Smart Health Prediction System with Data Mining

Ramavarapu Sai Rohith Bharadwaj, Muthyala Sathvik Reddy, 3Manoj Dyavathi, Maniteja Rachakonda, Madhi Reddy Keerthi Reddy

The digital technology era demands the world to provide an excellent health system, to ensure the citizen and community to be alive and healthy. Purpose: This study proposes the application of data mining algorithm for health prediction that can eventually shape a suitable health prediction system for patients. Although health care is available to everyone in the world, there is still no healthcare system that is completely reliable and accurate to carefully diagnose a patient on their current health issues. Even though some hospitals are well equipped to provide the best healthcare services to its citizens, some of the hospitals are still lacking in certain qualities. Consequently, patients are doubtful and uncertain when it comes to picking which hospital suits them. Problem: Numerous issues are faced by patients pertinent to hospitals such as being unable to provide medical services, insufficient number of qualified medical staffs, poor communication between doctors and patients, and unorganized health records and data. Eventually, these issues impede the opportunity for hospitals to handle both their management and their duties steadily to maintain the health of every citizen and community. Conclusion: Patients need treatment and diagnosis that are accurate and precise for them to be able to recover back for their proper health and medical staffs are required to be well-equipped in their clinical knowledge and communication skills to carefully assess their patients to ensure good health. Therefore, application of data mining in health prediction is considered in this paper as the best practice to facilitate better healthcare system.

Keywords: Machine Learning algorithms, health prediction, data mining.

1. Introduction

The healthcare industry collects huge amounts of healthcare data which, unfortunately, are not "mined" to discover hidden information for effective decision making.

Data mining has been a current trend for attaining diagnostic results. Huge amount of unmined data is collected by the healthcare industry in order to discover hidden information for effective diagnosis and decision making.

Data mining is the process of extracting hidden information from massive dataset, categorizing valid and unique patterns in data.

The health management system is an end user support and online consultation project. The proposed system allows users to get guidance on their health issues through an intelligent

health care online system.

Health prediction systems leverage the wealth of information embedded in healthcare databases, encompassing patient records, diagnostic data, lifestyle details, and genetic information.

By employing advanced data mining algorithms, these systems aim to unravel intricate patterns, correlations, and predictive trends hidden within these diverse datasets.

The overarching objective is to furnish accurate and timely predictions regarding individual health conditions, potential risks, and prognosis.

The digital technology era demands the world to provide an excellent health system, to ensure the citizen and community to be alive and healthy.

The current healthcare system faces several challenges that hinder its ability to deliver optimal care and fully leverage advancements in technology.

The one-size-fits-all approach may lead to suboptimal outcomes for certain individuals, as their unique needs and responses to treatment are not adequately considered.

There are currently a lot of health institutions that have been developed such as hospitals and medical centres which are crucial to maintain and improve the health of the community around us. It is a prime establishment of giving proper health care especially for every one of us who have ever lived.

For every illness and diseases that people may face today and sometime in the future, it is because of these medical institutions and all the doctors who worked at these places that have made our lives physically better and healthy.

Although hospitals now are well-equipped with their staffs working, there are still known issues that persist that cause the staffs to make the poor clinical decision that affects a patient's health such as the lack of qualified doctors, unorganized health information and poor communications between doctors and patients.

The primary objective of the smart health prediction system using data mining is to revolutionize healthcare delivery by leveraging advanced data analytics to anticipate, assess, and manage health-related outcomes.

The objective aligns with the broader vision of transforming healthcare from a reactive, one-size-fits-all model to a proactive, personalized, and resource-efficient system. By harnessing the power of data mining, the system seeks to enhance patient outcomes, promote preventive care, and contribute to the overall improvement of healthcare practices.

The system develops robust algorithms to identify potential health issues at their earliest stages, allowing for timely and effective interventions.

In addition the project implements predictive analytics to forecast and manage chronic diseases, improving the quality of life for affected individuals.

It also optimizes healthcare resource allocation by predicting patient needs, reducing unnecessary tests, and improving the efficiency of healthcare delivery.

2. Literature Survey

In recent years, the healthcare industry has witnessed a rapid increase in the volume and variety of data generated from various sources such as electronic health records (EHRs), medical imaging, wearable devices, and genomic data. This surge in healthcare data has provided an opportunity for the application of data mining techniques to extract valuable insights for improving patient care, disease diagnosis, treatment optimization, and healthcare management. This literature survey explores several studies and research papers that have contributed to the understanding and advancement of data mining in healthcare.

Saqib Qamar et al. provide a comprehensive review of data mining techniques in healthcare in their paper titled "Techniques of Data Mining In Healthcare: A Review" [11]. The authors discuss various methodologies, including classification, clustering, association rule mining, and anomaly detection, highlighting their applications in healthcare data analysis. They emphasize the importance of data mining in extracting meaningful patterns and trends from large healthcare datasets to support decision-making processes in clinical practice, disease management, and public health surveillance.

Basma Boukenze et al. focus on predicting chronic kidney failure disease using data mining techniques in their study [12]. By leveraging data mining algorithms, the authors analyze medical data to identify patterns and risk factors associated with chronic kidney disease. Their research aims to develop predictive models that can assist healthcare professionals in early detection and intervention, thereby improving patient outcomes and reducing healthcare costs associated with kidney disease management.

Casey C. Bennett et al. explore the integration of data mining with electronic health records (EHRs) in their paper titled "Data mining and electronic health records: Selecting optimal clinical treatments in practice" [13]. The authors highlight the potential of data mining techniques to analyze large volumes of EHR data and extract actionable insights for selecting optimal clinical treatments. By leveraging patient data and clinical outcomes, they demonstrate how data mining can support evidence-based decision-making in healthcare delivery and improve the quality and efficiency of patient care.

P Rajeswari et al. investigate the analysis of liver disorders using data mining algorithms in their research work [14]. Through the application of data mining techniques to medical data related to liver disorders, the authors aim to identify patterns and factors contributing to liver diseases. Their study contributes to early diagnosis and intervention strategies for liver disorders, thereby enhancing patient care and management of liver-related health issues.

Latha Parthiban et al. propose an intelligent heart disease prediction system using CANFIS (Cerebellar Adaptive Neuro-Fuzzy Inference System) and genetic algorithms in their paper [15]. By integrating computational intelligence techniques with data mining, the authors develop a predictive model for assessing the risk of heart disease based on patient data. Their research demonstrates the potential of data mining in predictive analytics for identifying individuals at high risk of heart disease and implementing preventive measures to mitigate cardiovascular health risks.

These studies collectively underscore the significance of data mining in healthcare and its potential to revolutionize various aspects of healthcare delivery, including disease prediction,

diagnosis, treatment, and management. By leveraging advanced analytical techniques and computational algorithms, healthcare organizations can extract actionable insights from diverse healthcare datasets, leading to improved clinical outcomes, patient satisfaction, and healthcare efficiency.

Furthermore, the integration of data mining with electronic health records (EHRs) presents opportunities for real-time data analysis, clinical decision support, and personalized medicine. By analyzing patient data stored in EHR systems, healthcare providers can identify trends, patterns, and correlations that may inform clinical practice guidelines, treatment protocols, and population health management strategies.

However, the widespread adoption of data mining in healthcare is not without challenges. Privacy concerns, data security issues, and ethical considerations surrounding the use of patient data remain significant barriers to overcome. Additionally, the complexity of healthcare data, including its heterogeneity, incompleteness, and variability, poses challenges for data mining algorithms and predictive modeling techniques.

In conclusion, the literature surveyed highlights the transformative potential of data mining in healthcare and its role in advancing evidence-based medicine, personalized healthcare, and population health management. Continued research and innovation in data mining techniques, coupled with interdisciplinary collaboration between healthcare professionals, data scientists, and policymakers, are essential to harnessing the full benefits of data-driven healthcare analytics.

Overall, data mining holds promise for improving patient outcomes, enhancing healthcare quality, and optimizing resource allocation in healthcare delivery systems. By leveraging the power of data analytics, healthcare organizations can move towards a data-driven paradigm that prioritizes proactive, preventive, and personalized approaches to healthcare management.

3. Methodology

i) Proposed Work:

The proposed smart health prediction system is designed to leverage data mining techniques to enhance healthcare outcomes through early detection, personalized risk assessment, and optimized treatment plans. Data mining can be described as a process of searching patterns or correlations from a large data sets to valuable information that can solve problems and predict outcomes. It involves analyzing certain amount of information to locate certain patterns of occurrence to predict future tendencies, using several processes of effective data collection, warehousing and computer processing The system is envisioned to be a comprehensive solution that integrates seamlessly into the existing healthcare infrastructure, providing valuable insights for both healthcare professionals and individuals. Smarter health systems create extensive and integrated ecosystems that support improvements in consumer and workforce experiences, better care outcomes and greater access to health care services. Health systems can increase productivity and efficiency, and provide better care to more people.
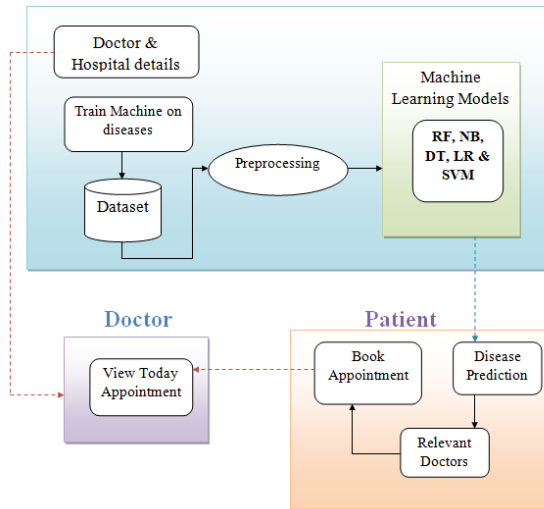
ii) System Architecture:

Fig 1 Proposed Architecture

iii) Admin Modules:

Admin login: The Admin Module allows administrators to log in to the application using the default credentials: username 'admin' and password 'admin'. Upon successful login, administrators gain access to a range of operations. These operations typically include managing user accounts, configuring system settings, monitoring application usage, and performing other administrative tasks essential for maintaining the functionality and security of the application. The Admin Module serves as a centralized control panel for overseeing various aspects of the application's operation and ensuring smooth functionality.

Add doctors: Within the Admin Module, administrators have the capability to add new doctors to the application. After adding a doctor, administrators assign unique usernames and passwords to them. With these credentials, doctors can access the application and log in securely. Once logged in, doctors can view their scheduled appointments with patients. This functionality enables seamless coordination between healthcare providers and ensures that doctors have access to their appointment schedules, facilitating efficient patient management and enhancing the overall workflow within the healthcare system.

View doctors: The "View Doctors" module within the Admin Module provides administrators with the ability to access a comprehensive list of all available doctors within the application. Through this module, administrators can effortlessly retrieve information such as the names, specialties, contact details, and availability of each doctor. This centralized repository of doctor profiles enables administrators to efficiently manage and organize the healthcare workforce, ensuring optimal coverage and distribution of medical expertise. Additionally, the "View Doctors" module facilitates seamless communication and collaboration between administrators and healthcare providers, streamlining administrative processes and enhancing overall efficiency in healthcare service delivery.

Train ML or data mining algorithms: The "Train ML or Data Mining Algorithm" functionality

empowers administrators to utilize disease and symptoms datasets to train a range of advanced algorithms including Random Forest, Support Vector Machine (SVM), Naïve Bayes, Decision Tree, and Logistic Regression. These algorithms are trained to correlate patient symptoms with specific diseases. Once trained, patients can input their symptoms into the models, which then predict potential diseases. This capability enhances diagnostic accuracy, enabling patients to receive preliminary insights into their health conditions promptly. By leveraging cutting-edge data mining techniques, healthcare providers can offer personalized healthcare solutions, fostering proactive disease prevention and management strategies.

iv) Doctor Modules:

Doctor login: Doctors can securely access the application through personalized login credentials provided by the administrator. By logging in, doctors gain access to their respective dashboards where they can view scheduled appointments from patients. This streamlined process ensures efficient management of patient appointments, allowing doctors to prioritize and prepare for upcoming consultations effectively. With seamless access to patient appointment information, doctors can optimize their workflow and provide timely and quality care to patients. The Doctor Login feature enhances communication and coordination between healthcare providers and patients, ultimately improving the overall patient experience within the healthcare system.

View today's appointment: Doctors have the capability to access a convenient "View Today's Appointments" feature within the application. This functionality enables doctors to effortlessly review and manage appointments scheduled for the current day. By accessing this feature, doctors can efficiently plan their daily schedules, prioritize patient consultations, and ensure timely delivery of healthcare services. This real-time visibility into today's appointments enhances doctor-patient communication, minimizes scheduling conflicts, and optimizes the utilization of medical resources. With a clear overview of their daily workload, doctors can focus on providing quality care to patients while maximizing their productivity within the healthcare setting.

v) Patient Modules:

Patient registration: The Patient Registration feature allows individuals to easily sign up and create accounts within the application. Patients can complete the registration process by providing essential information such as their name, contact details, date of birth, and other relevant personal details. Upon successful registration, patients gain access to the full range of functionalities offered by the application, including scheduling appointments, accessing medical records, and communicating with healthcare providers. This seamless registration process enhances patient engagement and empowers individuals to actively participate in managing their healthcare needs. By registering with the application, patients can conveniently access healthcare services and resources tailored to their needs.

Patient login: Patients can securely access the application through the "Patient Login" feature. Upon registration, patients receive login credentials which they can use to log in to their accounts. By logging in, patients gain access to personalized dashboards where they can manage appointments, view medical records, and communicate with healthcare providers. This streamlined login process ensures convenient access to healthcare services and

information, empowering patients to take control of their health and well-being. Through the Patient Login feature, individuals can conveniently access healthcare resources, track their medical history, and engage in proactive healthcare management within the application's user-friendly interface.

Predict disease: Through the "Predict Disease" functionality, patients can input their symptoms into the application to predict potential diseases. Leveraging advanced data mining algorithms trained on disease and symptoms datasets, the application analyzes the entered symptoms and generates predictions regarding possible underlying health conditions. This predictive capability empowers patients to gain insights into their health status, enabling early detection and proactive management of diseases. By providing patients with personalized disease predictions based on their symptoms, the application facilitates informed decision-making and encourages proactive healthcare-seeking behaviors, ultimately contributing to improved health outcomes and enhanced patient empowerment in managing their well-being.

View relevant doctors: Following disease prediction, patients can utilize the "View Relevant Doctors" feature to search for healthcare professionals specialized in treating the predicted condition. The application provides a user-friendly interface where patients can explore a list of relevant doctors based on their expertise and proximity. Additionally, patients can schedule appointments with preferred doctors directly through the application. Moreover, the integration of Google Maps allows patients to view the location of the selected doctor's clinic, enhancing convenience and facilitating navigation. This comprehensive feature empowers patients to access appropriate healthcare services efficiently, fostering a seamless and patient-centered healthcare experience.
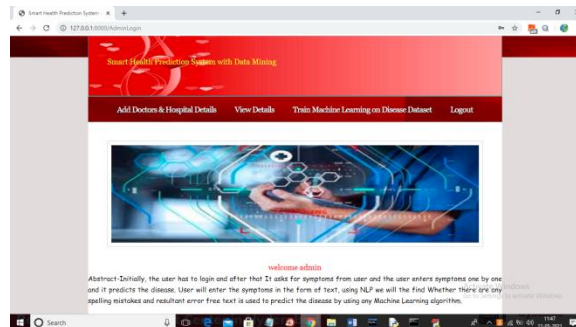
4. Experimental Results

Now double click on 'runServer.bat' file to start DJANGO server and then open browser and enter URL as http://127.0.0.1:8000/index.html and press enter key to get below home page


In above screen click on 'Admin' link to get below login screen
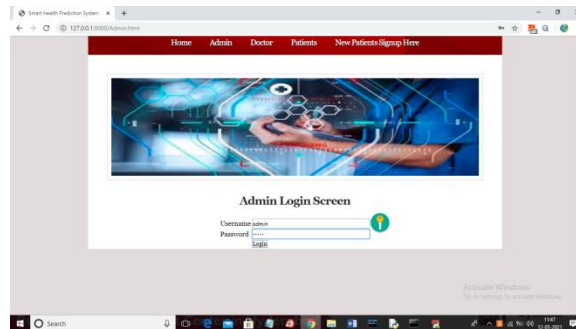
In above screen after login will get below admin screen

In above screen click on 'Add Doctors & Hospital Details' link to add new doctor and if you want you can take doctors details from 'Doctor.csv' file available inside code and this file has doctor username, password, latitude, longitude and address details. See below doctors.csv file
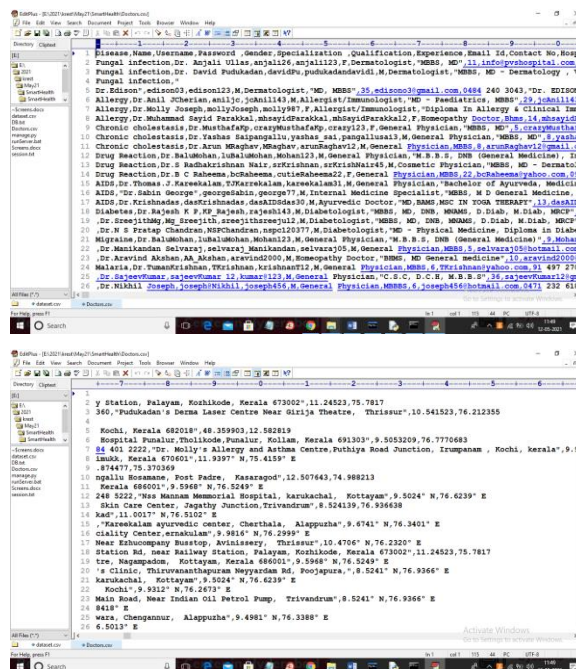
In above screen in last column you can see doctor latitude and longitude and you can use this files to add doctor details by copying content from it

In above screen I added doctor details from same doctors.csv file and if this file contains special symbols such "'s" then remove it otherwise mysql will not accept it and click on submit button to save doctor

Now admin can click on 'View Details' link to get all doctor details in below screen

In above screen admin can view all doctor details and now click on 'Train Machine Learning on Disease Dataset' link to train Random Forest and other machine learning algorithm
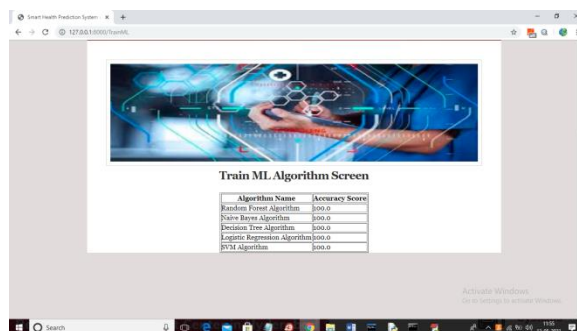
In above screen we trained various data mining algorithms and due to huge dataset size all algorithms got 100% prediction accuracy which means we will get accurate disease prediction and now logout and go for patient signup
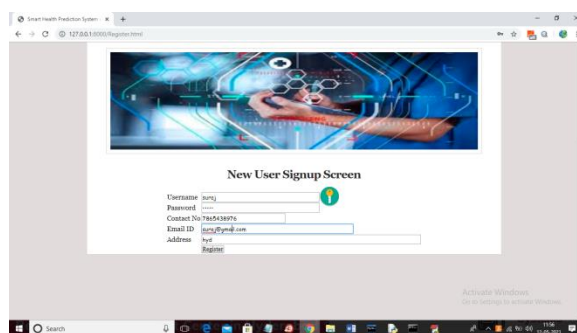
In above screen patients can signup and then click on 'Register' button to complete signup process and to get below screen
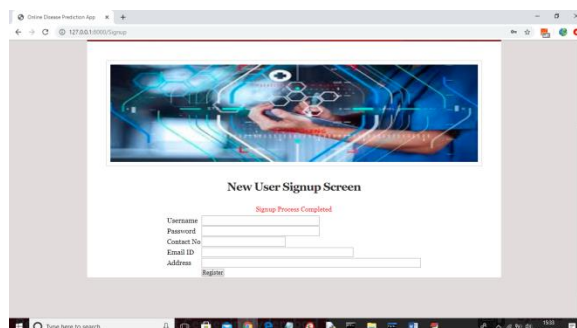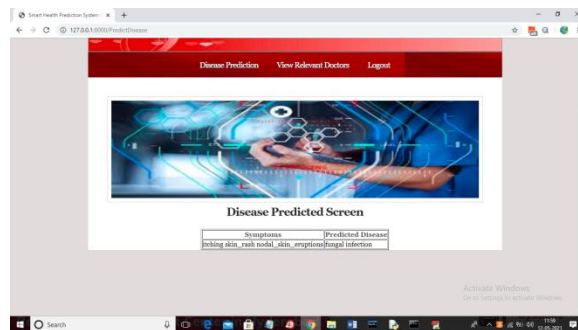

In above screen signup process completed and now patient can login in below screen
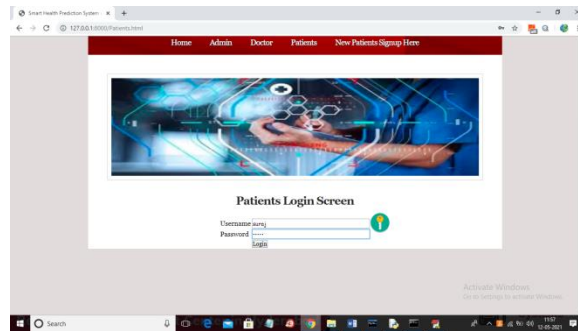
In above screen patient is logged in and after login will get below patient screen

In above screen patient can click on 'Disease Prediction' link to get below screen and then patient can enter symptoms separated by space and if you want you can copy symptoms from dataset file also


In above screen I entered symptom as 'itching skin_rash nodal_skin_eruptions' and then click on 'Submit' button to get below result
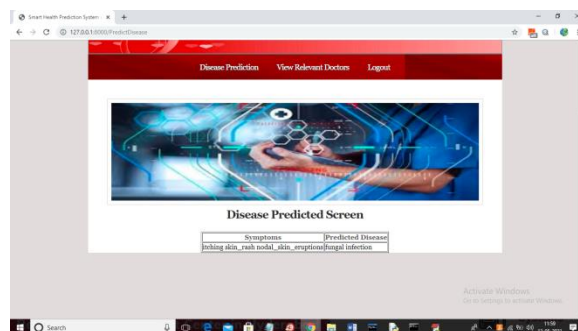

In above screen for given symptoms disease is predicted as 'fungal infection' and now patient can click on 'View Relevant Doctors' link to search and view doctors related to fungal infection
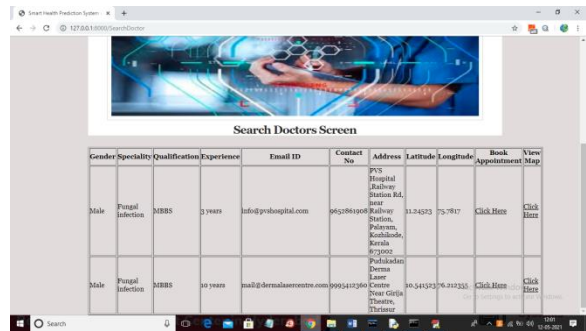
In above screen patient can see 2 doctors for fungal infection and then patient can click on 'Click Here' link under 'View Map' column to view doctor address in map like below screen
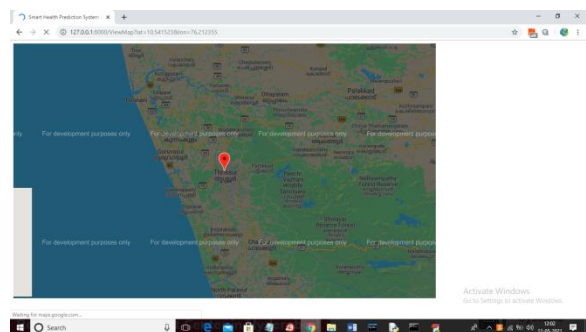

In above screen patient can view doctor address in map and now go back and click on 'Click Here' link under 'Book Appointment' column to book appointment to get below screen
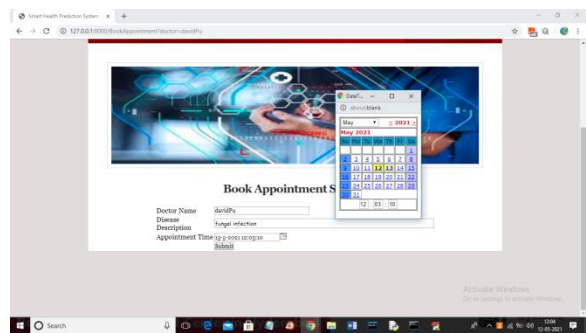

In above screen selected doctor name will be displayed and patient may enter disease name and then select appointment date and time to book appointment and then click submit button to get below screen
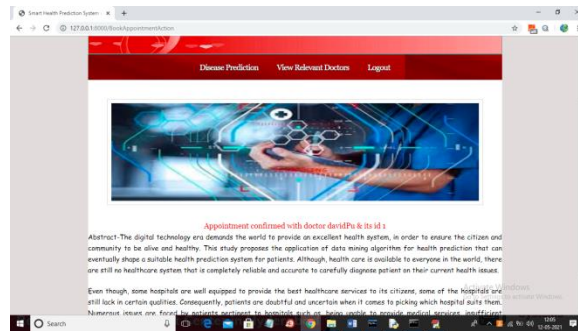
In above screen appointment is confirmed with 'davidPu' doctor and now logout and login as doctor to view appointment
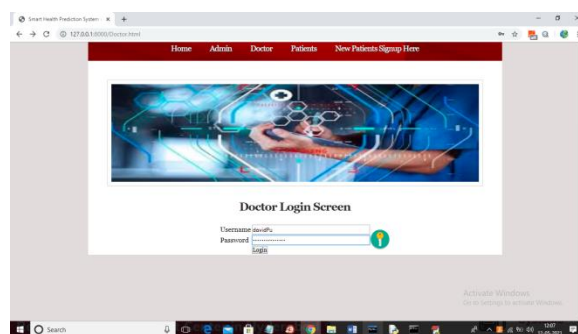
In above screen doctor is login and this login details are given by admin while adding doctors and after login doctor may get below screen
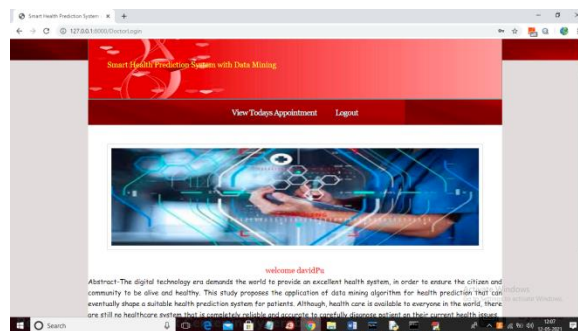
In above screen doctor can click on 'View Todays Appointment' button to view appointment and to get below appointment details from patients
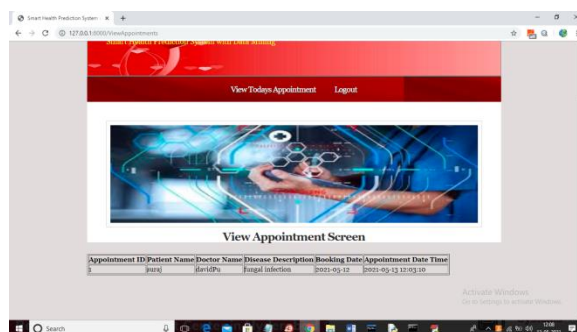
In above screen doctor can view patient appointment details.

Similarly you can add N number of patients and doctors and can perform disease prediction by giving symptoms from dataset.

5. Conclusion

Tendency for data mining application in healthcare today is great, because healthcare sector is rich with information, and data mining is becoming necessity.Healthcare organizations produce and collect large volumes of information on daily basis. They need to be collected and stored in the organized forms, and their integration enables forming of hospital information system. Healthcare data mining provides countless possibilities for hidden pattern investigation from these data sets. These patterns can be used by physicians to determine diagnoses, prognoses and treatments for patients in healthcare organizations By focusing on early disease detection, risk assessment, and optimized treatment plans, these systems aim to shift the healthcare paradigm from reactive to preventive. The ability to predict and manage chronic diseases enhances the quality of life for individuals, fostering a proactive approach to health management.The smart health prediction system stands at the forefront of a healthcare revolution, offering a glimpse into a future where data-driven insights empower individuals, support healthcare professionals, and contribute to a healthier society.

6. Future Scope

Integrating more advanced AI and machine learning techniques into the project holds significant promise for enhancing predictive accuracy and adaptability. By leveraging cutting-edge algorithms and models, the application can provide more precise disease predictions and personalized healthcare recommendations tailored to individual patient needs. Additionally, incorporating blockchain technology offers a robust solution for enhancing security and privacy within the healthcare ecosystem. Through blockchain, patient trust can be bolstered, and data sharing among healthcare entities can be securely facilitated, ensuring compliance with evolving privacy regulations such as GDPR and HIPAA. This future scope underscores the potential to revolutionize healthcare delivery by harnessing the power of AI, machine learning, and blockchain technologies to optimize patient care and data management processes.

References

[1] Scheffler RM, Arnold DR. Projecting shortages and surpluses of doctors and nurses in the OECD: what looms ahead. Health Econ Policy Law 2019 Apr;14(2):274-290.2.

[2] Chowdhury SBS. A Review Paper on Health Data Integration with Secured Record Linkage in National Health Data Warehouse and Privacy & Compatible Solution for its Privacy and Security Problems in Bangladesh and other developing countries. 2017. International Islamic University Chittagong (IIUC). Available at: http://works.bepress.com/sayem-binsarwarchowdhury/2/

[3] Tulsky JA, Beach MC, Butow PN, Hickman SE, Mack JW, Morrison RS, et al. A research agenda for communication between health care professionals and patients living with serious illness. JAMA Intern Med 2017 Sep;177(9):1361-1366.4.

[4] Kuziemsky, C. Decision-making in healthcare as a complex adaptive system. Healthcare management forum. Healthc Manage Forum 2016 Jan;29(1):4-7. 5.

[5] Lakshmi K, Ahmed DI, Kumar GS. A smart clinical decision support system to predict diabetes disease using classification techniques. IJSRSET 2018; 4(1):1520-1522. 6.

[6] Durairaj M, Ranjani V. Data mining applications in the healthcare sector: a study. Int J Sci Tech Res 2013; 2(10):29-35. 7.

[7] Chaudhary BV, Radhika Y. A Survey on Applications of Data Mining Techniques. Int J Appl Engg Res 2018; 13(7): 5384- 5392. 8.

[8] Kirubha V, Priya SM. Survey on data mining algorithms in disease prediction. Int J Comp Trends Tech 2016; 38(3): 124- 128. 9.

[9] Thakkar, B.A., Hasan, M.I. and Desai, M.A. Health care decision support system for swine flu prediction using naïve Bayes classifier. Int Conf Adv Recent Tech Comm Comp 2010. DOI:10.1109/ARTCOM.2010.98 10.

[10] Palaniappan S, Awang R. Intelligent heart disease prediction system using data mining techniques. Conference: Computer Systems and Applications, 2008. AICCSA 2008. DOI: 10.1109/ AICCSA.2008.4493524.

[11] Saqib Qamar , Syed Qasim Afser Rizvi, et. al., "Techniques of Data Mining In Healthcare: A Review" published in research gate open Access, available at https://www.researchgate.net/publication/278729826.

[12] Basma Boukenze , Abdelkrim Haqiq & Hajar Mousannif , et. al., "Predicting Chronic Kidney Failure Disease Using Data Mining Techniques" published in springer open Access, available at https://link.springer.com/chapter/10.1007/978-981-10-1627-1_55.

[13] Casey C. Bennett ,Thomas Doub   , et. al., "Data mining and electronic health records: Selecting optimal clinical treatments in practice" published in research gate open Access, available at https://www.researchgate.net/publication/51962845.

[14] P Rajeswari , G.Sophia Reena, et. al., "Analysis of liver disorder using data mining algorithm" published in research gate open Access, available at https://www.researchgate.net/publication/265082307.

[15] Latha Parthiban , R. Subramanian, et. al., "Intelligent heart disease prediction system using CANFIS and genetic algorithm" published in research gate open Access, available at https://www.researchgate.net/publication/238733259.