

PREDICTING HEART DISEASE USING MACHINE LEARNING

¹KHAJA NAZEERUDDIN,²M.MAHESH,³R. DHEERAJ

^{1,2,3}Assistant Professor

Department of CSE

Visvesvaraya College Of Engineering & Technology, Ibrahimpatnam, Telangana

ABSTRACT

The heart is an important organ in living things. The diagnosis and prognosis of heart-related disorders need more accuracy, perfection, and precision since even a little error may result in exhaustion or even death; the number of heart-related deaths is growing dramatically every day. In order to address the issue, a prediction system for illness awareness is vital. Artificial Intelligence (AI) has a branch called machine learning, which offers distinguished assistance in forecasting any kind of occurrence by using natural events as training data. In this study, we use the UCI repository dataset for training and testing to determine the accuracy of machine learning algorithms for heart disease prediction. The methods include k-nearest neighbor, decision tree, linear regression, and support vector machine (SVM). The Anaconda (jupyter) notebook is the greatest tool for implementing Python programming; it has a variety of libraries and header files that improve the accuracy and precision of the task.

I. INTRODUCTION

Heart is one of the most extensive and vital organ of human body so the care of heart is essential. Most of diseases are related to heart so the prediction about heart diseases is necessary and for this purpose comparative study needed in this field, today most of patient are died because their diseases are recognized at last stage due to lack of accuracy of instrument so there is need to know about the more efficient algorithms for diseases prediction.

Machine Learning is one of the efficient technology for the testing, which is based on training and testing. It is the branch of Artificial Intelligence(AI) which is one of broad area of learning where machines emulating human abilities, machine learning is a specific branch of AI. On the other hand machines learning systems are trained to learn how to process and make use of data hence the combination of both technology is also called as Machine Intelligence.

As the definition of machine learning, it learns from the natural phenomenon, natural things so in this project we uses the biological parameter as testing data such as cholesterol, Blood pressure, sex, age, etc. and on the basis of these, comparison is done in the terms of accuracy of algorithms such as in this project we have used four algorithms which are decision tree, linear regression, k-neighbour, SVM.

In this paper, we calculate the accuracy of four different machine learning approaches and on the basis of calculation we conclude that which one is best among them.

Section 1 of this paper consist the introduction about the machine learning and heart diseases. Section II described, the machine learning classification. Section III illustrated the related work of researchers. Section IV is about the methodology used for this prediction system. Section V is about the algorithms used in this project. Section VI briefly describes the dataset and their analysis with the result of this project. And the last Section VII concludes the summary of this paper with slight view about future scope of this paper.

II. LITERATURE SURVEY

1)Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques

AUTHORS: Senthil kumar mohan, chandrasegar thirumalai and Gautam Srivastva

Heart disease is one of the most significant causes of mortality in the world today. Prediction of cardiovascular disease is a critical challenge in the area of clinical data analysis. Machine learning (ML) has been shown to be effective in assisting in making decisions and predictions from the large quantity of data produced by the healthcare industry. We have also seen ML techniques being used in recent developments in different areas of the Internet of Things (IoT). Various studies give only a glimpse into predicting heart disease with ML techniques. In this paper, we propose a novel method that aims at finding significant features by applying machine learning techniques resulting in improving the accuracy in the prediction of cardiovascular disease. The prediction model is introduced with different combinations of features and several known classification techniques. We produce an enhanced performance level with an accuracy level of 88.7% through the prediction model for heart disease with the hybrid random forest with a linear model (HRFLM).

2)Prediction of Heart Disease using Machine Learning Algorithms: A Survey

AUTHORS: Himanshu Sharma and M A Rizvi

Health care field has a vast amount of data, for processing those data certain techniques are used. Data mining is one of the techniques often used. Heart disease is the Leading cause of death worldwide. This System predicts the arising possibilities of Heart Disease. The outcomes of this system provide the chances of occurring

heart disease in terms of percentage. The datasets used are classified in terms of medical parameters. This system evaluates those parameters using data mining classification technique. The datasets are processed in python programming using two main Machine Learning Algorithm namely Decision Tree Algorithm and Naive Bayes Algorithm which shows the best algorithm among these two in terms of accuracy level of heart disease.

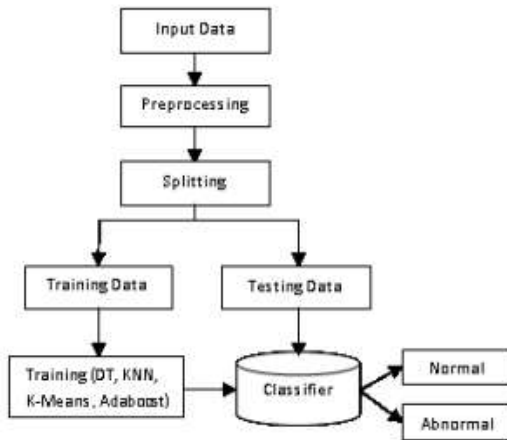
III. EXISTING SYSTEM

Very few systems use the available clinical data for prediction purposes and even if they do ,they are restricted by the large number of association rules that apply.Diagnosis of the condition solely depends upon the Doctor's intuition and patient's records.The decision support system and will prove to be an aid for the physicians with the diagnosis.The algorithm,Fuzzy c means uses clustering and makes use of clusters and data points to predict the relativity of an attribute .Each data point is associated with multiple clusters depending upon the membership degrees

IV. PROPOSED SYSTEM

To deal with the problem there is essential need of prediction system for awareness about diseases. Machine learning is the branch of Artificial Intelligence(AI), it provides prestigious support in predicting any kind of event which take training from natural events.we calculate accuracy of machine learning algorithms for predicting heart disease, for this algorithms are k-nearest neighbor, decision tree, linear regression and support vector machine(SVM) by using UCI repository dataset for training and testing

V. SYSTEM ARCHITECTURE



VI. IMPLEMENTATION

Users:

User add the data to the database and view the data to the view data and predict the heart disease using ml.

Data Collection:

First step for predication system is data collection and deciding about the training and testing dataset. In this project we have used 73% training dataset and 37% dataset used as testing dataset the system.

Attribute Selection:

Attribute of dataset are property of dataset which are used for system and for heart many attributes are like heart bit rate of person, gender of the person, age of the person and many more predication system.

Preprocessing of data:

Preprocessing needed for achieving prestigious result from the machine learning algorithms. For example Random forest algorithm does not support null values dataset and for this we have to manage null values from original raw data. For our project we have to convert some

categorized value by dummy value means in the form of “0”and “1” by using following code

Admin:

Admin will give authority to Users. In order to activate the users. the admin can Prediction Heart Disease.

VII. CONCLUSION

Since the heart is one of the body's most significant and crucial organs and heart disease prediction is a major worry for people, algorithm accuracy is one of the parameters used to analyze how well algorithms operate. The dataset that is utilized for testing and training determines how accurate the machine learning algorithms are. The KNN method is the best one, according to our examination of the algorithms using the dataset whose properties are shown in TABLE.1 and the confusion matrix.

REFERENCES

- [1] Santhana Krishnan J and Geetha S, “Prediction of Heart Disease using Machine Learning Algorithms” ICICT, 2019.
- [2] Aditi Gavhane, Gouthami Kokkula, Isha Panday, Prof. Kailash Devadkar, “Prediction of Heart Disease using Machine Learning”, Proceedings of the 2nd International conference on Electronics, Communication and Aerospace Technology(ICECA), 2018.
- [3] Senthil kumar mohan, chandrasegar thirumalai and Gautam Srivastva, “Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques” IEEE Access 2019.
- [4] Himanshu Sharma and M A Rizvi, “Prediction of Heart Disease using Machine Learning Algorithms: A Survey” International Journal on Recent and Innovation Trends in Computing and Communication Volume: 5 Issue: 8 ,IJRITCC August 2017.

- [5] M. Nikhil Kumar, K. V. S. Koushik, K. Deepak, "Prediction of Heart Diseases Using Data Mining and Machine Learning Algorithms and Tools" International Journal of Scientific Research in Computer Science, Engineering and Information Technology ,IJSRCSEIT 2019.
- [6] Amandeep Kaur and Jyoti Arora, "Heart Diseases Prediction using Data Mining Techniques: A survey" International Journal of Advanced Research in Computer Science , IJARCS 2015-2019.
- [7] Pahulpreet Singh Kohli and Shriya Arora, "Application of Machine Learning in Diseases Prediction", 4th International Conference on Computing Communication And Automation(ICCCA), 2018.
- [8] M. Akhil, B. L. Deekshatulu, and P. Chandra, "Classification of Heart Disease Using K- Nearest Neighbor and Genetic Algorithm," Procedia Technol., vol. 10, pp. 85–94, 2013.
- [9] S. Kumra, R. Saxena, and S. Mehta, "An Extensive Review on Swarm Robotics," pp. 140–145, 2009.
- [10] Hazra, A., Mandal, S., Gupta, A. and Mukherjee, " A Heart Disease Diagnosis and Prediction Using Machine Learning and Data Mining Techniques: A Review" Advances in Computational Sciences and Technology , 2017.
- [11] Patel, J., Upadhyay, P. and Patel, "Heart Disease Prediction Using Machine learning and Data Mining Technique" Journals of Computer Science & Electronics , 2016.
- [12] Chavan Patil, A.B. and Sonawane, P. "To Predict Heart Disease Risk and Medications Using Data Mining Techniques with an IoT Based Monitoring System for Post-Operative Heart Disease Patients" International Journal on Emerging Trends in Technology, 2017.
- [13] V. Kirubha and S. M. Priya, "Survey on Data Mining Algorithms in Disease Prediction," vol. 38, no. 3, pp. 124–128, 2016.
- [14] M. A. Jabbar, P. Chandra, and B. L. Deekshatulu, "Prediction of risk score for heart disease using associative classification and hybrid feature subset selection," Int. Conf. Intell. Syst. Des. Appl. ISDA, pp. 628–634, 2012.
- [15] <https://archive.ics.uci.edu/ml/datasets/Heart+Disease>