iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** XI **Month of publication:** November 2022

DOI: https://doi.org/10.22214/ijraset.2022.47434

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue XI Nov 2022- Available at www.ijraset.com

SmartCare: A Symptoms Based Disease Prediction Model Using Machine Learning Approach

Stepheny Lucas[1], Mitali Desai[1], Amisha Khot[1], Sincee Harriet[1], Nilambari Narkar[2]
[1]*Bachelor in Computer Engineering, Xavier Institute of Engineering, Mumbai*
[2] *Assistant Professor, Department of Computer Engineering, Xavier Institute of Engineering, Mumbai*

Abstract: ***The breakthrough on computer-based technology has resulted in storage of a lot of electronic data in the healthcare industry. Machine Learning technology has been proven beneficial in giving an immeasurable platform in the medical field so that health care issues can be resolved effortlessly and expeditiously. Prediction of disease at early stage could help people from getting the necessary treatment on time. These days many virtual prediction models are available for the same. The existing systems either made use of only one algorithm or prediction system were capable for predicting only one disease. The maximum accuracy of the existing systems range between 52% to 88%. The algorithms used in various prediction system consisted of Linear Regression, Decision Tree, Naïve Bayes, KNN, CNN, Random Forest Tree, etc. In our project i.e., "SmartCare: A Symptoms Based Disease Prediction Model Using Machine Learning Approach", it is possible to predict more than one disease at a time. So, the user does not need to traverse many models to predict the diseases. It will help to reduce the time and cost of predicting diseases at prior stages, so as to prevent the extremities of it and thus, there is a chance of reducing mortality rate.***
Keyword: Machine Learning, Decision Tree Algorithm, K-Nearest Neighbour Algorithm, Naïve Bayes Algorithm, Random Forest Algorithm, GPS Connectivity.

I. INTRODUCTION

Healthy lifestyle, healthcare and medicines are few of the essential elements of human lifestyles and economy. There is a tremendous change in the world we are living in now and the world that existed few months back. Everything has turned ugly and divergent. In this case, where the entirety has grown to become digital or let us say virtual, the doctors and nurses are giving their maximum efforts to keep people's lives and people's health even though they ought to danger their very own.

Even now in some parts of the world there are still some far-flung villages, remote places which lack clinical centers, health facilities. Machines have started to gain popularity and dependency by humans as, without any human mistakes, they could perform duties greater efficaciously and with a steady degree of accuracy.

A disease predictor is nothing but a virtual doctor, which can predict the disorder of any affected person without any human errors. The first disease prediction system focused on input of blood report values. Whereas The Symptoms Based Disease Prediction Model predicts the disease of the patient based on the input of symptoms. Depending on the disease being diagnosed a specialized doctor will be assigned for the patient.

The following algorithms are used in developing the Symptoms Based Disease Prediction Model: -

1) Decision Tree is a Supervised learning technique that can be used for both classification and Regression problems, but mostly it is preferred for solving Classification problems. It is a tree-structured classifier, where internal nodes represent the features of a dataset, branches represent the decision rules and each leaf node represents the outcome.[1]
2) Random Forest is a popular machine learning algorithm that belongs to the supervised learning technique. It can be used for both Classification and Regression problems in ML. It is based on the concept of ensemble learning, which is a process of combining multiple classifiers to solve a complex problem and to improve the performance of the model.[2]
3) The k-nearest neighbors' algorithm, also known as KNN or k-NN, is a non-parametric, supervised learning classifier, which uses proximity to make classifications or predictions about the grouping of an individual data point.[3]
4) Naïve Bayes algorithm is a supervised learning algorithm, which is based on Bayes theorem and used for solving classification problems. It is mainly used in text classification that includes a high-dimensional training dataset. Naïve Bayes Classifier is one of the simple and most effective Classification algorithms which helps in building the fast machine learning models that can make quick predictions.[4]

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue XI Nov 2022- Available at www.ijraset.com

II. LITERATURE REVIEW

1) *"Multi Disease Prediction Model by using Machine Learning and Flask API", "Akkem Yaganteeswarudu (2020)"* [5] - Mostly disease prediction model focuses on only one disease per analysis. Multi Disease prediction is a model which can be used to predict multiple diseases by using machine learning and Flask API. So, the user doesn't need to traverse through many models for disease prediction. The system analyzes heart disease and breast cancer analysis, Diabetes analysis, Diabetes Retinopathy analysis. The importance of this system is to research most diseases so that to monitor patient's condition and warn the patient beforehand to decrease the mortality ratio.
2) *"Symptoms based disease prediction using decision tree and electronic health record analysis", "Radhika, S and Shree, S Ramiya and Divyadharsini, V Rukhmani and Ranjitha (2021)"* [6] - The world is moving forward with fast speed and people are investing their whole time in working. So, they don't look after their body and hardly get time to go to hospital for checkup. Symptoms based disease prediction is developed as a full virtual heath hub which has four modules specifically admin, patient, doctor and disease prediction. This system works as per the symptoms entered by the user and it predicts if the disease is present or not.
3) *"Heart Disease Prediction Using Machine Learning Algorithms", "Archana Singh, Rakesh Kumar (2020)"* [7] - Heart is one of the most vital organs of human body. Therefore, prediction and diagnosis of heart related disease needs more accuracy and perfection because one mistake leads to death of a person. There is essential need of prediction system to deal with these problems and to create awareness about diseases. Heart disease prediction model is a system in which disease is predicted by using four algorithms which are k-nearest neighbor, decision tree, linear regression and support vector machines (SVM). It uses biological parameters like cholesterol, Blood pressure, etc. and prediction is done based on the accuracy of algorithms.

III. SYSTEM ANALYSIS

A. *Objective*

With the advancement of technology in almost every field, and the growing use of machine learning in various sectors, notably healthcare. The main goal of our project is to help people diagnose illnesses at an early stage so that the patient can receive the necessary treatment on time. The user can also predict the disease while sitting at home. It can also be used in hospitals to take appropriate precautions to avoid or reduce risk, thereby improving quality of care and avoiding potential hospital admissions.

B. Problem Statement

The "SmartCare: A Symptom-Based Disease Prediction Model Using Machine Learning Approach" does not focus on the prediction of a specific disease; instead, it predicts disease based on the symptoms given by the user. As a result, the user does not need to traverse many models to predict the disease. There is a probability of lowering the death rate due to the prediction of disease at an early stage. Utilizing machine learning methods, our goal is to create a symptom-based disease prediction model. The Frontend of the system would consist of a responsive website that can be accessed through any device. The website will be developed using HTML, CSS and JavaScript will be connected to the ML model using Flask. The user has to input the parameters for a specific disease and the model will detect if the disease is present or not. Based on more than 5000 records of patients, our goal is to create a prediction model that analyses the user's symptoms, determines the disease he or she is more likely to have, and directs the user to the closest hospitals based on location.

IV. SYSTEM REQUIREMENT

A. *Hardware Requirements*

1) Works on: Laptops, Computers, Tablets and Mobiles.
2) RAM: Minimum 2gb and recommended 8gb or More
3) CPU: Dual Core or Quad Core Processor
4) Bandwidth: 16Mbps or more

B. *Software Requirements*

1) Browsers supported: Chrome, Brave, Microsoft Edge, Mozilla Firefox
2) Technology used: HTML, CSS, JavaScript, JavaScript Libraries, Flask
3) Complier used: Sublime Text Code Editor and Google Collab

V. SYSTEM DESCRIPTION

Users can give various symptoms and the issues they are facing. The application takes the user's symptoms as inputs to check for various illnesses that could be associated with it using the algorithms. The system also provides the users with a list of hospitals near them which they could visit for further consultancy. The model will be available as a website for the user to use and is simple as well as easy to use. The traditional diagnosis approach demands an affected person visiting a doctor, undergoing many clinical assessments, and then reaching a conclusion. This whole process was very time consuming. This project proposes an automated disease prediction system using machine learning approach to save time and cost by predicting diseases at prior stages, so as to prevent the extremities of it and thus, there is a chance of reducing mortality rate.

VI. IMPLEMENTATION AND RESULT

A. System Architecture

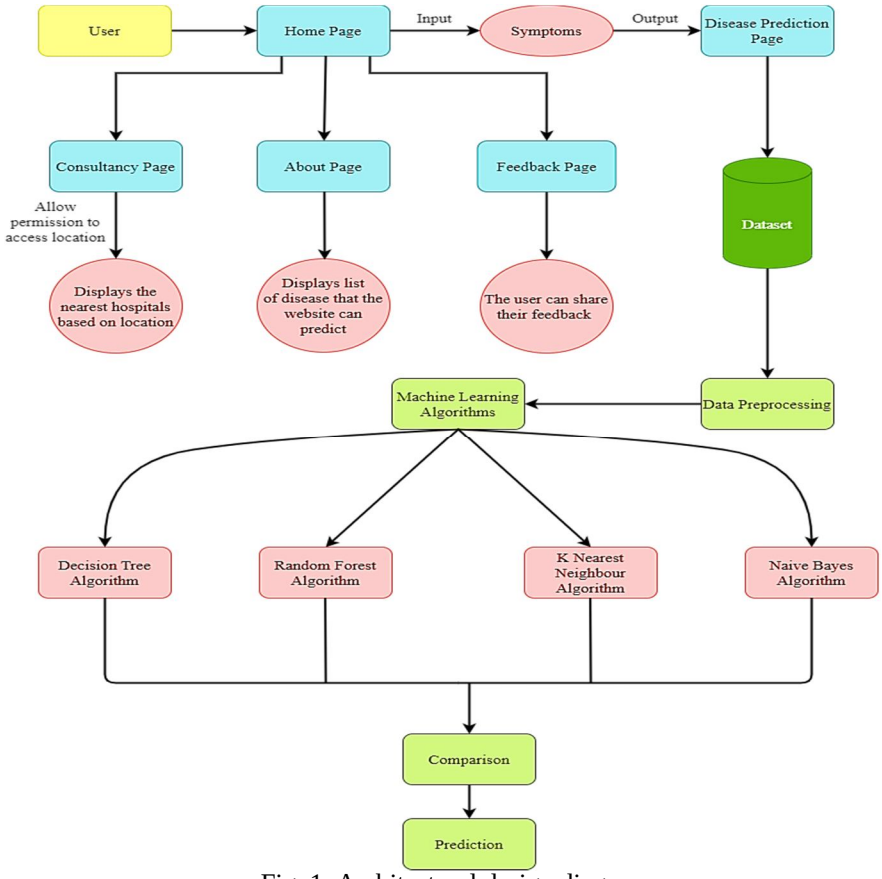


Fig. 1 Architectural design diagram

Fig. 1 illustrates architecture of the website where first the user will visit the Home page of the website and then go the Disease prediction where they will enter their symptoms. The four algorithms used for disease prediction include Decision Tree Algorithm, K-Nearest Neighbour Algorithm, Naïve Bayes Algorithm and Random Forest Algorithm. The user can also navigate from disease prediction page to the Consultancy page where on allowing his/her location they will be provided with list of all nearest hospitals. The user can also share their views with us through the feedback back. The about page gives a brief idea about the website to the user and also lists all the disease the website can predicted. All the pages are connected together with the help of the navbar. This will help the user to easily and efficiently navigate from one page to another.

The Frontend of the system would consist of a responsive website that can be accessed through any device. The website was developed using HTML, CSS and JavaScript and connected to the ML model using Flask. The user has to input the parameters for a specific disease and the model will detect if the disease is present or not, using four different algorithms which have significantly increased the accuracy rate of the system to 97%. They consist of Decision Tree algorithm, K-nearest neighbour algorithm, Naïve Bayes algorithm and Random Forest tree algorithm.

B. Modules Used

The website contains five pages i.e., Home Page, Disease Prediction Page, Consultant Page, About Page and Feedback Page.

1) *Home Page:* The user will see navbar in this area which contains the following buttons, "Home, Prediction, Consultancy, About, Feedback". These buttons are used to direct the user to the respective page.
2) *Prediction page:* The patient will be able to enter min 3 and max 5 symptoms and after clicking on submit, the four algorithms working in the backend will display the result.
3) *Consultancy page:* Based on the location of the user nearby hospitals will be recommended.
4) *About page:* Brief description about the website including the list of diseases that can be predicted by the website.
5) *Feedback page*: The user can give us their feedback through the feedback page.

C. Results

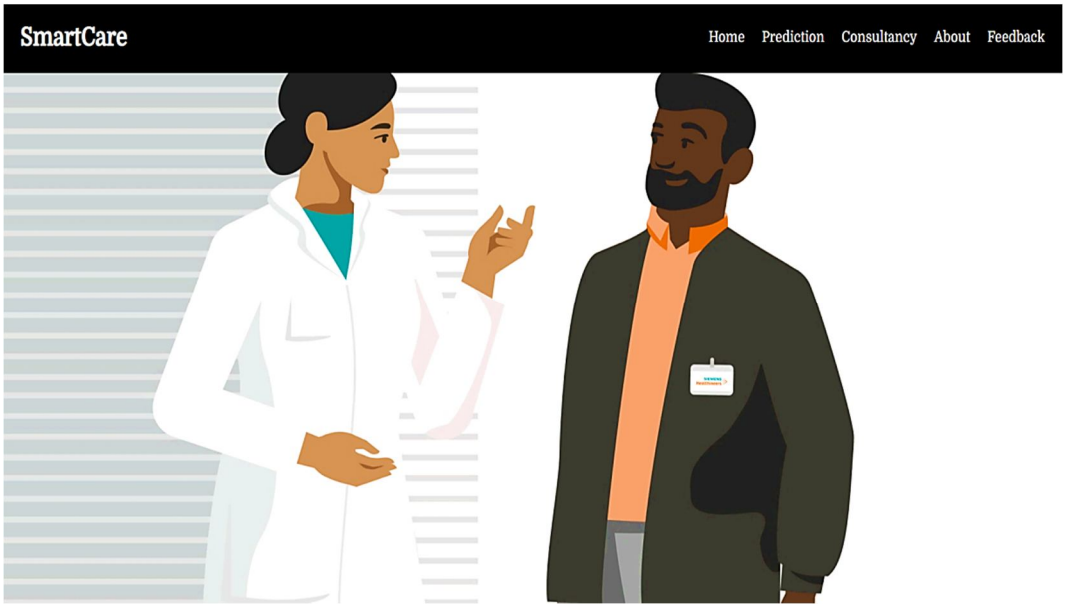


Fig. 2 Home Page

Fig. 2 depicts the Home Page of the system from which they can travel to any page of the website using the navbar that include the Prediction, Consultancy, About and Feedback.

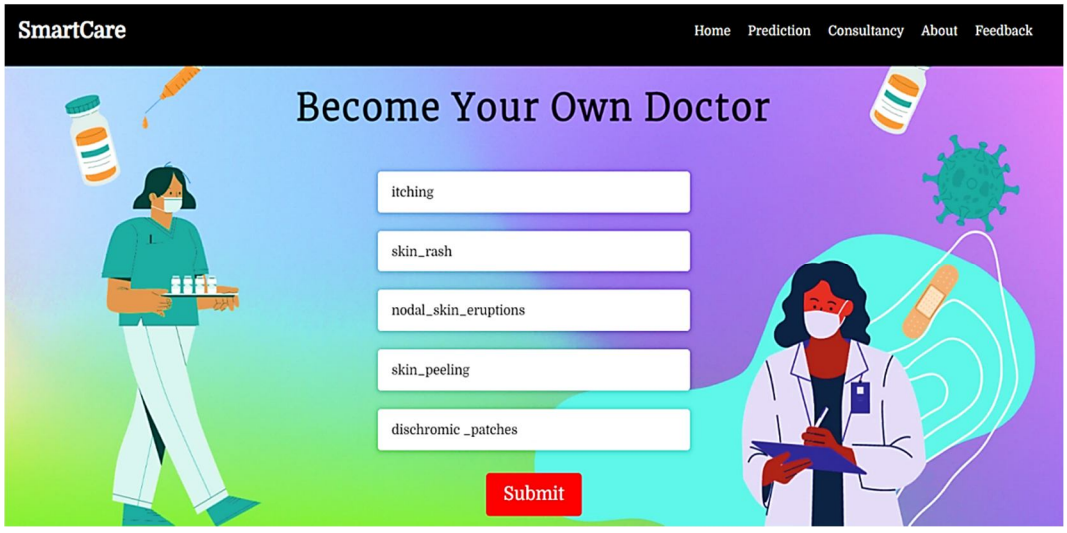


Fig. 3 Symptoms Page where the user has to input symptoms

Fig. 4 Disease Prediction Page

Fig. 4 represents the Prediction page where after giving five symptoms as input the predicted disease will be displayed with the help of four algorithms that consist of the Naive Bayer's, Decision Tree, K Nearest Neighbour and the Random Forest Algorithm.

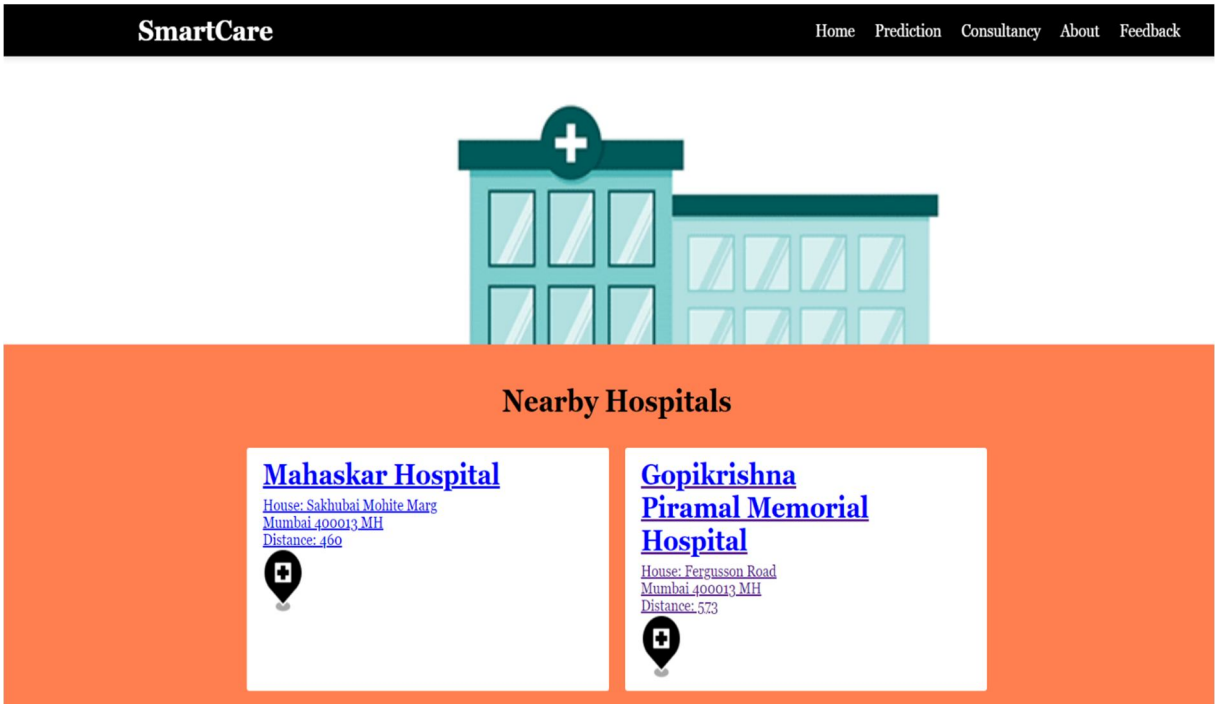


Fig. 5 Consultancy Page

Fig. 5 shows the Consultancy page where after allowing access to the user's current location, the GPS will track and display all the nearby hospitals immediately.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue XI Nov 2022- Available at www.ijraset.com

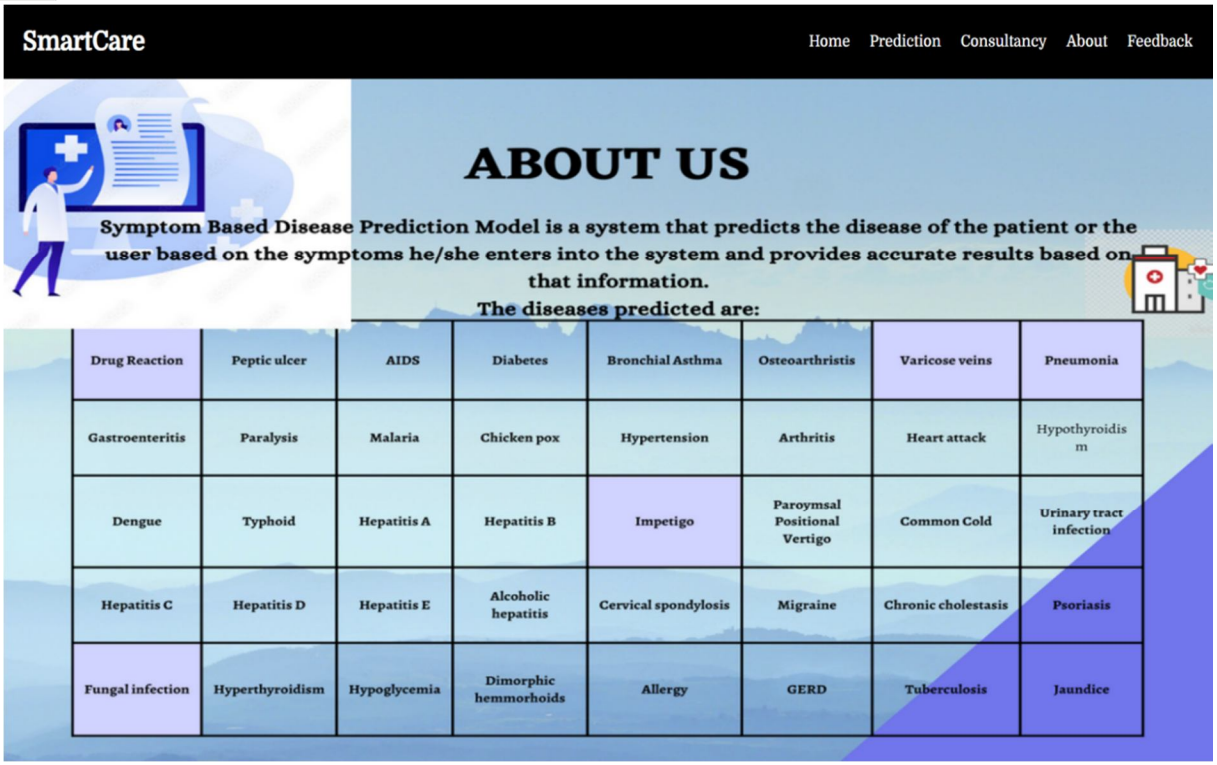


Fig. 6 About Us Page

Fig. 6 depicts the About page that gives the user a brief idea about our website and also displays the list of diseases that can be predicted by the system.

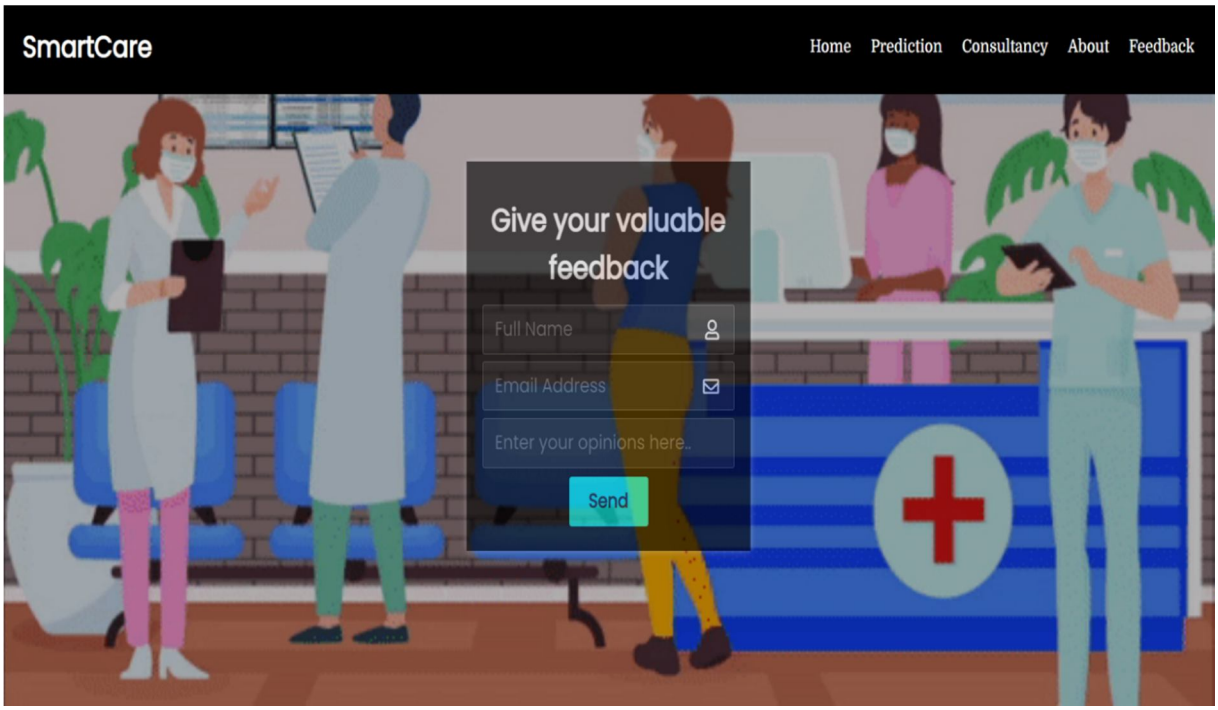


Fig. 7 Feedback Page

Fig. 7 represents the Feedback page where the user can share their views, comments or feedback.

VII. CONCLUSION

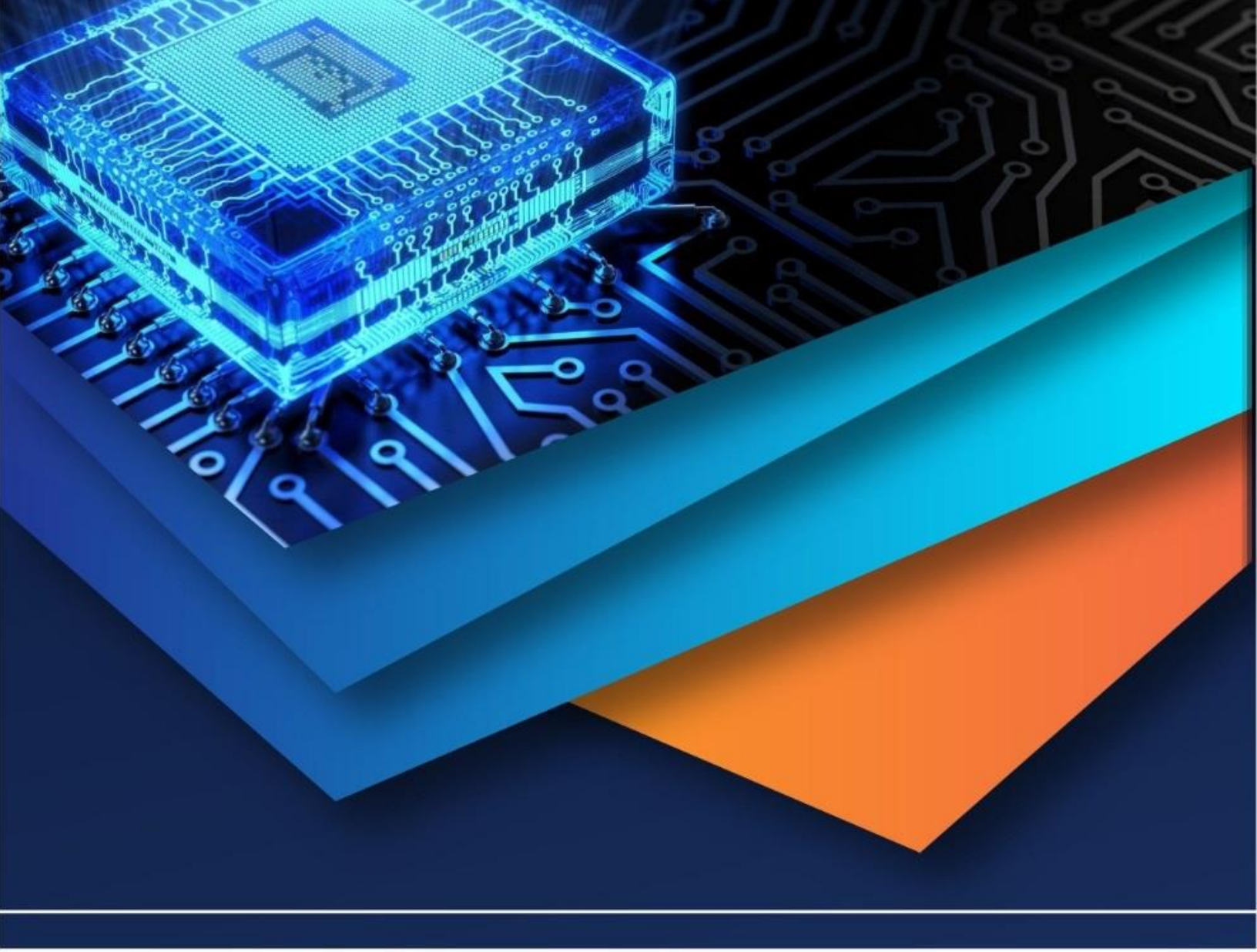
Making predictions from data is a strong use of machine learning. But it's crucial to keep in mind that machine learning is only as effective as the data used to train the algorithms. The website has been created in such a way that using it will be simple and easy for users. The project successfully implemented a website that could predict a disease with a 97% accuracy rate after comparing the four algorithms. Along with disease prediction, the website includes an about page, consultancy page and a feedback page where users can provide valuable feedback. In terms of future work, we intend to store the data of the users and use that information in the existing dataset and work on increasing the accuracy rate as well as include a greater number of diseases which the model can predict, generation of report and include services like virtual doctor appointments and online medicine shopping.

VIII. ACKNOWLEDGEMENT

We would like to express our sincere gratitude to our teacher Asst. Prof. Nilambari Narkar, our principal Dr. Y.D. Venkatesh, and our department head Dr. Vaishali Gaikwad for providing us the golden opportunity to do this wonderful project on the topic. The research carried out for the project helped us learn a lot of things and it also gave us practical experience with machine learning. We are really grateful to them.

REFERENCES

[1] studocu. What is a decision tree?
[2] Board Infinity. Explain random forest algorithm in machine learning?
[3] IBM. What is the k-nearest neighbors algorithm?
[4] Reshma. Classification using naive bayes.
[5] Akkem Yaganteeswarudu. Multi disease prediction model by using machine learning and flask api. In 2020 5th International Conference on Communication and Electronics Systems (ICCES), pages 1242–1246. IEEE, 2020.
[6] P Hamsagayathri and S Vigneshwaran. Symptoms based disease prediction using machine learning techniques. In 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), pages 747–752. IEEE, 2021.
[7] Archana Singh and Rakesh Kumar. Heart disease prediction using machine learning algorithms. In 2020 international conference on electrical and electronics engineering (ICE3), pages 452–457. IEEE, 2020.
[8] Yanli Liu, Yourong Wang, and Jian Zhang. New machine learning algorithm: Random forest. In International Conference on Information Computing and Applications, pages 246–252. Springer, 2012.
[9] Daniel Berrar. Bayes' theorem and naive bayes classifier. Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics, 403, 2018.
[10] Bahzad Charbuty and Adnan Abdulazeez. Classification based on decision tree algorithm for machine learning. Journal of Applied Science and Technology Trends, 2(01):20–28, 2021.
[11] Gongde Guo, Hui Wang, David Bell, Yaxin Bi, and Kieran Greer. Knn model-based approach in classification. In OTM Confederated International Conferences" On the Move to Meaningful Internet Systems", pages 986–996. Springer, 2003.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY