Intelligent Heart Disease Prediction System with Applications in Jordanian Hospitals

Mohammad Subhi Al-Batah[1], Mowafaq Salem Alzboon[2], Raed Alazaidah[3]
Faculty of Science and Information Technology, Jadara University, Irbid, Jordan[1, 2]
Faculty of Information Technology-Department of CS, Zarqa University, Zarqa-Jordan[3]

***Abstract*—Heart disease is the leading cause of mortality worldwide. Early identification and prediction can play a crucial role in preventing and treating it. Based on patient data, machine learning techniques may be used to construct cardiac disease prediction models. This work aims to investigate the usage of machine learning models for heart disease prediction utilizing a publicly available dataset. The dataset contains patient information on clinical and demographic characteristics and the presence or absence of cardiac disease. Based on classification performance, many machine learning methods were tested and compared. The findings reveal that machine learning models can predict cardiac disease with accuracy and AUC values. Furthermore, the developed system is used to examine some Jordanian patients, and the predictions of the results are satisfactory. The study's findings might have far-reaching consequences for the early identification and prevention of heart disease, as well as for improving patient outcomes and lowering healthcare expenditures.**

Keywords*—Heart disease; machine learning; predictive models; classification; clinical data; predictions

I. Introduction

Heart disease is a significant health concern that affects millions of people worldwide. Despite significant advancements in medical science, it remains a leading cause of death globally. Early detection and accurate heart disease prediction are crucial to prevent further complications and improve patient outcomes [1]. Machine learning techniques have shown promising results in several domains such as those in [2-6] and in predicting the occurrence of heart disease and other several diseases allowing for early intervention and better management [7]. Heart disease is a complex condition that involves various risk factors such as age, gender, heredity, smoking, hypertension, obesity, sedentary lifestyle, diabetes mellitus, metabolic syndrome, chronic renal failure, and stress. Identifying these risk factors and their interplay is critical in predicting the onset of heart disease. Traditional methods of heart disease prediction, such as clinical diagnosis, can be time-consuming, labor-intensive, and prone to errors. With the advent of machine learning techniques, accurate and efficient prediction of heart disease is now possible. The motivation behind this study is to develop a machine learning-based system that can effectively predict the occurrence of heart disease [2]. The system should be able to identify the most significant risk factors and their interplay, allowing for early intervention and better management. The study aims to improve upon existing methods of heart disease prediction and provide a reliable and efficient tool for healthcare professionals. The primary research question of this study is: Can machine learning techniques effectively predict the occurrence of heart disease? Specifically, the study aims to answer the following sub-questions: What are the most significant risk factors for heart disease, and how do they interplay?

To answer this sub-question, the study will comprehensively analyze various risk factors associated with heart disease. For instance, the study may use data from electronic health records (EHRs) containing information on patients' demographics, medical history, lifestyle factors, and laboratory results [8]. The study may use statistical techniques such as logistic regression or correlation analysis to identify the most significant risk factors and their interplay [9]. For example, the study may find that smoking and hypertension are strongly associated with heart disease, and their co-occurrence increases the risk of heart disease even further. Next, which machine learning algorithms are most effective in predicting heart disease occurrence? To answer this sub-question, the study will evaluate the performance of various machine learning algorithms such as decision trees, random forests, support vector machines, and neural networks [10]. The study may use a dataset of patients with and without heart disease and train the models to predict the occurrence of heart disease. The study may use performance metrics such as accuracy, Precision, recall, and AUC to evaluate the models' performance [11].

The rest of the paper is organized as follows. Section II presents the relevant works with the subject under consideration. Additionally, in Section III, dataset description and analysis of the methodology are provided. Then, Section IV discusses the acquired research results. Finally, conclusions and future directions are outlined in Section V.

II. Literature Review

To increase the accuracy of weak algorithms, an ensemble voting-based model combining many classifiers was presented. The proposed ensemble approach's power is appealing in increasing anemic classifier prognosis accuracy and establishing suitable performance in analyzing the risk of heart disease. An ensemble voting-based approach was used to achieve a remarkable gain in accuracy of 2.1 percent for anemic classifiers [12].This study aims to create a model that can accurately forecast cardiovascular disorders to lessen the number of people who die from them. The suggested model was applied to a real-world dataset of 70,000 Kaggle instances and achieved the following accuracy: XGBoost: 86.87 percent

(with cross-validation), random forest: 87.05 percent (with cross-validation), multilayer perceptron: 87.28 percent (with cross-validation), and 86.94 percent (with cross-validation) (without cross-validation). AUC values for the suggested models are 0.94 for XGBoost, 0.95 for the random forest, and 0.95 for multilayer perceptron. According to the findings of this study, the multilayer perceptron with cross-validation surpassed all other algorithms in terms of accuracy, with the greatest accuracy of 87.28 percent [13].

Based on Machine Learning methods, this study provides an efficient and accurate solution for diagnosing cardiac disease. Several cutting-edge Machine Learning methods are used to classify a cardiovascular dataset. Machine Learning methods such as Random Forest, Nave Bayes, and SVM are used to introduce the prediction model. The prediction model is intended to provide improved performance with high accuracy [11].This article examines the numerous machine learning algorithms utilized to accurately predict, diagnose, and treat various cardiac illnesses. The findings revealed that ANN had the highest average prediction accuracy (86.91 percent), whereas the C4.5 decision tree approach had the lowest average (74.0 percent). For automatic prediction, diagnosis, and treatment of heart disease, machine learning algorithms and techniques have been applied to several accessible heart disease datasets. The findings revealed that ANN had the highest average prediction accuracy (86.91 percent), whereas the C4.5 decision tree approach had the lowest average prediction accuracy (74.0 percent) [9]. Cardiovascular disease is a potentially fatal condition that has become more widespread in recent decades. Machine Learning tools and methodologies are employed to treat and diagnose this condition correctly. This study provides a survey of numerous models that accept such approaches and algorithms, as well as an analysis of their performance. Random Forest (RF), Decision Tree (DT), Naive Bayes, ensemble models, K-Nearest Neighbor (kNN), and Support Vector Machine are some common models (SVM) [14]. The heart is the human body's next main organ, and data analytics is utilized to anticipate the incidence of cardiac illnesses. To forecast the development of cardiac disorders, data mining and machine learning techniques such as Artificial Neural Network (ANN), Decision Tree, Fuzzy Logic, K-Nearest Neighbour (kNN), Nave Bayes, and Support Vector Machine (SVM) are utilized. This document includes an overview of known algorithms as well as a summary of previous work [14].According to the World Health Organization, heart disease kills 33 percent of the world's population. To address this, a UCI Machine Learning Repository dataset was analyzed and predicted heart disease classes. The key characteristics of each assembling technique were retrieved, filtered from the dataset, fitted to a logistic regression classifier, feature scaling, and fitted using logistic regression. Mean Squared Error (MSE), Mean Absolute Error (MAE), R2 Score, Explained Variance Score (EVS), and Mean Squared Log Error (MSLE) were used to analyze performance (MSLE). The feature significance retrieved from the AdaBoost classifier was successful before adding feature scaling, with an MSE of 0.04, MAE of 0.07, R2 Score of 92 percent, EVS of 0.86, and MSLE of 0.16. The feature significance retrieved from the AdaBoost classifier was successful after feature scaling, with an MSE of 0.09, MAE of 0.13, R2 Score of 91 percent, EVS of 0.93, and MSLE of 0.18 [15].

This research provides a machine learning framework that uses five algorithms to predict the likelihood of developing heart disease: Random Forest, Naive Bayes, Support Vector Machine, Hoeffding Decision Tree, and Logistic Model Tree (LMT). The Cleveland dataset is utilized for training and testing, and the findings demonstrate that Random Forest performs the best [16].The essential information in this book is that machine learning and deep learning techniques are utilized to predict the range of risks connected with this project. A dataset was constructed by integrating previously accessible datasets and categorizing them into eleven groups. In diagnosing cardiovascular disorders, machine learning techniques outperformed deep learning, and the PCA methodology was used to determine the relative significance of each of the dataset's 11 fields. Random Forest Classifiers, Decision Tree Classifiers, and Naive Bayes algorithms surpassed other MI algorithms regarding accuracy and recalled. In undeveloped, developing, and even industrialized nations, heart disease is the leading cause of mortality. This disease's death rate can be reduced if detected early and accurately predicted. Machine Learning is critical in predicting and preserving key data regarding cardiac illnesses. This paper examines numerous research studies that use datasets to use machine learning in the prediction of cardiac ailments [17]. This will assist medical professionals in taking corrective action.

Cardiovascular diseases (CVDs) are the world's leading cause of sudden mortality today, and valid, accurate, and practical ways to identify them are required. Machine learning algorithms (MLAs) have been created and proven useful and efficient in forecasting CVD issues based on historical data. This thesis proposes a novel methodology that focuses on finding appropriate features by using MLA techniques such as Deep Learning, Random Forest, Generalized Linear Model, Naive Bayes, Logistic Regression, Decision Tree, Gradient Boosted trees, Support Vector Machine, Vote, and HRFLM, with higher accuracy levels of 75.8 percent, 85.1 percent, 82.9 percent, 87.4 percent, 85 percent, 86.1 percent, 78.3 percent, 86.1 percent, [18]. This research investigates the differences in performance of multiple machine learning models on chronic renal disease and cardiovascular disease datasets using Principal Component Analysis dimensionality reduction approaches. Logistic Regression, K Nearest Neighbor, Naive Bayes, Support Vector Machine, and Random Forest Model were utilized to assess the models' performance with and without PCA. The authors discovered that the kNN classifier and logistic regression were the best approaches for predicting renal and heart illness, with 100% accuracy in chronic kidney disease and 85% in heart disease [19]. This research investigates the differences in performance of multiple machine learning models utilizing Principal Component Analysis dimensionality reduction approaches on Chronic Kidney and Cardiovascular Disease datasets. Logistic Regression, K Nearest Neighbor, Nave Bayes, Support Vector Machine, and Random Forest Model were used to examine the performance of the models with and without PCA. The

scientists discovered that the kNN classifier and logistic regression were the best approaches for predicting renal and heart illness, with 100 percent accuracy in chronic kidney disease and 85 percent in heart disease, respectively [11].

Machine Learning is the study of how computers can learn to discover answers without being explicitly programmed. It is a subset of Artificial Intelligence that allows practitioners to identify illnesses more quickly and efficiently. Machine learning enables the creation of models that correlate various characteristics with an illness. However, good analytic tools are scarce for uncovering underlying correlations and patterns in data. Machine learning is becoming more prevalent in the diagnosis field due to the advancement of classification and recognition algorithms in disease categorization [20]. Data analysis is an essential element of healthcare since it allows for extracting hidden information and predicting disease. This research evaluates several machine learning approaches to determine the superiority of the Random forest algorithm in predicting heart disease [21].

III. Methodology

The primary objective of this study is to identify an effective and predictive algorithm for the early detection of heart disease. To this end, we applied various machine learning models, including Logistic Regression (LR), Decision Tree, Neural Network, Naive Bayes (NB), the k-nearest neighbors (kNN), Support vector machines (SVM), AdaBoost (AB), Stochastic Gradient Descent (SGD), CN2 rule inducer, Constant and Random Forest (RF), to the Heart Disease Diagnostic dataset and evaluated the results. Fig. 1 outlines the planned architecture in detail.

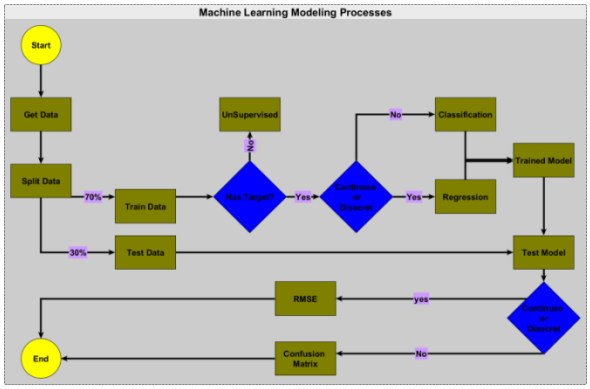

Fig. 1. Machine learning modeling processes flow diagram.

Our methodology begins with data collection, followed by preprocessing, which consists of four steps: data cleansing, attribute selection, target role setting, and feature extraction. Machine learning algorithms that can predict heart disease for a fresh set of measures are developed using the supplied data. To test the performance of an algorithm, we present the model with labeled new data. This is often accomplished by dividing the obtained labeled data into two sections using the Train test split function. Seventy-five percent of the data is referred to as the training data or training set and is utilized to construct our machine learning model. 25% of the data will be utilized to evaluate the model's performance, referred to as test data or test set. After evaluating the models, we compare the acquired data to select the algorithm with the highest accuracy and discover the most predictive algorithm for heart disease detection in Fig. 2.

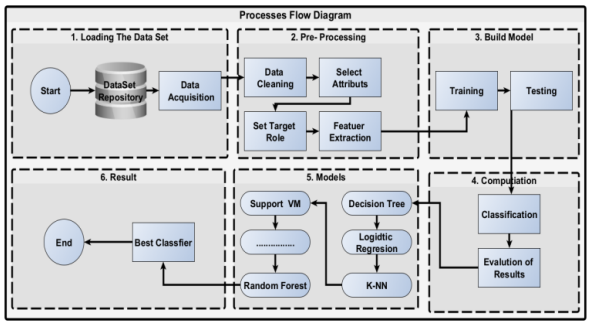

Fig. 2. Processes flow diagram.

A. Machine Learning Algorithms

The predictive analysis of machine learning algorithms is accomplished in our study. These are the machine learning algorithms used in our project.

Logistic regression is a statistical model used to estimate the probability of a binary outcome (such as yes/no) based on one or more independent variables. Logistic regression aims to determine the model that best represents the connection between the input features and the dependent variable. Logistic regression uses a logistic function to represent the relationship between the input features and the target variable, allowing for the calculation of the likelihood that the target variable would assume a specific value given the input information [22].

A decision tree model creates a tree-like structure by recursively dividing the data based on the input attributes. Each node in the tree indicates a determination based on a certain characteristic, while the leaves represent the final classification. Decision tree models are simple to read and depict, making them valuable for comprehending the model's decision-making process [23]. However, decision trees might be susceptible to overfitting if the tree is too complicated or if the data contains noise.

Neural Network: Neural networks are simulations of the structure and operation of the human brain. They are composed of layers of interconnected nodes (neurons) that identify data patterns and make predictions based on those patterns. Neural networks can learn intricate correlations between the input data and the target variable [24]. Neural networks can be challenging to comprehend and prone to overfitting if the model is too complicated or if there is insufficient data to train the model.

The Naive Bayes model calculates the probability of a certain class given the input features. It assumes that the characteristics are conditionally independent of one another, which simplifies the probability computation [25]. Naive Bayes is a simple and quick model that works well with tiny datasets, although, in some instances, it may not be as accurate as other models.

kNN is a non-parametric model that classifies new data points according to the class labels of the k nearest neighbours

in the training set. The value of k controls how many neighbours should be considered [26]. kNN is a simple and intuitive model that can perform well for low-dimensional data. Still, it can be computationally costly for high-dimensional data and may not perform well if the input is noisy or contains irrelevant features.

Support vector machines (SVM) are a model that locates a hyperplane that maximizes the difference between two classes in the data. Translating the input characteristics into a higher-dimensional space can handle both linearly and non-linearly separable data. SVM is a potent model that works well with high-dimensional data, but it is computationally intensive and may not scale well for huge datasets [27]. AdaBoost is a model that combines multiple weak classifiers to produce a powerful classifier. The final classification is based on the weighted combination of all the weak classifiers. AdaBoost is a potent model that can increase the accuracy of poor classifiers, but it is susceptible to noisy data and outliers [28]. SGD: Stochastic gradient descent (SGD) is a model that iteratively updates the model parameters using the gradient of the loss function for the model parameters. It modifies the parameters according to the gradients of the loss function to the parameters determined on a subset of the data (a mini batch). SGD is a quick and effective model that works well with large datasets but may require hyperparameter optimization for optimal performance [29]. CN2 is a model that learns decision rules from the input data. It employs a greedy search technique to determine the optimal collection of rules covering most situations while minimizing the number of rules. CN2 is a basic and interpretable model that works well with small datasets, but in some instances, it may not be as precise as other models [30]. Constant: The constant model is a model that always predicts the same class label for all instances. It is used as a baseline for comparison with other models and can be used to determine if a more complex model is necessary constant: The constant model predicts the same class label for all instances. It is used as a benchmark for comparing other models and to decide whether a more complex model is required [31]. Random Forest is a model that generates several decision trees based on random subsets of data and input attributes. The vote of most of all decision trees determines the highest classification. Random forest is a robust model that can handle high-dimensional data and noisy or irrelevant features. In addition, it is less susceptible to overfitting than decision trees. Random forest models can be challenging to interpret and computationally costly for large datasets [32]. Overall, each model has advantages and disadvantages, and the selection of a model depends on the nature of the problem being addressed and the data features [33]. It is essential to carefully pick and analyze the most suitable model for the work to get the highest potential performance.

B. Dataset Acquisition

Data Collection and Preprocessing: Data collection entails getting the dataset from a trustworthy source, such as a medical research database or a publicly accessible dataset repository [34]. It is critical to ensure that the dataset is both relevant to the study issue and of acceptable quality. Data preparation is preparing a dataset for analysis by identifying missing values, outliers, and inconsistencies. Missing values can be managed by imputation or removal, while outliers can be recognized and corrected via winsorization or trimming using statistical approaches or visualization techniques [35]. Inconsistencies can be fixed by looking for data input mistakes or integrating datasets. The models' performance on independent data can be evaluated by dividing the data into training and test sets. The training set trains models, while the test sets assess their performance [36]. It's critical to ensure the split is random and that the test set is representative of the entire dataset.

Extraction and Selection of Features: Categorical characteristics, such as gender or smoking status, have a restricted number of values. One-hot encoding is a method for representing category information as binary variables that may be fed into machine learning models [37]. Numeric characteristics are variables with continuous values, such as blood pressure or age. Scaling is a strategy to guarantee that the numeric characteristics have a consistent scale, preventing greater values from influencing the models [38]. A prominent scaling approach is standard scaling, which changes the data to have a mean of 0 and a standard deviation of 1.

Finding the most relevant characteristics for the classification task is known as feature selection. Correlation analysis may be used to find traits strongly connected to the target variable [39]. The feature importance ranking method may rank characteristics according to their relevance for the classification task.

Algorithms and Techniques for Machine Learning:Decision trees are a common machine-learning approach that builds a tree-like model of decisions and their potential outcomes [40]. Random forests are an ensemble approach that aggregates the forecasts of numerous decision trees.

Logistic regression is a machine learning approach that uses a logistic function to estimate the likelihood of a binary outcome. Support vector machines are machine learning algorithms that create a hyperplane to divide data into two classes. The K-nearest neighbours' method classifies data items based on the class of their k nearest neighbours [41]. Neural networks are a machine learning approach that uses a network of linked neurons to represent the connection between features and the goal variable.

Hyperparameter tuning entails picking the optimum values for the model parameters to maximise the model's performance on the data. This can be accomplished usinggrid search, random search, or Bayesian optimization techniques. Ensemble approaches integrate numerous models' predictions to increase their performance [42]. AdaBoost is an ensemble approach combining many weak models to get a stronger one. Bagging is an ensemble approach that uses bootstrap sampling to construct numerous models and then combines their predictions.

Metrics for Model Evaluation and Performance: Accurate is the proportion of correctly identified samples with the total number of samples. Precision is defined as the fraction of genuine positive samples among all positive samples. The

proportion of real positive samples out of the total number of positive samples is referred to as recall [43]. The harmonic mean of accuracy and recall is used to get the F1 score [44]. The area under the receiver operating characteristic curve (AUC-ROC score) evaluates the trade-off between genuine and false positive rates.

Cross-validation is a technique used to assess a model's capacity to generalize to new and previously unknown data [45]. It entails dividing the data into folds and testing the model on each fold while training it on the remaining folds [46].Statistical tests may be used to assess the performance of several models and see if there are any significant differences [46]. T-tests or ANOVA (analysis of variance) can be employed to compare the means of two or more groups.

To illustrate the performance of the models, ROC curves and confusion matrices can be employed. ROC curves compare the true positive rate against the false positive rate at various threshold levels [47]. Confusion matrices display the model's true positive, true negative, false positive and false negative predictions.

C. Experimental Environment

All machine learning algorithm tests mentioned in this research were conducted using Orange Data Mining version 3.35.0. The experimental machine was a Lenovo 20FES2FE0E with BIOS version N1GETA9W (1.88), an Intel Core i7-6600U CPU @ 2.60GHz (4 CPUs), and 8192MB of RAM. Windows 11 Pro 64-bit with build number 22621 was the operating system utilized for the experiment. The Heart Disease Diagnostic Datasets were used as the experimental dataset. The information supplied describes a dataset named "Heart Data Set." Here's a quick rundown of the dataset's characteristics:

The dataset has 1025 rows and 14 columns, which contain information about 1025 instances, each with 14 characteristics. The dataset has three categorical variables and ten numerical features, meaning that some features are qualitative (e.g., gender, kind of chest pain) and others are quantitative (e.g., age, resting blood pressure, serum cholesterol). The result variable is a categorical variable with two classes, indicating that the dataset is utilized for binary classification tasks. It isn't easy to interpret the dataset more thoroughly without further information about the target variable. It's worth mentioning that the dataset is popularly known as the Cleveland Heart Disease dataset and is frequently utilized for investigating predictive modeling jobs in the context of heart disease diagnosis.

IV. Results and Discussion

The information presented outlines a random sampling procedure used on a dataset of 1025 occurrences. Here's a quick rundown of the data:

The original dataset included 1025 occurrences, implying that it contained information about 1025 people. The random selection technique chose 75% of the data for inclusion, yielding a sample size of 769 occurrences. The sample is a subset of the original dataset that may be utilized for data analysis and modeling.

The remaining cases following the sampling procedure were not chosen for inclusion in the sample and were thus eliminated. In this example, the sample did not comprise 256 occurrences. It's worth noting that the rejected examples may include useful information that might impact the quality of any analysis or modeling conducted on the sample. As a result, it is critical to carefully analyze the sampling method employed and any potential biases imposed by the sampling process.

The material supplied provides a set of ten machine-learning models that may be utilized for various data analysis and modeling applications. Here's a quick rundown of each model:

Each model has advantages and disadvantages and may be better suited to tasks or datasets than others. Before picking the best model, it is critical to evaluate the problem at hand thoroughly, the nature of the data, and the limits of each model. Furthermore, it is critical to assess the model's performance using proper metrics and to interpret the findings with caution, considering any potential biases or restrictions caused by the data or modeling process as shown in Table I, Table II, and Table III.

TABLE I. Demonstrates the Accuracy Scores of Various Machine Learning Algorithms Applied to a Heart Disease Detection Dataset

Model	AUC	CA	F1	Prec	Recall	MCC	Spec	LogLoss
RF	1	1	1	1	1	1	1	0.023
LR	0.938	0.874	0.873	0.876	0.874	0.749	0.869	0.321
Tree	0.871	0.867	0.867	0.876	0.867	0.744	0.874	4.581
SVM	0.999	0.973	0.973	0.973	0.973	0.945	0.972	0.069
AB	1	1	1	1	1	1	1	0
NN	0.983	0.94	0.94	0.94	0.94	0.88	0.939	0.176
kNN	1	1	1	1	1	1	1	0
NB	0.936	0.867	0.867	0.867	0.867	0.734	0.866	0.39
CN2	1	1	1	1	1	1	1	0.073
SGD	0.81	0.802	0.799	0.836	0.802	0.64	0.817	6.827

TABLE II. Demonstrates the Accuracy Scores of Various Machine Learning Algorithms Applied to a Heart Disease Detection Dataset Target Class: 0

Model	AUC	CA	F1	Prec	Recall	MCC	Spec	LogLoss
RF	1	1	1	1	1	1	1	0.023
LR	0.938	0.874	0.861	0.901	0.825	0.749	0.918	0.321
Tree	0.871	0.867	0.87	0.812	0.937	0.744	0.804	4.581
SVM	0.999	0.973	0.971	0.973	0.97	0.945	0.975	0.069
AB	1	1	1	1	1	1	1	0
NN	0.983	0.94	0.936	0.947	0.926	0.88	0.953	0.176
kNN	1	1	1	1	1	1	1	0
NB	0.936	0.867	0.859	0.866	0.852	0.734	0.881	0.39
CN2	1	1	1	1	1	1	1	0.073
SGD	0.81	0.802	0.821	0.72	0.953	0.64	0.666	6.827

TABLE III. DEMONSTRATES THE ACCURACY SCORES OF VARIOUS MACHINE LEARNING ALGORITHMS APPLIED TO A HEART DISEASE DETECTION DATASET TARGET CLASS: 1

Model	AUC	CA	F1	Prec	Recall	MCC	Spec	LogLoss
RF	1	1	1	1	1	1	1	0.023
LR	0.938	0.874	0.884	0.853	0.918	0.749	0.825	0.321
Tree	0.871	0.867	0.864	0.934	0.804	0.744	0.937	4.581
SVM	0.999	0.973	0.974	0.973	0.975	0.945	0.97	0.069
AB	1	1	1	1	1	1	1	0
NN	0.983	0.94	0.944	0.934	0.953	0.88	0.926	0.176
kNN	1	1	1	1	1	1	1	0
NB	0.936	0.867	0.875	0.868	0.881	0.734	0.852	0.39
CN2	1	1	1	1	1	1	1	0.073
SGD	0.81	0.802	0.78	0.941	0.666	0.64	0.953	6.827

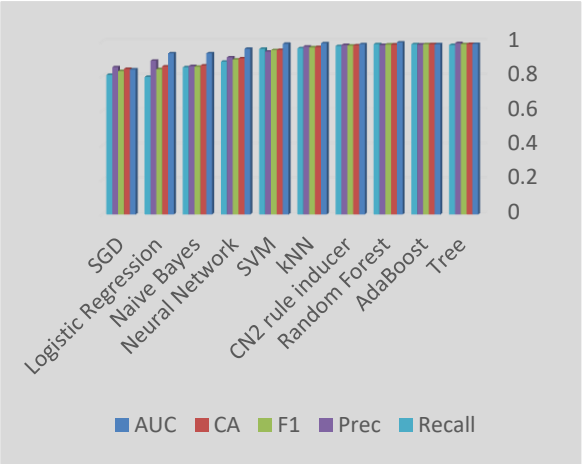

Fig. 3. Test and score, for the target class show: 0.

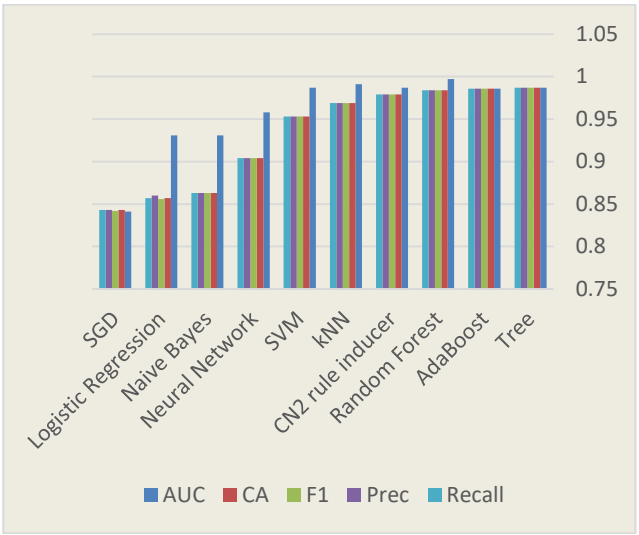

Fig. 4. Test and score, for the target class show average over classes.

The performance evaluation criteria for ten distinct machine learning models: interpretation and evaluation: The Random Forest model received excellent scores across all assessment measures, suggesting it did exceptionally well on the classification test. Fig. 4 shows test and score for the Target class Show: 0 whereas Fig. 3 shows the test and score for the Target class show average over classes. However, it is important to note that the model may have been overfitting the data because it earned excellent scores on the training data. More testing on independent test data is required to demonstrate that the model generalizes successfully to new and previously unknown data. Logistic Regression: The Logistic Regression model performed well in most assessment measures, indicating that it fits the classification problem well. However, several criteria, such as Specificity and LogLoss, fell short, indicating that there may be space for development in these areas.Tree: The Tree model scored poorer than other models across all assessment measures, notably LogLoss. This suggests it may not be the most appropriate model for this classification problem. SVM: The SVM model received excellent scores in all assessment parameters, including AUC, F1, Precision, Recall, MCC, and Specificity. This suggests that it might be a good model for the classification problem. The AdaBoost model received perfect scores across all assessment measures, suggesting it did exceptionally well on the classification test. However, like with the Random Forest model, it must be evaluated on independent test data to ensure it generalizes effectively to new and unknown data. The Neural Network model scored moderate to high in most assessment measures but fell short in others, such as Specificity and LogLoss. This suggests that these areas may have space for improvement. The kNN model received perfect scores on all assessment criteria, suggesting it did exceptionally well on the classification test. However, like with the Random Forest and AdaBoost models, the model must be evaluated on independent test data to ensure it generalizes effectively to new and unknown data. Finally, the SGD model had the lowest scores across all assessment measures, suggesting that it may not be appropriate for this classification job, as shown in Fig. 5.

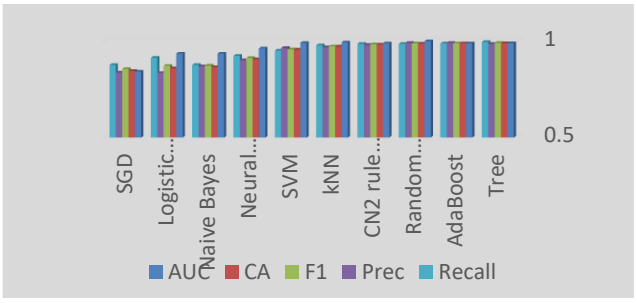

Fig. 5. Test and score, for the target class show: 1.

The information presented depicts the performance assessment metrics for 10 distinct machine learning models based on five evaluation metrics: AUC, CA, F1, Precision, and Recall. The models are assessed specifically on their ability to categorize data correctly and reliably as shown in Table IV.

According to the evaluation findings, the top-performing models are the Random Forest, AdaBoost, Tree, CN2 rule inducer, and kNN models, which received high scores across all assessment measures. The Random Forest model had the greatest AUC score, suggesting it has the best overall performance in rating the data. The Tree, AdaBoost, and CN2 rule inducer models all received perfect scores across all

assessment measures, suggesting they did exceptionally well on the classification test. The kNN model performed well across all assessment measures, indicating that it is good for the classification job.

TABLE IV. DEMONSTRATES THE ACCURACY SCORES OF VARIOUS MACHINE LEARNING ALGORITHMS APPLIED TO A HEART DISEASE DETECTION DATASET SHOW THE AVERAGE OVER CLASSES

Model	AUC	CA	F1	Prec	Recall
Tree	0.987	0.987	0.987	0.987	0.987
AB	0.986	0.986	0.986	0.986	0.986
RF	0.997	0.984	0.984	0.984	0.984
CN2	0.987	0.979	0.979	0.979	0.979
kNN	0.991	0.969	0.969	0.969	0.969
SVM	0.987	0.953	0.953	0.953	0.953
NN	0.958	0.904	0.904	0.904	0.904
NB	0.931	0.863	0.863	0.863	0.863
LR	0.931	0.857	0.856	0.86	0.857
SGD	0.841	0.843	0.842	0.843	0.843

The SVM model performed well in AUC, F1, Precision, and Recall, indicating that it might be a good model for the classification problem. However, it received a lower CA score, suggesting it may be less accurate in categorizing data than other models.

All assessment criteria gave the Neural Network model good ratings, indicating that it might be upgraded to increase its performance. Similarly, the Naive Bayes and Logistic Regression models performed well in most assessment criteria. Still, they performed poorly in others, such as Precision and Recall, showing space for development in these areas.

Finally, the SGD model had the lowest scores across all assessment measures, suggesting that it may not be appropriate for this classification job. The findings show that the Random Forest, AdaBoost, Tree, CN2 rule inducer, and kNN models are the best at this classification job. However, these models must be evaluated on independent test data to verify that they generalize effectively to new and unknown variables. Furthermore, it is critical to analyze the categorization task's unique goals and restrictions and choose the model that best satisfies those demands.

The information presented depicts the performance assessment metrics for 10 distinct machine learning models based on five evaluation metrics: AUC, CA, F1, Precision, and Recall. The models are assessed based on their ability to identify data appropriately and reliably. According to the assessment findings, the top-performing models are the Tree, Random Forest, AdaBoost, CN2 rule inducer, and kNN models, which received high scores across all evaluation measures. The Tree model had the greatest AUC score, suggesting it has the best overall performance in rating the data. All assessment metrics showed that the Random Forest, AdaBoost, and CN2 rule inducer models performed exceptionally well on the classification test. Except for F1, the kNN model had good scores in all assessment criteria, showing that it is viable for the classification job. The SVM model scored well in AUC, Recall, and F1, indicating that it might be a good model for the classification problem. It did, however, have a lower Precision score, indicating that it may be less exact in correctly identifying data than some of the other models.

TABLE V. DEMONSTRATES THE ACCURACY SCORES OF VARIOUS MACHINE LEARNING ALGORITHMS APPLIED TO A HEART DISEASE DETECTION DATASETTARGET CLASS: 0

Model	AUC	CA	F1	Prec	Recall
Tree	0.987	0.987	0.986	0.992	0.981
AB	0.985	0.986	0.985	0.984	0.986
RF	0.996	0.984	0.984	0.981	0.986
CN2	0.986	0.979	0.978	0.981	0.975
kNN	0.991	0.969	0.967	0.972	0.962
SVM	0.988	0.953	0.951	0.943	0.959
NN	0.959	0.904	0.897	0.91	0.885
NB	0.933	0.863	0.856	0.859	0.852
LR	0.933	0.857	0.841	0.89	0.797
SGD	0.84	0.843	0.83	0.853	0.808

All assessment measures yielded moderate to low scores for the Neural Network, Naive Bayes, Logistic Regression, and SGD models, indicating that they may not be the optimal models for this classification problem. The assessment findings show that the top-performing models for this classification job are the Tree, Random Forest, AdaBoost, CN2 rule inducer, and kNN models. However, these models must be evaluated on independent test data to verify that they generalize effectively to new and unknown variables. Furthermore, it is critical to analyze the categorization task's unique goals and restrictions and choose the model that best satisfies those demands as shown in Table V.

Looking at the individual assessment measures, the Tree model had the greatest AUC score, suggesting it performs the best overall regarding data ranking. All assessment measures showed that the Random Forest, AdaBoost, CN2 rule inducer, and kNN models performed exceptionally well on the classification test. The CN2 rule inducer model had the greatest Precision score, meaning it is the most accurate at accurately identifying data. The Naive Bayes and Logistic Regression models have the highest Recall ratings, indicating they are the best at detecting true positives. The Neural Network model had the lowest scores across all assessment measures, implying that it is not the best model for this classification job. It's crucial to note that the performance assessment metrics reported here were calculated using a single random sample of the dataset, and the models' performance may change with various samples or on anonymous data. As a result, it is critical to carefully analyze the sampling method employed and any potential biases imposed by the sampling process. Furthermore, the quality of the data used to train and assess the models is important to their performance, and without knowing more about the dataset's origin, collection, and preprocessing, it's difficult to draw definitive conclusions or make suggestions based on this data alone.

The top-performing models for this classification job are the Tree, Random Forest, AdaBoost, CN2 rule inducer, and kNN models, with high scores across all assessment measures. The SVM model is also suitable, with excellent AUC, Recall, and F1 scores. The Naive Bayes and Logistic Regression models are the best at recognizing true positives, while the CN2 rule inducer model is the best at accurately categorizing data. The Neural Network model had the lowest scores across all assessment measures, implying that it is not the best model for this classification job. However, more testing on independent test data is required to demonstrate that these models generalize effectively to new and previously unexplored data. Furthermore, it is critical to analyze the categorization task's unique goals and restrictions and choose the model that best satisfies those demands.

TABLE VI. DEMONSTRATES THE ACCURACY SCORES OF VARIOUS MACHINE LEARNING ALGORITHMS APPLIED TO A HEART DISEASE DETECTION DATASET TARGET CLASS: 1

Model	AUC	CA	F1	Prec	Recall
Tree	0.987	0.987	0.988	0.983	0.993
AB	0.985	0.986	0.986	0.988	0.985
RF	0.996	0.984	0.985	0.988	0.983
CN2	0.986	0.979	0.98	0.978	0.983
kNN	0.991	0.969	0.97	0.966	0.975
SVM	0.988	0.953	0.955	0.962	0.948
NN	0.959	0.904	0.91	0.899	0.921
NB	0.933	0.863	0.871	0.867	0.874
LR	0.933	0.857	0.87	0.833	0.911
SGD	0.84	0.843	0.854	0.835	0.874

The information presented depicts the performance assessment metrics for 10 distinct machine learning models based on five evaluation metrics: AUC, CA, F1, Precision, and Recall. The models are assessed based on their ability to identify data appropriately and reliably. According to the assessment findings, the top-performing models are the Tree, Random Forest, AdaBoost, CN2 rule inducer, and kNN models, which received high scores across all evaluation measures. The Tree model had the greatest AUC score, suggesting it has the best overall performance in rating the data. All assessment metrics showed that the Random Forest, AdaBoost, and CN2 rule inducer models performed exceptionally well on the classification test. Except for Precision, the kNN model received good scores in all assessment criteria, showing that it is viable for the classification job as shown in Table VI.

The SVM model received excellent AUC, Precision, and Recall scores, indicating that it might be a good model for the classification problem. However, it had a lower CA score, indicating that it may be less reliable in accurately categorizing data than other models. All assessment measures yielded moderate to low scores for the Neural Network, Naive Bayes, Logistic Regression, and SGD models, indicating that they may not be the optimal models for this classification problem. The assessment findings show that the top-performing models for this classification job are the Tree, Random Forest, AdaBoost, CN2 rule inducer, and kNN models. However, these models must be evaluated on independent test data to verify that they generalize effectively to new and unknown variables. Furthermore, it is critical to analyze the categorization task's unique goals and restrictions and choose the model that best satisfies those demands.Looking at the individual assessment measures, the Tree model had the greatest AUC score, suggesting it performs the best overall regarding data ranking. All assessment metrics showed that the Random Forest, AdaBoost, and CN2 rule inducer models performed exceptionally well on the classification test. The CN2 rule inducer model had the greatest Precision score, meaning it is the most accurate in accurately identifying data. The Neural Network model has the greatest Recall score, suggesting it is the best at detecting true positives. The Logistic Regression model has the greatest F1 score, suggesting a good mix of Precision and Recall.

The information supplied displays the classification results of 10 different machine-learning algorithms on a dataset of samples designated as malignant or benign as shown in Table VII. The findings are given as confusion matrices, which indicate the number of samples categorized as malignant or benign by the models and the samples' true labels. According to the assessment findings, the top-performing models include the Tree, Random Forest, AdaBoost, kNN, and CN2 models, which obtained high accuracy in accurately categorizing the data.

TABLE VII. SHOWS THE CONFUSION MATRIX OF SEVERAL MACHINE LEARNING ALGORITHMS APPLIED TO A HEART DISEASE DETECTION DATASET

Algorithms	Malignant	Benign	
Tree	358	7	Malignant
	3	401	Benign
Random Forest	360	5	Malignant
	7	397	Benign
Logistic Regression	291	74	Malignant
	36	368	Benign
SVM	350	15	Malignant
	21	383	Benign
AdaBoost	360	5	Malignant
	6	398	Benign
Neural Network	323	42	Malignant
	32	372	Benign
kNN	351	14	Malignant
	10	394	Benign
Naive Bayes	311	54	Malignant
	51	353	Benign
CN2	356	9	Malignant
	7	397	Benign
SGD	295	70	Malignant
	51	353	Benign

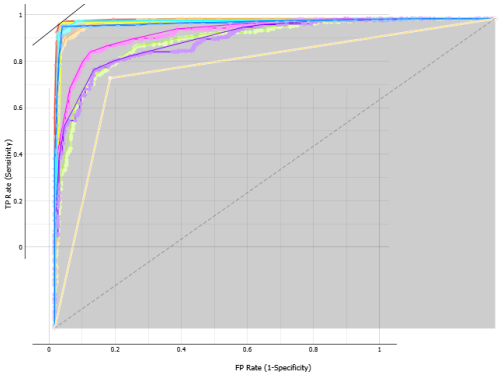

Fig. 6. The ROC curve to the total number of targets 0.

Fig. 6 depicts the ROC (Receiver Operating Characteristic) curve for target 0, a graphical depiction of a binary classification model's performance. The ROC curve is constructed by graphing the true positive rate (TPR) vs. the false positive rate (FPR) at different threshold levels.

The y-axis shows the TPR, also known as sensitivity or recall, which is the fraction of true positive cases accurately detected by the model. The x-axis depicts the FPR, the fraction of true negative cases the model mistakenly classifies as positive.

The ROC curve in the presented Fig. 6 is smooth and steep, indicating a well-performing model. The model's performance improves as the curve approaches the top-left corner of Fig. 6. This figure's curve is in the top-left corner, indicating that the model has a high TPR and a low FPR. In other words, the model can detect real positives while avoiding false positives.

The picture also includes the AUC (Area Under the Curve) score, which is a measure that highlights the overall performance of a binary classification model. The AUC score ranges from 0 to 1, with 1 representing perfect classification accuracy and 0.5 representing random guessing. In this scenario, the AUC value is near one, suggesting that the model performs well in categorizing the data with the goal 0.

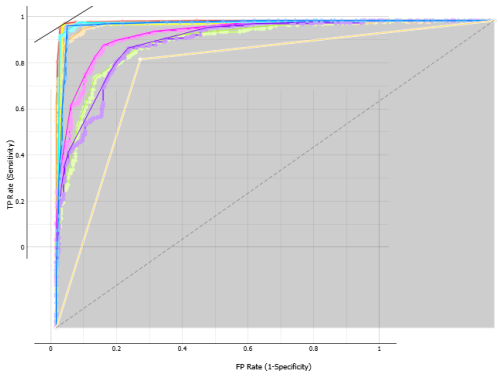

Fig. 7. The ROC curve to the total number of targets 1.

Fig. 7 depicts the ROC (Receiver Operating Characteristic) curve for target 1, a graphical depiction of a binary classification model's performance. The ROC curve is constructed by graphing the true positive rate (TPR) vs. the false positive rate (FPR) at different threshold levels.

V. Conclusion and Future Work

Finally, this work shows the promise of machine learning approaches for predicting heart disease using clinical and demographic data. The findings indicate that various machine-learning techniques may be utilized to create reliable predictive models for heart disease. The study's findings might have significant implications for the early identification and prevention of heart disease, as well as for improving patient outcomes and lowering healthcare expenditures. According to the findings of this study, machine learning models have the potential to be utilized as a tool for healthcare practitioners in the prediction of cardiac disease and the improvement of patient outcomes. These findings have major consequences for healthcare workers, patients, and healthcare systems. Early identification and prevention of cardiac disease can result in better patient outcomes and quality of life, lower healthcare costs, and more efficient resource allocation. Additional research and development are required to successfully integrate machine learning models in Jordanian hospitals. To ensure the generalizability and dependability of the predictive models, it is vital to evaluate the conclusions using separate datasets. Also required is research on the integration of machine learning models into existing healthcare systems and workflows. This includes addressing data quality, privacy, and interpretability issues, as well as designing user-friendly interfaces for healthcare professionals.

References

[1] E. Elbasi, &A. I. Zreikat, A. I.,"Heart Disease Classification for Early Diagnosis based on Adaptive Hoeffding Tree Algorithm in IoMT Data," INTERNATIONAL ARAB JOURNAL OF INFORMATION TECHNOLOGY, 20(1), 38-48. 2023.

[2] M. S. Al-Batah, M. Alzyoud, R. Alazaidah, M. Toubat, H. Alzoubi, &A. Olaiyat, "Early Prediction of Cervical Cancer Using Machine Learning Techniques," Jordanian Journal of Computers and Information Technology, 8(4). 2022.

[3] R. Alazaidah, A. Al-Shaikh, M. R. AL-Mousa, H. Khafajah, G. Samara, M. Alzyoud, ... &S. Almatarneh, "Website Phishing Detection Using Machine Learning Techniques," 2024.

[4] M. Haj Qasem, M. Aljaidi, G. Samara, R. Alazaidah, A. Alsarhan, &M. Alshammari, "An Intelligent Decision Support System Based on Multi Agent Systems for Business Classification Problem,"Sustainability, 15(14), 10977. 2023.

[5] R. Alazaidah, &M. A. Almaiah, "Associative classification in multi-label classification: An investigative study," Jordanian Journal of Computers and Information Technology, 7(2). 2021.

[6] R. Alazaidah, F. Thabtah, &Q. Al-Radaideh, "A multi-label classification approach based on correlations among labels," International Journal of Advanced Computer Science and Applications, 6(2), 52-59. 2015.

[7] K. S. Jagadeesh and R. R, "Heart Disease Prediction Using Machine Learning," Int. J. Res. Appl. Sci. Eng. Technol., 2022, doi: 10.22214/ijraset.2022.40918.

[8] Deb, M. S. Akter Koli, S. B. Akter, and A. A. Chowdhury, "An Outcome Based Analysis on Heart Disease Prediction using Machine Learning Algorithms and Data Mining Approaches," in 2022 IEEE World AI IoT Congress, AIIoT 2022, 2022, pp. 418–424. doi: 10.1109/AIIoT54504.2022.9817194.

[9] L. Riyaz, M. A. Butt, M. Zaman, and O. Ayob, "Heart Disease Prediction Using Machine Learning Techniques: A Quantitative Review," in Advances in Intelligent Systems and Computing, 2022, pp. 81–94. doi: 10.1007/978-981-16-3071-2_8.

[10] D. S. Priyadarsini, "Heart Disease Prediction Using Machine Learning Algorithm," Vol. ISSUE-10, AUGUST 2019, Regul. ISSUE, 2019, doi: 10.35940/ijitee.j9340.0881019.

[11] M. K. S. Ubale and D. P. N. Kalavadekar, "Effective Heart Disease Prediction Using Machine Learning Techniques," null, 2021, doi: null.

[12] Javid, A. K. Z. Alsaedi, and R. Ghazali, "Enhanced accuracy of heart disease prediction using machine learning and recurrent neural networks ensemble majority voting method," Int. J. Adv. Comput. Sci. Appl., vol. 11, no. 3, pp. 540–551, 2020, doi: 10.14569/ijacsa.2020.0110369.

[13] M. Bhatt, P. Patel, T. Ghetia, and P. L. Mazzeo, "Effective Heart Disease Prediction Using Machine Learning Techniques," Algorithms, vol. 16, no. 2, 2023, doi: 10.3390/a16020088.

[14] M. Marimuthu, M. Abinaya, K. S., K. Madhankumar, and V. Pavithra, "A Review on Heart Disease Prediction using Machine Learning and Data Analytics Approach," Int. J. Comput. Appl., vol. 181, no. 18, pp. 20–25, 2018, doi: 10.5120/ijca2018917863.

[15] S. Basheer, R. M. Mathew, and M. S. Devi, "Ensembling Coalesce of Logistic Regression Classifier for Heart Disease Prediction using Machine Learning," Int. J. Innov. Technol. Explor. Eng., doi: 10.35940/ijitee.l3473.1081219.

[16] P. Motarwar, A. Duraphe, G. Suganya, and M. Premalatha, "Cognitive Approach for Heart Disease Prediction using Machine Learning," in International Conference on Emerging Trends in Information Technology and Engineering, ic-ETITE 2020, 2020. doi: 10.1109/ic-ETITE47903.2020.242.

[17] Singh, D. Singh, and J. S. Samagh, "A Comprehensive Review Of Heart Disease Prediction Using Machine Learning," J. Crit. Rev., vol. 7, no. 12, 2020, doi: 10.31838/ jcr.07.12.63.

[18] L. Chandrika and K. Madhavi, "A Hybrid Framework for Heart Disease Prediction Using Machine Learning Algorithms," in E3S Web of Conferences, 2021. doi: 10.1051/e3sconf/202130901043.

[19] S. Mr, "A Detailed Analysis on Kidney and Heart Disease Prediction using Machine Learning," J. Comput. Neurosci., 2021, doi: 10.53759/181x/jcns202101003.

[20] Jayakrishnan, R. Visakh, and K. T. Ratheesh, "Computational Approach for Heart Disease Prediction using Machine Learning," ICCISc 2021 - 2021 Int. Conf. Commun. Control Inf. Sci. Proc., 2021, doi: 10.1109/ICCISc52257.2021.9485014.

[21] M. G, "Heart Disease Prediction Using Machine Learning Algorithms," Biochem. Biophys. Res. Commun., 2020, doi: 10.21786/bbrc/13.11/6.

[22] P. Singh and S. Agrawal, "Accuracy Prediction on Detection of Breast Cancer Using Machine Learning Classifiers," Int. Conf. Comput. Intell. Commun. Networks, 2022, doi: 10.1109/cicn56167.2022.10008345.

[23] M. Abbas and H. Ghous, "Early Detection of Breast Cancer Tumors using Linear Discriminant Analysis Feature Selection with Different Machine Learning Classification Methods," Comput. Sci. Eng. An Int. J., vol. 12, no. 1, pp. 171–186, 2022, doi: 10.5121/cseij.2022.12117.

[24] M. C. Irmak, M. B. H. Tas, S. Turan, and A. Hasiloglu, "Comparative breast cancer detection with artificial neural networks and machine learning methods," in SIU 2021 - 29th IEEE Conference on Signal Processing and Communications Applications, Proceedings, 2021. doi: 10.1109/SIU53274.2021.9477991.

[25] S. S. Olofintuyi, "Breast Cancer Detection With Machine Learning Approach," Fudma J. Sci., 2023, doi: 10.33003/fjs-2023-0702-1392.

[26] M. Mangukiya, "Breast Cancer Detection with Machine Learning," Int. J. Res. Appl. Sci. Eng. Technol., 2022, doi: 10.22214/ijraset.2022. 40204.

[27] M. S. Al-Batah, "Ranked features selection with MSBRG algorithm and rules classifiers for cervical cancer," Int. J. online Biomed. Eng., vol. 15, no. 12, pp. 4–17, 2019, doi: 10.3991/ijoe.v15i12.10803.

[28] R. Alazaidah, G. Samara, S. Almatarneh, M. Hassan, M. Aljaidi, &H. Mansur, "Multi-Label Classification Based on Associations," Applied Sciences, 13(8), 5081. 2023.

[29] M. Al-Batah, B. Zaqaibeh, S. A. Alomari, and M. S. Alzboon, "Gene Microarray Cancer classification using correlation based feature selection algorithm and rules classifiers," Int. J. online Biomed. Eng., vol. 15, no. 8, pp. 62–73, 2019, doi: 10.3991/ijoe.v15i08.10617.

[30] Alshraiedeh, S. Hanna, &R. Alazaidah, "An approach to extend WSDL-based data types specification to enhance web services understandability," International Journal of Advanced Computer Science and Applications, 6(3), 88-98. 2015.

[31] Quteishat, M. Al-Batah, A. Al-Mofleh, and S. H. Alnabelsi, "Cervical cancer diagnostic system using adaptive fuzzy moving k-means algorithm and fuzzy min-max neural network," J. Theor. Appl. Inf. Technol., vol. 57, no. 1, pp. 48–53, 2013.

[32] M. Alluwaici, A. K. Junoh, &R. Alazaidah, New problem transformation method based on the local positive pairwise dependencies among labels," Journal of Information & Knowledge Management, 19(01), 2040017. 2020.

[33] M. S. Al-Batah, A. Zabian, and M. Abdel-Wahed, "Suitable features selection for the HMLP network using circle segments method," Eur. J. Sci. Res., vol. 67, no. 1, pp. 52–65, 2011.

[34] R. Alazaidah, F. K. Ahmad, M. F. M. Mohsen, &A. K. Junoh, "Evaluating conditional and unconditional correlations capturing strategies in multi label classification," Journal of Telecommunication, Electronic and Computer Engineering (JTEC), 10(2-4), 47-51. 2018.

[35] M. S. Al-Batah,"Automatic diagnosis system for heart disorder using ESG peak recognition with ranked features selection,"International Journal of Circuits, Systems and Signal Processing, 13(June), 391–398, 2019.

[36] R. Alazaidah, F. K. Ahmad, &M. F. M. Mohsen, "A comparative analysis between the three main approaches that are being used to solve the problem of multi label classification," International Journal of Soft Computing, 12(4), 218-223. 2017.

[37] M. S. Al-Batah, M. S. Alkhasawneh, L. T. Tay, U. K. Ngah, Hj Lateh, H., and N. A. Mat Isa,"Landslide Occurrence Prediction Using Trainable Cascade Forward Network and Multilayer Perceptron. Mathematical Problems in Engineering,"2015, https://doi.org/10.1155/2015/512158

[38] K. Junoh, F. K. Ahmad, M. F. M. Mohsen, &R. Alazaidah, (2018, April). "Open research directions for multi label learning," In 2018 IEEE Symposium on Computer Applications & Industrial Electronics (ISCAIE) (pp. 125-128). IEEE. 2018.

[39] M. S. Alkhasawneh, U. K. Ngah, L. T.Tay, N. A.Mat Isa, and M. S. Al-Batah,"Determination of important topographic factors for landslide mapping analysis using MLP network,"The Scientific World Journal, 2013, https://doi.org/10.1155/2013/415023

[40] R. Alazaidah, F. K. Ahmad, &M. Mohsin, "Multi label ranking based on positive pairwise correlations among labels," The International Arab Journal of Information Technology, 17(4), 440-449. 2020.

[41] F. karimBaareh, A. Sheta, and M.S. Al-Batah,"Feature based 3D Object Recognition using Artificial Neural Networks,"International Journal of Computer Applications, 44(5), 1–7, 2012, https://doi.org/10.5120/6256-8402

[42] M. S.Al-Batah, "Integrating the Principal Component Analysis with Partial Decision Tree in Microarray Gene Data," IJCSNS International Journal of Computer Science and Network Security, 19(3), 24-29,2019.

[43] M. S. Al-Batah, S. Mrayyen, and M. Alzaqebah, "Arabic Sentiment Classification using MLP Network Hybrid with Naive Bayes Algorithm," Journal of Computer Science Science, 14(8), 1104–1114, 2018.

[44] R. Alazaidah, &F. K. Ahmad, "Trending challenges in multi label classification," International Journal of Advanced Computer Science and Applications, 7(10), 127-131. 2016.

[45] M. Almomani, A.Momani, A.B. Abdelnabi, R. Zaqeba, and M. Al-Batah, "Predicting the corrosion rate of tempered medium carbon steel Using artificial neural network,"Protection of Metals and Physical Chemistry of Surfaces, 58(2), 414-421, 2022.

[46] S.Mrayyen, M. S.Al-Batah, and M.Alzaqebah, "Investigation of Naive Bayes Combined with Multilayered Perceptron for Arabic Sentiment Analysis and Opinion Mining," International Journal of Mathematical Models and Methods in Applied Sciences, 12, 2018.

[47] M. S. Alzboon, M. S. Al-Batah,"Prostate Cancer Detection and Analysis using Advanced Machine Learning," (IJACSA) International Journal of Advanced Computer Science and Applications, 14(8), 388-396, 2023.