

```

# -*-
Coding
: utf-8 -*-
"""

@author: Ayush
"""

import numpy as np

# import matplotlib.pyplot for plotting our result
import matplotlib.pyplot as plt

# import pandas for importing csv files
import pandas as pd
from math import *

# Reading the csv file containing all the data using panda
df=pd.read_csv('Final Data.csv')

# Using sklearn libraries
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestRegressor
from sklearn.metrics import mean_squared_error

# Input variables in X
X=df[['Wind Speed (miles/h)', 'Temperature', 'Humidity', 'Pressure']]
# Output variables in Y
Y=df[['LV ActivePower (kW)']]

# Splitting training and testing data
X_train, X_test, y_train, y_test = train_test_split(X,Y,test_size=0.2,random_state=100)
y_train=y_train.values.ravel()

# Regression tree
regressor=RandomForestRegressor(n_estimators=100)
# Fitting the model
regressor.fit(X_train,y_train)

# Getting predictions using test data

```

```
y_predict=regressor.predict(X_test)
# Evaluating error
print('Error: ',sqrt(mean_squared_error(y_test, y_predict)))
import pickle
# Storing model in pickle for fast regression as it stores model in form of binary files
pickle.dump(regressor, open("model.pkl","wb"))
```