[STOCK MARKET ANALYSIS]

A Project Report Submitted in Partial Fulfillment of the Requirement for the Award of the degree of

Bachelor of Business Administration (Fintech)

Submitted By

[T. Hari Vamsi] Reg.No: <u>122014101003</u>

Under the guidance of

[Mr. V.V.L. Divakar Allavarapu] [Programme Coordinator]



GITAM SCHOOL OF BUSINESS GITAM (DEEMED TO BE University)

Rushikonda

Visakhapatnam-530045.

(2020 - 2023)

CERTIFICATE
This is to certify that the Internship Report entitled
"Stock Market Analysis"
is a genuine and bonafide work prepared in partial fulfillment of the requirements of BBA degree of
GITAM (Deemed to be University), Visakhapatnam was carried out by Mr./Miss. T. Hari Vamsi,
Reg No 12201401003 under my guidance and supervision.
Visakhapatnam
Date: [Signature]
[Guide name]

DECLARATION

I hereby declare that the project work/Internship report entitled "<u>Stock Market Analysis</u>" an original report submitted by me to words the partial fulfilment for the award of degree Bachelor of Business Administration, GITAM (Deemed to be University), Visakhapatnam and it is not submitted anywhere either in part or in full for degree or post-graduation of any university.

Place: Visakhapatnam

Date:

Signature

[T.Hari Vamsi] [122014101003]

Table Of Contents

- 1. Abstract
- 2. Introduction
- 3. Stock Price Prediction Model Process
 - a. Raw Data
 - b. Data Processing
 - c. Feature Extraction
 - d. Training Neural Network
 - e. Output Generation
 - f. Visualization
- 4. Stock Price Prediction Model Methodologies
 - a. Regression Based Model
- 5. Economic Factors
- 6. Conclusion
- 7. References

Stock Market Price Analysis

Abstract

In Stock Market Prediction, the aim is to predict the future value of the financial stocks index of Nifty50. The recent trend in stock market prediction technologies is the use of machine learning which makes predictions based on the values of current stock market indices by training on their previous values.

Introduction

The purpose of Stock Market Prediction is to forecast the value of Nifty50's financial stock in the future. The use of Artificial Intelligence and machine learning to create forecast values of current stock indices by training on their prior values is a recent trend in market forecasting technologies. Retail investors devote a significant amount of effort to researching investing alternatives. Professional financial counseling services are available to wealthy investors, but the charges are exorbitant for most regular investors. As a result, ordinary investors must educate themselves about the market and make their own decisions. In modern society, this makes investing extremely stressful. Humans, unfortunately, are illogical creatures. Decisions are impacted by cognitive biases or emotions without quantitative, data-driven models, resulting in avoidable losses. Even if investors are cautious enough, the majority lack the abilities necessary to evaluate the massive amounts of data required to make sound decisions. By researching the history of the specific stock market, precise predictions can be made. Artificial Intelligence and Machine learning is a good technique to express these types of operations. It forecasts a market value that is near to the tangible worth, which improves the accuracy.

Stock Price Prediction Model - Process

1. Raw Data: Raw market data feeds are generated by market exchanges and other trading venues for stocks, commodities, foreign currencies, options, futures, and other assets and financial instruments. Data feeds represent a market's activity from the perspective of the provider's internal software systems. It's critical that the sources of information used are reliable and well-constructed, so that the data collected (and later used as information) is of the highest possible quality. Use historical data collection and reporting to gather useful metrics about your managed network.

• Nifty 50

The NIFTY 50 is a benchmark Indian stock market index that represents the weighted average of 50 of the largest Indian companies listed on the National Stock Exchange. It is one of the two main stock indices used in India, the other being the BSE SENSEX.

Nifty 50 is owned and managed by NSE Indices (previously known as India Index Services & Products Limited), which is a wholly owned subsidiary of the NSE Strategic Investment Corporation Limited. NSE Indices had a marketing and licensing agreement with Standard & Poor's for co-branding equity indices until 2013. The Nifty 50 index was launched on 22 April 1996, and is one of the many stock indices of Nifty.

Crude Oil

Petroleum, also known as crude oil, or simply oil, is a naturally occurring yellowish-black liquid mixture of mainly hydrocarbons, [1] and is found in geological formations. The name petroleum covers both naturally occurring unprocessed crude oil and petroleum products that consist of refined crude oil. A fossil fuel, petroleum is formed when large quantities of dead organisms, mostly zooplankton and algae, are buried underneath sedimentary rock and subjected to both prolonged heat and pressure.

Petroleum has mostly been recovered by oil drilling. Drilling is carried out after studies of structural geology, sedimentary basin analysis, and reservoir characterisation. Recent developments in technologies have also led to exploitation of other unconventional reserves such as oil sands and oil shale. Once extracted, oil is refined and separated, most easily by distillation, into numerous products for direct use or use in manufacturing. Products include fuels such as gasoline (petrol), diesel and kerosene; asphalt; and chemical reagents used to make plastics, pesticides and pharmaceuticals. Petroleum is used in manufacturing a wide variety of materials, [2] and it is estimated that the world consumes about 100 million barrels (16 million cubic metres) each day. Petroleum production can be extremely profitable and was important for economic development in the 20th century, with some countries, so called "oil states", gaining significant economic and international power because of their control of oil production.

Gold Rate

Of all the precious metals, gold is the most popular as an investment. Investors generally buy gold as a way of diversifying risk, especially through the use of futures contracts and derivatives. The gold market is subject to speculation and volatility as are other markets. Compared to other precious metals used for investment, gold has been the most effective safe haven across a number of countries.

London Stock Exchange (L.S.E.)

London Stock Exchange (LSE) is a stock exchange in the City of London, England, United Kingdom. As of November 2021, the total market value of all companies trading on LSE was £3.9 trillion Its current premises are situated in Paternoster Square close to St Paul's Cathedral in the City of London. Since 2007, it has been part of the London Stock Exchange Group

Spot Rate

The spot market or cash market is a public financial market in which financial instruments or commodities are traded for immediate delivery. It contrasts with a futures market, in which delivery is due at a later date. In a spot market, settlement normally happens in T+2 working days, i.e., delivery of cash and commodity must be done after two working days of the trade date. A spot market can be through an exchange or over-the-counter (OTC). Spot markets can operate wherever the infrastructure exists to conduct the transaction.

Nasdaq 100

The Nasdaq-100 (^NDX) is a stock market index made up of 102 equity securities issued by 101 of the largest non-financial companies listed on the Nasdaq stock exchange. It is a modified capitalization-weighted index. The stocks' weights in the index are based on their market capitalizations, with certain rules capping the influence of the largest components. It is limited to companies from a single exchange, and it does not have any financial companies. The financial companies are in a separate index, the NASDAQ Financial-100.

S & P 500

The Standard and Poor's 500, or simply the S&P 500, is a stock market index tracking the stock performance of 500 large companies listed on stock exchanges in the United States. It is one of the most commonly followed equity indices. As of December 31, 2020, more than \$5.4 trillion was invested in assets tied to the performance of the index.

The S&P 500 index is a free-float weighted/capitalization-weighted index. As of August 31, 2022, the nine largest companies on the list of S&P 500 companies accounted for 27.8% of the market capitalization of the index and were, in order of highest to lowest weighting: Apple, Microsoft, Alphabet (including both class A & C shares), Amazon.com, Tesla, Berkshire Hathaway, UnitedHealth Group, Johnson & Johnson and ExxonMobil. The components that have increased their dividends in 25 consecutive years are known as the S&P 500 Dividend Aristocrats.

The index is one of the factors in computation of the Conference Board Leading Economic Index, used to forecast the direction of the economy.

The index is associated with many ticker symbols, including GSPC, .INX, and \$SPX, depending on market or website.

The S&P 500 is maintained by S&P Dow Jones Indices, a joint venture majority-owned by S&P Global, and its components are selected by a committee.

2. Data Processing: After the data has been acquired, the data preparation stage begins, when raw data is cleaned up and structured in preparation for the next stage of data processing. The pre-processing stage involves Data discretization, data transformation, data cleaning, data integration. After the dataset is transformed into a clean dataset, the dataset is divided into training and testing sets to evaluate. When data is collected and transformed into usable information, it is called data processing. Data processing is usually done by a data scientist or a team of data scientists, and it is critical that it is done correctly so that the end product, or data output, is not harmed. Data processing takes raw data and puts it into a more legible format (graphs, papers, etc.) so that computers can comprehend it and personnel throughout an organization can use it.

3. Feature Extraction: Feature extraction refers to the process of transforming raw data into numerical features that can be processed while preserving the information in the original data set. It yields better results than applying machine learning directly to the raw data. When you need to reduce the number of resources needed for processing without losing significant or relevant information, feature extraction comes in handy. Feature extraction can also reduce the amount of redundant data for a given analysis. Furthermore, the reduction of data and the machine's efforts in constructing variable combinations (features) speed up the learning and generalization stages of the machine learning process. Image Processing is one of the practical examples of feature extraction where the Algorithms are used to detect features such as shaped, edges, or motion in a digital image or video.

4 Main Features of Stock Exchange

- (1) Organised Market
- (2) Dealings in Securities Issued by Various Concerns
- (3) Dealing only through Authorised Members
- (4) Necessary to Obey the Rules and Bye-laws
- **4. Training Neural Network:** Backpropagation is the most common training algorithm for neural networks. It makes gradient descent feasible for multi-layer neural networks. steps for training a neural network:
- Data Set.
- Neural Networks.
- Training Strategy.
- Model Selection.
- Testing Analysis.
- Model Development.

5. Output Generation:

The output/interpretation stage is where non-data scientists can finally use the data. It is translated, readable, and frequently takes the shape of graphs, videos, photos, plain text, and other formats. Members of the firm or institution can now use the data for their own data analytics initiatives by self-serving it.

6. Visualization:

Data visualization uses algorithms to generate visuals from data so that humans may better comprehend and respond to it. The goal of artificial intelligence research is to create algorithms that can "understand" and respond to data as well as a human — if not better. Data storage is the ultimate stage of data processing. After all of the data has been analyzed, it is saved for future reference. While some information will be useful right away, most of it will be useful later. Furthermore, complying with data protection legislation necessitates correctly maintained data. When data is correctly saved, individuals of the organization may access it quickly and readily when they need it. A rolling analysis of a time series model is frequently used to determine the model's long-term stability. When using a statistical model to analyze financial time series data, one crucial assumption is that the model's parameters remain constant across time.

Machine Learning:

Machine learning (ML) is a field of inquiry devoted to understanding and building methods that 'learn', that is, methods that leverage data to improve performance on some set of tasks. [1] It is seen as a part of artificial intelligence. Machine learning algorithms build a model based on sample data, known as training data, in order to make predictions or decisions without being explicitly programmed to do so. Machine learning algorithms are used in a wide variety of applications, such as in medicine, email filtering, speech recognition and computer vision, where it is difficult or unfeasible to develop conventional algorithms to perform the needed tasks.

A subset of machine learning is closely related to computational statistics, which focuses on making predictions using computers, but not all machine learning is statistical learning. The study of mathematical optimization delivers methods, theory and application domains to the field of machine learning. Data mining is a related field of study, focusing on exploratory data analysis through unsupervised learning. Some implementations of machine learning use data and neural networks in a way that mimics the working of a biological learning. In its application across business problems, machine learning is also referred to as predictive learning.

Stock Price Prediction Model – Methodology

Regression Based Model: The regression equation is solved to find the coefficients, by using those coefficients we predict the future price of a stock. Regression analysis is a statistical tool for investigating the relationship between a dependent or response variable and one or more independent variables.

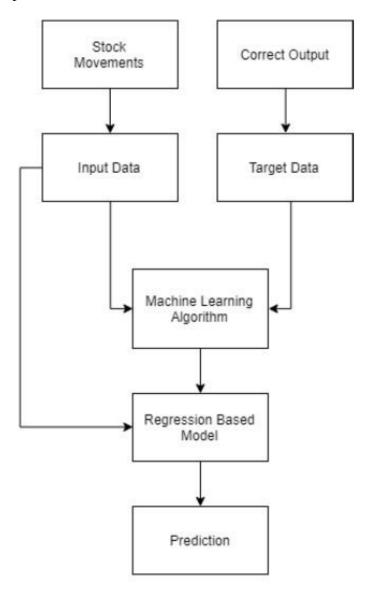


Figure 1. Flow Chart for Regression Based Model

<u>Neural Networks</u>: Artificial Neural Network and Recurrent Neural Network are used to explain the stock price prediction using AI in this report.

The organization of Neural Networks can imitate the interaction of organic nervous systems with real-world objects. NN are broadly parallel linked networks made of simple adaptable components. Each neuron in a biological neural network is connected to other neurons, and when a neuron is activated, it transfers chemicals to the related neurons, changing their potential. When a neuron's potential exceeds a certain level, it is activated and sends chemicals to neighboring neurons. This neuron model, for example, can be expressed as an M-P network with one neuron and three input nodes when abstracted. The neuron takes input x and transfers it through the link with weight w, as can be shown. The whole input signal is then compared to the neuron threshold to determine if it is engaged via activation function processing.

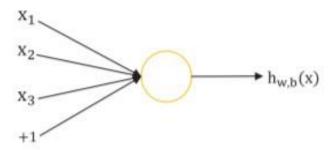


Figure 1. M-P neural network with one neuron and three input nodes

Artificial neural networks (ANNs) have advanced rapidly in recent decades thanks to their usage of neural architecture. ANNs can learn and generalize from their experience building complicated neural networks to deal with non-linear interactions between complex time series, making them an excellent approximator in predicting situations.

ECONOMIC FACTORS

Economic factors affect the economy, including interest rates, tax rates, laws, policies, wages, and governmental activities. These factors are not directly related to the business but influence the investment value in the future. There are three types of economic factors they are

- Macro Economic Factors
- Micro Economic Factors
- Social Economic Factors
- 1. Interest Rate
- 2. Exchange Rate
- 3. Inflation Rate
- 4. GDP Growth Rate
- 5. Unemployment Rate

Macroeconomic factors: •Unemployment rate •GDP growth rate •Interest rates •Inflation rate • Exchange rates Microeconomic factors: Level of education Gender •Income level Debts Social factors: • Materialism •Self-control, impatience Financial awareness Health conditions Religious beliefs

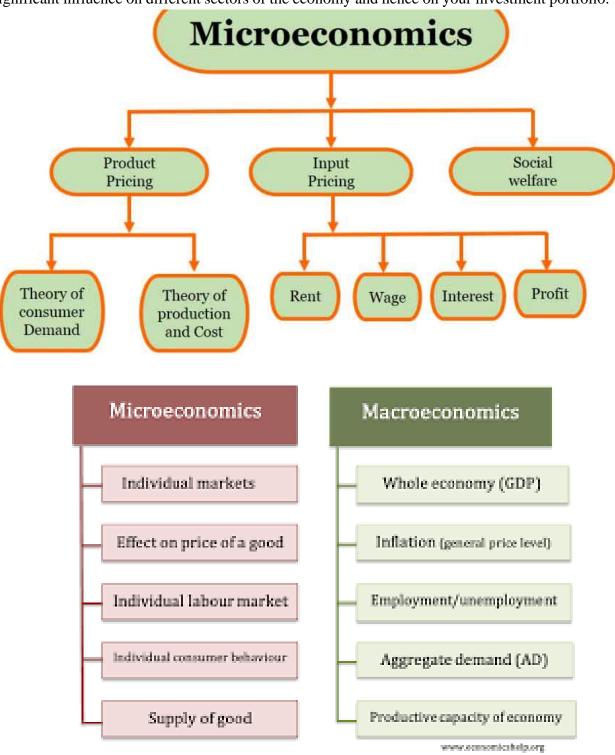
Macroeconomic factor:

Macroeconomic factors tend to impact wide swaths of populations, rather than just a few select individuals. Examples of macroeconomic factors include economic outputs, unemployment rates, and inflation. These indicators of economic performance are closely monitored by governments, businesses and consumers alike.



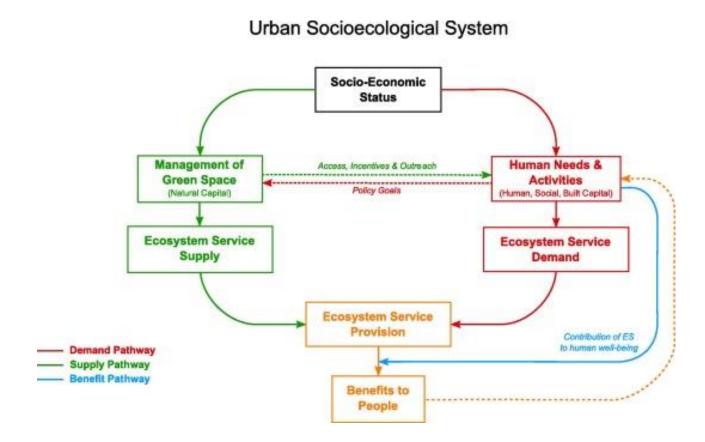
Microeconomic factor:

Microeconomic factors such as supply and demand, taxes and regulations, and macroeconomic factors such as gross domestic product (GDP) growth, inflation, and interest rates, have a significant influence on different sectors of the economy and hence on your investment portfolio.



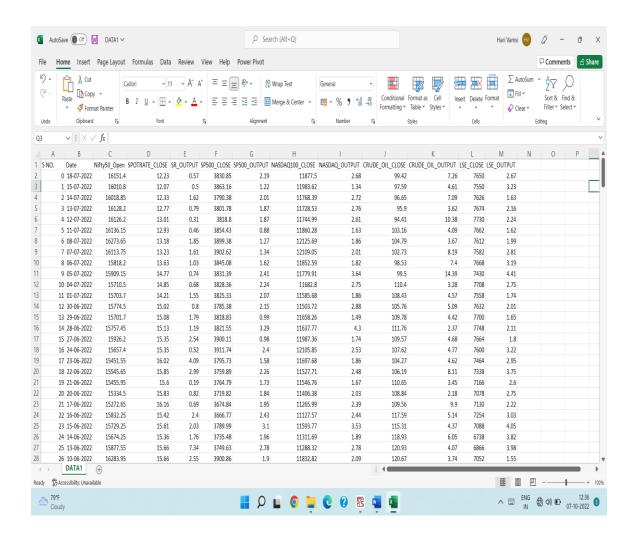
Social and economic factor:

Social and economic factors s, such as income, education, employment, community safety, and social supports can significantly affect how well and how long we live. These factors affect our ability to make healthy choices, afford medical care and housing, manage stress, and more.



STOCKS:

- Nifty 50
- Crude Oil
- Gold Rate
- London Stock Exchange (L.S.E.)
- Spot Rate
- Nasdaq 100
- S & P 500



CODE:

```
import pandas as pd
x=pd.read csv("D:\\New data\\DATA1.csv")
df = pd.DataFrame(x)
Y = df['Nifty50\_Open']
X = df[[ 'SPOTRATE_CLOSE',
   'SP500_CLOSE', 'SP500_OUTPUT',
   'NASDAQ_OUTPUT', 'CRUDE_OIL_CLOSE',
   'LSE_CLOSE']]
from sklearn.model_selection import train_test_split
X_{train}, X_{test}, Y_{train}, Y_{test} = train_test_split(X, Y, test_size=0.2,
random_state=101)
from sklearn.linear_model import LinearRegression
#using Stat model
from statsmodels.api import OLS
model_lin =
OLS.from formula("Nifty50 Open~SPOTRATE CLOSE+SR OUTPUT+SP500 CLOS
E+SP500_OUTPUT+NASDAQ100_CLOSE+NASDAQ_OUTPUT+CRUDE_OIL_CLO
SE+CRUDE OIL OUTPUT+LSE CLOSE+LSE OUTPUT",data=df)
result_lin = model_lin.fit()
result_lin.summary()
lm = LinearRegression(copy_X=True, fit_intercept=False)
lm.fit(X_train,Y_train)
print(lm.intercept_)
print(lm.coef_)
print(lm.score(X_train,Y_train))
```

```
import pandas as pd
x=pd.read_csv("D:\\New data\\DATA1.csv")
      df= pd.DataFrame(x)
      Y = df['Nifty50_Open']
     13
              'LSE_CLOSE']]
      from sklearn.model_selection import train_test_split
      X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size=0.2, random_state=101)
from sklearn.linear_model import LinearRegression
      #using Stat model
      from statsmodels.api import OLS
      model_lin = OLS.from_formula("Nifty50_Open~SPOTRATE_CLOSE+SR_OUTPUT+SP500_CLOSE+SP500_OUTPU
      result lin = model lin.fit()
      result_lin.summary()
      lm = LinearRegression(copy_X=True, fit_intercept=False)
      lm.fit(X_train,Y_train)
print(lm.intercept_)
      print(lm.coef_)
      print(lm.score(X_train,Y_train))
```

OUTPUT:

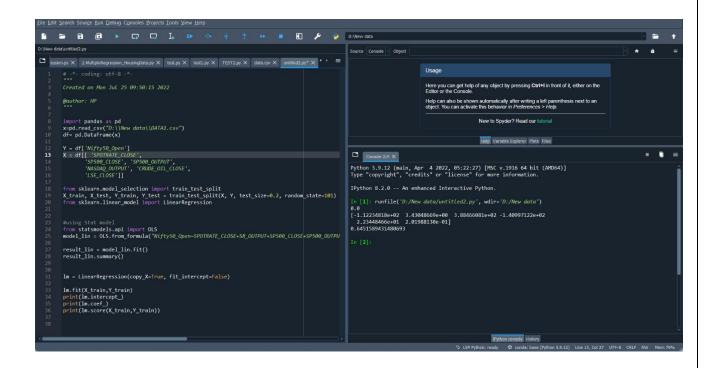
Intercept = 0.0

Coefficient = [-1.12234818e+02 3.43048669e+00 3.88466081e+02 -1.40997122e+02 2.23448466e+01 2.01988130e-01]

Score = 0.6451589431480693

```
In [1]: runfile('D:/New data/untitled2.py', wdir='D:/New data')
0.0
[-1.12234818e+02  3.43048669e+00  3.88466081e+02 -1.40997122e+02
  2.23448466e+01  2.01988130e-01]
0.6451589431480693
```

Execution:



TOOLS USED FOR PROGRAMMING: SPYDER:

Spyder is a free and open-source scientific environment written in Python, for Python, and designed by and for scientists, engineers and data analysts. It features a unique combination of the advanced editing, analysis, debugging, and profiling functionality of a comprehensive development tool with the data exploration, interactive execution, deep inspection, and beautiful visualization capabilities of a scientific package.



CONCLUSION:

In this proposed report, we have discussed Regression Based methodology used for prediction of the stock price. This report mentions the process of collecting and analyzing the historical data in order to achieve our goal. A correct prediction of stocks can lead to huge profits. While data cleaning we need to be specific on all terms like 'Date', 'Close Price', 'Open Price', 'Output' etc. as the predicted value can be differed is any mistakes or changes occur in those elements and expected prediction could not be found.

By implementing the methodologies discussed above we can conclude that these are reasonably effective as they are efficient in recognizing the patterns that exist in the domain of stock market. This demonstrates that there is a fundamental structure that is present in all stock markets. Therefore, use the above methodologies to diversify your risk, and build a portfolio of multiple stocks.

References:

- M. Billah, S. Waheed and A. Hanifa, "Stock market prediction using an improved training algorithm of neural network," 2016 2nd International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE), Rajshahi, 2016, pp. 1-4. [5] H. L. Siew and M. J. Nordin, "Regression"
- 2. https://www.researchgate.net/publication/331345883 <a href="https://www.researchgate.net/publication/assatze.n
- 3. (PDF) Stock Market Prediction Using Machine Learning. Available from: https://www.researchgate.net/publication/331345883_Stock_Market_Predict ion_Using_Machine_Learning [accessed Oct 24 2021].
- 4. https://www.google.com/search?q=regularization+in+deep+learning&rlz= 1C1GCEA_enCA831CA831&sxsrf=AOaemvLCpUiu3WAQoAjE2T-wsUVO CPCGUw%3A1635134682823&ei=2ix2YaLaMbOD4-EPvuu0gAU&oq=regularization+in+deep+learning&gs_lcp=
- 5. Wang J.Z., Wang J.J., Zhang Z.G., Guo S.P. Forecasting stock indices with back propagation neural network Expert Systems with Applications, 38 (1) (201), pp. 14346-14355
- 6. Guresen E., Kayakutlu G., Daim T.U. Using artificial neural network models in stock market index prediction Expert Systems with Applications, 38 (8) (201), pp. 10389-10397
 - S. Hochreiter and J. Schmidhuber, "Long Short-term Memory", Neural Computation, vol. 9, no. 8,