DEPARTMENT OF INFORMATION SCIENCE AND ENGINEERING

A

MINI PROJECT REPORT

ON

"MULTIPLE LINEAR REGRESSION"

Submitted in the partial fulfillment of the requirements in the 6th semester of

BACHELOR OF

ENGINEERING IN

INFORMATION SCIENCE AND ENGINEERING

BY

VARNA MURALI

1NH17IS126

Under the guidance of

Mrs. KARTHIYAYINI J,

Senior. Assistant Professor,

Dept. of ISE, NHCE

DEPARTMENT OF INFORMATION SCIENCE AND ENGINEERING NEW HORIZON COLLEGE OF ENGINEERING

(Autonomous College Permanently Affiliated to VTU, Approved by AICTE, Accredited by NBA & NAAC with 'A' Grade)

Ring Road, Bellandur Post, Near Marathalli,

Bangalore-560103, INDIA

NEW HORIZON COLLEGE OF ENGINEERING

Autonomous College Permanently Affiliated to VTU, Approved by AICTE & UGC
Accredited by NAAC with 'A' Grade. Accredited by NBA

nirf TOP ENGINEERING COLLEGE OF INDIA as per NATIONAL INSTITUTIONAL RANKING FRAMEWORK, 2018
Ministry of Human Resource Development, Government of India

DEPARTMENT OF INFORMATION SCIENCE AND ENGINEERING

CERTIFICATE

I hereby certify that, the report entitled **"Multiple Linear Regression"** as a part of Mini Project Component in partial fulfillment of the requirements during 6th semester Bachelor of Engineering in Information Science and Engineering during the year 2019-2020(Mar 2018-May 2018) is an authentic record of my own work carried out by XXXX (1NH17is126), a bonafied student of NEW HORIZON COLLEGE OF ENGINEERING.

Name & Signature of Student

(Ms. Varna Murali)

Name & Signature of Guide

(Mrs. Karthiyayini J)

Name & signature of HOD

(Dr. R J Anandhi)

ACKNOWLEDGEMENT

Any achievement, be it scholastic or otherwise does not depend solely on the individual efforts but on the guidance, encouragement and cooperation of intellectuals, elders and friends. A number of personalities, in their own capacities have helped me in carrying out this project. I would like to take an opportunity to thank them all.

First and foremost I thank the management, **Dr. Mohan Manghnani,** Chairman, New Horizon Educational Institutions for providing necessary infrastructure and creating good environment.

I would like to thank **Dr. Manjunatha**, Principal, New Horizon College of Engineering, Bengaluru, for his constant encouragement and facilities extended to us towards completing my project work.

I extend my sincere gratitude to **Dr. R J Anandhi**, Head of the Department, Information Science and Engineering, New Horizon College of Engineering, Bengaluru for her valuable suggestions and expert advice.

I deeply express my sincere gratitude to my guide **Mrs. Karthiyayini J Sr. Assistant Professor**, Department of ISE, NHCE, Bengaluru, for her able guidance, regular source of encouragement and assistance throughout this project.

I thank my Parents, and all Faculty members of Department of Information Science and Engineering for their constant support and encouragement.

Last, but not the least, I would like to thank my peers and friends who provided me with valuable suggestions to improve my project.

Varna Murali

(1NH17IS126)

ABSTRACT

Machine Learning is the field of study that gives computers the ability to learn without being explicitly programmed. ML has become one of the most exiting technologies in today's developing world. Data Science uses many machine learning algorithms to discover patterns in big data that leads to actionable insights. Multiple linear regression is a multivariate statistical technique for examining the linear correlation between two or more independent variables and one dependent variable.

This project is used to predict the relationship between the dependent variable that is stock_index_price and the independent variables unemployment_rate and interest_rate. All the libraries are imported. The dataset for performing multiple linear regression analysis is imported from pandas library. Before doing the multiple linear regression analysis the linearity between the dependent variable and independent variable should be checked. If there exists a linear relationship only then multiple linear regression can be performed. In this project there exists a linear relation in positive slope between stock_index_price and interest_rate and a linear relationship between stock_index_price and unemployment_rate. This linearity check is done using the scatter diagram utilizing the matplotlib library. The multiple linear regression is performed using the sklearn library and the statsmodels. The output is displayed in GUI format using the Tkinter library. The output consists of two parts one is the graphical representation of the relationship between the stock_index_price and the unemployment_rate and the graph for relationship between the stock_index_price and the interest_rate. The second part asks the user to enter interest rate and unemployment rate. The present stock index price is calculated using the formula intercept+ (interest rate)*X_1 + (unemployment rate)*X_2 where X1 and X2 are coefficients. Interest rate and unemployment rate are found using the sklearn library.

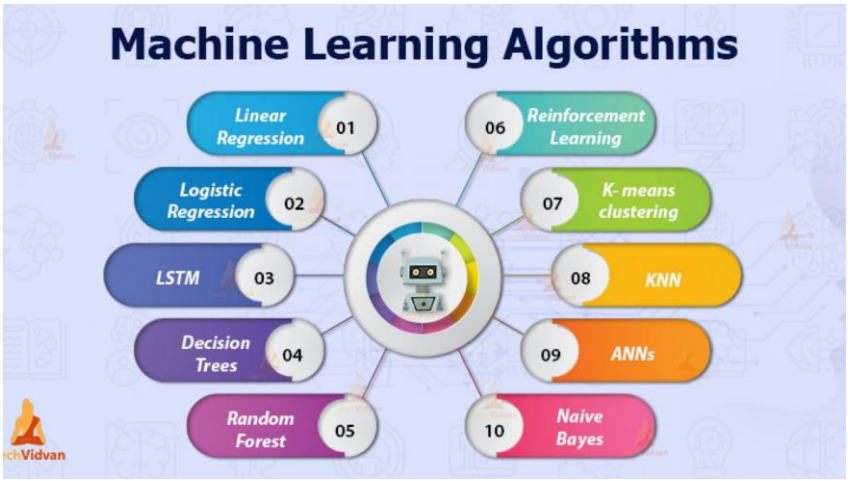
TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

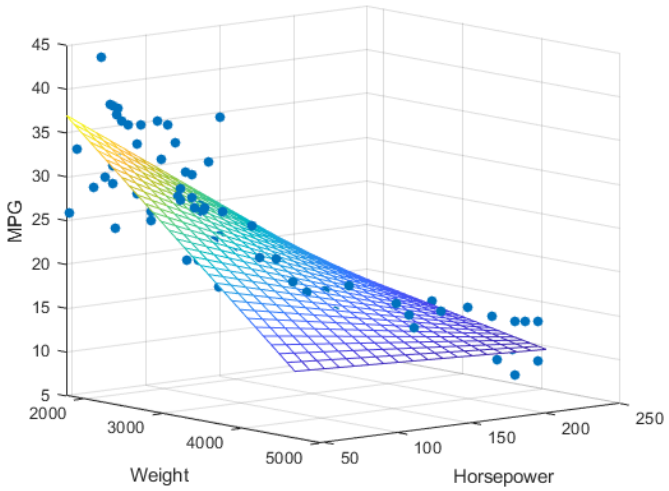
INTRODUCTION

Machine Learning is the field of study that gives computers the ability to learn without being explicitly programmed. ML has become one of the most exiting technologies in today's developing world. Data Science uses many machine learning algorithms to discover patterns in big data that leads to actionable insights. Machine learning algorithms can be classified as supervised learning algorithm and unsupervised learning algorithm. Machine learning can be classified in to two types supervised and unsupervised. Supervised learning is the task of learning that maps an input to an output based on example input-output pairs. Multiple linear regression falls under supervised learning. Regression is a supervised machine learning algorithm that essentially allows one to establish a relationship within a dataset.



1.1 Machine learning algorithm types

Multiple linear regression is a multivariate statistical technique for examining the linear correlation between two or more independent variables and one dependent variable. Multiple linear regression can be visualised as path diagrams and/or Venn diagrams. Multiple regression is an extension of simple linear regression. It is used when we want to predict the value of a variable based on the value of two or more other variables. The variable we want to predict is called the dependent variable (or sometimes, the outcome, target or criterion variable).



1.2 Multiple Linear Regression

Multiple linear equation = $y_i=\beta_0+\beta_1 x_{i1}+\beta_2 x_{i2}+...+\beta_p x_{ip}+\epsilon$

Where, for i=n observations:

y_i=dependent variable

x_i=explanatory variables

β_0=y-intercept (constant term)

β_p=slope coefficients for each explanatory variable

ϵ=the model's error term (also known as the residuals)

1.1 Purpose of Study:

In Multiple linear regression a linear combination of two or more predictor variables is used to explain the variation in a response. As a predictive analysis the multiple linear regression is used to explain the relationship between one continuous dependent variable and two or more independent variables. The independent variable can be continuous or categorical. The main purpose of multiple linear regression is to predict the value of a variable based on the value of other variables.

1.2 Problem Statement:

Multiple linear regression is used to find the mathematical relationship among a number of random variables. When there is a need to predict the output of a random variable based on other variable then multiple linear regression is used. Using MLR we can graphically represent the relationship between one input variable and more than output variables.

1.3 Motivation :

The object of multiple linear regression is to develop a prediction equation that permits the estimation of the value of the dependent variable based on the knowledge of multiple independent variables. It increases the understandability of the relationship between variables as they are graphically represented. Seeing the graph one can predict if the variables are directly related to each other or inter dependent on each other.

1.4 Methodology:

Multiple linear regression is a multivariate statistical technique for examining the linear correlation between two or more independent variables and one dependent variable. Multiple linear regression can be visualised as path diagrams and/or Venn diagrams. Multiple regression is an extension of simple linear regression.

For example, an analyst may want to know how the movement of the market affect the price of Exxon Mobil (XOM). In this case, his linear equation will have the value of the S&P 500 index as the independent variable, or predictor or the price of XOM as the dependent variable. In reality there are multiple factors that predict the outcome of an event. In reality, there are multiple factors that predict the outcome of an event. The price movement of Exxon Mobil, for example, depends on more than just the performance of the overall market. Other predictors such as the price of oil, interest rates, and the price movement of oil futures can affect the price of XOM and stock prices of other oil companies. To understand a relationship in which more than two variables are present, a multiple linear regression is used.

The formula according to example:

- y_i = dependent variable: price of XOM
- x_{i1} = interest rates
- x_{i2} = oil price
- x_{i3} = value of S&P 500 index
- x_{i4}= price of oil futures
- B_0 = y-intercept at time zero
- B1= regression coefficient

This is how multiple linear regression works. The above example to analyse the market effect of Exxon mobile shows how multiple linear regression works. The dependent variables and independent variables are substituted in the formula given and the multiple linear regression analysis is done.

This project is used to predict the relationship between the dependent variable that is stock_index_price and the independent variables unemployment_rate and interest_rate. All the libraries are imported. The dataset for performing multiple linear regression analysis is imported from pandas library. Before doing the multiple linear regression analysis the linearity between the dependent variable and independent variable should be checked. If there exists a linear relationship only then multiple linear regression can be performed. In this project there exists a linear relation in positive slope between stock_index_price and interest_rate and a linear relationship between stock_index_price and unemployment_rate. This linearity check is done using the scatter diagram utilizing the matplotlib library. The multiple linear regression is performed using the sklearn library and the statsmodels. The output is displayed in GUI format using the Tkinter library. The output consists of two parts one is the graphical representation of the relationship between the stock_index_price and the unemployment_rate and the graph for relationship between the stock_index_price and the interest_rate. The second part asks the user to enter interest rate and unemployment rate. The present stock index price is calculated using the formula intercept+ (interest rate)*X_1 + (unemployment rate)*X_2 where X1 and X2 are coefficients. Interest rate and unemployment rate are found using the sklearn library.

CHAPTER 2

SYSTEM REQUIREMENTS AND LANGUAGE USED

Every computer system requires certain set of software components as well as hardware components for the effective use and efficient performance. Thus, it becomes necessary that for any form of development the system requirements are met. These are prerequisites are often used as guideline and need not be absolute rules. There are two sets of system requirements: minimum and recommended, it essential to utilize minimum requirements for development and to contain recommended set of system requirements so as to maintain proper functioning of the system. It is also important to make sure that any software or hardware being used are safe for the system and have been standard certified.

2.1 HARDWARE AND SOFTWARE REQUIREMENTS:

Hardware Configuration:

Processor: Intel Core i3-380M dual-core processor

Ram: 1GB (32 bit) o2 3GB (64 bit)

Hard disk: 16GB

Software System Configuration:

Operating System: windows

Programming Language: python

Compiler: python idle

2.2 About the language:

Python is an interpreted high level programming language created by Guido Van Rossum. It was first released in 1991. Its design philosophy emphasizes code reusability with its notable use of significant uses of white spaces. Its language constructs and object oriented approach help programmers to write clear, logical code for small codes for large-scale products.Python is dynamically typed and garbage-collected. It supports multiple programming, paradigms, including structured object-oriented and functional programming.

Python today has multiple implementations including Jython, scripted in Java language for Java Virtual Machine*;* IronPython written in C# for the Common Language Infrastructure, and PyPy version written in RPython and translated into C. To be noted, Cpython which is written in C and developed by Python Software Foundation is the default and most popular implementation of Python.

Python features**:**

- ### Broad Standard Library:

  Python has huge collection of defined library which makes very easy to code in python. Its library is portable and compatible with all platforms like Macintosh, UNIX, and Windows. You don't have to write your own code for each and every thing as it provides rich sets of modules and functions. It has various libraries for web browsing, regular expressions, etc.

- ### Interpreted Language:

  Python is one of the Interpreted Language as its code is executed line by line at a time. It is not required to compile our code like in other languages like java, c++, etc. which makes it easier to debug our code. The source code of python is converted into an immediate form called byte code.

- ### Supports for GUI Programming:

  Python provides various modules like PyQt, Tkinter, wxPython through which user can created Graphical User Interface for mobile applications. The most popular for creating graphical apps using python is PyQt5. Tkinter also provides all of the required options to create a beautiful user interface even Gaffer is made importing Tkinter module. This user interface can be connected to the backend using any one of the DBMS also supported by python, makes it more beautiful.

- ### Object Oriented Programming Language:

Python is an object oriented programming language which include the concept of class and object. It supports all OOPs concepts like inheritance, data abstraction, polymorphism, encapsulation etc.

- Scalable and Extendable:

  Python provides a better structure and support for large programs than shell scripting. You can add low-level modules to the Python interpreter. These modules enable programmers to add to or customize their tools to be more efficient

CHAPTER 3

SYSTEM DESIGN

System design is the process of defining architecture, modules, interfaces, and data for a system to satisfy specified requirements. There is some overlap with the disciplines of system analysis, systems architecture and system engineering**.** System Analysis is the process that decomposes a system into its component pieces for the purpose of defining how well those components interact to accomplish the set requirements. This chapter consists of architecture, flowchart, and algorithm.

The libraries used are:

Pandas: pandas are a high level manipulation tool developed by Wes McKinney. It is built on the numpy package and its key data structure is called DataFrame. DataFrame allows you to store and manipulate tabular data in rows of observation and columns of variables. Pandas is a software library in computer programming. They are used in python to deal with data analysis and manipulation. Pandas is a fast powerful, flexible and easy to use open resource data analysis tool.

Sklearn: Scikit-learn is a free machine algorithm in python. It features various algorithms like support vector machine, random forests, and k-neighbouring, and it also supports python numerical and scientific libraries like Numpy and Scipy. It is probably the most useful library for machine learning in python. The sklearn library contains a lot of tools for machine learning and statistical modelling including classification, regression, clustering and dimensionality reduction.

Matplotlib: Matplotlib is a plotting library for the python programming language and its numerical mathematics extension of numpy. It provides an object-oriented API for embedding plots into applications using general purpose GUI toolkits like Tkinter, wxPython, QT or GKT+. It is used for 2D graphics in python. Matplotlib is a brainchild of John Hunter who along with its main contributors have put an immeasurable amount of time and effort into producing a piece of software used by thousands of scientists worldwide.

Statsmodels: Statsmodels is a python module that provides classes and functions for the estimation of many different statistical models, as well as conducting many statistical tests and statistical data exploration. An extensive list of result statistics are available for each estimator. The result are tested against existing statistical package to ensure that they are correct. Stats model supports specifying models using R-style formulas and pandas DataFrame.

Tkinter: The tkinter package is the standard python interface to the TK GUI toolkit. Both tk and Tkinter are available in most UNIX platform as well as on windows systems. Python when combined with Tkinter as fast and easy way to create GUI applications. Tkinter provides a powerful object oriented interface to the tk GUI toolkit.

3.1 Architecture:

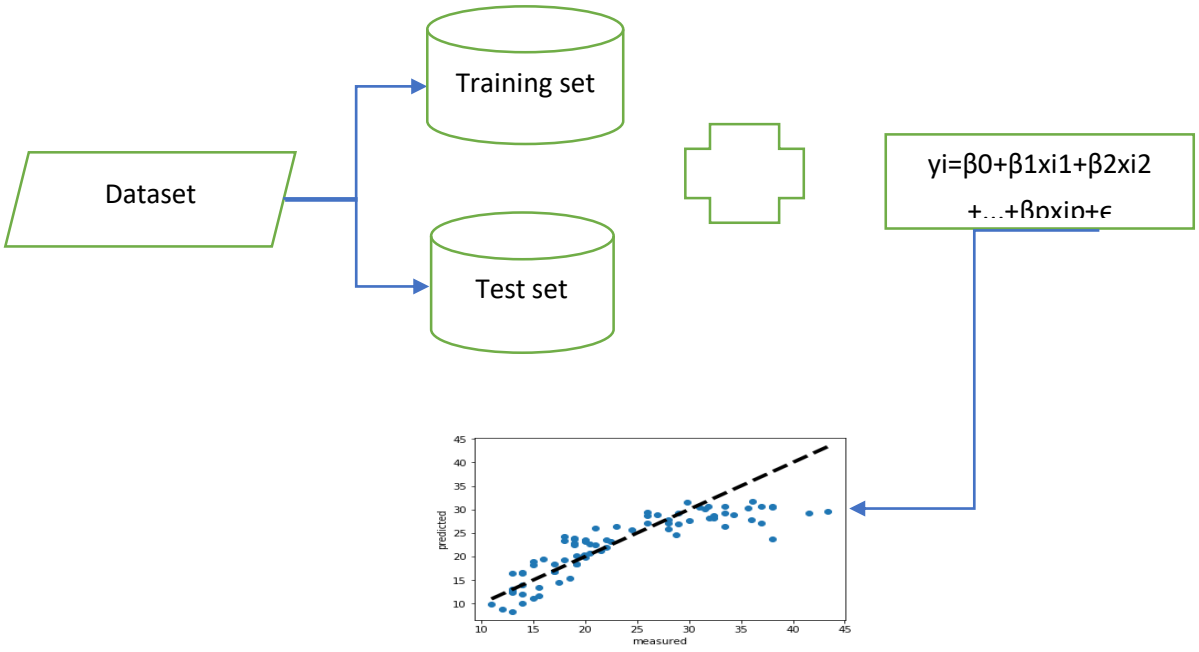


Figure 3.1 Architecture for multiple linear regression

This project is used to predict the relationship between stock_index_price and the unemployment_rate and stock_index_price and interest_rate. The dataset is imported from pandas library. Linearity between the dependent variable and the independent variables is checked using matplotlib library. The multiple linear regression analysis is done using sklearn and statsmodels. The relationship between the dependent variable and independent variables is displayed using tkinter.

3.2 Algorithm:

Step 1: Start

Step 2: import the libraries matplotlib, pandas, sklearn and statsmodels.

Step 3: Import the dataset from pandas library.

Step 4: Check linearity between stock_index_price and unemployment_rate.

Step 5: Check the linearity between stock_index_price and interest_rate.

Step 5: Fit the multiple linear regression model using sklearn library.

Step 6: Calculate the stock_index_price using the coefficients of interest rate and Unemployment rate

Step 7: Display the graphical plot for stock_index_price and unemployment_rate and The graphical plot for stock_index_price and interest_rate

Step 7: Stop

3.3 Flowchart:

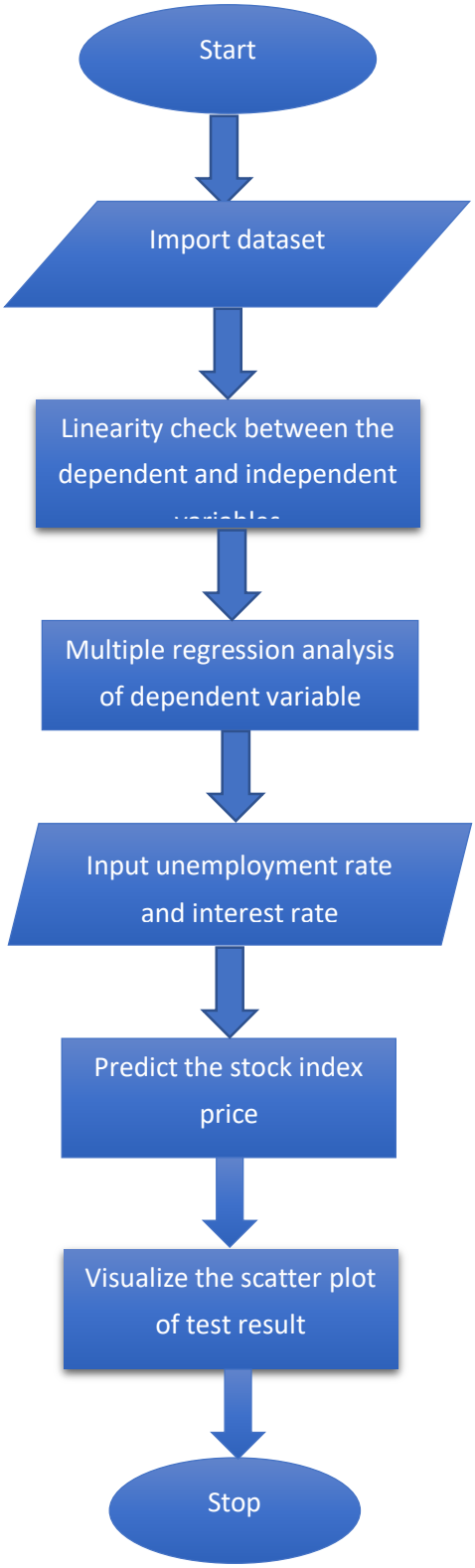


Figure 3.2 Flowchart of multiple linear regression

The above figure represents the diagrammatic representation of the algorithm. First import the libraries required. Then import the dataset from pandas library. Find the linearity between the stock_index_price (dependent variable) and unemployment_rate (independent variable) and the linearity between stock_index_price and interest_rate using matplotlib library. Predict the multiple linear regression using sklearn. Input the interest rate and unemployment rate and find the stock_index_price using the formula intercept+ (interest rate)*X_1 + (unemployment rate)*X_2 where X1 and X2 are coefficients. Display the graphical plot of stock_index_price and interest_rate and the stock_index_price and unemployment_rate.

3.4 Code:

```
From pandas import DataFrame

from sklearn import linear_model

import tkinter as tk

import matplotlib.pyplot as plt

from matplotlib.backends.backend_tkagg import FigureCanvasTkAgg

Stock_Market={'Year':[2017,2017,2017,2017,2017,2017,2017,2017,2017,2017,2017,2016,2016,2016,2016,2016,2016,2016,2016,2016,2016,2016],

 'Month': [12, 11,10,9,8,7,6,5,4,3,2,1,12,11,10,9,8,7,6,5,4,3,2,1],

'Interest_Rate':[2.75,2.5,2.5,2.5,2.5,2.5,2.5,2.25,2.25,2.25,2,2,2,1.75,1.75,1.75,1.75,1.75,1.75,1.75,1.75,1.75],

'Unemployment_Rate':[5.3,5.3,5.3,5.3,5.4,5.6,5.5,5.5,5.5,5.6,5.7,5.9,6,5.9,5.8,6.1,6.2,6.1,6.1,6.1,5.9,6.2,6.2,6.1],

'Stock_Index_Price':[1464,1394,1357,1293,1256,1254,1234,1195,1159,1167,1130,1075,1047,965,943,958,971,949,884,866,876,822,704,719]

 }

Df =DataFrame (Stock_Market, columns= ['Year','Month','Interest_Rate',' Unemployment_rate', 'Stock_Index_Price'])

X = df [['Interest_Rate', 'Unemployment_Rate']].astype (float) # here we have 2 input variables for multiple regression. If you just want to use one variable for simple linear regression, then use X = DF ['Interest_Rate'] for example. Alternatively, you may add additional variables within the brackets

Y = df ['Stock_Index_Price'].astype (float) # output variable (what we are trying to predict)

# with sklearn
```

```
Regr = linear_model.LinearRegression ()

regr.fit(X, Y)

Print ('Intercept: \n', regr.intercept_)

Print ('Coefficients: \n', regr.coef_)

# Tkinter GUI

Root= tk.Tk ()

canvas1 = tk.Canvas (root, width = 500, height = 300)

canvas1.pack ()

# with sklearn

Intercept_result = (‘intercept: ‘regr.intercept_)

label_Intercept = tk.Label (root, text=Intercept_result, justify = 'center')

canvas1.create_window (260, 220, window=label_Intercept)

# with sklearn

Coefficients_result = ('Coefficients: ‘regr.coef_)

label_Coefficients = tk.Label (root, text=Coefficients_result, justify = 'center')

canvas1.create_window (260, 240, window=label_Coefficients)

# New_Interest_Rate label and input box

label1 = tk.Label (root, text='Type Interest Rate: ')

canvas1.create_window (100, 100, window=label1)

entry1 = tk.Entry (root) # create 1st entry box

canvas1.create_window (270, 100, window=entry1)
```

```
# New_Unemployment_Rate label and input box

label2 = tk.Label (root, text=' Type Unemployment Rate: ')

canvas1.create_window (120, 120, window=label2)

entry2 = tk.Entry (root) # create 2nd entry box

canvas1.create_window (270, 120, window=entry2)

def values ():

global New_Interest_Rate #our 1st input variable

 New_Interest_Rate = float (entry1.get ())

 global New_Unemployment_Rate #our 2nd input variable

  New_Unemployment_Rate = float (entry2.get ())

  Prediction_result = ('Predicted Stock Index Price: ‘regr.predict ([[New_Interest_Rate, New_Unemployment_Rate]]))

 label_Prediction = tk.Label (root, text= Prediction_result, bg='orange')

  canvas1.create_window (260, 280, window=label_Prediction)

 button1 = tk.Button (root, text='Predict Stock Index Price’, command=values, bg='orange')

# button to call the 'values' command above

canvas1.create_window (270, 150, window=button1)


#plot 1st scatter

figure3 = plt.Figure (figsize= (5, 4), dpi=100)

ax3 = figure3.add_subplot (111)

ax3.scatter (df ['Interest_Rate'].astype (float), df ['Stock_Index_Price'].astype (float), color = 'r')

scatter3 = FigureCanvasTkAgg (figure3, root)
```

```
scatter3.get_tk_widget ().pack (side=tk.RIGHT, fill=tk.BOTH)

ax3.legend ()

ax3.set_xlabel ('Interest Rate')

ax3.set_title ('Interest Rate vs. Stock Index Price')


#plot 2nd scatter

figure4 = plt.Figure (figsize= (5, 4), dpi=100)

ax4 = figure4.add_subplot (111)

ax4.scatter (df['Unemployment_Rate'].astype(float),df['Stock_Index_Price'].astype(float),
color = 'g')

scatter4 = FigureCanvasTkAgg (figure4, root)

scatter4.get_tk_widget ().pack (side=tk.RIGHT, fill=tk.BOTH)

ax4.legend ()

ax4.set_xlabel ('Unemployment_Rate')

ax4.set_title ('Unemployment_Rate vs. Stock Index Price')

root.mainloop ()
```

CHAPTER 4

RESULTS AND DISCUSSION

Results and discussion has an important role in documentation. This chapter consists the summary of the results obtained and output of the project. From this chapter the idea of the project and what the code actually does can be determined. A clear idea about the code and its output will be received from this chapter.

4.1 Summary Of Results Obtained:

This project is used to predict the relationship between a dependent variable and two independent variables. The dependent variable is stock_index_price and the independent variables are unemployment_rate and interest_rate. First check the linearity between the dependent variable and the independent variables. Multiple linear regression can be performed only if there exists a linear relationship. This is done using matplotlib library. There exists a linear relationship in positive slope between stock_index_price and interest_rate and a linear relationship in negative slope between stock_index_price and unemployment_rate. Multiple linear regression is calculated using sklearn and statsmodels. The unemployment rate and interest rate is entered by the user and the stock_index_price is calculated using the formula intercept+ (interest rate)*X_1 + (unemployment rate)*X_2 where X1 and X2 are coefficients. The relationship and stock_index_price is displayed through GUI using Tkinter.

Year	Month	Interest_Rate	Unemployment_Rate	Stock_Index_Price
2017	12	2.75	5.3	1464
2017	11	2.5	5.3	1394
2017	10	2.5	5.3	1357
2017	9	2.5	5.3	1293
2017	8	2.5	5.4	1256
2017	7	2.5	5.6	1254
2017	6	2.5	5.5	1234
2017	5	2.25	5.5	1195
2017	4	2.25	5.5	1159
2017	3	2.25	5.6	1167
2017	2	2	5.7	1130
2017	1	2	5.9	1075
2016	12	2	6	1047
2016	11	1.75	5.9	965
2016	10	1.75	5.8	943
2016	9	1.75	6.1	958
2016	8	1.75	6.2	971
2016	7	1.75	6.1	949
2016	6	1.75	6.1	884
2016	5	1.75	6.1	866
2016	4	1.75	5.9	876
2016	3	1.75	6.2	822
2016	2	1.75	6.2	704
2016	1	1.75	6.1	719

4.1 Dataset used for the project

4.2 Output Snapshot:

4.2 Graphical representation

The relationship between the dependent variable i.e. stock_index_price and the independent variables i.e. unemployment_rate and interest _rate is displayed. There exists a linear relationship in a positive slope between stock_index_price and interest_rate. There exists a linear relationship in negative slope between stock_index_price and unemployment_rate.

4.3Calculate the stock_index_price

The user is asked to enter the interest rate and unemployment rate and the stock index price is calculated using the formula intercept+ (interest rate)*X_1 + (unemployment rate)*X_2 where X1 and X2 are coefficients.

CHAPTER 5

CONCLUSION

Multiple linear regression is the most common form of linear regression analysis. As a predictive analysis, the multiple linear regression is used to explain the relationship between one continuous dependent variable and two or more independent variables. Multiple Linear Regression Analysis consists of more than just fitting a linear line through a cloud of data points. It consists of three stages: 1) analysing the correlation and directionality of the data, 2) estimating the model, i.e., fitting the line, and 3) evaluating the validity and usefulness of the model. Multiple linear regression should be able to identify the strength of the effect that the independent variables has on dependent variable. It is also used to forecast the effect or impact of changes that is multiple linear analysis helps us to find out how much the dependent variable changes when there is a change in the dependent variable. Multiple linear regression analysis can also predict the future trends or changes. The project helps us to predict the relationship between stock index price and the unemployment rate and the stock index price and interest rate. With the help of coefficients it predicts the stock index price.

REFERENCE

[1] Towardsdatascience.com

[2] www.researchgate.net

[3] www.geeksforgeeks.org

MULTIPLE LINEAR REGRESSION

ORIGINALITY REPORT

25%	%	25%	%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Satheesh Abimannan, Yue-Shan Chang, Chi-Yeh Lin. "Chapter 25 Air Pollution Forecasting Using LSTM-Multivariate Regression Model", Springer Science and Business Media LLC, 2020 Publication	3%
2	Manoj S. Sankhe, Kamalakar D. Desai. "Chapter 39 Fetal Heart Rate Variability: Multiple Regression Models Using Autoregressive Analysis and Fast Fourier Transform", Springer Science and Business Media LLC, 2016 Publication	2%
3	Sofianita Mutalib, Nor Aina Azman, Shuzlina Abdul-Rahman. "Predicting patients survival using supervised techniques", 2011 11th International Conference on Hybrid Intelligent Systems (HIS), 2011 Publication	2%
4	Thomas Haslwanter. "An Introduction to Statistics with Python", Springer Science and	2%

Business Media LLC, 2016
Publication

---

5 Vandana Bhagat, Bastin Robins, Pallavi M.O. "Sparx - Data Preprocessing Module", 2019 IEEE 5th International Conference for Convergence in Technology (I2CT), 2019 — 1%
Publication

---

6 Badreddine Boudriki Semlali, El Amrani Chaker. "Towards Remote Sensing Datasets Collection and Processing", International Journal of Embedded and Real-Time Communication Systems, 2019 — 1%
Publication

---

7 Sachin Kamley, Shailesh Jaloree, R. S. Thakur. "chapter 10 Bombay Stock Exchange of India", IGI Global, 2017 — 1%
Publication

---

8 James Liech Gor. "Agent-based interoperability system in health insurance", 2017 IST-Africa Week Conference (IST-Africa), 2017 — 1%
Publication

---

9 Xiong Xiong, Ding Nan, Yang Yang, Zhang Yongjie. "Study on Market Stability and Price Limit of Chinese Stock Index Futures Market: An Agent-Based Modeling Perspective", PLOS ONE, 2015 — 1%
Publication

10	P.S. Maya Gopal, R. Bhargavi. "A novel approach for efficient crop yield prediction", Computers and Electronics in Agriculture, 2019 Publication	1%
11	Santosh Kumar Nanda, Arun Arun, Athira G S G S, Niranjan Kumar Ray. "Development of Novel Ensemble Machine Learning Architecture for Forecasting Unresponsive Server State", 2019 International Conference on Information Technology (ICIT), 2019 Publication	1%
12	Hasan U. Zaman, Saif Mahmood, Sadat Hossain, Iftekharul Islam Shovon. "Python Based Portable Virtual Text Reader", 2018 Fourth International Conference on Advances in Computing, Communication & Automation (ICACCA), 2018 Publication	1%
13	M Kavya, C Kishore, K S Rajath Kumar. "Verification and validation of data on Wi-Fi Over The Air(OTA)", 2018 3rd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), 2018 Publication	1%
14	Xiaofei Zhang, Jiajin Huang, Yang Yang, Xiang He, Ruohao Liu, Ning Zhong. "Applying Python	1%

in Brain Science Education", 2019 International Joint Conference on Information, Media and Engineering (IJCIME), 2019
Publication

---

15 Andreas François Vermeulen. "Industrial Machine Learning", Springer Science and Business Media LLC, 2020 — 1%
Publication

---

16 Shu-Hong Zhu, Jichao Sun, Suzanne C Billings, Won S Choi, Ann Malarcher. "Predictors of smoking cessation in U.S. adolescents", American Journal of Preventive Medicine, 1999 — 1%
Publication

---

17 Fu-Yun Zhao, Guang Shen, Ke-Jun Liu, Ying Xu, Di Liu, Han-Qing Wang. "Room airborne pollutant separation by the use of air curtains in the large building enclosure: Infiltration efficiency and partial enclosure ventilation rate", Journal of Building Engineering, 2018 — 1%
Publication

---

18 Maria Schuld, Ilya Sinayskiy, Francesco Petruccione. "An introduction to quantum machine learning", Contemporary Physics, 2014 — 1%
Publication

---

19 Andreas François Vermeulen. "Practical Data Science", Springer Science and Business Media LLC, 2018 — 1%

Publication

20	Vighnesh, Sanjana D. "Traffic Prediction Using a Supervised Learning Approach", 2018 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT), 2018 Publication	<1%
21	"Proceedings of International Conference on Remote Sensing for Disaster Management", Springer Science and Business Media LLC, 2019 Publication	<1%
22	Tang, Wenlong, Ting Zhang, and Edward Sazonov. "The Application of Machine Learning in Monitoring Physical Activity with Shoe Sensors", Cyber-Physical Systems Integrated Computing and Engineering Design, 2013. Publication	<1%
23	Santiago L. Rovere, Michael J. North, Guillermo P. Podesta, Federico E. Bert. "Practical Points for the Software Development of an Agent-Based Model of a Coupled Human-Natural System", IEEE Access, 2016 Publication	<1%
24	Xiong Xiong, Juan Liang, Yian Cui, Wei Zhang, Yongjie Zhang. "Analysis of the Spot Market's T+1 Trading System Effects on the Stock Index	<1%

Futures Market", EURASIA Journal of Mathematics, Science and Technology Education, 2017
Publication

25 Lu Xiao, Yichao Lang, George Christakos. "High-resolution spatiotemporal mapping of PM2.5 concentrations at Mainland China using a combined BME-GWR technique", Atmospheric Environment, 2018 <1%
Publication

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	On		