**DETECTING PARKINSON’S DISEASE USING IBM WATSON MACHINE LEARNING**

**1. INTRODUCTION**

**1.1 Overview**

More than 10 million people are living with Parkinson’s Disease worldwide, according to the Parkinson’s Foundation. While Parkinson’s cannot be cured, early detection along with proper medication can significantly improve symptoms and quality of life

The researchers found that the drawing speed was slower and the pen pressure is lower among Parkinson’s patients. One of the indications of Parkinson’s is tremors and rigidity in the muscles, making it difficult to draw smooth spirals and waves. It is possible to detect Parkinson’s disease using the drawings alone instead of measuring the speed and pressure of the pen on paper. Our goal is to quantify the visual appearance(using HOG method) of these drawings and then train a machine learning model to classify them. In this project, We are using, Histogram of Oriented Gradients (HOG) image descriptor along with a Random Forest classifier to automatically detect Parkinson’s disease in hand-drawn images of spirals and waves.

**1.2 Purpose**

In this Python machine learning project, we will build a model to detect Parkinson’s disease using one of the Classifier techniques known as RandomForestClassifier as our output contains only 1’s and 0’s. We’ll load the dataset, get the features and targets, split them into training and testing sets, and finally pass them to RandomForestClassifier for prediction.

**2. LITERATURE SURVEY**

**2.1 Existing problem**

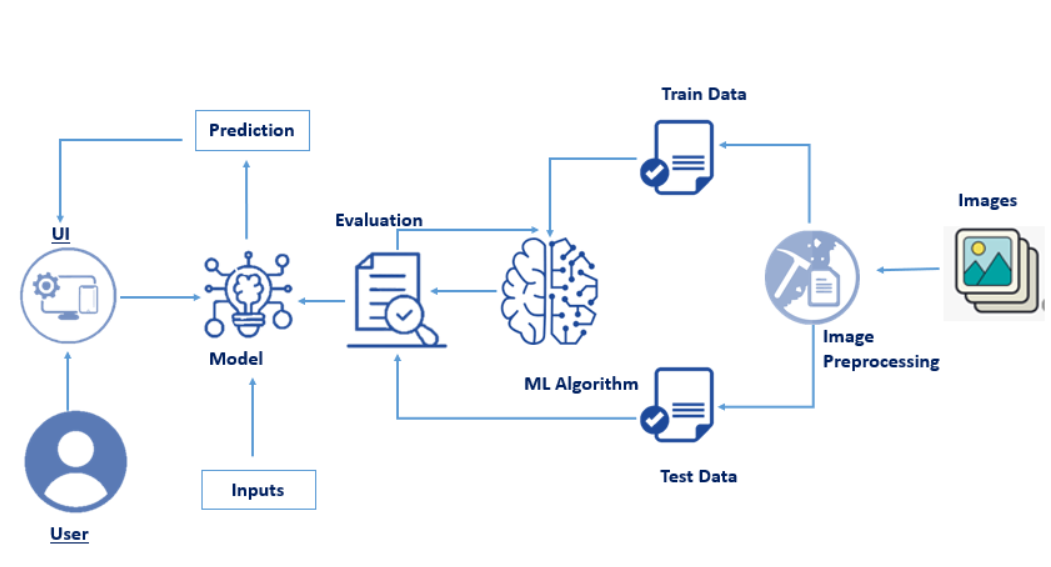
In existing system, PD is detected at the secondary stage only (Dopamine deficiency) which leads to medical challenges. Also doctor has to manually examine and suggest medical diagnosis in which the symptoms might vary from person to person so suggesting medicine is also a challenge. Thus the mental disorders are been poorly characterized and have many health complications. PD is generally diagnosed with the following clinical methods as, • MRI or CT scan - Conventional MRI cannot detect early signs of Parkinson's disease • PET scan - is used to assess activity and function of brain regions involved in movement • SPECT scan - can reveal changes in brain chemistry, such as a decrease in dopamine This results in a high misdiagnosis rate (up to 25% by non-specialists) and many years before diagnosis, people can have the disease. Thus existing system is not effective in early prediction and accurate medicinal diagnosis to the affected people

**2.2 Proposed solution**

The proposed system for predictive analytics is a mixture of clustering of K-means and a decision tree used to gain insights from patients. The problem can be addressed with reduced error rate with the application of machine learning techniques. Our proposed system also produces accurate results by combining the spiral drawing inputs of patients impacted by common and Parkinson’s. From these drawings, the principal component analysis algorithm (PCA) for extraction of the feature from the spiral drawings and support vector machine is used for classification. UCI machine learning platform voice data collection in Parkinson's disease is used as feedback. Thus, our study results will show early detection of the disorder can promote the therapeutic care of the elderly and increase the chances of their life span and healthier lifestyle living peaceful life.

**3. THEORITICAL ANALYSIS**

**3.1 Block Diagram**



**3.2 Hardware / Software designing**

***Software Requirements:***

To complete this project we have used the following software and packages -

**Anaconda Navigator :**

Anaconda Navigator is a free and open-source distribution of the Python and R programming languages for data science and machine learning related applications. It can be installed on Windows, Linux, and macOS.Conda is an open-source, cross-platform,  package management system. Anaconda comes with so very nice tools like JupyterLab, Jupyter Notebook,

QtConsole, Spyder, Glueviz, Orange, Rstudio, Visual Studio Code. For this project, we will be using Jupyter notebook and Spyder

**To build Machine learning models you must require the following packages**

* **Numpy**:
* It is an open-source numerical Python library. It contains a multidimensional array and matrix data structures and can be used to perform mathematical operations
* **Scikit-learn:**
* It is a free machine learning library for Python. It features various algorithms like support vector machine, random forests, and k-neighbours, and it also supports Python numerical and scientific libraries like NumPy and SciPy
* **Scikit-image**
* Scikit-image, or skimage,  is an open-source Python package designed for image preprocessing.

* **Install imutils**
* Imutils are a series of convenience functions to make basic image processing functions such as translation, rotation, resizing, and displaying Matplotlib images easier with OpenCV
* Open anaconda prompt and type command

“pip install imutils”

* **OpenCV**
* [OpenCV](https://en.wikipedia.org/wiki/OpenCV) is a library of programming functions mainly aimed at real-time computer vision. Here, OpenCV is used to capture frames by accessing the webcam in real-time.
* Open anaconda prompt and type command

“pip install opencv-contrib-python”

* **Flask:**

Web framework used for building Web applications

If you are using **anaconda navigator**, follow below steps to download required packages:

* Open anaconda prompt.
* Type “pip install numpy” and click enter.
* Type “pip install scikit-image” and click enter.
* Type “pip install imutils” and click enter.
* Type “pip install scikit-learn” and click enter.
* Type “pip install opencv-contrib-python” and click enter.
* Type “pip install Flask” and click enter.

***Hardware Requirements:***

* ~Operating system: window 7 and above with 64bit
* ~ RAM: 4Gb and above
* ~Hard disk: min 100GB

**4. EXPERIMENTAL INVESTIGATIONS**

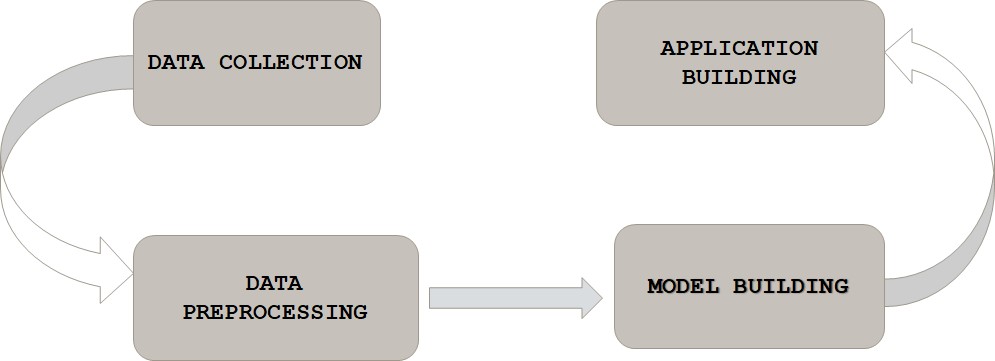
In general we prefer to use Deep learning and Convolution Neural Network for image data. In CNN we use data augmentation to pre-process the data. But as our data is very sensitive, improper use of data augmentation may lead to predicting a healthy person’s drawing look like Parkinson’s patient(drawing).

Therefore, applying computer vision to this problem is an appropriate choice.

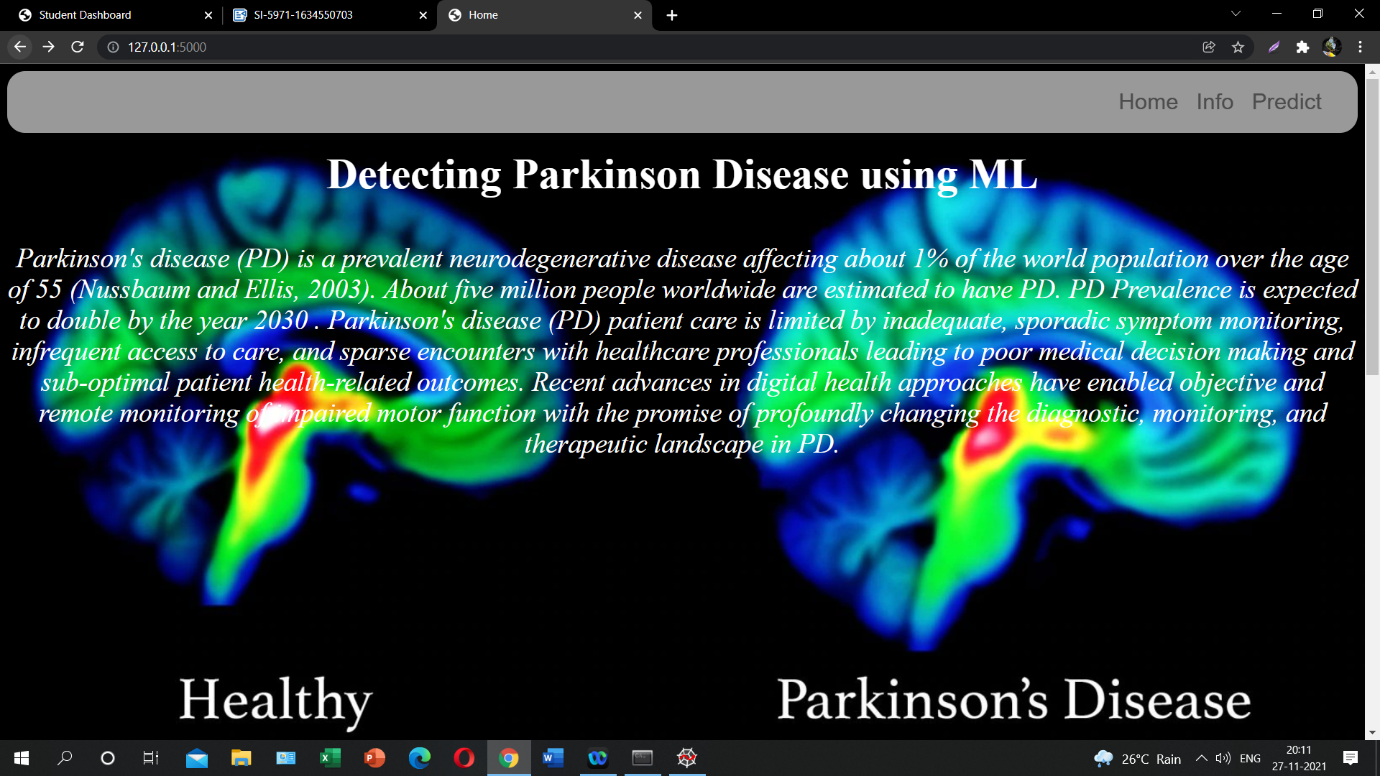
**By the end of this project:**

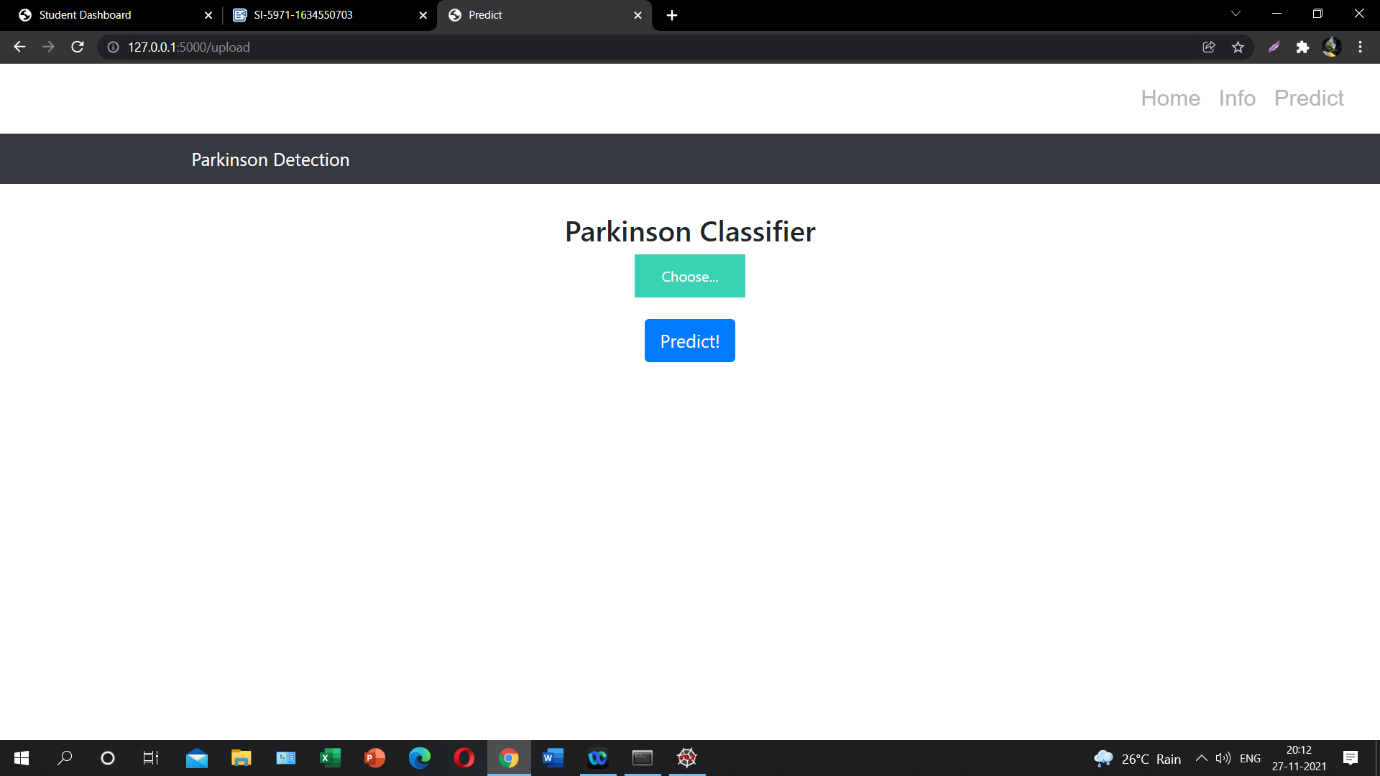
* You’ll be able to understand the problem to classify if it is a regression or a classification kind of problem.
* You will be able to know how to pre-process the image by using different data pre-processing techniques.
* you will learn how to use OpenCV and machine learning to automatically detect Parkinson’s disease in hand-drawn images of spirals and waves
* You will be able to know how to find the accuracy of the model.
* You will be able to Build web applications using the Flask framework.

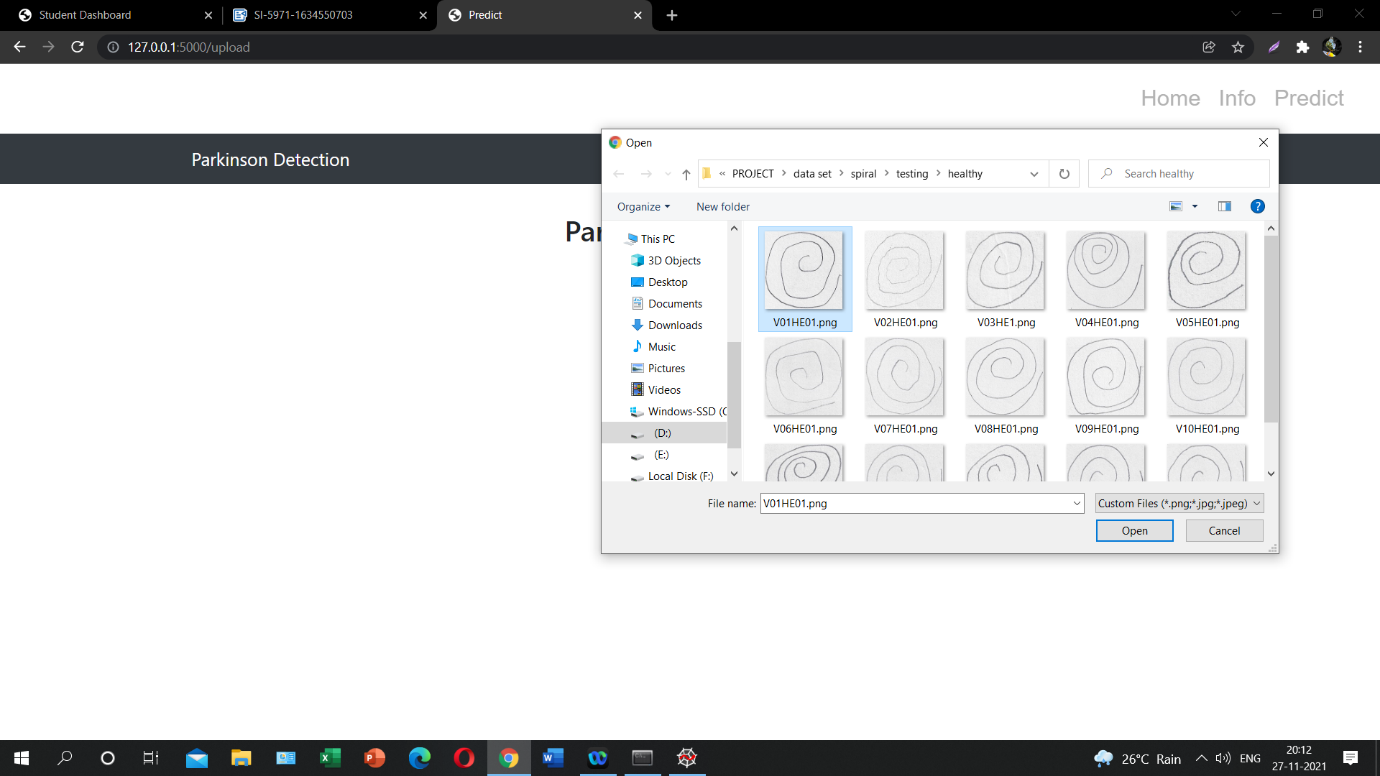
**5. FLOWCHART**

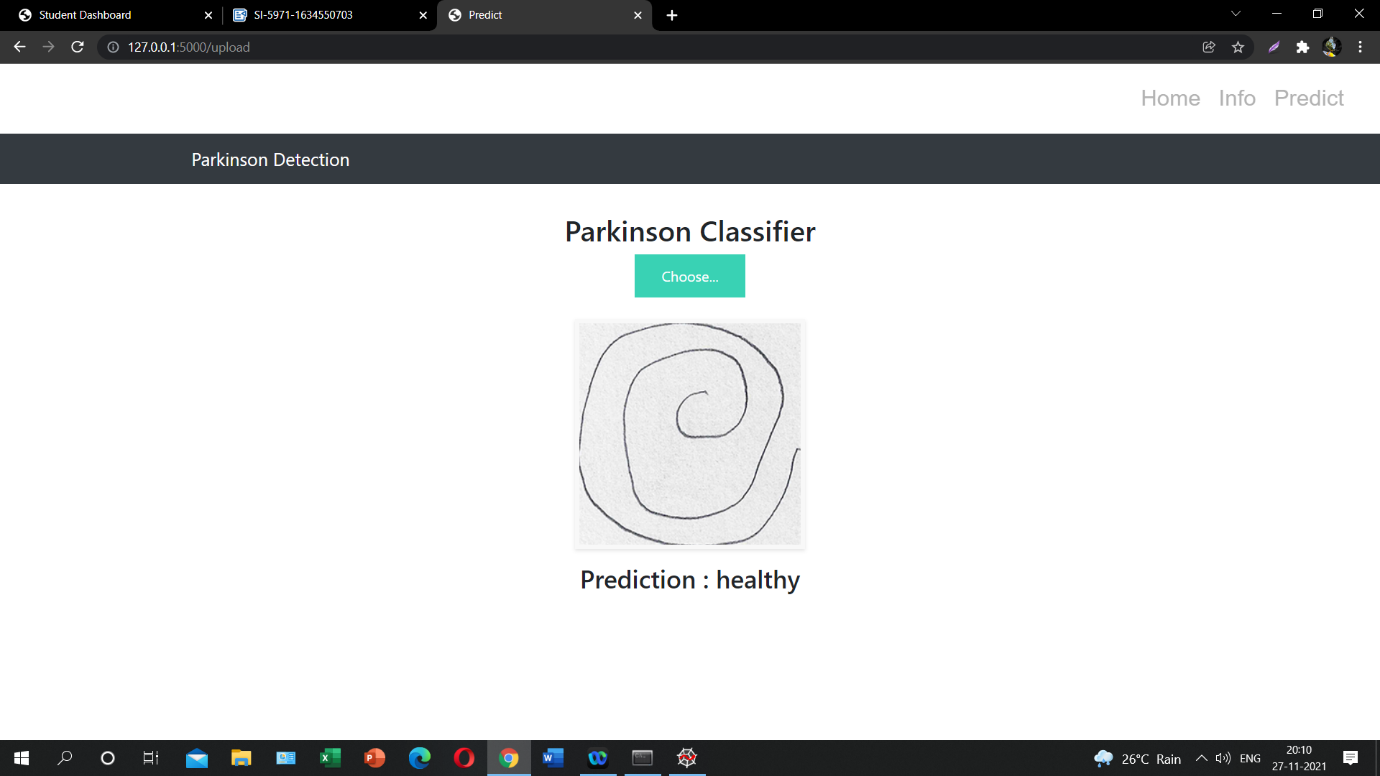
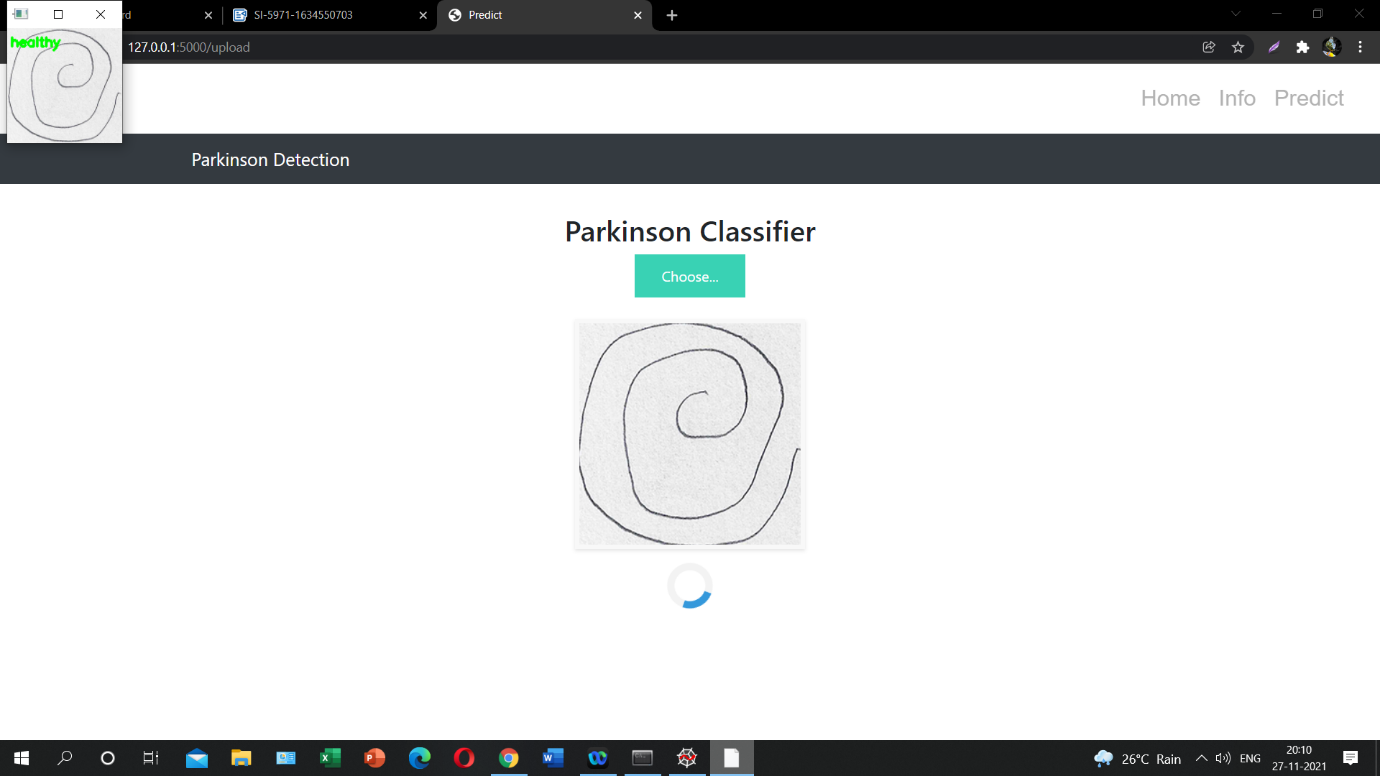
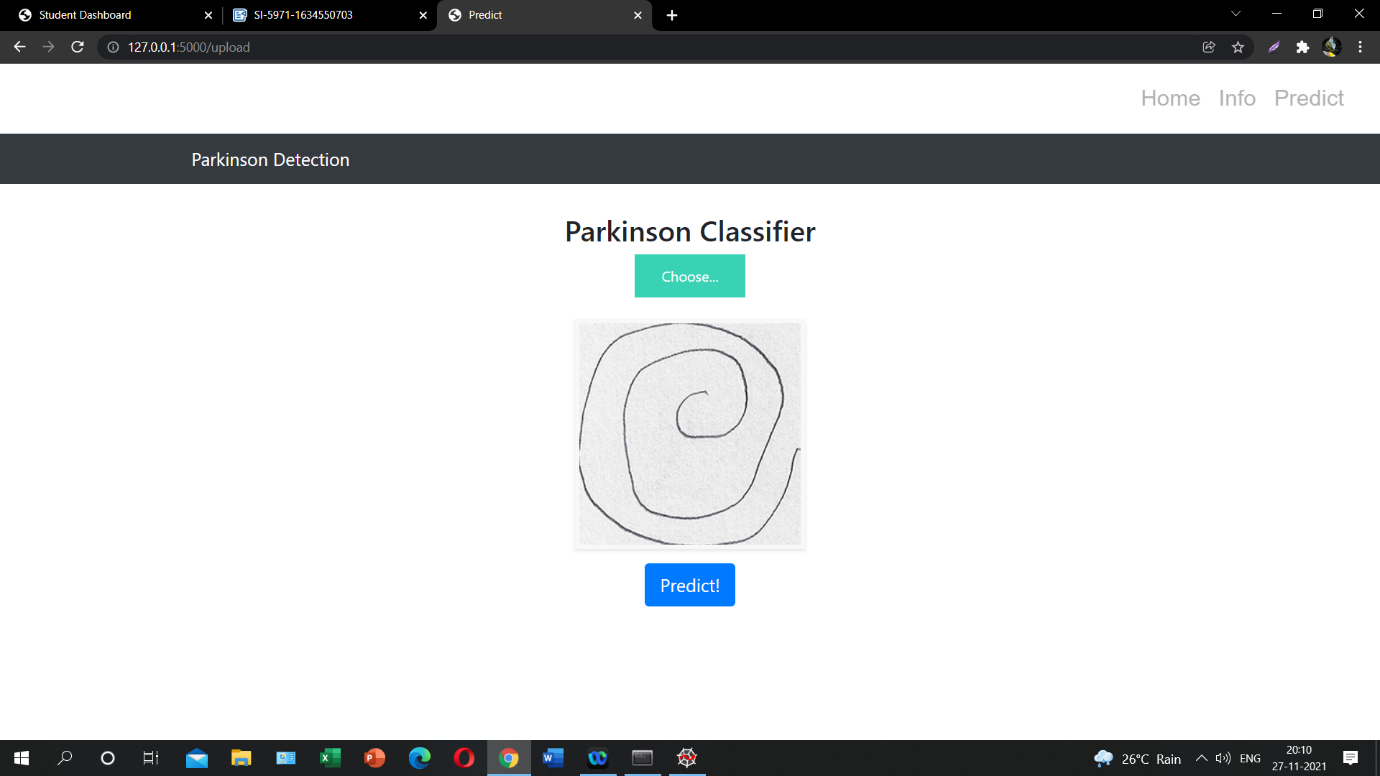


**6. RESULT**









**7. ADVANTAGES & DISADVANTAGES**

***Advantages:***

Parkinson’s disease detection using machine learning models could be very effective, cheap and scalable especially with the advent of transfer learning and pre-trained models which work quite well even with constraints like less data.

It reduces images to a form which is easier to process without losing features which are critical. Image pre-processing required is much less compared to other algorithms.

Machine learning does not require the design of handcrafted features, which is o of its biggest advantages.

***Disadvantages:***

There has been a lot of fine-tuning of these methods to make them perform better for blood smear images, and more so for the image analysis methods than for machine learning.

Despite the large number of publications, the performance numbers that have been published are very unsatisfying from a clinician’s point of view.

It requires a large training data. It requires appropriate model

**8. APPLICATIONS**

There has been a lot of fine-tuning of these methods to make them perform better for blood smear images, and more so for the image analysis methods than for machine learning.

There is certainly the potential that some of these methods gain importance outside malaria diagnosis, in particular for pre-processing and for detecting a segmenting red blood cells in other applications.

Working on this project lead us to believe that this work can play a part toward building a fully automated system for malaria parasite detection which may be useful in resource-constrained areas in the foreseeable future.

The same methods used in this project , can help In medical science, by making more models to work for many other diseases diagnoses. With this, medical technology can grow faster and be able to build 3D models that can predict accurately.

The increasingly growing number of applications of machine learning in healthcare allows us to glimpse at a future where data, analysis, and innovation work hand-in-hand to help countless patients without them eve realizing it. Soon, it will be quite common to find ML-based applications embedded with real-time patient data available from different healthcare systems in multiple countries, thereby increasing the efficiency of new treatment options which were unavailable before.

**9. CONCLUSION**

In this project, we built a machine learning model that can detect and class parkinson disease. A web application is integrated into the model, from where u can upload an spiral image and see the analyzed results on User Interface. A GUI based application which uses a custom CNN model to predict if uploaded spiral image is parasitized or uninfected. System will read the image uploaded by the user, augment it and will use the saved custom model to detect whether the disease is present or not in the patient and thus display the res in a user-friendly language.

Parkinson’s disease affects the CNS of the brain and has yet no treatment unless it’s detected early. Late detection leads to no treatment and loss of life. Thus its early detection is significant. For early detection of the disease, we utilized machine learning algorithms such as XGBoost and Random Forest. We checked our Parkinson disease data and find out XGBoost is the best Algorithm to predict the onset of the disease which will enable early treatment and save a life.

So, more automations must be brought in this field, and make the detection of diseases more effective and accurate.

**10. FUTURE SCOPE**

Classical approaches to feature computation and selection run the serious danger of being superseded soon by techniques not relying on handcrafted features, such as machine learning in particular.

In this project, we have designed & developed a machine learning model which automatically classifies and predicting infected person.

Deep learning is the latest trend in machine learning, which has already boost the performance in many nonmedical areas. Deep learning typically requires large training sets. This is the reason why medical applications have been among the last applications to adopt deep learning, as annotated training images are significantly harder to obtain because of expert knowledge requirements and privacy concerns. But through machine learning, there can be a drastic improvement in the Medical field, which will in turn help the world grow. In many cases, it will also enable early discovery and treatment accessibility in remote or developing places, as ML can significantly reduce costs and necessity for doctor appointments.

Overall, ML in healthcare is an incredible development that will increase efficiency and accuracy in disease detection.

**11. BIBILOGRAPHY**

1. <https://en.wikipedia.org/wiki/KNN>

2. <https://keras.io/api/preprocessing/image/>

3. <https://en.wikipedia.org/wiki/parkinson-disease>

4. <https://en.wikipedia.org/wiki/machine-learning>

**APPENDIX**

**Source Code**

