Mehran University Research Journal of Engineering and Technology

https://doi.org/10.22581/muet1982.3268

2025, 44(2) 155-163

Impact of computing platforms on classifier performance in heart disease prediction

Beenish Ayesha Akram [a, *], Muhammad Irfan [a], Amna Zafar [b], Sidra Khan [a], Rubina Shaheen [a]

[a] *Department of Computer Engineering, University of Engineering and Technology, Lahore, Pakistan*

[b] *Department of Computer Science, University of Engineering and Technology, Lahore, Pakistan*

[*] Corresponding author: Beenish Ayesha Akram, Email: beenish.ayesha.akram@uet.edu.pk

Received: 13 May 2024, Accepted: 27 March 2025, Published: 01 April 2025

KEYWORDS

Heart disease prediction
Support Vector Machine
Classification
Machine learning
Classification metrics
Platform Comparison

ABSTRACT

Prediction and classification, a supervised learning technique in machine learning, addresses various challenges related to finding useful patterns present in data. This work explores how different computing platforms influence the accuracy of classification results when employing the same models. Heart disease, a widespread global health issue affecting both men and women, results from a complex interplay of lifestyle factors and genetics. Through visual representations, we examined the diverse factors influencing heart stroke occurrences. We employed multiple classification methods such as Logistic Regression, K Nearest Neighbour (KNN), Support Vector Machine (SVM), Naïve Bayes, and Decision Tree (DT), assessing their accuracy using WEKA and Google Colab (using Scikit-Learn library). Our evaluation revealed that SVM achieves 77% accuracy when implemented using Scikit-Learn, demonstrating superiority over other methods. However, when using WEKA, both logistic regression and SVM demonstrated nearly 91% accuracy using the exact same hyperparameters. This research demonstrated the significance of platform selection in influencing classifier performance, offering valuable insights on how results reported in research can be impacted by the selection of the software and tools, using heart disease prediction as a use case scenario.

1. Introduction

The process of extracting and learning the underlying patterns in data to identify the category of the sample at hand is known as classification. It had been extensively utilized to tackle numerous real-world problems. Classification is a function that creates different classes of data by separating the data points. Several algorithms are used in classification. Linking a variable of interest with other features/attributes is the main goal of classification. It establishes a connection between the variable and prediction using that connection.

The medical industry makes extensive use of machine learning techniques and tools. Machine learning (ML) methods are a cost-effective way for numerous health organizations to address problems with quick diagnosis and treatment quality. Heart disease is a general term used to describe a variety of heart-related illnesses, such as blood vessel disease, heart attack, stroke, heart failure, etc.

The terms “heart disease” and “cardiovascular disease", are often confused by the public. The latter term defines circumstances that might result in a heart attack, chest discomfort, or stroke. According to a 2016 study conducted by the WHO [1], over 17 million deaths worldwide were attributed to cardiovascular disease, accounting for almost 30% of all deaths. According to a different study [2] by the World Health Organization, 38% of Morocco's annual deaths are attributed to cardiovascular disease, and

14% are caused by cancer. If no precautions are taken, these numbers are predicted to rise dramatically by 2030 [3]. There are many factors of heart attacks, but lifestyle habits such as eating habits, physical inactivity, and obesity are considered as major reasons for heart attacks [4], [5]. Some more factors have been described in [6], which are family history, smoking, poor diet, high blood pressure, high blood cholesterol, obesity, physical inactivity, and hypertension.

It is imperative to highlight that in medical-related problems; accurate diagnosis is very essential. The machine learning-based heart disease prediction is being employed by researchers. According to a study [7], machine learning algorithms are crucial for disease classification and radiological applications because they can automatically recognize intricate patterns and support radiologists in making critical decisions.

Heart disease has been identified as one of the biggest causes of death in the developed countries [6]. The fact that heart attacks are discovered much later in life contributes to the death toll from this deadly illness. But by foreseeing cardiac disease in its early stages, machine learning can be a huge assist in solving this issue.

Classification is becoming increasingly important in the healthcare industry thanks to the ever-increasing accuracy and precision exhibited by machine learning models and algorithms. The models are being trained for classification tasks either by using unsupervised learning like clustering to find facts that were previously unknown [8] or by using supervised techniques, such as predicting neurological diseases like Parkinson's [9], cardiovascular diseases like diabetes [10], or chronic diseases like diabetes [11]. Diagnosing heart disease is a complex task, often influenced by the patient's current health condition. Effective diagnosis and appropriate treatment have the potential to save many lives [12]. However, without access to technology, diagnosing and treating heart diseases becomes notably challenging, with the risk of inaccurate results posing a danger to patients [13].

ML model accuracy is directly impacted by anomalies and inconsistencies in the data. This data is used to build accurate classifiers. There are a lot of free ML platforms accessible, which lead to better models, but the decision of picking up the best one is tough. Along with that, the research community has a variety of platforms and tools to build ML classifier models. To be able to compare the results of numerous research efforts focusing on disease classification, a thorough investigation is required where the impact of the platform is evaluated on the classifier's performance and accuracy. Therefore, this study emphasizes the significance of platform selection in classifier outcomes, particularly in addressing critical problems like heart disease diagnosis.

The main focus of this research is to present the comparison of ML platforms, including WEKA and Google Colab (using the Scikit-Learn library), that classify cardiovascular disease using machine learning classifiers. The performance measures obtained from both platforms can be then compared to understand their influence on the classifier’s results.

The rest of the paper is organized as follows: Section 2 is related to literature review. Section 3 details the dataset description and exploratory data analysis. Section 4 discusses the methodology and classifiers utilized for this research. In Section 5, we analyze the results, and Section 6 discusses the conclusion.

2. Literature Review

Classification refers to finding underlying patterns in the data, which can help distinguish one category of data apart from others. This multifaceted procedure involves several key steps, including data collection, cleaning, processing, evaluation, visualization, and extraction of meaningful insights. Thus, classification encompasses a comprehensive approach aimed at harnessing the potential of data to derive valuable information and facilitate informed decision-making [14]. The crux of the process is to predict the dependent class label by using the independent set of values. It has vast applications in the field of medicine to handle very large datasets, as these datasets cannot be handled manually.

Classifiers such as neural networks, decision trees, regression, support vector machines, and Naïve Byes classifiers are a few of the techniques utilized to classify and identify the presence of disease as well as the type of disease. Researchers have employed various machine learning classifiers for disease prediction with high accuracy, which can help health experts in efficient and rapid diagnosis [15].

When it comes to the disease's contributing factors, heart disease is thought to be influenced by age, smoking, hypertension, and diabetes [16]. Researchers have utilized several ML classifiers, such as decision trees, neural networks, associations, and naïve Bayes, to predict diseases with high accuracy [8], [17].

Some classification techniques have shown that decision trees provided considerably good results [8]. Another comparative study evaluated Naïve Byes algorithms as the most accurate ML classifier for disease prediction, followed by neural networks and then decision trees [17]. Artificial neural networks are another ML model that is extensively used for predicting heart disease [12]. A combination of genetic algorithms and neural networks was incorporated in [18] to obtain 99.7% accuracy.

In [5], the authors applied two classification methods—J48 on the Hungarian data set and Naïve Bayes on the echocardiography database [19]—using WEKA to predict cardiac disease. Performance metrics and a confusion matrix were employed to assess the categorization models. The latter dataset [20] consists of 132 records and 12 features, whereas the former dataset contains 14 features with a target variable with 5 values. For feature selection, the authors used J48 and Naïve Bayes algorithms, with all attributes and a set of selected attributes for comparison. A classification accuracy of 82.3% was obtained with all features using the first dataset, outperforming the 65.64% accuracy achieved with a subset of selected attributes. Based on the second dataset, the findings indicate that a classification accuracy of 93.24% with the chosen attributes and 98.64% with all attributes was achieved.

There are multiple factors that impact heart disease, like family history, smoking, poor diets, high blood pressure, high blood cholesterol, obesity, physical inactivity, and hypertension. Using the dataset with the aforementioned attributes, three classification models, including Naïve Byes, Decision Tree, and Neural Network, were used on the dataset, and results show that accuracy is 90.74%, 99.62% and 100%, respectively [14].

In a research article [21], the authors used six machine learning classification algorithms on the Statlog heart disease dataset [20]. The performance evaluation was carried out using tenfold cross-validation. The results indicated that the highest rate of false positives and misclassifications were shown by the decision tree, whereas, the highest specificity and F1 score was achieved by logistic regression.

A genetic algorithm combined with recurrent fuzzy neural networks for the diagnosis of heart disease was proposed. The results showed that accuracy of 97.78% was obtained with the help of this model [22].

Another research work combined genetic algorithms and KNN for disease classification. Experimental results show that accuracy was increased by combining the aforementioned models [13].

In [23], for prediction of heart failure, supervised machine learning methods were applied on the Cleveland heart disease dataset. The authors trained machine learning methods on Orange and WEKA tools using 14 features out of 76 and 303 records in the data set. Firstly, the data was pre-processed to remove missing values. Afterward, they evaluated the performance of four classification methods on the two systems using precision and recall metrics. Although both tools produced encouraging results, higher values of precision and recall were achieved on WEKA than Orange.

Another research utilized support vector machines, Naïve Bayes, logistic regression, K-Nearest Neighbour, Decision Tree, and artificial neural networks. It was shown that this system based on the ensemble of these classifiers provided improved accuracy than the accuracy of individual classifiers. Such systems can easily be implemented in healthcare for the identification of heart disease [24].

A comparative analysis was done in [6] to determine how ensemble technique can be applied to improve the accuracy of the classifiers. This work focused on accuracy as well as on the prediction to determine the heart disease at an early stage of life. Experimental results show that 7% accuracy can be increased by combining weak classifiers.

The authors, in [25], conduct an experimental investigation into real-world problems using the WEKA implementations of machine learning algorithms. The research primarily focused on classification tasks and the comparative evaluation of the relative performance of various algorithms under specific criteria. The main issue addressed in this work was the prediction of rainfall, specifically determining whether it would rain or not. To this end, several classifiers were evaluated and compared, including TreesJ48, TreesJ48graft, Random Tree, OneR, ZeroR, Decision Table, Naïve Bayes, Bayes Net, Naïve Bayes Simple, Bayes Updatable, Multilayer Perceptron, Logistic, RBF Network, and Simple Logistic. The study aimed to provide insights into the effectiveness of these algorithms in addressing rainfall prediction while also highlighting their strengths and limitations under the defined evaluation framework. The multilayer perceptron showed the best performance with ~80 percent accuracy.

There were several other studies where the results and performance of various tools were compared, but in the context of shoppers' intentions and botnet attack identification. However, our work is evaluating the performance of various tools, but for identification of heart disease and dataset feature importance.

In [19], the study focused on classifying online shoppers' spending intentions using machine learning algorithms such as Naïve Bayes, Multi-Layer

Perceptron, Support Vector Machine, Random Forest, and J48 Decision Trees. The algorithms were evaluated and compared using WEKA and Scikit-learn, with performance metrics including F1-score, accuracy, Kappa statistics, and mean absolute error. Discrepancies were noted between the tools, particularly for the Support Vector Machine algorithm. The results demonstrated that Random Forest outperformed the others, making it the most suitable algorithm for classifying online shoppers' intentions.

The authors, in [26], focused on botnet attack identification utilizing different tools and platforms. To identify whether packet traffic constitutes a malware attack, machine learning classification methods were employed. This study implemented four algorithms—Ada Boost, Decision Tree, Random Forest, and Naïve Bayes—using the WEKA and Scikit-learn analysis tools. Experiments were conducted to evaluate the performance of these algorithms based on accuracy, execution time, and false positive rate (FPR). The results indicated that the WEKA tool provided more accurate and efficient classifications. However, Scikit-learn demonstrated superior performance in terms of false positive rate, offering better results in this specific metric.

The authors worked on brain MRI scans for disease identification in [27]. They evaluated the performance of machine learning tools such as WEKA and Python, particularly in conjunction with appropriate filters and classifiers, which have become essential for advancing medical imaging. Previous studies have highlighted the effectiveness of these tools in image classification tasks, though a comparative analysis specific to brain tumour MRI scans remains an area of interest. This study builds existing research by systematically evaluating and deploying the most effective machine learning model to brain tumour classification, aiming to enhance diagnostic outcomes and streamline clinical workflows.

3. Dataset Description and Exploratory Data Analysis

The dataset was obtained from Kaggle “heart disease” [28]. This dataset has 319795 instances. Exploratory data analysis (EDA) was done by using both WEKA and Scikit-Learn to analyze different factors that contribute to heart disease and stroke.

3.1 Data Description

Data analysis was performed by identifying attributes that are used as predictors of heart disease in this work. Therefore, before building any classifier in Scikit-Learn the critical step is to identify data types of predictors and target features. There are a total of 18 features in the dataset 13 have data type objects, while 5 features are float64.

Null values can create problems while selecting and processing the data. The main issue with null values is that it affects the performance with forward lookup. They are removed by careful examination in Scikit-Learn. If there were missing values in a sample, the sample was removed from the dataset. As being a medical record of people, replacing missing values with commonly used approaches such as missing value replacement by mean or mode could have introduced bias in the dataset. The exploratory data analysis also revealed that it is a highly unbalanced dataset, with fewer samples of Class= ‘Heart Disease Present’. Hence, 10-fold cross-validation was employed for thorough and reliable results.

3.2. Visualization of Dataset and Classification

Exploratory data analysis was done to understand various features of the dataset, their values, spread, and frequency of values as well as correlations between different features of dataset. Heart disease depends on various factors; therefore, Scikit-Learn and WEKA are used to analyze this relationship visually too. Moreover, feature importance is analyzed to identify the most contributing factors in heart disease.

Analysis, using Scikit-learn, is shown in Fig. 1 to Fig. 3. It can be observed that individuals with high body mass index (BMI) also have a risk of heart attacks, while those with lower BMI face reduced risk. It is evident from the figure that people with a high BMI have a higher likelihood of experiencing heart attacks. Physical health is one of the most important factors affecting heart health. If a person is unwell and their illness persists, the risk of heart disease also increases. If an individual is physically fit, they need not worry about heart health. The mental health impact on heart stroke is shown in Fig.2.

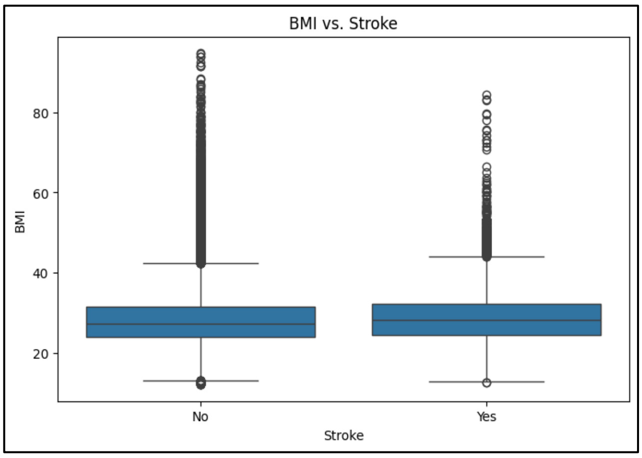

Fig. 1. Impact of BMI on Heart Disease

The occurrence of heart disease also varies depending on gender and age. As age increases, the likelihood of heart attacks also rises. This relationship is directly proportional. Individuals aged 80 or older have the highest risk of heart attacks. However, they can reduce this risk by taking precautionary measures.

Engaging in physical activity keeps your body fit and strengthens your heart over time. The graph illustrates that individuals who are physically active experience fewer heart attacks, whereas those with lower levels of physical activity face a higher risk of heart attacks. Consequently, athletes have fewer heart attacks due to their high level of physical activity.

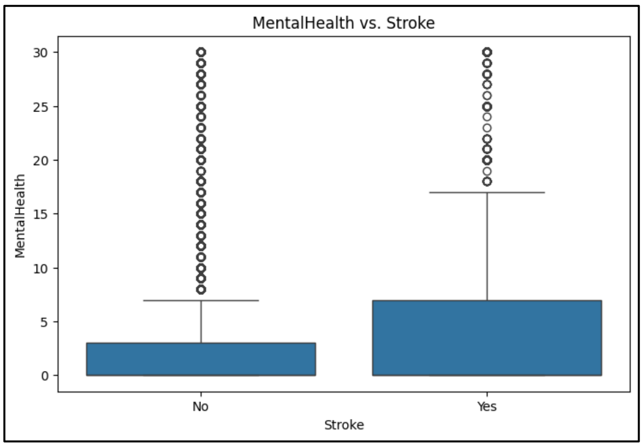

Fig. 2. Role of Mental Health in Heart Disease Prediction

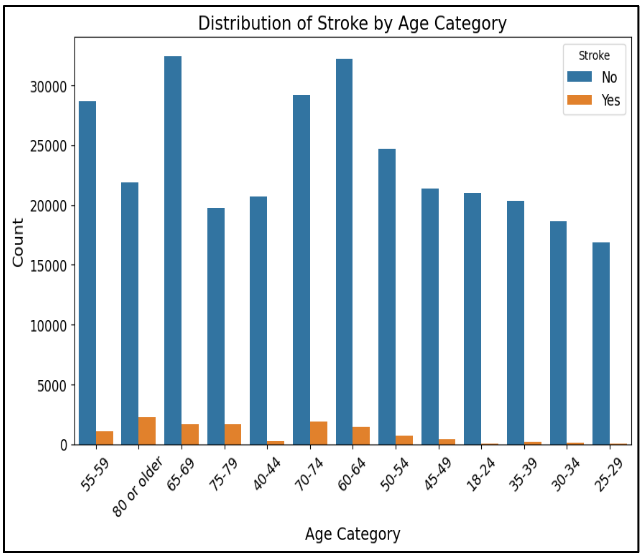

Fig. 3. Impact of Person's Age on Heart Disease Indicating Higher Chances of Heart Disease with Growing Age; Y-Axis Shows the Sample Count of Different Age Brackets with Class Label Yes/No Indicating Heart Disease Presence

Detailed analysis using WEKA showed that the overall health of an individual also significantly influences the likelihood of heart attacks. Poor health increases the risk of heart attacks, whereas good health reduces it. As a person's health improves, it concurrently enhances heart health and reduces the chances of heart attacks.

Fig. 4 shows that individuals with kidney disease are at a higher risk of experiencing heart attacks compared to those without kidney disease.

Similarly, the same holds true for skin cancer. The figure indicates that cancer patients have a higher likelihood of heart attacks compared to non-cancer patients. Fig. 5 clearly illustrates the significant role of smoking in increasing the likelihood of heart attacks. Individuals who smoke are at a higher risk of heart attacks, while non-smokers have better health prospects. Similarly, individuals who consume alcohol have a higher incidence of heart disease compared to non-drinkers.

Fig. 4. Visualization of Relation of Various Features to Heart Disease Using WEKA; Sample Count Per Feature Along Y-Axis with Colour Coded Heart Disease Class Label, Yes in Blue and No in Red

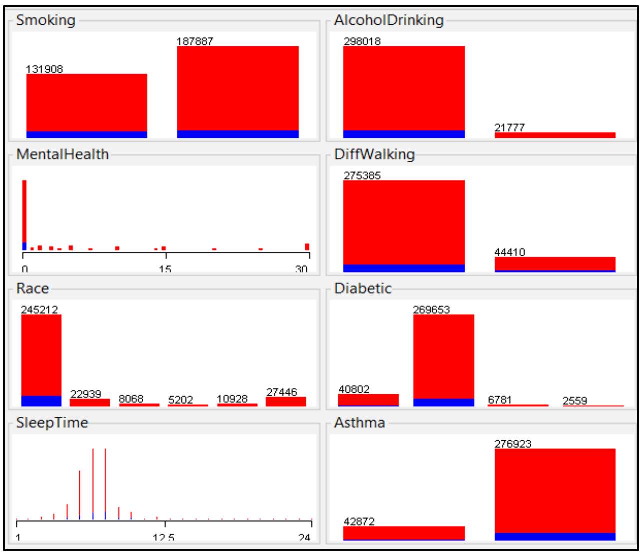

Fig. 5. Impact of Smoking, Drinking, Ethnicity, Sleep and Other Disease Presence On Heart Disease Using WEKA; Sample Count Per Feature Along Y-Axis with Colour Coded Class Label Yes in Blue and No in Red

Regular physical activity, particularly walking, is essential for maintaining physical and mental fitness and combating heart diseases, as shown in Fig. 5. Individuals who engage in regular walking are less likely to experience heart disease as their overall fitness and heart strength improve. Thus, there exists an inverse relationship between walking and heart disease. When considering the relationship between race and heart disease, it is surprising to note that both white individuals and American Indians have the highest risk compared to other racial groups.

Sleeping patterns play a significant role in heart attack risk in an intriguing manner. Fig. 5 indicates that sleeping for 7 hours is optimal. Sleeping fewer or more than 7 hours increases the likelihood of heart attacks.

Analysis of the dataset using WEKA has revealed that individuals with asthma are more likely to experience heart attacks compared to those without the condition.

Fig. 5 illustrates that individuals with diabetes have a higher risk of heart attacks. Conversely, those without diabetes have a lower percentage of heart attacks.

4. Methodology

The following approach was utilized to evaluate the impact of computing platforms on classifier performance using heart disease prediction as use case scenario.

For the experimental evaluation, we have used Scikit-Learn in Google Colab and WEKA v3.8.6, both of which are frequently used for ML tasks. Multiple classification models were trained and tested on these tools to generate results. After applying the techniques, results were compared based on accuracy. On both platforms, the same pre-processing and data analysis tasks were performed, and commonly used classifiers for disease prediction such as logistic regression, k-nearest neighbours, support vector machines, Naïve Bayes, and decision trees were employed, having exactly the same hyperparameters. That way, the performance measures generated by each classifier are comparable, as except the platform, all other conditions are kept the same. Fig. 6 summarizes the steps graphically for a quick overview.

10-fold cross validation was employed to obtain results, and averaged results over the 10 folds were reported. The details about hyperparameters of various models and the training-testing split are described in Table 1. Afterward, the performance measures, namely accuracy, were compared to what was achieved in Google Colab vs. WEKA by each and every classifier. Interesting results reveal and verify the impact of platform on the final accuracy achieved when everything, including the classifier, the hyperparameters, and the seed for the pseudo random number generator, were kept exactly the same.

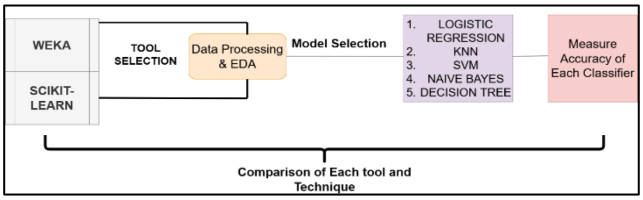

Fig. 6. Methodology

Table 1

Summarization of Hyper-Parameters of various models

Classifiers (10-fold CV)	Hyper-Parameters
	Kept Same in Weka and Colab (Scikit-Learn)
Logistic Regression	Optimization algorithm 'lbfgs' (Limited-memory Broyden–Fletcher–Goldfarb–Shanno), L2 regularization, maximum iterations allowed =1000
K Nearest Neighbor	K = 23 Distance Measure = Euclidean Weight to Neighbors = equal weight
Decision Tree	criterion='gini', splitter='best', max_depth=None, min_samples_split=2, min_samples_leaf=1
Naïve Bayes	priors=None, var_smoothing=1e-09
Support Vector Machine	Kernel = ‘rbf’ (Radial Basis Function), C=1.0, degree=3, gamma='scale'

5. Results And Discussions

WEKA and Scikit-Learn allow us to compare the accuracy of several classifiers. Table 2 makes it evident that the support vector machine classifier, when built with Scikit-Learn in our experiment, has the highest accuracy of 76.585% followed by logistic regression with accuracy of 76.55%. Naïve Bayes comes in third place with an accuracy of 74.910%. KNN, with an accuracy of 74.320%, holds the 4th position, while Decision Tree (DT) comes in last with 67.185% accuracy. These represent the top five classifiers, and it is clear that SVM stands out as the most accurate classifier available for this dataset.

WEKA was likewise utilized with the same classifiers and their respective hyperparameters, and the outcomes are significantly better. With a maximum accuracy of 90.67%, LR and support vector machines rank highest, while decision trees come in second with 90.483 percent accuracy. With 87.841 % accuracy, Naïve Bayes and KNN classifiers have 84.606 % accuracy. It is evident that the accuracy of the classifier increases when WEKA is used but decreases when Scikit-Learn is used.

Fig. 7 shows that using WEKA increased the classifier's accuracy, while using Scikit-Learn reduced it. Additionally, it is evident that using WEKA in our trials maximized the accuracy of both SVM and logistic regression.

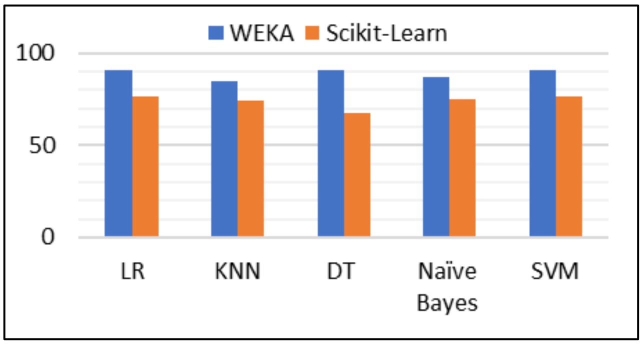

Fig. 7. Pictorial Depiction of Classifier Accuracy (0-100 Percent) on WEKA Vs Google Colab (Scikit-Learn) with same Hyper-Parameters

Table 2

Comparison of classifier performance in WEKA Vs Google Colab (Scikit-Learn) with same hyperparameters

Classifiers	Accuracy	
	WEKA	Scikit-Learn
Logistic Regression	90.675 %	76.555 %
K Nearest Neighbour	84.606 %	74.320 %
Decision Tree	90.483 %	67.185 %
Naïve Bayes	87.841 %	74.910 %
Support Vector Machine	90.675 %	76.585 %

The results of the experiment provide intriguing insights into the performance of different classifiers and both the platforms. The most prominent observation is the significant difference in accuracy between WEKA and Scikit-Learn for the same classifiers and hyperparameters. This discrepancy can be attributed to several factors, including implementation differences, randomness in cross-validation, and algorithm-specific optimizations provided by the platforms.

WEKA, has been optimized for machine learning tasks over many years. Its implementations of algorithms like SVM, Logistic Regression (LR), and Decision Trees (DT) may contain additional optimizations, or internal parameter tuning that are not explicitly visible to the user. In comparison, Scikit-Learn, a Python library, offers further transparency and control over the implementation of algorithms. Nevertheless, this likewise means that the user is in control for ensuring proper pre-processing, hyperparameter tuning, and other optimizations. The lower accuracy in Scikit-Learn could be due to differences in default settings such as scaling. For example, SVM and KNN are sensitive to feature scaling, and if the data is not scaled in Scikit-Learn, it could lead to suboptimal performance. Furthermore, though both platforms use 10-fold cross-validation, the manner in which the folds are split might differ, leading to disparities in results.

The performance of the classifiers varies significantly between the two platforms, but some trends are consistent. Support Vector Machine (SVM) achieved the highest accuracy in WEKA (90.67%), indicating that WEKA's implementation is highly optimized for this dataset. In Scikit-Learn, SVM achieved 76.585%, which is significantly lower, likely due to the lack of proper scaling or suboptimal hyperparameters. Logistic Regression (LR) also performed well in WEKA, achieving 90.67%, tied with SVM, while in Scikit-Learn, it achieved 76.55%. LR is a robust and interpretable algorithm, and its performance is less sensitive to pre-processing compared to SVM. However, the difference in accuracy suggests that WEKA's implementation might include additional regularization or optimization.

Naïve Bayes achieved 87.841% in WEKA and 74.910% in Scikit-Learn. Naïve Bayes is a simple and fast algorithm, but it assumes independence between features, which might not hold true for this dataset. The higher accuracy in WEKA could be due to better handling of continuous features, such as using kernel density estimation instead of Gaussian assumptions. K-Nearest Neighbours (KNN) achieved 84.606% in WEKA and 74.320% in Scikit-Learn. KNN is sensitive to feature scaling and the choice of distance metric, and the lower accuracy in Scikit-Learn suggests that the data might not have been scaled properly, or the value of k (number of neighbours) might not have been optimized. Decision Trees (DT) achieved 90.483% in WEKA and 67.185% in Scikit-Learn. Decision trees are prone to overfitting, and their performance can vary significantly based on hyperparameters like maximum depth and minimum samples per leaf. WEKA's implementation seems to handle these aspects better, possibly by using pruning or other regularization techniques.

WEKA's strengths lie in its user-friendly interface with built-in pre-processing and optimization, highly optimized implementations of algorithms, and suitability for users who want quick results and an out-of-the-box solution without deep technical expertise. However, it has less flexibility and control compared to Scikit-Learn, limited support for deep learning and advanced techniques, and is Java-based, which might not integrate well with Python-based workflows. On the other hand, Scikit-Learn is highly flexible and customizable, integrates well with other Python libraries, and has extensive documentation and community support. However, it requires more manual effort for pre-processing and hyperparameter tuning, and its default implementations might not be as optimized as WEKA's.

6. Conclusion

A thorough examination of the dataset with Weka and Scikit-Learn has demonstrated the dependence of heart disease on multiple parameters. Heart disease problems have been categorized using a number of classifiers that have been used extensively in research. Scikit-Learn is used to build five classifiers using various hyperparameter combinations. On the other hand, the identical classifiers are constructed in WEKA with the same set of hyperparameters, and accuracy is used as the performance metric to assess them.

It is evident from Table 2 and Fig. 7 that the accuracy of various classifiers varied when the underlying platform was changed. This accuracy is partly dependent on the tools we have used for our research. When utilizing WEKA, the highest accuracy is approximately 91% for SVM and LR, and approximately 90.5% for DT. Similar to this, SVM has a maximum accuracy of 76.585% when built utilizing Scikit-Learn, and LR has a maximum accuracy of 76.555%. Looking at Fig. 7, we can conclude that all the classifiers provided higher accuracy when built using WEKA compared to Scikit-Learn.

Therefore, the underlying platform for building ML models plays a crucial role and raises an important question whether reported performance measures in various research works are directly comparable? When the dataset, ML model, and the hyperparameters remain the same, the implementation tool/platform can significantly impact the performance of the classifier.

We conclude that while drawing comparisons of one research's performance measures with another similar work but on a different platform, the direct one-on-one performance measure comparison should be considered with a grain of salt.

7. References

[1] "WHO / cardiovascular diseases (CVDs)", Accessed: Mar. 09, 2024. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)

[2] WHO, "Noncommunicable diseases country profiles 2018", Accessed: May 09, 2024. [Online]. Available: https://www.who.int/publications/i/item/9789241514620

[3] K. Mohtadi et al., "Current Risk Factors of Ischemic Cardiovascular Diseases Estimated in a Representative Population of Casablanca", Endocrinol Metab Syndr, vol. 07, no. 01, 2018, doi: 10.4172/2161-1017.1000284.

[4] B. J. W. Chow et al., "Interpreting Wide-Complex Tachycardia With the Use of Artificial Intelligence," Canadian Journal of Cardiology, vol. 40, no. 10, pp. 1965–1973, Oct. 2024, doi: 10.1016/j.cjca.2024.03.027.

[5] B. Martins, D. Ferreira, C. Neto, A. Abelha, and J. Machado, "Data Mining for Cardiovascular Disease Prediction", J Med Syst, vol. 45, no. 1, p. 6, Jan. 2021, doi: 10.1007/s10916-020-01682-8.

[6] V. K and J. Singaraju, "Decision Support System for Congenital Heart Disease Diagnosis based on Signs and Symptoms using Neural Networks", IJCA, vol. 19, no. 6, pp. 6–12, Apr. 2011, doi: 10.5120/2368-3115.

[7] S. Wang and R. M. Summers, "Machine learning and radiology", Medical Image Analysis, vol. 16, no. 5, pp. 933–951, Jul. 2012, doi: 10.1016/j.media.2012.02.005.

[8] J. Soni, U. Ansari, D. Sharma, and S. Soni, "Predictive Data Mining for Medical Diagnosis: An Overview of Heart Disease Prediction", IJCA, vol. 17, no. 8, pp. 43–48, Mar. 2011, doi: 10.5120/2237-2860.

[9] A. Benba, A. Jilbab, and A. Hammouch, "Discriminating Between Patients With Parkinson's and Neurological Diseases Using Cepstral Analysis", IEEE Trans. Neural Syst. Rehabil. Eng., vol. 24, no. 10, pp. 1100–1108, Oct. 2016, doi: 10.1109/TNSRE.2016.2533582.

[10] I. Kavakiotis, O. Tsave, A. Salifoglou, N. Maglaveras, I. Vlahavas, and I. Chouvarda, "Machine Learning and Data Mining Methods in Diabetes Research", Computational and Structural Biotechnology Journal, vol. 15, pp. 104–116, 2017, doi: 10.1016/j.csbj.2016.12.005.

[11] P. I. Dorado-Díaz, J. Sampedro-Gómez, V. Vicente-Palacios, and P. L. Sánchez,

"Applications of Artificial Intelligence in Cardiology. The Future is Already Here", Revista Española de Cardiología (English Edition), vol. 72, no. 12, pp. 1065–1075, Dec. 2019, doi: 10.1016/j.rec.2019.05.014.

[12] Q. K. Al-Shayea, "Artificial Neural Networks in Medical Diagnosis", vol. 8, no. 2, 2011.

[13] M. A. Jabbar, B. L. Deekshatulu, and P. Chandra, "Classification of Heart Disease Using K- Nearest Neighbor and Genetic Algorithm", Procedia Technology, vol. 10, pp. 85–94, 2013, doi: 10.1016/j.protcy.2013.12.340.

[14] C. S.Dangare and S. S. Apte, "Improved Study of Heart Disease Prediction System using Data Mining Classification Techniques", IJCA, vol. 47, no. 10, pp. 44–48, Jun. 2012, doi: 10.5120/7228-0076.

[15] Y. Xing, J. Wang, Z. Zhao, and andYonghong Gao, "Combination Data Mining Methods with New Medical Data to Predicting Outcome of Coronary Heart Disease", in 2007 International Conference on Convergence Information Technology (ICCIT 2007), Gyeongju-si, Gyeongbuk, Korea: IEEE, Nov. 2007, pp. 868–872. doi: 10.1109/ICCIT.2007.204.

[16] X. Liu et al., "A Hybrid Classification System for Heart Disease Diagnosis Based on the RFRS Method", Computational and Mathematical Methods in Medicine, vol. 2017, pp. 1–11, 2017, doi: 10.1155/2017/8272091.

[17] F. Tasnim and S. U. Habiba, "A Comparative Study on Heart Disease Prediction Using Data Mining Techniques and Feature Selection", in 2021 2nd International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST), DHAKA, Bangladesh: IEEE, Jan. 2021, pp. 338–341. doi: 10.1109/ICREST51555.2021.9331158.

[18] J. Nahar, T. Imam, K. S. Tickle, and Y.-P. P. Chen, "Computational intelligence for heart disease diagnosis: A medical knowledge driven approach", Expert Systems with Applications, vol. 40, no. 1, pp. 96–104, Jan. 2013, doi: 10.1016/j.eswa.2012.07.032.

[19] Y. Christian, "Comparison of Machine Learning Algorithms Using WEKA and Sci-Kit Learn in Classifying Online Shopper Intention", JITE, vol. 3, no. 1, pp. 58–66, Jul. 2019, doi: 10.31289/jite.v3i1.2599.

[20] "UCI Machine Learning Repository. University of California, School of Information and Computer Science, Irvine, CA.", [Online]. Available: https://scirp.org/reference/referencespapers?referenceid=2607575

[21] A. K. Dwivedi, "Performance evaluation of different machine learning techniques for prediction of heart disease", Neural Comput & Applic, vol. 29, no. 10, pp. 685–693, May 2018, doi: 10.1007/s00521-016-2604-1.

[22] K. Uyar and A. İlhan, "Diagnosis of heart disease using genetic algorithm based trained recurrent fuzzy neural networks", Procedia Computer Science, vol. 120, pp. 588–593, 2017, doi: 10.1016/j.procs.2017.11.283.

[23] R. Shishodia and P. Anand, "Analysis Of Heart Disease Using Orange And R", 2019.

[24] J. P. Li, A. U. Haq, S. U. Din, J. Khan, A. Khan, and A. Saboor, "Heart Disease Identification Method Using Machine Learning Classification in E-Healthcare", IEEE Access, vol. 8, pp. 107562–107582, 2020, doi: 10.1109/ACCESS.2020.3001149.

[25] A. Desai and D. S. Rai, "Analysis of Machine Learning Algorithms using Weka", International Journal of Computer Applications.

[26] Susanto, D. Stiawan, M. A. S. Arifin, Mohd. Y. Idris, and R. Budiarto, "IoT Botnet Malware Classification Using Weka Tool and Scikit-learn Machine Learning", in 2020 7th International Conference on Electrical Engineering, Computer Sciences and Informatics (EECSI), Yogyakarta, Indonesia: IEEE, Oct. 2020, pp. 15–20. doi: 10.23919/EECSI50503.2020.9251304.

[27] B. O. Awojoyogbe and M. O. Dada, Digital Molecular Magnetic Resonance Imaging. in Series in BioEngineering. Singapore: Springer Nature Singapore, 2024. doi: 10.1007/978-981-97-6370-2.

[28] https://www.kaggle.com/code/ricksan4ez/heartdisease-visualisation-classification/input