A Survey on Prediction Techniques of Cardiovascular Disease using Machine Learning

Divya Raja Gupta
Department of Computer Science & Eng
Chouksey Engineering College
Lal Khadan, Masturi Road, Bilaspur (C.G)

Neetin Kumar
Department of Computer Science & Eng
Chouksey Engineering College
Lal Khadan, Masturi Road, Bilaspur (C.G)

Abstract— Cardiovascular 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 Cardiovascular Diseases. Some symptoms like chest pain, faster Cardiovascular beat, discomfort in breathing are recorded. This data is analysed on regular basis. In this review, an overview of the Cardiovascular 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 Cardiovascular 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 Cardiovascular problem easily.

Keywords— Machine Learning, Prediction, Classification Technique, Decision Tree, Accuracy.

I. INTRODUCTION

Cardiovascular disease is the kind of disease which can cause the death. Each year too many peoples are dying due to Cardiovascular disease. Cardiovascular disease can be occurred due to the weakening of Cardiovascular muscle. Also, the Cardiovascular failure can be described as the failure of Cardiovascular to pump the blood. Cardiovascular disease is also called as coronary artery disease (CAD). CAD can be occurred due to insufficient blood supply to arteries. Cardiovascular disease can be detected using the symptoms like: high blood pressure, chest pain, hypertension, cardiac arrest, etc. There are many types of Cardiovascular diseases with different types of symptoms. Like: 1) Cardiovascular disease in blood vessels: chest pain, shortness of breath, pain in neck throat., 2)Cardiovascular disease caused by abnormal Cardiovascularbeats :slow Cardiovascularbeat, discomfort, chest pain., etc. Most common symptoms are chest pain, shortness of breath, discomfort, chest pain., etc. Most common symptoms are chest pain, shortness of breath, fainting. Causes of Cardiovascular disease are defects you're born with, high blood pressure, diabetes, smoking, drugs, alcohol. Sometimes in Cardiovascular disease the

infection also which affects the inner membrane which is identified by symptoms like fever, fatigue, dry cough, skin rashes. Causes of Cardiovascular infection are bacteria, viruses, parasites. Types of Cardiovascular disease: Cardiac arrest, Hypertension, Coronary artery disease, Cardiovascular failure, Cardiovascular infection, Congenital Cardiovascular disease, Slow Cardiovascular beat, Stroke type Cardiovascular disease, angina pectoris. Now a days there are too many automated techniques to detect the Cardiovascular disease like data mining, machine learning, deep learning, etc. So, in this paper we will brief introduction about machine learning techniques. In this we train the datasets using the machine learning repositories. There are some risk factors on the basis of that the Cardiovascular disease is predicted. Risk factors are: Age, Sex, Blood pressure, Cholesterol level, Family history of coronary illness, Diabetes, Smoking, Alcohol, Being overweight, Cardiovascular rate, Chest Pain.

II. LITERATURE REVIEW

There is number of works has been done related to disease prediction systems using different machine learning algorithms in medical Centre's.

Senthil Kumar Mohan et al,[1] proposed Effective Cardiovascular Disease Prediction Using Hybrid Machine Learning Techniques in which strategy that objective is to finding critical includes by applying Machine Learning bringing about improving the exactness in the expectation of cardiovascular malady. The expectation model is created with various blends of highlights and a few known arrangement strategies. We produce an improved exhibition level with a precision level of 88.7% through the prediction model for Cardiovascular disease with hybrid random forest with a linear model (HRFLM) they likewise educated about Diverse data mining approaches and expectation techniques, Such as, KNN, LR, SVM, NN, and Vote have been fairly famous of late to distinguish and predict Cardiovascular disease.

Sonam Nikhar et al [2] has built up the paper titled as Prediction of Cardiovascular Disease Using Machine Learning Algorithms by This exploration plans to give a point by point portrayal of Naïve Bayes and decision tree classifier that are applied in our examination especially in the prediction of Cardiovascular Disease. Some analysis has been led to think about the execution of prescient data mining strategy on the equivalent dataset, and the result uncovers that Decision Tree beats over Bayesian classification system.

Aditi Gavhane, Gouthami Kokkula, Isha Pandya, Prof. Kailas Devadkar (PhD), [3] Prediction of Cardiovascular Disease Using Machine Learning", In this paper proposed system they used the neural network algorithm multilayer

perceptron (MLP) to train and test the dataset. In this algorithm there will be multiple layers like one for

input, second for output and one or more layers are hidden layers between these two input and output layers. Each node in input layer is connected to output nodes through these hidden layers. This connection is assigned with some weights. There is another identity input called bias which is with weight b, which added to node to balance the perceptron. The connection between the nodes can be feedforwarded or feedback based on the requirement.

Abhay Kishore et al,[4] developed Cardiovascular Attack Prediction Using Deep Learning in which This paper proposes a Cardiovascular attack prediction system using Deep learning procedures, explicitly Recurrent Neural System to predict the probable prospects of Cardiovascular related infections of the patient. Recurrent Neural Network is a very ground-breaking characterization calculation that utilizes Deep Learning approach in Artificial Neural Network. The paper talks about in detail the significant modules of the framework alongside the related hypothesis. The proposed model deep learning and data mining to give the precise outcomes least blunders. This paper gives a bearing and point of reference for the advancement of another type of Cardiovascular attack prediction platform. Prediction stage.

Lakshmana Rao et al,[5] Machine Learning Techniques for Cardiovascular Disease Prediction in which the contributing elements for Cardiovascular disease are more (circulatory strain, diabetes, current smoker, high cholesterol, etc..). So, it is difficult to distinguish Cardiovascular disease. Different systems in data mining and neural systems have been utilized to discover the seriousness of Cardiovascular disease among people. The idea of CHD ailment is bewildering, in addition, in this manner, the disease must be dealt with warily. Not doing early identification, may impact the Cardiovascular or cause sudden passing. The perspective of therapeutic science furthermore, data burrowing is used for finding various sorts of metabolic machine learning a procedure that causes the framework to gain from past information tests, models without being expressly customized. Machine learning makes rationale dependent on chronicled information.

Mr. Santhana Krishnan.J and Dr. Geetha.S, [6] Prediction of Cardiovascular disease using machine learning algorithm This Paper predicts Cardiovascular disease for Male Patient using Classification Techniques. The detailed information about Coronary Cardiovascular diseases such as its Facts, Common Types, and Risk Factors has been explained in this paper. The Data Mining tool used is WEKA (Waikato Environment for Knowledge Analysis), a good Data Mining Tool for Bioinformatics Fields. The all three available Interface in WEKA is used here; Naive Bayes, Artificial Neural Networks and Decision Tree are Main Data Mining Techniques and through this techniques Cardiovascular disease is predicted in this System. The main Methodology used for prediction is Decision Trees like CART, C4.5, CHAID, J48, ID3 Algorithms, and Naive Bayes Techniques.

Avinash Golande et al,[7] proposed Cardiovascular Disease Prediction Using Effective Machine

Learning Techniques in which Specialists utilize a few data mining strategies that are available to support the authorities or doctors distinguish the Cardiovascular disease. Usually utilized methodology utilized are decision tree, k- closest and Naïve Bayes. Other unique characterization-based strategies utilized are packing calculation, Part thickness, consecutive negligible streamlining and neural systems, straight Kernel self- arranging guide and SVM (Bolster Vector Machine). The following area obviously gives subtleties of systems that were utilized in the examination.

V.V. Ramalingam et Al,[8] proposed Cardiovascular disease prediction using machine learning techniques in which Machine Learning algorithms and techniques have been applied to various medical datasets to automate the analysis of large and complex data. Many researchers, in recent times, have been using several machine learning techniques to help the health care industry and the professionals in the diagnosis of Cardiovascular related diseases. This paper presents a survey of various models based on such algorithms and techniques and analyse their performance. Models based on supervised learning algorithms such as Support Vector Machines (SVM), K- Nearest Neighbour (KNN), Naïve Bayes, Decision Trees (DT), Random Forest (RF) and ensemble models are found very popular among the researchers and systems have been applied to different clinical datasets to robotize the investigation of huge and complex information. Numerous scientists, as of late, have been utilizing a few Machine Learning algorithms and techniques have been applied to various medical datasets to automate the analysis of large and complex data. Many researchers, in recent times, have been using several machine learning techniques to help the health care industry and the professionals in the diagnosis of Cardiovascular related diseases. This paper presents a survey of various models based on such algorithms and techniques and analyze their performance. Models based on supervised learning algorithms such as Support Vector Machines (SVM), K- Nearest Neighbour (KNN), Naïve Bayes, Decision Trees (DT), Random Forest (RF) and ensemble models are found very popular among the researchers. strategies to enable the wellbeing to mind industry and the experts in the analysis of Cardiovascular related sicknesses. This paper presents a review of different models dependent on such calculations and methods and analyze their exhibition. Models in light of directed learning calculations, for example, Support Vector Machines (SVM), K- Nearest Neighbour (KNN), Naïve Bayes, Decision Trees (DT), Random Forest (RF) and group models are discovered extremely well known among the scientists.

TABLE I. A COMPARATIVE STUDY OF VARIOUS ALGORITHMS IN LITERATURE REVIEW

YEAR	AUTHOR	PURPOSE	TECHNIQUES USED	ACCURACY

2016	Sonam Nikhar[2]	Prediction of Cardiovascular Disease Using Machine Learning Algorithms	1) Naïve Bayes Classifier 2)Decision tree	**Decision tree** has better accuracy as compared to naïve Bayes classifier.
2018	Aditi Gavhane[3]	Prediction of Cardiovascular Disease Using Machine Learning	1) Multi-layer perceptron algorithm	The **MLP** gives best Accuracy
2018	V.V. Ramalingam[8]	Cardiovascular disease prediction using machine learning techniques	1) Naïve Bayes 2) Support Vector Machine 3) K – Nearest Neighbour 4) Decision Tree 5) Random Forest 6) Ensemble Model	**SVM** has more accuracy than other techniques.
2019	Abhay Kishore1[4]	Cardiovascular Attack Prediction Using Deep Learning	1) RNN	RNN accuracy **92%**
2019	Mr. Santhana Krishnan.J[6]	Prediction of Cardiovascular Disease Using Machine Learning Algorithm s.	1)Naïve Bayes 2)Decision Tree	Decision Tree (**91%** Accuracy)
2019	Avinash	Cardiovascular	1)Decision	Decision tree

	Golande[7]	Disease Prediction Using Effective Machine Learning Technique s	Tree 2)KNN, 3)K-mean clustering 4)Adaboost	(86.60% Accuracy)
2019	Senthilkumar Mohan [1]	Effective Cardiovascular Disease Prediction Using Hybrid Machine Learning Technique s	1) Decision tree 2)Language Model 3)Support Vector Machine 4)Random Forest 5)Naïve Bayes 6)Neural Networks 7)KNN 8)HRFLM	HRFLM (88.4% Accuracy)
2019	A. Lakshmanarao[5]	Machine Learning Technique s For Cardiovascular Disease Prediction	1)Random Over sampling 2)Synthetic Minority Oversampling 3)Adaptive synthetic sampling approach	For random oversampli ng, SVM given the best accuracy of 82.30%. For Synthetic Minority Oversampli ng, Random

				Forest given the best accuracy of **91.3%** For Adaptive synthetic sampling, Random Forest (**90.3%** Accuracy)

III. CONCLUSION AND FUTURE WORK

We have summarized different types of machine learning algorithms for prediction of Cardiovascular disease. We elaborated various machine learning algorithms and worked towards finding the best algorithm by analysing their features. Every algorithm has given different result in different situations. Further it is analysed only marginal accuracy is achieved for predictive model of Cardiovascular disease and hence more complex models are needed to increase the accuracy of predicting the early Cardiovascular disease. In future we will propose methodology for early prediction of Cardiovascular disease with high accuracy and minimum cost and complexity.

REFERENCES

[1]   Senthilkumar Mohan, Chandrasegar Thirumalai, Gautam Srivastava―Effective Cardiovascular Disease Prediction Using Hybrid Machine Learning Techniques‖,   Digital Object Identifier 10.1109/ACCESS.2019.2923707, IEEE Access, VOLUME 7, 2019

[2]   S.P. Bingulac, ―On the Compatibility of Adaptive Controllers,‖ Proc. Fourth Ann. Allerton Conf. Circuits and Systems Theory, pp. 8-16, 1994. (Conference proceedings)

[3]   Sonam Nikhar, A.M. Karandikar" Prediction of Cardiovascular Disease Using Machine Learning Algorithms" International Journal of Advanced Engineering, Management and Science (IJAEMS)

Infogain Publication,[Vol-2, Issue-6, June- 2016].I.S. Jacobs and C.P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G.T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271-350.

[4]   Aditi Gavhane, Gouthami Kokkula, Isha Pandya, Prof. Kailas Devadkar (PhD)," Prediction of Cardiovascular Disease Using Machine Learning", Proceedings of the 2nd International conference on Electronics, Communication and Aerospace Technology (ICECA 2018).IEEE Conference Record # 42487; IEEE Xplore ISBN:978-1- 5386-0965-1

[5]   Abhay Kishore1, Ajay Kumar2, Karan Singh3, Maninder Punia4, Yogita Hambir5," Cardiovascular Attack Prediction Using Deep Learning", International Research Journal of Engineering and Technology (IRJET), Volume: 05 Issue: 04 | Apr-2018.

[6]   A.Lakshmanarao, Y.Swathi, P.Sri Sai Sundareswar," Machine Learning Techniques For Cardiovascular Disease Prediction", International Journal Of Scientific & Technology Research Volume 8, Issue 11, November 2019.

[7]   Mr.Santhana Krishnan.J, Dr.Geetha.S," Prediction of Cardiovascular Disease Using Machine Learning Algorithms",2019 1st International Conference on Innovations in Information and Communication Technology(ICIICT),doi:10.1109/ICIICT1.2019.8741465.

[8]   Avinash Golande, Pavan Kumar T," Cardiovascular Disease Prediction Using Effective Machine Learning Techniques", International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8, Issue-1S4, June 2019.

[9]   V.V. Ramalingam, Ayantan Dandapath, M Karthik Raja," Cardiovascular disease prediction using machine learning techniques: a survey", International Journal of Engineering & Technology, 7 (2.8) (2018) 684-687.

[10]   V. Manikantan and S. Latha, "Predicting the analysis of Cardiovascular disease symptoms using medicinal data mining methods", International Journal of Advanced Computer Theory and Engineering, vol. 2, pp.46-51, 2013.

[11]   M. S. Amin, Y. K. Chiam, K. D. Varathan,''Identification of significant features and data mining techniques in predicting Cardiovascular disease,'' Telematics Inform., vol. 36, pp. 82–93, Mar.2019.

[12]   S. M. S. Shah, S. Batool, I. Khan, M. U. Ashraf, S. H. Abbas, and S. A. Hussain,''Feature extraction through parallel probabilistic principal component analysis for Cardiovascular disease diagnosis,'' Phys. A, Stat. Mech. Appl.,vol. 482,   pp.   796–807,2017. doi:10.1016/j.physa.2017.04.113.

[13]   Stephen F. Weng, Jenna Reps, Joe Kai1, Jonathan M. Garibaldi, Nadeem Qureshi,―Can machine-

learning improvecardiovascular risk prediction using routine clinical data?‖, PLOS ONE | https://doi.org/10.1371/journal.pone. 0174944 April 4, 2017.

[14] N. Al-milli, ‗‗Backpropagation neural network for prediction of Cardiovascular disease, ‗‗J. Theor. Appl.Inf. Technol., vol. 56, no. 1, pp.131–135, 2013.

[15] A. S. Abdullah and R. R. Rajalaxmi, ''A data mining model for predicting the coronary Cardiovascular disease using random forest classifier,'' in Proc. Int. Conf. Recent Trends Comput. Methods, Commun. Controls, Apr. 2012, pp. 22–25.