**Detecting Parkinson’s Disease using Machine Learning**

**Solution Architecture**

****

**WORKING OF THE ARCHITECTURE:**

Spiral drawing datasets of PD affected and unaffected patients collected by neurologists are obtained from Machine Learning repository. These are stored into the python environment as Testing and Training datasets and imported usingnecessary packages. Python is an open source dynamic, high level, free and interpreted programming language. This supports objectoriented programming and procedural programming.Python is currently the most popular progra mming language for Machine Learning research and development.PyCharm is an integrated development environme nt (IDE) primarily for the Python language, used in computer programming. Microsoft Visual Studio is a development environment by Microsoft. It is used to develop computer programs, websites, web applications, web services, and mobile apps. 36 different programming languages are supported by Visual Studio which includes C#, C++, etc., and allows the code editor and debugger to support nearly any programming language, provided a language-specific service exists. Built-in languages include C, C++, C++/CLI, Visual Basic .NET, C#, JavaScript, Typescript, XML, HTML, and CSS.

1. **Importing datasets into PyCharm/Visual studio code –**

Spiral drawing datasets of PD affected and unaffected patients collected by neurologists are obtained from Machine Learning repository. These are stored into the python environment as Testing and Training datasets and imported using necessary packages.

1. **Pre-Processing –**

It involves image acquisition, pre-processing and segmentation. Preprocessing image is a way to improve image quality, so that the resulting image is better than the original one. The goal of image acquisition is to collect images having low noise when compared to HD images. The main advantage of this module is to have images with better clarity, low noise and distortion.The aim of segmentation is to make the representation of an image simpler or more easily analyzable.

1. **Feature extraction –**

In this project, mean filter and median filter are presented for processing of selecting the images. The median filter is a non-linear tool, while linear is the average filter. Mean filtering of smoothing images is fast, intuitive and easy to implement i.e. reduces the amount of variation in intensity between one pixel and the next. The median filter is normally used in a picture to reduce salt-and-pepper noise. It often does a better job than maintaining useful information in the picture than the mean filter. The median is determined by first sorting Vol-6 Issue-2 2020 IJARIIE-ISSN(O)-2395-4396 11591 www.ijariie.com 509 all the pixel values in numerical order from the surrounding area and then replacing the pixel that is considered with the middle pixel value. If there are even number of pixels in the neighborhood under consideration the sum of the two middle pixel values is used. Both mean and median filters are used to remove noise. This is used as the input for further analysis.

**4**.**OpenCV library function –**

OpenCV (Open Source Computer Vision Library) was developed to provide an interface for computer vision applications and to facilitate the use of machine perception in the real time

**5. Classification (Random Forest) –**

It is a supervised learning algorithm used for classification. Random forest algorithm builds decision trees on data samples, then obtains the prediction from each and finally sel ects the best solution by voting.It is an ensemble approach that is better than a single decision tree, as it eliminates o verfitting by averaging the outcome. Where we can find the confusion matrix with the help of confusion\_matrix() function of sklearn, which is nothing but a table with two dimensions viz. “Actual” and “Predicted” and furthermore, both the dimensions have “True Positives (TP)”, “True Negatives (TN)”, “False Positives (FP)”, “False Negatives (FN)”,which calculates accuracy, specificity and sensitivity