SVKM’s NMIMS

Mukesh Patel School of Technology Management & Engineering

Department of Electronics and Telecommunication Engineering

Subject: Machine Learning Program: B.Tech/MBA.Tech

Sem: III/V ACAY: 2020-21

EXPERIMENT NO. 5

Aim:

1. To be able to perform multiple linear regression using sklearn and statsmodels.
2. To be able to interpret the results obtained from multiple linear regression.

Software: PYTHON.

Prerequisite:

|  |  |
| --- | --- |
| Sr. No | Concepts |
| 1. | Knowledge of multiple linear regression |

Outcome:

After successful completion of this experiment students will be able to:

1. Implement multiple linear regression by using sklearn package statsmodels.
2. Interpret the results obtained from different models and choose the best model for the given data set.

Theory:

* The data which we will be using for our multiple **linear regression** example is in a .csv file called: ‘MLR\_data.csv.
* For loading the data use the command: df=pd.read\_csv(MLR\_data.csv'

TO BE COMPLETED BY STUDENTS

* Students must upload the soft copy of the program in the given format.

|  |
| --- |
| Name of the Experiment: Experiment 5 – Multiple Linear Regression |
| Roll No.: N049 Name: Tarun Tanmay |
| Program : MBATech CE Semester: 5 |
| Date of Performance: 21/8/20 Date of Submission: 20/9/20 |

**Step 1: Importing the Relevant Libraries**

#Experiment Number 5

#Multiple Linear Regression

import pandas as pd

import statsmodels.api as sm

import matplotlib.pyplot as plt

from sklearn import linear\_model

import numpy as np

### ****Step 2: Loading the Data****

df=pd.read\_csv('MLR\_data.csv')

df

### ****Step 3: Visualizing the Data Frame****

### A screenshot of a cell phone Description automatically generated

### ****Step 4: Exploring the Data****

#Relationship between independent variable & dependent variable:

#Stock index price: dependent on interest rate and unemployment rate or not?

plt.figure(figsize=(10,10))

x1=df['Interest\_Rate']

x2=df['Unemployment\_Rate']

y=df['Stock\_Index\_Price']

plt.subplot(2,2,1)

plt.scatter(x1,y,color='red')

plt.xlabel('interest Rate')

plt.ylabel('Stock\_Index\_Price')

plt.title('IV(x1) to DV(y)')

plt.subplot(2,2,2)

plt.scatter(x2,y,color='green')

plt.xlabel('Unemployment Rate')

plt.ylabel('Stock\_Index\_Price')

plt.title('IV(x2) to DV(y)')

plt.subplot(2,2,3)

plt.scatter(x1,x2,color='blue')

plt.xlabel('Unemployment Rate')

plt.ylabel('Stock\_Index\_Price')

plt.title('IV(x1) to IV(x2)')

### ****Step 5: Plot the scatter plots of IV vs DV and DV vs DV****

**A screenshot of a social media post

Description automatically generated**

**Step 6: Fit a multiple linear regression model using sklearn**

X=df[['Interest\_Rate', 'Unemployment\_Rate']]

Y=df['Stock\_Index\_Price']

regr=linear\_model.LinearRegression()

regr.fit(X,Y)

print('Intercept:',regr.intercept\_)

print('Coefficent:',regr.coef\_)

b0=regr.intercept\_

b1=regr.coef\_[0]

b2=regr.coef\_[1]

New\_Interest\_Rate=2.75

New\_Unemployment\_Rate=5.3

New\_Stock\_Index\_Price=b0+(b1\*New\_Interest\_Rate)+(b2\*New\_Unemployment\_Rate)

print('New\_Stock\_Index\_Price: ', New\_Stock\_Index\_Price)

new\_stock\_price\_pred=regr.predict([[New\_Interest\_Rate,New\_Unemployment\_Rate]])

print('New Stock Index Price: ', new\_stock\_price\_pred)

A picture containing table

Description automatically generated



### ****Step 7: Get the model statistics using statsmodel****

X=sm.add\_constant(X)

model=sm.OLS(y,X).fit()

predictions=model.predict(X)

print\_model=model.summary()

print(print\_model)

A screenshot of text

Description automatically generated

### ****Step 8: Interpret the results and implement other models****

x1=np.array(x1)

x1=x1.reshape((-1,1))

regr1=linear\_model.LinearRegression()

regr1.fit(x1,y)

regr1\_b0=regr1.intercept\_

regr1\_b1=regr1.coef\_

print('Intercept:',regr.intercept\_)

print('Coefficent:',regr.coef\_)

New\_Interest\_Rate=2.75

New\_Unemployment\_Rate=5.3

New\_Stock\_Index\_Price=regr1\_b0+(regr1\_b1\*New\_Interest\_Rate)

print('New\_Stock\_Index\_Price: ', New\_Stock\_Index\_Price)

A picture containing table

Description automatically generated

model1=sm.OLS(y,x1).fit()

predictions1=model1.predict(x1)

model1.summary()

### A screenshot of text Description automatically generated

x2=np.array(x2)

x2 = x2.reshape((-1,1))

regr2 = linear\_model.LinearRegression()

regr2.fit(x2,y)

regr2\_b0 = regr2.intercept\_

regr2\_b1 = regr2.coef\_[0]

print('Intercept: ',regr2\_b0)

print('Slope: ',regr2\_b1)

New\_Unmployment\_Rate = 5.3

New\_Stock\_Index\_Price2 = regr2\_b0 + (regr2\_b1 \* New\_Unmployment\_Rate)

print('New Stock Index Price: ',New\_Stock\_Index\_Price2)

### A screenshot of a cell phone Description automatically generated

### ****Step 9: Choose the best possible model****

model2 = sm.OLS(y,x2).fit()

predictions2 = model2.predict(x2)

print\_model2 = model2.summary()

print(print\_model2)

### A screenshot of text Description automatically generated

### Step 10: Conclusion

### A screenshot of a cell phone Description automatically generated