PARAMETER BASED DISEASE PREDICTION AND RECOMMENDATION

Miss.S.V.Gunjal, Raka Darshan, Pawar Rohit, Halnor Vaibhav, Ambre Uday

1 Lecturer, Cloud Computing and Big Data, P.Dr.V.V.P Polytechnic Loni, Maharashtra, India
2 Student, Cloud Computing and Big Data, P.Dr.V.V.P Polytechnic Loni, Maharashtra, India
3 Student, Cloud Computing and Big Data, P.Dr.V.V.P Polytechnic Loni, Maharashtra, India
4 Student, Cloud Computing and Big Data, P.Dr.V.V.P Polytechnic Loni, Maharashtra, India
5 Student, Cloud Computing and Big Data, P.Dr.V.V.P Polytechnic Loni, Maharashtra, India

ABSTRACT

Hospital recommendation System is considered as an important factor in health care Sector for managing the administrative, financial and clinical aspects of a hospital. Due to data mining progress in biomedical and healthcare communities, accurate study of medical data benefits early disease recognition, patient care and community services. When the quality of medical data is incomplete the exactness of study is reduced. Moreover, different regions exhibit unique appearances of certain regional diseases, which may results in weakening the prediction of disease outbreaks. In the proposed system, it provides machine learning algorithms for effective prediction of various disease occurrences in disease- frequent societies. It experiment the altered estimate models over real-life hospital data collected. To overcome the difficulty of incomplete data, it use a latent factor model to rebuild the missing data. It experiment on a regional chronic illness of cerebral infarction. Using structured and unstructured data from hospital it use Machine Learning Decision Tree algorithm. It predicts probable diseases and hospitals by mining data sets. To the best of our knowledge in the area of medical big data analytics none of the existing work focused on both data types. Compared to several typical estimate algorithms, the calculation exactness of our proposed algorithm reaches 94.8% with a convergence speed which is faster than that of the Decision tree disease risk prediction algorithm.

Keywords: Disease, Suggestion, Healthcare, Android App

1. INTRODUCTION:

According to the statistics of National Health and Family Planning Commission, there are a total of 991,632 medical institutions in China until Nov. 2016 where new institutions are increased from Nov. 2015. Nevertheless, the difficulty of getting medical care remains as one of China's major livelihood issues. A survey on Peking University First Hospital, which is one of the most famous hospitals in Beijing, indicates that more than 45% outpatients have to wait for over two hours after the registration, whereas 85% of them have less than 10 minutes for the doctor's inquiry. This phenomenon is actually common in China's 776 Top-Class hospitals1.

The reason is that many hospitals lack publicity and are not familiar to most patients. Having no way of knowing whether there is a good enough and less crowded hospital nearby, people have no choice but go to those congested famous hospitals regardless of the severity of the disease. In fact, most outpatients with a mild disease may expect a quick treatment yet have a low
requirement for the hospital's treatment ability. It is imperative to find a simple and convenient way to access the crowd status and basic information of the neighbouring hospitals. There are several standard ways for crowd counting, such as approaches based on video or beacon.

However, these methods rely on surveillance data or wireless network data, and any company or non-governmental organization can hardly gather these data of all hospitals even in one city, not to mention in any larger range.

Fortunately, location-based service (LBS) big data offers a potential solution to this dilemma. LBS data have two unique properties: a) LBS data naturally belongs to the service providers.

1.1 PROBLEM DEFINATION:

The problem definition for a parameter based disease prediction and recommendation uses navigating the intricacies of Parameter-Based Disease Prediction and Recommendation involves overcoming the complexities of integrating multifactorial health parameters.

This requires addressing challenges in predictive accuracy, data privacy, and ethical considerations, while striving for clinical adoption and patient engagement. Developing a robust and adaptable system, capable of continuous learning, is paramount to harnessing the depth of healthcare data for personalized, effective, and ethically sound disease predictions and recommendations.

1.2 PROPOSED SYSTEM

The proposed system for Parameter-Based Disease Prediction and Recommendation aims to revolutionize healthcare by introducing a comprehensive and adaptive framework. Integrating diverse health parameters, the system employs advanced predictive algorithms to accurately forecast disease risks. Personalized recommendations, grounded in ethical considerations and data privacy, empower individuals to proactively manage their health.

The system emphasizes clinical adoption, seamless interoperability, and continuous learning, fostering a scalable solution that aligns with evolving medical knowledge and technological advancements. Ultimately, the proposed system aspires to redefine healthcare delivery by providing actionable insights, enhancing patient outcomes, and contributing to a proactive approach in disease prevention and management.

- **Comprehensive Framework:** Develop a comprehensive system integrating diverse health parameters for a holistic understanding of individual health profiles.
- **Predictive Algorithms:** Utilize advanced predictive algorithms to accurately forecast disease risks based on multifactorial parameters.
- **Personalized Recommendations:** Generate personalized health recommendations, prioritizing ethical considerations and ensuring robust data privacy.
- **Clinical Adoption Focus:** Emphasize seamless integration with existing healthcare systems to promote clinical adoption and facilitate interoperability.
- **Continuous Learning Mechanism:** Implement a continuous learning mechanism to adapt to evolving medical knowledge and technological advancements.
- **Scalable Solution**: Design the system with scalability in mind to accommodate a growing volume of health data and diverse healthcare settings.
- **Actionable Insights:** Provide actionable insights for individuals, empowering them to proactively manage their health and mitigate potential risks.

2. LITERATURE SURVEY:

Paper-1: Artificial Intelligence and Internet of Things Enabled Disease Di- agnosisModel for Smart Healthcare Systems(17 March 2021)

Authors: Romany Fouad Mansour ,Adnen El Amraoui , Issam Nouaouri ,Vicente Garc´ıa D´ıaz ,Deepak Gupta , Sachin Kumar

Journal Name: IEEE Access

Summary: In this paper surveys the Designing and development of a novel AI andIoT convergence-based disease diagnosis model for the smart health system, Proposed a CSO- CLSTM model for diagnosing diabetes and heart dis- ease, incorporates iForest technique-based outlier detection process to improve the classification results, Performed parameter tuning of LSTM model using CSO Algorithm Also Validated the performance of the CSO-LSTM model on two benchmark datasets.

Paper-2: Prediction of Heart Disease Using a Combination of Machine Learning and Deep Learning
Authors: Rohit Bharti , Aditya Khamparia , Mohammad Shabaz , Gaurav Dhiman , Sagar Pande and Parneet Singh
Journal Name: Hindawi

Summary: In this paper three methods in which comparative analysis was done and promising results were achieved. The proposed approach was applied to the dataset in which firstly the dataset was properly analyzed and then different machine learning algorithms consisting of linear model selection in which Logistic Regression was used.

Paper-3: Machine Learning Based Diabetes Classification and Prediction for Healthcare Applications.
Authors: Umair Muneer Butt , Sukumar Letchmunan , Mubashir Ali , Fadratul Hafinaz Hassan , Anees Baqir and Hafiz Husnain Raza Sherazi 10
Journal Name: Hindawi AVCOE, Department of Computer Engineering 2022-23 V

Summary: In this paper The proposed diabetes classification and prediction system has exploited different machine learning algorithms. performed with state-of-the-art algorithms. -e experimental results show the supremacy of the proposed algorithm as compared to state-of-the- art algorithms. -e details of the dataset, performance measures, and comparative analysis performed are described in the following sections.is study has also proposed the architecture of a hypothetical diabetic monitoring system for diabetic patients. -e proposed hypothetical system will enable a patient to control, monitor, and manage their chronic conditions in a better way at their homes.

2.1 EXISTING SYSTEM:

In the realm of existing healthcare systems, the predominant frameworks include traditional health records, whether paper-based or electronic, which primarily focus on historical medical data without extensive predictive capabilities. Additionally, there are disease-specific predictive models designed for particular conditions, often lacking integration with diverse health parameters and limiting the scope of predictive insights. The landscape is characterized by static recommendation systems that provide generic health suggestions without dynamically adjusting to changes in individual health parameters.

Fragmented interoperability among various health systems poses challenges in efficient data exchange and the creation of comprehensive patient profiles. Privacy concerns and data security issues, particularly with electronic health records, remain prominent. User engagement in existing systems is typically low, with patients often playing a passive role in their healthcare management. Manual efforts are often required for the integration of diverse health parameters, introducing inefficiencies and potential inaccuracies. The prevailing healthcare approach tends to be reactive, addressing diseases after onset rather than offering proactive measures for prevention.

Predictive models in use often demand significant computational resources, limiting scalability and accessibility in various healthcare settings. Existing systems struggle with adapting to changes in patient health parameters over time, resulting in static recommendations. Furthermore, there is a limited emphasis on incorporating lifestyle factors in disease prediction, overlooking crucial elements that contribute to an individual's overall health. Recognizing these limitations in the current healthcare landscape underscores the necessity for the development of an improved and comprehensive Parameter-Based Disease Prediction and Recommendation system

2.2 DRAWBACK OF EXISTING SYSTEMS:

Limited Predictive Capabilities: Existing systems often lack robust predictive capabilities, relying more on historical data than on advanced algorithms to foresee potential future health issues.
Fragmented Data Integration: Many systems struggle with the seamless integration of diverse health parameters, resulting in fragmented data and an incomplete understanding of the patient's overall health profile**.**
Privacy and Security Concerns: The prevailing concern of privacy and security breaches in existing systems, especially with electronic health records, poses a risk of unauthorized access to sensitive patient information.
Reactive Approach to Healthcare: Most systems follow a predominantly reactive approach, addressing health issues after their onset rather than adopting a proactive stance toward disease prevention**.**
Limited Patient Engagement: Low levels of patient engagement are a common drawback, with existing systems often failing to actively involve individuals in their healthcare management.

Resource-Intensive Models: The resource-intensive nature of some predictive models hinders scalability and accessibility, making them impractical for deployment in various healthcare settings.
Inflexibility to Parameter Changes: The inflexibility of existing systems to adapt to changes in patient health parameters over time results in static recommendations that may not align with evolving health conditions.
Inadequate Lifestyle Factor Consideration: Many systems inadequately incorporate lifestyle factors into their predictive models, overlooking essential elements that significantly contribute to an individual's overall health

3. REQUIREMENT ANALYSIS:

We will see all the requirements used for this project.

3.1 FUNCTIONAL REQUIREMENTS:

Registration

- System must allow user to create new account using email address and password.
- System must allow user to login with there email address and password.
- System must allow user to reset password by clicking on “I forget my password” and receiving link to there verified email address.

Patient Details

- System must allow user to enter their personal details.
- System must allow patient to enter their medical history.

Display Parameters

- System must display disease parameters to be asked to patient.
- System must allow patient to answer parameters display on the screen.
- System must save the patient response into the database.

Predictions

- System must perform predictions based on the answered parameters.
- System must display the results of prediction to the patient.
- System must allow user to download their report.

Recommendations

- System must recommend doctor based on predicted disease.
- System must allow show patient report to the doctor.

3.2 NON- FUNCTIONAL REQUIREMENTS:

Non-functional requirements for Parameter-Based Disease Prediction and Recommendation encompass critical characteristics that define the overall behavior, performance, and quality attributes of the system. These requirements are integral to ensuring the effectiveness, reliability, and user experience of the predictive healthcare solution. Here are key categories of non-functional requirements for such a system:

Response Time: Specify the maximum allowable time for the system to provide predictions and recommendations, ensuring timely responses.
Throughput: Define the system's capacity to handle a certain number of prediction and recommendation requests concurrently.
User Interface Design: Specify guidelines for a user-friendly interface, ensuring ease of navigation and understanding for healthcare professionals and end-users.

Accessibility: Ensure the system is accessible to users with disabilities, complying with relevant accessibility standards.
Authentication: Define secure user authentication mechanisms to prevent unauthorized access to sensitive health data.
Data Encryption: Mandate the use of encryption for data transmission and storage to protect patient information.
Availability: Specify the acceptable system uptime to ensure that predictions and recommendations are consistently accessible.
Fault Tolerance: Define measures for the system to recover gracefully from potential failures without compromising its overall performance.
Code Maintainability: Define coding standards and practices to facilitate easy maintenance and future enhancements.
Documentation: Mandate comprehensive documentation, including system architecture, code comments, and user manuals, to aid in system maintenance.
Platform Compatibility: Specify the platforms and environments on which the system should be compatible, ensuring flexibility and adaptability.
Horizontal Scalability: Define how the system can scale horizontally by adding more instances or nodes to handle increased demand.
Vertical Scalability: Specify how the system can scale vertically by enhancing the capacity of existing components.

3.3 DESIGN REQUIREMENTS:

Figure 3.1: Gantt chart

SOFTWARE REQUIREMENTS:

- **Server:** Linux-based OS (Ubuntu Server, CentOS) or Windows Server OS (Windows Server 2016 or later).

Backend Development:

- Android Studio
- Firebase
- Java

HARDWARE REQUIREMENTS:

- **Processor:** Quad-core processor (Intel Core i5 or equivalent)
- **RAM:** 8 GB or higher
- **Storage:** SSD with at least 256 GB storage space

- **Network Interface:** Gigabit Ethernet

4. DESIGN:

4.1 Steps Involved in Supervised Learning:

- First Determine the type of training dataset
- Collect/Gather the labeled training data.
- Split the training dataset into training dataset, test dataset, and validation dataset.
- Determine the input features of the training dataset, which should have enough knowledge so that the model can accurately predict the output.
- Determine the suitable algorithm for the model, such as support vector machine, decision tree, etc.
- Execute the algorithm on the training dataset. Sometimes we need validation
- sets as the control parameters, which are the subset of training datasets.
- Evaluate the accuracy of the model by providing the test set.

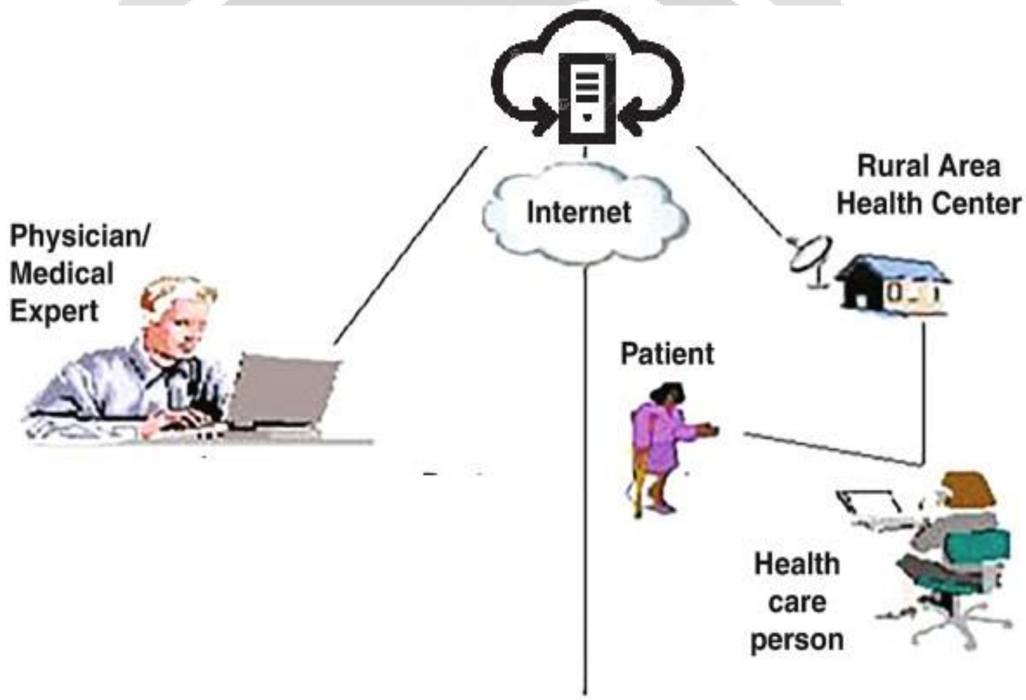

Fig 4.1 : Block Diagram of Parameter based Disease Prediction and Hospital Recommendation.

Regression algorithms are used if there is a relationship between the input variable and the output variable. It is used for the prediction of continuous variables, such as Weather forecasting, Market Trends, etc.

1. Linear Regression
2. Regression Trees
3. Non-Linear Regression
4. Bayesian Linear Regression
5. Polynomial Regression

4.2 FLOWCHART:

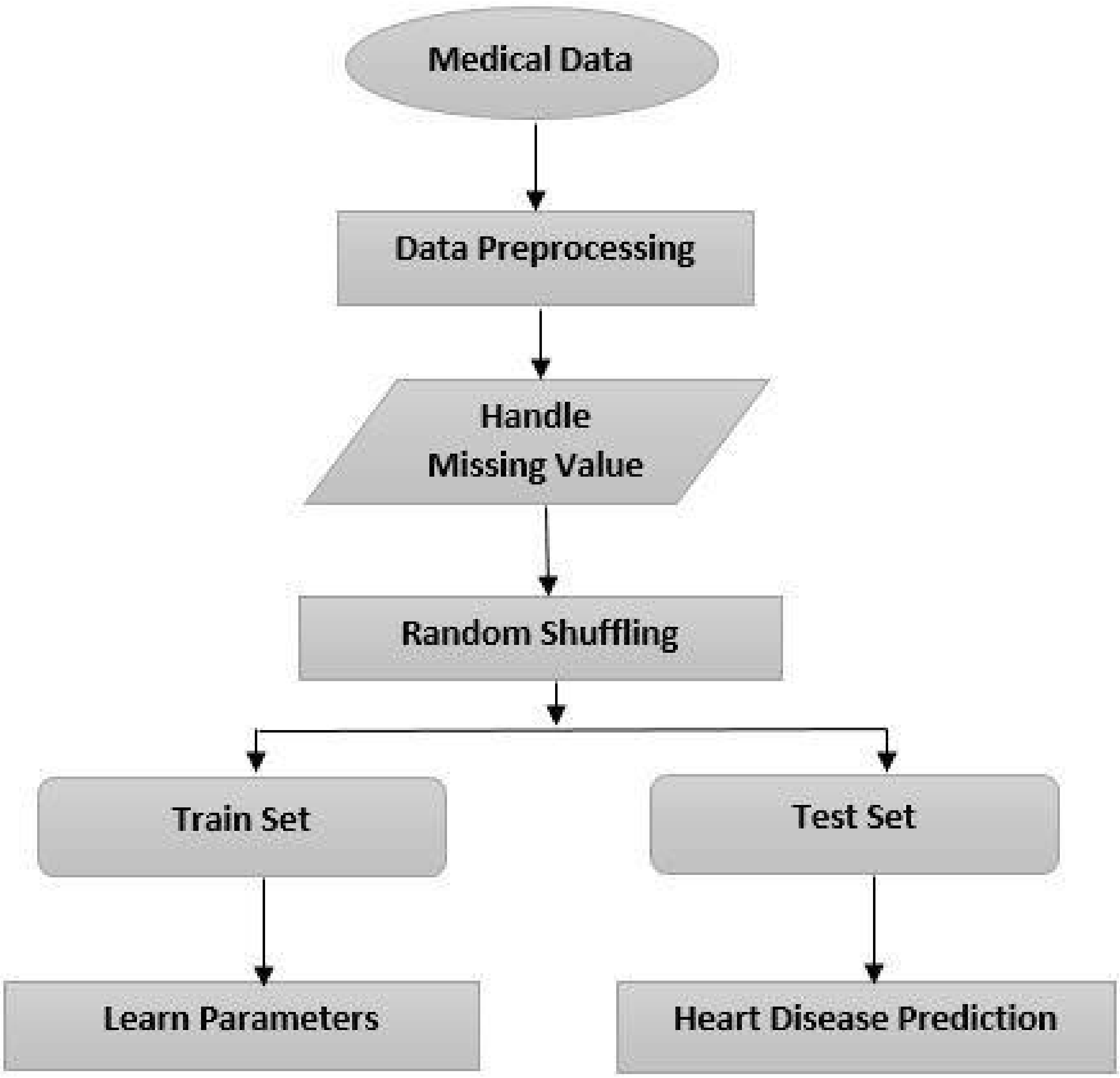

Fig 4.2 Flowchart of the Project

4.3 ALGORITHM :

ALGORITHM USED

There are various algorithms in Machine learning, choosing the best algorithm for the given dataset and problem is the main point to remember while creating a machine learning model. Below are the two reasons for using the Decision tree:

- Decision Trees usually mimic human thinking ability while making a decision, so it is easy to understand.
- The logic behind the decision tree can be easily understood because it shows a Tree- like structure.

Decision Tree Terminologies:

- Root node is from where the decision tree starts. It represents the entire dataset, which further gets divided into two or more homogeneous sets.
- Leaf nodes are the final output node, and the tree cannot be segregated further after getting a leaf node.
- Splitting is the process of dividing the decision node/root node into sub-nodes according to the given conditions.
- A tree formed by splitting the tree known as branch tree.
- Pruning is the process of removing the unwanted branches from the tree.

5. CONCLUSION:

Parameter-based disease prediction and recommendation systems offer a promising avenue for enhancing healthcare outcomes. By leveraging various parameters such as demographic data, medical history, lifestyle factors, and genetic predispositions, these systems can accurately forecast the likelihood of individuals developing specific diseases. Additionally, by integrating advanced algorithms and machine learning techniques, these systems can provide personalized recommendations for preventive measures, early detection strategies, and tailored treatment plans.

Furthermore, the implementation of parameter-based disease prediction and recommendation systems has the potential to revolutionize healthcare delivery by enabling proactive interventions, reducing healthcare costs, and improving overall patient outcomes. However, challenges such as data privacy concerns, algorithm bias, and the need for continuous validation and refinement must be addressed to ensure the effectiveness and ethical integrity of these systems.

In summary, parameter-based disease prediction and recommendation systems represent a valuable tool in the arsenal of modern healthcare, offering the promise of early detection, personalized intervention, and improved health outcomes for individuals and communities alike. Through continued research, development, and collaboration between healthcare professionals, researchers, and technology experts, these systems can play a pivotal role in advancing preventative medicine and transforming the future of healthcare.

6. FUTURE SCOPE:

Our prototype is the first user should put patient details and laboratory reports. The parameters are based on the report of the patient. the data set of results that will be gained by applying the algorithm to that dataset. After applying the ML dataset will get to know whether the user has diabetes or not. The UML diagrams for the project are stated in which class diagram represents which are the required objects for the particular application that helps for constructing executable code of the software application.

The data flow diagram provides information regarding the flow of the data. Sequence diagram simply shows how objects communicate with each other which helps to understand how the system is working internally. Activity diagrams help to know the workflow from one activity to another, use case diagram provides information related to the functionality of the system.

Continued advancements in technology, including the integration of artificial intelligence (AI) and machine learning algorithms, will further improve the accuracy and precision of disease prediction models. By incorporating a wider range of parameters, such as genomic data, environmental factors, and real-time health monitoring data from wearable devices, these systems can provide more comprehensive and personalized predictions.

7. REFERENCES

[1] Yuji Mizuno; Teppei Baba; Yoshito Tanaka; Fujio Kurokawa; Masaharu Tanaka; Ilhami Colak; Nobumasa

Matsu “A New Load Prediction Method and Management of Distributed Power System in Island Mode of a Large Hospital”/2018

[2] B. Hu and M. Ester, “Social topic modeling for point-of-interest recommendation in location-based social networks,” in 2014 IEEE International Conference on Data Mining (ICDM), pp. 845–850, IEEE, 2014.

[3] Mohammad Reza Khoie, Tannaz Sattari Tabrizi,Elham Sahebkar Khorasani, Shahram Rahimi and Nina Marhamati “A Hospital Recommendation System Based on Patient Satisfaction Survey/2017”

[4] E. Velasco, T. Agheneza, K. Denecke, G. Kirchner, and T. Eckmanns, “Social media and internet-based data in global systems for public health surveillance: a systematic review,” Milbank Quarterly, vol. 92, no. 1, pp. 7–33, 2014.

[5] Z. Yao, Y. Fu, B. Liu, Y. Liu, and H. Xiong, “Poi recommendation: A temporal matching between poi popularity and user regularity,” in 2016 IEEE 16th International Conference on Data Mining (ICDM), pp. 549– 558, IEEE, 2016

[6] Min Chen, Yixue Hao, Kai Hwang, Lu Wang. “Disease prediction by Machine Learning over big data from Healthcare Communities” IEEE-2015

[7] Bo Hu and Martin Ester “Social topic modelling for point of interest recommendation in location based social networks”/2014 IEEE

[8] Yuji Mizuno; Teppei Baba; Yoshito Tanaka; Fujio Kurokawa; Masaharu Tanaka; Ilhami Colak; Nobumasa Matsu “A New Load Prediction Method and Management of Distributed Power System in Island Mode of a Large Hospital”/2018