International Journal of Scientific Research & Engineering Trends
Volume 5, Issue 6, Nov-Dec-2019, ISSN (Online): 2395-566X

Survey on Prediction of Heart Disease using Machine Learning

M.Tech. Scholar Yashita Agarwal Asst. Prof. Deepak Singh Tomar Hod Dr. Bhupesh Gour
Department of Computer Science And Engineering,
Lakshmi Narain College Of Technology And Science
Bhopal, (M.P), India

Abstract **- Machine Learning algorithms provide the facility to find new facts in the form of patterns from the history of records stored in databases. Predicting the diseases in earlier stages is a major task of data mining algorithms. The number of costly tests is required for patients to analyze the symptoms and causes for effective disease diagnosis. On the other hand, this number of tests on patients can be eliminated using data mining machine learning algorithms. This minimizes number of tests which plays considerable outcomes in time and accuracy level of prediction. Heart disease prediction is essential since it permits healthcare professionals to analyze the attributes important for diagnosis like blood pressure, diabetes, age, height, weight, etc., effectively.**

Keywords- **Heart Disease, Machine Learning, ECG, deep learning classification.**

I. INTRODUCTION

Data mining is concerned with the practice of computationally extracting hidden and unknown knowledge patterns from large data repositories of known information to build predictive models. It is a step to discover knowledge using the history of databases. This knowledge discovered can be exploited by the top level peoples of the organization to make strategic decisions for improving the performance and profit.

Data mining is the way of examining data from various viewpoints and summarizing it into helpful knowledge with minimum time consumption. The data mining methodologies were used in various fields of applications. It seems to be the most efficient in building healthcare support systems. Data mining algorithms have the ability to learn configurations and relationships from the clinical data set. Data mining provides a cluster of algorithms to discover hidden or unknown pattern from data which helps in transforming the heaps of data into useful information for decision making.

There is a massive volume of records available in the medical industry. Even though, there is a deficiency of efficient health care analysis tools to identify hidden knowledge in patient data, Several data mining tools and techniques are available to analyze the huge volume of healthcare data to predict several life threatening diseases like Cancer, Diabetes, Liver diseases and Heart diseases. Many data mining algorithms were introduced by many researchers to diagnose and predict these life threatening diseases to save the life of human beings. The existing Researches and the data mining approaches predict the diseases roughly and not precisely. It is necessary to study, analyze and renovate existing algorithms to improve the predicting behavior to get enhanced results. Particularly, heart disease is one of the most life threatening diseases over the globe. It is the foremost source of monotony and deaths in current modern life. Agreeing with the World Health Organization (WHO) assessment till 2030, closely 23.6 million personages will lose their life due to heart disease. In order to reduce the hazard, the probability of getting heart disease should be broken down [1]. The further most difficult consignment in medical amenities is discovery of true condition.

In the year 2008, the United States spent around 2.2 trillion dollars for healthcare and it was 15.5 percent of its Gross Domestic Product (GDP). Also, 31 percent of this outflow is credited to hospitals. Manifestly, smooth and modest reductions in hospital care cost problem. The hospital cost during 2006 was almost 30.8 billion dollars with heart diseases being accounted for around 31 percent of that extent.

II. LITERATURE SURVEY

Aishwarya et al. (2013) have illustrated in their paper, machine learning has been one of the standard and improving techniques with strong methods for classification and reorganization based on recursive learning. It allows training and test classification system, with Artificial Intelligence. Machine learning in recent years has been the evolving, reliable and supporting tools in medical domain and has provided greatest support for predicting disease with correct case of training and testing. Automatic learning has fetched a greater amount

© 2019 IJSRET

of interest in medical domain due to less amount of time for detection and less interaction with patient, saving time for patients care.

According to Buvaneshwari et al. (2012) machine learning is to build computer systems that can adapt and learn from their experience. It is the domain of research and recently it has developed in medical domain. The domain is automatically learn some task of healthcare information, medical management, patient health management etc., Application of machine learning methods to large databases is called data mining. But machine learning is not just a database problem. It is also a part of artificial intelligence. If the system can learn and adapt to such changes, the system designer need not foresee and provide solutions for all possible situations.

Asha Rajkumar et al. (2010) has illustrated that, naïve Bayes classifier as a term dealing with a simple probabilistic classifier based on application of Bayes theorem with strong independence assumptions. It assumes that the presence or absence of particular feature of a class is unrelated to the presence or absence of any other feature. It is based on conditional probabilities. It uses Bayes' theorem which finds the probability of an event occurring, given the probability of another event that has already occurred. If B represents the dependent event and A represents the prior event, Bayes' theorem can be stated as follows:

$$\text{Prob (B given A)} = \text{Prob(A and B)/Prob(A)}$$

Sudha et al. (2012) have described the decision tree as a popular classifier and prediction method for handling high dimensional data and it looks like a tree structure. It is one of the successful data mining techniques used in the diagnosis of heart disease. It applies a straightforward idea to solve the classification problem and is a very simple and easy way for handling dataset. It divides the dataset into multiple groups by evaluating individual data record, which can be described by its attributes. It requires no domain knowledge or parameter setting and can handle high dimensional data. It is also simple and easy to visualize the process of classification where in the predicates return discrete values and can be explained by a series of nested if-then-else statements. The results obtained from decision trees are easy to read and interpret.

According to Buvaneshwari et al. (2012) machine learning is to build computer systems that can adapt and learn from their experience. It is the domain of research and recently it has developed in medical domain. The domain is automatically learn some task of healthcare information, medical management, patient health management etc., Application of machine learning methods to large databases is called data mining. But machine learning is not just a database problem. It is also a part of artificial intelligence. If the system can learn and adapt to such changes, the system designer need not foresee and provide solutions for all possible situations. In recent times, the arena of data mining has realized a booming attention in all the directions and disciplines. It is a method of collecting and determining a remarkable amount of intelligence from the huge volume of databases. Classification problems for different categories of applications are the primary tasks that data mining techniques need to be implemented for the analysis and prediction [5]. Decision tree classification technique is a prevalently used technique whose procedures are best appropriate in medicinal analysis. C4.5 decision tree algorithm is one of the most popular and effectively used classifier.

A significant feature of the precise treatment determination is the development of tools that can investigate and deduce “big data” sets in a computerized and adaptive approach while providing precise and actionable medical evidence. The purpose of this research was to introduce algorithms for the identification of heart disease and the prognostication of mortality risk and to determine whether such models perform better than conventional analyses.

Machine learning systems can produce accurate disease classification and prediction that outperform existing models. Such approach can be employed as potentially more accurate. Easily automated ways of identifying risk of patients for conditions. Good risk prediction methods do not exist and some existing methods give very poor performance. Future studies should attempt to test, automate, and validate the models, with the aim of reducing the occurrence of undiagnosed disease and the burden of hostile results related to delays in preventive interventions [6].

III. HEART DISEASE AND RISKS

The heart related diseases consumes around a million lives of peoples every year, creating this as the primary reason. One death among three is due to heart disease in the United States (US). In the year 2016, around 9,20,000 peoples had heart attacks, and nearly half of them occurred suddenly without prior symptoms. Sudden death is the only symptom for heart disease. A Heart attack happens when the internal organs of the heart or muscles started to expire though the person survives for a certain period of time.

This will result in severe problems in the parts of the heart, leading to health issues like increased risk of immediate cardiac arrest. Current estimates tell that more number of heart disease encumbrances will be in India than all other countries. One death among five is due to heart problems in India. During the next three years, it

© 2019 IJSRET

will be one death among every three deaths. 6 Miserably, most of them belonging to young age, particularly in India. In India, heart disease happens1 to 1.5 decades in advance when compared to the western countries. An estimation reports that there are around 45,000,000 will be affected by heart problems. Day to day more youngsters in India are affected by heart disease. As the millions of people addicted to machine life, the upcoming generations seem to be unattractive. There seems to be a stable rise in hypertension pervasiveness for the previous 5 decades, which is extra in city areas than in pastoral zones. It is 25 to 30 percent in urban and 10 to 50 percent in rural zones. Inactive lifestyle is a main reason of death, disease and disability which multiplies the danger of heart disease.

IV. CONCLUSIONS

1. Several surveys have come up with the result that heart diseases are among the top five reasons of deaths worldwide.
2. Conventional diagnosis of heart diseases relying on symptoms and standard tests like the ECG highly depends upon the experience of the physician and the fluctuations in the test data.
3. Errors arising out of any of the above factors may have serious repercussions.
4. In this previous research present a technique in which consists of data pre-processing using the wavelet transform, classifying using the Euclidean Distance Classifier.

REFERENCES

[1]. Akhil Jabbar, M, Deekshatulu, BL & Priti Chandra 2013, 'Classification of Heart Disease using Artificial Neural Network and Feature Subset Selection', Global Journal of Computer Science and Technology, vol.13, no.3, pp. 5-14.

[2]. Anbarasi, M, Anupriya, E & Iyengar, N 2010, 'Enhanced Prediction of Heart Disease with Feature Subset Selection using Genetic Algorithm', International Journal of Engineering Science and Technology, vol.2, no.10, pp. 5370-5376.

[3]. Anchana Khemphila & Veera Boonjing 2011, 'Heart disease Classification using Neural Network and Feature Selection', In Proceedings of 21st International Conference on Systems Engineering, IEEE computer society, Washington, USA pp. 406-409.

[4]. Anooj, PK 2012, 'Clinical Decision Support System : Risk Level Prediction of Heart Disease Using Decision Tree Fuzzy Rules', Asian Transactions on Computers Journal of Engineering, Computing, Sciences & Technology, vol.2, no.4, pp.1-11.

[5]. Anuja Kumari, V & Chitra, R 2013, 'Classification of Diabetes Disease Using Support Vector Machine'International Journal of Engineering Research and Applications, vol. 3, no. 2, pp. 1797-1801.

[6]. Aishwarya, R, Gayatri, P & Jaisankar, NI 2013, 'A Method for Classification Using Machine Learning Technique for Diabetes', International Journal of Engineering and Technology, vol. 5, no. 3, pp.2903-2908.

[7]. Asha Rajkumar & Sophia Reena, G 2010, 'Diagnosis of Heart Disease Using Data mining Algorithm', Global Journal of Computer Science and Technology, vol. 10, no. 10, pp. 38-43.

[8]. Asha Gowda Karegowda, Manjunath AS & Jayaram MA 2011, 'Application of Genetic Algorithm Optimized Neural Network Connection Weights for Medical Diagnosis of Pima Indians Diabetes', International Journal on Soft Computing, vol. 2, no. 2, pp.15-23.

[9]. Aqueel Ahmed & Shaikh Abdul Hanan 2012, 'Data Mining Techniques to Find Out Heart Diseases An Overview', International Journal of Innovative Technology and Exploring Engineering, , vol. 1, no. 4, pp. 18-23.

[10]. Balasundar, V, Devi, T & Saravanan, N 2012, 'Development of Data Clustering Algorithm for Predicting Heart', International Journal of Computer Applications, vol. 48, no. 7, pp. 8-13.

© 2019 IJSRET