Design & Implement Neural network models (Perceptron,SVM,LVQ,SOM) on Parkinson's Disease Classification

Arpit Shrma
11812449
KM108 INT246
A26
Project link- Parkinson's Disease Classification

Introduction-

Parkinson’s disease (PD) manifests as the death of dopaminergic neurons in the substantia nigra pars compacta within the midbrain. This neurodegeneration leads to a range of symptoms including coordination issues, bradykinesia, vocal changes, and rigidity. Dysarthria is also observed in PD patients; it is characterized by weakness, paralysis, and lack of coordination in the motor-speech system: affecting respiration, phonation, articulation, and prosody. Since symptoms and the disease course vary, PD is often not diagnosed for many years. Therefore, there is a need for more sensitive diagnostic tools for PD detection because, as the disease progresses, more symptoms arise that make PD harder to treat.
By these symptoms we can predict that person is suffering from this disease.

Problem Description-

There is currently no objective method for diagnosing PD. It can take months to get a reliable PD diagnosis, and symptoms need to be carefully monitored. Even then the probability of an inaccurate diagnosis is approximately 25%.

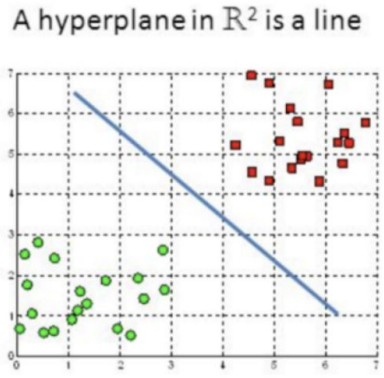
Solution to this problem-

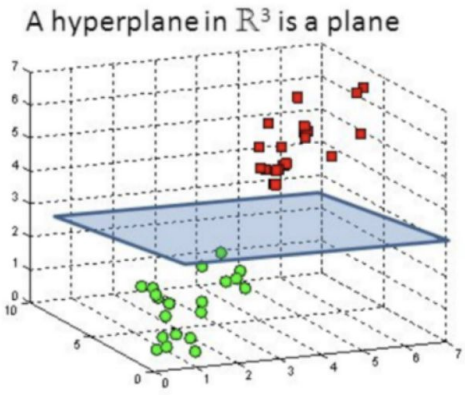
We used the Parkinson's disease classification database,.The results show that person is diagnosing and monitoring PD. To Make the solution I have used the SVM model. To train our algorithm I take data set from UCI machine learning repository.

What is SVM-

The objective of the support vector machine algorithm is to find a hyperplane in an N-dimensional space(N — the number of features) that distinctly classifies the data points.

To separate the two classes of data points, there are many possible hyperplanes that could be chosen. Our objective is to find a plane that has the maximum margin, i.e the maximum distance between data points of both classes. Maximizing the margin distance provides some reinforcement so that future data points can be classified with more confidence.





I've used the SVM algorithm from sklearn library.

Dataset Description-

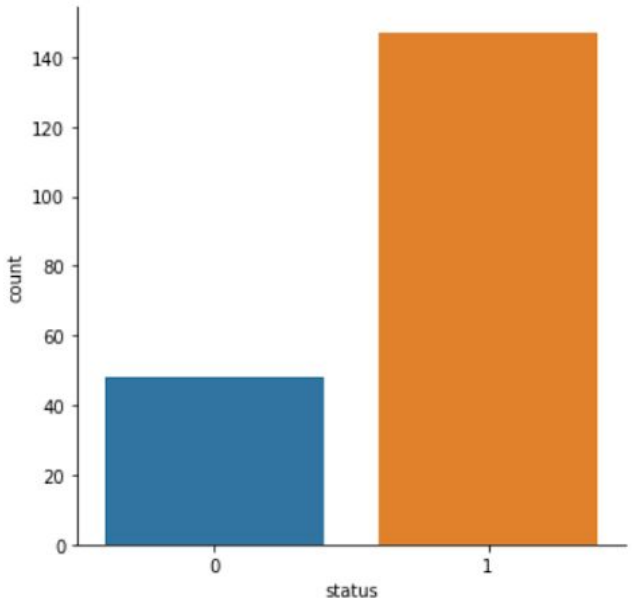
In UCI data set, there are 24 attributes in which 22 are independent features , one for name and one for status of the person for Parkinson's.

```
In [49]: a.dtypes

Out[49]: name                 object
         MDVP:Fo(Hz)         float64
         MDVP:Fhi(Hz)        float64
         MDVP:Flo(Hz)        float64
         MDVP:Jitter(%)      float64
         MDVP:Jitter(Abs)    float64
         MDVP:RAP            float64
         MDVP:PPQ            float64
         Jitter:DDP          float64
         MDVP:Shimmer        float64
         MDVP:Shimmer(dB)    float64
         Shimmer:APQ3        float64
         Shimmer:APQ5        float64
         MDVP:APQ            float64
         Shimmer:DDA         float64
         NHR                 float64
         HNR                 float64
         status                int64
         RPDE                float64
         DFA                 float64
         spread1             float64
         spread2             float64
         D2                  float64
         PPE                 float64
         dtype: object
```

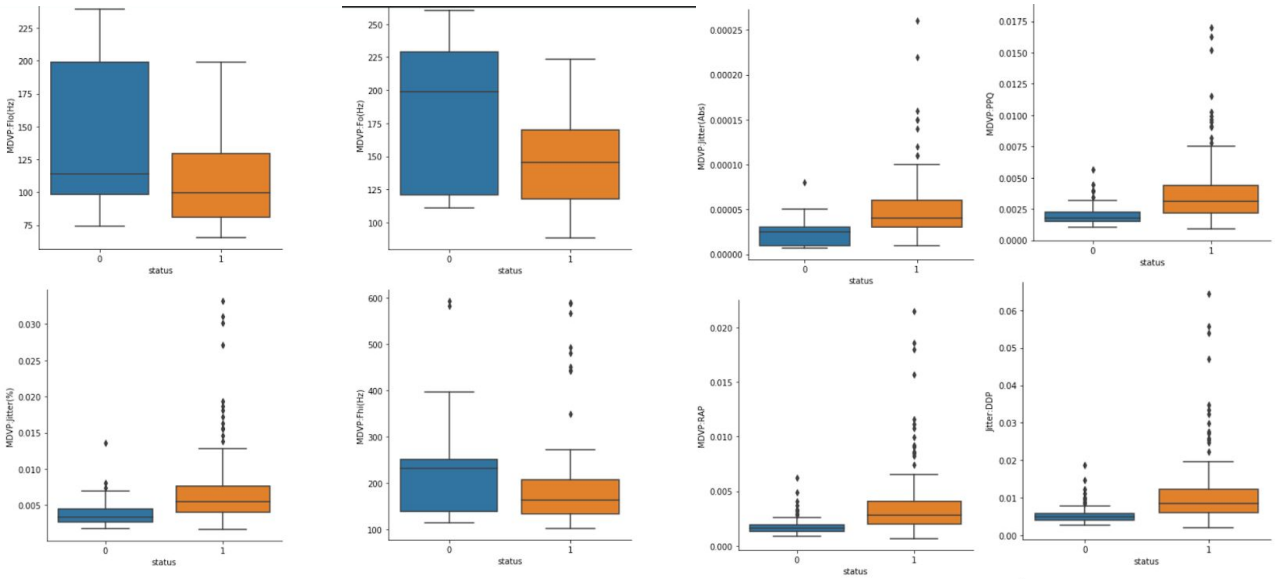
The bar plot shown below clearly shows that most of the observations given in the dataset have a status of 1 i.e most of the people represented in the dataset have been prone to Parkinson's disease.

```
In [50]: sns.catplot(x='status',kind='count',data=a)
Out[50]: <seaborn.axisgrid.FacetGrid at 0x1592f7d6f10>
```



Since 0 represent negative for Parkinson's disease, and 1 represent positive for Parkinson's disease.

The boxplot shown below helps in identifying the difference in values with respect to the 'status' of the patient.



Result-

After training the algorithm we get about 86% accuracy that is good enough to be diagnosed for Parkinson's Disease.

Tools I’ve Used-

- Jupyter Notebook
- Python
- UCI Dataset
- Seaborn
- sklearn
- Pandas
- Numpy