

Analysis of Machine Learning Models for Heart Disease Prediction using Different Algorithms: A Review

Neeraj¹, Jagtar Singh²

¹ NCCE Israna, Department of ECE, Panipat, Haryana, India

² NCCE Israna, Department of ECE, Panipat, Haryana, India

neeruahalawat4753@gmail.com

jagtar_nit@rediffmail.com

Abstract

Now a days the heart diseases are growing very rapidly making it an important and apprehensive task of prediction of these kinds of diseases in advance. The diagnosis is also a tough chore because it has to be performed in a precise and efficient manner. The emerging technology in modern life style integrated with internet of thing which having sensors and huge amount of data is sent to various clouds for further investigation using different algorithms to fetch out precise information for various domains. Across the world approximately 3 quintillion bytes/day information generated and this data stored for further examination. As data is in huge quantity therefore, appropriate methods applied to examine the perfect analysis so that prediction can be carried out optimally. Clinical decision making is dominant to all patient care happenings which includes choosing a deed, between replacements. These days emerging field like Machine Learning play prime role in healthcare to analyze and predict the diseases. After investigating numerous research article on Machine Learning, it was found that for same data set accuracy was different for various algorithms. In our research work different machine learning techniques will be implemented and will be tested for various parameters like accuracy, precision, recall on validated dataset. ML and Neural Networks are more capable in supporting deciding and predicting from the enormous data formed by health care systems.

Keywords: Machine Learning, Cardiac Arrest, Neural Network, SVM, Logistic Regression

1. INTRODUCTION

The critical determination of ML methods in the medical industry is to explore enormously high-volume datasets related to health and to recognize useful and easily understandable outlines from health histories. This process supports in prior discovery and diagnosis of lethal illnesses. It provides the many benefits to patients like long life and precise cost-effective treatment. This dissertation seeks to explain the development of existing research on utilizing computational intelligence techniques in heart diseases diagnosis.

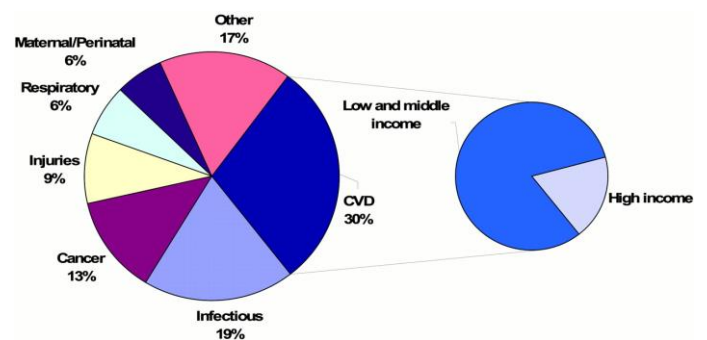


Fig.1: Different disease status in developing world [3]

This disorder extremely malignant ailment, over 1.7 billion demise all over the world. Uncertainty, primordial prognosis and diagnosis of the diseases are significant. The recent advancement in medical technology, higher computational techniques, reduced cost of storage techniques and internet connectivity enables the digitalization of diagnostic systems in the present world. Medical practitioners use diagnostic tests to reduce uncertainty about the presence of heart disease. These tests are usually expressed with various

statistical measures. Anatomic structure of heart made of soft tissue muscle contrived into 4 compartments unglued by blood vessels separated as pair of divisions. Respective divisions are called as artium and ventricle. Atriam accumulate blood flock together, and the ventricles pump out the flow of blood out of the organ. oxygenated circulation provides energy to the body. This paper seeks to explain the development of existing research on utilizing computational intelligence techniques in heart diseases diagnosis. This disorder extremely malignant ailment, over 1.7 billion demise all over the world. Uncertainty, primordial prognosis and diagnosis of the diseases are significant. The recent advancement in medical technology, higher computational techniques, reduced cost of storage techniques and internet connectivity enables the digitalization of diagnostic systems in the present world.

which other parts of body unable to do proper function. Figure 4 represents type of heart arrest due to high pressure.

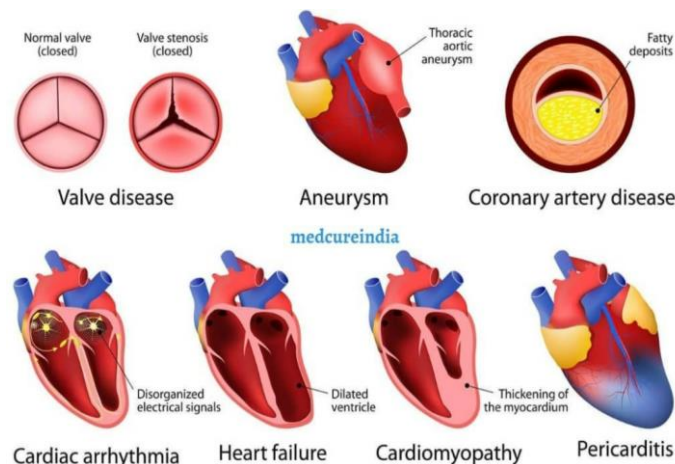


Fig.3: Types of Cardiac Disease [6].

Coronary Artery Disease: These circumstances came into existence when blood circulation becomes low then coronary artery disease take place. Veins are destroyed when there will be limited supply in arteries along this it also imbalance various parameters like diastolic function and regular systolic of the heart.

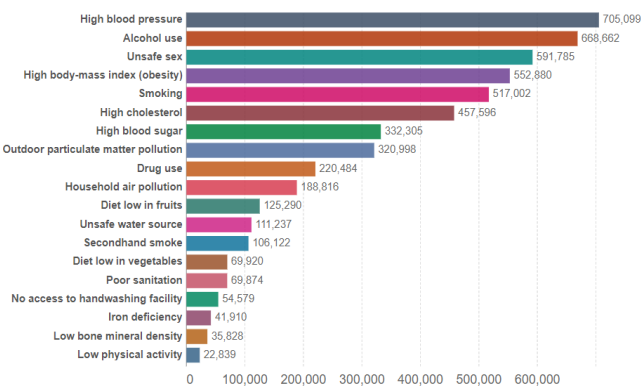


Fig.2: Number of deaths by risk factor aged 15-49 [4].

Types of Cardiac Disease: Heart disease classified into different classes and figure 3 depicts various class of heart disease as per clinical circumstances. These classes are generally categorized as:

- Valve Disease
- Aneurysm
- Coronary Artery Disease
- Cardiac Arrhythmia
- Pericarditis
- Cardiomyopathy

Acute myocardial infarction: Myocardial infarction is another name of a cardiac arrest in clinical language. When fatty material will be available in our blood then chances of cardiac arrest increased tremendously and due to fatty material, it blocks the proper blood supply to arteries and hence result in tissue damage on arteries. In this circumstance, obstruction arteries unable to provide proper oxygenated blood supply to rest of the body and due to

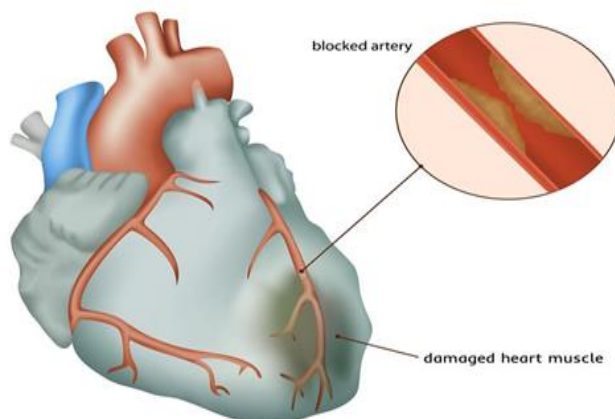


Fig. 4 Acute Myocardial Infarction [9]

Chest Pain (Angina): Angina's clinical name is also known as chest pressure. It is very major medical care requirement to deal the patients. In this circumstance patients must be deal with ventilators instantly. Due to blockage of arteries blood supply affected and it produce pressure on blood walls and finally blood vessels affected severely. After that it produce pressure on blood vessels which results in angina. Figure 1.7. shows typical chest pressure produced in the coronary vessel.

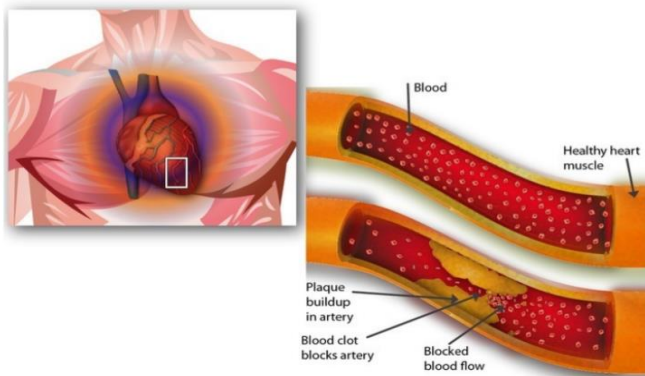


Fig. 5: Angina [10]

2. LITERATURE SURVEY

Dinesh Kumar G et al (2018): In this article first of all various techniques are implemented to pre-process the data so that machine learning algorithms can be applied effectively.

- Elimination of noisy data
- Elimination of missing data
- Sorting of features for prediction and many more.

The developed model, efficacy investigated with different parameters which are listed below: Accuracy

- Classification
- Sensitivity

This research work projected a forecast system which can predict that a person has to heart illness or not. Different machine learning algorithm are applied for same dataset for various parameters and out of them that optimized algorithm for higher accuracy will be best suited model for forecasting [6].

Zhang et al. (2017): Identifies a classifier model of forwarding feature inclusion combined with back-elimination of various information sets namely Arrhythmia, and heart disease datasets. Experiential results illustrated that the feature selections enhanced classification techniques accurately and scaled down the quantity of intakes. The provided arrhythmia dataset delivers enhanced performance of seventy-eight percentage with abridged prominently reduced from features of nineteen. The next information set provides the enhanced performance of eighty-five percentage and the quantity of features downsized to four. The above Investigation demonstrates that reduce features enhance the classifier performance [8].

Yang et al (2014): An ANN-based FIS is explained by which provides a predictive methodology for endanger information and prediction using ANFIS and linear LDA methods. The inference system makes use of the hybrid method of classification, and the results show that it

provides comparatively better accuracy than existed method. It also assists in earlier detection and prevention of coronary heart diseases [24].

Aditi Gavhane et al (2018): In this paper, researchers develop a model which can forecast that person is susceptible to heart attack based on few important parameters like pulse rate, sex, age etc. There are numerous algorithms available in machine learning to do such a tedious task. But out of them neural network implemented by the researchers and which gives satisfactory result. With pace of time extensive surge in the heart attack cases around the world, starting from childish ages, therefore there must be a robust system in our healthcare that can identify such types of severe disease with some basic information and at effective cost in early stage and can be treated with help of medical procedures. In developing countries like India, it is impossible for a person to undergo various test like ECG at regular interval of time. Therefore, there must be a reliable and economical system that can detect this lethal disease at early stage [7].

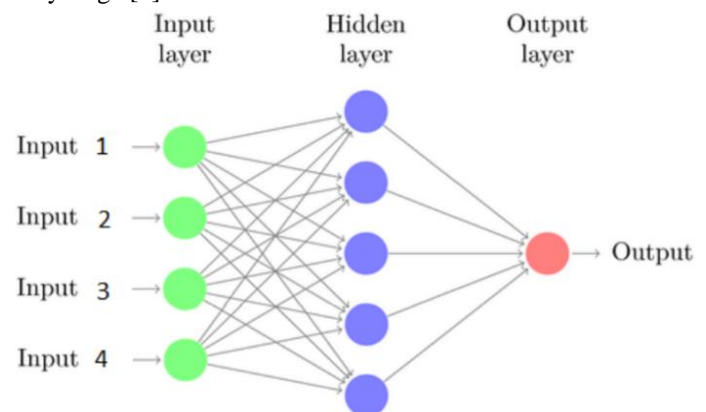


Fig.6: Logical Structure of MLP [7]

Amin Khatami et al (2017): Demonstrates a cooperative neural network ensemble to train the individual neural networks in an ensemble model. This model makes use of a constructive approach to improve the accuracy measures to determine the hidden nodes in an individual neural network. Each neural network in a multi-layered model is trained separately in an incremental manner using the negative correlation concepts. The use of harmful correlation learning methods maintains the property of diversity across multiple layers of the neural networks. This method is extensively tested across various machine learning problems such as prediction of heart illness, diabetes and recognition of letter datasets. The results show it provides a useful NN technique with a higher grade of simplification ability [9].

Ashwini Shetty et al (2016): Projected a model of an accurate prediction system utilization of Fire-Fly algorithm

based upon rough sets. The integration of roughest theory and fuzzy system reduces the challenges of doubts and high-dimensional factors of heart disease datasets. The implemented model assists in easier identification of optimal solutions with lesser computation measures. The results are more than when compared with the ANN and SVM for heart disease prediction and the medication process [10].

Bayasi N et al (2016): Emphasized an approach to heart disease diagnosis using naive Bayes classification. This work highlights the impacts of heart disease in modern life style across the universe. It provides a combinational approach of statistical methods integrated with NB classifier to perform efficient prediction and diagnosis of heart diseases. It employs data pre-processing techniques to effectively deal with the large and complex set of medical data. A discretization algorithm is used to map various heart diseases into its corresponding categories. The output of implemented system state that it provides better accuracy measures than existing techniques [12].

Wang et al (2017): explained linear integrated technique to identify correct hyperplane. It enables easier identification of appropriate hyperplane. SVM cataloging systems are listed below:

- Soft Margin Classifier
- Maximal Margin Classifier

These classifiers use the concept of margin distance measures. Through which it performs an efficient classification process [23].

3. MACHINE LEARNIG

The emerging technology in modern life style integrated with internet of thing which having sensors and huge amount of data is sent to various clouds for further investigation using different algorithms to fetch out precise information for various domains. Across the world approximately 3 quintillion bytes/day information generated and this data stored for further examination. As data is in huge quantity therefore, appropriate methods applied to examine the perfect analysis so that prediction can be carried out optimally.

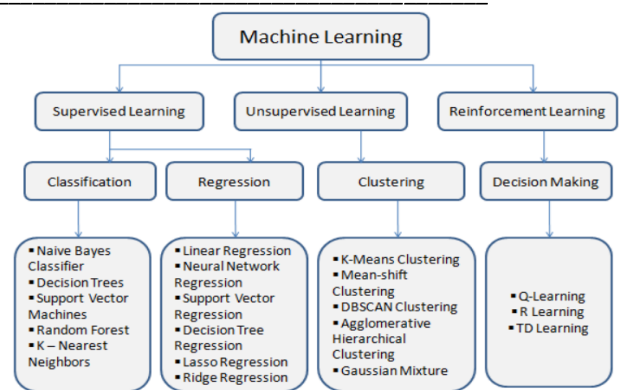


Fig.7: Machine Learning Classification.

Machine learning divided into two main group which are listed below:

- Supervised Learning
- Unsupervised Learning.

Supervised Learning: In this, algorithm shapes a mathematical model from a set of data which embed desired output as well as inputs.

Unsupervised Learning: In this, algorithm shapes a mathematical model from a set of data which contain only input, but there is no provision of desired output. as well as inputs. As in our research work, prime objective is to forecast the probability of having heart illness which directly dependent on various body parameters therefore, inputs as well as output both are required. In such a case we have to opt supervised learning.

There are various machine learning life cycle to process the data which are listed below:

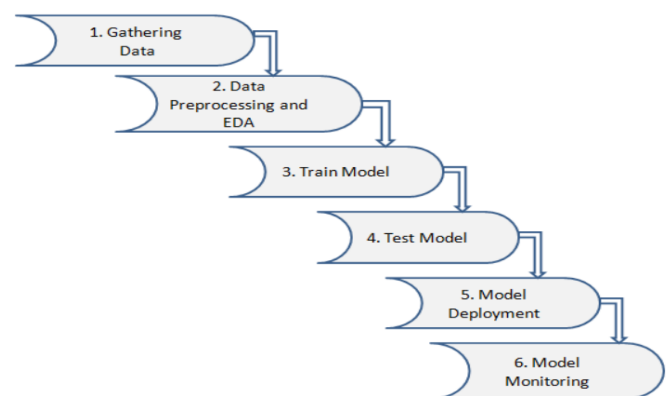


Fig.8: Machine Learning Life cycle

4. MOTIVATION

The critical determination of ML methods in the medical industry is to explore enormously high-volume datasets related to health and to recognize useful and easily understandable outlines from health histories. This process supports in prior discovery and diagnosis of lethal illnesses. It provides the many benefits to patients like long life and precise cost-effective treatment.

To fulfil the objective of research, following stages are implemented.

- Fetch dataset from reliable resource
- Pre-processing the data for further investigation
- Execute different models on cleaned dataset
- Examine the result obtained from different models
- Deduce the final result

5. CONCLUSION

Cardiac ailments are one of the most vivacious reasons of deceases. The early-stage disease estimation is a very grave task for medical data studies. Digitalization of medical data has generated a new era towards the diagnostic field. With the massive growth of digital information, these unprocessed patients' medical information is extremely essential to analyze, explore and utilize with various classification techniques. The health informatics combines mathematical models, algorithm and analysis to provide improved quality of healthcare services to the electronic health users. Across, the branch of health informatics machine learning plays a crucial role in health data analysis and management processes. This dissertation mainly focuses on classification based on heart disease predictions. These days emerging field like Machine Learning play prime role in healthcare to analyze and predict the diseases. After investigating numerous research article on Machine Learning, it was found that for same data set accuracy was different for various algorithms. In our research work different machine learning techniques will be implemented and will be tested for various parameters like accuracy, precision, recall on validated dataset. ML and Neural Networks are more capable in supporting deciding and predicting from the enormous data formed by health care systems.

REFERENCES

- [1]. H. Jindal, S. Agrawal, R. Khera, R. Jain and P. Nagrath, "Heart disease prediction using machine learning algorithms", *ICCRDA 2020, IOP Conf. Series: Materials Science and Engineering*, 1022 (2021) 012072, DOI:10.1088/1757-899X/1022/1/012072.
- [2]. P. Motarwar, A. Duraphe, G. Suganya, M. Premalatha, "Cognitive Approach for Heart Disease Prediction using Machine Learning", *International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE)*, IEEE, 2020, DOI: 10.1109/ic-ETITE47903.2020.242.
- [3]. V. Sharma, S. Yadav, M. Gupta, "Heart Disease Prediction using Machine Learning Techniques", *2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN)*, IEEE, 18-19 Dec. 2020, DOI: 10.1109/ICACCCN51052.2020.9362842.
- [4]. A. Nikam, S. Bhandari, A. Mhaske, S. Mantri, "Cardiovascular Disease Prediction Using Machine Learning Models" *IEEE Pune Section International Conference (PuneCon)*, IEEE, 16-18 Dec. 2020, DOI: 10.1109/PuneCon50868.2020.9362367.
- [5]. S. Mohan, C. Thirumalai, G. Srivastava, "Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques", *IEEE Access, Special Section on Smart Caching, Communications, Computing and Cybersecurity for Information-Centric Internet of Things*, 2019, DOI: 10.1109/ACCESS.2019.2923707.
- [6]. D. Kumar, S. Kumar, K. Arumugaraj, V. Mareeswari, "Prediction of Cardiovascular Disease Using Machine Learning Algorithms", *IEEE International Conference on Current Trends toward Converging Technologies*, Coimbatore, 2018.
- [7]. A. Gavhane, G. Kokkula, I. Pandya, K. Devadkar, "Prediction of Heart Disease Using Machine Learning", *Proceedings of the 2nd International conference on Electronics, Communication and Aerospace Technology*, IEEE Conference, 2018.
- [8]. Zhang, Shuai, Y.-L. S. A., "Deep learning-based recommender system: a survey and new perspectives", *Journal of ACM Computing Surveys* 1(1), 1–35, 2017.
- [9]. A. Khatami, A. Khosravi, "Medical image analysis using wavelet transform and deep belief networks", *Journal of Expert Systems with Applications*, 3(4), 190–198, 2017.
- [10]. A. Shetty, C. Naik, "Different data mining approaches for predicting heart disease", *International journal of innovative research in science, engineering and technology* 3(2), 277–281, 2016.
- [11]. Aydin, S., "Comparison and evaluation data mining techniques in the diagnosis of heart disease", *Indian journal of science and technology*, 6(1), 420–423, 2016.
- [12]. N. Bayasi, T. Tekeste, "Low-power ECG-based processor for predicting ventricular arrhythmia", *Journal of IEEE transactions on very large-scale integration systems*, 24(5), 1962–1974, 2016.

- [13]. GB. Berikol, O. Yildiz, "Diagnosis of acute coronary syndrome with a support vector machine", *Journal of Medical System*, 40(4), 11–18, 2016.
- [14]. Z. Wang, X. Liu, "Identification of metabolic biomarkers in patients with type-2 diabetic coronary heart diseases based on metabolomic approach", 6(30), 435–439, 2016.
- [15]. M. Singh and Martins, "Building a cardiovascular disease predictive model using structural equation model and fuzzy cognitive map", *Journal of fuzzy system*, 02(6), 1377–1382, 2016.
- [16]. S.Prabhavathi, "Analysis and prediction of various heart diseases using DNFS techniques", *International journal of innovations in scientific and engineering research*, 2(7), 678–684, 2016.
- [17]. R. Sali, M. Shavandi, "A clinical decision support system based on support vector machine and binary particle swarm optimisation for Cardiovascular disease diagnosis", *International Journal of Data mining and Bio-informatics*, 15(1), 312–327, 2016.
- [18]. P. Ghadge, K. Prajakta, "Intelligent heart attack prediction system using big data", *International journal of recent research in mathematics computer science and information technology*, 2(2), 73–77, 2016.
- [19]. G. Purusothaman, K. Krishnakumari, "A survey of data mining techniques on risk prediction: heart disease", *Indian journal of science and technology*, 8(5), 643–651, 2015.
- [20]. A. Richter, J. Listing, M. Schneider, T. Klopsch, A. Kapelle, J. Kaufmann, A. Zink, A. Strangfeld, "Impact of treatment with biologic dmards on the risk of sepsis or mortality after serious infection in patients with rheumatoid arthritis", *National Center for Biotechnology Information*, pp. 147–153, 2015.
- [21]. S.H. Mujawar, P.R. Devale, "Prediction of Heart Disease using Modified K-means and by using Naive Bayes", *International Journal of Innovative Research in Computer and Communication Engineering*, 3(10), Oct 2015.
- [22]. M. H. Vafaie, M. Ataei, "Heart diseases prediction based on ECG signals classification using a genetic-fuzzy system", *Journal of biomedical signal processing and control*, 14(5), 291–296, 2014.
- [23]. J. Wang, W. Wang, "Study on qi-deficiency syndrome identification modes of Coronary heart disease based on metabolomic biomarkers", *Journal of evidence-based complementary and alternative medicine*, 24(16), 192–198, 2014.
- [24]. X. Yang, M. Li, Y. Zhang, J. Ning, "Cost-sensitive naive bayes classification of uncertain data", *Journal of Scientific World*, 9(8), 1897–1904, 2014.
- [25]. D. Chandna, "Diagnosis of heart disease using data mining algorithm", *International journal of computer science and information technologies*, 5(2), 1678–1680, 2014.



Neeraj

M. Tech Scholar

I am Neeraj. I received my B.Tech degree from Geeta Engineering College, Naultha (Panipat). I am pursuing M.Tect from N.C College of engineering, Israna (Panipat).My Area of interest is Wireless Communication.



Jagtar Singh

Assistant Professor

Dr. Jagtar Singh received his B.Tech degree from NIT Kurukshetra, Haryana, M.Tech from N.C. College of Engineering, Israna, Panipat and Ph.D. from Guru Jambheshwar University of Science & Technology, Hisar, Haryana, India. His research interest mainly focuses on advances in MIMO technology, Digital Communication, Beyond Fourth generation (B4G) wireless communications systems.

Corresponding author. E-mail: jagtar_nit@rediffmail.com.

