Cardiovascular Disease Prediction using Various Machine Learning Algorithms

Debabrata Swain, Badal Parmar, Hansal Shah, Aditya Gandhi, Manas Ranjan Pradhan, Harprith Kaur and Biswaranjan Acharya

Department of Computer Science and Engineering, Pandit Deendayal Energy University, Gandhinagar, India
Department of Information Technology, Skyline University College Sharjah, United Arab Emirates
Faculty of Data Science and Information Technology, INTI International University, Nilai, Malaysia
Department of Computer Engineering-AI, Marwadi University, Rajkot, India

Abstract: Almost one-third of all deaths caused around the world were caused due to cardiovascular diseases. Even if death was not the result, much cost is incurred during the treatment of such diseases. But much of these deaths and treatments could have been prevented with prior action. Advance knowledge of the symptoms and consequently proper care can lead us to avoid such diseases. Thus, current research proposes a highly effective model to predict the presence of heart diseases. Bad eating habits, smoking, stress, and genetics are some of the factors that influence our body mechanisms, which actually cause various irregularities in our hearts and thus adversely affect our bodies. The body mechanisms influenced by external factors have been included to prepare an efficient model to predict the probability of cardiovascular diseases. UCI repository dataset has been utilized for the training and testing purpose in our model. Then accordingly, five different algorithms namely Logistic Regression, Support Vector Machine, Multi-Layer Perceptron (MLP) Classifier with Principal Component Analysis (PCA), Deep Neural Network, Bootstrap Aggregation using Random Forests are executed on our filtered dataset to find which one is the optimum out of all of them. Pre-processing techniques have been extensively used to filter out the dataset. The data processing along with the different models employed make this a sound paper, which could be utilized for real-world cases without any prior modification. Different places around the world would take different factors into account, hence our model can be used as it takes all critical factors from several datasets.

Keywords: Cardiovascular Disease Prediction, Aggregated Dataset, Machine Learning Algorithms, Deep Learning, Bootstrap Aggregation using Random Forests, Logistic Regression, Deep Neural Network, MLP with PCA, SVC

Introduction

With the advancement in technology, though longevity has increased, many lives are still lost to causes that could have been prevented with proper care. Heart attacks are responsible for a fair share of deaths around the world and even the number of deaths that have occurred due to heart attacks are on the rise (Motarwar et al., 2020). Moreover, heart attack in the younger ages, an incident seldom seen before, is frequently occurring in the current times. One of the more terrifying attributes of such attacks is that there are not many symptoms or predicaments beforehand. This is a dire situation that needs to be addressed with utmost priority. The better step to take would be to prevent such diseases or at least have a forewarning. Hence, the initial stage toward a healthier body would be the prediction of heart attacks. Current research proposes a system to predict whether heart disease will exist or not (Terrada et al., 2020). Numerous relevant factors ergo Age, chest pain, cholesterol, fasting sugar, resting BP, etc., have been incorporated to produce a sound model. Even the data included for this model has been integrated from various places, to give a better simulation comparable to the real-world cases. For data processing, imputing the missing values, then the output column, which indicates the chances of heart-related
issues, had to be transformed too along with scaling. Then this filtered data was passed through five different models—namely Logistic regression, Support Vector Machine (SVM), MLP Classifier with Principal Component Analysis, Deep Neural Network, and finally Bootstrap aggregation using Random Forest; to test which one gave us the best performance (Chakarverti et al., 2019). Promising accuracy of 97.67% was achieved from Bootstrap aggregation, out of all the models.

**Literature Review**

Neural Networks require much experimenting with a series of parameters to get the best performance, (Yazid et al., 2018) the paper suggests a parameter tuning framework, for the Artificial Neural Network. Statlog heart disease dataset along with the Cleveland dataset was considered to train the model. The inclusion of a variety of datasets would have increased the scope of the model, making it more acceptable (Yazid et al., 2018). The accuracy obtained is 90% for the Star log dataset, while 90.9% for the Cleveland one, suggesting that more improvements could be made. In our paper, the current model is built while considering data from various sources to make it more versatile and achieve higher accuracy of 97.67%.

Diwakar et al. (2021) in paper explain that diagnosis of the disease before its occurrence can probably save people’s lives. A thorough study of the classification methods for machine learning and image fusion has been depicted in this study (Diwakar et al., 2021). Age, Sex, FBP-Fasting blood pressure, hypertension, smoking, and many more attributes have been taken into consideration by different methods to diagnose the diseases. Then a basic explanation of various techniques along with outlining of the whole process sums up the constituent of this study, no actual method or model to diagnose or predict the occurrence of the disease has been shown here. While the model proposed by us calculates the probability of occurrence and gives a definite answer whether the disease will occur or not, by considering numerous factors.

As the number of problems related to the heart is increasing with each passing day, problems can even cause death. Hence the authors, (Singh and Kumar, 2020) have suggested a model for different algorithms (Singh and Kumar, 2020). A casual mention of the working of different Machine learning methods, along with their definitions and basic execution of four models namely k-Nearest Neighbor-KNN, decision tree, linear regression, and Support Vector Machine (SVM) has been done. The dataset on which these methods were employed is the UCI repository dataset. Further data processing and tweaking and tuning of the parameters could have resulted in much better performance. Our paper goes in-depth about these workings, which ultimately results in a more sophisticated model, that too with good accuracy of 97.67%.

One of the most critical threats to human beings in the current times is heart-related diseases. The paper (Ambesange et al., 2020) uses normalization and outlier detection was done to improve the data, obtained from the UCI repository. Application of several feature selection methods, like the Extra tree’s classifier, Random search, and other techniques are made for tuning. Seven models are developed with various features (Ambesange et al., 2020). One of the models for which the dimension reduction technique namely Kernel Principal Component Analysis (PCA) was employed on the dataset and then the Grid Search method was used to give 100% accuracy, which suggests the model is over-fitted, thus requiring much more processing. No such anomalies will be found in the paper written by us.

Heart diseases impose a great threat to mankind and one of the first steps to dealing with such diseases is to detect them. Thus, a model is proposed by Yadav et al. (2020) to predict the state of heart diseases. UCI dataset has been included to build up the model. A comparison of binary and multi-class classification, along with an explanation regarding various machine learning algorithms has been included here (Yadav et al., 2020). Regularization is used to overcome over-fitting. Fuzzy KNN is working better than the other algorithms. Pre-processing of the dataset-imputing missing values, normalization, etc., should have been done to make the dataset better.

Prediction of patterns occurring in the medical sector is a challenge. Prediction of the occurrence of diseases is quite a complicated task in itself. Motarwar et al. (2020) employ five algorithm models Random Forest, Naïve Bayes, Support Vector Machine (SVM), Hoeffding Decision Tree, and Logistic Model Tree (LMT) for predicting the occurrence of heart diseases (Motarwar et al., 2020). The dataset used for training and testing the model is the Cleveland dataset. Out of 76 attributes present, 13 are selected. Using only one dataset, without any further processing limits the functionality of this model for real-world issues. A brief explanation of the working of each algorithm is done. No methods were used for cleaning the data. The application of such filtering methods would have resulted in a much better efficiency of the model. Thus, much work in detail would have to be performed on this model to make it worthy for the predicting of diseases for any actual scenario working.

**Proposed Work**

**Dataset Description**

The proposed system aims to collect and interpret the data from various clinical databases related to heart disease originating from different geographical locations like Switzerland, Hungary, Long Beach, and Cleveland. The different datasets are available at the UCI repository (UCI, 1990).
B. Data Preprocessing

Data Cleaning

The four different datasets cumulatively contribute to a different number of records as shown in Table 1.

The dataset consists of 13 independent features and 1 dependent feature. The dependent feature is a categorical type that indicates the presence (denoted by value 1) or absence (denoted by value 0) of heart disease. The features/attributes mentioned in Table 2 and Fig. 1 depict the whole architecture of the proposed system.

A total of 920 records. The dataset contained a large amount (>50%) of missing values for Thalassemia, Peak exercise, and Vessels colored by fluoroscopy features. The number of records with missing values for each feature is shown in Table 3.

The missing values for Cholesterol are imputed using the mean strategy, while the null values for Vessels colored, fasting sugar, Slope, and Thalassemia are imputed using the mode strategy. The missing values for the Resting ECG, induced angina, Maximum rate, Resting BP, and ST Depression features were negligible and hence the records were removed from the dataset resulting in a total of 854 rows in the dataset.

The target variable of the dataset is categorized into 5 different labels numbered from 0 to 4 (inclusive). The digits 1 to 4 indicate different stages of heart disease. Since the proposed system is based on binary classification, all the labels within the range (Terrada et al., 2020; Yazid et al., 2018) are replaced with a common digit 1, indicating the presence of the disease.

Heatmaps for features before and after imputing the missing values are shown in Fig. 2 and 3 respectively.

Transforming the Labels

The target variable of the dataset had five unique values in the range [0, 4], where 0 represents no risk of heart attack and the values 1, 2, 3, and 4 indicate the increased risk of a heart attack. To make the dataset suitable for binary classification, the labels in the range (Swain et al., 2019ab; 2020) are all replaced by 1. Hence, they now indicate whether the person has a risk of heart attack or not. It no longer gives information about the amount of risk. The count of each label in the dataset is shown in Fig. 4.

Numerical Feature Selection

From all the available numerical features, the top 4 were selected using the ANOVA measure. ANOVA (Analysis of Variance) is a parametric statistical hypothesis procedure used to determine whether the means of two or more samples of data emerge out of the same distribution. This F-statistic score is widely used when the feature is numerical and the target is categorically similar to the problem at hand (Sthle and Wold, 1989). The larger the value of ANOVA more significant the numeric feature. The test concluded that the features 'Age', 'Cholesterol', 'Maximum Heart Rate, and 'ST Depression' are the most significant numerical features.

Table 1: The number of data samples from different datasets

| Dataset    | Records |
|------------|---------|
| Cleveland  | 303     |
| Hungary    | 294     |
| Switzerland| 123     |
| Long Beach | 200     |

Table 2: Features of the dataset

| Type       | Feature          | Data type  |
|------------|------------------|------------|
| Independent| Age              | Numeric    |
|            | Chest pain       | Categorical|
|            | Cholesterol      | Numeric    |
|            | Fasting sugar    | Categorical|
|            | Induced angina   | Categorical|
|            | Maximum rate     | Numeric    |

Table 3: Number of missing values for individual features

| Feature          | Missing values |
|------------------|----------------|
| Age              | 0              |
| Chest pain       | 0              |
| Sex              | 0              |
| Resting ECG      | 2              |
| Cholesterol      | 30             |
| Induced angina   | 55             |
| Maximum rate     | 55             |
| Resting BP       | 59             |
| ST Depression    | 62             |
| Fasting sugar    | 90             |
| Slope            | 308            |
| Thalassemia      | 477            |
| Vessels colored  | 606            |

Fig. 1: System architecture

Fig. 2: Missing values before imputation (Yellow lines indicate missing values)
Categorical Feature Selection

The top 4 categorical features were selected using the Chi-Squared statistic. It is a statistical reliability coefficient that presupposes (the null hypothesis) that the measured and predicted frequencies for a categorical variable match as mentioned and applied in (Spencer et al., 2020). In the context of the chi-squared distribution with the required number of degrees of freedom, we can interpret the test statistic as follows:

- If Statistic \( (S_0) > \) Critical Value \( (C_v) \), the result is significant and the null Hypothesis \( (H_0) \) is rejected
- Otherwise, the null Hypothesis \( (H_0) \) cannot be rejected

The test concluded that the features ‘Chest Pain’, ‘Induced Angina’, ‘Vessels Colored’, and ‘Thalassemia’ are the most significant categorical features.

Standard Scaling

The original values of the continuous features have very different scales. While modeling the data, some features tend to dominate others because of their higher range. Hence, to reduce the variance between the features and scale all of them to a certain range, standardization is used. Standardization expects the data to have Gaussian distribution and as observed, the features of the dataset have close to the Gaussian distribution. Standard scaling centers the data points by subtracting them from the mean and scales by dividing them by the standard deviation. Thus standardization is also referred to as 'center scaling'.

Equations 1, 2, and 3 are the equations for Mean, Standard deviation and Standard Scaled values respectively:

\[
\text{Mean} = \frac{1}{n} \sum_{i=1}^{n} x_i
\]  

\[
\text{Standard deviation} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)^2}
\]  

\[
\text{Standard scaled value} = \frac{(x_i - \mu)}{\sigma_i}
\]

C. Proposed Models

Various models employing different algorithms were trained on the same training dataset and tested on the same test dataset. The models are as follows:

1) Logistic regression  
2) Support vector machine  
3) MLP classifier with principal component analysis  
4) Deep neural network  
5) Bootstrap aggregation using random forest

Logistic Regression

At the core of logistic regression, we have the logistic function, also called the sigmoid function (Wright, 1995). This function can map any real-valued input into the range of (0,1). The function has an S-shaped curve as shown in Fig. 5.

\[
\text{Sigmoid} = \frac{1}{1 + e^{-x}}
\]

The output of a logistic regression model indicates the probability of the record belonging to a default class in binary classification. The closer the output is to 1, the higher the probability. The equation used in logistic regression is shown in Eq. 5:
where,
\[ x_1, x_2, \ldots, x_n \] = Input record
\[ \beta_0 = \text{The bias or intercept term} \]
\[ \beta_1, \beta_2, \ldots, \beta_n = \text{Are the coefficients for different feature inputs} \]

**Support Vector Machine (SVM)**

SVM is a soft margin classifier (derived from the Maximal-Margin Classifier) that learns a hyperplane from the training dataset (Noble, 2006). Tuning parameter \( C \) denotes the amount of relaxation that is allowed across all the dimensions of the hyperplane. All the training instances lying with the predefined margin are considered to be the support vectors. The smaller the value of \( C \), the more sensitive the algorithm is to the training dataset and vice versa. SVM algorithm is practically implemented using a kernel. Three different types of kernels are used in SVM:

1) Linear
2) Polynomial
3) Radial

In general, all the SVMs try to find the equation of a hyperplane that maximizes the marginal distance. Doing so allows the algorithm to classify the unknown data samples with higher accuracy. For the given dataset, an SVM model with a linear kernel has been deployed.

**Multi-Layer Perceptron (MLP) Classifier with Principal Component Analysis (PCA)**

Higher the number of features, the more dimensions and the larger the volume of space. If the number of features is considerably higher concerning the number of data samples, then the sample space does not represent the \( n \)-dimensional space with all varieties. This might lead to a lack of information and can lead to unreliable models. To overcome this issue, Principal Component Analysis (PCA) is adopted. It is a technique from linear algebra that uses feature projection methods to reduce the number of features while still maintaining the effectiveness of the data in the dataset. It is often referred to as one of the data compression techniques. PCA adopts the eigen decomposition to calculate the eigenvalues and eigenvectors that are also called the principal components. These components are then sorted in descending order of their eigenvalues. More the eigenvalue, the more impactful the eigenvector (Abdi and Williams, 2010; Krishna and Reddy, 2019).

After reducing the dimensions of the dataset down to only 3, using PCA, a Multi-Layer Perceptron (MLP) is trained using the reduced features. An MLP classifier is a fully connected, feed-forward neural network that utilizes backpropagation for updating the weight matrices (Sonawane and Patil, 2014; Terrada et al., 2020).

**Deep Neural Network**

The architecture of the neural network is shown in Fig. 6. The model consists of 3 hidden layers each with 64 neurons and ReLU as their activation function. A dropout layer is used between every hidden layer to prevent overfitting. The output layer has one neuron and it uses a sigmoid as its activation function (Karayılan and Kılıç, 2017). Other parameters tuned for the model are listed in Table 4.

**Bootstrap Aggregation using Random Forest**

Bootstrap aggregation also called bagging, is a powerful ensemble method. The method combines the predictions from multiple machine learning algorithms to provide an even more accurate prediction. In bootstrap aggregation, multiple models are fed and trained on the resamples of the same training dataset with replacement. The individually trained models are expected to learn different patterns from the same dataset which are then combined to give an accurate prediction.

Random forests are also a powerful bootstrap aggregation algorithm. The algorithm utilizes multiple decision trees that are trained on the resamples of the training dataset. Decision trees work greedily, i.e., at each split point, they select the feature that most optimally minimizes the error. This nature can in turn produce a high correlation between the predictions of the individual decision trees. Bagging algorithms rely upon independent models that have high variance and low bias so that there is a very weak correlation between their predictions. To overcome this issue, random forests limit the search space of the features that a decision tree can go through at a split point. This little tweak allows random forests to generate weakly correlated decision trees (Hastie et al., 2009).

![Fig. 5: Curve of the sigmoid function](image-url)
Fig. 6: Neural network architecture

Fig. 7: Classification report for bootstrap aggregation using random forests

Fig. 8: Confusion matrix for bootstrap aggregation using random forests

Fig. 9: Precision-Recall (PR) curve of logistic regression model

Result and Performance Analysis

The proposed approach is evaluated using several splitting models and analysis is then undertaken. The confusion matrix in Table 3 is used to illustrate the model’s performance. Classification Accuracy is defined as the percentage of correctly categorized points to the total test data points (Swain et al., 2019a; 2020). The accuracy of the proposed classifier was 97.67%. The model’s Cohen Kappa Score (E1. 6), R2 Score (Eq. 7), and AUC Score (Eq. 8) is determined to be 95.34, 90.67, and 97.77%, respectively. Furtherly, as illustrated in Fig. 7 the classification report
shows the F1 score which is used to evaluate the model's testing accuracy in terms of recall and precision for binary classification (Swain et al., 2020; 2018). For sensitive applications where an error-free diagnosis is a must, recall and precision are primarily seen as critical performance aspects.

Cohen Kappa ($\kappa$):

$$\kappa = \frac{(\rho_0 - \rho_e)}{1 - \rho_e}$$

(6)

where,

$\rho_0 = $ Empirical probability of agreement on the label assigned to the sample (the observed agreement ratio)

$\rho_e = $ Expected agreement when both annotators assign labels randomly

R2 score:

$$R^2 = 1 - \frac{s_{res}}{s_{tot}}$$

(7)

where,

$s = $ Sum of squares of the residual errors

$s = $ Total sum of the errors

$$AUC = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} p_i > p_j$$

(8)

where,

$p_i > p_j = $ Probability score by the classifier to data points $i$ and $j$

$1 p_i > p_j = $ Indicator Function (Output 1 if the condition is satisfied) $i$ run over all $m$ data points with true label 1 and $j$ runs over all $n$ data points with true label 0

The classification report and confusion matrix for each of the deployed models are generated, as shown in Fig. 7 and 8 for Bootstrap Aggregation using Random Forests. Accuracy, Precision, Recall, and F1 Score measurements are used to evaluate the system's performance. The robust performance of the proposed classifier is demonstrated by the significant number of these metrics.

The Precision-Recall (PR) curves for different models are illustrated in Fig. 9-13.

It can be observed that for the majority of the models, which include Logistic Regression, SVM, MLP with PCA, and Bootstrap aggregation, the 90-10 split of the dataset gave the best PR curve among the four splits. The results of the PR curves for these four models are as expected since the model always learns better when exposed to more training data. The exception of the DNN model having a slightly better PR curve for the 80-20 split than the 90-10 split, is likely to be the result of overfitting in the 90-10 split DNN model. Among the five models, the PR curve corresponding to the 90-10 test dataset split for the Bootstrap aggregation model is the best.
for the DNN, trained on a 90-10 split training dataset outperformed the ones trained with other split proportions. The Bootstrap aggregation model produces the best AUC score (0.9778) among the other models with all their splits (Jerome, 2006; Hamel, 2009).

**Fig. 12:** Precision-Recall (PR) curve of the deep neural network model

**Fig. 13:** PR curve of bootstrap aggregation using random forests model

Receiver Operating Characteristic (ROC) curve values for all the models are illustrated in Fig. 14-18. A trend similar to that for the PR curves can be observed for the ROC curves. All the models, except

**Fig. 14:** ROC curve for logistic regression model history

**Fig. 15:** ROC curve of Simple Vector Machine (SVM) model history
The development of artificial intelligence in the field of medical sciences has sparked a plethora of studies aimed at lowering the death rate by using various approaches to machine learning. The proposed model compared the outcomes of all the aforementioned procedures and arrived at the 97.674% accurate decision to employ the Bootstrap Aggregation with Random Forest technique which is illustrated in Table 6.

Results of five evaluation metrics for all the models trained on different splits are mentioned in Table 5.

Table 5: Performance of all the models

| Model                  | Split | Accuracy  | Cohen Kappa | R2 Score | AUC   |
|------------------------|-------|-----------|-------------|----------|-------|
| Logistic Regression    | 90-10 | 0.95348   | 0.90677     | 0.81355  | 0.9533|
|                        | 80-20 | 0.94736   | 0.89450     | 0.78888  | 0.9475|
|                        | 70-30 | 0.95719   | 0.91434     | 0.82835  | 0.9580|
|                        | 60-40 | 0.94152   | 0.88255     | 0.76492  | 0.9416|
| Support Vector Machine (SVC) | 90-10 | 0.91860   | 0.83667     | 0.67371  | 0.9178|
|                        | 80-20 | 0.92982   | 0.85908     | 0.71851  | 0.9290|
|                        | 70-30 | 0.89105   | 0.78067     | 0.56308  | 0.8887|
|                        | 60-40 | 0.88888   | 0.77519     | 0.55335  | 0.8850|
| MLP with PCA           | 90-10 | 0.91860   | 0.83703     | 0.67371  | 0.9189|

Table 4: Hyperparameters used for the model

| Hyperparameters                        | Value/type          |
|----------------------------------------|---------------------|
| Activation function for the hidden layers | ReLU                |
| Activation function for the output layer | Sigmoid             |
| The technique used to prevent overfitting | Early stopping      |
| Optimizer                               | Stochastic gradient Descent |
| Loss function                           | Binary cross entropy |
| Batch size                              | 8                   |
| epochs                                  | 1000                |

Fig. 16: ROC curve of Multi-layer Perceptron (MLP) Classifier with Principal Component Analysis (PCA) model history

Fig. 17: ROC curve of deep neural network model history

Fig. 18: ROC curve of bootstrap aggregation using random forests model history
The logistic model employed along with the sigmoid function gives a surprising accuracy of 95.38%. Multi-Layer Perceptron (MLP) classifier with Principal Component Analysis (PCA) is used to overcome when some information may not be included. The average accuracy of 91.86% is obtained, which has increased to these levels due to backpropagation techniques. The Deep neural network has three layers and is using ReLu as its activation function while giving an optimal accuracy of 95.32% among all splits. Here, the 90-10 split gets overfitted a little, hence the 80-20 split performs better overall. On the other hand, the SV Classifier performed worst when employed with the 60-40 (train-test) dataset split at 88.88% accuracy.

The hyper-parameters of all the models were finely tuned using random search to optimize the model. The Bootstrap Aggregation model performed exceedingly well as compared to the other systems, the reason being it had the advantage of multiple model predictions. This allowed the model to use the "crowd-vote" system to understand the dynamics of the dataset better and cancel out any bias that an individual model may possess. The purpose behind the building up of these models is to start predicting accurately
the occurrence of the diseases, so we can start on the next step of treating or even preventing them from happening.

Acknowledgment

The authors of this manuscript would like to express their appreciation and gratitude to their universities for supporting this research.

The authors would like to thank the editors for their efforts in handling the manuscript and all reviewers for the constructive comments which improved the original submission.

Author’s Contributions

Debabrata Swain: Problem identification, Implementation.

Badal Parmar, Hansal Shah and Aditya Gandhi: Implementation.

Manas Ranjan Pradhan, Harprith Kaur and Biswaranjan Acharya: Data Set Identification and Preprocessing.

Ethics

This study is original and innovative and contains unpublished material. The corresponding author confirms that all the other authors have read and approved the manuscript and no ethical issues involved or conflicts of interest to release.

References

Abdi, H., & Williams, L. J. (2010). Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics, 2*(4), 433-459. https://doi.org/10.1002/wics.101

Ambesange, S., Vijayalaxmi, A., Sridevi, S., & Yashoda, B. S. (2020, July). Multiple heart disease prediction using logistic regression with ensemble and hyper parameter tuning techniques. In 2020 Fourth World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4) (pp. 827-832). IEEE. https://ieeexplore.ieee.org/abstract/document/9210404

Chakarverti, M., Yadav, S., & Rajan, R. (2019, July). Classification Technique for Heart Disease Prediction in Data Mining. In 2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT) (Vol. 1, pp. 1578-1582). IEEE. https://ieeexplore.ieee.org/abstract/document/8993191

Diwakar, M., Tripathi, A., Joshi, K., Memoria, M., & Singh, P. (2021). Latest trends on heart disease prediction using machine learning and image fusion. *Materials Today: Proceedings, 37*, 3213-3218. https://doi.org/10.1016/j.matpr.2020.09.078

Hamel, L. (2009). Model assessment with ROC curves. In *Encyclopedia of Data Warehousing and Mining, Second Edition* (pp. 1316-1323). IGI Global. https://www.igi-global.com/chapter/model-assessment-roccurves/10992

Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). *The elements of statistical learning: Data mining, inference and prediction* (Vol. 2, pp. 1-758). New York: Springer. https://link.springer.com/book/10.1007/978-0-387-21606-5

Jerome, F. S. U. (2006, January). Understanding receiver operating characteristics. Canadian Journal of Emergency Medicine, Volume 8(Issue 1), 19-20.

Karaylan, T., & Kılıç, Ö. (2017, October). Prediction of heart disease using neural network. In 2017 *International Conference on Computer Science and Engineering (UBMK)* (pp. 719-723). IEEE. https://ieeexplore.ieee.org/abstract/document/8093512

Krishna, C. L., & Reddy, P. V. S. (2019, February). An efficient deep neural network multilayer perceptron based classifier in healthcare system. In 2019 3rd *International Conference on Computing and Communications Technologies (ICCCT)* (pp. 1-6). IEEE. https://ieeexplore.ieee.org/abstract/document/8824913

Motarwar, P., Durape, A., Suganya, G., & Premalatha, M. (2020, February). Cognitive approach for heart disease prediction using machine learning. In 2020 *International Conference on Emerging Trends in Information Technology and Engineering (icETITE)* (pp. 1-5). IEEE. https://ieeexplore.ieee.org/abstract/document/9077680

Noble, W. S. (2006). What is a support vector machine? *Nature Biotechnology, 24*(12), 1565-1567. https://www.nature.com/articles/nbt1206-1565

Singh, A., & Kumar, R. (2020, February). Heart disease prediction using machine learning algorithms. In 2020 *International conference on electrical and electronics engineering (ICE3)* (pp. 452-457). IEEE. https://ieeexplore.ieee.org/abstract/document/9122958

Sonawane, J. S., & Patil, D. R. (2014, February). Prediction of heart disease using multilayer perceptron neural network. In *International conference on information communication and embedded systems (ICICES2014)* (pp. 1-6). IEEE. https://ieeexplore.ieee.org/abstract/document/7033860

Spencer, R., Thabtah, F., Abdelhamid, N., & Thompson, M. (2020). Exploring feature selection and classification methods for predicting heart disease. *Digital health, 6*, 2055207620914777. https://doi.org/10.1177/2055207620914777
Sthle, L., & Wold, S. (1989). Analysis of variance (ANOVA). *Chemometrics and Intelligent Laboratory Systems*, 6(4), 259-272. https://doi.org/10.1016/0169-7439(89)80095-4

Swain, D., Ballal, P., Dolase, V., Dash, B., & Santhappan, J. (2020). An efficient heart disease prediction system using machine learning. In *Machine Learning and Information Processing* (pp. 39-50). Springer, Singapore. https://link.springer.com/chapter/10.1007/978-981-15-1884-3_4

Swain, D., Pani, S. K., & Swain, D. (2018, December). A metaphoric investigation on prediction of heart disease using machine learning. In 2018 *International conference on advanced computation and telecommunication (ICACAT)* (pp. 1-6). IEEE. https://ieeexplore.ieee.org/abstract/document/8933603

Swain, D., Pani, S. K., & Swain, D. (2019a). An efficient system for the prediction of coronary artery disease using dense neural network with hyper parameter tuning. *Int. J. Innov. Technol. Explor. Eng.*, 8(6), 689-695.

Swain, D., Pani, S., & Swain, D. (2019b). Diagnosis of coronary artery disease using 1-D convolutional neural network. *Int. J. Recent Technol. Eng. (IJRTE)*, 8(2).

Terrada, O., Cherradi, B., Hamida, S., Raihani, A., Moujahid, H., & Bouattane, O. (2020, September). Prediction of patients with heart disease using artificial neural network and adaptive boosting techniques. In 2020 *3rd International Conference on Advanced Communication Technologies and Networking (CommNet)* (pp. 1-6). IEEE. https://ieeexplore.ieee.org/abstract/document/9199620

UCI. (1990). Heart disease dataset. https://archive.ics.uci.edu/ml/datasets/heart+disease

Yadav, S. S., Jadhav, S. M., Nagrale, S., & Patil, N. (2020, March). Application of machine learning for the detection of heart disease. In 2020 *2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)* (pp. 165-172). IEEE. https://ieeexplore.ieee.org/abstract/document/9074954

Yazid, M. H. A., Satria, M. H., Talib, S., & Azman, N. (2018, October). Artificial neural network parameter tuning framework for heart disease classification. In 2018 *5th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)* (pp. 674-679). IEEE. https://ieeexplore.ieee.org/abstract/document/8752821