International Journal of Current Engineering and Technology
E-ISSN 2277 – 4106, P-ISSN 2347 – 5161

Available at http://inpressco.com/category/ijcet

Research Article

An Efficient Method of Health Risk Prediction System

Miss. Prachiti Khilari and Prof. Vikas Mapari

Department of Computer Engineering. D Y Patil clg of engineering, Ambi Prachiti

Received 10 Nov 2020, Accepted 10 Dec 2020, Available online 01 Feb 2021, **Special Issue-8 (Feb 2021)**

Abstract

Data evaluation plays a noteworthy function in managing a massive quantity of facts within the healthcare. The previous scientific researches based on handle and assimilate a big quantity of hospital records rather than prediction. Due to a significant amount of information boom inside the biomedical and healthcare area the correct analysis of scientific statistics turns into propitious for in advance detection of disorder and patient care. With the event of society and economy, people pay more attention to their own health. The demand of more personalized health service is gradually rising. However, due to the shortage of experienced doctors and physicians, most healthcare organizations cannot meet the medical demand of public. Due to that public want the Medical Treatment online with accuracy. Now a day's, public has no time to urge the doctor physically, then search the web hospital near about the present location. With the widespread use of hospital data system , there's huge amount of generated data which may be wont to improve healthcare service. Thus, more and more data processing applications are developed to supply people more customized healthcare service. In EHRs, users may be a health data owner (i.e., patients) or a requester (i.e., doctors or pharmacists), servers, in turn could be local or cloud servers that store, process and analyze the gathered health data. Networks, on the other hand, act as the bridge connecting between patients and the medical staff to support the transmitting and sharing of data. So, it is necessary to ensure patients feel fully confident to use the system and have their own privacy control over it. To this end, in this paper, we conduct an in-depth survey study to analyze the healthcare system's security and privacy threats.In the proposed paper we use different algorithm to extend the safety of sensitive information of hospital management includes doctors, patients then on.We also propose a novel security model that captures the scenario of data interoperability and supports the security fundamental of EHR along with the capability of providing finegrained access control.

Keywords: *Disease diagnosing, Security, Data Mining, Biomedical and Healthcare, Hospital organization.*

Introduction

With the rapid economic improvement and the development of the national income, many fitness-care subjects have lately been turning into more essential. The reason is that people need to pay more attention to care their personal health states. With rapid improvement of statistics mining technology, the related strategies had been implemented to diverse realistic programs, such as scientific analysis, mobile information analysis, multimedia statistics analysis, and so forth. In the sphere of clinical applications, the developed clinical systems have been evolved the clever scientific systems, together with the wise auto-diagnosing body structure signal.

Data mining is the core step, which consequences within the discovery of hidden however beneficial knowledge from massive databases. Data mining technology affords a user-oriented approach to novel and hidden patterns in the statistics. The found know-how can be used by the healthcare administrators to improve the first-rate of service. The discovered knowledge can also be utilized by the medical practitioners to reduce the wide variety of damaging drug effect, to indicate less expensive therapeutically equivalent alternatives. Anticipating patient's future behavior on the given records is one in all the vital programs of records mining techniques that may be used in fitness care management. A major undertaking dealing with healthcare organizations (hospitals, clinical centers) is the supply of fine offerings at low-priced costs. Quality service implies diagnosing sufferers efficaciously and administering treatments which are effective. Poor clinical choices can prompt tragic outcomes which are hence unsatisfactory.

Emergency clinics should likewise limit the expense of clinical tests. They can accomplish these outcomes by utilizing proper PC based data as well as choice emotionally supportive networks. Social insurance information is enormous. It incorporates tolerant driven information, asset the executives information and changed information. Social insurance associations must have capacity to investigate information. Treatment records of a large number of patients can be

put away and mechanized and information mining strategies may help in noting a few significant and basic inquiries identified with medicinal services.

In this specific situation, electronic medicinal services frameworks (EHRs) representative such guidelines and along these lines were classified as security basic frameworks. These frameworks are separated in one significant angle to different frameworks, the adjusting among privacy and accessibility. The pressure between these objectives is clear: while all the patient's information ought to be accessible to be shared and checked to convey proficient human services administrations; for security reasons, some portion of the information might be viewed as private and must not be available. In EHRs, clients might be a health related information proprietor (i.e., patients) or a requester (i.e., specialists or drug specialists), servers, thusly could be neighborhood or cloud servers that store, process and investigate the assembled health information. Systems, then again, go about as the scaffold interfacing among patients and the therapeutic staff to help the transmitting and sharing of information.

In this way, it is important to guarantee patients feel completely certain to utilize the framework and have their very own protection authority over it. To this end, in this paper, we lead an inside and out study concentrate to break down the medicinal services framework's security and protection dangers. At that point, we propose a novel security model that catches the situation of information interoperability and supports the security principal of EHR alongside the capacity of giving fine-grained get to control.

B. Motivation

The Motivation behind that is to manage a tremendous amount of different ailment data and on that the hazard forecast of ailment can be analyzed. With the extraordinary utilization of emergency clinic insights framework, there is a lot of created measurements which can be utilized to upgrade wellness care administrations, in this manner creating actualities mining applications to offer individuals increasingly specially crafted social insurance administration.

C. Objectives

- To implement Naive Bayes Classifier that classifies the disease as per the input of the user.
- To develop web interface platform for the forecast of the sickness
- To provide a web platform to predict the occurrences of ailment on the idea of various signs and symptoms. The addict can choose different side effects and can discover the ailments.
- To construct a completely useful framework so as to accomplish effectiveness in earlier health treatment and online consultation framework.

Review of Literature

In this paper, the author has provided an wise and powerful coronary heart attack prediction methods the usage of records mining. Firstly, it provided an efficient method for the extraction of giant styles from the coronary heart disorder information warehouses for the efficient prediction of coronary heart assault Based at the calculated significant weight age, the frequent patterns having value greater than a predefined threshold had been selected for the precious prediction of coronary heart attack. In this paper the drawbacks are for predicting coronary heart attack appreciably 15 attributes are indexed. Besides the 15 indexed in clinical literature we can also contain other records mining techniques, e.G., Time Series, Clustering and Association Rules. [1]

In this paper, creator presented a middleware solution method to support data and community safety over eHealthcare system sing medical sensor networks. It has been shown that a masquerade assault can be launched to the device and patients 'records are in danger. We proposed this middleware to counter this form of assault wherein a person and all devices into the healthcare network are mutual authenticated. Finally a performance analysis has been performed in regards to masquerade assault and the end result well-known shows the green of the proposed solution. [2]

This work has analyzed prediction structures for Heart disorder using extra range of input attributes. The paintings makes use of clinical phrases inclusive of sex, blood stress, ldl cholesterol like thirteen attributes to expect the chance of patient getting a Heart disease. Until now, 13 attributes are used for prediction. This studies work introduced two greater attributes i.E. obesity and smoking. The statistics mining classification algorithms, specifically Decision Trees, Naive Bayes, and Neural Networks are analyzed on Heart sickness database. [3]

Medical Diagnosis Systems play important function in scientific practice and are used by medical practitioners for prognosis and remedy. In this work, a clinical prognosis device is described for predicting the hazard of cardiovascular sickness. This system is built by means of combining the relative blessings of genetic approach and neural community. Multilayered feed forward neural networks are particularly adapted to complicated class problems. The weights of the neural network are determined using genetic approach because it reveals acceptably excellent set of weights in much less number of iterations. [4]

In this paper, the author designed cloud-primarily based fitness –Cps device in which manages the large quantity of biomedical statistics. This device performed various operations on cloud-like data evaluation, tracking and prediction of records. With the help of this system, a person receives more statistics about the way to take care of and manage the big quantity of biomedical information in the cloud. Also, the various

services associated with healthcare recognise by this machine. [5]

In this paper, the author proposes an efficient decision tree algorithm technique for coronary heart disease prediction. To achieve correct and cost effective treatment laptop-primarily based systems can be evolved to make good selection. . Data mining is a effective new technology for the extraction of hidden predictive and actionable facts from huge databases, the principle objective of this venture is to increase a prototype which can determine and extract unknown knowledge (patterns and relations) associated with heart ailment from a past coronary heart disease database record. [6]

In this paper, the authors used clinical information mining strategies like association rule mining, type, clustering I to investigate the different sorts of heart primarily based problems. Decision tree is made to demonstrate every possible final results of a decision. Different regulations are made to get the fine outcome. In this studies age, sex, smoking, overweight, alcohol intake, blood sugar, hear charge, blood stress are the parameters used for making the decisions. [7]

In this paper, the author have proposed for predicting the coronary heart disease the usage of association rule mining technique, they have generated a huge quantity of policies when affiliation regulations are implemented to dataset .Frequent Item set Mining is used to discover all frequent item sets. Association rule mining strategies like Apriority and FP increase are used frequently. [8]

In this paper, the author has evolved the laptop aided coronary heart disorder prediction system that helps the doctor as a tool for coronary heart ailment diagnosis. From the evaluation it's far concluded that neural network with offline training is right for disease prediction in early level and excellent performance can be acquired by using pre-processed and normalized dataset. [9]

This work focuses a new approach for applying affiliation guidelines inside the Medical Domain to find out Heart Disease Prediction. The health care enterprise collects big amount of health care information which, unfortunately aren't mined to find out hidden information for effective choice making. Discovery of hidden styles and relationships frequently is going unexploited. Data mining techniques can help remedy this situation. Data mining have found several applications in Business and Scientific domains. Association rules, category, clustering are major areas of interest in information mining. [10]

In this, the author proposes a novel two-step predictive framework for ECG signal processing, where a global classifier recognizes severe abnormalities (red alarms) by comparing the signal against a universal reference model. The seemingly normal signal samples undergo a subsequent deviation analysis and yellow alarms are called by identifying mild and yet informative signal morphology distortions comparing to the learned patient-specific baseline that can be indicative of upcoming heart conditions. This methodology can be integrated with wearable heart monitoring devices to be used by health-care providers for elderly and highrisk heart patients. This software provides warnings about potential upcoming heart problems ahead of time to guide the patients to take precautions and preventive actions that are planned and approved by respective experts and physicians. The authors finally note that this methodology is applicable to other biomedical signals, including photoplethysmography (PPG), pulse oximeter (Pleth), and electroencephalogram (EEG). [11]

The main aim of the project has been to provide early detection of obesity related diseases. This work aims to overcome the above-mentioned limitations by developing a state-of-the-art system that streamlines machine learning algorithms for the effective prediction of Obesity and its related diseases considering the population of India. The government indulgent regulation on food restriction provides easy accessibility to unhealthy, processed food. Also, India's growing career-oriented lifestyle gives rise to irregular biological patterns especially in younger generation who prefer indoor games rather than playing outside. The behavioral and socio-psychological factors such as sleep, stress, ethnicity in addition to hormonal disorders are influential in giving rise to obesity in today's era. The research on Obesity helped us in providing a model specific for Indian population which detects the obesity in a person by considering various parameters rated by medical practitioners, and categorizes the person based on obesity. If the person is healthy, an optional Lipid Profile test can be taken to check if any further risk persists and that can be detected. If the person is obese, a test is conducted by taking the Symptoms of the patient as input to examine which obesity related disease the person is more prone to have such as Heart attack, Hypertension, Varicose Veins, Diabetes, Osteoarthritis. The developed system will undeniably be beneficial for predicting obesity, its related diseases and for the future betterment of an individual. [12]

Proposed Methodology

In our society, humans give greater attention to their very own health. Personalized health carrier is gradually rising. The lack of experienced docs and physicians, maximum healthcare businesses can't meet the clinical call for of public. Public need more accurate and immediately result. Thus, greater and more facts mining applications are evolved to provide humans greater custom designed healthcare service. It is a great solution for the mismatch of inadequate medical resources and rising clinical demands.

The proposed system we build which leverages data processing methods to reveal the connection between the regular physical examination records and therefore the potential health risk given by the user or public. this technique uses the Machine learning and

data processing algorithms like Na¨ıve Bayes, Support Vector Machine are used for the disease prediction and for the storage of the info the system used MYSQL database. The system provides a user-friendly interface for various users and doctors. during this paper, an efficient and privacy-preserving disease risk prediction scheme for e-healthcare is proposed. within the existing paper there's a drawback of security associated with patient's information. So within the proposed work we are getting to use encryption technique to supply security to the sensitive information of the patients. Compared with the prevailing work we are getting to use Na¨ıve Bayes algorithm to look the info , and used encryption algorithm to supply security and SVM algorithm for predict the diseases.

Advantages of Proposed System:

1.Increases human-pc interactions.
2. Recommends the clinic and physician to patient according to sicknesses predicted.
3. Shortest distance is provided

Modules:

In this different modules are used:

Module 1:
Patient: Patient will first register to the system and then login to the system with valid username and password. Then patient will search the doctor or hospital by their profession or by their name. The patients will enter the symptoms and according to that symptoms the patients will get the results
Module 2:
Doctor: Doctor will first register and login to the system by using valid username and password. Doctor can view the personal details of the user. Doctor can view the symptoms of the user and according to that symptom he/she will predict the disease through the system.
Module 3:
Admin: Admin will login to the system by using valid username and password. After login, Admin main page displays view users and view doctor functions. From where Admin can view all the information of the user and Doctors.

A. Architecture

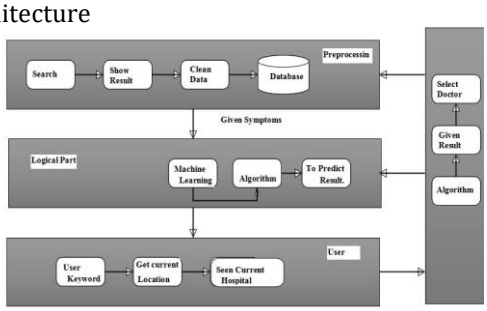

Fig. 1. Proposed System Architecture

Explanation:
In this framework, first we are login to the framework by utilizing OTP. After login to the framework we are going to look through the medical clinic or specialist. At that point enter the side effects or answer the inquiries that is created by the framework. At that point preparing of the calculations will be done and, finally, the anticipated sickness will be get as the conclusive outcome from the framework.
B. Algorithms
1. Naive Bayes Algorithm:
Naive Bayes algorithm is the algorithm that learns the probability of an object with certain features belonging to a particular group/class. In short, it is a probabilistic classifier. The Naive Bayes algorithm is called "naive" because it makes the assumption that the occurrence of a certain feature is independent of the occurrence of other features. Here we classify the heart disease based on heart check up attributes. Naive Bayes or Bayes' Rule is the basis for many machine learning and data mining methods. The rule (algorithm) is used to create models with predictive capabilities. It provides new ways of exploring and understanding data.
Why to prefer naive Bayes implementation:

- When the data is high.
- When the attributes are independent of each other.
- When we expect more efficient output, as compared to other methods output.

Based on all these information and steps we classify to predict the disease depending on its check up attributes.
Steps:
1. Given training dataset D which consists of documents belonging to different class say Class A and Class B 2. Calculate the prior probability of class A=number of objects of class A/total number of objects Calculate the prior probability of class B=number of objects of class B/total number of objects

3. Find NI, the total no of frequency of each class

Na=the total no of frequency of class A Nb=the total no of frequency of class B

4. Find conditional probability of keyword occurrence givena class:

P (value 1/Class A) =count/ni (A)
P (value 1/Class B) =count/ni (B)
P (value 2/Class A) =count/ni (A)
P (value 2/Class B) =count/ni (B)
.....
.....
.....
P (value n/Class B) =count/ni (B)

5. Avoid zero frequency problems by applying uniformdistribution
6. Classify Document C based on the probability p(C/W)
   a. Find P (A/W) =P (A)*P (value 1/Class A)* P (value 2/Class A)..... P(value n /Class A)
   b. Find P (B/W) =P (B)*P (value 1/Class B)* P (value 2/Class

B)..... P(value n /Class B)
7. Assign document to class that has higher probability.
C. Mathematical Model
1. Mathematical equation:
The algorithm implemented in this project is describe as:

$$P(class|features) = \frac{P(class) * P(features|class)}{P(features)} \quad (1)$$

where,

- P(class/features) : Posterior Probability
- P(class) : Class Prior Probability
- P(features/class) : Likelihood
- P(features) : Predictor Prior Probability

A. Normal distribution

The probability density of the normal distribution is:

$$f(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} {}^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (2)$$

Where

- 'μ' is the mean or expectation of the distribution,
- 'σ' is the standard deviation, and
- 'σ^2' is the variance.

IV. RESULT AND DISCUSSION

Experiments are done by a personal computer with a configuration: Intel (R) Core (TM) i5-6700HQ CPU @ 2.60GHz, 16GB memory, Windows 7, MySQl Server 5.1 and Jdk 1.8.

In our system, We compared the proposed disease prediction accuracy on number of samples and show the result graphically. Let see the following graph and table shows the disease prediction accuracy result based on classification technique. The overall accuracy of proposed classification technique is 95perc compare to existing techniques. So our proposed system prediction accuracy is better than existing system. So this works gives better disease prediction compare to existing method.

A. Results and Performance

Classification between Algorithms:

S.No	Classification Methods	No. of Samples	Predicted Samples	Rate of prediction
01	Proposed Classification	25	23	95%
02	Existing Classification	25	20	91%

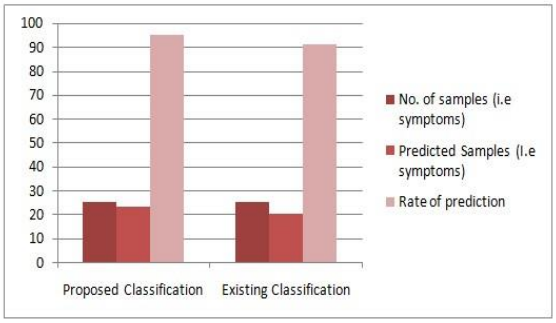

Fig. 2. Classification graph

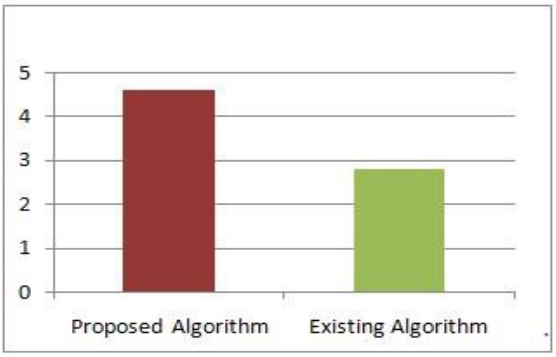

Fig. 3. Algorithms Comparison

Conclusion

In this paper, we proposed a privacy-keeping disease predicting machine which can assist physicians make proper prognosis of disease and offer fitness services for sufferers anytime everywhere in a privacy-maintaining manner. This task implements a ailment danger prediction machine which leverages data mining methods to show the relationship among the normal physical examination facts and the capacity health hazard given by way of the person or public. Different machine learning algorithms are carried out to predict physical status of examinee that will be at risk for physical decay one year from now. In our method person or affected person seek the sanatorium and the consequences given are consistent with the nearest region of present day place of person/patient. User / Patient give signs and symptoms and the machine predicts the disease and offers the medicines. The advantages of privacykeeping prognosis are to keep the balance between safety and efficiency which need to be taken into consideration firstly. Therefore, a way to optimize the model training the use of set for efficiency improvement and finding an effective way of introducing a few different advanced gadget studying methods to build the privacy-retaining disease prediction machine are worth of investigation.

Acknowledgment

The authors would like to thank the researchers as well as publishers for making their resources available and for their guidance. We are thankful to the authorities of University of Pune and concern members of CPGCON 2020 conference, organized by, for their constant guidelines and support. We are also thankful to the reviewer for their valuable suggestions. We also thank the college authorities for providing the required infrastructure and support. Finally, we would like to extend a heartfelt gratitude to friends and family members.

References

[1] Srinivas K, Rani B K, Govrdhan A. "Applications of Data Mining Techniques in Healthcare and Prediction of Heart Attacks". International Journal on Computer Science Engineering, 2010.

[2] Alessandra Toninelli, Rebecca Montanari, And Antonio Corradi "Enabling secure service discovery in mobile healthcare enterprise networks", IEEE Wireless Communications Volume: 16, Issue: 3, June 2009.

[3] Chaitrali S Dangare "Improved Study Of Heart Disease Prediction System Using Data Mining Classification Techniques", International Journal Of Computer Applications, Vol.47, No.10 (June 2012).

[4] Amma, N.G.B "Cardio Vascular Disease Prediction System using Genetic Algorithm", IEEE International Conference on Computing, Communication and Applications, 2012.

[5] IM. Chen, Y. Ma, Y. Li, D. Wu, Y. Zhang, and C. Youn, " Wearable 2.0: Enable human-cloud integration in next generation healthcare system," IEEE Commun. , vol. 55, no. 1, pp. 54–61, Jan. 2017.

[6] Lakshmishree J, K Paramesha, "Prediction of Heart Disease Based on Decision Trees", International Journal for Research in Applied Science Engineering Technology (IJRASET) , May 2017.

[7] M.A.Nishara Banu, B Gomathy, "Disease Predicting System Using Data Mining Techniques", International Journal of Technical Research and Applications, Volume 1, Issue 5 (Nov-Dec 2013), PP. 41-45.

[8] Hnin Wint Khaing, "Data Mining based Fragmentation and Prediction of Medical Data", International Conference on Computer Research and Development, ISBN: 978-1-61284-840- 2, 2011.

[9] Chitra R and Seenivasagam V, "Review of Heart Disease Prediction System Using Data Mining and Hybrid Intelligent Techniques", ICTACT Journal on Soft Computing, July 2013, Volume: 03, Issue: 04, 2013.

[10] MA.Jabbar, B.L Deekshatulu, Priti Chandra, "An evolutionary algorithm for heart disease prediction", CCIS,PP 378-389 , Springer(2012).

[11] Jiaming Chen, Ali Valehi, Abolfazl Razi, "Smart Heart Monitoring: Early Prediction of Heart Problems Through Predictive Analysis of ECG Signals", IEEE Access 2019.

[12] Naomi Christianne Pereira, Jessica D'souza, Parth Rana, Supriya Solaskar, "Obesity Related Disease Prediction From Healthcare Communities Using Machine Learning", 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT).

.