PROF : Sylvain Jaume

NAME : Spoorthi Kondagadapu

Date : Tuesday, Feb 28

Lab : 6

WEATHER DATA

We can do only linear Regression on weather data

Step1 : uploading the datasets to Zeppelin

Code : ~~imports~~

```
1) import.org.apache.spark.SQL.functions
2) import org.joda.time.Format.DateTimeFormat
3) Val InputPath = "/C:/Users/spoor/Desktop.
                    finaldatasets/*"
4) Val. option("header", "true")
5)     .option("delimiter", ";")
6)     .option("inferSchema", "true")
7)     .load(inputPath)
```

NAME: SPOORTHI KONDAGHADAPU

DATE: TUESDAY Feb 28

8) import csv
9) import sys
10) import matplotlib.pyplot as plt
11) import numpy as np
12) import pandas as pd
13) from sklearn import datasets, linear
14) model [/code]

15) first_x-parameter, first_y-parameter
16) second_x-parameter, second_y-parameter
17) def get_data(finaldatasets)
18) data = pd.read_csv(final datasets)
19) first_x_parameter = [ ]
20) first_y_parameter = [ ]
21) second_x_parameter = [ ]
22) second_y_parameter = [ ]

```
23) for x1, y1 in
24) ~~zip~~(finaldataset['finaldatasets'], data)
25) first_x_parameter.append([float(x1)])
26) first_y_parameter.append([float(y1)])
27) return first_x_parameter,
28)              first_y_parameter
29) def Reg(x1,y1)
30) reg1 = linear_model.LinearRegression()
31) regr1.fit(x1,y1)
32) Predicted_Value1 = reg1.predict(1)
33) Print Predicted_Value1
```