INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

ISSN
INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 7.488

Machine Learning for Prediction of Cardio Vascular Illnesses

Priyanka M C, Kruthika R , Roshni H S, Vandana C M, Varshitha H P

Assistant Professor, Dept. of ISE, Malnad College of Engineering, Hassan, India

UG Students, Dept. of ISE, Malnad College of Engineering, Hassan, India

ABSTRACT: Heart diseases furthermore referred to as cardio vascular illnesses embody a large type of conditions that have an impact on the coronary heart. These variety from blood vessel illnesses, heart rhythm issues to coronary heart defects that one is born with. It is the number one motive for loss of lifestyles international throughout the previous couple of a few years. It is the want of the hour to attain accurate and dependable technique to advantage early assessment of the disease thru using automating the task and for this reason apprehend inexperienced control of it. Data Science plays an vital in processing massive quantities of facts inside the situation of scientific sciences. Researchers employ several Machine Learning Techniques to investigate massive devices of information and beneficial useful resource within the proper prediction of coronary heart illnesses. This paper analyzes the supervised analyzing models of, K-Nearest Neighbors, Support Vector Machine to predict the heart attack in stage vice and give the precaution based on the stages.

I. INTRODUCTION

Life is simply relying on efficient strolling of the coronary heart. The term Heart infection refers to infection of coronary heart blood vessel device interior tithe coronary heart is an essential organ of human body. If the blood movement to the frame is insufficient, the organs of the body this is mind and coronary heart save you running and demise takes region in short whilst. Heart infection is a main reason of lack of lifestyles worldwide from beyond 15 years. The not unusual hazard elements related are identified as age, circle of relatives records, Sex, Stress, excessive cholesterol ,Heart price, smoking, alcohol consumption, obese, bodily the us of the us of no interest, chest ache kind and lousy diet plan. Information acquired thru studying the statistics document of the affected person, it is possible to isolate the document and provide file on HD if it is awesome or terrible.

The number one challenge recall is prediction the usage of tool analyzing techniques. Machine learning is extensively used now a days in loads of enterprise applications like e exchange and plenty of extra. Prediction is one in every of region where in this device studying used, our mission be counted is ready prediction of coronary heart attack with the useful resource of processing affected individual's dataset and a facts of patients to whom we want to expect the danger of prevalence of a coronary heart illness.

Heart contamination is the most not unusual motive of lack of lifestyles globally. Many health facility statistics structures are designed to useful resource affected man or woman billing, inventory control and technology of easy statistics. Some hospitals use choice guide systems, however they may be in large issue confined. Mining is a manner of exploring big gadgets of facts to take out styles which might be hidden and previously unknown relationships and facts detection to help the higher know-how of clinical facts to prevent coronary heart infection. Classification of coronary Heart Disease can be treasured for the clinical practitioners within the event that it's miles automated with the prevent purpose of brief locating and specific surrender cease result.

Presence of coronary heart sickness precisely can spare patients living days. The artwork includes the education of Heart Disease the usage of Support Vector Machine (SVM) and KNN.

II. RELATED WORK

[1] Mohammed Abdul Khaleel has given paper inside the Survey of Techniques for mining of facts on Medical Data for Finding Frequent Diseases domestically. This paper consciousness on dissect information mining strategies which may be required for medicinal statistics mining specifically to find out regionally go to illnesses, as an example, coronary coronary coronary coronary heart infirmities, lung malignancy, bosom infection etcetera. Information mining is the manner within the course of extricating information for locating inactive examples which Vembandasamy et al. Done a chunk, to research and find out coronary coronary coronary heart illness. In this the set of regulations used changed into Naive Bayes set of guidelines. In Naïve Bayes set of guidelines they used Bayes theorem. Hence Naive Bayes has a completely power to make assumption independently. The used information-set is received from a diabetic studies institutes of Chennai, Tamilnadu this is crucial institute. There are more than 500 sufferers inside the dataset. The tool used is Weka and class is finished via the usage of 70% of Percentage Split. The accuracy supplied with the useful resource of Naive Bayes is 86.419%.

[2] Costas Sideris, Nabi Alshurafa, Haik Kalantarian and Mohammad Pourhomayoun have given a paper named Remote Health Monitoring Outcome Success prediction the use of First Month and Baseline Intervention Data. RHS structures are powerful in saving prices and lowering infection. In this paper, they painting an upgraded RHM framework, Wanda- CVD that is cellular phone based totally truly and imagined to offer some distance flung education and social help to individuals. CVD counteractive movement measures are perceived as a crucial popularity with the aid of social insurance associations round the arena.

[3]. L.Sathish Kumar and A. Padmapriya has given a paper named Prediction for similarities of infection with the beneficial useful resource of using ID3 set of pointers in television and mobile cellphone . This paper offers a programmed and concealed way to cope with recognize designs which may be included up of coronary contamination. The given framework rent information mining strategies, for example, ID3 set of guidelines. This proposed method permits the people not first-class to apprehend approximately the illnesses but it may moreover help's to lessen the dying price and bear in mind sizeable shape of contamination affected humans.

[4]. M.A.Nishara Banu and B.Gomathy has given a paper named Disease Predicting gadget the usage of records mining strategies. In this paper they communicate approximately MAFIA (Maximal Frequent Item set set of rules) and K- Means clustering. As type is essential for prediction of a disorder. The class based on MAFIA and KMeans effects in accuracy.

[5]. Wiharto and Har Kusnanto have given a paper named Intelligence System for Diagnosis Level of Coronary Heart Disease with K-Star Algorithm. In this paper they display off an expectation framework for coronary heart contamination the usage of Learning vector Quantization neural device calculation The neural gadget in this framework acknowledges thirteen clinical includes as information and predicts that there may be a nearness or nonattendance of coronary contamination within the affected person, along numerous execution measures.

III. PROPOSED METHOD

DATA PRE-PROCESSING

Cleaning: Data that we want to manner will now not be easy that is it can embody noise or it is able to encompass values lacking of we method we cant get precise consequences a great way to accumulate specific and exceptional outcomes we want to eliminate all this, the approach to do away with all this is statistics cleaning. We will fill lacking values and can dispose of noise via manner of the usage of some strategies like filling with maximum not unusual rate in missing location.

Transformation: This includes converting information format to one shape to big this is making them maximum understandable with the beneficial resource of doing normalization, smoothing, and generalization, aggregation techniques on statistics.

Integration: Data that we want not way might not be from a single deliver from time to time it is able to be from unique belongings we do now not combine them it may be a hassle at the same time as

processing so integration is taken into consideration surely one in all vital phase in records pre-processing and unique problems are taken into consideration right proper right here to

Reduction: When we paintings on records it could be complex and it can be tough to recognize every now and then if you want to cause them to apprehend-able to tool we're capable of lessen them to required layout truely so we are able to gather properly effects.

Fig 1:Flow Chart

IV. PROPOSED ALGORITHM

SVM algorithm

The SVM is a learning set of regulations for kind. It tries to discover the advanced setting apart hyper plane such that the predicted splendour errors for unseen styles is minimized. For linearly non-separable statistics the input is mapped to excessive- dimensional function area where in they are able to be separated through a hyper aircraft. This projection into high dimensional function place is successfully completed thru using kernels. More precisely, given a tough and fast of education samples and the corresponding preference values -1, 1 the SVM wants to discover the extraordinary keeping apart hyper plane given through the equation WTx+b that maximizes the gap a number of the 2 commands.

KNN algorithm

K-nearest neighbours (KNN) set of policies uses 'feature similarity' to anticipate the values of recent datapoints which further manner that the modern- day records point may be assigned a fee based totally on how carefully it suits the factors within the training set. We can understand its walking with the assist of following steps

PSEUDO CODE

Step 1: Input: Heart disease data set

Step 2: Output : Classification of data set into patients with heart disease and normal Step 3: Input the data set

Step 4: Apply pre-processing techniques-Fill in missing values

Step 5: select the features based on values obtained after applying. Step 6: Discard redundant features (features with low values of PSO) Step 7: Apply (KNN) and SVM on Predominant features

Step 8: Measure the performance of the KNN model SVM model. Step 9: heart attack prediction

V. DATA SET

The information is collected from the UC tool studying repository. The data set is referred to as Heart Disease DataSet and can be determined indoors the UC tool reading repository. The UC tool analyzing repository is composed of a massive and numerous quantity of datasets which embody datasets from several area names. These information are significantly finished through tool reading network from novices to experts to recognize records empirically. Various academic papers and researches had been completed the use of this repository.

Table 1. Attributes of the Heart disease dataset

Attribute	Representation	Information Attribute	Description
Age	Age	Integer	Age in years (29 to 77)
Sex	Sex	Integer	Gender instance (0 = Female, 1 = Male)
ChestPainType	Cp	Integer	Chest pain type (1: typical angina, 2: atypical angina, 3: non- anginal pain, 4: asymptomatic)
RestBloodPressure	Trestbps	Integer	Resting blood pressure in mm Hg[94, 200]
SerumCholestoral	Chol	Integer	Serum cholesterol in mg/dl[126, 564]
FastingBloodSugar	Fbs	Integer	Fasting blood sugar > 120 mg/dl (0 = False, 1= True)
ResElectrocardiographic	Restecg	Integer	Resting ECG results (0: normal, 1: ST-T wave abnormality, 2: LV hypertrophy)
MaxHeartRate	Thalach	Integer	Maximum heart rate achieved[71, 202]
ExerciseInduced	Exang	Integer	Exercise induced angina (0: No, 1: Yes)
Oldpeak	Oldpeak	Real	ST depression induced by exercise relative to rest[0.0, 62.0]
Slope	Slope	Integer	Slope of the peak exercise ST segment (1: up-sloping, 2: flat, 3: down-sloping)
MajorVessels	Ca	Integer	Number of major vessels coloured by fluoroscopy (values 0 – 3)
Thal	Thal	Integer	Defect types: value 3: normal, 6: fixed defect, 7: irreversible defect
Class	Class	Integer	Diagnosis of heart disease (1: Unhealthy, 2: Healthy)

VI. RESULT AND ANALYSIS

The reason of this studies is to research the general ordinary usual performance of numerous type algorithms and in doing so find out the maximum accurate set of regulations for predicting whether or no longer a affected man or woman may additionally boom and coronary heart attack or no longer. This studies have emerge as completed the usage of strategies Support Vector Machine, K-Nearest Neighbor at the UCI dataset. Dataset have end up cut up into education and take a look at records and models were informed.
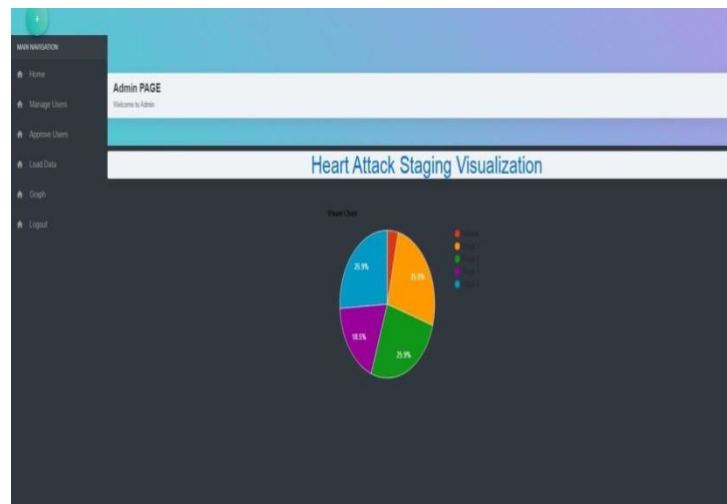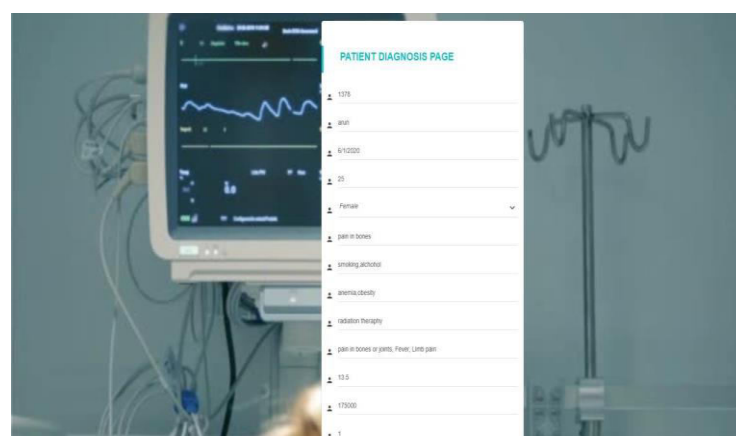
Fig 2: Home Page

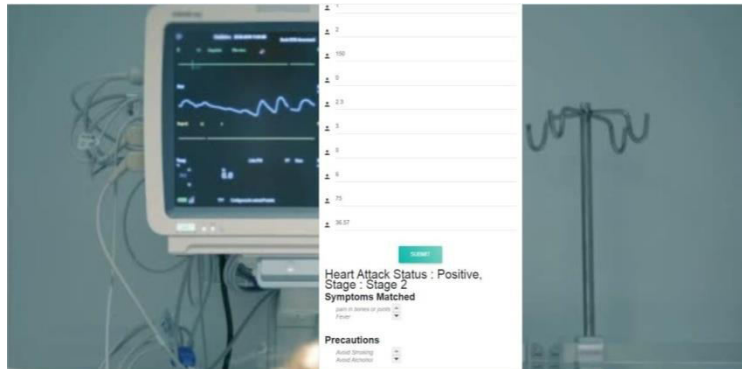
Fig 3:Graph visualization

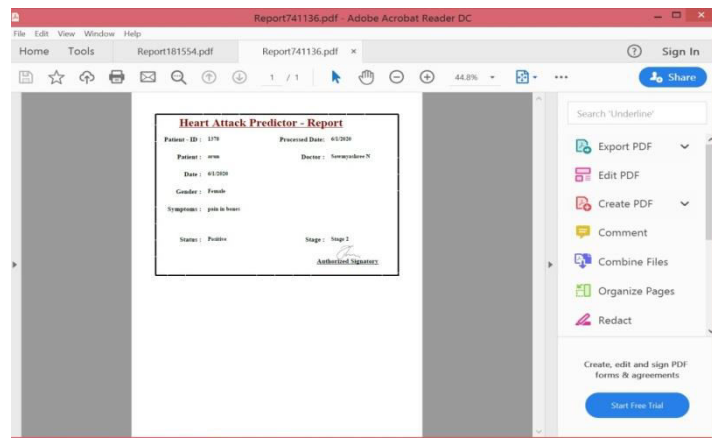Fig 4:Collecting Patient data

Fig 5:Result

Fig 6 :Report

VII. CONCLUSION AND FUTURE WORK

Heart Disease is a deadly disease through its nature. This infection makes a existence threatening complexities which encompass coronary heart assault and shortage of life. The importance of Machine Learning inner the Medical Domain is found out and steps are taken to use relevant techniques inside the Disease Prediction. We are imposing a device so you can help to are searching in advance to coronary heart infection counting on the sufferers medical records associated with the detail related to coronary heart infection. By the use of medical dataset of the patients such as age, sex, blood strain, overweight and blood sugar and thru appling KNN and SVM classifier we are able to assume that the sufferers getting a coronary heart contamination or now not . In addition kind accuracy, sensitivity and specificity of the SVM were decided to be excessive consequently making it a advanced possibility for the assessment. We are additionally doing evaluation on the records from which we've got become at which age it regularly rise up or which region receives inspired via the usage of that contamination. So precaution can be taken to keep away from the shortage of lifestyles because of the coronary heart infection..

REFERENCES

1. Senthilkumar Mohan, Chandrasegar Thirumalai, Gautam Srivastava ―Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques‖, Digital Object Identifier 10.1109/ACCESS.2019.2923707, IEEE Access, VOLUME 7, 2019 S.P. Bingulac, ―On the Compatibility of Adaptive Controllers,‖ Proc. Fourth Ann. Allerton Conf. Circuits and Systems Theory, pp. 8-16, 1994. (Conference proceedings)

2. Sonam Nikhar, A.M. Karandikar" Prediction of Heart Disease Using Machine Learning Algorithms" International Journal of Advanced Engineering, Management and Science (IJAEMS) Infogain Publication,[Vol-2, Issue-6, June2016].I.S. Jacobs and C.P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G.T. Rado and H.Suhl, Eds. New York: Academic, 1963, pp. 271-350.|

3. Aditi Gavhane, Gouthami Kokkula , Isha Pandya, Prof. Kailas Devadkar (PhD)," Prediction of Heart Disease Using Machine Learning", Proceedings of the 2nd International conference on Electronics, Communication and Aerospace Technology (ICECA 2018). IEEE Conference Record # 42487; IEEE Xplore ISBN:978-1- 5386-0965-1

4. Abhay Kishore1, Ajay Kumar2, Karan Singh3, Maninder Punia4, Yogita Hambir5," Heart Attack Prediction Using Deep Learning", International Research Journal of Engineering and Technology (IRJET), Volume: 05 Issue: 04 | Apr2018.

5. A.Lakshmanarao, Y.Swathi, P.Sri Sai Sundareswar," Machine Learning Techniques For Heart Disease Prediction", International Journal Of Scientific & Technology Research Volume 8, Issue 11, November 2019.

6. Mr.Santhana Krishnan.J, Dr.Geetha.S," Prediction of Heart Disease Using Machine Learning Algorithms" ,2019 1st International Conference on Innovations in Information and Communication Technology (ICIICT), doi:10.1109/ICIICT1.2019.8741465.

7. Avinash Golande, Pavan Kumar T," Heart Disease Prediction Using Effective Machine Learning Techniques", International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8, Issue-1S4, June 2019.

8. V.V. Ramalingam, Ayantan Dandapath, M Karthik Raja," Heart disease prediction using machine learning techniques: a survey", International Journal of Engineering & Technology, 7 (2.8) (2018) 684-687.

9. V. Manikantan and S. Latha, "Predicting the analysis of heart disease symptoms using medicinal data mining methods", International Journal of Advanced Computer Theory and Engineering, vol. 2, pp.46-51, 2013.

10. M. S. Amin, Y. K. Chiam, K. D. Varathan,''Identification of significant features and data mining techniques in predicting heart disease,'' Telematics Inform., vol. 36, pp. 82– 93, Mar.2019

INTERNATIONAL JOURNAL
OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

📱 9940 572 462 ⊙ 6381 907 438 ✉ ijircce@gmail.com