Disease Prediction And Security In Healthcare Using Data Mining

Mrs.B.MeenaPreethi[1], Ms.Sasirevathi.N[2], Mr.S.Thinesh[3]
[1, 2, 3] Dept of SS
[1, 2, 3] Sri Krishna Arts and Science college

Abstract- *Data mining is a subfield of Computer Science that uses already existing data in different databases to transform it into new researches and results. The major objective is to use the data mining techniques is to use data mining techniques effectively. This can be done by taking the data provided as input by the users and providing them with the suspected disease and also with the remedies of the obtained disease. Medical researchers hope to utilize clinical data to discover knowledge that is lying completely in individual patient health records. Data used for this will be extracted from data set in unstructured format and by using open source tools which will be later converted to structured format. The Clustering of data is done according to requirements and by applying association rule disease is predicted and remedies are given.The patients records are compared with the data sets which is uploaded beforehand to generate result. The data mining technique is used here which correlates the data from the patient records with the data set to predict the possibility of future occurrence of disease. This prediction will help in maintaining good health by avoiding risks in future.*

Keywords- Healthcare, Disease prediction, security, Symptoms, Data mining, Clustering, Association.

I. INTRODUCTION

Data mining is the method used for finding unknown values from huge amount of data. Since the patients population increases the medical databases also keeps on increasing day by day. The transactions and investigation of those medical information is tough while not the pc based mostly analysis system. The computer primarily based analysis system indicates the mechanized diagnosis system. This mechanized designation system support the health care provider to form smart call in treatment and malady. Data mining is the huge platform for the doctors to handle the huge amount of patient's data set in many ways such as make sense of complex diagnostic tests, interpreting previous results, and mixing the dissimilar information along.

In today's computerized world considering automatic and dynamic requirements healthcare system should be more efficient by predicting the disease and providing appropriate medications through user friendly applications. This study aims principally for the maintenance of data electronically and securely. It also aims in providing easy access of data and searching remedies for the diseases predicted.Health care is changing and so are the tools used to coordinate better care for patient's privacy and security. To improve this objective Electronic Health Records are used in which records are maintained electronically and send via internet. With this comes the opportunity for patients to receive improved and coordinated care from providers for easier access to their health related information. It's a useful way to make it easier for every human to be better informed and involved in their health care. However, for many reasons Electronic Health Records also come with questions and concerns about the privacy and security of our health information.

II. LITERATURE SURVEY

In the paper "A study on data mining prediction techniques in healthcare sector" the fields which discussed are, Knowledge Discovery Process (KDD) is the process of changing the low-level data into high-level knowledge. Hence, KDD refers to the nontrivial removal of implicit, previously unknown and potentially useful information from data in databases. The Knowledge Discovery in Databases process comprise of a few steps leading from information collections to some sort of new data. The iterative process consists of the following steps: Data cleaning, Data integration, Data selection, Data transformation, Data mining, Pattern evaluation, Knowledge. Healthcare data mining prediction based on data mining techniques are as follows: Neural network, Bayesian Classifiers, Decision tree, Support Vector Machine. The paper states the comparative study of different healthcare predictions, Study of data mining techniques and tools for prediction of heart disease, various cancers, diabetes, eye disease and dermatological conditions. Data mining based mostly prediction system reduces the human effects and value effective one. Few limitations are that if attributes are not related then Decision trees prediction is less accurate and ANN is computationally intensive to train also it does not leads to specific conclusion.

The paper "Predicting Disease by Using Data Mining Based on Healthcare Information System" applies the data mining process to predict hypertension from patient medical

records with eight other diseases. The data was extracted from a real world Healthcare Information System database containing medical records. Under sampling technique has been applied to come up with coaching knowledge sets, and data mining tool Weka has been used to generate the Naive Bayesian and J-48 classifiers created to improve the prediction performance, and rough set tools were used to reduce the ensemble based on the idea of second order approximation. Experimental results showed a bit improvement of the ensemble approach over pure Naive Bayesian and J-48 in accuracy, sensitivity and F-measure. Initially they had a classification and then ensemble the classifiers and then the reduction of Ensemble Classifiers is used. But the decision trees generated by J-48 is sometimes lacking in the balancing so the overall improvement of using ensemble approach is less.

The paper "An approach to devise an Interactive software solution for smart health prediction using data mining" aims in developing a computerized system to check and maintain your health by knowing the symptoms. It has a symptom checker module which actually defines our body structure and gives us liability to select the affected area and checkout the symptoms. Technologies implemented in this paper are: The front end is designed with help of HTML, JavaScript and CSS. The back finish is meant victimization MySQL that is employed to style the databases. This paper also contains the information of testing like Alpha testing which is done at server side or we can say at the developer's end, this is an actual testing done with potential users or as an independent testing process at server end. And Beta testing is done after performing alpha testing, versions of a system or software known as beta versions are given to a specific audience outside the programming team. Only the limitation of this paper is it suggests only the award winning doctors and not the nearby doctors to the patient.

III. EXISTING SYSTEM

Whenever the computerized records are used, the organization using this system have to be careful about protecting the data from unauthorized access. The adoption to electronic record is that extreme diligence is required to shield sensitive data from cyber criminals and malicious hackers. The records which contain the details of the user are mishandled. This data threat leads to a huge break between the patient and the respective medical organization. Some records like X-ray, ECG and radio therapy records are handled manually by many medical organization which leads to a great disaster.

IV. PROPOSED SYSTEM

Proposed system which generates Electronic health records incorporate a vast amount of patient information and diagnostic data, most of which is considered protected health information. With the advancement of technology, the emergence of advanced cyber threats has escalated, which hinders the privacy and security of health information systems such as Electronic Health Records. To avoid such risks security methods are included through algorithms. Prediction and solution to various health risk are made possible and security is improved to maintain the trust and relationship between the patients and organization is improved. Proposed system which generates Electronic health records incorporate a vast amount of patient information and diagnostic data, most of which is considered protected health information. With the advancement of technology, the emergence of advanced cyber threats has escalated, which hinders the privacy and security of health information systems such as Electronic Health Records. To avoid such risks security methods are included through algorithms.

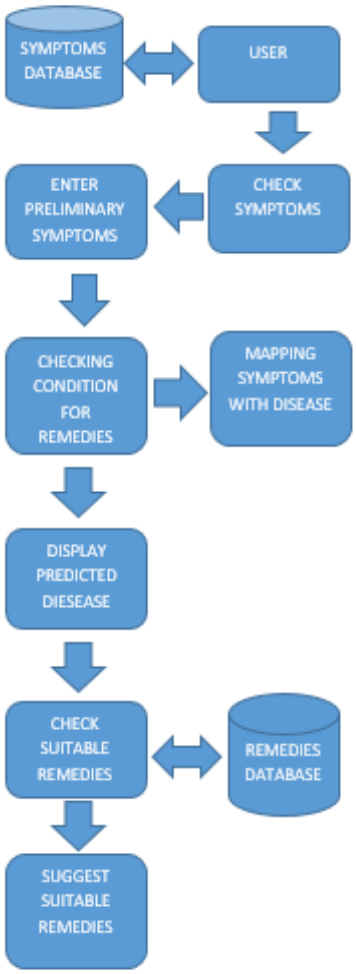

Fig 1: Diagram For Proposed System

V. ALGORITHMS USED

Data Mining concept generally has many algorithms. The two essential algorithms used here are:

a) Random Number Generation Algorithm

The Random Number Generation (RNG) is important for using cryptography. A Random Number Generator is an important algorithm designed to generate a sequence of numbers and alphabets that does not have any pattern, therefore it appears unique and random. In this project this algorithm is used for security purpose by creating random sequenced passwords.

b) Genetic Algorithm

Genetic Algorithms (GAs) are search based algorithms based on the concepts of natural selection. GAs area unit a set of a far larger branch of computation called biological process Computation. In GAs, user have a pool or a population of possible solutions to the given problem. These solutions then bear recombination and mutation like in natural genetic science, manufacturing new kids, and also the method is perennial over numerous generations. In this way user keep evolving better solutions till user reach a stopping criterion. Genetic Algorithms area unit sufficiently randomized in nature, however they perform far better than random native search within which we have a tendency to simply strive numerous random solutions, keeping track of the most effective so far and also they exploit historical information as well.

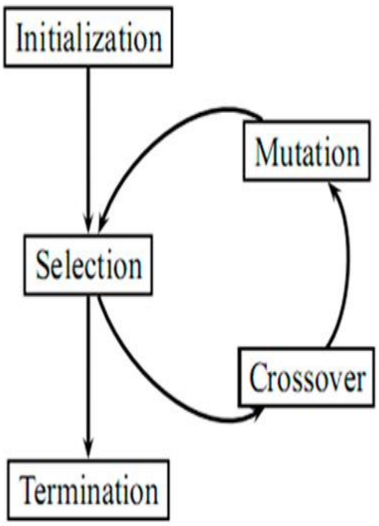

Fig 2: Genetic Algorithm

VI. OBJECTIVE

Health care is changing and so are the tools used to coordinate better care for patient's privacy and security. To improve this objective Electronic Health Records are used in which records are maintained electronically and send via internet. With this comes the opportunity for patients to receive improved and coordinated care from providers for easier access to their health related information. It's a useful way to make it easier for every human to be better informed and involved in their health care. However, for many reasons Electronic Health Records also come with questions and concerns about the privacy and security of our health information.

VII. CONCLUSION

This paper contains patient-centered records that make information available instantly and securely to authorized users. While these electronic records does contain the medical and treatment histories of patients, this system is built to go beyond digitalizing report to prediction of disease which might occur in future and its solution strategy based on users record. One of the key features of this system is that health information of the patient can be created and managed by authorized providers in a digital format and capable of being shared with other providers across more than one health care organization by hiding personal information of the patients and providing solution and remedies to their problems.

REFERENCES

[1] Aditya Tomar, "Disease Prediction System using data mining techniques",in International Journal of Advanced Research in computer and Communication Engineering,ISO 3297, July 2016.

[2] Dr.B.Srinivasan, K.Pavya, "A study on data mining prediction techniques in healthcare sector", in International Research Journal of Engineering and Technology(IRJET), March-2016.

[3] Megha Rathi, Vikas Pareek, "An integrated hybrid data mining approach for healthcare" , in IRACST International Journal of Computer Science and Information Technology Security (IJCSITS), ISSN: 2249-9555 , Vol.6, No.6,Nov-Dec 2016.

[4] Feixiang Huang, Shengyong Wang, and Chien-Chung Chan, "Predicting Disease By Using Data Mining Based on Healthcare Information System" , in IEEE 2012.

[5] M.A. Nishara Banu,B Gomathy, "An approach to devise an Interactive software solution for smart health prediction using data mining, in International Journal of

Technical Research and Applications , eISSN, Nov-Dec 2013.

[6] Akshay Raul, Atharva Patil, Prem Raheja, Rupali Sawant, "Knowledge Discovery, Analysis And Prediction in Healthcare using Data Mining And Analytics", in IEEE,2nd International Conference on Next Generation Computing Technologies,2016.

[7] M.A. Nishara Banu, B.Gomathy, "Disease Forecasting System Using Data Mining Methods" ,in IEEE,2014 International Conference on Intelligent Computing Applications.

[8] Sellappan Palaniappan,RafiahAwang, "Intelligent Heart Disease Prediction System Using Data Mining Techniques",in IEEE 2008.

[9] Theresa Princy, R Research Scholar, J.Thomas,"Human Heart Disease Prediction System using Data Mining Techniques",in 2016 International Conference on Circuit, Power and Computing Technologies [ICCPCT], @2016 IEEE.

[10] RanganathaS.1 ,Pooja Raj H.R.2 , Anusha C.3 , Vinay S.K,, "Medical Data Mining And Analysis For Heart Disease Dataset Using Classification Techniques",in IEEE 2009

[11] Ms. Chaitrali S. Dangare, Dr. Mrs. Sulabha S. Apte, "A data mining approach for prediction of heart disease using neural networks, international journal of computer engineering and technology",2012

[12] M.A.Nishara Banu and B.Gomathy," Disease Forecasting System Using Data Mining Methods", 2014

[13] Aqueel Ahmed, Shaikh Abdul Hannan, "Data Mining Techniques to Find Out Heart Diseases", International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-1, Issue-4, September 2012.

[14] Ms. Ishtake S.H, Prof. Sanap S.A., "Intelligent Heart Disease Prediction System Using Data Mining Techniques", International J. of Healthcare & Biomedical Research, 2013

[15] Chitra R and Seenivasagam V, "REVIEW OF HEART DISEASE PREDICTION SYSTEM USING DATA MINING AND HYBRID INTELLIGENT TECHNIQUES", ISSN: 2229-6956(ONLINE) ICTACT JOURNAL ON SOFT COMPUTING, JULY 2013, VOLUME: 03, ISSUE: 04, 2013

[16] Nidhi Bhatla and Kiran Jyoti, "An Analysis of Heart Disease Prediction using Different Data Mining Techniques", International Journal of Engineering Research & Technology (IJERT), ISSN: 2278-0181,Vol. 1 Issue 8, October – 2012