**Smart Healthcare Prediction Implemented With Machine Learning**

# Prof. Shegar S.R1, Pingat Pravin2, Hadawale Niraj3,Yendhe Bhavesh4

1,2,3,4, Department of Computer Engineering, Samarth Group Of Institutions College of Engineering, Belhe, India, 412410.

**Abstract:**

Based on the symptoms that users enter into the system, the "Smart Health Prediction Using Machine Learning" system use predictive modeling to forecast the illness of users or patients. There are three ways to log onto the application: admin, doctor, and user/patient. depending on an algorithmic prediction, the tool determines the probability of the condition depending on the symptoms entered by the user or patient. The Nave Bayes Classifier is used to produce intelligent health predictions. The Nave Bayes Classifier uses every characteristic it learned during the training phase to calculate the sickness% likelihood. Accurate interpretation of disease data gives users and patients a clear picture of the condition and aids in early disease prediction.

**Keywords:** Symptoms, Naïve Bayes, Machine Learning and Prediction Analysis.

**1. Introduction:**

The generative process of machine learning creates prediction models based on specific instances. It expands on the notion that computer programs can search for patterns in data, learn from them, and make deductions with little help from people. Using test data or previously acquired data, machine learning is a programming technique that accurately optimizes outcomes.

Phases one and two of the machine learning algorithm include investigation and preparation. Users' or patients' sign and symptom diaries are used to forecast the rate of illness and health. Machine learning provides a multitude of applications in the medical field to address health disease prediction, based on user/patient experience. We monitor all illnesses and symptoms using machine learning. Models of prediction examined with the aid of technology for machine learning. Technology improves patient care by empowering users/patients to decide for themselves whether to seek medical attention for certain symptoms. The large amount of data that was acquired is examined using the Gaussian Nave Bayes classifier. For every sector of disease prediction, the administrative, clinical, academic, and educational components of disease prediction from symptoms were also illustrated. There are multiple approaches to gathering data.

There are several resources available to prepare for smart health. However, research has focused on severe disease, and a risk threshold has been determined. These approaches are not commonly used to predict disease in general.

using a method for machine learning algorithms that is ideal for using to analyze patient symptoms in order to diagnose a variety of illnesses. Both of them time and resources are saved.

**2. Literature Survey:**

Several machine learning techniques and algorithms have been used in a number of research on the subject of illness prediction, which can be used by medical institutions.

A machine learning algorithm-based health prediction system has been proposed by Naveen Kumar and his colleague. He employed a variety of algorithms, including Decision Tree, K-Nearest Neighbor, and Naive Bayes, to predict the system. The accuracy of this suggested technique was 94%[1].Aditi Gavhane and her colleague proposed using machine learning to predict cardiac disease. In this system, the multi-layer perceptron model is employed. Based on common symptoms like age, sex, pulse rate, etc., this algorithm makes predictions about heart disease. This suggested technique has a 91% accuracy rate [2].The naive bayes algorithm is used in the heart prediction system that Gupta A. and his colleague devised. The naïve bayes algorithm has a 93% accuracy rate.Based on the basic symptoms of shortness of breath, jaw pain, neck pain, and chest pain, this approach makes predictions about heart problems [3].D Dahiwade and his colleague employed machine learning techniques, such as CNN and KNN, to create a model for disease prediction. This study proposes disease prediction, that is, diagnosis based on patient symptoms. CNN has a 92% accuracy rate, while KNN has a 94% accuracy rate[4].Using data mining techniques, N. Shabaz Ali and his colleague created a model for disease prediction. This research describes how machine learning and data mining approaches are combined to forecast diseases based on user symptoms[5].A mechanism that Shubham Salunke and his associate have created for using the naive bayes algorithm to diagnose the illnesses that patients are experiencing[6].IoT and machine learning are used by H. Pandey and his colleague to monitor healthcare. Smart health monitoring systems have been made possible by the Internet of Things (IoT). These health monitoring systems are able to monitor an individual's stress, anxiety, and hypertension levels in order to track their overall physical and mental well-being[7].Using data mining techniques, Monika Gandhi and her colleague have devised a methodology for predicting heart disorders. This work uses algorithms to examine data mining techniques, such as Naive Bayes, Neural Networks, and Decision Tree Algorithms, on medical data sets[8].Innovation in healthcare has been made possible by S. Ananth and his colleague's suggested IoT-based system for health prediction. The precision around 82% of the system[9].A model for the prognosis system for health and disease was proposed by S.S. and his colleague. In order to forecast diseases more accurately and efficiently, several machine learning methods are examined, including Random Forest, Naive Bayes, Logistic Regression, Support Vector Machine, K-Nearest Neighbors, Decision Tree, and Gradient Boosting. For illness prediction, the most accurate model is kept for later. Particularly, this method is employed in early disease prediction [10].A method for predicting numerous diseases has been proposed by Rudra A. and his collaborator. The largest flaw in the model is the additional emergence of prescription medications and illness medication in this system. The system's accuracy is roughly 85%[11].M. Asia searched for a scalable solution with his a colleague who uses the Random Forest Algorithm to predict various diseases. This approach is compared to the Naïve-Bayes classifier, however Random Forest produces results that are 98% more accurate[12].A system titled disease prediction based on classification algorithm was created by Sneha R. and a collaborator. This illness prediction model is constructed using classification methods including Decision Tree, Support Vector Machine (SVM), K Nearest Neighbor (KNN), Random Forest, Logistic Regression, and Naive Bayes.

Because it has the best accuracy (97%), the Naive Bayes algorithm is used to forecast diseases[13].A health prediction system employing multilinear regression and support vector machines was built by Farooqui and his collaborator. Up to 87% accuracy is achieved in the output produced by the suggested system[14].A system for predicting health using several machine learning algorithms, such as logistic regression, random forest, and support vector machines, has been proposed by N. Kosarkar and his colleague. The suggested model has an accuracy of 82%[15].

1. **Project Analysis :** 
   1. **Existing Method :**

The model forecasts chronic illnesses based on a given population and geographic region. Only particular ailments are included in the disease prediction. This method forecasts the probability of acquiring an illness by using convolutional neural networks and massive amounts of data. For S-type data, the method uses machine learning techniques like Decision Trees and K-nearest neighbors. Under certain circumstances, the accuracy rate of the approach can reach 94.8%. In a previous work, we examined machine learning techniques to forecast the occurrence of chronic illness epidemics in susceptible populations. We are leveraging real hospital data from a few selected regions/areas to create improved prediction models. We present a novel multimodal structured and unstructured sickness risk prediction method for convolutional neural networks, based on patient/user input.

* 1. **Proposed Method:**

After being diagnosed with a disease, a person needs to see a doctor or physician, which is costly and time-consuming. At times, users may find it extremely difficult to visit the hospital, which makes it difficult to make an independent diagnosis of sickness. On the other hand, the patient would gain a great deal and the user and patient would save time and resources if the previously indicated procedure were performed using an electronic software application. Everything about the surgery would proceed smoothly. The "Smart Health Care Prediction" web application predicts a user's illness depending on the symptoms the user or patient encounters. The Smart Health Prediction Framework data set was collected from multiple health-related websites. Considering the stated symptoms The user would be able to predict when the ailment would manifest itself on the web application. The objective of this project is to develop an online platform that can predict disease episodes based on a range of symptoms. Users can choose from a range of symptoms and diagnose diseases by using probabilistic estimations and conditions.

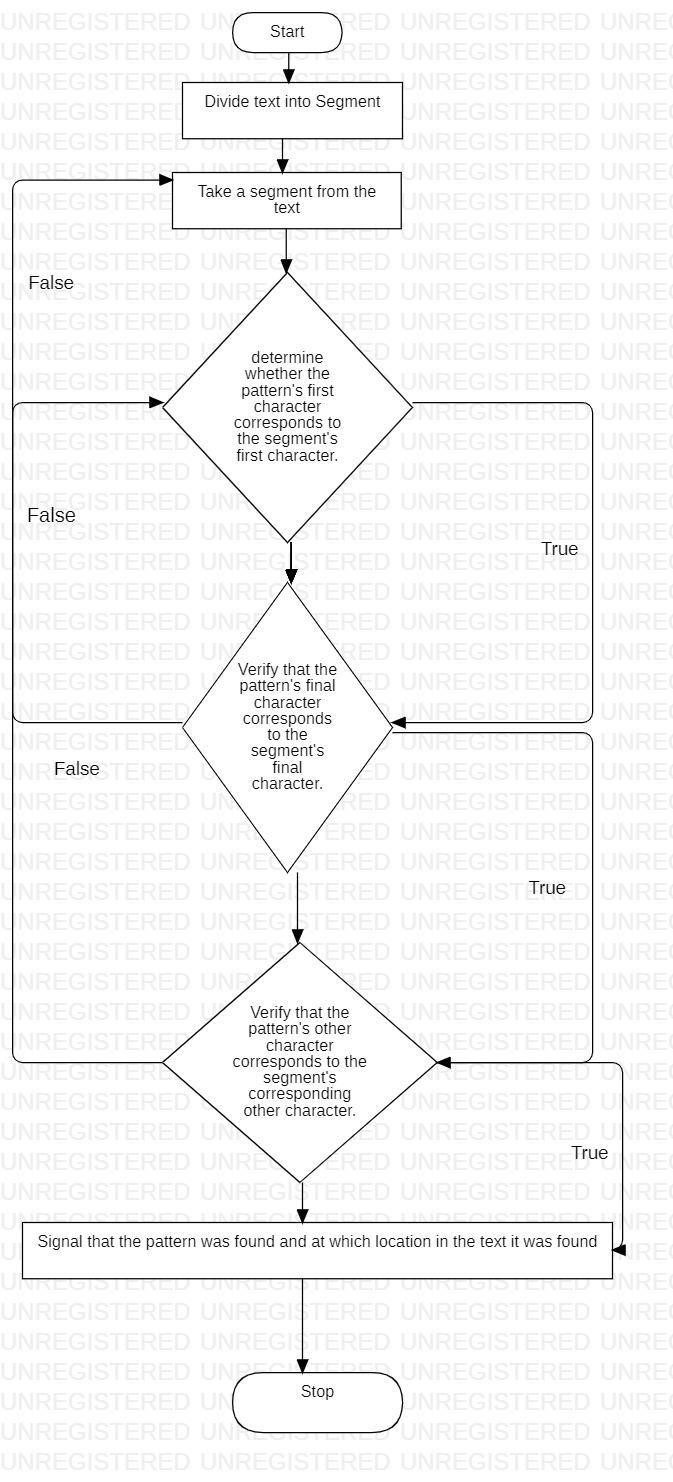
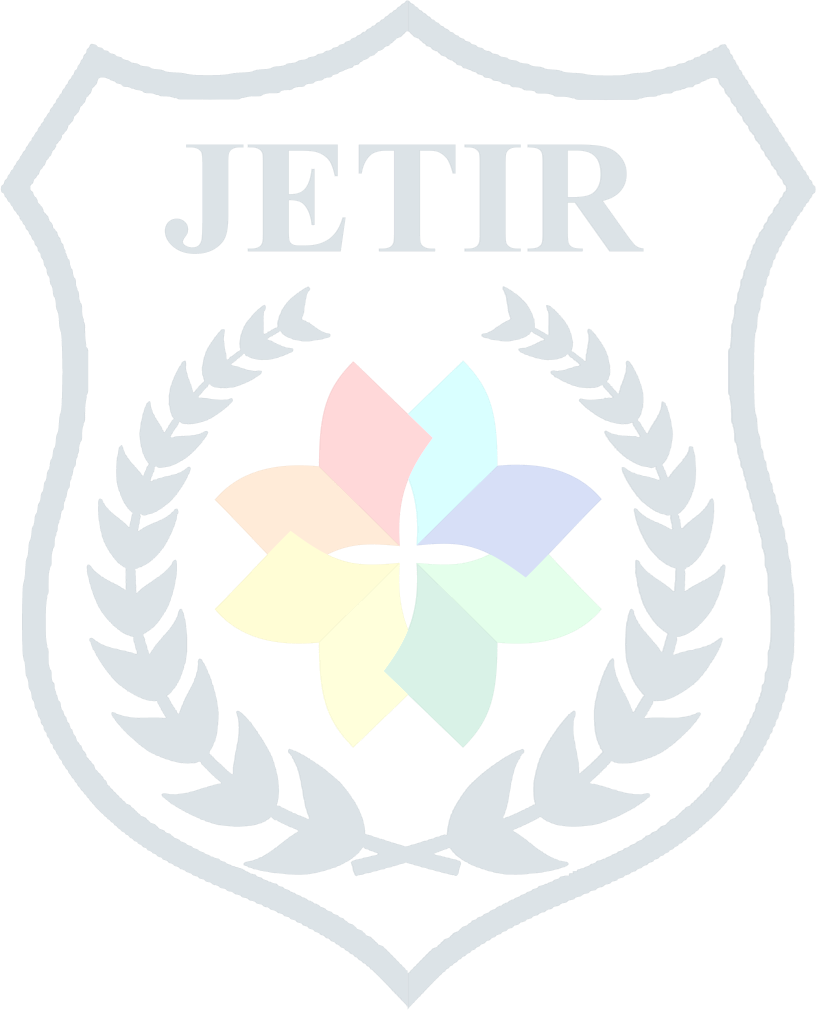
**Efficiency comparison TABLE**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Diseases** | **DT** | **RF** | **SVC** | **NB** |
| **Breast Cancer Wise** | **86.25** | **93.45** | **96.48** | **97.18** |
| **Breast Cancer** | **75.36** | **69.48** | **74.36** | **77.83** |
| **Dermatomegaly** | **88.36** | **97.59** | **97.25** | **98.13** |
| **Echo Chambers** | **75.61** | **95.69** | **95.98** | **96.10** |
| **Liveries** | **57.12** | **69.48** | **71.29** | **72.39** |
| **Pimaricin Diabetes** | **73.89** | **75.98** | **76.27** | **77.19** |
| **Hematidrosis** | **72.39** | **85.65** | **79.45** | **81.34** |
| **Heart-statlog** | **73.69** | **85.69** | **74.56** | **81.59** |
| **Heart-b** | **75.74** | **84.58** | **78.85** | **81.29** |
| **Hepatitis** | **82.98** | **84.25** | **80.25** | **79.85** |
| **Lung Cancer** | **54.28** | **49.85** | **45.69** | **41.08** |
| **Lymph’s** | **66.94** | **79.25** | **84.29** | **79.68** |
| **Osteoporosis** | **69.38** | **63.84** | **65.89** | **82.36** |
| **Tumor** | **38.84** | **43.65** | **39.71** | **43.58** |

We looked at the accuracy of a few of the approaches in the above table, and the results show that the Naive Bayes algorithm performs better than the others in terms of timing and accuracy. We applied the naïve bayes algorithm on patient data. Medical data is growing at an exponential rate, and in order to predict the precise disease from its symptoms, analysis is required. We were able to precisely anticipate the likelihood of acquiring a general illness as an output using a user/patient record as the input, which allowed us to comprehend the degree of disease risk prediction. Thanks to technology, it is now possible to forecast risk factors and diseases in a rapid, easy, and economical way.

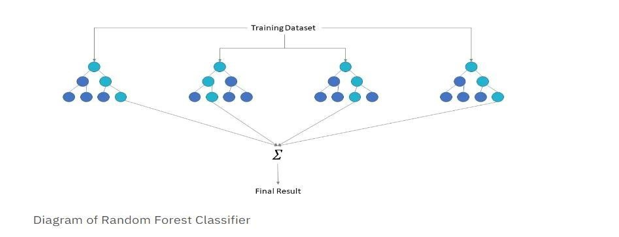
1. **Algorithm and Architecture** 
   1. **Naïve Bayes Algorithm**

A simple dynamic technique for creating models to categorize issue instance occurrences and map them to objects is the Naive Bayes algorithm. Class names are chosen from a restricted set of possibilities. It is a collection of algorithms rather than a single technique, based on a general idea. According to this theory, the function value of any Naive Bayes Classifier is unrelated to the value of any other feature. For example, a fruit that is round, orange, and has a diameter of 10 to 15 cm could be called an orange. The Naive Bayes algorithm takes into account each trait to determine whether the fruit is an orange or not. There are n distinct probability models; nonetheless, the performance of supervised learning for some of themof the Naive Bayes method is the best.



* 1. **Random Forest Algorithm:**

The basis of random forest is the bagging method, which is a kind of ensemble learning strategy. The random forest algorithm mixes the output of the decision trees and uses several of them. Both the overfitting problem and accuracy can be avoided by doing this. A forest is composed of many distinct types of trees; the more species present, the more diverse the forest's composition.

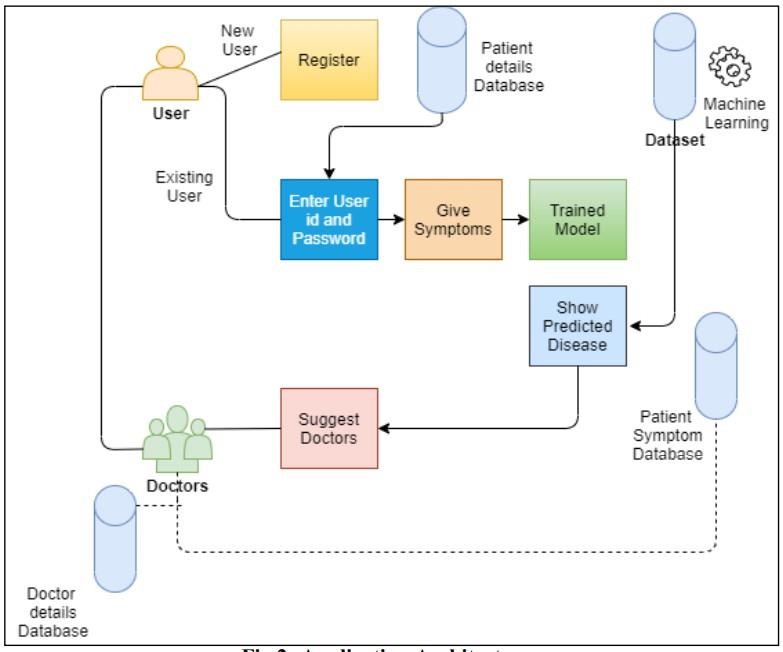


**4.3. Support Vector Classifier- SVC:**

• One of the most popular supervised learning algorithms, Support Vector Machine, or SVM, is used to solve Classification and Regression issues. However, machine learning classification problems are where it is mostly applied.

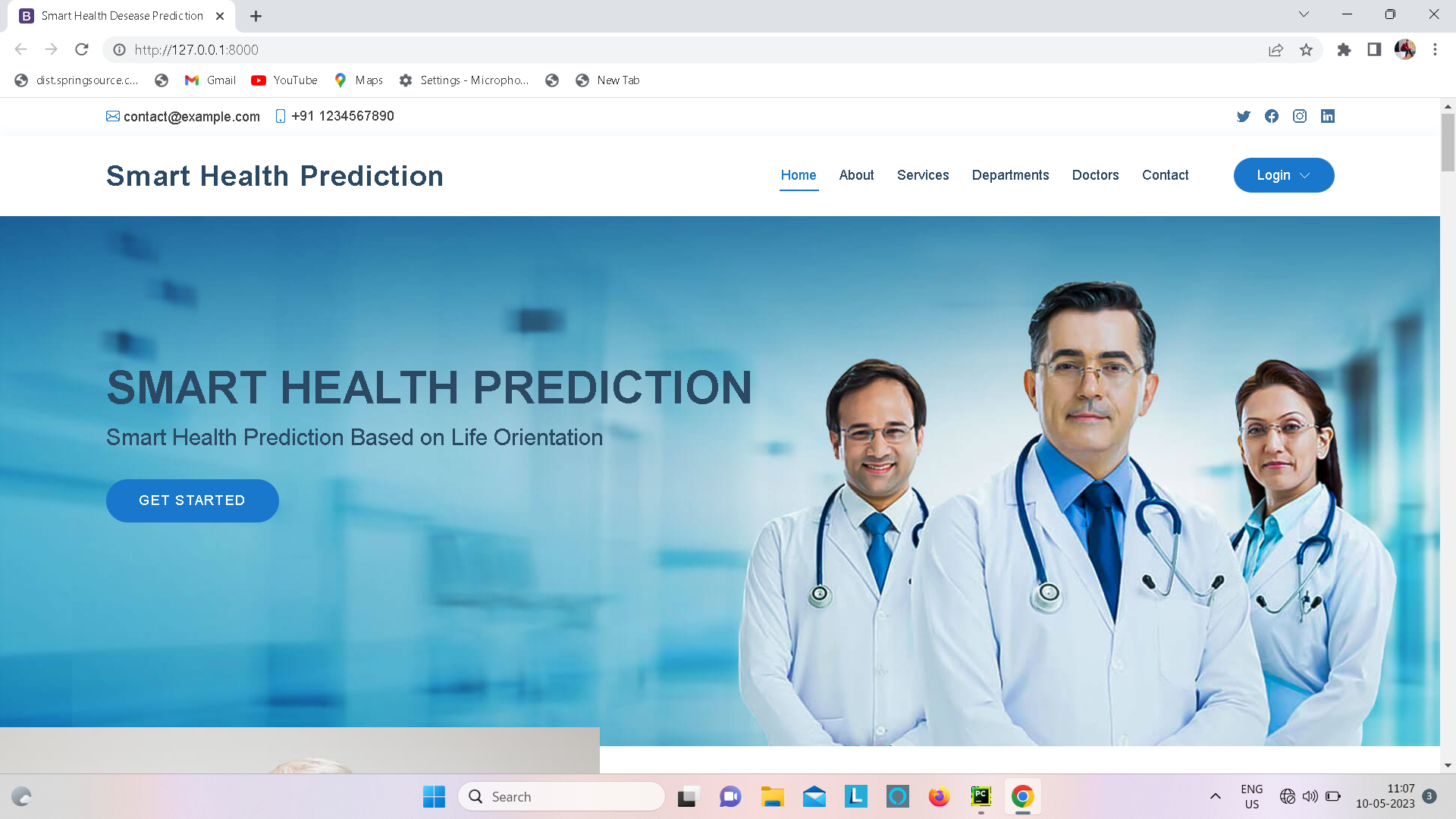
• The SVM method seeks to identify the best line or decision boundary that can divide n-dimensional space into classes in order to quickly categorize fresh data points in the future. This ideal decision boundary is known as a hyperplane.

**Architecture:**



1. **Outputs:**

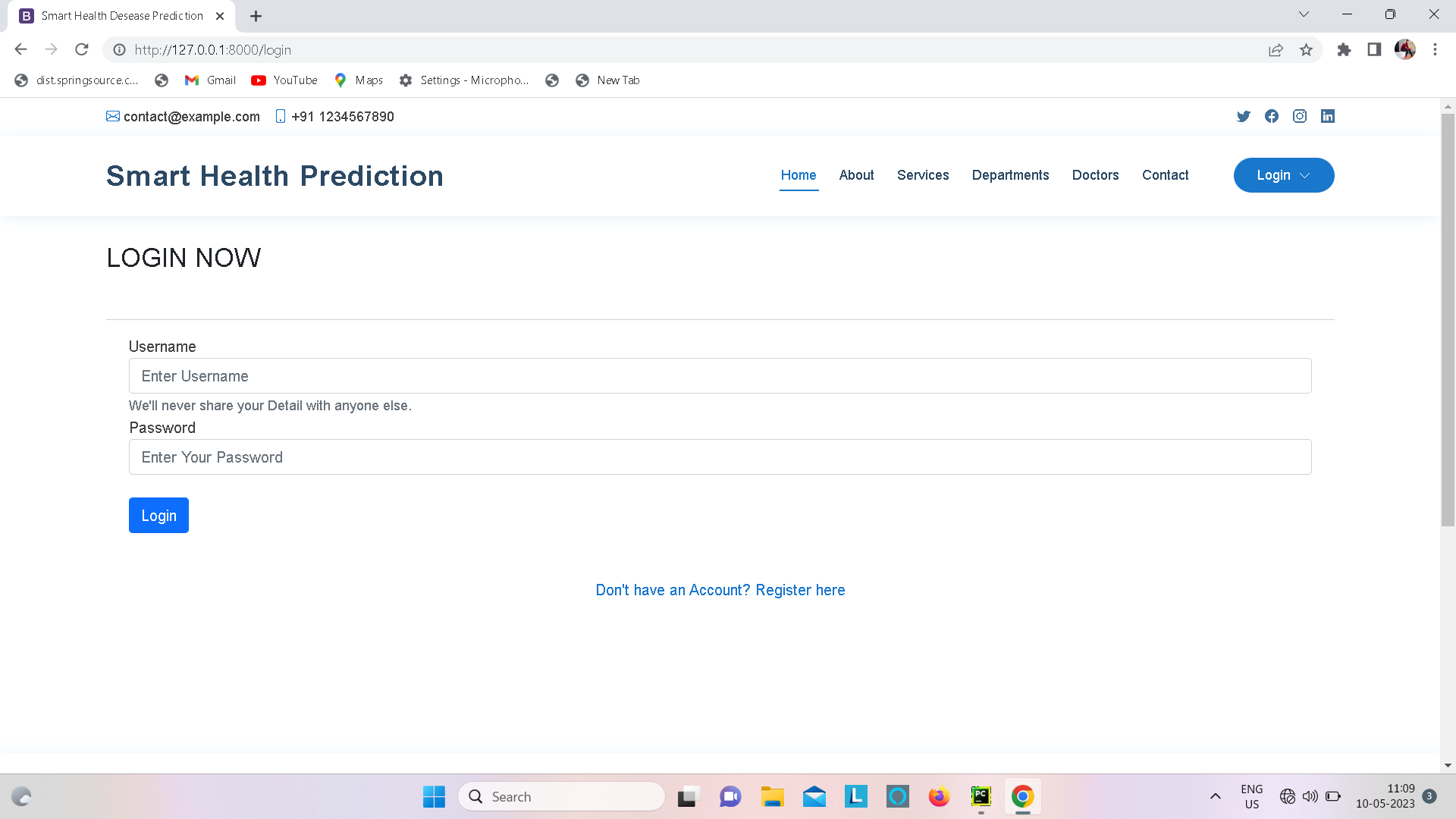
**5.1: Home Page**

****

**5.2:User Registration Page**

****

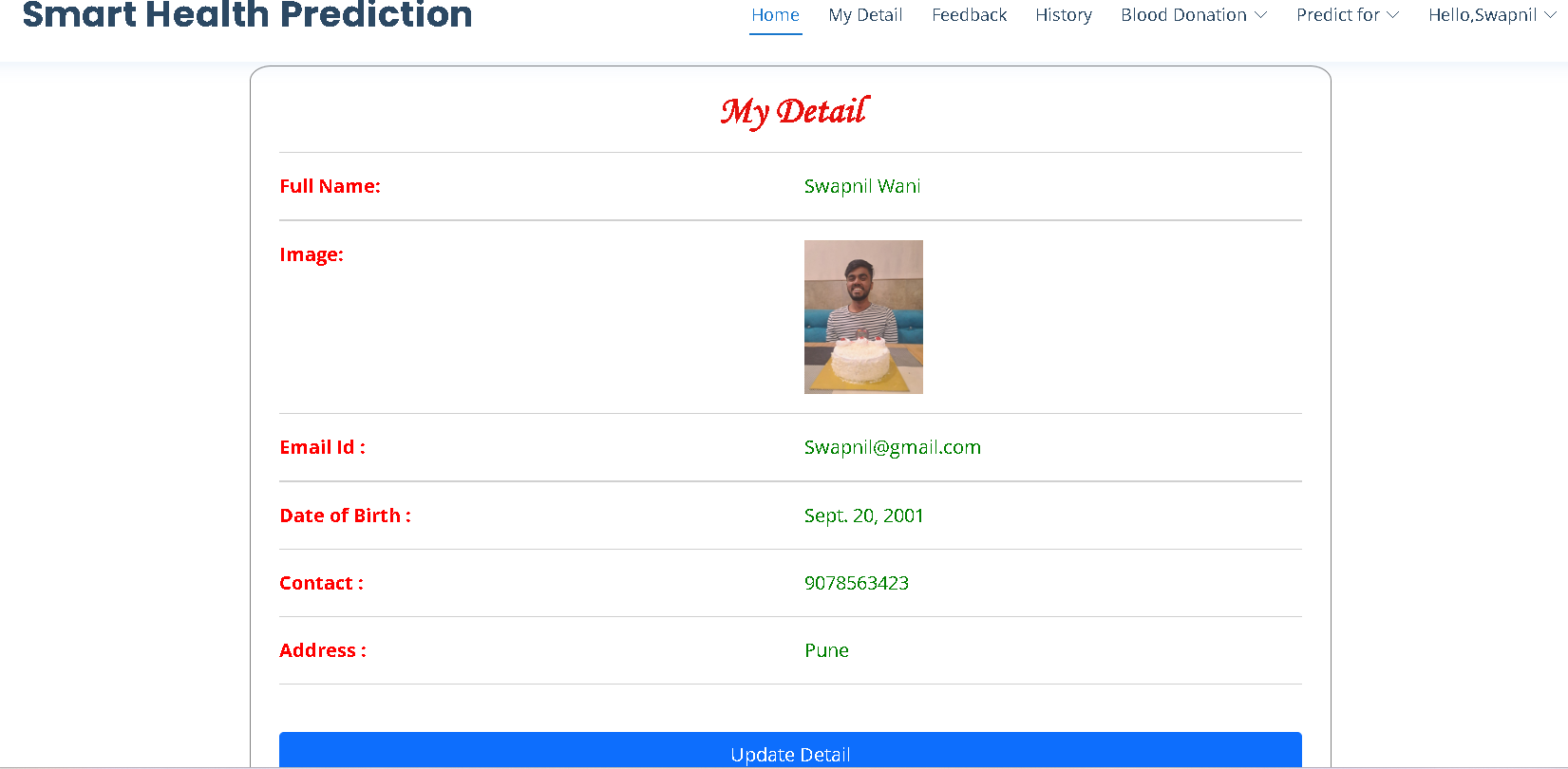
**5.3:Login Page**

****

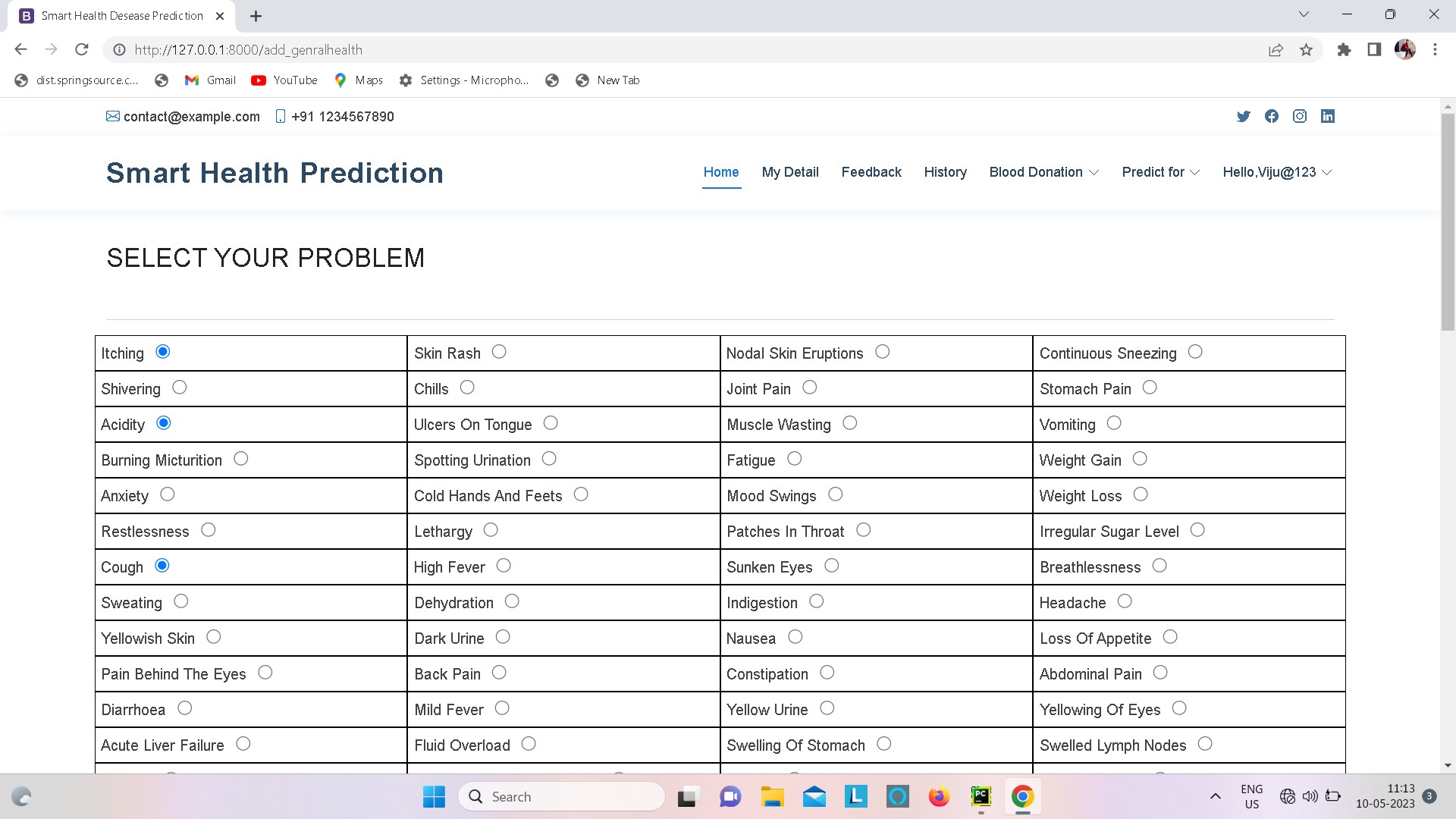
**5.4:Patient Home Page**

****

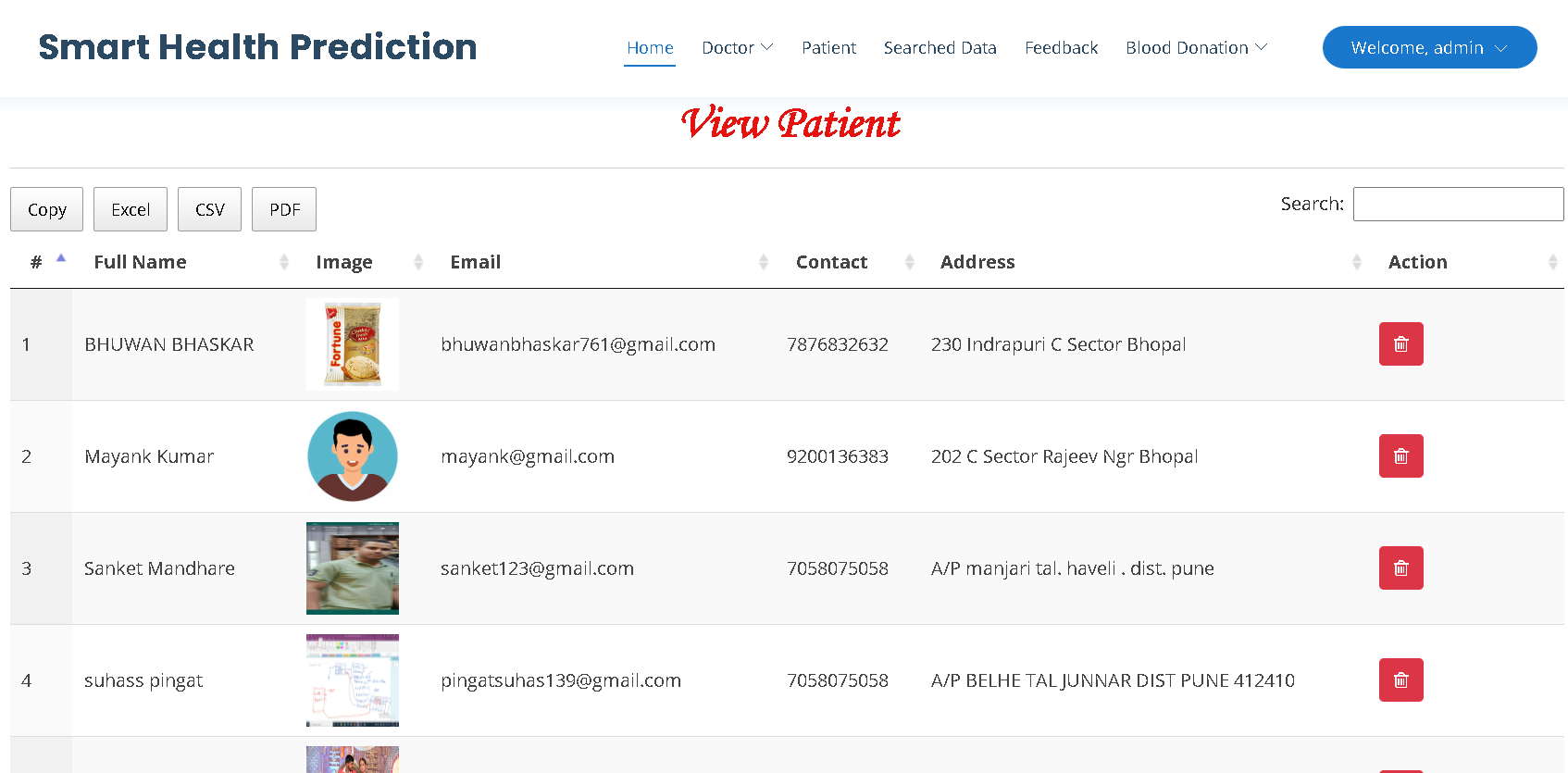
**5.5: User Profile Page**

****

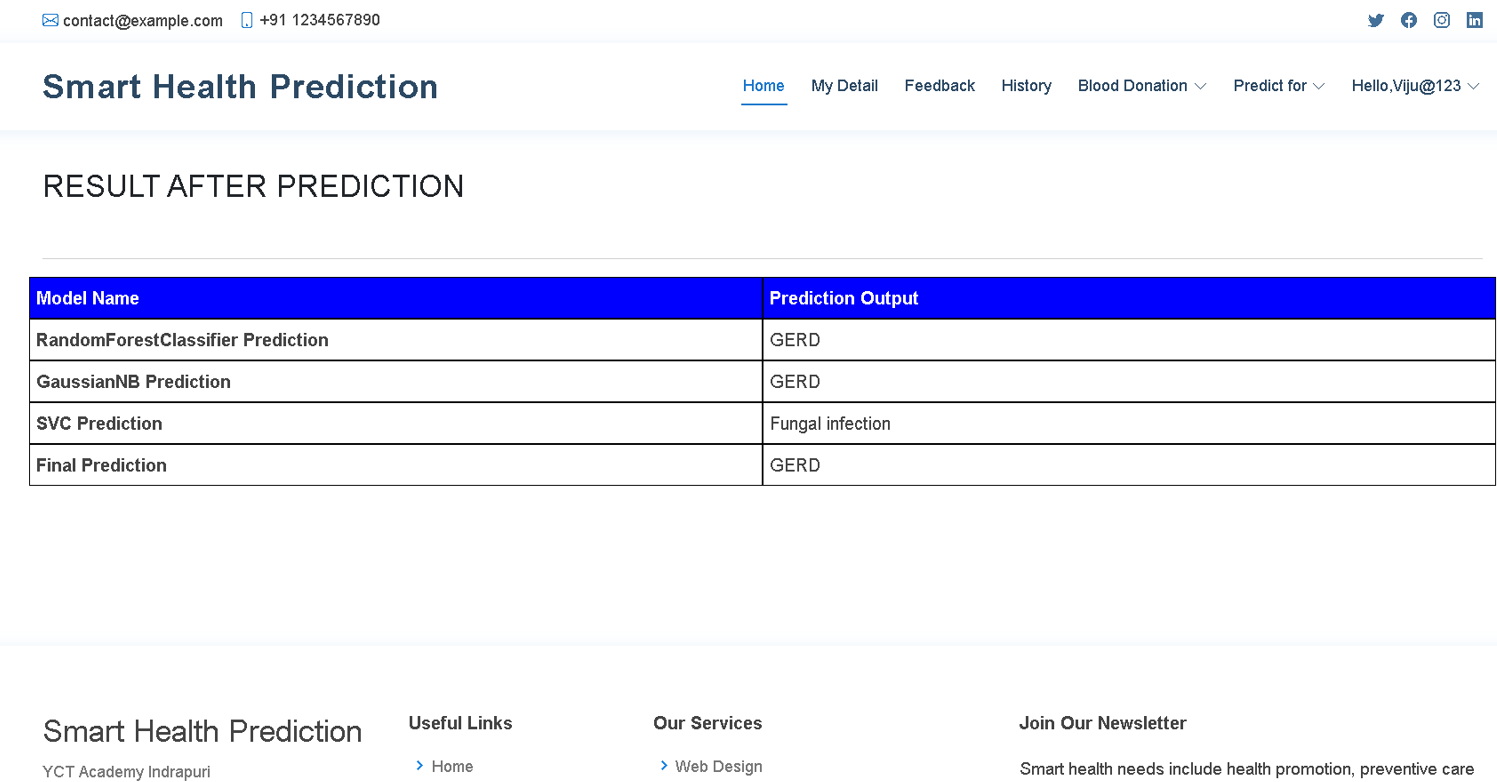
**5.6:Input User Health Parameter Page**

****

**5.7: View Searched Data Page**

****

**5.8:Result Page**

****

1. **Conclusion:**

In order to improve the performance of handwritten digit recognition, we looked at convolutional neural network modifications in this work that do away with complex ensemble (classifier combination) methods, costly feature extraction, and extensive pre-processing of a standard recognition system. We were able to obtain a 99.89% recognition rate for the MNIST database using the Adam optimizer. The present study is distinct in that it conducts a thorough analysis of every CNN architecture configuration that yields the highest recognition accuracy for an MNIST dataset. Using ensemble CNN network topologies for the same dataset resulted in greater recognition accuracy, albeit at the cost of more computationally demanding testing.

**7. References:**

[1] Naveenkumar S1 , Kirubhakaran R2 , Jeeva G3 , Shobana M4 , Sangeetha K5,” Smart Health Prediction Using Machine Learning: A survey”,3 march 2021.

[2] Aditi Gavhane, Gouthami Kokkula, Isha Pandya and Kailas Devadkar, “Prediction of Heart Disease Using Machine Learning” IEEE Xplore ISBN: 978-1-5386-0965-1, pp. 1275-1278, 2018.

[3] Gupta A., Kumar L., Jain R., Nagrath P. (2020) Heart Disease Prediction Using Classification (Naive Bayes). In: Singh P., Pawłowski W., Tanwar S., Kumar N., Rodrigues J., Obaidat M. (eds) Proceedings of First International Conference on Computing, Communications, and Cyber-Security (IC4S 2019). Lecture Notes in Networks and Systems, vol 121. Springer, Singapore.

[4] D. Dahiwade, G. Patle and E. Meshram, "Designing Disease Prediction Model Using Machine Learning Approach," 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 2019, pp. 1211-1215

[5] N. Shabaz Ali, G. Divya,”Prediction of Diseases in Smart Health CareSystem using Machine Learning” 5 January 2020.

|  |  |
| --- | --- |
| Shubham Salunke, ShubhamRajiwade, Deepak Yadav, S.K.Sabnis, "smarthealth | |
| prediction system using machinelearning",March 2020. |  |

[6]

[7] H.Pandey and S. Prabha, "Smart Health Monitoring System using IOT and Machine Learning Techniques," 2020 Sixth International Conference on Bio Signals, Images, and Instrumentation (ICBSII), Chennai, India, 2020, pp. 1-4.

[8] Monika Gandhi and Shailendra Narayan Singh "Predictions in heart disease using techniques of data mining" 2015 International Conference on Futuristic Trends on Computational Analysis and Knowledge Management (ABLAZE) pp. 520-525 2015.

[9] S. Ananth, P. Sathya, P. Madhan Mohan,” Smart Health Monitoring System through IOT”,In 2019 International Conference on Communication and Signal Processing (ICCSP),IEEE April 2019.

[10] S. S, V. S and A. R, "Health and Disease Prognosis System using Machine Learning," 2021 IEEE International Conference on Intelligent Systems, Smart and Green Technologies (ICISSGT), Visakhapatnam, India, 2021, pp. 192-196.

[11] Rudra A. Godse Smita S. Gunjal Karan A. Jagtap Neha Mahamuni and Suchita Wankhade "Multiple Disease Prediction Using Different Machine Learning Algorithms Comparatively" International Journal of Advanced Research in Computer and Communication Engineering vol. 8 no. 12 Dec 2019.

[12] R. Delshi Howsalya Devi P. Sreevalli Keerthana Prathyusha and M. Asia "Prediction of Diseases Using Random Forest Classification Algorithm" Zeichen Journal vol. 6 no. 5 2020.

[13] Sneha R. Monisha S. Jahnavi C. and S. Nandini "Disease Prediction Based On Symptoms Using Classification Algorithm" Journal of Xi'an University of Architecture &amp; Technology vol. 12 no. 4 2020.

[14] Farooqui, Md. Ehtisham and Ahmad, Dr. Jameel, Disease Prediction System using Support Vector Machine and Multilinear Regression (August 13, 2020). International Journal of Innovative Research in Computer Science & Technology (IJIRCST) ISSN: 2347-5552, Volume, 8, Issue, 4, July, 2020.

[15] N. Kosarkar, P. Basuri, P. Karamore, P. Gawali, P. Badole and P. Jumle, "Disease Prediction using Machine Learning," 2022 10th International Conference on Emerging Trends in Engineering and Technology - Signal and Information Processing (ICETET-SIP-22), Nagpur, India, 2022, pp. 1-4.