



Available online at <http://www.advancedscientificjournal.com>

<http://www.krishmapublication.com>

IJMASRI, Vol. 2, issue 12, pp.- 770 - 773, Dec. -2022

<https://doi.org/10.53633/ijmasri>

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY ADVANCED SCIENTIFIC RESEARCH AND INNOVATION (IJMASRI)

ISSN: 2582-9130

IBI IMPACT FACTOR 1.5

DOI: 10.53633/IJMASRI

RESEARCH ARTICLE

HEART FAILURE PREDICTION

Abhaya Gupta

Department of Information Technology, Maharaja Agrasen Institute of Technology affiliated to Guru Gobind Singh Indraprastha University, Rohini, Delhi Email: abhaygpt496@gmail.com

Abstract

Heart is one of the most important Part of the body. It helps to purify and circulate blood to all parts of the body. Most number of deaths in the world are due to heart diseases. Some symptoms like chest pain, faster heartbeat, discomfort in breathing is recorded. This data is analyzed on regular basis. In this review, an overview of the heart disease and its current procedures is firstly introduced. Furthermore, an in-depth analysis of the most relevant machine learning techniques available on the literature for heart disease prediction is briefly elaborated. The discussed machine learning algorithms are Decision Tree, SVM, ANN, Naive Bayes, Random Forest, KNN. The algorithms are compared on the basis of features. We are working on the algorithm with best accuracy. This will help the doctors to assist the heart problem easily.

Keywords: Machine Learning, Prediction, Classification Technique, Decision Tree, Accuracy

Introduction

Heart disease is one of the prevalent disease that can lead to reduce the lifespan of human beings nowadays. Each year 17.5 million people are dying due to heart disease (Animesh Hazra et al., 2017). Life is dependent on component functioning of heart, because heart is necessary part of our body. Heart disease is a disease that affects on the function of heart [2]. An estimate of a person's risk for coronary heart disease is important for many aspects of health promotion and clinical medicine. A risk prediction model may be obtained through multivariate regression analysis of a longitudinal study [3]. Due to

digital technologies are rapidly growing, healthcare centres store huge amount of data in their database that is very complex and challenging to analysis. Data mining techniques and machine learning algorithms play vital roles in analysis of different data in medical centres. The techniques and algorithms can be directly used on a dataset for creating some models or to draw vital conclusions, and inferences from the dataset. Common attributes used for heart disease are Age, Sex, Fasting Blood Pressure, Chest Pain type, Resting ECG (test that measures the electrical activity of the heart), Number of major vessels colored by fluoroscopy, Threst Blood Pressure (high blood

pressure), Serum Cholesterol (determine the risk for developing heart disease), Thalach (maximum heart rate achieved), ST depression (finding on an electrocardiogram, trace in the ST segment is abnormally low below the baseline), painloc (chest pain location (substernal=1, otherwise=0)), Fasting blood sugar, Exang (exercise included angina), smoke, Hypertension, Food habits, weight, height and obesity.

Related background

(Polaraju *et al.*, 2017), proposed Prediction of Heart Disease using Multiple Regression Model and it proves that Multiple Linear Regression is appropriate for predicting heart disease. Marjia *et al.*, developed heart disease prediction using KStar, j48, SMO, and Bayes Net and Multilayer perception using WEKA software. S. Seema *et al.* focuses on techniques that can

predict chronic disease by mining the data containing in historical health records using Naïve Bayes, Decision tree, Support Vector Machine (SVM) and Artificial Neural Network (ANN). Chala Beyene *et al.*, recommended Prediction and Analysis the occurrence of Heart Disease Using Data Mining Techniques. R. Sharmila *et al.*, proposed to use non-linear classification algorithm for heart disease prediction. S.Prabhavathi *et al.*, proposed Decision tree based Neural Fuzzy System (DNFS) technique to analyse and predict of various heart disease. Sharan Monica.L *et al.* proposed an analysis of cardiovascular disease. Sharma Purushottam *et al.*, proposed c45 rules and partial tree technique to predict heart disease.

Table related background table for Heart failure prediction

Study	Purpose	Techniques Used	Accuracy
Sharma Purushottam (2015)	Efficient Heart Disease Prediction System using Decision Tree.	Decision tree classifier	86.3% for testing phase. 87.3% for training phase.
Boshra Brahmi (2015)	Prediction and Diagnosis of Heart Disease by Data Mining Techniques.	J48, Naïve Bayes, KNN, SMO	J48 gives better accuracy than other three techniques.
Sairabi Mujawar (2015)	Prediction of Heart Disease using Modified K-means and by using Naïve Bayes	Modified k-means algorithm, naive bayes algorithm.	Heart Disease detection=93% Heart Disease undetection=89%
Sharma Purushottam (2015)	Heart Disease Prediction System Evaluation using C4.5 Rules and Partial Tree	C4.5 rules and Naive Bayes algorithm	C4.5 gives better accuracy than Naive Bayes.
S. Prabhavathi (2016)	Analysis and Prediction of Various Heart Diseases using DNFS Techniques.	Decision tree, c4.5, SVM, naïve bayes	Accuracy according to the types of heart disease. CVD Diagnosis= between 85% and 99%.

Sharan Monica (2016)	Analysis of Cardiovascular disease Prediction using Data Mining Techniques.	J48 Naïve Bayes Simple CART	91.4% 88.5% 92.2%
Jayami Patel (2017)	Heart disease Prediction using Machine Learning and Data mining Technique.	LMT, UCI	UCI gives better accuracy, compared to LMT.
P. Sai Chandrasekhar Reddy (2017)	Heart disease prediction using ANN algorithm in data mining.	ANN	Accuracy proved in JAVA.
Chala Bayen (2018)	Prediction and Analysis the occurrence of heart disease using data mining techniques.	J48, Naïve Bayes, Support Vector Machine.	It gives short time result which helps to give quality of services and reduce cost to individuals.
R. Sharmila (2018)	A conceptual method to enhance the prediction of heart diseases using the data techniques.	SVM in parallel fashion	SVM provides better and efficient accuracy of 85% and 82.35%. SVM in parallel fashion gives better accuracy than sequential SVM.

Challenges Faced

Data plays a significant role in the machine learning process. One of the significant issues that machine learning professionals face is the absence of good quality data. Unclean and noisy data can make the whole process extremely exhausting. We don't want our algorithm to make inaccurate or faulty predictions. Hence the quality of data is essential to enhance the output. Therefore, we need to ensure that the process of data pre-processing which includes removing outliers, filtering missing values, and removing unwanted features, is done with the utmost level of perfection. This process occurs when data is unable to establish an accurate relationship between input and output variables. It simply means trying to fit in undersized jeans. It signifies the data is too simple to establish a precise relationship. To overcome this issue: Maximize the training time, Enhance the complexity of the model, Add more features to the data. Overfitting refers to

a machine learning model trained with a massive amount of data that negatively affect its performance. It is like trying to fit in Oversized jeans. Unfortunately, this is one of the significant issues faced by machine learning professionals. This means that the algorithm is trained with noisy and biased data, which will affect its overall performance. The machine learning industry is young and is continuously changing. Rapid hit and trial experiments are being carried on. The process is transforming, and hence there are high chances of error which makes the learning complex. It includes analysing the data, removing data bias, training data, applying complex mathematical calculations, and a lot more. The most important task you need to do in the machine learning process is to train the data to achieve an accurate output. Less amount training data will produce inaccurate or too biased predictions. Let us understand this with the help of an example. Consider a machine learning algorithm like training a child. This is one of the common issues faced by machine learning professionals. The

machine learning models are highly efficient in providing accurate results, but it takes a tremendous amount of time. Slow programs, data overload, and excessive requirements usually take a lot of time to provide accurate results. Further, it requires constant monitoring and maintenance to deliver the best output. So, you have found quality data, trained it amazingly, and the predictions are concise and accurate. Yay, you have learned how to create a machine learning algorithm!! But wait, there is a twist; the model may become useless in the future as data grows. The best model of the present may become inaccurate in the coming Future and require further rearrangement. So, you need regular monitoring and maintenance to keep the algorithm working. This is one of the most exhausting issues faced by machine learning professionals. One other challenge we faced is that we needed to learn all the gradient boosting algorithms first and then implemented them.

Conclusion

We may fairly say that every approach of these discoveries has been explored, along with all possible implementations, after examining the flaws in the already utilized. Heart Failure Prediction Analysis. We found numerous flaws and contradictions in how these research articles were conducted. By implementing a substantially improved version of the Heart Failure analysis system, we hope to eliminate these shortcomings

Reference

1. Animesh Hazra., Arkomita Mukherjee, Amit Gupta, Asmita Mukherjee. (2017). "Heart Disease Diagnosis and Prediction Using Machine Learning and Data Mining Techniques: A Review", Research Gate Publications, July 2017, pp.2137-2159.
2. Krishnaiah, V. Narsimha, G and Subhash Chandra,N. (2016). "Heart Disease Prediction System using Data Mining Techniques and Intelligent Fuzzy Approach: A Review", International Journal of Computer Applications, February 2016.
3. Guizhou H and Martin M. Root. (2005). "Building Prediction Models for Coronary Heart Disease by Synthesizing Multiple Longitudinal Research Findings", European Science of Cardiology, 10May 2005.
4. Mythili, T. Dev Mukherji. Nikita Padaila and Abhiram Naidu, (2013). "A Heart Disease Prediction Model using SVMDecision Trees-Logistic Regression (SDL)", International Journal ofComputer Applications, vol. 68, 16 April 2013.
5. <https://www.medicalnewstoday.com/articles/257484.php>.
6. Nimai Chand Das Adhikari, Arpana Alka, and rajat Garg (2014). "HPPS: Heart Problem Prediction System using Machine Learning".
7. Polaraju, K and Durga Prasad, D. (2017). "Prediction of Heart Disease using Multiple Linear Regression Model", International Journal of Engineering Development and Research Development, ISSN: 2321-9939, 2017.
8. Marjia Sultana and Afrin Haider, (2017). "Heart Disease Prediction using WEKA tool and 10-Fold cross-validation", The Institute of Electrical and Electronics Engineers, March 2017.
9. Seema Shedole S and Kumari Deepika. (2016). "Predictive analytics to prevent and control chronic disease", <https://www.researchgate.net/punlication/316530782>, January 2016.
10. Ashok kumar Dwivedi. (2016). "Evaluate the performance of different machine learning techniques for prediction of heart disease using ten-fold cross-validation", Springer, 17 September 2016.
11. Megha Shahi, R and Kaur Gurm,(2017). "Heart Disease Prediction System using Data Mining Techniques", Orient J. Computer Science Technology, vol.6 2017, pp.457-466.
