

A SURVEY PAPER ON A NOVEL APPROACH FOR FORECASTING HEALTH CARE USING DATA MINING

Mr.Sudesh Rao,Akshaya Shetty,Ruksana Banu,Ian Gladwin Alva,Simran Sheikh

Assistant Professor,Dept.of CSE,SUCET Mukka

UG Students, Dept.of CSE, SUCET Mukka

Abstract:

Data Mining is one of the interesting Areas of Research which is extremely popular in Health care Industry.This is used to extract the hidden patterns from junk of raw data's available and predict the output.Since,social-media is very much popular,Health care system is also implemented through online so that it can help to maintain patients health in day to day life.Here we find which medication is best and also its side effects.Then the symptoms of the disease is displayed.By taking the suggestions of the doctor,medication of any disease is listed out.The data is updated daily in order to help the Health-care centers.

Keywords— Data Mining, Forecasting, Health-care system, Medication, Symptoms.

I. INTRODUCTION

Data Mining is the process of extraction of large data Which involves Machine Learning, Database systems and so on. Machine Learning provides the ability to learn automatically and improve from past experience without being explicitly programmed. Its algorithm has two passes: Training and Testing.

Here by using Patients symptoms, prediction of disease is made. The data generated by Health care system is very vast. Hence Data Mining Technique is used to solve this issue. This website is designed in such a way that any one can use this from any corner of their residence. Hence the user can also use their website to know the Symptoms on their own, without the involvement of Doctor. This proposed Health care system will help the Patients to get diagnosed with disease easily without being physically going to the hospital. Hence, the Data

Mining Technique is used in this system to provide user Friendly application.

Here the user needs to login into system by registering into the website. After logging, the users are given with

unique token id's. The user can enter the symptoms and predict the disease accordingly. The doctor will suggest the medication by using user token id and the prescription is given to that patient. The patient can also give the feedback regarding their experience.

II. PROPOSED SYSTEM

Login, the user will login to the first page using his/her registered name and password. After logging successfully,

Authorized page will be displayed else error message is shown.

Registration, the user must register into the page in order to maintain separate account for each individual. Registration is must before log in.

Doctor Registration, only admin is capable to register new doctor. Admin can enter doctor details in this module such as his/her name, specialization and other related information.

Approve user, the admin can only approve the user who registered into the website in order to login. Admin will approve the user by viewing user details.

View feedback, admin can view the feedback in this page.

View patient details, in this page doctor can view patient details based on token id given to each user.

Provide medication, in this page doctor will prescribe the medications to the patient.

View Doctor, here patients can view all the doctors present in the hospital and their related information.

Feedback, the user can give the feedback based on their experience.

III. EXISTING SYSTEM

System can predict diseases, but not its subtypes caused by one disease. The system fails to predict all the subtypes caused by a particular disease. Existing system handles only the data which is structured. The prediction system is vast and complex. In current past prediction systems were only affordable by richer people. When it comes to common folks its not easily affordable. So far the guess systems are non-specific, that is , Machine can predict the disease but not its sub types.

IV. FUNCTIONAL REQUIREMENTS

Modules:

Login: The user will login to the main page using his/her registered name and password. After logging successfully,

Authorized page will be displayed else error message is shown.

Registration: the user must register into the page in order to maintain separate account for each individual. Registration is must before log in.

Doctor Registration: only admin is capable to register new doctor. Admin can enter doctor details in this module such has his/her name, specialization and other related information.

Approve user: the admin can only approve the user who registered into the website in order to login. Admin will approve the user by viewing user details.

View feedback: admin can view the feedback in this page. Which user has given what feedback can be viewed by Admin.

View patient details: in this page doctor can view patient details based on token id given to each user.

Provide Medication: in this page doctor will prescribe the medications to the patient.

View Doctor: Here, patients can view all the doctors present in the hospital and their related information.

Feedback: the user can give the feedback based on their experience.

V. HARDWARE REQUIREMENTS

- RAM: 1GB or above
- Hard disk: 10GB or above
- Processor: 2.4GHZ or above

VI. SOFTWARE REQUIREMENTS

- Language: Python, .net, c#, javascript, HTML.

- Operating Source: Windows 7, 8, 10
- Back end: Mysql
- Front end: Visual code, visual studio

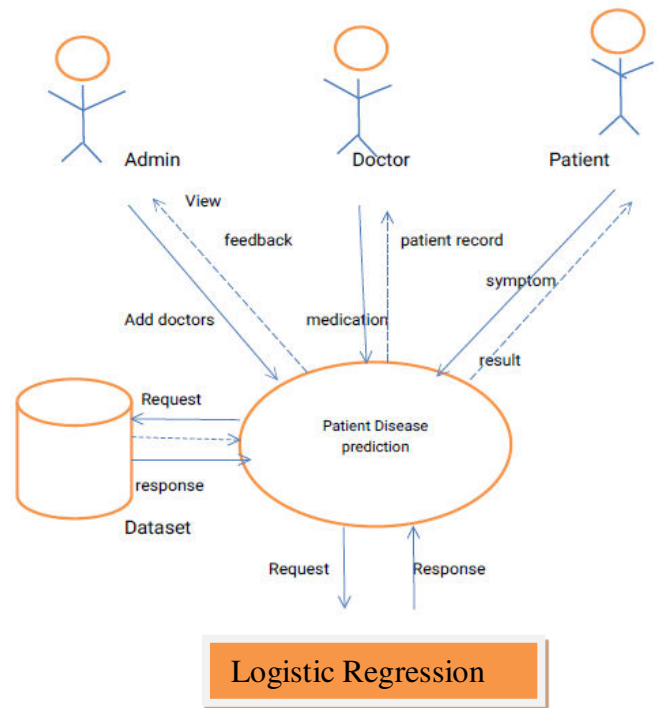
VII. LITERATURE SURVEY

In 2010, Apache Hadoop defines big data as “ data sets which could not be apprehended, succeeded and managed by general computers within an okay scope.” On this basis in May 2011 , McKinsey & Company, a global accessing help said Big Data as the next edge for improvement, this includes two associations: First, data sets that obey to the big data are shifting and may cultivate or with scientific developments. Second, data sets measurements that adapt to the ordinary big data. Clinical records and dealing with patients must have a bigger research. Data mining of health record provides new patient-stratification doctrines and for unknown disease links. Further medical examination and care using EHR data and tasks that must be performed. The Health care of many countries are very limited. For example: 80% of the people living in China doesn't have adequate source to health care system and 80% medical shops are located in big cities. Construction of Health care system by using Internet of Things (IOT), big data and cloud computing is the way to solve the above difficulties. Big Health is a talented industry, which is differentiated by people-center, managing a person's health from birth to death and from anticipation to rehabilitation. This field covers health goods such as drugs, medical devices and health service such as medical services, income services and fitness real estate such as pension and health finance which includes health protection, financial products. Chinese Herbal products are introduced to the patients who suffer from hyperlipidemia. The study of population of about 1,000,000 folks from National Wellbeing Insurance Exploration. It observed that 30,784 folks are linked with hyperlipidemia and collected their medical records. In this chapter it witness the use of recurrent neural networks. In this method, the weight is computed by weighted sum of the hidden states of the RNN at each word according to their scores. The system achieves end-to-end RNN and user click logs. The worker click logs are sampled from commercial search engine. It highlights the two aspects: query rewriting and a modified BM25 metric. The system explains that using learned attention scores one will be able to produce better qualities of sub-queries. To improve its performance evaluated by AUC. Abstract Traditional wearable devices have various drawbacks, such as uncomfortableness for long-term wearing, and insufficient accuracy. Health care using traditional wearable is hard to sustain. In order to manage healthcare the system design “Smart Clothing” is enabled. To offer good Smart Clothing erection, cloud stand is constructed by using mobile internet. This paper introduces design, requirements, and implementation methods. Claims made by Smart Clothing and big data clouds are presented such as medical backup

response, emotion care, disease diagnosis. Here, new deep learning manner Bi-CNN-MI for paraphrase identification. PI needs two sentences which is created on the vision, which learns multigranular images using convolutional neural network and model boundary features. These are input to logistic classifier for PI. All limits of the model are straight optimized. In order to address the lack of training data, the system provides the network in a novel method using language modeling task. MSRP corpus is used to estimate lung cancer using the double dispensation system. The challenge in this system is the recognition of tiny nodes which earlier used to cancer finding. The knobs in the lungs can be found using recognition algorithm. In this system it generates gene information, DNA methylation and miRNA. The Data Mining algorithms classify the data and extract knowledge. It also discusses the difficulties in algorithms like Support Vector Machine, Naïve Bayesian classification, Rough set theory, Decision Tree and does the study on decision trees and their behavior. Tree node splitting is based on relevant feature selection which is a key in decision tree learning. The partitioning of recursive nodes results in geometric reduction of data quantity of the leaf nodes which causes complex in model and data over fitting. The author presented a novel architecture called a Decision Stream. The paper is based on the behavior of smoker. The e-cigarette has a coiled wire in 1.5 ohms which is connected to positive and negative poles. When its button is pressed, the coil is connected to electrical supply immersing “E-liquid”. The coil gets heated up and transforms to vapor, which is inhaled by smokers. This system tracks the behavior of the smoker in order to prevent the user from cancer.

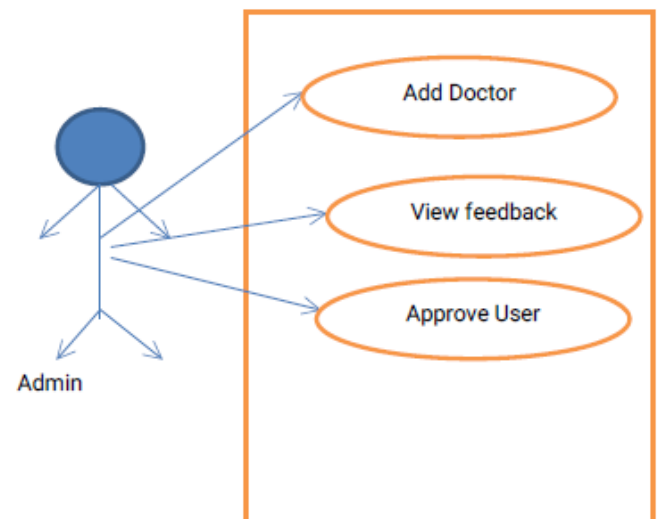
VIII. ARCHITECTURAL DESIGN

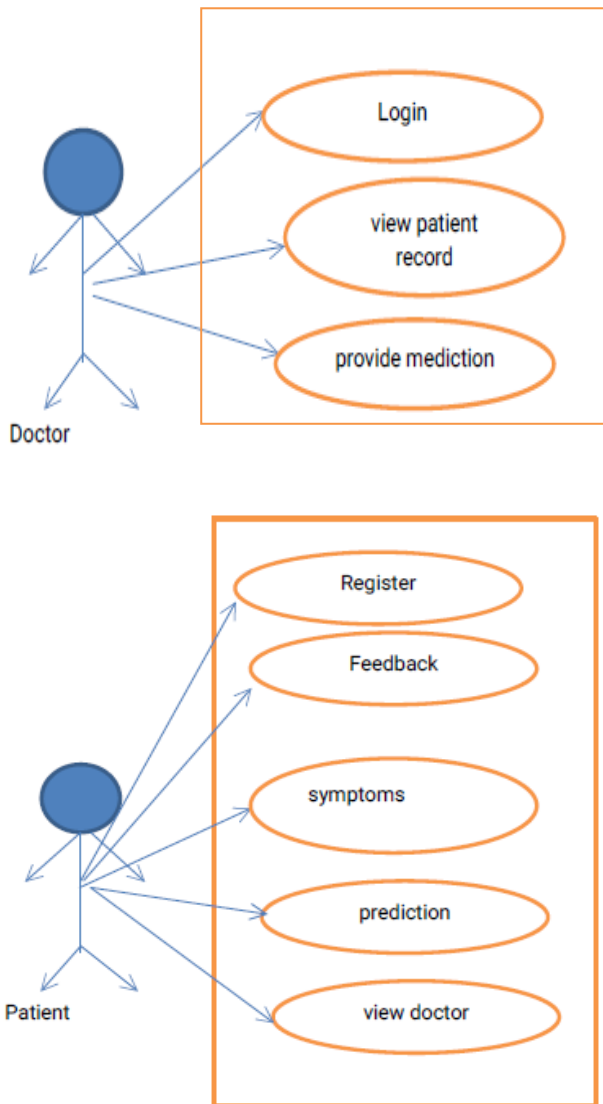
This design is a graphical representation of set of theories, which is a part of Architecture, including their principles, elements and components. There are different kinds of architecture diagram, system architecture diagram, security architecture diagram, application architecture diagram and software architecture diagram and so on. Architecture is coherent set of theories formed to structure.



Use Case Diagram:

The purpose of use-case diagram is to find the dynamic aspect of a system. This is used to gather the requirements of a system. These requirements are design requirements. Use cases are prepared and actors are identified.





IX. SYSTEM IMPLEMENTATION

This is the stage where theoretical design is converted into working system, by replacing the existing system. The system is implemented by using the following:

Anaconda: with a list of python packages, editors, Python distributions include python interpreter. Anaconda is one among python distributions. New distribution of python is python and R data science package. Anaconda has more than 100 new packages. The latest version of Anaconda 5.0.1 is released in October 2017.

Python: Python is an interpreted, high-level, general-purpose programming language. Founded by Guido van Russom,

released in 1991. It constructs object-oriented approach to help to write clear, small code for large projects.

SQLite: (Structured Query Language) is a domain specific Domain-specific language which is used in programming and designed for managing data in relational database management system or for stream processing in relational data stream management system. This is useful in handling structured data.

X. TESTING

Software Testing helps to identify correctness, Completeness, security and quality of computer software. This is the process of executing programs in order to find Errors. The testing phase evaluate the software which has been developed to confirm that it produces safe outputs. It will inherent the errors occurred here. The user is Informed so that they can follow the guidelines. While testing, the program is executed with set of test cases, and the output of the program with test case is evaluated, to check if the program is performing as expected.

The first step in testing is determining the errors in a Program. In this phase, the errors occurred from the previous phases must be detected.

A. Objectives of Testing:

- Testing is a process of executing a program with the intention of finding errors.
- A successful is one that un-covers an as yet undiscovered error.

B. Test Case Design:

There are two types of testing. They are:

- White Box Testing
- Black Box Testing

White Box Testing: This method is used to control the Structure of procedural design to drive test case. All logical decisions on true and false side execute all loops at their boundaries

Black Box Testing: This will focus on functional requirements of the software.

Black Box Testing finds errors in following stages:

- Incorrect or missing function.
- Interface errors.
- Errors in data structures or external database access.
- Performance errors.

XI. CONCLUSION

Data Mining is useful in Health Care Organizations. Health Care is the top sector that is highly beneficial. This system Will definitely reduce human efforts, reduce in cost and can manage time efficiently. Data Mining is quite risky since the data's are junk and complex. But its tools help us to seem it in an interesting way. Data Mining Techniques can also be used in other fields as well. In this system the user can predict the disease by inputting the symptoms without the intervention of the doctor. Basically this project detects the disease on the basis of symptoms. The system is designed in such a way that it takes the symptoms from the user as Input and then produces output i.e, it will predict the disease. In this paper, Machine Learning Logistic Regression algorithm is bid by using structured and unstructured data from hospital. And also uses Logistic Regression algorithm for partitioning data. The accuracy Of our proposed algorithm reaches 94.8% with regular speed which is faster than CNN-based unimodel disease risk prediction Logistic Regression algorithm and produces output.

REFERENCES

- [1] Min Chen, Yixue Hao, Kai Hwang, Fellow, IEEE, LuWang, and Lin Wang "Disease Prediction by Machine Learning over Big Data from Healthcare Communities"(2017)
- [2] .Mr. Chala Beyene, Prof. Pooja Kamat, "Survey on Prediction and Analysis the Occurrence of Heart Disease Using Data Mining Techniques", International Journal of Pure And Applied Mathematics, 2018.
- [3] S. Patel and H. Patel, "Survey of data mining techniques used in healthcare domain," Int. J. of Inform .Sci. and Tech., Vol. 6, pp. 53-60, March 2016 .
- [4] S.-H. Wang, T.-M. Zhan, Y. Chen, Y. Zhang, M. Yang, H.-M. Lu, H.-N. Wang, B. Liu, and P. Phillips, "Multiple sclerosis detection based on biorthogonal Wavelet transform, rbf kernel principal component analysis, and logistic regression," IEEE Access, vol. 4, pp. 7567–7576, 2016.