection algorithm.

Keywords-data mining; classification; multilayer perceptron; Naive Bayes; decision tree; prediction

I. INTRODUCTION

When the blood supply to a myocardium section is diminished or halted, it frequently results in Myocardial Infarction (MI), which can subsequently cause circulatory decline and abrupt death. This study centers around examining MI complications through data gathered at the Krasnoyarsk Interdistrict Clinical Hospital in Russia [1]. The primary aim of this data selection was to investigate the MI complications that contribute to lethal outcomes. Consequently, one suggestion is to extract insights from these complication records for determining the key attributes directly influencing the lethal outcomes. It is crucial to classify MI complications and predict lethal consequences in order for necessary measures to be taken to decrease the condition prevalence. Machine learning algorithms offer potential solutions for classifying complication attributes that may result in death due to MI and for predicting the lethal outcomes. This study focusses on the three machine learning algorithms, Multi-layer Perceptron (MLP), Naive Bayes (NB), and Decision Tree (DT).

In closely related work, machine learning algorithms have been increasingly utilized in the cardiology field. For instance, authors in [2] compared the use of DT and NB for modeling cardiovascular disease. The dataset they used contains 299 patients with 13 attributes. Their results showed that both DT and NB offer high accuracy. Authors in [3] analyzed DT, NB, and Neural Networks (NNs) to design a predictive model for heart disease data containing 336 patients with 24 attributes. They found that all algorithms provided superior performance. Authors in [4] compared Support Vector Machine (SVM), Random Forest (RF), K-Nearest Neighbors (KNN), and Logistic Regression (LR) algorithms for performing myocardial infraction prediction, employing a dataset based on 303 patients with 14 attributes. Their findings yielded that LR outperformed the other classifiers. Other studies regarding the use of using machine learning algorithms in cardiology can be found in [5-7]. This paper, however, aims to compare the performance of the MLP, NB, and DT classifiers in two scenarios: using the full dataset of MI complications once and utilizing the same dataset again, after implementing the WEKA attribute selection algorithm. The classifiers are applied to build a classification model for the MI complications in order to predict lethal outcomes. These classifiers were chosen due to their robust theoretical foundations, spanning various schools of thought. The WEKA attribute selection technique [8] is applied to identify the most relevant to fatal outcomes attributes.

II. METHODOLOGY

A. Multilayer Perceptron

The MLP algorithm basically is a class of feed-forward Artificial Neural Networks (ANNs) that consist of several layers working together to process input data, and produce an output for various applications like pattern recognition, classification, and prediction tasks [9-11].

B. Naive Bayes

The NB algorithm is a widely recognized probabilistic classifier that fundamentally utilizes the probability concept. Its primary function is to address classification issues in a straightforward and efficient manner. By applying Bayes' theorem principles, the algorithm learns the probability of individual objects, their features and the groups to which these features belong [12-14].

C. Decision Tree

The DT algorithm is a classification method for modeling the decision-making process by constructing a tree structure based on specific characteristics of the dataset being analyzed [13-15]. It serves as a tree-diagram-based technique, featuring a root node at the top and leaf nodes at the bottom. Each leaf node possesses a target class attribute.

D. Data Description

The data were amassed via a survey study carried out at the Krasnoyarsk Interdistrict Clinical Hospital in Russia, and are currently accessible through the UCI machine learning repository [1]. Comprised of multivariate characteristics, the dataset includes 1700 patients with 123 attributes. Among these attributes, 111 served as input features, while 12 were considered complications. Nine of these 111 input features were assessed before the conclusion of days one, two, and three, while 102 input features were evaluated upon admission to the hospital. This database primary aim was to explore MI complications in the patient population. Chronic heart failure patients displayed the highest rate (23.18%) of complications, whereas supraventricular tachycardia-afflicted patients exhibited the lowest rate (1.18%). The results due to MI complications were categorized as a diverse outcome (1429 alive cases and with 271 dead cases). The major death causes were identified within 7 categories: Cardiogenic shock, pulmonary edema, myocardial rupture, progression of congestive heart failure, thromboembolism, systole, and ventricular fibrillation. Utilizing the previously mentioned machine learning algorithms, this study primarily concentrates on classifying MI complications and predicting lethal outcomes.

E. Pre-processing Stage

Like other big data, the myocardial infraction data used in this study suffer from some common irregularities that may be found in any real-world database. The pre-processing phase addresses these issues, rendering the data suitable for subsequent classification and prediction tasks. Patient IDs have been eliminated from the analytical process due to their irrelevance in the analysis. Additionally, attributes not directly connected to predicting lethal outcomes, such as the recurrence of pain post-hospital admission, have been excluded. Approximately 8% of the remaining dataset contains missing information. To avoid biased results in predictive modeling caused by missing data, the Multiple Imputation (MI) technique is employed via the MICE function [16], proving to be more effective than conventional methods like complete and available case techniques. The 3 classifiers were applied to the myocardial infraction data in 2 scenarios: using the full dataset once, and using the dataset again, after implementing the WEKA attribute selection algorithm. Sixteen attributes, which included the lethal outcome, were selected after applying the WEKA algorithm. These attributes primarily relate to complications and input features measured at the time of hospital admission. The selected attributes exhibiting a strong correlation with lethal outcomes are presented in Table I:

TABLE I. SELECTED ATTRIBUTES AFTER WEKA

Features measured at the time of admission	Complications
Age	Ventricular fibrillation
Quantity of MIs in the anamnesis	Myocardial rupture
Cardiogenic shock at the time of admission to intensive care unit	Chronic heart failure
Presence of an anterior MI (left ventricular)	MI relapse
Presence of a right ventricular MI	Post-infarction angina
ECG rhythm at the time of hospital admission – sinus (with a 60-90 heart rate)	
Complete RBBB on ECG at the time of hospital admission	
White blood cell count (billions per liter)	
Time elapsed from the beginning of the CHD attack to the hospital	
Use of liquid nitrates in the ICU	

F. *Measures for Performance Assessment*

In this study, three classifiers - MLP, NB, and DT - are evaluated based on several metrics:

1. Accuracy, which denotes the proportion of accurate predictions out of the total.
2. Error rate, referring to the fraction of incorrect predictions within the total of predictions.
3. The Kappa statistic is a normalized measure of agreement [17]. Kappa values range between -1 and 1. It is important to note that a Kappa value of 1 represents perfect agreement, whereas a value of 0 indicates agreement not better than random chance. Furthermore, if the Kappa value is greater than 0, the classifier performs superiorly compared to random chance. A negative Kappa value signifies an agreement inferior to that expected by mere chance.
4. F-Measure: This metric amalgamates recall and precision measures into a singular performance indicator.
5. Matthew's Correlation Coefficient (MCC) is a means for calculating the difference between predicted and actual values [18].
6. The area under the Receiver Operating Characteristic (ROC) curve is commonly used for measuring the classifier performance [19]. The ROC curve is drawn by plotting the True Positive Rate (TPR) on the y-axis and the False Positive Rate (FPR) on the x-axis. The classifier performs better when the ROC curve is closer to the upper left corner.

Although accuracy and error rate are deemed crucial and effective metrics, they possess certain limitations in the context of imbalanced data. Consequently, alternative measures such as Kappa statistic (k), F-measure, and MCC offer more insightful evaluations as they account for class balance and diverse error types. Given the MI complications, the derived data are imbalanced. These alternative metrics were proven to be essential for this research.

III. RESULTS AND DISCUSSION

A. *Results of MLP, NB, and DT for All Attributes*

Figure 1 presents the accuracy and error rate results. The accuracy metrics obtained by the MLP, NB, and DT classifiers were 90%, 79%, and 88%, correspondingly. The MLP classifier consistently delivered superior performance.

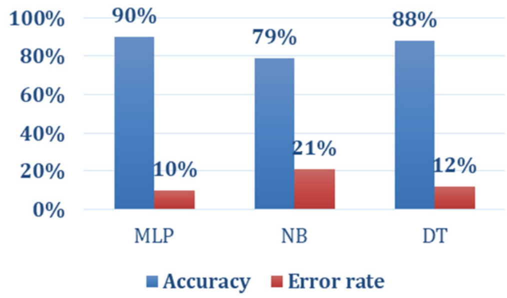

Fig. 1. Accuracy and error rate for MLP, NB, and DT.

Table II displays the Kappa statistic results obtained for the three classifiers. The values for these classifiers were 0.6292, 0.4036, and 0.4654. Upon evaluating these outcomes, it can be inferred that the 3 classifiers employed in our model exhibited effective performance as their Kappa statistics demonstrated values exceeding 0.

TABLE II. KAPPA STATISTIC VALUES

MLP	NB	DT
0.6292	0.4036	0.4654

Nevertheless, the recall percentage for NB was observed to be adversely affected. It is crucial to highlight that any variations in the level of recall measurement inhibit linear alterations in precision measurement. This phenomenon can perhaps be attributed to the substitution of False Positives (FP) and False Negatices (FN), within the denominator of the precision metric. It is well-established that higher precision and F-measure values are preferable. Consequently, as depicted in Table III, the highest F-measure value corresponded to MLP, signifying its effectiveness as a classification instrument. Furthermore, the findings revealed that the MCCs corresponding to MLP, NB, and DT were all positive, indicating satisfactory predictions by every classifier. Among them, MLP held a comparative advantage due to its superior correlation coefficient. Succinctly put, since a positive association between actual and predicted classifications prevailed in all instances, there was a high likelihood that most predicted values aligned with their true counterparts.

TABLE III. DETAILED ACCURACY BY CLASS

TP rate	FP rate	Precision	Recall	F-Measure	MCC
MLP					
0.901	0.259	0.883	0.901	0.892	0.682
NB					
0.791	0.227	0.861	0.791	0.822	0.493
DT					
0.881	0.513	0.832	0.881	0.851	0.489

As illustrated in Table III, the specific accuracy of every class-classification employing MLP, NB, and DT classifiers is presented. The results demonstrated that the recall and precision percentages of all the classifiers were relatively close to one another as their values did not exhibit a significant disparity. Figure 2 represents the ROC curves for the three classifiers. It can be observed that the MLP classifier demonstrates the best performance, as its ROC curve is closer to the top-left corner, followed closely by DT.

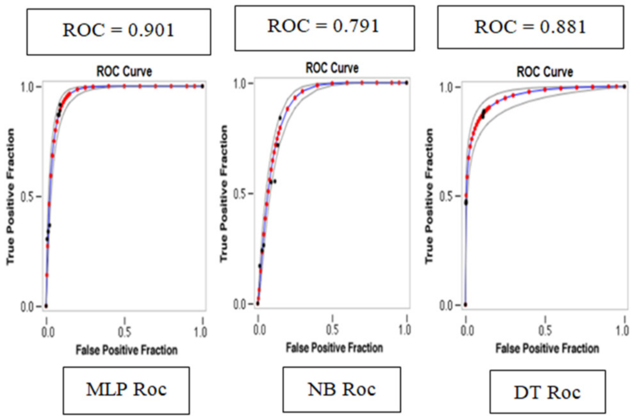

Fig. 2. Roc curves for MLP, NB. and DT algorithms for all attributes.

A. *Results after using the WEKA Attribute Selection Technique*

Figure 3 displays the accuracy and error rates for MLP, NB, and DT when reducing the dataset through the use of the attribute selection technique. The DT classifier exhibited a higher accuracy percentage, signifying its effectiveness as a classification instrument. Interestingly, the accuracy percentages of MLP and NB were quite comparable. Consequently, the DT classifier demonstrated lower error rate as opposed to those achieved by MLP and NB. It is noteworthy that MLP lower accuracy percentage was not observed when analyzing the full dataset.

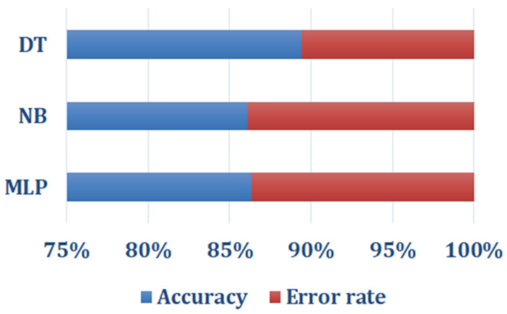

Fig. 3. Accuracy and error rate for MLP, NB, and DT.

The results of the Kappa statistics, as determined by the three classifiers, are exhibited in Table IV. All values for this metric were positive, denoting that the classifiers displayed efficient performance within the analytical model. In regard to the full dataset, the MLP classifier yielded the highest Kappa statistic. However, when employing the attribute selection method, this value became the lowest.

TABLE IV. KAPPA STATISTIC VALUES

MLP	NB	DT
0.4717	0.4843	0.4875

Table V illustrates the comprehensive accuracy of each class-classification technique, using the MLP, NB, and DT classifiers. Based on these findings, it is evident that the MLP and NB classifiers produced nearly identical percentage results. Both MLP and NB demonstrated high levels of precision and recall with a preference for the DT classifier. Specifically, the most elevated percentages were attributed to the DT classifier. Regarding the F-measure criterion, all classifiers displayed effective performance as their F-measure percentages were consistently higher in every instance. Among the classifiers employed in this research, the DR classifier was deemed more favorable. In relation to the MCC measure, all the algorithms exhibited positive correlation coefficients. Notably, DT's MCC surpassed those achieved by MLP and NB.

TABLE V. DETAILED ACCURACY BY CLASS FOR MLP, NB AND DT

TP rate	FP rate	Precision	Recall	F-Measure	MCC
MLP					
0.864	0.400	0.834	0.864	0.849	0.510
NB					
0.861	0.378	0.845	0.861	0.853	0.512
DT					
0.894	0.540	0.968	0.894	0.921	0.547

The ROC curves obtained by the three classifiers are displayed in Figure 4. The DT classifier introduced the best performance, as its ROC curve was closer to the upper left corner. Yet, the performance of the MLP and NB classifiers did not differ significantly.

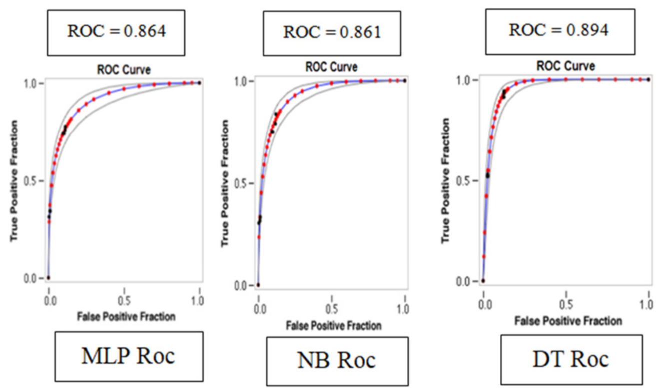

Fig. 4. ROC curves for MLP, NB and DT algorithms after using the WEKA attribute selection technique.

Upon evaluating all three classifiers concurrently, the MLP classifier showed optimal performance for the full MI complications dataset. This finding is well documented in previous studies [3]. In contrast, the DT classifier emerged as a more favorable classifier when the dataset sample size was diminished through the attribute selection algorithm, followed by NB. The DT and NB findings are in line with the findings in [2, 3], according to which both classifiers are appropriate under a small number of attributes.

IV. CONCLUSION

This study has examined the myocardial infarction complications that could potentially lead to lethal consequences. The primary objective of this article was to evaluate the effectiveness of MPL, NB, and DT classifiers in categorizing myocardial infarction complications and predicting lethal outcomes. A comparison between classifiers was made, either employing an attribute selection algorithm or not. The findings provided an overall perspective on the efficacy of various machine learning models in distinguishing myocardial infarction complications and predicting mortality resulting from these complications. It was evident from the results that classifier performance may fluctuate based on multiple factors such as data sample size and their features. Therefore, these aspects must be considered when deciding which classifiers are most suitable for handling data concerning myocardial infarction complications.

In the context of analyzing myocardial infarction complications, and using the full dataset (large sample size), it was observed that the MPL had a superior performance over the NB and DT classifiers, with DT being a close second. Among the classification algorithms employed in this research, MPL emerged as the most effective, achieving an accuracy rate of up to 90%. NB exhibited the lowest accuracy rate, translating to moderate performance in classifying myocardial infarction complications. While NB may still be applicable to the myocardial infarction data utilized in this study, it is crucial to acknowledge that this classifier might be a less favorable option for predicting lethal outcomes when compared to more sophisticated classifiers like MPL for large data samples.

Upon employing the WEKA attribute selection algorithm, DT was consistently proved to be the most superior classifier in terms of accuracy, demonstrating the highest accuracy rate with a smaller sample size. This suggests that the implementation of attribute selection algorithms augmented DT accuracy and subsequently heightened its ability to categorize myocardial infarction complication cases. Furthermore, it can be discerned that DT performance experienced significant improvement as the sample size diminished. Additionally, it is important to note that MLP values were lower than those derived from other classifiers. This signifies that when data instances (sample size) decrease, the classification correctness rate dependent on MPL also decreases. Hence, MLP performance was more effective on large datasets.

ACKNOWLEDGEMENT

The authors extend their appreciation to the Deanship of Scientific Research at Northern Border University, Arar, KSA for funding this research work through the project number "NBU-FFR-2024-2877-01."

REFERENCES

[1] S. A. Golovenkin *et al.*, "Myocardial infarction complications Database." University of Leicester, Mar. 30, 2020, https://doi.org/10.25392/leicester.data.12045261.v3.

[2] V. S. K. Reddy, P. Meghana, N. V. S. Reddy, and B. A. Rao, "Prediction on Cardiovascular disease using Decision tree and Naïve Bayes classifiers," *Journal of Physics: Conference Series*, vol. 2161, no. 1, Jan. 2022, Art. no. 012015, https://doi.org/10.1088/1742-6596/2161/1/012015.

[3] N. Chaithra and B. Madhu, "Classification Models on Cardiovascular Disease Prediction using Data Mining Techniques," *Cardiovascular Diseases & Diagnosis*, vol. 6, no. 6, pp. 1–4, 2018, https://doi.org/10.4172/2329-9517.1000348.

[4] P. Maindarkar and S. S. Reka, "Machine Learning-Based Approach for Myocardial Infarction," in *International Conference on Artificial Intelligence and Sustainable Engineering*, Singapore, 2022, pp. 17–27, https://doi.org/10.1007/978-981-16-8542-2_2.

[5] B. Trstenjak, D. Donko, and Z. Avdagic, "Adaptable Web Prediction Framework for Disease Prediction Based on the Hybrid Case Based Reasoning Model," *Engineering, Technology & Applied Science Research*, vol. 6, no. 6, pp. 1212–1216, Dec. 2016, https://doi.org/10.48084/etasr.753.

[6] R. Ramesh and S. Sathiamoorthy, "A Deep Learning Grading Classification of Diabetic Retinopathy on Retinal Fundus Images with Bio-inspired Optimization," *Engineering, Technology & Applied Science Research*, vol. 13, no. 4, pp. 11248–11252, Aug. 2023, https://doi.org/10.48084/etasr.6033.

[7] A. K. Dubey, A. K. Sinhal, and R. Sharma, "An Improved Auto Categorical PSO with ML for Heart Disease Prediction," *Engineering, Technology & Applied Science Research*, vol. 12, no. 3, pp. 8567–8573, Jun. 2022, https://doi.org/10.48084/etasr.4854.

[8] R. Bouckaert *et al.*, *WEKA manual for version 3-6-0*. Hamilton, New Zealand: The University of WAIKATO, 2008.

[9] L. B. Almeida, "Multilayer Perceptrons," in *Handbook of Neural Computation*, IOP Publishing Ltd. and Oxford University Press, 1997.

[10] P. Cortez, "Data Mining with Neural Networks and Support Vector Machines Using the R/rminer Tool," in *Advances in Data Mining. Applications and Theoretical Aspects*, Berlin, Heidelberg, 2010, pp. 572–583, https://doi.org/10.1007/978-3-642-14400-4_44.

[11] A. A. Heidari, H. Faris, S. Mirjalili, I. Aljarah, and M. Mafarja, "Ant Lion Optimizer: Theory, Literature Review, and Application in Multi-layer Perceptron Neural Networks," in *Nature-Inspired Optimizers: Theories, Literature Reviews and Applications*, S. Mirjalili, J. Song Dong, and A. Lewis, Eds. Cham, Germany: Springer International Publishing, 2020, pp. 23–46.

[12] P. Adriaans and D. Zantinge, *Data Mining*, 3rd ed. New York, NY, USA: Addison-Wesley, 1996.

[13] F. Gorunescu, *Data Mining*, Berlin, Heidelberg, Germany: Springer, 2011.

[14] I. H. Witten, E. Frank, and M. A. Hall, *Data Mining: Practical Machine Learning Tools and Techniques*, 3rd ed. Burlington, MA, USA: Morgan Kaufmann, 2011.

[15] B. V. Chowdary, A. Gummadi, U. N. P. G. Raju, B. Anuradha, and R. Changala, "Decision Tree Induction Approach for Data Classification Using Peano Count Trees," *International Journal of Advanced Research inComputer Science and Software Engineering*, vol. 2, no. 4, pp. 475–479, 2010.

[16] S. van Buuren and K. Groothuis-Oudshoorn, "mice: Multivariate Imputation by Chained Equations in R," *Journal of Statistical Software*, vol. 45, pp. 1–67, Dec. 2011, https://doi.org/10.18637/jss.v045.i03.

[17] J. Cohen, "A Coefficient of Agreement for Nominal Scales," *Educational and Psychological Measurement*, vol. 20, no. 1, pp. 37–46, Apr. 1960, https://doi.org/10.1177/001316446002000104.

[18] B. W. Matthews, "Comparison of the predicted and observed secondary structure of T4 phage lysozyme," *Biochimica et Biophysica Acta (BBA) - Protein Structure*, vol. 405, no. 2, pp. 442–451, Oct. 1975, https://doi.org/10.1016/0005-2795(75)90109-9.

[19] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, Jun. 2006, https://doi.org/10.1016/j.patrec.2005.10.010.