Visualizing and Predicting Heart Diseases with an Interactive Dashboard

K. Mathan et al. (2018) proposed a decision tree data mining method with a neural network classifier for the prediction of heart disease. Among the different prediction models, the Gini index and neural networks models produce reliable predictions. A multi-layer perceptron neural networks (MLPNN) is utilized.The calculation depends on the decision trees.The most noteworthy precision accomplished is 86.1% by the equivalent width Gain ratio decision tree.

Youness Khourdifi et al.(2018) proposed a machine learning algorithms optimized by Particle Swarm Optimization and Ant Colony Optimization for predicting and classifying heart disease.They exploited the Fast Correlation-Based Feature Selection (FCBF) method to filter redundant features in order to improve the quality of heart disease classification.Each algorithm worked better in some situations and worse in others. KNearest Neighbour K-NN, and Random Forest RF and Artificial Neural Network MLP are the models likely to work best in the data set used in this study.A maximum classification accuracy of 99.65% using the optimized model proposed by FCBF, PSO and ACO.

S. Mohan et al. (2019) suggested a hybrid machine learning technique to predict heart disease which enhances performance level with an accuracy level of 88.7% through the prediction model for heart disease with the hybrid random forest with a linear model (HRFLM). The severity of the disease is classified based on various methods like K-Nearest Neighbor Algorithm (KNN), Decision Trees (DT), Genetic algorithm (GA), and Naive Bayes (NB). The performance of this method can be estimated from the accuracy in the outcome results based on ECG data.

Abderrahmane Ed-daoudy et al. (2019) proposed a real time heart disease prediction system based on Apache Spark which stands as a strong large scale distributed computing platform that can be used successfully for streaming data event against machine learning through in-memory computations. The system consists of two main sub parts, namely streaming processing and data storage and visualization. The first uses Spark MLlib with Spark streaming and applies a classification model on data events to predict heart disease. The seconds uses Apache Cassandra for storing the large volume of generated data.

Mohammad Shafenoor Amin et al. (2018) proposed data mining techniques that can improve the accuracy of predicting cardiovascular disease. Prediction models were developed using different combination of features, and seven classification techniques: k-NN, Decision Tree, Naive Bayes, Logistic Regression (LR), Support Vector Machine (SVM), Neural Network and Vote (a hybrid technique with Naïve Bayes and Logistic Regression). Experiment results show that the heart disease prediction model developed using the identified significant features and the best-performing data mining technique (i.e. Vote) achieves an accuracy of 87.4% in heart disease prediction.

C Latha et al. (2019) proposed machine learning technique improving the accuracy of prediction of heart disease risk based on ensemble classification techniques, to improve the performance, weak classifiers and ensemble algorithms are used, this work has used ensemble algorithms such as bagging, boosting, voting, and stacking. Some of the techniques used for such prediction problems are the Support Vector Machines (SVM), Neural Networks, Decision Trees, Regression and Naive Bayes classifiers. This work has used ensemble algorithms such as bagging,boosting, voting, and stacking.

Ali A. Samir et al. (2021) used a CNN-jSO approach for the prediction of heart (cardiac) diseases, in which the jSO optimization algorithm is employed to tune those CNN hyperparameters. The performance of the designed system is tested on the PhysioNet heart sound and Kaggle heartbeat sounds datasets. The proposed CNN-jSO is compared with other algorithms and shown to be better than them. The CNN-jSO system was implemented in Python and yielded 97.76% training accuracy and 94.12% testing accuracy.

Ahmed M. Khedir et al. (2021) suggested an efficient Association Rule Mining From Distributed Medical Databases for Predicting Heart Diseases.The proposed association rule mining methodology will be used for heart disease prediction using real heart disease data. These real data exist at different clinics and cannot be moved to a central site. The proposed model protects the patient data privacy and achieves the same results as if the data are moved and joined at a central site. They also validate the extracted association rules from all the data providers using an independent test datasets.

Md Mamun Ali et al.(2021) found that utilizing a dataset on heart disease obtained from Kaggle, three classification methods based on k-nearest neighbor (KNN), decision tree (DT), and random forests (RF) algorithms were used. The RF approach was shown to reach 100% accuracy as well as 100% sensitivity and specificity. As a result, there is a high incidence of instances that are misdiagnosed, a problem that could be solved by building a reliable early-stage cardiac disease prediction system using digitized patient information to enhance clinical decision-making.

R. Valarmathi et al. (2021) proposed a prediction technique to identify cardiac disease that involves hyper parameter adjustment of the Random Forest Classifier and XGBoost Classifier models. The evaluation process makes use of the Z-Alizadeh Sani and Cleveland Heart Disease datasets. Based on the Gaussian process, Bayesian optimization is used to evaluate the algorithm's performance. Various techniques, including Grid Search, Randomized Search, and Tpot Classifier, are used to fine-tune the parameters. For the CHD dataset, the random forest model with TPOT classifier has the maximum accuracy (97.52%).