Title Of The Thesis

Application Of Machine Learning In The Field Of Medicine:Symptom to Disease Prediction and Chest X-Ray Scan Diagnosis

Group Number 19

Group Members Name

AVIHARSH SHUKLA **(IPG2016-022)**
SANJAL DHIR KATIYAR **(IPG2016-090)**
SANYAM GUPTA **(IPG2016-093)**

Supervisor Names

Prof. Pankaj Srivastava

Dr. Prasenjit Chanak

Background

Our project aims to imitate a doctor's role to some extent. Our system takes the user(patient) input as symptoms and learns on them to predict the corresponding disease. So the area of our interest is basically diagnosis, pharmaceuticals and therapeutics. Also we are mapping the disease predicted by our system to its

description and possible precautions that would help our users(patients) to get a better understanding of their health condition.
We are also providing an additional feature for the users to upload a chest X-ray image as an input to our system and to diagnose and determine whether the lungs are in a healthy condition or not.

Motivation

Why this problem is important?

- Usually, with the onset of the symptoms, patients are in a dilemma as to what could be ailing them. This system can provide an initial understanding to the users about their symptoms and the corresponding disease.

- This problem is concerned with the most important aspect of our lives i.e, health and well-being.

- As it affects our lives significantly so more study and research should be done in this direction.

Why did we choose this problem?

- This problem is chosen by us owing to our civic sense mindset and social concern. This would be highly beneficial for the people living in remote areas especially village people

who do not have access to proper medical facilities.

- Open source and high-quality dataset is available which can easily be used after proper cleaning and pre-processing.

Why should others be interested in this problem?

- Usually, similar kinds of systems only have one feature, that is to predict the disease, given symptoms from the user as inputs. What makes our system different from others is that, apart from learning and predicting disease from the given symptoms, we are also providing an additional feature for the users to upload a chest X-ray image as an input to our system and to provide a diagnosis of the condition while determining whether the lungs are in a healthy condition or not.

- Our system is built in the form of a web application which uses our trained machine learning models.

- Our system boasts of a user-friendly interface as well as it predicts the condition of the patient with a good level of confidence.

Objectives /Mathematical Model /Description Of Our Thesis

- The first part of our problem consists of taking the symptoms as inputs from the user (patient) and using a Decision Tree Classifier to learn on them and predict the corresponding disease.Our system is capable of taking 132 different symptoms as inputs and predicting 41 different diseases.

- A Decision Tree Classifier is a tree-like hierarchical structure where each node corresponds to the feature or characteristic on which we split (in our case the ‘patient symptom’) , the branch represents the splitting decision norm and each leaf node represents the predicted result of the splitting.

- The second part of our problem consists of mapping the disease predicted by our system to its description and possible precautions that would help our users(patients) to get a better understanding of their health condition.

- For this we have used a hash data structure consisting of key and value pairs. In our case, the keys are the predicted diseases and the values are its corresponding description and precautions to be administered.

- The third part of our problem addresses a relatively more complex issue i.e. taking the chest X-ray scan image from the user as input and providing a diagnosis of the condition while determining whether the lungs are in a healthy condition or not.Our system is capable of predicting 15 different conditions or states of the patient’s lungs.

- To achieve this task, we have used a deep Convolutional Neural Network. We have used transfer learning on a pre-trained VGG19 model (which consists of 19 layers) by adding three extra layers (2 convolutional and 1 fully connected) and using ‘dropout’ for regularization.

- All the trained models are built into a FLASK API. Its endpoints are used in our web application to send the user(patient) selected symptom or X-ray scan to our model and get the corresponding output.

Timeline (Gantt Chart)