

# Disease Predictor Application

Project Report Submitted in Partial Fulfilment of the Requirements for  
the Degree of

## **Bachelor of Engineering** *in* **Computer Science & Engineering**

*Submitted by*

Rahul Bairwa: (Roll No. 19UCSE4016)

Puru Soni: (Roll No. 19UCSE4013)

*Under the Mentorship of*  
Dr. Alok Singh Gehlot  
Assistant Professor

&

*Under the Guidance of*  
Dr. Shrawan Ram  
Assistant Professor



Department of Computer Science & Engineering  
MBM University, Jodhpur  
**July, 2022**

# Disease Predictor Application

Project Report Submitted in Partial Fulfilment of the Requirements for the  
Degree of

## **Bachelor of Engineering** *in* **Computer Science & Engineering**

*Submitted by*

Rahul Bairwa: (Roll No. 19UCSE4016)

Puru Soni: (Roll No. 19UCSE4013)

*Under the Mentorship of*

Dr. Alok Singh Gehlot  
Assistant Professor

*&*

*Under the Guidance of*

Dr. Shrawan Ram  
Assistant Professor



Department of Computer Science & Engineering  
MBM University, Jodhpur  
July, 2022



## Department of Computer Science & Engineering

M.B.M. University,  
Ratanada, Jodhpur, Rajasthan, India –342011

### CERTIFICATE

This is to certify that the work contained in this report entitled “**Disease Predictor Application**” is submitted by the group members Mr. Rahul Bairwa (Roll. No: 19UCSE4016) and Mr. Puru Soni, (Roll. No: 19UCSE4013) to the Department of Computer Science & Engineering, M.B.M. University, Jodhpur, for the partial fulfilment of the requirements for the degree of **Bachelor of Engineering in Computer Science & Engineering**.

They have carried out their work under my guidance. This work has not been submitted elsewhere for the award of any other degree or diploma.

The project work in our opinion, has reached the standard fulfilling the requirements for the degree of Bachelor of Engineering in Computer Science in accordance with the regulations of the Institute.

**Dr. Shrawan Ram**

Assistant professor

(Guide)

Dept. of Computer Science & Engg.

M.B.M. University, Jodhpur

**Dr. Alok Singh Gehlot**

Assistant professor

(Mentor)

Dept. of Computer Science & Engg.

M.B.M. University, Jodhpur

## DECLARATION

We, *Rahul Bairwa and Puru soni*, hereby declare that this project titled “*Disease Predictor Application*” is a record of original work done by us under the supervision and guidance of *Dr. Shrawan Ram*.

We further certify that this work has not formed the basis for the award of the Degree/Diploma/Associateship/Fellowship or similar recognition to any candidate of any university and no part of this report is reproduced as it is from any other source without appropriate reference and permission.

SIGNATURE OF STUDENT

**(Rahul Bairwa)**  
**8<sup>th</sup> Semester, CSE**  
Enroll. - 18R/06197  
Roll No. - 19UCSE4016

SIGNATURE OF STUDENT

**(Puru Soni)**  
**8<sup>th</sup> Semester, CSE**  
Enroll. - 18R/33061  
Roll No. - 19UCSE4013

## ACKNOWLEDGEMENT

We take the opportunity to express our gratitude to all who have provided their immense support and their valuable time in guiding us. It is due to their support, Guidance, Supervision and Encouragement that We have successfully completed our Project Report on “*Disease Predictor Application*”. We are highly indebted to our guide Dr. Shrawan Ram , Assistant Professor and mentor Dr. Alok Singh Gehlot, Assistant Professor, for their guidance and constant supervision as well as providing necessary information regarding this Project.

## **ABSTRACT**

This Disease Predictor Application project is implemented in the Flask framework. The main aim of this project is to create an online Disease predictor application . Using this application, users can predict his/her disease by selecting the symptoms He/She is experiencing.

DoctorX is a web application that is used to diagnose the disease as per the respective symptoms provided by the user. The model is built in a sequential neural network consisting of 131+ symptoms and 40+ diseases that the model can predict with 92 percent accuracy.

Due to the current situation of Covid-19, people are suffering from various kinds of disease. This project aims to target those people who are suffering from well-known diseases, then it will predict the disease accurately according to provided symptoms and the user can give treatment to himself according to the predicted disease.

With an increase in biomedical and healthcare data, accurate analysis of medical data benefits early disease detection and patient care. By using linear regression and decision tree we are predicting diseases like Diabetes, Malaria, Jaundice, Dengue, and Tuberculosis.

# Contents

<b>Chapter 1</b>	<b>1</b>
<b>Introduction</b>	<b>1</b>
1.1. Overall Description	1
1.2. Purpose	2
1.2.1. Functionalities	3
1.3. Motivation and Scope	3
1.3.1. Aim	3
1.3.2. Future Scope	4
<b>Chapter 2</b>	<b>5</b>
<b>Technology Used</b>	<b>5</b>
2.1. Flask Framework	5
2.1.1 Why is Flask a good web framework choice?	5
2.2. Git/GitHub	6
2.2.1 Why do we need Git ?	6
2.3. Python	7
2.4. Neural Network - Deep Learning	8
2.5. Amazon Web Services (AWS)	9
<b>Chapter 3</b>	<b>11</b>
<b>Work Done</b>	<b>11</b>
3.1. Overview	11
3.2. Components	11
3.2.1 SignUp	11
3.2.2 SignIN	12
3.2.3 Home	12
3.2.4 Predict	13

3.3. Architecture	14
3.3.1. Network Architecture	15
3.3.2. Confusion Matrix	16
3.4. Results	16
<b>Chapter 4</b>	<b>17</b>
<b>Conclusion &amp; Future Work</b>	<b>17</b>
4.1. Conclusion	17
4.2. Future Work	17
<b>References</b>	<b>19</b>



# List of Figures

<b>2.1 Flask .....</b>	<b>5</b>
<b>2.2 Github .....</b>	<b>6</b>
<b>2.3 Python .....</b>	<b>7</b>
<b>2.4 Neural networks .....</b>	<b>7</b>
<b>2.5 AWS .....</b>	<b>8</b>
<b>3.1 SignUp .....</b>	<b>10</b>
<b>3.2 SignIn .....</b>	<b>10</b>
<b>3.3 Home .....</b>	<b>11</b>
<b>3.4 (a) Predict Page .....</b>	<b>12</b>
<b>3.4 (b) Predict Page .....</b>	<b>12</b>

# Chapter 1

## Introduction

Machine Learning is the domain that uses past data for predicting. Machine Learning is the understanding of computer systems under which the Machine Learning model learns from data and experience. The machine learning algorithm has two phases: 1) Training & 2) Testing. To predict the disease from a patient's symptoms and from the history of the patient, machine learning technology has struggled from past decades. Healthcare issues can be solved efficiently by using Machine Learning Technology. We are applying complete machine learning concepts to keep track of patient's health.

### 1.1. Overall Description

Disease prediction using patient treatment history and health data by applying data mining and machine learning techniques has been an ongoing struggle for the past decades. Many works have been applied to data mining techniques to pathological data or medical profiles for prediction of specific diseases. These approaches tried to predict the reoccurrence of disease. Also, some approaches try to do prediction on control and progression of disease. The recent success of deep learning in disparate areas of machine learning has driven a shift towards machine learning models that can learn rich, hierarchical representations of raw data with little pre-processing and produce more accurate results. With the development of big data technology, more attention has been paid to disease prediction from the perspective of big data analysis; various researches have been conducted by selecting the characteristics automatically from a large amount of data to improve the accuracy of risk classification rather than the previously selected characteristics.

ML model allows us to build models to get quickly cleaned and processed data and deliver results faster. By using this system doctors will make good decisions related to patient diagnoses and according to that, good treatment will be given to the patient, which increases improvement in patient healthcare services. To introduce machine learning in the medical field, healthcare is the prime example. To improve the accuracy of large data, the existing work will be done on unstructured or textual data. For the prediction of diseases, the existing will be done on linear, KNN, Decision Tree algorithm.

## **1.2. Purpose**

The main focus is on using machine learning in healthcare to supplement patient care for better results. Machine learning has made it easier to identify different diseases and diagnose correctly. Predictive analysis with the help of efficient multiple machine learning algorithms helps to predict the disease more correctly and help treat patients.

The healthcare industry produces large amounts of healthcare data daily that can be used to extract information for predicting disease that can happen to a patient in future while using the treatment history and health data. This hidden information in the healthcare data will be later used for effective decision making for patient's health. Also, these areas need improvement by using informative data in healthcare.

One such implementation of machine learning algorithms is in the field of healthcare. Medical facilities need to be advanced so that better decisions for patient diagnosis and treatment options can be made. Machine learning in healthcare aids the humans to process huge and complex medical datasets and then analyze them into clinical insights. This then can further be used by physicians in providing medical care. Hence machine learning when implemented in healthcare can lead to increased patient satisfaction.

### **1.2.1. Functionalities**

Functionalities provided by the Disease Predictor System are as follows:

- At the starting of WebApp User can login/signup to the portal
- The user will provide their details which are symptoms
- Then the model will predict the disease according to symptoms and show to the user his/her basic information
- These symptoms will be stored in the MySQL database for future reference.

## **1.3. Motivation and Scope**

In this Project we had created a Disease Predictor App which asks users about the symptoms he/she is experiencing on the basis of this our App will predict the disease the user might have with 92% Accuracy. In this App he used Sequential Neural Network Model to predict the Disease and this Model is Trained for 131 symptoms and can predict 40+ Disease with 92% Accuracy and created an Flask App on AWS Server.

### **1.3.1. Aim**

Our project aims at Business process automation, i.e. we have tried to computerize various processes of the Online Disease Predictor System.

- To utilize resources in an efficient manner by increasing their productivity through automation.
- It satisfies the user requirement.
- Be easy to understand by the user and operator
- Be easy to operate.
- Have a good user interface
- Be expandable
- Delivered on schedule within the budget.

### **1.3.2. Future Scope**

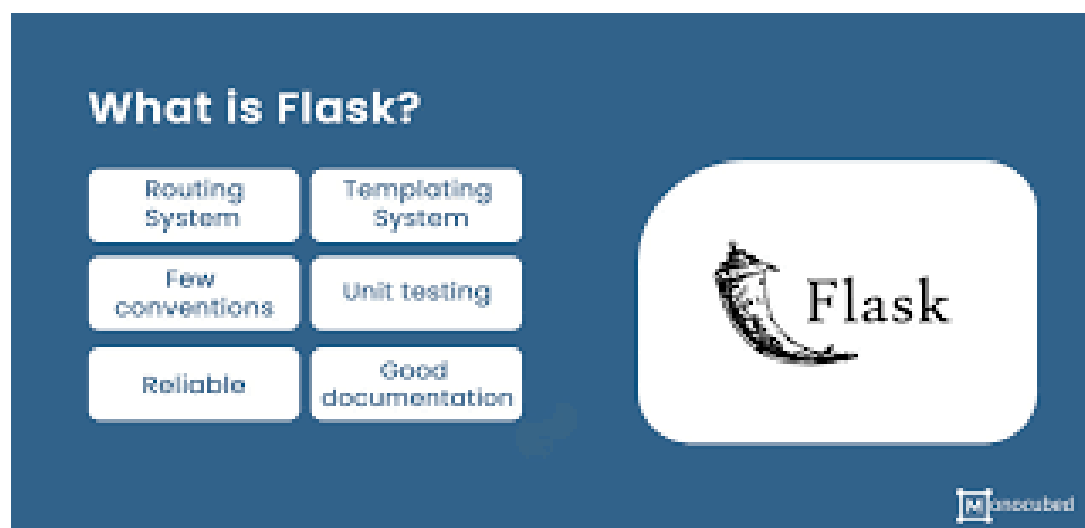
In Future we can extend this Project to tell users about what medicines he/she can take along with the Disease to get rid of the disease.

# Chapter 2

## Technology Used

### 2.1. Flask Framework

Flask is a small and lightweight Python web framework that provides useful tools and features that make creating web applications in Python easier. It gives developers flexibility and is a more accessible framework for new developers since you can build a web application quickly using only a single Python file.



**Figure 2.1** What is Flask?

#### 2.1.1 Why is Flask a good web framework choice?

Flask is considered more Pythonic than the Django web framework because in common situations the equivalent Flask web application is more explicit. Flask is also easy to get

started with as a beginner because there is little boilerplate code for getting a simple app up and running.

## **2.2. Git/GitHub**

GitHub is a code hosting platform for version control and collaboration. It lets you and others work together on projects from anywhere. This tutorial teaches you GitHub essentials like repositories, branches, commits, and pull requests.

Simply put, Git is a version control system that lets you manage and keep track of your source code history. GitHub is a cloud-based hosting service that lets you manage Git repositories. If you have open-source projects that use Git, then GitHub is designed to help you better manage them.

### **2.2.1 Why do we need Git ?**

Git makes it easy to contribute to open source projects. Nearly every open-source project uses GitHub to manage their projects. Using GitHub is free if your project is open source, and it includes a wiki and issue tracker that makes it easy to include more in-depth documentation and get feedback about your project.

If you want to contribute, you just fork (get a copy of) a project, make your changes, and then send the project a pull request using GitHub's web interface. This pull request is your way of telling the project you're ready for them to review your changes.



**Figure 2.2 :- Github**

## **2.3. Python**

Python is a computer programming language often used to build websites and software, automate tasks, and conduct data analysis. Python is a general-purpose language, meaning it can be used to create a variety of different programs and isn't specialized for any specific problems.

Python is an interpreted, object-oriented, high-level programming language with dynamic semantics. Its high-level built in data structures, combined with dynamic typing and dynamic binding, make it very attractive for Rapid Application Development, as well as for use as a scripting or glue language to connect existing components together. Python's simple, easy to learn syntax emphasizes readability and therefore reduces the cost of program maintenance. Python supports modules and packages, which encourages program modularity and code reuse. The Python interpreter and the extensive standard library are available in source or binary form without charge for all major platforms, and can be freely distributed.

Often, programmers fall in love with Python because of the increased productivity it provides. Since there is no compilation step, the edit-test-debug cycle is incredibly fast.



Debugging Python programs is easy: a bug or bad input will never cause a segmentation fault. Instead, when the interpreter discovers an error, it raises an exception. When the program doesn't catch the exception, the interpreter prints a stack trace. A source level debugger allows inspection of local and global variables, evaluation of arbitrary expressions, setting breakpoints, stepping through the code a line at a time, and so on. The debugger is written in Python itself, testifying to Python's introspective power. On the other hand, often the quickest way to debug a program is to add a few print statements to the source: the fast edit-test-debug cycle makes this simple approach very effective.



**Figure 2.3 :- Python**

## **2.4. Neural Network - Deep Learning**

Neural networks, also known as artificial neural networks (ANNs) or simulated neural networks (SNNs), are a subset of machine learning and are at the heart of deep learning algorithms. Their name and structure are inspired by the human brain, mimicking the way that biological neurons signal to one another.

In a neural network, we have the same basic principle, except the inputs are binary and the outputs are binary. The objects that do the calculations are perceptrons. They adjust themselves to minimize the loss function until the model is very accurate. For example, we can get handwriting analysis to be 99% accurate.

There are three main components: an input layer, a processing layer, and an output layer. The inputs may be weighted based on various criteria.

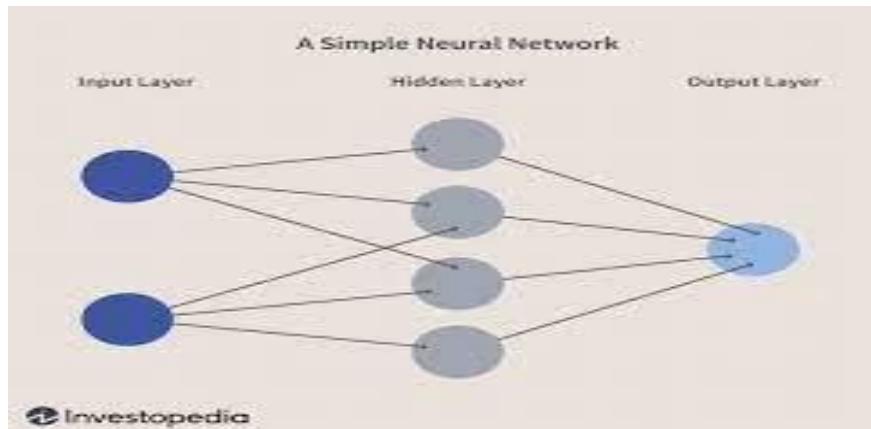


Figure 2.4 -: Neural Network

## 2.5. Amazon Web Services (AWS)

AWS (Amazon Web Services) is a comprehensive, evolving cloud computing platform provided by Amazon that includes a mixture of infrastructure as a service (IaaS), platform as a service (PaaS) and packaged software as a service (SaaS) offerings.

AWS is architected to be the most flexible and secure cloud computing environment available today. Our core infrastructure is built to satisfy the security requirements for the military, global banks, and other high-sensitivity organizations.

Amazon Web Services provides services from dozens of data centers spread across availability zones (AZs) in regions across the world. An AZ is a location that contains multiple physical data centers. A region is a collection of AZs in geographic proximity connected by low-latency network links.

Amazon Simple Storage Service (S3) provides scalable object storage for data backup, collection and analytics. An IT professional stores data and files as S3 objects -- which can range up to 5 gigabytes (GB) -- inside S3 buckets to keep them organized. A business can save money with S3 through its Infrequent Access storage tier or by using Amazon Glacier for long-term cold storage.



**Figure 2.5 :- AWS**

# Chapter 3

## Work Done

### 3.1. Overview

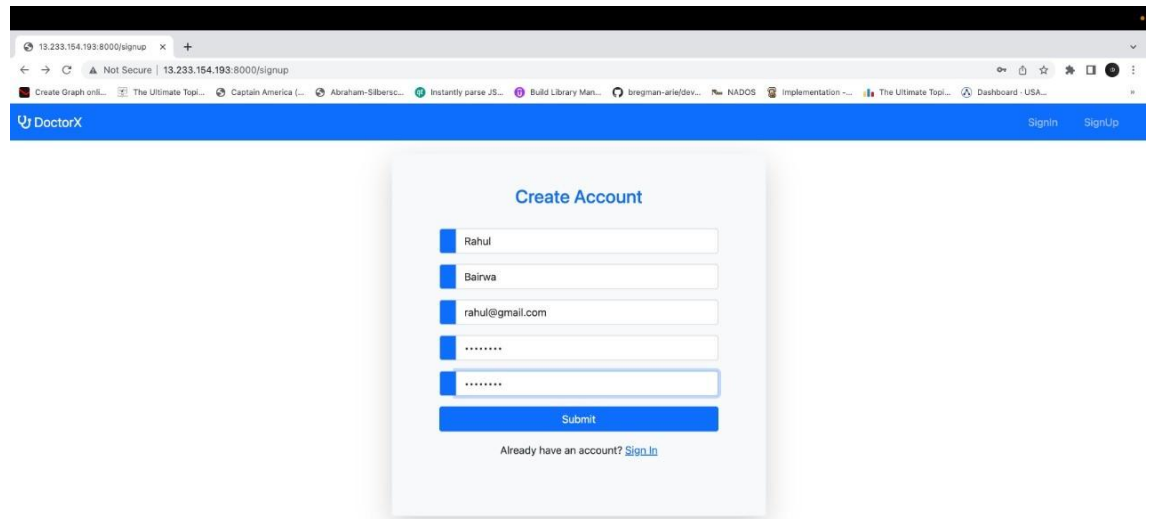
In this project we have implemented a standard signup/login portal which will allow users to sign in to their respective account. After signing, the user will be required to provide the symptoms he/she is facing, these symptoms will be saved in the database for further reference and then our ML model will predict the disease based on provided symptoms . For a user friendly environment Flask App is used for frontend.

### 3.2. Components

Below are components with usage:

#### 3.2.1 SignUP

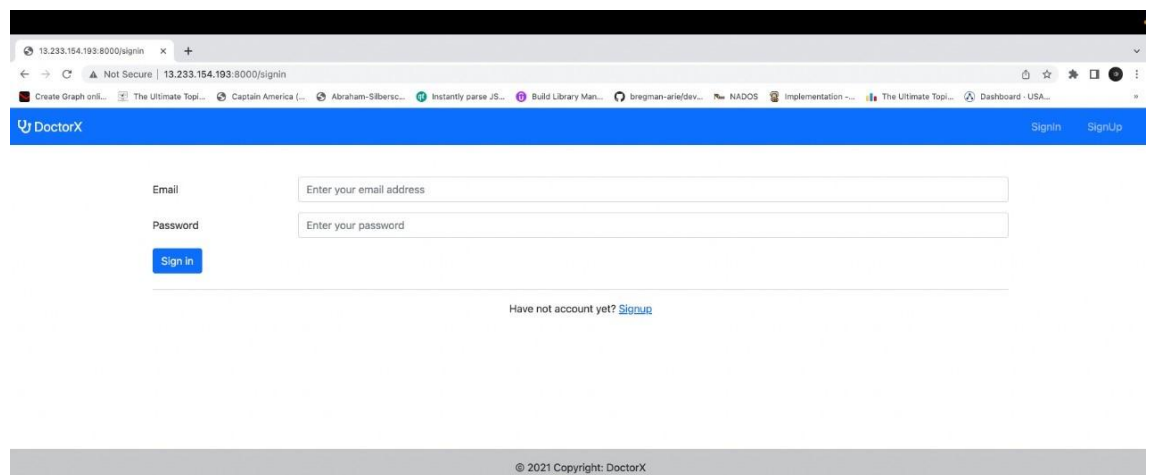
This is the first step of our app. In this page a user can register himself/herself by providing his/her required details such as Name, EmailId, Password, etc. The account details of the user will be stored in our Sqlite Database so that user can easily login himself/herself.



**Figure 3.1 :- SignUp Page**

### 3.2.2 SignIN

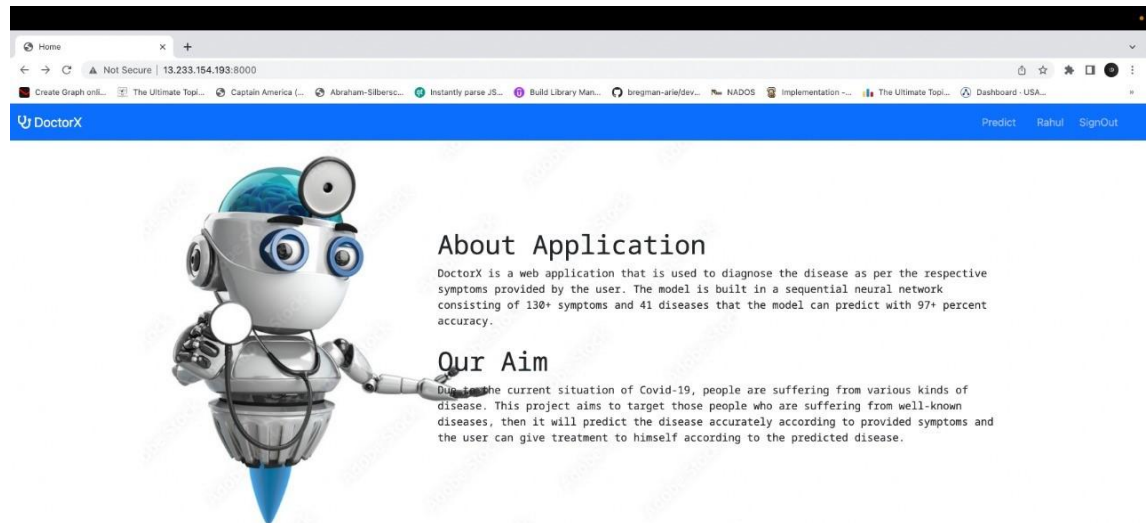
Before proceeding to this page, the user is required to SignUp. For Login users have to provide their details such as EmailId and Password, after this user's account will be signed in.



**Figure 3.2 :- SignIn Page**

### 3.2.3 Home

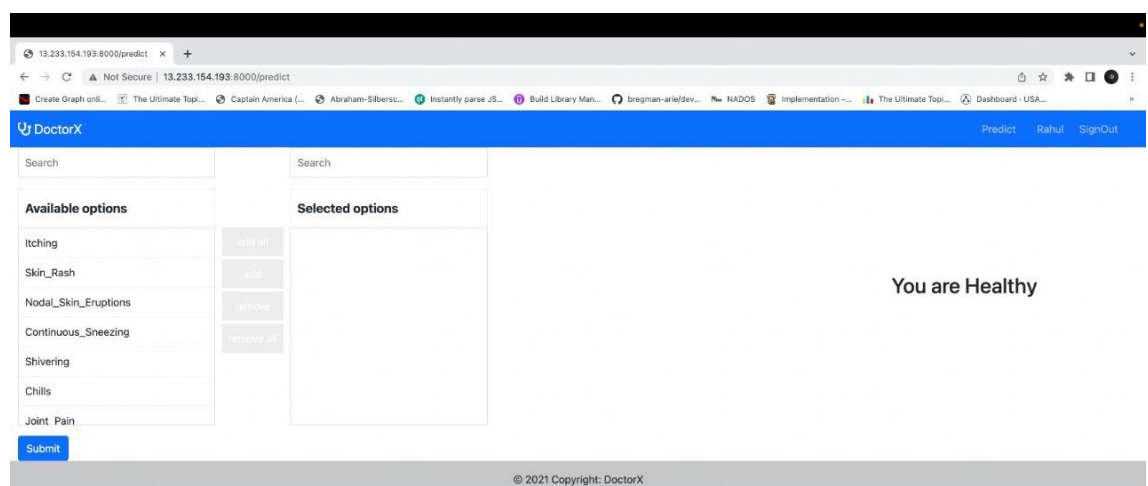
This is our HomePage section which is describing the product and it's work.



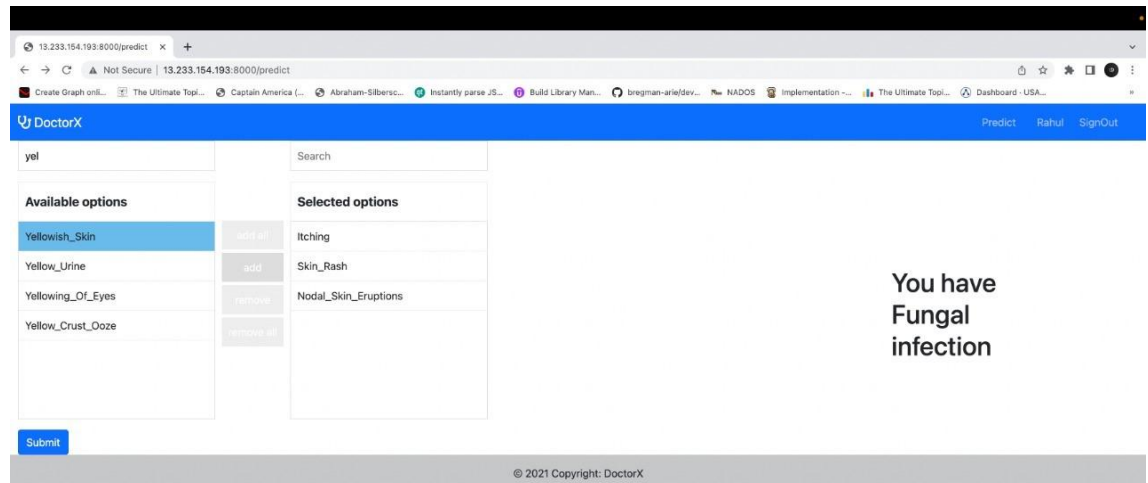
**Figure 3.3 :- Home Page**

### 3.2.4 Predict

This is the main component of our App. On the left most side, the app provides all the symptoms that it contains to the user to select and add them to the prediction list. Users can also search for any symptoms that are available in the list. After this user needs to submit the form and then predicted disease will be shown on the right most side of the screen as shown in figure.



**Fig 3.4 (a) : Predict Page**



**Figure 3.4 (b): Predict Page**

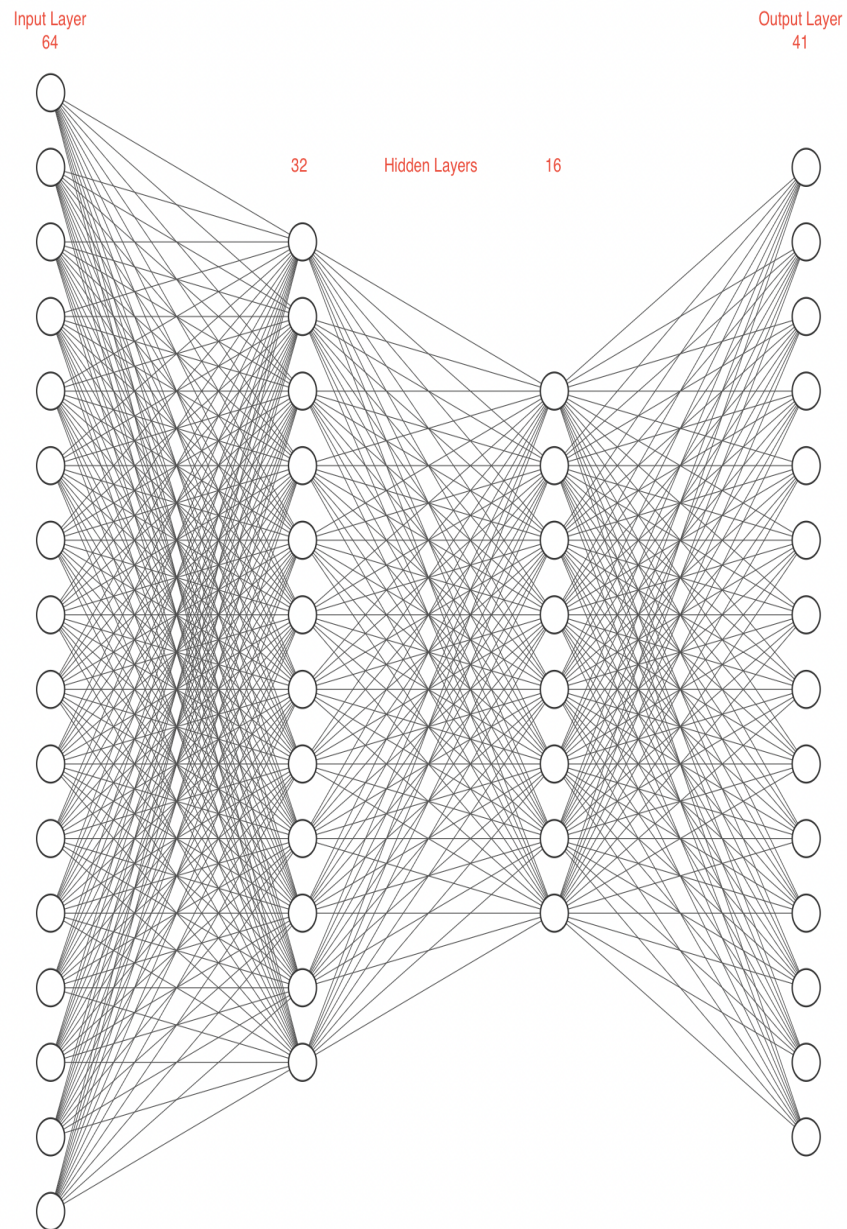
### 3.3. Architecture

We have followed the Sequential model for neural network processing and the model consist of 4 layers. First Dense layer is the input layer which is of size 64 although we have 132 as input shape, it will fit according to this. After this there are two more hidden layers which is of 32 and 16. After this we have output layer which is of 41 as we have 41 prediction classes.

Model: "sequential"

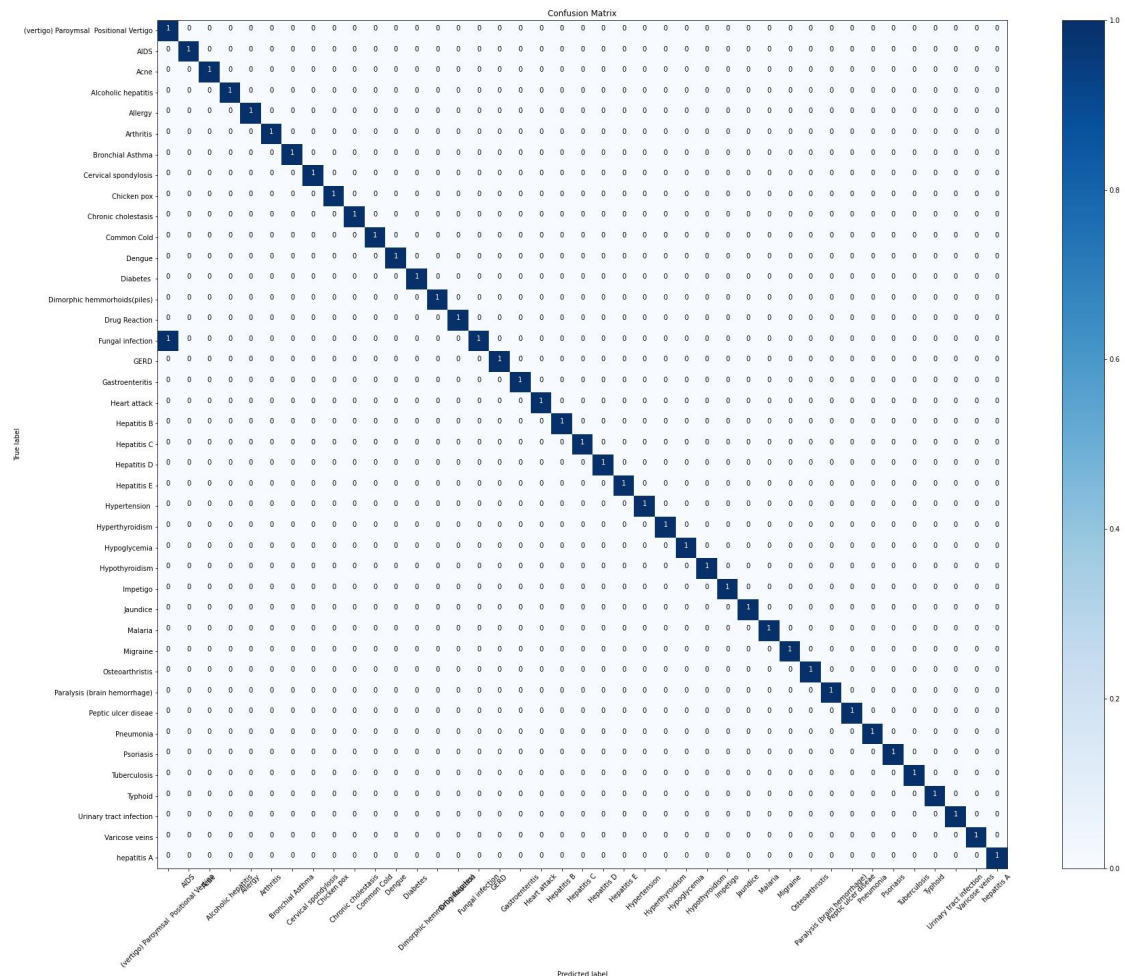
Layer (type)	Output Shape	Param #
dense (Dense)	(None, 64)	8512
dense_1 (Dense)	(None, 32)	2080
dense_2 (Dense)	(None, 16)	528
dense_3 (Dense)	(None, 41)	697
Total params: 11,817		
Trainable params: 11,817		
Non-trainable params: 0		

### 3.3.1. Network Architecture





### 3.3.2. Confusion Matrix



As we can see in confusion we have found accuracy around 92% which is quiet good.

### 3.4. Results

The Prediction Engine provides an optimal performance with the right dataset and efficient training of the classifier models considering all aspects and a lot of learning from the previous experiences. The implemented Prediction Engine is capable of predicting the presence of Disease with an accuracy of 92% accuracy. This will help anyone who wants to predict disease based on their symptoms.

## Chapter 4

### Conclusion & Future Work

#### 4.1. Conclusion

In Today's World due to current situation of CORONA people can use this app to predict there disease without need of going to hospital for every small disease and there are more chances to get affected by the virus at the hospital and due to lockdown condition there are many boundations on people for going out so they can use this app to check there disease at their home residence. And can take Safety measures according to it. The manuscript presented the technique of predicting the disease based on the symptoms, age, and gender of an individual patient. The Weighted Sequential model gave the highest accuracy of 92% for the prediction of diseases using the above-mentioned factors. Almost all the ML models gave good accuracy values. As some models were dependent on the parameters, they couldn't predict the disease and the accuracy percentage was quite low. Once the disease is predicted, we could easily manage the medicine resources required for the treatment. This model would help in lowering the cost required in dealing with the disease and would also improve the recovery process.

#### 4.2. Future Work

1. In the future we can extend this Project to tell users about what medicines he/she can take along with the Disease to get rid of the disease.
2. In the future we will try to increase the accuracy of our model so that it can predict disease more accurately.

3. In the future we will try to train our model on more symptoms.
4. In the future we will try to train our model for more diseases.

## References

- [1]. A. Mir, S.N. Dhage, in 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA) (IEEE, 2018), pp. 1–6
- [2]. Y. Khourdifi, M. Bahaj, Heart disease prediction and classification using machine learning algorithms optimized by particle swarm optimization and ant colony optimization, *Int. J. Intell. Eng. Syst.* 12(1), 242 (2019)
- [3]. S. Vijayarani, S. Dayananda, Liver disease prediction using svm and naïve bayes algorithms, *International Journal of Science, Engineering and Technology Research (IJSETR)* 4(4), 816 (2015)
- [4]. S. Mohan, C. Thirumalai, G. Srivastava, Effective heart disease prediction using hybrid machine learning techniques, *IEEE Access* 7, 81542 (2019)
- [5]. T.V. Sriram, M.V. Rao, G.S. Narayana, D. Kaladhar, T.P.R. Vital, Intelligent parkinson disease prediction using machine learning algorithms, *International Journal of Engineering and Innovative Technology (IJEIT)* 3(3), 1568 (2013)
- [6]. A.S. Monto, S. Gravenstein, M. Elliott, M. Colopy, J. Schweinle, Clinical signs and symptoms predicting influenza infection, *Archives of internal medicine* 160(21), 3243 (2000)
- [7]. R.D.D.P. Sreevalli, K.P.M. Asia, Prediction of diseases using random forest classification algorithm