

# Parkinson's Disease Prediction



# Objective:

The main goal of this project is to perform extensive Exploratory Data Analysis(EDA) on the Parkinson's Disease Dataset and build an appropriate Deep Learning Model That will to Predict Parkinson's Disease.

# Benefits:

1. Using organisation data into real world Business use-case.
2. Predicting Parkinson's Disease and other general objective.
3. Helps increase profits to organisation.
4. Optimum services provided by app.

# Data Sharing Agreement:

- \* sample file name (disease.csv) and source of the file is

<https://www.kaggle.com/datasets/debasisdotcom/parkinson-disease-detection>

- \* shape of the data 195 x 24.

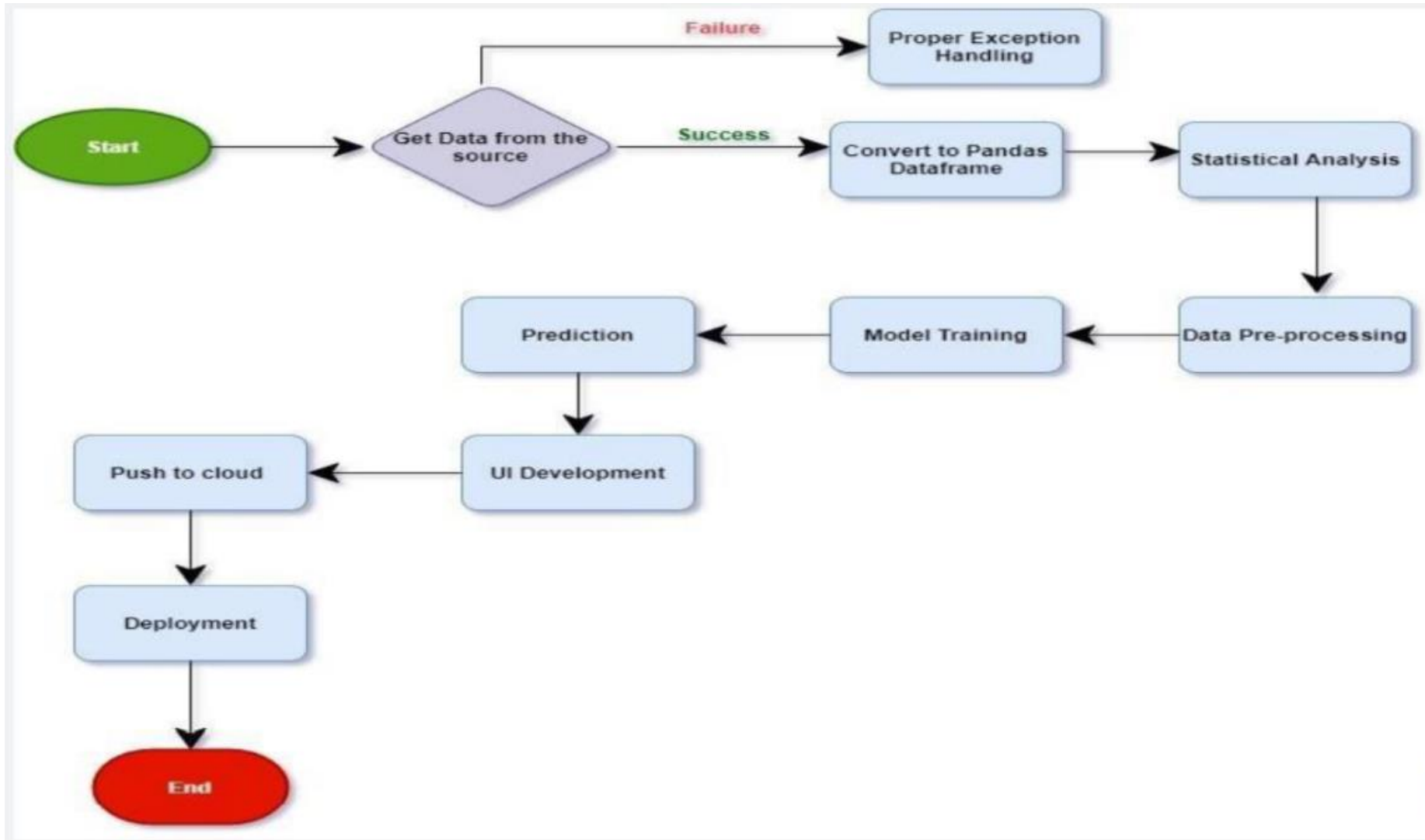
- \* 195 rows.

- \* 24 columns.

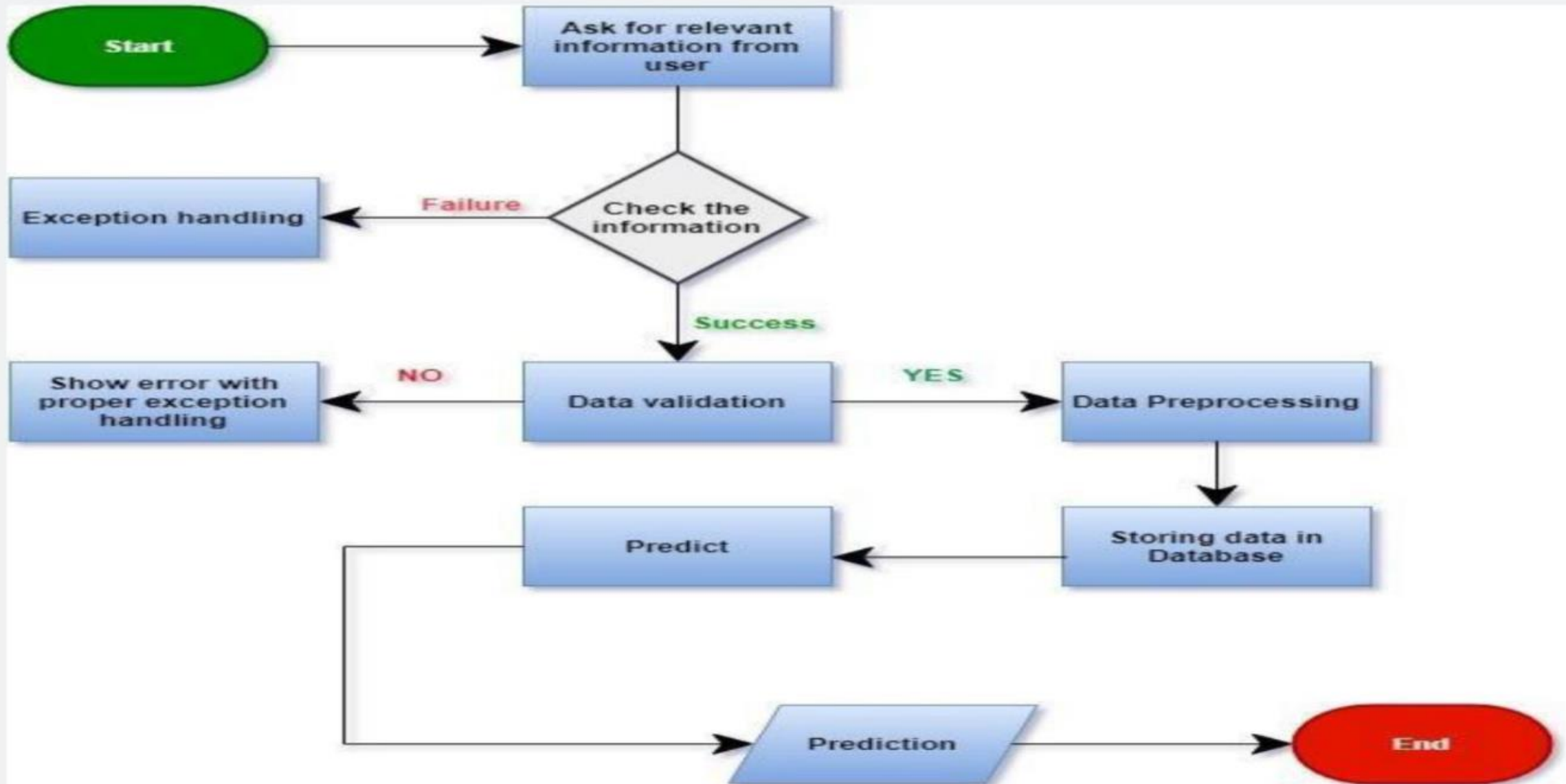
- \* Column data type where int64 , float64.

# Architecture

## Deep Learning Model



## Input Output Flow of Project



# Model Training

- Data Export From CSV:

Loading CSV data using python pandas and extracting all the data into dataframe in python file

- Data Preprocessing

- Performing EDA to get insight of data like identifying distribution , outliers ,trend among data etc.
- Check for null values in the columns. If present impute the null values.
- Perform Feature Selection and extract all the necessary features from the data

# Feature Selection:

\* We have selected featured based on Threshold value, by giving threshold value 3 we have selected features.

```
Index(['MDVP:Fo(Hz)', 'MDVP:Fhi(Hz)', 'MDVP:Flo(Hz)', 'MDVP:Shimmer(dB)',  
      'HNR', 'status', 'RPDE', 'spread1', 'D2'],  
      dtype='object')
```

# Train and Test Split:

- \* Train data is 70% of whole data.
- \* Test data is 30% of whole data.
- \* Data is randomly spitted.
- \* There is only train and test set data.

## • Model Selection:

- As for the data we have used Deep Learning Model and we have got 87% we have 2 input layers and 1 output layer.
- We have used Relu activation function for two inputs and sigmoid activation function for output layer.
- We have used Adam optimizer and binary cross entropy for loss.

```
classifier = Sequential()  
classifier.add(Dense(units = 6, kernel_initializer='he_uniform',activation = 'relu', input_dim =8 ))  
classifier.add(Dense(units = 6, kernel_initializer='he_uniform',activation = 'relu', input_dim = 8))  
classifier.add(Dense(units = 1, kernel_initializer='glorot_uniform',activation = 'sigmoid'))
```

```
classifier.compile(optimizer = 'adam', loss = 'binary_crossentropy', metrics = ['accuracy'])
```



# Data Insertion In Data Base:

- \* After Model is created and prediction is done the result is inserted into Cassandra database in order to keep records of result.
- \* Database : Database is created with name Parkinson's, scalable, flexible format.

# Q&A

1. What is the source of the data.

Ans. The data for training is provided by client (ineuron) in form of csv and source of the file is

<https://www.kaggle.com/datasets/debasisdotcom/parkinson-disease-detection>

2. What are type of data.

Ans. The data was combination of numerical values. There are no null values in this data.

3. What the complete flow you followed in this Project.

Ans. Refer slide 4<sup>th</sup> and 5<sup>th</sup> for better understanding.

4. How are logs are managed.

Ans. We are using different logs as per the steps that we follow in validation and modelling like File validation log, Data Insertion , Model Training Log, Prediction Log.

5. What are techniques were you using for data pre-processing.

Ans.

- \* Visualization relation of independent variables with each other and output variables.
- \* Removing unwanted features.
- \* Checking and changing distribution of data.
- \* Cleaning the data.
- \* Selecting Features.

6. How training was done and model was used.

Ans.

- \* Before diving the data in training and test set we performed pre-processing order to get better data.
- \* As per the model the training and test were divided.
- \* The random split was performed over training and validation data.
- \* deep learning model is used.

7. what are different stages of Deployment.

Ans

- \* When the model is ready we deployed it local environment, where UAT is performed.
- \* Then project uploaded in GitHub account.
- \* deployed in Heroku Cloud Platform.