**PRACTICAL NO – 1**

**TITLE : SIMPLE LINEAR REGRESSION**

**THEORY:**

**Simple linear regression** is a statistical method that allows us to summarize and study relationships between two continuous (quantitative) variables:

One variable, denoted *x*, is regarded as the **predictor**, **explanatory**, or **independent** variable.

The other variable, denoted *y*, is regarded as the **response**, **outcome**, or **dependent** variable.

The line for a simple linear regression model can be written as:

Y=b0+b1\*X

where b0 and b1 are the coefficients we must estimate from the training data. Once the coefficients are known, we can use this equation to estimate output values for y given new input examples of x.

**APPLICATIONS:-**

1) Linear Regression can be used to predict the sale of products in the future based on past buying behaviour.

2) An organisation can use linear regression to figure out how much they would pay to a new joining based on the years of experience.

3) Linear regression analysis can help a builder to predict how much houses it would sell in the coming months and at what price.

**PROBLEM STATEMENT**- BUILD A LINEAR REGRESSION MODEL (simple linear regression)

**ALGORITHM-**

1. **import libraries:**
2. NumPy (numerical Processing)
3. Matplotlib.pyplot (Plotting graphs)
4. Pandas (for data set processing)
5. **import dataset (excel file converted to .csv format)**
6. **dividing the data set into X and Y**
7. **Divide each attribute into training and testing datasets.**
8. **Fit the SLR model on the training set.**
9. **Predict the test set result.**
10. **Visualize the training set results.**
11. **Visualize the test set results.**
12. **End.**

**CODE: -**

# -\*- coding: utf-8 -\*-

"""

Created om Mon 02-Jul-2018

Simple linear Regression

@Author-RoKu

"""

**#step 1**

#import libraries

import pandas as pd

import matplotlib.pyplot as plt

**#step 2**

#import dataset

dataset = pd.read\_csv('Dataset\_flwr.csv');

X = dataset.iloc[:,:-1]

y = dataset.iloc[:,1]

**#step3**

#divide the dataset

from sklearn.model\_selection import train\_test\_split

x\_train,x\_test,y\_train,y\_test = train\_test\_split(X,y,test\_size=0.20,random\_state=42)

**#step 4**

#fit the model

from sklearn.linear\_model import LinearRegression

regressor = LinearRegression();

regressor.fit(x\_train,y\_train)

**#step 5**

#predict on test dataset

y\_pred = regressor.predict(x\_test);

y\_pred2 = regressor.predict(x\_train);

#miscellenous

#accuracy

accuracy = regressor.score(x\_test,y\_test)

**#step 6**

#visualizing the trained set result #run both together for getting combined

plt.scatter(x\_train,y\_train,color='red') #plots the visual result

plt.plot(x\_train,regressor.predict(x\_train),color='blue') ##plots the model line

plt.title('Sepal Length V/S Petal Length for training set')

plt.xlabel('Petal Length')

plt.ylabel('Sepal Length')

plt.show()

#visualizing the test set result #run both together for getting combined

plt.scatter(x\_test,y\_test,color='red') #plots the visual result

plt.plot(x\_train,regressor.predict(x\_train),color='blue') ##this is same as object regressor is the same

#NO need to train again |^^

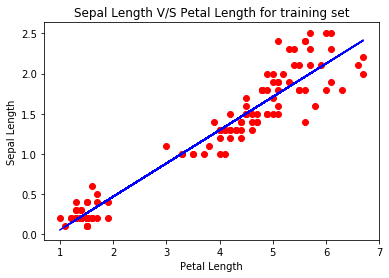
plt.title('Sepal Length V/S Petal Length for training set')

plt.xlabel('Petal Length')

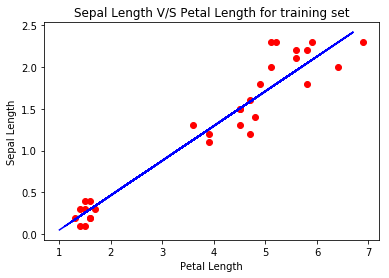
plt.ylabel('Sepal Length')

plt.show()

**RESULT:**



**Training Set**



**Testing Set**

**DISCUSSIONS:**

The graph of training set consists of 66% of total available samples and a graph has been plotted. After applying simple linear regression model, a line passes through the sample points indicating the optimum output level of the model.

The line is drawn in such a way that error is minimised. So, it tries to pass through points from which we get minimum error rate.

Same is the case for testing dataset. After learning the relationships from training datasets, it tries to pass a line through testing datapoints with minimal error.

**CONCLUSIONS:** Successfully acquired the required datasets, compiled the code and got the respective results. The values have been predicted using simple linear regression.

|  |  |
| --- | --- |
| **SUBMISSION DATE**-  8-02-19 | **SIGN OF COURSE INSTRUCTOR**- |
| **ROLL NO OF THE STUDENT-**  **01** | **NAME OF THE STUDENT-**  **Rohit Kulkarni** |