Myocardial Infraction Predication using Machine Learning

Mounica Manasa P[1], Priya Darshini S[2], Naveen Kumar V[3], Maheswari M[4], Roselin Mary S[5]

1. Student, Computer science and engineering, Anand Institute of Higher Technology,*Chennai, India.*
2. *Student, Computer Science and Engineering, Anand Institute of Higher Technology,Chennai, India.*
3. *Assistant Professor, Computer Science and Engineering, Anand Institute of HigherTechnology, Chennai, India.*
4. *Assistant Professor, Computer Science and Engineering, Anand Institute of HigherTechnology, Chennai, India.*
5. *Head of Department, Computer Science and Engineering, Anand Institute of HigherTechnology, Chennai, India.*

ABSTRACT

Heart disease is one of the complex diseases and globally many people suffered from this disease. On time and efficient identification of heart disease plays a key role in healthcare, particularly in the field of cardiology. We proposed an efficient and accurate system to diagnosis heart disease and the system is based on machine learning techniques. The system is developed based on classification algorithms includes Support vector machine, Logistic regression, Artificial neural network, K-nearest neighbor, Naïve bays, and Decision tree while standard features selection algorithms. The research paper mainly focuses on which patient is more likely to have a heart disease based on various medical attributes. We prepared a heart disease prediction system to predict whether the patient is likely to be diagnosed with a heart disease or not using the medical history of the patient. The strength of the proposed model was quiet satisfying and was able to predict evidence of having a heart disease in a particular individual by using random forest, KNN and Logistic Regression which showed a good accuracy in comparison which is used prediction. So, a quiet significant amount of pressure has been lift off by using the given model in finding the probability of the classifier to correctly and accurately identify the heart disease and got 91% Accuracy.

.

Keywords: Logistic Regression (LR), Machine Learning (ML), Features, Classifiers , Myocardial Infraction (MI), Electrocardiogram (ECG).

I. INTRODUCTION:

Myocardial infarction, usually called a coronary heart attack, is a severe medical situation which can have life-threatening consequences if now not handled promptly. Gadget studying can play a critical position in helping healthcare specialists to discover patients at high hazard of myocardial infarction, enhance prognosis, and expand customized treatment plans.

The use of machine getting to know algorithms, it's far feasible to analyse massive volumes of affected person information, consisting of medical history, scientific symptoms, and take a look at results, to pick out patterns and predictive factors related to myocardial infarction. By developing predictive models, healthcare experts can become aware of sufferers who're at higher chance of experiencing a coronary heart attack and enforce preventative measures to lessen their danger.

Moreover, gadget gaining knowledge of also can be used to enhance diagnostic accuracy through reading clinical pix and detecting subtle adjustments that could imply the presence of a coronary heart assault. This could assist healthcare experts to make a timely diagnosis and initiate appropriate treatment.

II. PROBLEM STATEMENT:

Heart disease is a leading reason of demise global, and early detection can improve patient outcomes and decrease healthcare charges. Gadget gaining knowledge of techniques may be used to predict the chance of heart ailment, based on a aggregate of affected person traits and medical take a look at results.

The trouble announcement for heart disorder prediction using gadget learning entails growing a predictive version which can as it should be pick out people coronary heart ailment. The model have to be taught on a dataset of affected person facts.

The purpose of the version is to are expecting the likelihood of coronary heart disease based totally on these enter features. This may assist healthcare vendors identify sufferers who might also advantage from similarly checking out or treatment and take proactive steps to prevent or manipulate heart disease.

To obtain this intention, the version have to learn and evaluated the usage of appropriate system mastering algorithms and strategies and optimized to gain excessive accuracy and reliability in predicting coronary heart disease. The resulting model may be used as a tool for healthcare vendors to improve patient effects and decrease healthcare expenses related to coronary heart disease.

III. RELATED WORKS:

Myocardial infarction(MI), usually known as a coronary heart assault, is a existence-threatening condition that happens whilst the blood deliver to the coronary heart is interrupted, resulting in harm to the coronary heart muscle. [1]Early detection of MI is crucial for well timed and effective remedy, as not on time treatment can result in further damage to the heart and boom the threat of mortality.

Machine learning (ML) has shown amazing potential in enhancing the accuracy and velocity of MI detection.[3] ML algorithms can analyse large datasets of affected person facts, which include medical records, lab consequences, and imaging research, to identify styles and chance factors associated with MI. These algorithms can then be used to broaden predictive fashions that could help clinicians in making correct diagnoses and remedy decisions.

One application of ML is MI detection is the use of electrocardiogram (ECG) information.[5] ECG's interpretation can be tough, even for knowledgeable clinicians.ML algorithms can examine ECG information to discover diffused adjustments that can be indicative of MI ,taking o consideration earlier detection and remedy . Every other area in which ML is being carried out is inside the evaluation of imaging research, inclusive of coronary angiograms and cardiac magnetic resonance imaging (MRI).[6]ML algorithms can analyze those pix to pick out structural abnormalities and blockages inside the coronary arteries which could cause MI.This could assist clinicians in making more correct diagnoses and treatment choices.

ML is also getting used to develop risk prediction fashions for MI[9].Those fashions can examine big datasets of affected person information to perceive danger factors and broaden customized chance profiles for sufferer .This could assist clinicians in identifying sufferers who can be at expanded danger of MI and growing targeted prevention techniques.ML has shown awesome ability in improving the accuracy and pace of MI detection.[10]ML algorithms can examine big datasets of affected person facts to perceive patterns and hazard elements associated with MI,and broaden predictive fashions that may assist clinicians in making correct diagnoses and remedy choices.ML is probable to play an increasing number of critical function within the early detection and prevention of ML within destiny.

IV. EXISTING SYSTEM:

In present medical systems, together with health center management structures and decision-making systems, attention on collecting and mining the entire clinical records. The whole patient records are loaded, and all factors are taken into consideration .The C4.5 model that we formerly evolved. On the premise of this new degree, a information mining set of rules was advanced to mine the causal dating among capsules and their

associated with chance for heart disease. Coronary heart musicality issues(arrhythmias); and heart abandons you're introduced into the world with (inherent coronary heart deserts), amongst others. The expression "coronary contamination" is frequently utilized reciprocally with the expression "cardiovascular contamination ."Prescient demonstrating makes use of measurements to foresee consequences. Most frequently the event one needs to assume is later on, yet prescient showing may be implemented to an obscure occasion, paying little mind to passed off.

V. PROPOSED SYSTEM:

In the proposed work user will search for the heart disorder analysis (heart ailment and remedy related information) with the aid of giving signs as a query inside the search engine. Those signs are preprocessed to make the similarly system simpler to locate the signs and symptoms key-word which helps to perceive the coronary heart ailment quickly. Wherein the characteristic is most effective approximated regionally, and all computation is deferred till type. This option has been recognized as the maximum appropriate for the existing system. In this project we're the usage of aid vector classifier which presents excessive accuracy above 85%.

VI.SYSTEM DESIGN:

A. SYSTEM MODEL

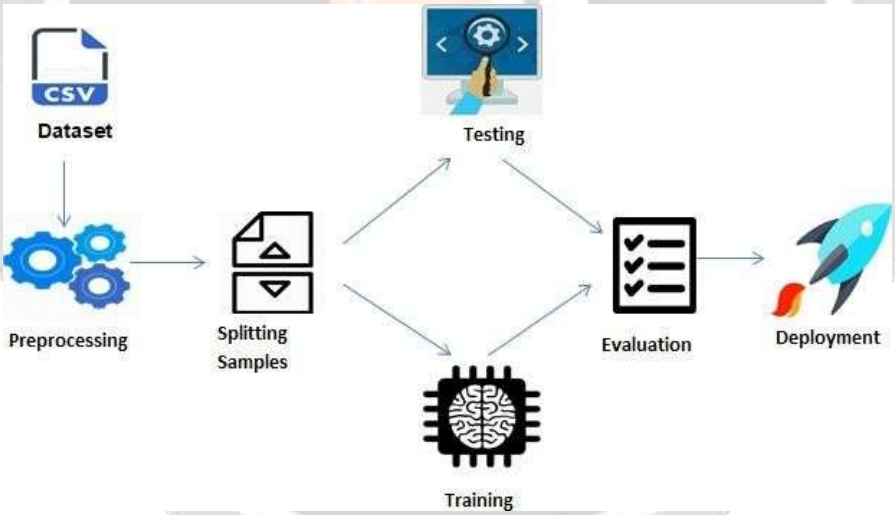

Figure 6.1 System Architecture

B.MODULES

The system is divided into 4 sections. Dataset evaluation is the first module, and it is the method of comparing, cleansing, remodelling, and modelling records with the goal of identifying relevant facts via informing conclusions and assisting desire -making. The second one module is Dataset Pre-processing to cleaning the statistics, as a way to increase the accuracy and efficiency of a machine analysing version the usage of synthetic Minority Oversampling approach (SMOTE). The zero.33 is model checking out and schooling, on this module we use supervised class algorithms for training and trying out our dataset using system mastering algorithms. The fourth module is model Deployment, in this Module we advanced consumer interface.

1. DATA ANALYSIS:

Data evaluation is a manner for acquiring uncooked records and converting it into data beneficial for selection-making by way of users. Facts are amassed and analyzed to answer questions, take a look at hypotheses or

disprove theories. Using pre-gathered records, via manner of datasets from Kaggle, UCI(college of California Irvine device) those are utilized in our survey. Our dataset contains of 919 rows, containing records from Angina and eleven different rows containing features. As it's miles mention in above point ,this dataset has been extensively used to hit upon the risk of coronary heart disorder.

2. DATA PRE-PROCESSING:

Dataset pre-processing is required for cleansing the facts, which increases the accuracy and performance of a device mastering version. those dataset has to be wiped clean from duplicates, correct errors, to deal with missing values, facts type conversions. Dataset pre-processing entails obtaining dataset, importing libraries, finding lacking statistics, encoding categorical facts, dataset splitting into training and checking out .

3. MODEL TRAINING &TESTING:

The goal of training is to reply a query or make a prediction successfully as frequently as possible. on this module we use supervised class algorithms for education our dataset consisting of Random forest, assist Vector gadget, Gaussian Naïve Bayes, k-Nearest neighborhood, Linear regression to educate the version on the cleaned dataset.

4. MODEL DEPLOYMENT:

Integrated Module we've used HTML five for built-in the front end and we imported flask framework as our server, pickle and NumPy as our library programs. right here we used Python 3.eleven as IDE and Spyder integrated Anaconda Navigator as Interpreter. We used localhost http://127.0.0.1:5000/ for host built integrated our the front stop website. Google Chrome browser is used for load integrated our browser.

VII. RESULT AND DISCUSSION:

In project the research of various kinds prediction algorithms are used. The techniques consist of collaborative filtering, content material based filtering and the blessings and drawbacks are defined .The comparative study of diverse techniques mentioned above is provided on this report. As a end result the anyone can input the attributes like BP, sugar stage ,and so on. to locate whether or not they have chances of getting heart diseases or no longer. This additionally allows to are expecting the chances of heart illnesses formerly by means of the usage of machine getting to know algorithms. this is carried out as internet site which is consumer friendly and clean to use with none previous information.

OUTPUT:

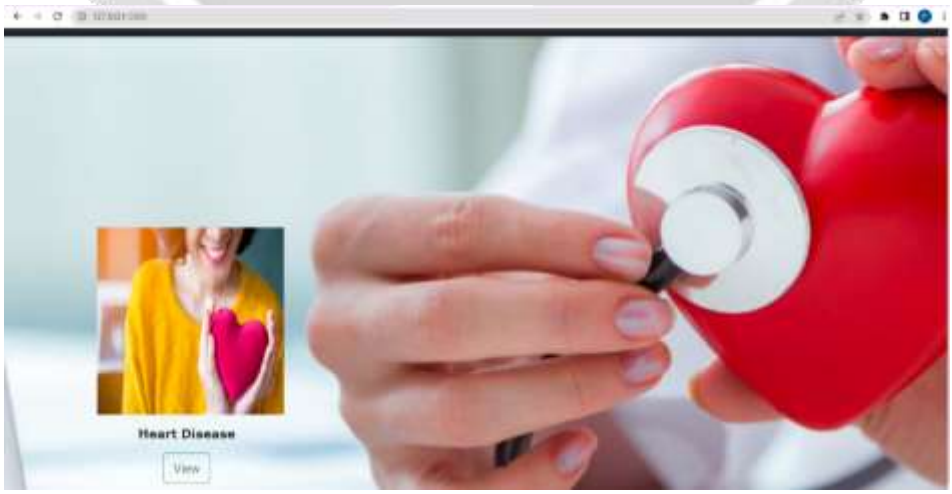

figure 7.1 Heart Disease Webpage

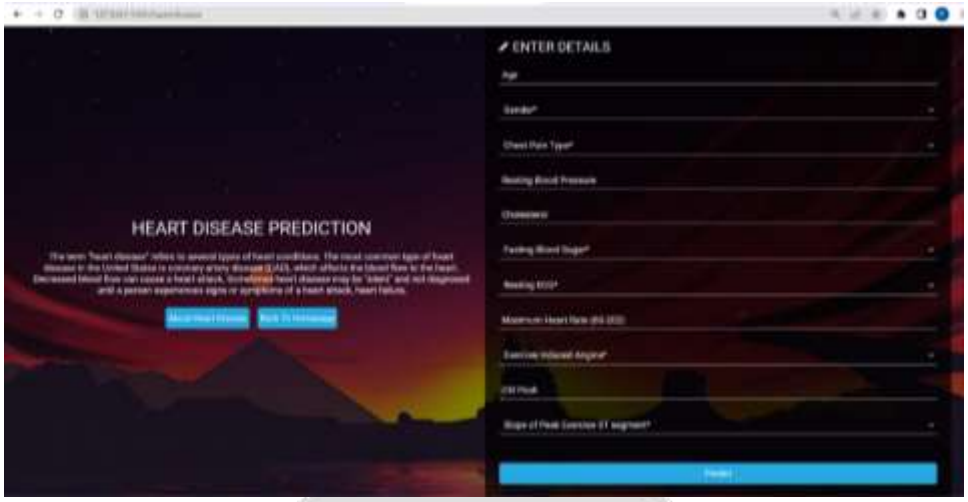

Figure 7.2 User Interface for data entry

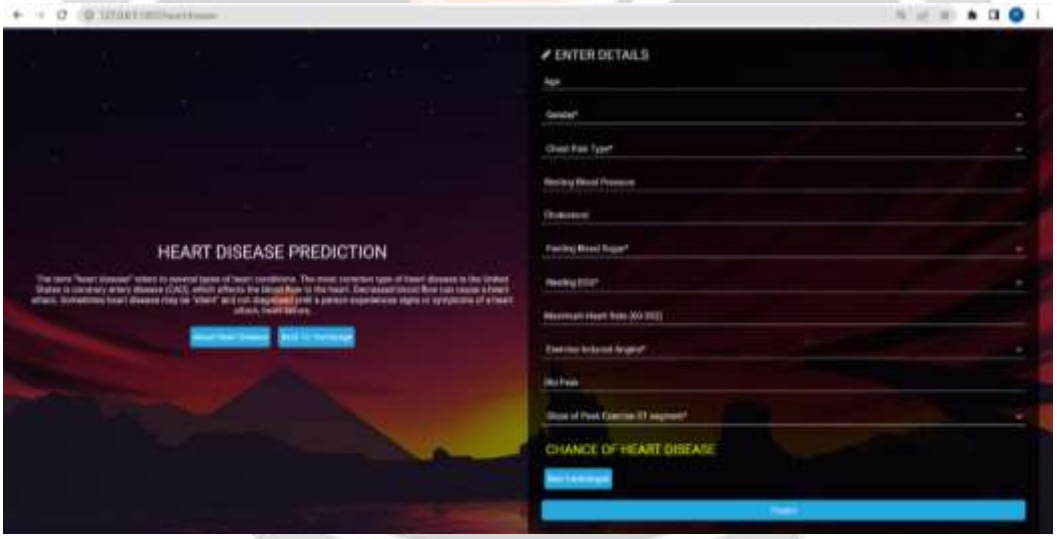

Figure 7.3 User Interface for detecting Chance of Heart disease

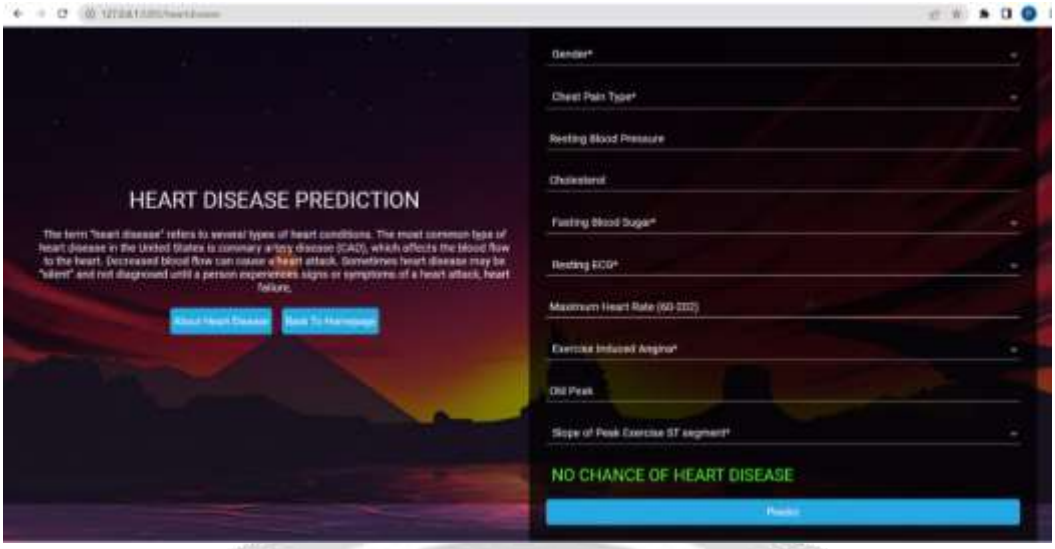

Figure 7.4 User Interface for detecting No Chance of Heart disease

VIII CONCLUSION AND FUTURE WORKS:

In our survey, which we've got performed a survey of comparative analysis of ML techniques to detect Heart disease early. In the process of doing so, by testing all of our classifiers by dropping a couple of features in combinations, Logistic Regression generated the highest accuracy of 91.1% amongst all others. As a result, the Logistic Regression emerges as the best option for detecting Heart disease in its earlier stages.

IX FUTURE WORKS:

As a scope for future enhancement, it is easy to improve the techniques utilized in building the general prediction models, by the usage of custom designed classifiers because of a fusion of two or extra exceptional current ones, which have ended in an ideal accuracy to date. further, there may be a room for improvement while it's miles approximately the general layout of person interface and user experience. a good deal finely tuned dataset, with issue to availability, can act as a catalyst in developing prediction fashions for detecting the presence of coronary heart's disorder, an awful lot in advance and simpler than current models at use. This enhancement might also make it an awful lot greater optimized and powerful to use, for conformity checks, in regard with the recognized sickness.

X REFERENCES:

[1] A. L. Bui, T. B. Horwich, and G. C. Fonarow, ''Epidemiology and risk profile of heart failure,'' Nature Rev. Cardiol., vol. 8, no. 1, p. 30, 2011.

[2] M. Durairaj and N. Ramasamy, ''A comparison of the perceptive approaches for preprocessing the data set for predicting fertility success rate,'' Int. J. Control Theory Appl., vol. 9, no. 27, pp. 255–260, 2016.

[3] L. A. Allen, L. W. Stevenson, K. L. Grady, N. E. Goldstein, D. D. Matlock,R. M. Arnold, N. R. Cook, G. M. Felker, G. S. Francis, P. J. Hauptman,E. P. Havranek, H. M. Krumholz, D. Mancini, B. Riegel, and J. A. Spertus,

''Decision making in advanced heart failure: A scientific statement from the American heart association,'' Circulation, vol. 125, no. 15,pp. 1928–1952, 2012.

[4] S. Ghwanmeh, A. Mohammad, and A. Al-Ibrahim, ''Innovative artificial neural networks-based decision support system for heart diseases diagno-sis,'' J. Intell. Learn. Syst. Appl., vol. 5, no. 3, 2013, Art. no. 35396.

[5] Q. K. Al-Shayea, ‘‘Artificial neural networks in medical diagnosis,’’ Int. J.Comput. Sci. Issues, vol. 8, no. 2, pp. 150–154, 2011.

[6] J. Lopez-Sendon, ‘‘The heart failure epidemic,’’ Medicographia, vol. 33,no. 4, pp. 363–369, 2011.

[7] P. A. Heidenreich, J. G. Trogdon, O. A. Khavjou, J. Butler, K. Dracup,M. D. Ezekowitz, E. A. Finkelstein, Y. Hong, S. C. Johnston, A. Khera,D. M. Lloyd-Jones, S. A. Nelson, G. Nichol, D. Orenstein, P. W. F. Wilson,and Y. J. Woo, ‘‘Forecasting the future of cardiovascular disease in theunited states: A policy statement from the American heart association,’’Circulation, vol. 123, no. 8, pp. 933–944, 2011.

[8] A. Tsanas, M. A. Little, P. E. McSharry, and L. O. Ramig, ‘‘Nonlinearspeech analysis algorithms mapped to a standard metric achieve clinically useful quantification of average Parkinson’s disease symptom severity,’’J. Roy. Soc. Interface, vol. 8, no. 59, pp. 842–855, 2011.

[9] S. I. Ansarullah and P. Kumar, ‘‘A systematic literature review oncardiovascular disorder identification using knowledge mining and machine learning method,’’ Int. J. Recent Technol.Eng., vol. 7, no. 6S,pp. 1009–1015, 2019.

[10] S. Nazir, S. Shahzad, S. Mahfooz, and M. Nazir, ‘‘Fuzzy logic based decision support system for component security evaluation,’’ Int. Arab J.Inf. Technol., vol. 15, no. 2, pp. 224–231, 2018.

[11] R. Detrano, A. Janosi, W. Steinbrunn, M. Pfisterer, J.-J. Schmid, S. Sandhu,K. H. Guppy, S. Lee, and V. Froelicher, ‘‘International application of a new probability algorithm for the diagnosis of coronary artery disease,’’ Amer.J. Cardiol., vol. 64, no. 5, pp. 304–310, Aug. 1989.

[12] J. H. Gennari, P. Langley, and D. Fisher, ‘‘Models of incremental concept formation,’’ Artif. Intell., vol. 40, nos. 1–3, pp. 11–61, Sep. 1989.

[13] Y. Li, T. Li, and H. Liu, ‘‘Recent advances in feature selection and itsapplications,’’ Knowl. Inf. Syst., vol. 53, no. 3, pp. 551–577, Dec. 2017.

[14] J. Li and H. Liu, ‘‘Challenges of feature selection for big data analytics,’’IEEE Intell. Syst., vol. 32, no. 2, pp. 9–15, Mar. 2017.

[15] L. Zhu, J. Shen, L. Xie, and Z. Cheng, ‘‘Unsupervised topic hypergraph hashing for efficient mobile image retrieval,’’ IEEE Trans. Cybern., vol. 47,no. 11, pp. 3941–3954, Nov. 2017.