Student Exam Marks Prediction - Linear Regression

import pandas as pd #Read Write Operations on CSV Files that we used

import numpy as np #Dealing with Numerical Tasks

import matplotlib.pyplot as plt #Displaying the Results in Graphical Plot Format

dataset=pd.read\_csv('student\_scores.csv') #import studentscores.csv

dataset.shape

(25,2) #Rows 25 and 2 Columns

dataset.head() #First 5 Rows it will show

dataset.describe() #It will apply Default Statistical Calculations in our DataSet

dataset.plot(x='Hours', y='Scores', style="\*") #x - Hours, Y = Scores

plt.title('Student Mark Predition') # 2 Variables in Linear Regession

plt.xlabel('Hours') #Student Studied

plt.ylabel('Percentage Marks') #Percentage he Gets

plt.show() #Ploting process of DataSet, it will Shows in Graphical

X=dataset.iloc[:,:-1].values #Data into Independent Variables, Attributes

Y=dataset.iloc[:,:-1].values #Dependent Variables, Labels

from sklearn.model\_selection import train\_test\_split #Split this DataSet into Training and Testing Sets

X\_train, X\_test, Y\_train, Y\_test=train\_test\_split(X,Y,test\_size=0.2,random\_state=0) #80% Training Set and 20% Testing Set

from sklearn.linear\_model import LinearRegression #Train our Algorithm we Use Linear Regression

regressor=LinearRegression() #Linear Regression Class, Best for Intercept and Slope

regressor.fit(X\_train, Y\_train) #Fit Method along Training Data, Best Fits the Data

print(regressor.intercept\_) #Linear Regression is the Best Method for Intercept and Slope

print(regressor.coef\_) #If Student Studied more than 1 hour, they can expect 9.91% of Marks in the Exam

y\_pred=regressor.predict(X\_test) #Make some Predictions, and see how accurate the algorithm predicts the % score

df=pd.DataFrame(('Actual':Y\_test,'Predicted':Y\_pred))

df