International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 03 | March - 2022 ISSN: 2582-3930

Heart Disease Prediction

Khan Nasreen, Mhaske Utkarsha, Nimbalkar Rutuja , Gite Pratiksha

Khan Nasreen Computer Engineering & Anantrao Pawar College Of Engineering And Research
Mhaske Utkarsha Computer Engineering & Anantrao Pawar College Of Engineering And Research
Nimbalkar Rutuja Computer Engineering & Anantrao Pawar College Of Engineering And Research
Gite Pratiksha Computer Engineering & Anantrao Pawar College Of Engineering And Research

---

Abstract - Heart disease (HD) is one of the most common diseases nowadays, and an early diagnosis of such a disease is a crucial task for many health care providers to protect their patients from such a disease and to save lives. The treatment cost of heart disease is not affordable for most of the patients. So we can reduce this problem by a Heart Disease Prediction System (HDPS). It is helpful for earlier diagnosis of heart disease The system can predict the likelihood of patients getting a heart disease by using medical profiles such as age, sex, blood pressure, cholesterol and blood sugar. Also, the performance will be compared by calculation of confusion matrix. This can help to calculate accuracy, precision, and recall. The overall system provides high performance and better accuracy. The heart is one of the main organs of the human body. It pumps blood through blood vessels of the circulatory system. The circulatory system is extremely important because it takes care of activities like transportation of blood, oxygen and other materials to the different organs of the body. Heart plays the most crucial role in the circulatory system. If the heart does not function properly then it will lead to serious health conditions sometimes even to death. The diagnosis of heart disease is based on signs, symptoms and physical examination of the patient. There are several factors that increase the risk of heart disease, such as smoking habit, body cholesterol level, family history of heart disease, obesity, high blood pressure and lack of physical exercise. The existing system does not manage the clinical details. Existing system leads to failure in case of any inconsistencies and missing data. The current automated system is used to identify the key patterns or features from the medical data by using the main classifier model.

***Key Words*:** Cleveland Heart Disease Database, Decision Trees, Random forest, Hybrid algorithm, Machine learning

1.INTRODUCTION

Data mining is useful for studying and understanding a large amount of data. It is used for the extraction of data and to make the decision for further applications. The most common techniques covered under data mining are clustering, association rule mining, and classifications. There are plenty of algorithms available for implementing these data mining techniques. Though there are tools like weka are available for simulations, Python programming is emerging with these algorithms built with scikit learn packages. Thus, the real-time implementation of data mining concepts is more reliable than ever. Machine learning usage is growing vastly in the medical diagnosis industry, where the manual error can be reduced with computer analysis, and accuracy is improved. The diagnosis of a disease is highly reliable with machine learning techniques. Disease such as heart disease, liver disease, diabetes, tumor predictions is done through machine learning concepts [18]. classification algorithms such as decision tree, naïve bayes and SVM (Support Vector Machine) are available;

similarly, regression algorithms, namely Random forest, lasso, and logistic regressions, were used in the medical industry. In most of the tumor predictions, deep learning algorithms are largely used in the medical diagnosis field. As per survey reports, each year, nearly 17 million deaths occurred due to cardiovascular diseases (CVD). The early detection of disease may save many lives, and mortality can be reduced if the patients take their treatments on time [19]. Cardiovascular diseases include many threats such as heart disease, and all etc. With the lack of physical activity due to the lifestyle changes, these diseases are becoming very common even in the lesser age groups. Smoking, lack of physical exercise, high cholesterol food, junk food, living habits are the leading causes of heart disease.

2.MOTIVATION

Heart disease prediction is an android-based machine learning application, trained by a dataset. The user inputs its specific medical details to get the prediction of heart disease for that user. The algorithm will calculate the probability of presence of heart disease. The result will be displayed on the mobile app itself. Thus, minimizing the cost and time required to predict the disease. The system is implemented using KNN algorithm. The algorithm will be trained using the dataset. Some steps will be taken for optimizing the algorithms thereby improving the accuracy. These steps include cleansing the dataset and data pre processing. The main application is an android application which accepts the various parameters from the user as input and computes.

3.LITRETURE SURVEY

Ankita Dewan; Meghna Sharma“Prediction of heart disease using a hybrid technique in data mining classification” [1]This study demonstrates a low-cost smart device that can ensure women's safety. The device can protect ladies who are out on their own. Women's participation in school and work will rise as a result of the use of this technology. It's a smart device with additional functions that make it more appealing. There are certain drawbacks to it. Some algorithms are tweaked after a trial run in a minor way. If the user makes a mistake and sets off a false alert, she can send a corrective SMS to the authority as soon as she realises the error. By following the law and following the laws, the prize gun is the most effective self-defense weapon. This smart device is a one-stop shop for all of a woman's security concerns. Women will feel confident and free to go outside without fear.

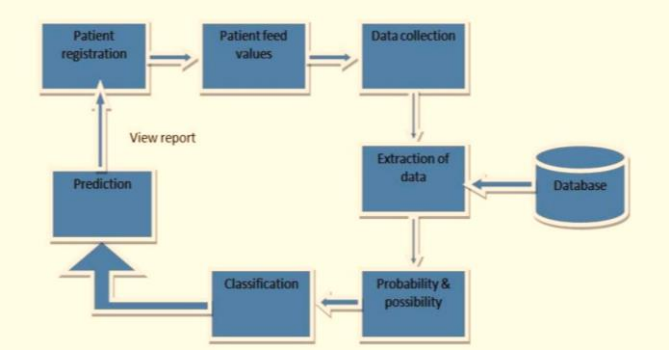
Md. Touhidul Islam; Sanjida Reza Rafa; Md. Golam Kibria “Early Prediction of Heart Disease Using PCA and Hybrid Genetic Algorithm with k-Means” [2]According to study, millions of lives are lost each year due to heart disease. Massive amounts of data on heart disease are produced by the healthcare industry, but they are tragically not utilised to locate secret information for effective decision-making. Detecting cardiac disease at an early stage is one of the most critical components right now. The UCI Machine Learning heart disease dataset has been subjected to a variety of approaches. Many academics have attempted to apply advanced methodologies to this dataset, which currently lacks extensive analyses. Principal Component Analysis (PCA) was used to decrease characteristics in this article. Aside from the final clustering, a hybrid genetic algorithm (HGA) with k-means was applied. The k-means approach is commonly used for clustering data. Because this is a heuristic strategy, this form of clustering can get caught in local optima. To avoid this issue, we employed the Hybrid Genetic Algorithm (HGA) for data clustering. Our proposed technique has a 94.06 percent accuracy in predicting early cardiac disease.

Aditi Gavhane; Gouthami Kokkula; Isha Pandya; Kailas Devadkar "Prediction of Heart Disease Using Machine Learning" [3]With the rampant increase in the heart stroke rates at juvenile ages, we need to put a system in place to be able to detect the symptoms of a heart stroke at an early stage and thus prevent it. It is impractical for a common man to frequently undergo costly tests like the ECG and thus there needs to be a system in place which is handy and at the same time reliable, in predicting the chances of a heart disease. Thus we propose to develop an application which can predict the vulnerability of a heart disease given basic symptoms like age, sex, pulse rate etc. The machine learning algorithm neural networks has proven to be the most accurate and reliable algorithm and hence used in the proposed system.

Chittampalli Sai Prakash; Myneni Madhu Bala; Attluri Rudra "Data Science Framework - Heart Disease Predictions, Variant Models and Visualizations" [4]Most of the poultry farms in Brunei are manually monitored and controlled. Temperature, humidity, air quality level, lighting, ventilation and food feeding are the important factors that are needed to be monitored and controlled. These factors are directly related to the poultry production. Currently, the mortality rate of broiler chicken in Brunei is higher than the normal rate. This research project is aim to produce healthy chickens and reduce the mortality rate of chicken to improve the productivity in Brunei by automating the process of monitoring and maintaining the temperature, humidity, air quality level and food feeder effectively using Internet of Things (IoT) and Wireless Sensor Networks (WSN). A prototype was created using IoT and WSN technologies and the above parameters were tested against threshold values. When these parameters exceeded the threshold values, corrective processes are initiated automatically that can help to reduce the mortality rate of chickens in the farm. This system also sends automatic alert notification to the user through SMS, Email and WhatsApp. A Web interface is also created to monitor and display these parameters.

M. Kavitha; G. Gnaneswar; R. Dinesh; Y. Rohith Sai; R. Sai Suraj "Heart Disease Prediction using Hybrid machine Learning Model"[5] Heart disease causes a significant mortality rate around the world, and it has become a health threat for many people. Early prediction of heart disease may save many lives; detecting cardiovascular diseases like heart attacks, coronary artery diseases etc., is a critical challenge by the regular clinical data analysis. Machine learning (ML) can bring an effective solution for decision making and accurate predictions. The medical industry is showing enormous development in using machine learning techniques. In the proposed work, a novel machine learning approach is proposed to predict heart disease. The proposed study used the Cleveland heart disease dataset, and data mining techniques such as regression and classification are used. Machine learning techniques Random Forest and Decision Tree are applied. The novel technique of the machine learning model is designed. In implementation, 3 machine learning algorithms are used, they are 1. Random Forest, 2. Decision Tree and 3. Hybrid model (Hybrid of random forest and decision tree). Experimental results show an accuracy level of 88.7% through the heart disease prediction model with the hybrid model. The interface is designed to get the user's input parameter to predict the heart disease, for which we used a hybrid model of Decision Tree and Random Forest.

4.SYSTEM ARCHITECTURE



5.PROPOSED SYSTEM

Our aim with this system is to be the connecting bridge between doctors and patients without wasting time. We are proposing such a system which will flaunt a simple user interface with a rapid solution which consumes less time. The main feature will be the system allows users to share their report numbers. It then processes the user's values to check for various risks that could be associated with it. In this we use the machine learning techniques, in which we will be using Random Forest as well as K-Nearest Neighbors (KNN) Algorithms to gain accurate prediction. As we know both techniques are simple, easy-to-implement supervised machine learning algorithm that can be used to solve both classification and regression problems After delivering the results, our system will also suggest the user about doctors to avoid the risk. System provide and easy communication interface for patient to get instant diagnosis.

6.LIMITATIONS

•It does not recommend medications for the disease.

•It requires active internet connection.

•Incorrect Input Values may result in wrong prediction

7. CONCLUSIONS

We set out to develop a system that can forecast sickness based on input parameters. By providing early warning, such a system can help India reduce the rate of late treatment. The use of a communication interface can make it easier to grasp specific concerns from any location. Additionally, our system offers a user-friendly interface. It also includes a number of graphic representations of the data and outcomes gathered.

ACKNOWLEDGEMENT

I am sincerely thankful to all my teachers for their guidance for my seminar. Without their help it was tough job for me to accomplish this task. I am especially very thankful to my guide…..for his consistent guidance, encouragement and motivation throughout the period of this work. I also want to thank our Head of the Department -------- for providing me all necessary facilities.

REFERENCES

[1] Jabbar, M. A., B. L. Deekshatulu, and Priti Chandra. "Intelligent heart disease prediction system using random forest and evolutionary approach." Journal of Network and Innovative Computing 4.2016 (2016): 175-184.
[2] Alkeshuosh, Azhar Hussein, et al. "Using PSO algorithm for producing best rules in diagnosis of heart disease." 2017 international conference on computer and applications (ICCA). IEEE, 2017.
[3] Al-Milli, Nabeel. "Backpropagation neural network for prediction of heart disease." Journal of theoretical and applied information Technology 56.1 (2013): 131-135.
[4] Mythili, T., et al. "A heart disease prediction model using SVM-Decision Trees-Logistic Regression (SDL)." International Journal of Computer Applications 68.16 (2013).
[5] Detrano R. VA Medical Center, Long Beach and Cleveland Clinic Foundation: Robert Detrano, MD (Doctoral dissertation, Ph. D., Donor: David W. Aha, 1998.
[6] Xing, Yanwei, Jie Wang, and Zhihong Zhao. "Combination data mining methods with new medical data to predicting outcome of coronary heart disease." 2007 International Conference on Convergence Information Technology (ICCIT 2007). IEEE, 2007.
[7] Chen, Jianxin, et al. "Predicting syndrome by NEI specifications: a comparison of five data mining algorithms in coronary heart disease." International Conference on Life System Modeling and Simulation. Springer, Berlin, Heidelberg, 2007.

[8] Soni, Jyoti, et al. "Predictive data mining for medical diagnosis: An overview of heart disease prediction." International Journal of Computer Applications 17.8 (2011): 43-48.
[9] Singh, A., et.al (2020, February). Heart Disease Prediction Using Machine Learning Algorithms. In 2020 International Conference on Electrical and Electronics Engineering (ICE3) (pp. 452-457). IEEE.
[10] Hashi, E.K. and Zaman, M.S.U., 2020. Developing a Hyperparameter Tuning Based Machine Learning Approach of Heart Disease Prediction. Journal of Applied Science & Process Engineering, 7(2), pp.631-647.
[11] Shouman, Mai, Tim Turner, and Rob Stocker. "Using data mining techniques in heart disease diagnosis and treatment." 2012 Japan-Egypt Conference on Electronics, Communications and Computers. IEEE, 2012.
[12] Mohan, Senthilkumar, Chandrasegar Thirumalai, and Gautam Srivastava. "Effective heart disease prediction using hybrid machine learning techniques." IEEE Access 7 (2019): 81542-81554.
[13] Ramalingam, V. V., Ayantan Dandapath, and M. Karthik Raja. "Heart disease prediction using machine learning techniques: a survey." International Journal of Engineering & Technology 7.2.8 (2018): 684-687.
[14] Polat, Kemal, Seral Şahan, and Salih Güneş. "Automatic detection of heart disease using an artificial immune recognition system (AIRS) with fuzzy resource allocation mechanism and k-nn (nearest neighbour) based weighting preprocessing." Expert Systems with Applications 32.2 (2007): 625-631.
[15] Palaniappan, Sellappan, and Rafiah Awang. "Intelligent heart disease prediction system using data mining techniques." 2008 IEEE/ACS international conference on computer systems and applications. IEEE, 2008.
[16] Das, Resul, Ibrahim Turkoglu, and Abdulkadir Sengur. "Effective diagnosis of heart disease through neural networks ensembles." Expert systems with applications 36.4 (2009): 7675-7680.
[17] Jonnavithula, et.al (2020, October). Role of machine learning algorithms over heart biseases prediction. In AIP Conference Proceedings (Vol. 2292, No. 1,p.040013). AIP Publishing LLC. 4. Michalewicz, Z.: Genetic Algorithms + Data Structures = Evolution Programs. 3rd edn. Springer-Verlag, Berlin Heidelberg New York (1996)