CHOOSELYF

Parth Adokar
Department of Information Technology Engineering
Atharva College of Engineering
Mumbai, India

Shubham Awasthi
Department of Information Technology Engineering
Atharva College of Engineering
Mumbai, India

Atharva Dhopadkar
Department of Information Technology Engineering
Atharva College of Engineering
Mumbai, India

Navkar Kumath
Department Of Information Technology Engineering. Atharva College of Engineering
Mumbai, India

Prof. Nilemma Pathak
Department of Information Technology Engineering, Atharva College Of Engineering
Mumbai. India

Abstract **-- ChooseLyf is a phenomenon that offers an application for smartphones to ease healthcare assistance via offering consumers with efficient and effective answers. Because the utilization of mobile phones has increased significantly, the model will provide users with accessibility to all services at any moment and from any location. In this contemporary and digital age, it is critical to handle problems in a computerized swifter manner, so the requirement during a crisis must be met in a user-friendly manner. This application includes a symptom tester, which analyzes the inputted patient's symptoms and recommends medicines or physicians based on the signs and symptoms. It is difficult to recall every detail in today's world, which is why the medication reminder maintains the individual using it updated on the prescribed dose without forgetting any. The accuracy of information is of the utmost importance due to the various kinds of user accounts, which aid in safeguarding information by only granting access to information to which the user is allowed. This application also includes a conversation area where the user can ask a Chatbot about their symptoms. A machine learning program that detects three main diseases, namely heart disease, diabetes, and Parkinson's disease, has also been developed.**

Keywords — Diabetes, Heart, Parkinson, Linear Regression, Machine Learning, Android.

I. INTRODUCTION

Mobile devices have become increasingly popular among both healthcare professionals and the broader population in the past few decades. The cell phone is a novel technology that integrates mobile contact and processing in a portable device, allowing for computing capabilities at the level of treatment. The primary goal of this research is to identify and describe smartphone-based medical equipment in the available literature in relation to their functionality. In this respect, we offer an organized examination of the available literature. To the greatest extent of our understanding, this is the first study in a structured literature review style for organizing and reviewing health care applications for cellphones. The meaning of "good health" can be viewed from a number of perspectives. Recognizing how different people view health on an individual level could provide specialists with useful information about what affects behavior with regard to of happiness and mental health in the general population.

We all know that wellness is prosperity. Health has evolved into a crucial component for everyone on the globe. An individual's well-being is always an important factor in his or her capacity to thrive in life. Most people had undoubtedly overlooked this aspect since their daily lives are filled with busy schedules.

In recent years, the use of machine learning methods for illness prediction has grown in prominence in healthcare. Diabetes, heart disease, and Parkinson's disease are just a few of the most prevalent long-term conditions that impact millions of people around the globe. Early identification and precise prediction of these illnesses can enhance patients' quality of life and allow healthcare workers to offer timely treatment and prevention measures.

Machine learning algorithms provide an efficient and effective method for predicting diseases based on a variety of criteria such as demographic data, medical background, lifestyle routines, and genetic variables. These algorithms can analyze large quantities of data and find trends that human specialists could have overlooked.

To address such circumstances, we are creating an Android program called "ChooseLyf" that includes a catalogue of diseases and their associated symptoms. Other tools include a BMI estimator, a medication reminder, and health tips. The user can look for local hospitals and clinics, as well as a machine learning system that uses linear regression analysis to identify Parkinson's, diabetes, and heart conditions.

II. MOTIVATION

The present healthcare system confronts a number of challenges, including limited access to treatment, growing healthcare expenses, and an expanding chronic illness load. Patients frequently encounter barriers to prompt and high-quality healthcare, leading to poor health results. Furthermore, healthcare workers are overburdened with growing patient loads, making personalized and effective treatment difficult.

To handle these issues, a complete healthcare smart application that uses the latest innovations such as AI chatbots, ML algorithms, and medical reminders is required. Such an application enables patients to effectively access healthcare services, while healthcare professionals may deliver personalized treatment to patients.

Making a healthcare smart application, on the other hand, poses several challenges, such as assuring patient confidentiality and data security, integrating with current healthcare systems, and overcoming regulatory hurdles.

As a result, the problem statement for developing a healthcare smart application is to create an exhaustive, intuitive, and secure system that utilizes the most recent innovations to provide individuals with personalized and efficient health care while maintaining information safety and compliance with laws and regulations. In order to handle the issues that the present healthcare system faces, the application must be linked to current healthcare systems and obtainable to a differed variety of patients.

III. LITERATURE REVIEW

1. "Understanding the Lives of People to Identify the Causes of Diabetes Utilising Data Mining," by Gavin Pinto, Sunil Jangid, and Radhika Desai.

Heart disease prediction was designed by Marjia et al. [8] employing KStar, j48, SMO, Bayes Net, and Multilayer awareness with Weka tool. Using kfold classification algorithm, SMO and Bayes Net surpass KStar, Multilayer Perception, and J48 method depend on results from various MOTfactors. These algorithms' precision results remain under acceptable. As a result, the consistency's efficiency is improved further to provide superior diagnosis decisions

2.M.Marimuthu, S.Deivarani ,R.Gayatri, "Analysis of Heart Disease Prediction using Machine Learning Techniques".

Nave Bayesian, Decision Trees, Support Vector Machines (SVM), and neural network models are used to pull information about prior medical files in order to study techniques to anticipate chronic condition (ANN). To find out which categorization actually works on an exact accomplishment, a comparative study is done. SVM has the highest accuracy rate in this experiment, whereas Nave Bayes has the highest accuracy for diabetes..

3. Doctor Chat-bot – Smart Health Prediction, IJSRST 2021

The mass of chatbots analyze the information using tools like WEKA or MATLAB for Intelligence and data mining techniques like SVM, Logistic Regression, Evolutionary Algorithms, Colonnaded Bayesian, Pattern Recognition, Xml, and much more.. After being given raw information, some algorithms were examined and contrasted on the basis of accuracy. AIML was used for the QA framework in some Chatbots. The Conversation Interface specified in this article is Dialog flow. The Support Vector Machine method was the data mining model that gave the highest accuracy for forecasting these illnesses from a heart disease collection in the medical realm. AIML is additionally available to create Chatbots by storing queries and responses in XML format. For analysis, MATLAB and Weka tools were used, including pre-processing, grouping, and graph visualization.

IV. PROPOSED SOLUTION

The healthcare smartphone app created with Android Studio called 'ChooseLyf' is a complete utility intended to meet the requirements of both patients and healthcare providers. The app includes revolutionary features such as an AI chatbot, medical alerts, ML algorithms, and a search tool for local hospitals.

The AI chatbot is a clever virtual helper that interacts with patients and answers their healthcare-related questions using spoken word processing. The chatbot is programmed to respond to patient inquiries in a customized manner, to give rudimentary health care guidance, and to refer patients to suitable healthcare providers if required.

Patients can plan medicine and meeting notes using the health care reminders tool. Customers receive timely alerts from the application, assuring that they administer medicines on schedule and do not skip any meetings.

The application's ML algorithms provide predictive analytics to assist patients and healthcare workers in predicting and preventing the onset of chronic illnesses such as diabetes, heart disease, and Parkinson's disease. To spot possible health hazards and provide early intervention, the algorithms analyze patient data such as demographic data, medical histories, living routines, and genetic variables.

Patients can use the local hospitals search tool to find hospitals near them, clinics, and medical providers, as well as contact information, ratings, and reviews. This feature is particularly helpful for people who are new to a community and require to locate a healthcare practitioner fast.

To address the problem statement of creating a comprehensive, user-friendly, and secure healthcare smartphone application, we propose the following solution using Android Studio:

AI Chatbot: Implement an AI chatbot that uses natural language processing to interact with patients and assist them with their healthcare-related queries. The chatbot should be able to provide personalized responses to patient queries, offer basic medical advice, and direct patients to appropriate healthcare providers if necessary.

ML Algorithms: Integrate ML algorithms that analyze patient data, including demographic information, medical history, lifestyle habits, and genetic factors, to predict potential health risks and provide early intervention for Parkinson's disease, diabetes, and heart diseases.

Medical Reminders: Develop a medical reminders feature that enables patients to schedule their medication and appointment reminders. The application should send timely notifications to patients, ensuring that they take their medications on time and do not miss any appointments.

Appointment Scheduler: Implement an appointment scheduler that enables patients to schedule appointments with healthcare providers. The application should enable patients to view healthcare providers' availability, select a suitable time slot, and receive confirmation of their appointment.

Data Security: Ensure patient privacy and data security by implementing robust security measures, such as encryption and authentication protocols. The application should comply with all relevant regulations and standards to ensure the confidentiality and integrity of patient data

V. DESIGN

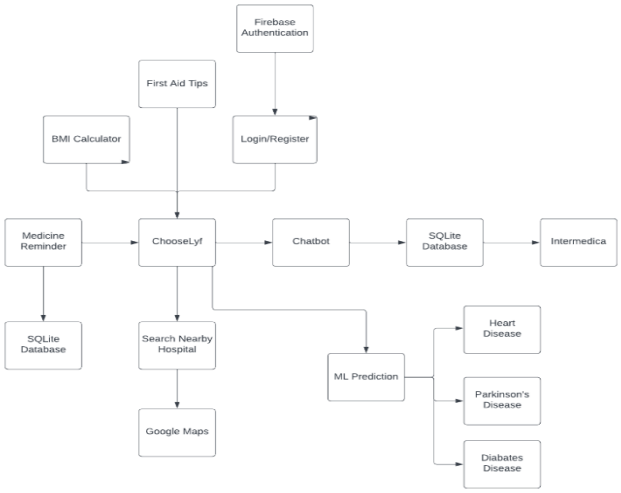


Figure 5.1: Block Diagram

In figure 5.1, we have experimented on three diseases that is heart, diabetes and liver as these are correlated to each other. The first stage is to import the UCI dataset, PIMA dataset, and Indian liver information for heart disease, diabetes illness, and liver disease, correspondingly. Once we have imported the dataset then visualization of each inputted data takes place. After visualization pre-processing of data takes place where we check for outliers, missing values and also scale the dataset then on the updated dataset and we divide the information into two parts: training and testing..

On the training dataset, we used KNN, XGboost, and random forest algorithms, and on the assessment dataset, we used information on the categorized method. After implementing our expertise, we will select the method with the highest efficiency for each illness. Then we create a condiment submit for each illness and then.

5.2 User Interface Design

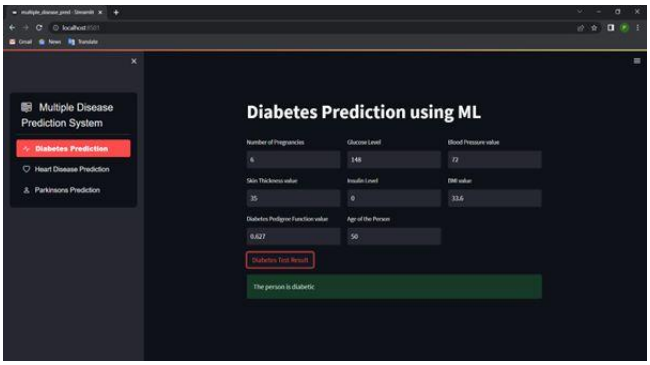


Figure 5.2: Graphical User Interface

VI. IMPLEMENTATION

6.1 Linear Regression

The working of the Linear Regression algorithm is as followed:

Step-1: We have a dataset with multiple features that can help predict the diseases, and each instance in the dataset has labels indicating the presence or absence of diabetes, heart disease, and Parkinson's disease.

Step-2: Preprocessing the dataset. Divide the information into two stages: testing and training.

- Normalize or scale the features to ensure that all features are on a similar scale and have the same importance during training.
- One-hot encode the labels to convert them into binary values (0 or 1).

Step-3: Initialize the parameters:

- Initialize the weights and bias with zeros or small random values.
- Set an instruction rate and repetition count.

Step-4: rain the model:

- To maximize the balance of weights and skew, use descend gradients.
- Update the weights and bias for each feature by computing the partial derivative of the loss function with respect to the weight and bias.
- Use the sigmoid function to compute the predicted probabilities for each disease.
- Estimating the gradient descent to discover the difference between the actual identifiers and the projected odds.
- Update the weights and bias in the opposite direction of the gradient to minimize the loss.

Step-5: Then assign the new point to the category having maximum number of neighbors. For example Category A has highest number of neighbor so we will assign the new data point to category A.

Step-6: Test the model:

- Use the trained model to predict the disease probability for the testing set.
- To assess the model's success, determine the reliability, precision, recollection, and F1 value.

6.2 Random Forest Algorithm

Dealing with a random forest is feasible in two stages: the initial phase is to construct the random forest by combining N decision trees, and the subsequent phase is to make predictions for each tree that were generated in the first step. The working of the random forest is as follows:

Step-1: It will first choose K unique data values from the training collection.

Step 2: Generate the classification trees corresponding to the chosen k data values after.

Step 3: Prefer N as that of the number of decision trees you'd like to develop.

Step-4: Identifying each decision tree's forecasts and giving the latest information to the group with the most votes.

6.3 XGBoost Algorithm

The following describes how the XGBoost algorithm operates.

Step 1: Initially make a single-leaf tree.

Step 2: Then, for the initial branch, we should always estimate the mean of the data point as an estimate, and thereafter determine the residuals using preferred gradient descent. Additional trees then employ the approximation from the previous tree to determine their multicollinearity.

Step 3: Using an algorithm, compute the similarity score:

$$Similarity\ Score = Gradient \frac{Gradient^2}{Hessian + \lambda}$$

where Hessian is the number of residuals Gradient2 = residual sum squared; is a normalization hyper parameter.

Step 4: Using the similarity value, we choose the right node. The greater the similarity number, the greater the uniformity.

Step 5: We decide the Knowledge Advantage by implementing the Togetherness Significance. Supervised learning reveals the variation among old and new resemblance and shows the level of homogenisation achieved by separating the terminal at a selected spot. According to the following algorithm, it is determined:

$$Information\ Gain = Left\ Similarity + Right\ Similarity - Similarity\ for\ Roots$$

Step 6: Creating the tree of desired length using the above method pruning and regularization can be done by playing with the regularization hyper parameter.

Step 7: Then, utilizing the Decision Tree you created, one can forecast the leftover numbers. The new collection of residuals is computed as follows:

Step 8: Then go back to step 1 and repeat the process for all the trees.

$$New\ Residuals = Old\ Residulas + \rho \sum Predicted\ Residuals$$

VII. RESULTS AND DISCUSSIONS

The KNN algorithm was used in the system to forecast diabetes disease, the XGboost algorithm was used to predict heart disease, and the random forests algorithm was used to predict liver disease. Whenever the client uploads the disease-specific measure, it will indicate the degree to which the client has the illness in question. The components will show the range of values that are needed, and if the principle embodies falls outside of that range, is wrong, or is missing, a warning message will appear asking that the correct value be entered.

ACCURACY FOR EACH DISEASE:

Table No 7.1: Diabetes Disease

ALGORITHM	Diabetes
Linear Regression	88%
XGBoost	84%

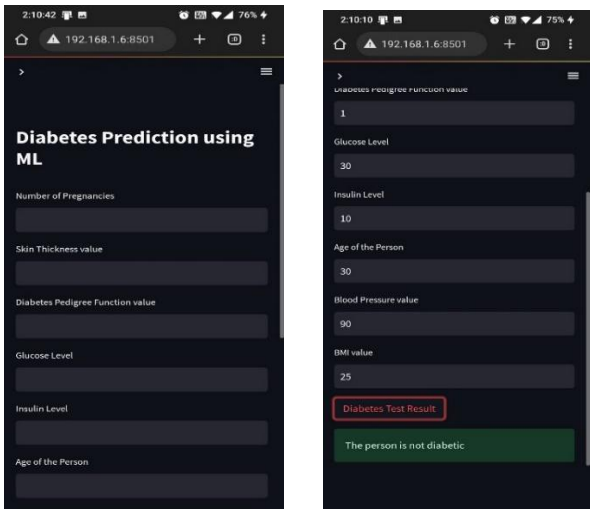
Table No 7.2: Heart Disease

ALGORITHM	Heart
Linear Regression	85%
Random Forest	77%

Table No 7.3: Parkinson's disease

ALGORITHM	Liver
Linear Regression	73%
XGBoost	68%

2. Diabetes Disease:



A. Figure No 7.1: Diabetes Disease

3. Heart Disease/Parkinson's:

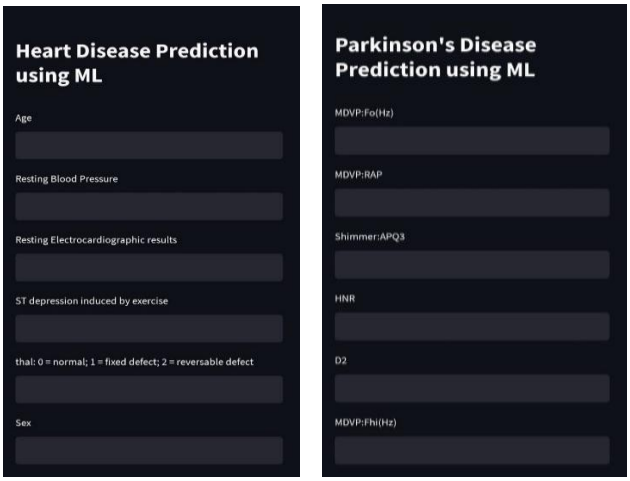


Fig 7.2 Heart Disease/Parkinson's Disease

Figure 7.3: ChooseLyf Interface / Chatbot

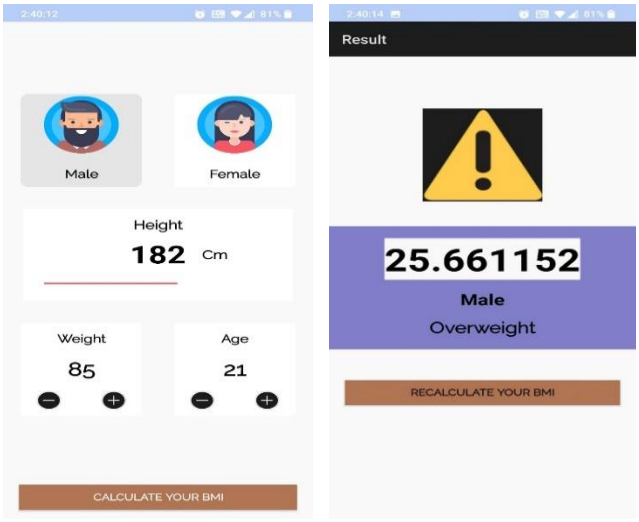


Figure 6.3: BMI page and result

VIII.CONCLUSION

ChooseLyf was designed with user-friendliness in mind, allowing users to take advantage of machine learning in both routine tasks and emergency situations. The patient's symptoms are analyzed using a symptom analyzer, which enables doctors to treat patients more effectively.

Numerous machine learning, data mining, and other algorithms have been created to anticipate the incidence of cardiovascular disease. Determine how well each precisely define, and afterwards incorporate the proposed solution to the extent required. Significantly raise the precision of methodologies by employing more exact methods for picking features. There are multiple options for therapy if an individual is diagnosed with a specific sort of coronary disease. Data analysis with a suitable data source can yield significant understanding.

Diabetic prediction methods in the system used the Clustering techniques, heart problem uses the XGBoost automated system, and liver situation utilises the arbitrary forests technique because the outcomes were most precise. When the client adds the illness variable, this would demonstrate if the client has an illness or not based on the illness chosen. The variables will display the required price point, and if the valuation is outside of those distances, is invalid, or is vacant, it will display a warning logo that enhances a correct quantity.

As a consequence, as disclosed by the review article, cascaded and more advanced features are needed to enhance the precision of predicting the early symptoms of heart problems, since only slight advancements have been made in the development of forecasting analytics for people with heart disease. As even more information is entered into the directory, the scheme are becoming extremely advanced.

IX. REFERENCES

[1] Trends in coronary Heart Disease Epidemiology

[2] Centre for Disease Control and Prevention (Heart Disease Facts).

[3] Asian Pacific Journal of Global Trend of Cancer Mortality rate: A 25-year study.

[4] Times of India: Cancer cases upswing 10% in 4 years to 13.9 lakh.

[5] International Diabetes Federation: Expenditure and deaths related to diabetes.

[6] Epidemiology of Diabetes: A report of Indian Heart Association.

[7] Naveen Kishore G, V .Rajesh, A.Vamsi Akki Reddy, K.Sumedh,T.rajesh Sai Reddy, "Prediction Of Diabetes Using Machine Learning Classification Algorithms".

[8] Gavin Pinto, Sunil Jangid, Radhika Desai,"Understanding the Lifestyle of people to identify the reasons of Diabetes using data mining".

[9] M.Marimuthu, S.Deivarani ,R.Gayatri, "Analysis of Heart Disease Prediction using Machine Learning Techniques".

[10] Purushottam, Richa Sharma, Dr. Kanak Saxena,"Efficient Heart Disease Prediction System".

[11] Adil Hussain She, Dr. Pawan Kumar Chaurasia," A Review on Heart Disease Prediction using Machine Learning Techniques".

[12] M. Chinna Rao, K. Ramesh, G. Subbalakshmi,"Decision Support in Heart Disease Prediction System using Naïve Bayes".

[13] Amandeep Kaur, Jyothi Arora," Heart Disease Prediction using data mining Techniques: A survey".

[14] Noreen Fatima, Li Liu, Sha Hong, Haroon Ahmed ,"Prediction of Breast Cancer, Comparitive Review Of Machine Learning Algorithms and their analysis".

[15] Ch .Shravya ,K.Pravallika , Shaik Subhani, "Prediction of Cancer using supervised machine learning Algorithms".

[16] Nikita Rane, Jean Sunny, Rucha Kanade, Sulochana Devi," Breast Cancer classification and prediction using Machine learning ".

[17] Deepti Sisodia, Dilip Singh Sisodia," Prediction of Diabetes using classification Techniques".

[18] Dr.B.Santhosh Kumar, T.Daniya, Dr. J.Ajayan," Breast Cancer Prediction using Machine Learning Algorithms".

[19] Mümine KAYA KELEŞ,"Cancer Prediction using and Detection using Machine Learning Algorithms: A Comparative Study".

[20] Heart Disease Dataset" by UCI.

[21] Pima Indians Diabetes Dataset" by Kaggle.

[22] Ravuvar, H. N., Goda, H., R, S., & Chinnasamy, P. (2020). Smart Health Predicting System Using K-Means Algorithm. 2020 International Conference on Computer Communication and Informatics (ICCCI).

[23] Mathew, R. B., Varghese, S., Joy, S. E., & Alex, S. S. (2019). Chatbot for Disease Prediction and Treatment Recommendation using Machine Learning. 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI).

[24] Vijava Shetty, S., Karthik, G. A., & Ashwin, M. (2019). Symptom Based Health Prediction using Data Mining. 2019 International Conference on Communication and Electronics Systems (ICCES).

[25] Siddhi Pardeshi, Suyasha Ovhal, Pranali Shinde, Manasi Bansode, Anandkumar Birajdar. A survey on Different Algorithms used in Chatbot. 2020 International Research Journal of Engineering and Technology (IRJET)

[26] Ravi Aavula, M.Kruthini, N.Ravi teja, K.Shashank. Smart Health Consulting Android System. 2017 International Journal of Innovative Research in Science, Engineering and Technology

[27] Chetty, N., Vaisla, K. S., & Patil, N. (2015). An Improved Method for Disease Prediction Using Fuzzy Approach.

[28] Prashant Tiwari, Aman Jaiswal, Narendra Vishwakarma, Pushpanjali Patel. SMART HEALTH CARE (AN ANDROID APP TO PREDICT DISEASE ON THE BASIS OF SYMPTOMS). 2017 International Research Journal of Engineering and Technology (IRJET)

[29] Devashri Deshmukh, Ulhas B. Shinde, Shrinivas R. Zanwar. Android Based Health Care Monitoring System. 2017 INTERNATIONAL JOURNAL OF ADVANCE SCIENTIFIC RESEARCH AND ENGINEERING TRENDS.

[30] Kumar, M. A., & Sekhar, Y. R. (2015). Android based health care monitoring system. 2015 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS).

[31] Ms. Dipali S. Shinde, Ms. Bhagyashree N. Tupe, Ms. Pooja R. Raut, Ms. Supriya P. Gore, Mrs. Jagruti Wagh. Android Application for Health-care System with Data Warehouse. 2017 IJSTE - International Journal of Science Technology & Engineering