

HEARTDISEASEANALYSIS

R.SriRagulR

UGFinalYearDepartmentofComputerScience
SriKrishnaArts&ScienceCollegeCoimbatore

Abstract—This document is an aspect of applying system in the diagnosis and therapy for heart diseases and also to improve patients safety using a simple data mining technique.

Keywords—medical errors, Data mining, heart disease.

I. INTRODUCTION

There are various reasons for the occurrence of Heart Diseases, which can be frequently investigated through the Attribute Set related to different test results of Patients. The different sources of medical data are Medical Analysis, Diagnostic Centres, past Case Sheets, Doctor Prescriptions. Heart diseases can be predicted through the analysis made on some attributes like age, sex, chest pain type, blood pressure, cholesterol, fasting blood sugar, Maximum heart rate achieved.

K-means clustering algorithm is used for validating the accuracy of medicinal data. These algorithms can be used to optimize the data storage for practical and legal purposes.

I. METHOD OF USE

A. Existing:

Numerous works for heart disease prediction is achieved by artificial neural network and data mining techniques. A requires hardwares like radiator, scanners and manpower. There is no existing software that works alone without a doctor and hardwares.

Accuracy Issues: A manual system alone does not ensure accuracy, and the warehouse data is only as good as the data entry that created it. By manually because the amount is very large so some data mining and machine learning techniques can be used.

B. Proposal:

- The User can search for doctor's help at any point of time.
- User can talk about their Heart Disease and get instant diagnosis.
- Doctors get more clients online.
- Very useful in case of emergency.

I. ALGORITHM USED

K-Means is a good option (fast, robust and easy to understand) for local searchability. It handles missing values and investigates each possibility. If an attribute has more than 5% missing values then the record should not be deleted and it is advisable to impute. Meanwhile, it produces different results at different initial partitions. K-means clustering algorithm is used for validating the accuracy of medicinal data. These algorithms can be used to optimize the data storage for practical and legal purposes.

	Height	Weight
Person 1	H1	W1
Person 2	H2	W2
.	.	.
.	.	.
.	.	.
.	.	.
.	.	.
Person 50	H50	W50

Fig. Example of person height and weight analysis

Data with Multiple Features is a type of clustering algorithm method which is used to analyse data remaining coresponding to the input data given.

A. Patient observation

Heart diseases can be predicted through the analysis made on some attributes like age, sex, chest pain type, blood pressure, cholesterol, fasting blood sugar, Maximum heart rate achieved.

- └ Patients register their details like name, email, phone, username and password.
- └ Patient logs into the system with registered username and password.
- └ Patient enter their disease history attribute details like age, gender, smoking habit, drinking habit, fast food intake habit, whether they are overweight, whether they are mentally stressed,, whether they have pin in chest, blood pressure detail, blood sugar details.
- └ Patient disease history attribute values are analysed with k-means algorithm and finally Disease prediction is show to user.

B. Attributes Required

Basic attributes required to predict a heart disease are:

- └ Diabetes
 - └ Sugar level of the patient
 - └ Cholesterol level
 - └ Age
 - └ 60% of disease affected patients will be above age of 40
 - └ Sex
 - └ Female patients affect more in heart disease
 - └ Habbits
 - └ Risk factors

└ Does affected of any kind of diseases Number Consequently if they don't know about the above mentions factors, we can analyse their medical reports on particular checkups (like sugar, cholesterol, feminism checkups.)

Attributes like name, age, sex, id is mandatory and analysis is started only after these logins. C. Some common non aspects

Patients who does not know their

- └ Age
- └ Medical status
- └ On the risk stages of disease

└ Notvisitedadoctorforawhile



Fig2:Howhealthyareyou-statement.

IV. REMEDYOFNONASPECTS

Ageofapatient(whodoesn'tknows)canbepartialypredictedbytheirappearanceandphysic.It doesn'tmatersthatnotvisitingadoctorbutmatersontheymustknowabouttheirphysicand medicalstatus.Anormalmedicalcheckupisenough[1]Height/weight [2]Bloodpressure[3] Cholesterollevel[4]Bloodsugartest[5]Throatcheck[6]EarandEyecheck.“ElectroCardiogram”for thoseatahigherriskofheartdisease.



Fig3:Doctorcheckingone'spressure

A. Doctor'spointofview

Bioscienceisthemajortechniqueused bydoctors.Clinicaltestreportsarepredicted by effectivenessofhugeandcomplexclinicaldatainordertodiagnosedisease.Theseinformationsare minedundereffectivemedicalassistance.

SomepowerfulknowledgeonheartdiseasepredictionusinganapproachofANNwithLVQ achievedaccuracyof80-85%andspecificityofabout70%.

(ANN-Artificialneuralnetwork,LVQ-LearningVectorQuantization)

- └ Doctorregisterstheirdetailslikename,email,phone,specialistin,usernameandpassword.
- └ Doctorlogsintothesystemwithusernameandpassword
- └ Doctorcanviewpatientdiseasehistoryattributedetailslikeage,gender,smokinghabit,drinking habit,fastfoodintakehabit,whethertheyareoverweight,whethertheyarementallystressed, whethertheyhavepininchest,bloodpressuredetail,bloodsugardetails.
- └ Doctorcanviewoverallclusteredreport.
 - └ Theclusterreportwilsaylargepercentageofheartatackdiseaseisduetosmokinghabit ordinkinghabitorfastfoodintakehabitorstressfactor,etc.
 - └ Thereportsayswhetherthepatientsthehighriskofheartdisease,diabetesorcancer.
 - └ Thisdataandanalysiscanbeextractedasreportandhandedovertothepatients.

B. Tables

a) PatientLogin: It contains patient details

Atributes	Fields	Action
1. User name	Username can be of their choice	created
2. Password	Password must be created	created

Table 1: Example of a Patient Login table

Fig4: Patient login

Patient must create their own login by using their username and password.

Username may be of his/her choice. Password will be saved and can be noted by admin for medical purpose and data mining.

Although the details will be kept secured and only can be viewed by permitted doctors and admin handlers. Details of the patients must be mandatory and true on behalf of the patients.

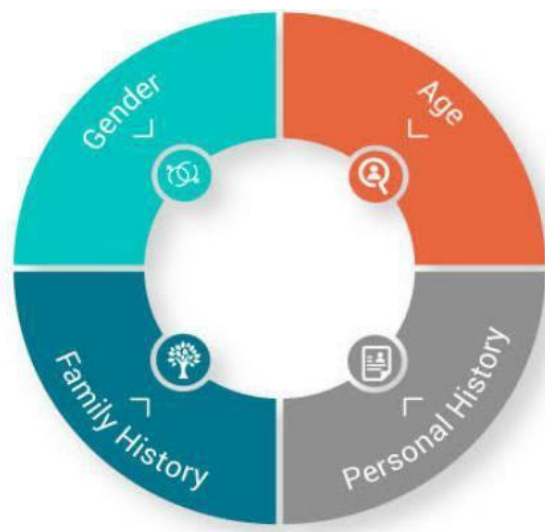


Fig5: Important attributes to predict a disease

b) PatientDetails:itcontainsPatientData

Atributes	Fields
1.Age and gender	Ageofthepatient
2.Habbit	Smoking/Drinking
3.Obesity	Foodhabbits/weight
4.Sugarlevel	Sugarlevelinblood ?/180ml
5.Cholesterol	Cholesterollevel?/120ml inblood
6.Intakes	Medicationsforanykindof disease

Table2:Exampleofpatientdetailstable

SamefortheDoctorloginpageandhe/shecantrackthepatient'sstageofatributestopredicttheir presence.It doesn't mean that the doctor predicts the disease. Disease is mined using the data fields given by patients. Doctors visit the data to ensure the result.

Table 2.— Male Cohort Mean Daily Dietary Intake: Framingham 16-Year Analysis		
Variable	Mean (± SD) Age	
	45-55 y (n = 420)	56-65 y (n = 393)
Energy intake, kJ	11 195 ± 3155	10 671 ± 2833*
Total fat, g	118.4 ± 41.9	109.3 ± 39.2*
% of energy intake	39.7 ± 8.4	38.3 ± 8.9†
Saturated fatty acid, g	45.3 ± 18.2	42.3 ± 17.6†
% of energy intake	15.2 ± 4.4	14.8 ± 4.5
Polyunsaturated fatty acid, g	16.5 ± 8.7	15.0 ± 7.5*
% of energy intake	5.5 ± 2.5	5.4 ± 2.4
Monounsaturated fatty acid, g	48.5 ± 18.3	44.3 ± 17.2‡
% of energy intake	16.2 ± 4.0	15.5 ± 4.0*
Polyunsaturated-saturated ratio	0.4 ± 0.2	0.4 ± 0.2
Dietary cholesterol, mg	529.6 ± 269.1	531.8 ± 291.3

Fig6:Exampleofapatient'smedicationanddiabetesandcholesterolreport

Theabovetableisanexampleofmedicationundertakenbyapatientandhis/herlevelofcholesterol anddiabetes.

V. HeartDiseaseFacts

- 1)Heartdiseaseistheleadingcauseofdeathforbothmenandwomen.Morethanhalfofthe deathsdue toheartdiseasein2009wereinmen.
- 2)About610,000Americansdiefromheartdiseaseeachyear—that's1inevery4deaths.
- 3)Coronaryheartdiseaseisthemostcommontypeofheartdisease,kilingmorethan370,000 peopleannually.
- 4)IntheUnitedStates,someonehasaheartatackevery43seconds.Eachminute,someonein theUnitedStatesdiesfromaheartdisease-relatedevent.
- 5)Heartdiseaseistheleadingcauseofdeathforpeopleofmostracial/ethnicgroupsinthe UnitedStates,includingAfricanAmericans,Hispanics,andwhites.
- 6)ForAsian AmericansorPacificIslanders andAmerican IndiansorAlaskaNatives,heart diseaseissecondonlytocancer.
- 7)CoronaryheartdiseasealonecoststheUnitedStates\$108.9bilioneachyear.Thistotal includesthecostofhealthcareservices,medications,andlostproductivity

VI. CONCLUSION

Thispaperpresentedonbasisofheartdiseasepredictionforhealthcaresystemusingdatamining techniques.Numerousworksforheartdiseasepredictionisachievedbyartificialneuralnetworkand dataminingtechniques.Infuture,weareplanningtoproposeaneffectivediseasepredictionsystem topredictheheartdiseasewithbeteraccuracyusingdifferentdataminingclassificationtechniques suchasNaïveBayesandDecisionTree.

References:

- [1] www.cdc.gov/heartdisease/facts.htm
- [2] Variouswebsitesonsearchofheartdiseases
- [3] Dr.BimalChhajer,"REVERSALOFHEARTDISEASE"
- [4] CaldwellB.Esselstyen"PREVENTANDREVERSE THEHEARTDISEASE"Book...
- [5] Junjiwu"BOOKOFKMEANSCLUSTERING".
- [6] DataminingtechniquesunderguidanceofstaffsofComputerSciencedepartment,Skasc.
- [7] ChaitraliS.Danagre,SulabhaS.Apte,Ph.D,ImprovedStudyofHeartDiseasePredictionSystemusing DataminingClassificationTechniques,IJCA,June2012.
- [8] SelappanPalaniappan,RafiahAwang,"IntelligentHeartDiseasePredictionSystemusingdata MiningTechniques",IJCSNS,Vol.8No.8,August2008
- [9] FeixiangHuang,ShengyongWang,andChien-ChungChan,"PredictingDiseaseByUsingData MiningBasedonHealthcareInformationSystem",IEEEComputing,2012.
- [10] Blake,C.L.,Mertz,C.J.: "UCIMachineLearningDataset",<http://mllearn.ics.uci.edu/databases/heartdisease/>,2004.