**ML LAB PROGRAMS:**

**Linear-Regression**

import pandas as pd

import matplotlib.pyplot as plt

df = pd.read\_csv("C:/Users/Asus/Desktop/students.csv")

df

df.plot(x="Hours", y="Scores", style="o")

plt.show()

x\_mean = df["Hours"].mean()

y\_mean = df["Scores"].mean()

print(x\_mean, y\_mean)

df["x"] = df["Hours"] - x\_mean

df["y"] = df["Scores"] - y\_mean

df["x\*y"] = df["x"] \* df["y"]

df["x^2"] = df["x"]\*\*2

df["y^2"] = df["y"]\*\*2

df

summation\_x\_y = df["x\*y"].sum()

summation\_x\_squared = df["x^2"].sum()

summation\_y\_squared = df["y^2"].sum()

print(summation\_x\_y, summation\_x\_squared, summation\_y\_squared)

correlation = summation\_x\_y / (summation\_x\_squared \* summation\_y\_squared)\*\*0.5

correlation

def getMean(numbers):

if len(numbers) == 0:

return None

else:

current\_sum = 0

for i in numbers:

current\_sum += i

current\_avg = current\_sum/len(numbers)

return current\_avg

def getStandardDeviation(numbers):

if len(numbers) == 0:

return 0

else:

mean = getMean(numbers)

std\_deviation = 0

for i in numbers:

std\_deviation += (i - mean)\*\*2

return (std\_deviation/len(numbers))\*\*0.5

std\_deviation\_x = getStandardDeviation(df["x"].tolist())

std\_deviation\_y = getStandardDeviation(df["y"].tolist())

print(std\_deviation\_x, std\_deviation\_y)

m = correlation \* (std\_deviation\_y / std\_deviation\_x)

m

c = df["Scores"].mean() - m \* df["Hours"].mean()

c

df["y\_prediction"] = m \* df["Hours"] + c

df

plot1 = plt.scatter(df["Hours"], df["Scores"])

plot2 = plt.scatter(df["Hours"], df["y\_prediction"])

plt.show()

output:

**Error:**

df["error"]=df["y"]-df["y\_prediction"]

df

**SSE:**

df["SSE"]=df["error"]\*df["error"]

df

**Y-ymean:**

df["y\_-ymean"]=df["y"]-y\_mean

df

**SST:**

df["SST"]=df["y\_-ymean"]\*df["y\_-ymean"]

df

**y prediction:**

df["yprediction"]=df["y\_prediction"].mean()

df

**Square of y prediction:**

df["yprediction-yprediction^2"]=df["y\_prediction"]-df["yprediction^2"]

df

**SSR:**

df["SSR"]=df["yprediction-yprediction^2"]\*df["yprediction-yprediction^2"]

df

**R^2:**

df["R^2"]=df["SSR"]/df["SST"]

df

**sum of SSR:**

df["sum of SSR"]=df["SSR"].sum()

df

out put















