



Disease Prediction Using Machine Learning

Saidulu Pallapati, Tarunkumar Parisa, Vinay Challa,
Naga Sai Saravanan Yanamadala and Himadri Vegad

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.

March 28, 2024

DISEASE PREDICTION USING MACHINE LEARNING

PALLAPATI SAIDULU

*Computer Science and Engineering
PIET, Parul University
Vadodara, India
200303125059 @paruluniversity.ac.in*

PARISA TARUNKUMAR

*Computer Science and Engineering
PIET, Parul University
Vadodara, India
200303125062 @paruluniversity.ac.in*

CHALLAVINAY

*Computer Science and Engineering
PIET, Parul University
Vadodara, India
200303125030 @paruluniversity.ac.in*

Y.NAGA SAI SARAVANAN

*Computer Science and Engineering
PIET, Parul University
Vadodara, India*

Himadri Vegad

*Computer Science and Engineering
PIET, Parul University
Vadodara, India*

200303125070 @paruluniversity.ac.in

himandri.vegad21143@paruluniversity.ac.in

Abstract

The types of viruses and bacterial infections are increasing exponentially every year. These viruses and bacteria are air transmitted. People are mostly attacked by these when they go to some clinic for a minor check up or doctor suggestion. To counter this problem we are trying to design a efficient machine learning algorithm to predict the problem you might have in turn decreasing the need for a common man to visita hospital. We are taking our project to next level by adding modules to book doctors slot online and buy medicines online which will also help in minimizing the need to go near a hospital. Our idea has a lot of future scope as we can make this data into logs and use them to find the areas which are affected with the viruses and can alert people to keep their distance from that area.

Keywords

Disease, prediction, Machine-learning, Training, Testing, symptoms.

Introduction

Machine learning is an arising approach that helps in determination of diseases using model information or past data. The Machine learning algorithm has two passes: Training and Testing. A Forecast of a disease by using a patient's symptoms and history machine learning technology is battling for past many years. Machine Learning technology gives a decent stage in the clinical field, so a medical care issue can be solved proficiently. We are applying machine learning to maintain complete hospital data Machine learning technology which allows building models to get rapidly examine information and convey results quicker. With the use of machine

learning technology doctors can make good decision for patient diagnoses and treatment options, which prompts improvement of patient medical care administrations. Machine Learning technology gives a decent platform in medical field, so that a healthcare issue can be tackled proficiently. The fundamental target of our project is to make the user Interface as apart of front end and connect it with ml models using Django python web framework. We will predict the disease in 2 variants i.e., in general and for specific disease. There is a need to study and make a framework which will make it simple for an end user to anticipate the constant sicknesses without visiting a specialist or a doctor for the treatment and also predicts the type of doctors they should visit. Machine Learning has great power to analyse and cope up with different diseases so that prediction is more accurate and it is cost effective in the treatment Computers give us information; they engage us and help us in a lot of manners. A chatbot is a program intended to counterfeit smart communication on a text or speech. Yet, this paper concentrates only on text. These systems can learn themselves and restore their knowledge using human assistance or using web resources.

Literature Survey

We propose a Wearable 2.0 healthcare system to improve QoE and QoS of the next generation healthcare system In the proposed system, washable smart clothing, which consists of sensors, electrodes, and wires, is the critical component to collect users' physiological data and receive. [1]. Conversational service for psychiatric counseling that is adapted methodologies to understand counseling contents based on of high-level natural language understanding (NLU), and emotion recognition based on multi-modal approach The methodologies enable continuous observation of emotional changes

sensitively. [2]. The first build of Rattata focuses on using a method of scanning for keywords within the input, then pull a reply with the most matching keywords. The database is an xml-based dataset from stack-exchange. [3]. They can also be used in the field of medical to solve health related problems. Chatbots are going to explode and can be really dominating in future. Chatbots can provide a new and flexible way for users. They are giving AI something better to do. [4]. This is a medical application where the goal is to predict the risk of patients on certain disease using Electronic Health Records (EHR). They proposed an interactive patient risk prediction method, which actively queries medical experts with the relative similarity of patients. [5]. They implemented work that can be deduce that the factors affecting person's health can be analyzed very closely. They implemented work that can be deduce that the factors affecting person's health can be analyzed very closely [6]. They said The final comparison of the results obtained by the different algorithms employed for prediction of disease prediction from hospital data, revealed that CART model or simply said, decisiontree gave the highest performance, followed by logistic regression, KNN, Naïve Bayes, SVM, Random Forest, and LDA. RFE method of feature selection even though employed on only two models, made a major impact. [7]. Study results are encouraging because they demonstrate significant improvement in the standard of care for acute ischemic stroke patients across the United States. Improvements in performance metrics were more frequently observed in older patients and resulted in narrowing of age-related treatment gaps. However, most of the hospitals participating in the GWTC Stroke program are self-selected and probably have more interest in quality improvement, thereby challenging the generalizability of the results to the entire nation. [8]. They used a Hybrid Machine Learning Algorithm that combined Support Vector Classification and Random Forest Classification. They used a Hybrid Machine Learning Algorithm that combined Support Vector Classification and Random Forest Classification was developed for disease predictions like heart and diabetes. [9]. Seven computational intelligence techniques named as Logistic regression (LR), Support Vector Machine (SVM), Deep Neural Network (DNN), Decision Tree (DT), Naïve Bayes (NB), Random Forest (RF), and K-Nearest Neighbor (K-NN). In this paper, we compared a number of computational intelligence techniques for the prediction of coronary artery heart disease. [10]. There are some popular mental counseling applications in the markets, most of which are based on one mental survey or examination. they noticed that this can be done in a more natural/intelligent way-texting, talking or even videoing. Our application can interact with users by texting, talking and videoing, which are based on natural language processing (NLP) and integrated continuous emotion dialogue analysis

and sentence/audio generation. [11]. While a black box approach is used, by controlling the communication structure, to and from the web-service, the web-service allows all types of clients to communicate to the server from any platform. [12]. The chatbot consists of core and interface that is accessing that core in relational database management systems (RDBMS). The database has been employed as knowledge storage and interpreter has been employed as stored programs of function and procedure sets for pattern-matching requirement. [13]. This Chatbot is very simple and user friendly. It is not very complicated like other Chatbots. The working of the Chatbot is simple and can be easily understood by any person. In other Chatbots, the working is very complicated. Many classes are used which is difficult to understand. [14]. R programming Language can be further framed as a query with the help of Artificial Intelligent Markup language (AIML) script. Along with AIML script of the query is unsuccessful then this will be framed as query with help of SQL lite. R programming Language can be further framed as a query with the help of Artificial Intelligent Markup language (AIML) script. Along with AIML script of the query is unsuccessful then this will be framed as query with help of SQL lite. [15]. To get better outcomes, the system might be trained with a bigger and more thorough dataset. Future development might include extending the system to include more languages. Deep Learning algorithms may be used to improve illness categorization accuracy and outcomes. Natural Language Generation may be used to improve chatbot replies by training a model on diverse conversational datasets. [16]. Chatbot is a computer program designed to simulate interactive conversations or communication to users. In this study, chatbot was created as a customer service that functions as a public health service in Malang. This application is expected to facilitate the public to find the desired information. The method for user input in this application used N-Gram. [17]. So by using SVM algorithm and disease symptoms system can predict disease. The system takes a plain text as input and answering all type of questions output by qualified user is the output. [18]. They proposed system the doctors and patients are brought to one place which is handled by admin and the admin allows patients to register and search for the doctors basing on the location the list of doctors will be shown and patient can book by selecting the time slots and the admin will confirm the booking so everything is computerized and done very fast which will save time. [19]. In a simulation study that considers a model clinic, which is created using data obtained from an actual clinic, we find that the heuristics proposed outperform all the other benchmark policies, particularly when the patient load is high compared with the regular capacity. This paper develops a framework and proposes heuristic dynamic policies for scheduling patient appointments. This paper develops a

framework and proposes heuristic dynamic policies for scheduling patient appointments, taking into account the fact that patients may cancel or not show up for their appointments.[20].

Requirements

A. HARDWARE

- System : Intel(R) Core(TM) i3 7020U CPU @ 2.30GHz
- Hard Disk : 1 TB.
- Input Devices : Keyboard, Mouse
- Ram: 4GB.

B. SOFTWARE

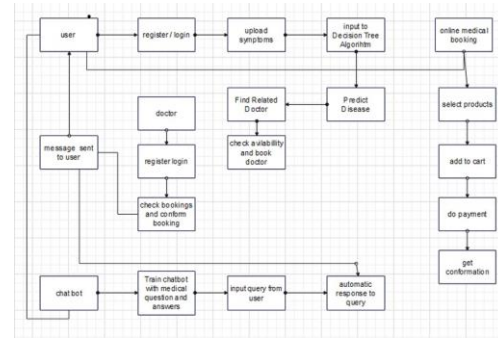
- Operating system : Windows XP/7/10.
- Programming language : Python
- Interface : Flask Framework.
- Tools : Anaconda.
- Dataset : Disease Dataset, Chatbot dataset.
- Database : Mysql.

Methodology

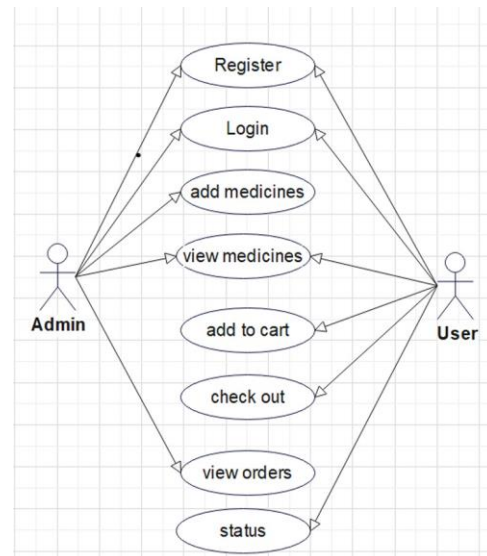
Predicting diseases using machine learning involves a systematic methodology aimed at harnessing data-driven insights for accurate and timely diagnosis. It begins with a precise definition of the problem at hand, specifying the target diseases for prediction and identifying relevant features crucial for accurate assessments. Subsequently, a comprehensive collection of data from diverse sources such as electronic health records, medical imaging data, genetic information, and lifestyle factors is undertaken, ensuring data integrity and representativeness. The collected data then undergoes meticulous preprocessing, encompassing tasks such as handling missing values, cleaning outliers, selecting pertinent features, and transforming data to a suitable format for analysis. Interpreting models is a crucial step, offering insights into how predictions are made and shedding light on the most influential features driving these predictions. Techniques like SHAP values or LIME aid in understanding model decisions, contributing to model transparency and trustworthiness. Once models are thoroughly evaluated and interpreted, they are primed for deployment into production environments, such as healthcare systems or mobile applications, enabling real-time disease prediction and intervention.

System Architecture

A system architecture or systems architecture is the conceptual model that defines the structure, behavior, and more views of a system. an architecture description is a formal description and representation of a system. organized in a way that supports reasoning about the structures and behaviors of the system.



Usecase Diagrams:



Conclusions

We were able to build a ML algorithm that accurately predict the diseases with six symptoms input and also help the customer in booking and buying medicine and help him get in contact with doctor. In this application machine learning based medical helper application is developed by testing disease dataset with multiple machine learning algorithms and most accurate algorithm is used to predict disease which is used in flask web framework. Using this framework health website is designed which has doctor appointment booking, chat bot helper, medicine booking, disease prediction all health-related services are integrated in single application. and

make the fantasy to work out as expected

Future Work

In future disease data for different disease are collected and trained using Deep learning methods to get more effective results and accuracy. Segmentation of MRI scans can be applied to dataset can be integrated to website. We can collect the disease data in logs and try marking areas with large virus spreads and alert users. This tech will also help government and make their jobs easier.

References

- 1) Author: M. Chen, Y. Hao, K. Hwang, L. Wang, and L. Wang, Title: "Disease prediction by machine learning over big data from healthcare communities" Year:2017.
- 2) Author: B. Qian, X. Wang, N. Cao, H. Li, and Y.-G. Jiang, Title: "A relative similarity based method for interactive patient risk prediction," Year: 2015.
- 3) Author: IM. Chen, Y. Ma, Y. Li, D. Wu, Y. Zhang, and C. Youn, Title: " Wearable 2.0: Enable human-cloud integration in next generation healthcare system," Year: 2017.
- 4) Author: Y. Zhang, M. Qiu, C.-W. Tsai, M. M. Hassan, and A. Alamri, Title: "HealthCPS: Healthcare cyberphysical system assisted by cloud and big data," Year: 2017.
- 5) Author: L. Qiu, K. Gai, and M. Qiu, Title: "Optimal big data sharing approach for telehealth in cloud computing," Year: 2016
- 6) Author: Ajinkya Kunjir, Harshal Sawant, NuzhatF.Shaikh, Title: "Data Mining and Visualization for prediction of Multiple Diseases in Healthcare," Year:2017.
- 7) Author: Shanthi Mendis, PekkaPuska, Bo Norrving, Title: Global Atlas on Cardiovascular Disease Prevention and Control. Year: 2011
- 8) Author: Amin, S.U.; Agarwal, K.; Beg, R., Title: "Genetic neural network based data mining in prediction of heart disease using risk factors", Year: 2013.
- 9) Author: K.Oh,D.Lee,B.KoandH.Choi,Title:"A Chatbot for Psychiatric Counseling in Mental Healthcare Service Based on Emotional Dialogue Analysis and Sentence Generation", Year: 2017
- 10) Author: Bayu Setiaji and Ferry Wahyu Wibowo, Title: "Chatbot Using a Knowledgein Database: Human to-Machine Conversation Modeling", Year: 2016
- 11) Author: Menal. Dahiya, Title:"A Tool of Conversation: Chatbot", INTERNATIONAL JOURNAL OF COMPUTERSCIENCE-SANDEENGINEERING, Year: 2017.
- 12) Author: C.P. Shabariram, V. Srinath and C.S. Indhuja, Vidhya, Title: "Ratatta: Chatbot Application Using Expert System", Year: 2017.
- 13) Author: Mrs Rashmi Dharwadkar and Neeta A. Deshpande, Title: "A Medical ChatBot", Year: 2018.
- 14) Author: Farheen Naaz and Farheen Siddiqui, Title: "modified n-gram based model for identifying and filtering near-duplicate documents detection", Year: 2017.
- 15) Title: N-gram Accuracy Analysis in the Method of Chatbot Response" Year:2018.
- 16) Author: V.K Shukla and A Verma, Title: "Enhancing LMS Experience through AIML Base and Retrieval Base Chatbot using R Language", Year: 2019
- 17) Author: F. Talib and Z. Rahman, Title: Current health of Indian healthcare and hospitality industries: a demographic study, Year: 2013.
- 18) Title: S. Sri Gowthem and K.P. Kaliyamurthie, Title: "Smart Appointment Reservation System", Year: 2015.
- 19) Author: Y. Jiang, H. Zhan and Z. Huang, Title: "Design and Implementation of A General Web-based Course Teaching Management System", Year: 2012.
- 20) Author: N. Liu, S. Ziya and V. Kulkarni, Title: "Dynamic Scheduling of Outpatient Appointments Under Patient No-Shows and Cancellations", Year:2010.