International Journal of Scientific Research & Engineering Trends
Volume 8, Issue 6, Nov-Dec-2022, ISSN (Online): 2395-566X

Machine Learning Algorithm Based Health Care Monitoring System

M.Tech. Scholar Sonal Shrivastava , Prof. Rajesh Kumar Boghey
Department Of Computer Science Engineering
Technocrats Institute of Technology, Bhopal

Abstract- The regular measurement of vital signs enables early diagnosis and warning of developing problems. Furthermore, it allows closer monitoring of the effects of medication and lifestyle, making more personalized treatment plans possible. The system contains a patient loop interacting directly with the patient to support the daily treatment. It shows the health development, including treatment adherence and effectiveness. An educated and motivated patient can improve his/her treatment compliance and health. The system also contains a professional loop involving medical professionals (e.g. alerting to revisit the care plan). The patient loop is securely connected with hospital information systems, to ensure optimal personalized care. Big data analytics provides services to various organizations, especially in the healthcare field. The medical field contains a large amount of data and is well suited for data analysis. Medical big data is mainly used for clinical data, and chronic disease monitoring and health monitoring are mainly used to detect changes in patients' health. First, you must process the data to remove unnecessary data and provide effective prediction results. The second is the data analysis process - this is the process of cleaning, transforming and modeling data for the purpose of discovering useful information. In this process, we propose privacy protection to keep patient information secure. And support vector Machine learning algorithms are mainly used to predict diseases and provide more efficient prediction results. Finally, our system will predict the disease based on the patient's symptoms and show the treatment to the patient.

Index Terms—medical big data, chronic diseases, health monitoring, data analysis, data visualization

I. INTRODUCTION

With the continuous development of medical information, the expansion of medical data accelerates and the coverage area increases. The advent of the big data era creates new opportunities for health and medical domains. As it's known, big data contain 4Vs, which are volume, velocity, variety or veracity. Medical big data has all four features. For volume, a commonly cited statistic from EMC said that 4.4 zettabytes of data existed globally in 2013. That number is predicted as 44 zettabytes by 2020 as it more than doubles each year. For velocity, the health monitoring data is generating every second. For variety, the medical domain contains the many potential big data sources, for instance, the digital medical record, MRI, CT, health monitoring data, genome data. For veracity, a medical data might be incomplete, biased, or even filled with noise. And users can't utilize insights.

To analyze the data, data pre-processing, data modelling, data visualization and security are needed. Ambiguous information, repetition, noise and ultra-high dimensions influence the medical data. Therefore it is necessary to pre-process the data. Medical big data pre-processing, integration involves data ETL (Extract, Transform, Load), multi data source integration, and unified. Typical algorithms and tools based on existing big data platform are mainly used to make data analysis more convenient and effective. Also, medical big data visualization including treemap, circle-packing, sunburst etc. is the most effective tool when faced with complex medical data and growing medical needs. The type of big data visualization is information visualization, interaction techniques and architectures, modeling techniques, multiresolution methods, visualization algorithms and techniques and volume visualization. Moreover, there are some customized analytics based on visualization which has been listed in this paper. Medical big data is mainly used in clinical data, monitoring and early warning of chronic diseases, daily activities and physical characteristics index detection and collection nowadays Big data can play an important role in determining causal relationships in patient symptoms, predicting the risks of disease occurrence or recurrence, and improving the quality of primary care. Using systematic reviews to create the classification results for the healthcare big data utilizing[1]

showed that the data was split into four broad categories: management and childbirth, clinical decision support, consumer behaviour and support services. Two well-known examples of using big data in health are Google Flu Trends and HealthMap. By analyzing the medical status of Americans[2], it has been found that technological advances make it easier to collect and

analyze patient record information from multiple sources, as individual patient data may come from different payers, hospitals, laboratories and doctor's offices. By digitizing, consolidating and effectively utilizing big data, significant benefits could be realized from the organization of single-physician offices and multiple health-care facilities to large-scale hospital networks and the care providers responsible for care facilities[3]. A survey by IBM found that big data analytics will help healthcare in many ways, such as evidence-based medicine, genomic analysis, pre-trial fraud analysis, device/remote monitoring, and patient file analysis. Big data helps reduce clinical procedures, research and development, public health waste and inefficiencies. Data collected by sensors and smartphones identify the individual risk factorsB. Chronic disease monitoring and early warning Diabetic Mellitus is a chronic metabolic disease that requires active and sustained participation by diabetes patients, its careers and physicians for management and good control[4]. Through an 11-year follow-up of Southall diabetes among South Asians and Europeans[5], they found that more than 20% of middle-aged and older South Asians worldwide had diabetes mellitus, whereas South Asian adult diabetics compared to Europeans had a marked increase in their propensity for cardiovascular disease, especially among young people. With big data processing technology, all the technologies related to chronic diseases are extracted, converted, loaded and stored in HFDS.

II. LITERATURE SURVEY

This paper analyzes the concept of classification analysis and the properties of decision tree, and gives the implementation process of ID3 algorithm. The medical examination data of medical examination information management system of Xi'an Shiyou University Hospital from 2013 to 2020 are selected as the training sample set and discretized, and a direct data model suitable for classification analysis is designed. ID3 algorithm is employed to classify and analyze the sampled data set, and the classification rules are extracted. Using the prediction conclusions of these classification rules, physical examination doctors can quickly and scientifically predict the possibility of chronic diseases of each university teacher. It can provide information technology support for the screening and prediction of chronic diseases and personalized intervention of chronic diseases. Zhang (Quancheng et.al.2021)[1]

[6] big data often has high values in volume, velocity, variety, variability, value, complexity, and sparseness. Big data has the potential of applications in healthcare which include disease surveillance, epidemic control, clinical decision support, population health management, etc. [7] Big Data in healthcare can provide significant benefits such as detecting diseases at an early stage. The inclusion of Big Data analytics in smart healthcare systems brings innovative electronic and mobile health (e/m-health) that increase efficiency and save medical costs Predictive analytics can be used in predicting pharmaceutical outcomes, identifying patients who benefit the most from pharmacist interventions, providing pharmacists with a better understanding of the risks of specific medication-related problems, and delivering interventions tailored to patients' needs [8]Precision medicine deals with data ranging from collection and management (such as data storage, sharing, and privacy) to analytics (such as data integration, data mining, and visualization). Complex biomedical data with a huge volume are becoming available due to advances in biotechnologies. Big Data analytics is required to use these heterogeneous data and it covers application areas such as health informatics, sensor informatics, bioinformatics, imaging informatics, etc. [9] Veracity is crucial for Big Data analytics. Personal health records (PHRs) may contain abbreviations, typographical errors, and cryptic notes. Ambulatory measurements are possibly completed under uncontrolled and less reliable environments compared with clinical data which is collected by trained practitioners in a clinical setting. Using spontaneous unmanaged data from social media may result in inaccurate predictions. In addition, data sources are sometimes biased [10] 'Noise' data is a massive problem especially when it grows fast. Databases with various degrees of completeness and quality lead to heterogeneous results, which increase the possibility of false discoveries and 'biased fact-finding excursions'. Low data quality and biases due to the absence of randomization are two major problems. Efforts in increasing the value of big data are often made through linking different databases and analyzing all existing and related data [11-15]

III . PROPOSED SYSTEM

Big data can play an important role in determining the causal relationship between patients' symptoms, predicting the risk of illness or recurrence, and improving the quality of primary care. Big data helps reduce clinical procedures, research and development, public health waste and inefficiency. To overcome the security issues in the existing system and provide more accurate results, introduced this system which will provide better prediction results. In this process, we propose a privacy protection for securely maintaining the patient's Information. The patient details include BMI, Glucose, Diabetes_ Pedigree_ Function, Age, Blood Pressure, Skin Thickness and Insulin Level. After entering their details it will find whether the patient having the diabetes or not. Suppose, if the patient having the diabetes it will check the type and display the treatment to the user.

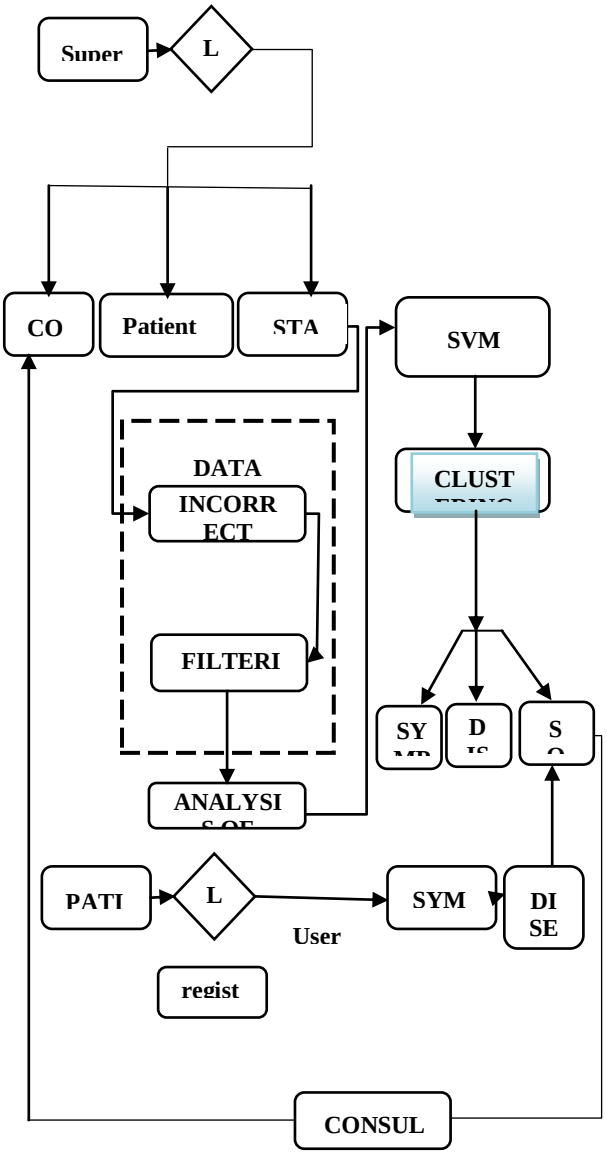


Fig 1 flow diagram.

Admin and user login-The entire process is maintained by the administrator. The administrator can log in with his username and password (if valid) and it will be redirected to another form. The administrator manages all patient information, provides consultation to patients, and loads data sets for prediction. If it is a new user, you can log in with their username and password and the user must register before logging in. Users must enter their detailed information for disease prediction. After entering the details, it shows the user the type and processing method.

Data Selection and Input-Data extraction technology is part of the business information and uses specific algorithms to detect hidden patterns and relationships in large data sets. In this module, a data set is loaded to predict the patient's health.

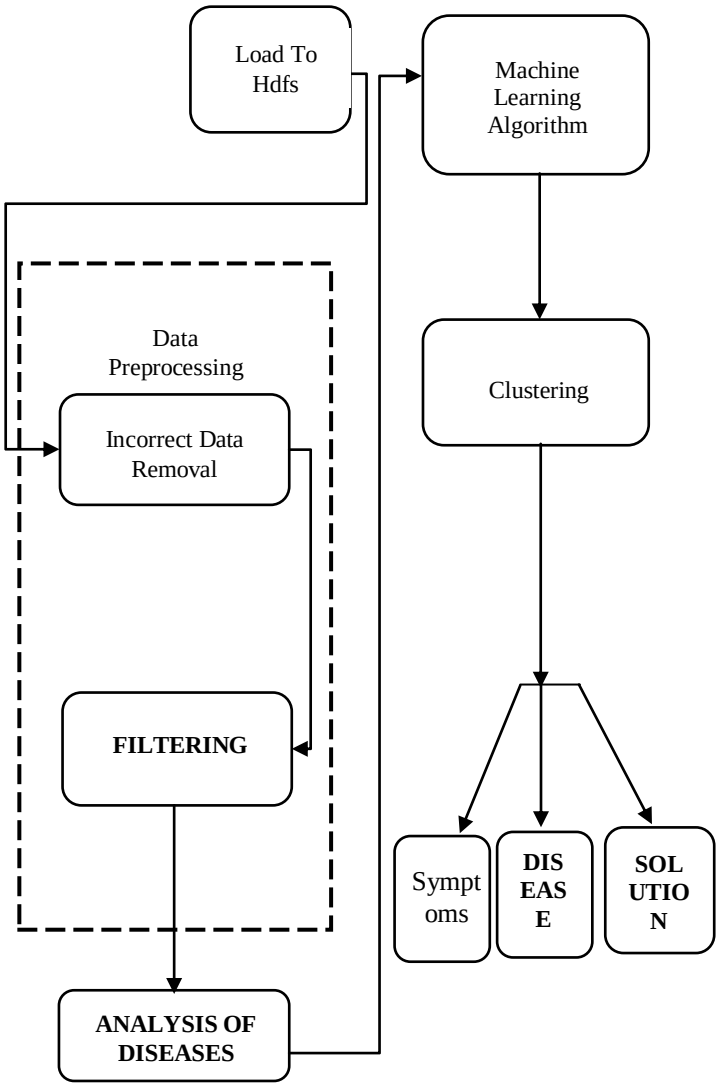
Data Pre-Processing-Initial step of this process is data pre-processing. The data is pre-processed to remove the unwanted data

Analysis of Disease and Classification

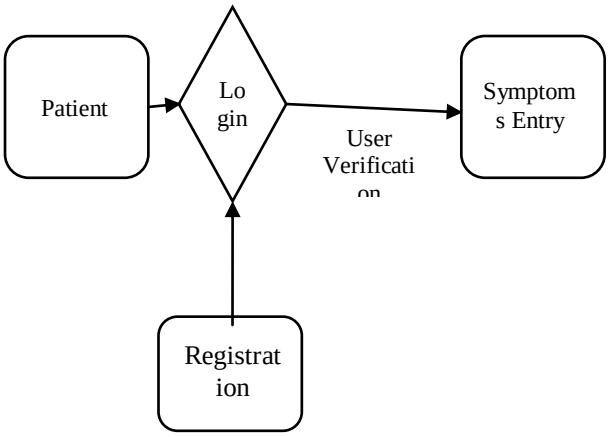
Support vector machine classification algorithm is mainly used for classification and prediction. It analyzes and classifies the data based on whether or not the patient has diabetes.

Patient indtræden

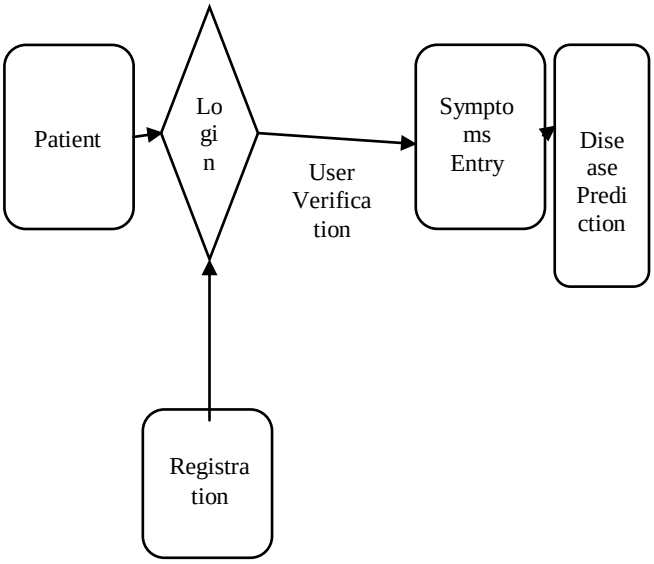
Patient details include BMI, glucose, Diabetes_ Pedigree _Function, age, blood pressure, skin thickness and insulin level. After entering the information, it will find out if the patient



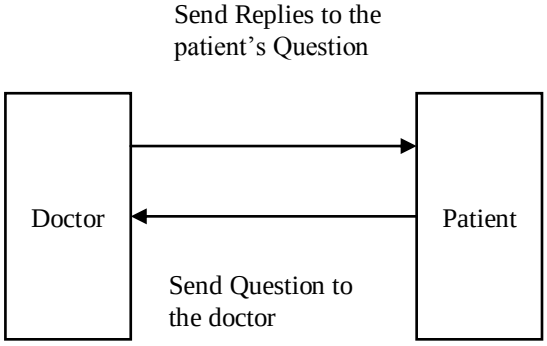
Having the diabetes or not. Suppose, if the patient having the diabetes it will check the type and display the treatment to the user.



Disease Prediction In this module, the patient disease is predicted based on their information. In this project the support vector algorithm is used for predicting for patient disease.



Doctor Consultant
After, finding the type and treatment of disease, the patient will communicate with the doctor for getting more information about their problems.



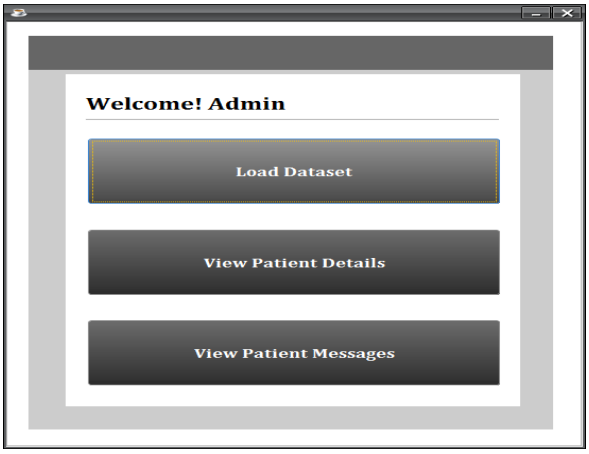


Fig. 3 Admin Login.

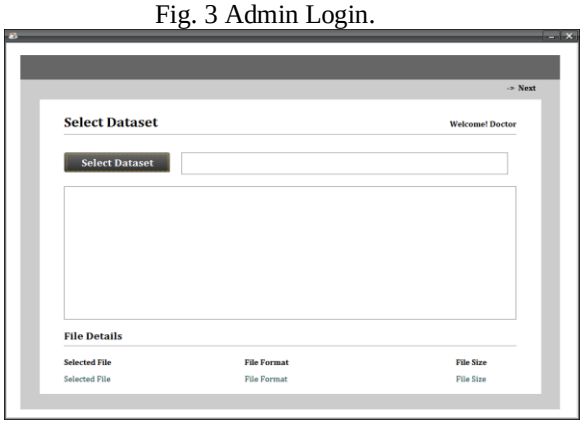


Fig. 4. Load Dataset

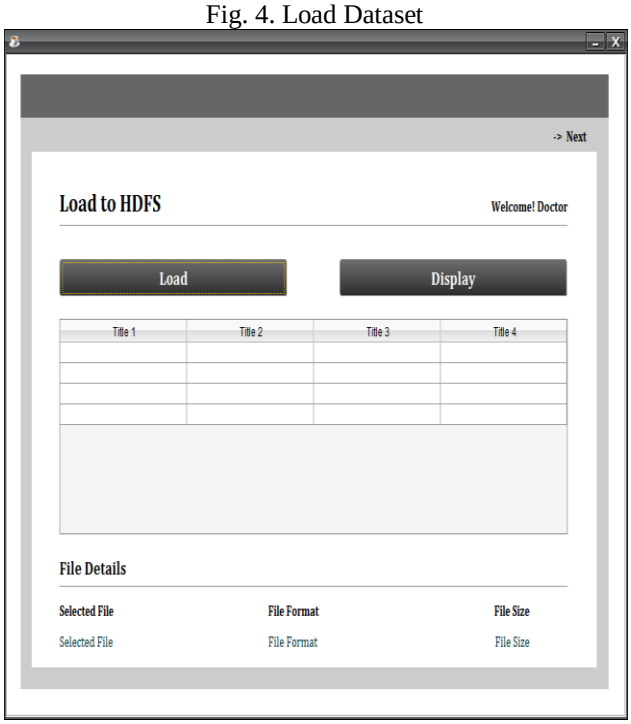


Fig 5 Load to HDFS

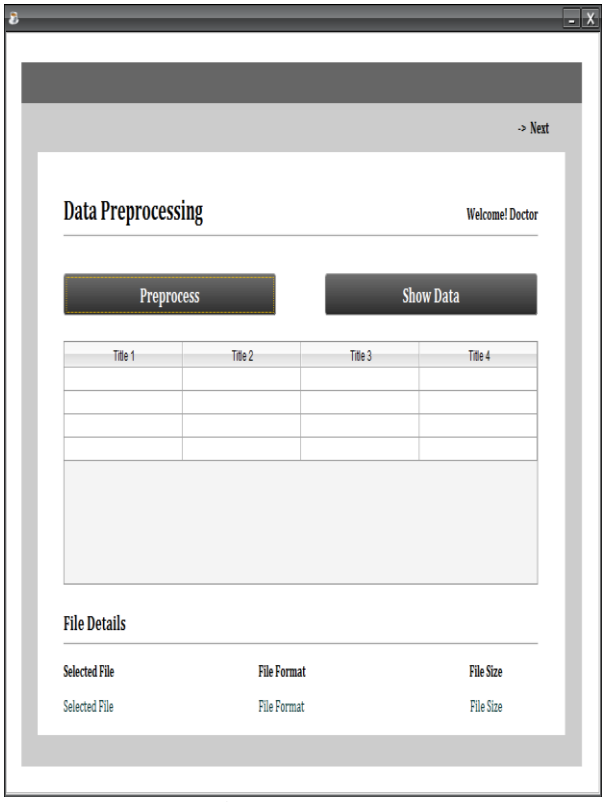


Fig 6 Data Pre-processing

Fig 7 Data Pre-processing

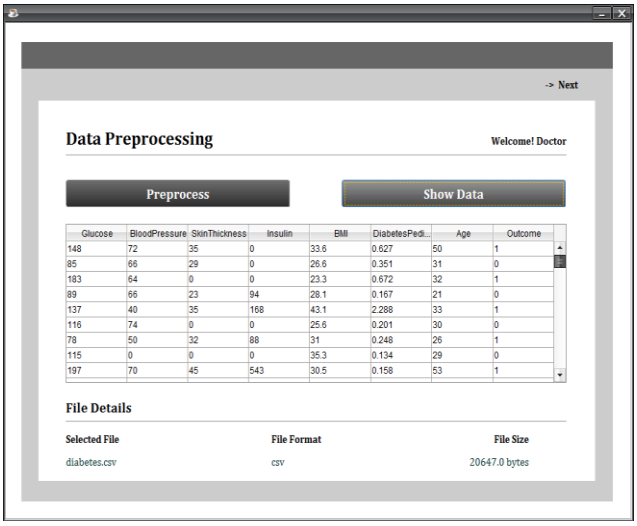


Fig 8 Group by Disease

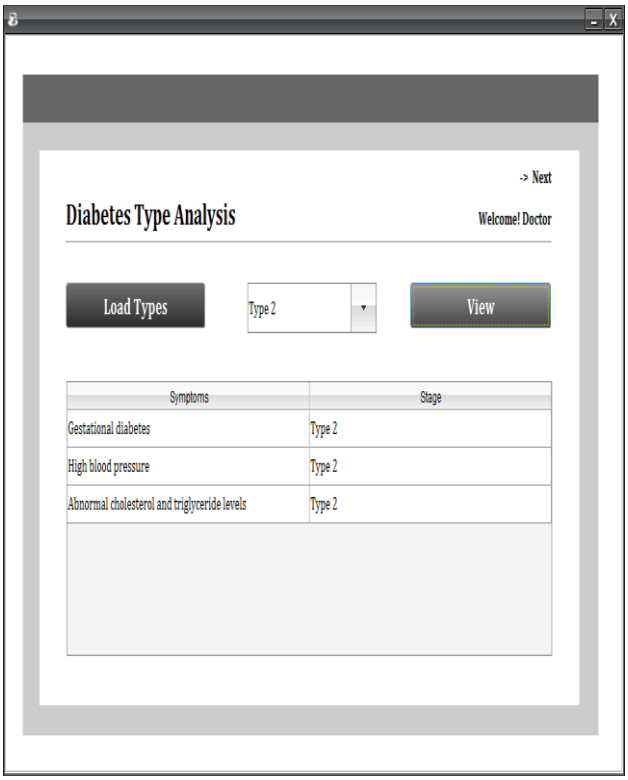


Fig 9 Diabetes Type Analysis

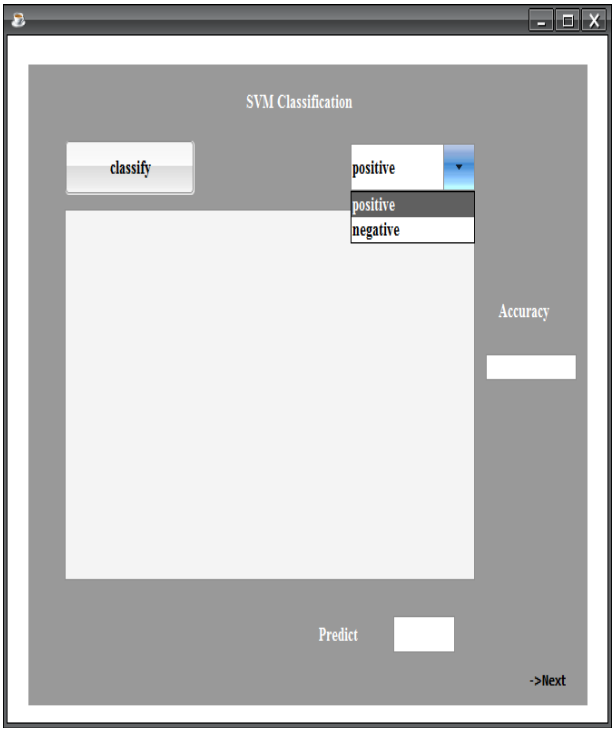


Fig.10 View Symptoms and Type

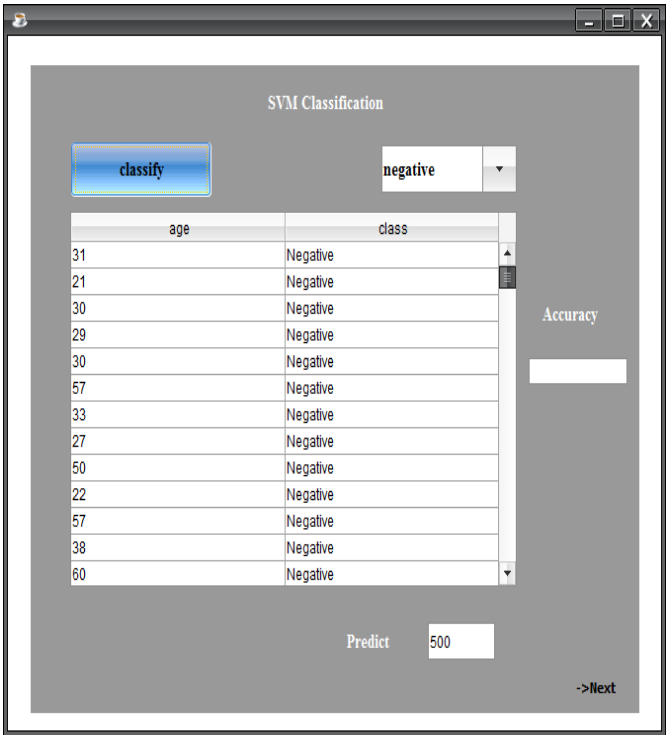


Fig 12 View Classified Data

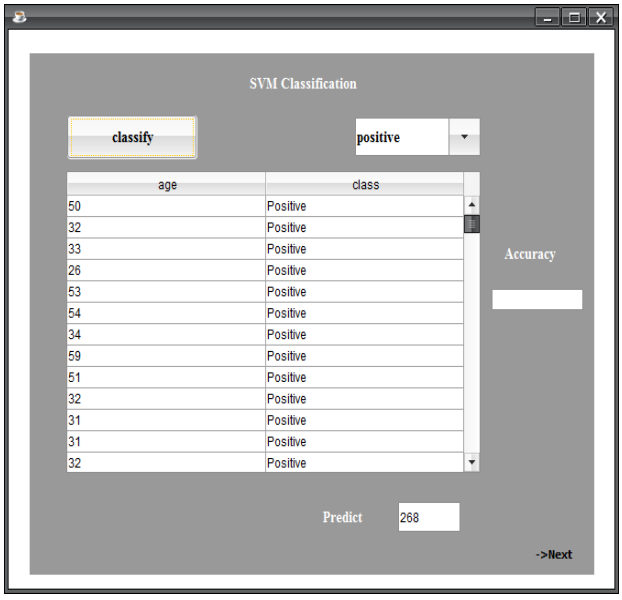


Fig 11 SVM Classification

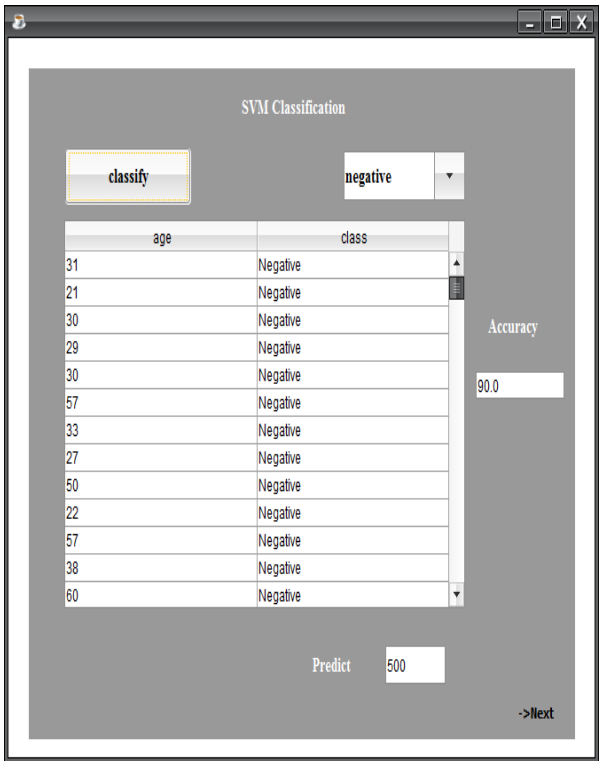


Fig 13 View Classified Data

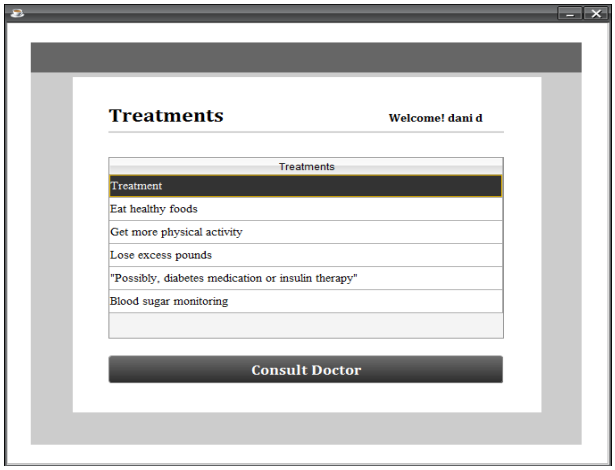


Fig 14 Treatment

Fig 15 Patient Page

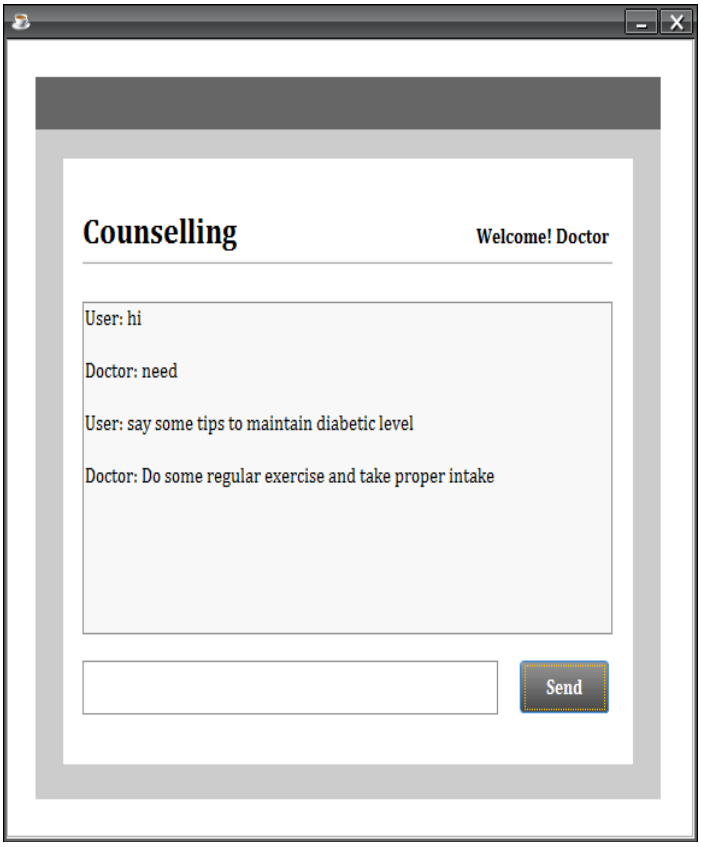


Fig. 16 Doctor Counselling Page

V. CONCLUSIONS

This study introduces the history and methods of developing medical big data for diseases and health monitoring. Using big data, based on analysis of similar patients and the response to these methods, more personalized medicine using specific data will be more. The real problem is not that we want a lot of data. This is the importance of the data you want to process. Today, much of the cost and time spent on developing drugs is due to failed formulations. By enabling researchers to identify compounds that are more likely to succeed, big data can help reduce the cost of new drugs and shorten time to market. By integrating knowledge learned from medical data at an early stage of development, researchers will now also be able to customize drugs to fit pooled patient data.At present, privacy information is the major obstacle to the adoption of big data in the healthcare field. Another reason is the lack of powerful analytical solutions sufficient to collect large amounts of unstructured health data, perform complex analyzes quickly, and trigger meaningful solutions, such as gathering all data from ICU monitors not yet stored, so it must be saved. In the cloud, important medical patterns that have not yet been discovered decipher and trigger medical actions, not just alarms.By summarizing the current state of big data applications in the healthcare field, this research paper

examines the current challenges facing governments and healthcare stakeholders. All big data projects in leading countries and healthcare have similar overall goals, such as the provision of practical and equal public services and better citizen healthcare. However, each government or health interest has its own priorities, opportunities and threats depending on the unique environment of the country. Second, a top-down approach to medical data that crosses departmental boundaries is needed to effectively manage and integrate big data. Third, real-time analysis of moving big data needs to be done while protecting privacy and security. Therefore, the government and healthcare stakeholders should explore new areas of technology, such as cloud computing, advanced analytics, security technologies and regulations. Finally, the limitation of the study is that the practical use of big data for the examination of health issues has not been fully demonstrated due to lack of practice. For future research, practitioners and researchers should carefully review and accumulate information on the practical use of big data to determine the best way to use big data in health issues.In the future, we can implement this process on the Internet of Things. Eg. The system automatically detects the patient's health and provides a better solution.

REFERENCES

1. Zhang Quancheng* He Jingbin Research on Data Mining of Physical Examination for Risk Factors of Chronic Diseases Based on Classification Decision Tree Research on Data Mining of Physical Examination for Risk Factors of Chronic Diseases Based on Classification Decision Tree 978-1-6654-0413-6/21/$31.00 ©2021 IEEE
2. Christian Bizer, Peter Boncz, Michael L. Brodie, and OrriErling. The meaningful use of big data: four perspectivesfour challenges. ACM SIGMOD Record, 40(4):56–60, 2012.
3. Nitesh V. Chawla and Darcy A. Davis. Bringing big data to personalized healthcare: a patient-centered framework. Journal of general internal medicine, 28(3):660–665, 2013.
4. Yin Zhang, MeikangQiu, Chun-Wei Tsai, Mohammad Mehedi Hassan, and AtifAlamri. Health-CPS: Healthcare cyber-physical system assisted by cloud and big data. IEEE Systems Journal, 11(1):88–95, 2017.
5. Oliver M. Duschka, Michael R. Genesereth, and Alon Y. Levy. Recursive query plans for data integration. The Journal of Logic Programming, 43(1):49–73, 2000.
6. Ralph Kimball and Joe Caserta. The Data Warehouse? ETL Toolkit: Practical Techniques for Extracting, Cleaning, Conforming, and Delivering Data. John Wiley & Sons, 2011.
7. Panos Vassiliadis. A survey of Extracttrans formLoad technology. International Journal of Data Warehousing and Mining (IJDWM), 5(3):1–27, 2009.
8. Qin Yao, Yu Tian, Peng-Fei Li, Li-Li Tian, Yang-Ming Qian, and Jing-Song Li. Design and development of a medical big data processing system based on hadoop. Journal of medical systems, 39(3):23, 2015.
9. Min Chen, YixueHao, Kai Hwang, Lu Wang, and Lin Wang. Disease Prediction by Machine Learning over Big Data from Healthcare Communities. IEEE Access, 2017.
10. Ivo D. Dinov. Methodological challenges and analytic opportunities for modeling and interpreting Big Healthcare Data. Gigascience, 5(1):12, 2016.
11. Ruogu Fang, Samira Pouyanfar, Yimin Yang, Shu-Ching Chen, and S. S. Iyengar. Computational health informatics in the big data age: a survey. ACM Computing Surveys (CSUR), 49(1):12, 2016.
12. M. Shamim Hossain and Ghulam Muhammad. Healthcare Big Data Voice Pathology Assessment Framework. IEEE Access, 4:7806–7815, 2016.
13. Jigna Ashish Patel and Priyanka Sharma. Big data for better health planning. In Advances in Engineering and Technology Research (ICAETR), 2014 International Conference on, pages 1–5. IEEE, 2014.
14. AislingODriscoll, JurateDaugelaite, and Roy D. Sleator. Big data, Hadoop and cloud computing in genomics. Journal of biomedical informatics, 46(5):774–781, 2013.
15. Farid, D. M., Nowe, A., & Manderick, B. (2016, December). A feature grouping method for ensemble clustering of high-dimensional genomic big data. In Future Technologies Conference (FTC) (pp. 260-268). IEEE