Parkinsons Diseases Detection

PARKINSON'S DISEASE

Reasons

heredity

head injuries

age

Treatment

Lorem ipsum dolor sit amet, consectetuer adipiscing elit, sed diam nonummy nibh euismod tincidunt ut laoreet dolore magna aliquam erat volutpat. Ut wisi enim ad minim veniam, quis nostrud

Men suffer from Parkinson's disease more often than women

Parkinson's disease affects about 9 million people worldwide

Prevention

Lorem ipsum dolor sit amet, consectetuer adipiscing elit, sed diam nonummy

physical exercise

walks in the open air

no smoking

no alcohol

Symptoms

hand shake

speech impairment

impaired body coordination

chronic fatigue

small steps

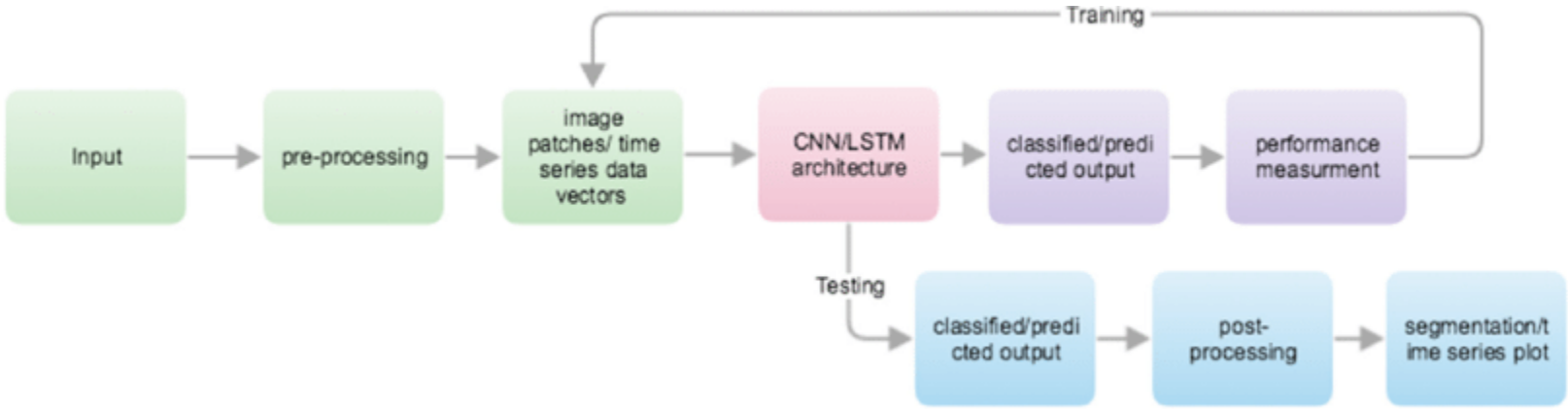
constipation and urination problems

Lorem ipsum dolor sit amet, consectetuer adipiscing elit, sed diam nonummy nibh euismod tincidunt ut laoreet dolore magna aliquam erat volutpat.

Objective

1. Accurately detecting Parkinson’s disease (PD) at an early stage is certainly indispensable for slowing down its progress and providing patients the possibility of accessing to disease-modifying therapy.

2. Towards this end, the premotor stage in PD should be carefully monitored. An innovative deep-learning technique can be apply as early to uncover whether an individual is affected with PD or not based on premotor features.



Data Validation and Data Transformation :

Name Validation - Data Validation is used to validate the data. checking the accuracy and quality of source data before training a new model version.

Number of Columns – Validation of number of columns present in the files.

Name of Columns - The name of the columns is validated and should be the same

as given in the schema file.

Data type of columns - The data type of columns is given in the schema file. It is validated when we insert the files.

Null values in columns - If any of the columns in a file have all the values as NULL or missing, we discard such a file

Model Selection –

After the clusters are created, we find the best model for each cluster. By using Algorithms ‘’Random Forest”, “Linear Regression”, “AdaBoost”, “Gradient Boost”, “SVM” and "XGBoost". For each cluster both the hyper tuned.
algorithms are used. We calculate the AUC scores for both models and select the model with the best score. Similarly, the model is selected for each cluster. All the models for every cluster are saved for use in prediction

Prediction:

The testing files are shared in the batches and we perform the same Validation operations ,data transformation and data insertion on them.

The accumulated data iis exported in csv format for prediction

We perform data pre-processing techniques on it.

Based on the cluster number respective model is loaded and is used to predict the data for that cluster.

Once the prediction is done for all the clusters. The predictions are saved in csv format and shared.


Thank
you