SAVITRIBAI PHULE PUNE UNIVERSITY

A REPORT ON

"AI To Predict Price Movements in the Stock Market"

SUBMITTED TOWARDS THE
PARTIAL FULFILLMENT OF THE REQUIREMENTS OF

PROJECT BASED LEARNING(TE COMPUTER)

BY

WANI SHUBHAM GOPAL   Exam No:

UNDER THE GUIDANCE OF
PROF. Mrs.V.A.Khairnar

DEPARTMENT OF COMPUTER ENGINEERING
LATE G.N. SAPKAL COLLEGE OF ENGINNERING
ANJINERI, NASHIK.
YEAR 2022-23

KCT's
LATE G.N. SAPKAL COLLEGE OF ENGINNERING

DEPARTMENT OF COMPUTER ENGINEERING

CERTIFICATE

This is to certify that the Project Entitled

"AI To Predict Price Movements in the Stock Market"

Submitted by

WANI SHUBHAM GOPAL   Exam No:

is a bonafide work carried out by Students under the supervision of **Prof. Mrs.V.A.Khairnar** and it is submittedtowards the partial fulfilment of the requirement of Project Based Learning (SE.Computer)course of SavitribaiPhule Pune University, Pune in the academic year 2021-22.

Prof. V.A.Khairnar   **Dr. N.R Wankhade**
Internal Guide   **H.O.D**
Computer Department   **Computer Department**

DECLARATION

I hereby declare that the Seminar entitled, **"AI To Predict Price Movements in the Stock Market"** was carried out and written by me under the guidance of Prof. V.A.Khairnar, Department of Computer Engineering, L.G.N.Sapkal College of Engineering,Anjeneri , Nashik. This work has not been previously formed the basis for the award of any degree or diploma or certificate nor has been submitted elsewhere for the award of any degree or diploma.

Place: Nashik   Shubham gopal wani
Date;

Acknowledgment

It gives us great pleasure in presenting the seminar report on **'AI To Predict Price Movements in the Stock Market***'.*

I would like to take this opportunity to thank my guide **Prof, V.A.Khairnar** *for giving me all the help and guidance I needed. I am really grate-ful to them for their kind support. Their valuable suggestions were veryhelpful.*

I am also grateful to **Dr.N.R Wankhade***, Head of Computer EngineeringDepartment, L.G.N.Sapkal College of Engineering,Anjeneri Nashik for his indispensable support, suggestions.*

In the end our special thanks to **Prof, V.A.Khairnar** *for providing various resources such as laboratory with all needed software platforms, continuous Internet connection, for Our Seminar.*

Shubham Gopal Wani
(T. E. Computer Engg.)

Abstract

A study on how to integrate recurrent neural networks in creating an algorithm that can predict stock market price changes. Trading in the stock market can be overwhelming due to its volatility in price changes. As a result, traders become indecisive and can miss several gaining opportunities offered by the market. The dataset for this research includes existing historical price data of the Philippine Stock Exchange index, which comprises the weighted mean of the top 30 publicly traded companies in the Philippines. The dataset is utilized for creating the predictor model using recurrent neural network algorithms. The researchers constructed a model that can potentially help forecast future price movements of the stock index by dividing the datasets for testing and training, and by using regressions and long short-term memory network (LSTM). This research is beneficial not only to the academic community but will also bring great value to the traders and investors of different markets.

There are two types of stocks. You may know of intraday trading by the commonly used term "day trading." Intraday traders hold securities positions from at least one day to the next and often for several days to weeks or months. LSTMs are very powerful in sequence prediction problems because they're able to store past information. This is important in our case because the previous price of a stock is crucial in predicting its future price. While predicting the actual price of a stock is an uphill climb, we can build a model that will predict whether the price will go up or down.

Keywords: LSTM, CNN, ML, DL, Trade Open, Trade Close, Trade Low, Trade High

Contents

List of Figure

Chapter 1

Introduction

1.1. What is stock market prediction?

A stock market prediction is an attempt to forecast the future value of an individual stock, a particular sector or the market, or the market as a whole. These forecasts generally use fundamental analysis of a company or economy, or technical analysis of charts, or a combination of the two.

1.2. What is stock marketing.

Stock market is one of the major fields that investors are dedicated to, thus stock market price trend prediction is always a hot topic for researchers from both financial and technical domains. In this research, our objective is to build a state-of-art prediction model for price trend prediction, which focuses on short-term price trend prediction.

As concluded by Fama in , financial time series prediction is known to be a notoriously difficult task due to the generally accepted, semi-strong form of market efficiency and the high level of noise. Back in 2003, Wang et al. in already applied artificial neural networks on stock market price prediction and focused on volume, as a specific feature of stock market. One of the key findings by them was that the volume was not found to be effective in improving the forecasting performance on the datasets they used, which was S&P 500 and DJI. Ince and Trafalis in targeted short-term forecasting and applied support vector machine (SVM) model on the stock price prediction. Their main contribution is performing a comparison between multi-layer perceptron (MLP) and SVM then found that most of the scenarios SVM outperformed MLP, while the result was also affected by different trading strategies. In the meantime, researchers from financial domains were applying conventional statistical methods and signal processing techniques on analyzing stock market data.

The optimization techniques, such as principal component analysis (PCA) were also applied in short-term stock price prediction. During the years, researchers are not only focused on stock price-related analysis but also tried to analyze stock market transactions such as volume burst risks, which expands the stock market analysis research domain broader and indicates this research domain still has high potential. As the artificial intelligence techniques evolved in recent years, many proposed solutions attempted to combine machine learning and deep learning techniques based on previous approaches, and then proposed new metrics that serve as training features such as Liu and Wang. This type of previous works belongs to the feature engineering domain and can be considered as the inspiration of feature extension ideas in our research. Liu et al. in proposed a convolutional neural network (CNN) as well as a long short-term memory (LSTM) neural network based model to analyze different quantitative strategies in stock markets. The CNN serves for the stock selection strategy, automatically extracts features based on quantitative data, then follows an LSTM to preserve the time-series features for improving profits.

The latest work also proposes a similar hybrid neural network architecture, integrating a convolutional neural network with a bidirectional long short-term memory to predict the stock market index [4]. While the researchers frequently proposed different neural network solution

architectures, it brought further discussions about the topic if the high cost of training such models is worth the result or not.

There are three key contributions of our work

(1) A new dataset extracted and cleansed.

(2) A comprehensive feature engineering.

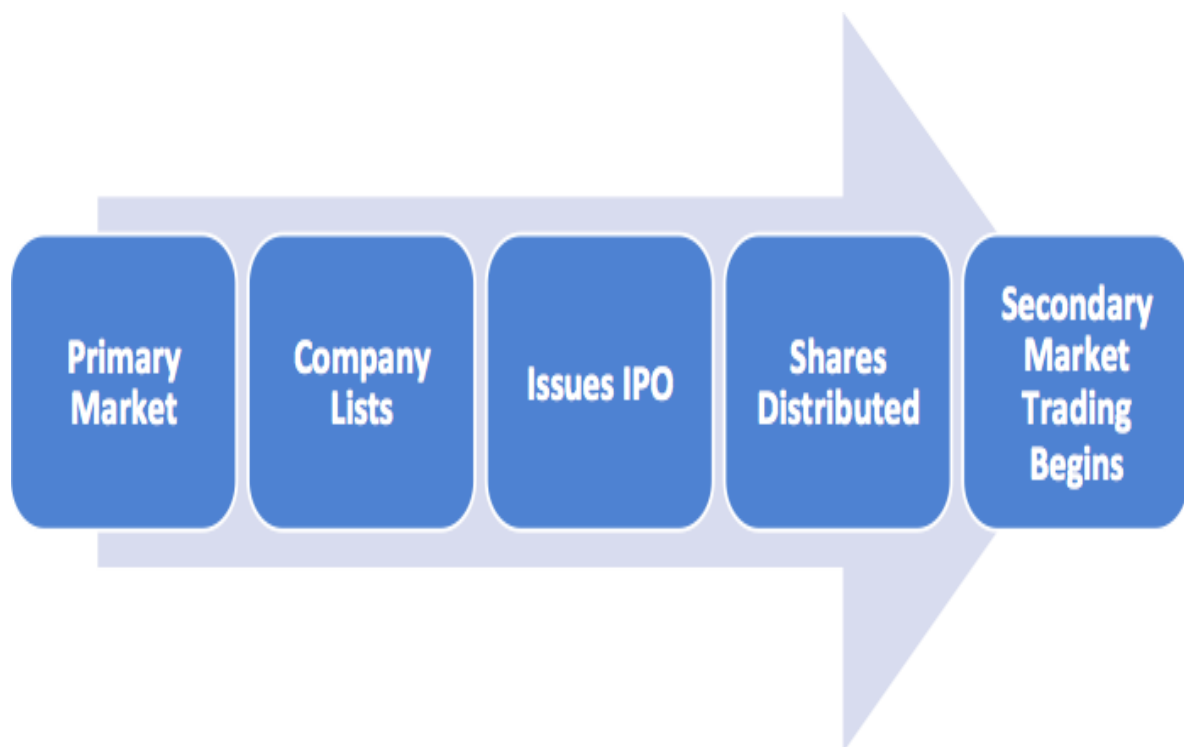(3) A customized long short-term memory (LSTM) based deep learning model.

Fig: - working of stock market

Chapter 2

Literature Survey

2.1 Paper Surveys

"What other people think" has always been an important piece of information for most of us during the decision-making process. The Internet and the Web have now (among other things) made it possible to find out about the opinions and experiences of those in the vast pool of people that are neither our personal acquaintances nor well-known professional critics — that is, people we have never heard of. And conversely, more and more people are making their opinions available to strangers via the Internet. The interest that individual users show in online opinions about products and services, and the potential influence such opinions wield, is something that is driving force for this area of interest. And there are many challenges involved in this process which needs to be walked all over in order to attain proper outcomes out of them. In this survey we analysed basic methodology that usually happens in this process and measures that are to be taken to overcome the challenges being faced.

1. Alvin Ken Steven, Ellysa D. Pua, Elmer P. Dadios, Robert Kerwin C. Billones. AI To Predict Price Movements in the Stock Market [1].
   This paper presents a study on how to integrate recurrent neural networks in creating an algorithm that can predict stock market price changes. Trading in the stock market can be overwhelming due to its volatility in price changes. As a result, traders become indecisive and can miss several gaining opportunities offered by the market. The dataset for this research includes existing historical price data of the Philippine Stock Exchange index, which comprises the weighted mean of the top 30 publicly traded companies in the Philippines. The dataset is utilized for creating the predictor model using recurrent neural network algorithms. The researchers constructed a model that can potentially help forecast future price movements of the stock index by dividing the datasets for testing and training, and by using regressions and long short-term memory network (LSTM). This research is beneficial not only to the academic community but will also bring great value to the traders and investors of different markets.

2. Ashkan Safari, Amir Aminzadeh Ghavifekr. International Stock Index Prediction Using Artificial Neural Network (ANN) and Python Programming [2].
   The stock market is one of the best channels for financial development that requires a high accuracy prediction of the trades. This subject needs some technical skills and experience to achieve the best result. This paper represents a tuned Python console program based on the Neural Network (NN), and the Artificial Intelligence (AI) to predict future price in a qualified and quantized way with high accuracy and close to real. New ideas implemented in this paper are combining AI and NN model in the Python console system with a security shell that works with voice and PIN to authenticate the user. It has Cross-Platform capability and supports cryptocurrencies price and their predictions. This program enables the user to have a duplication of the final data in his/her given email. The proposed approach presents the influence of AI and Machine learning in nearly future predictions. This system can be used in the all kinds of subjects that include past time databases.

3. Ambarish Gadgil Aditya Desity, Prasanna Asole, Harsh Dandge, Spurti Shinde. Stock Price Prognosticator using Machine Learning Techniques [3].

Stock market price prediction is one of the favourite research topics under consideration for professionals from various fields like mathematics, statistics, history, finance, computer science engineering etc., as it requires a set of skills to predict variation of price of shares in a very volatile and challenging share market scenario. Share market trading is mostly dependent on sentiments of investors and other factors like economic policies, political changes, natural disasters etc., Many theories were forwarded, mathematical and statistical applications in conjunction with probability, to simplify the complex process. After the advent of computers, it got further simplified but still challenging due to various external influential factors ruling the volatility of the market prices. Thus, AI and ML algorithms were being developed, but for only for next day using Linear Regression procedures.

Chapter 3

Working

3.1 How Does Stock Market Prediction Works?

In this section, we present the proposed methods and the design of the proposed solution. Moreover, we also introduce the architecture design as well as algorithmic and implementation details.

Stock price analysis has been a critical area of research and is one of the top applications of machine learning. This tutorial will teach you how to perform stock price prediction using machine learning and deep learning techniques. Here, you will use an LSTM network to train your model with Google stocks data.
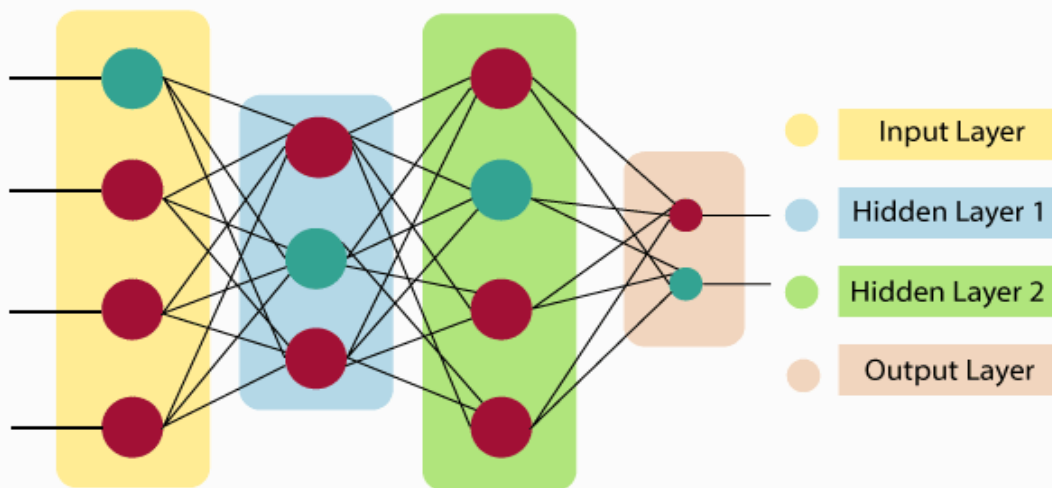

Fig: -ANN Architecture

3.2 Long Short Term Memory (LSTM): -

LSTM is a form of recurrent neural network that addresses sequential problems. The network model is created by building several layers for the LSTM structure. The variation can be used to solve multiple complex problems, including the prediction of patterns based from the input data of the sequence over time to achieve a high prediction accuracy.

Backpropagation in neural networks is an algorithm that computes the gradient of a model's error function, with consideration on its weights. The order for the gradient's computation starts from the last layer and ends with the first most layer, representing a backward process. Although the algorithm is widely used in neural networks, it has the tendency to produce lags during the process, most likely due to deficient and decaying error backflow. For this reason, one of the main purposes of LSTM is to avoid the lag time, giving the system a faster processing time. LSTM RNN is known to be more effective over other forms of neural network combination algorithms. A sample of the LSTM RNN architecture is shown as follows.
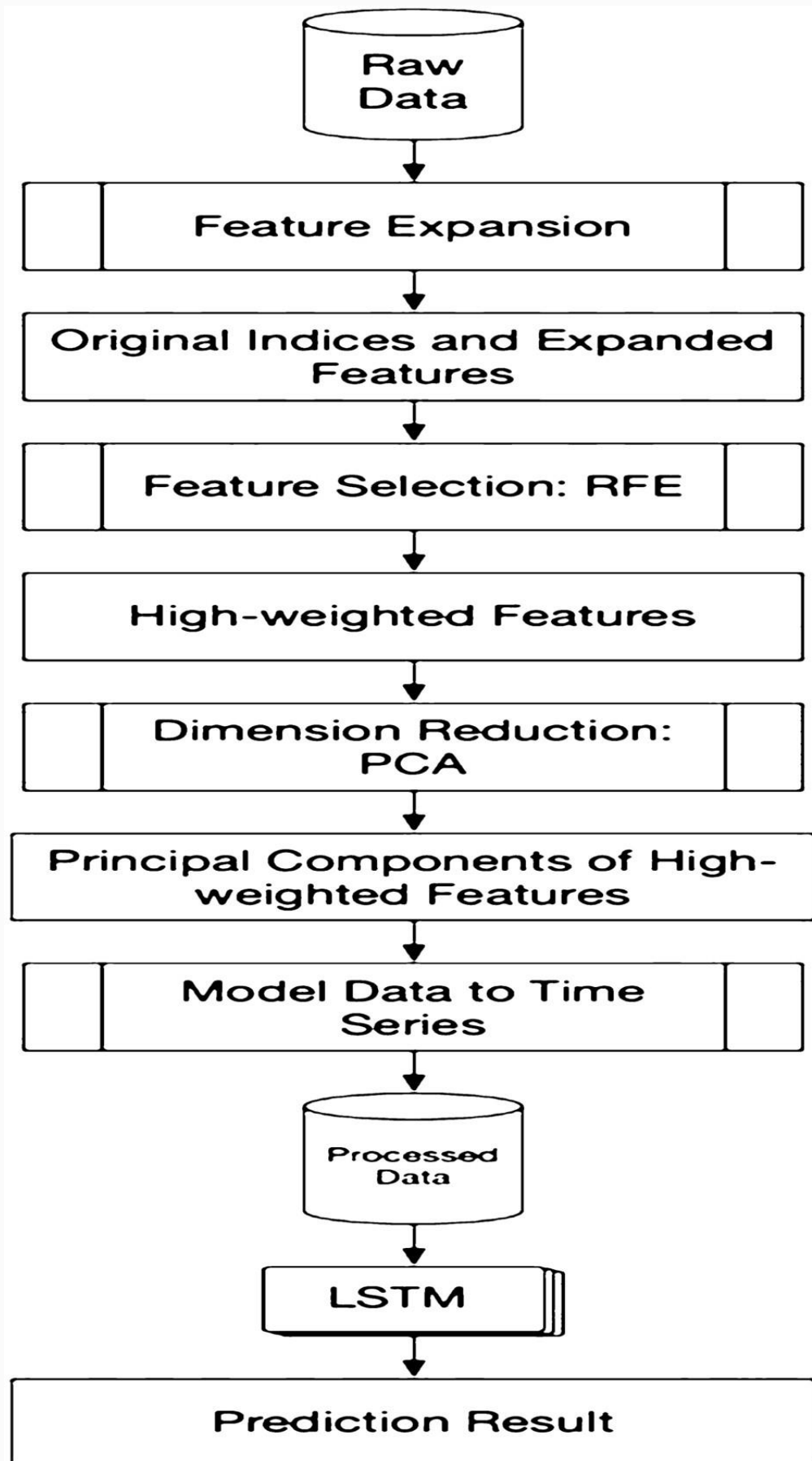
Fig: - LSTM Architecture.

3.3 RESULTS AND DISCUSSION

The implementation of the RNN-LSTM approach is seen to be a success (Figure 6). The model was trained using 100 epochs with 24 data per batch. By visual analysis, the predicted model contains almost the same trends as the real model. Upon further analysis, two problems can be seen here. The first one is the price fall at around the 50th index wherein the predicted model projected around 5000 while the real price is much lower at 4000. The second one is at the end of the graphs where the real stock prices are projecting a rise while the predicted stock prices are projecting the opposite. A lot can go wrong with this as being informed of a false positive may cause a big loss for investors.

From the figure, it can also be seen that a lot of the points contain values that were over-predicted, which is supported by the negative percent error. Having MPEs less than zero implies over-prediction of values as it can only turn negative when a significant number of predicted values are more than their true value counterparts. For this case, the values predicted are approximately 3.58% above the true values. This is not much of a problem as it can be seen by visual inspection that this overshoot is consistent throughout many different points. This means that further tweaking and optimization on the code can adjust the predicted points downwards without compromising the validity of the study.
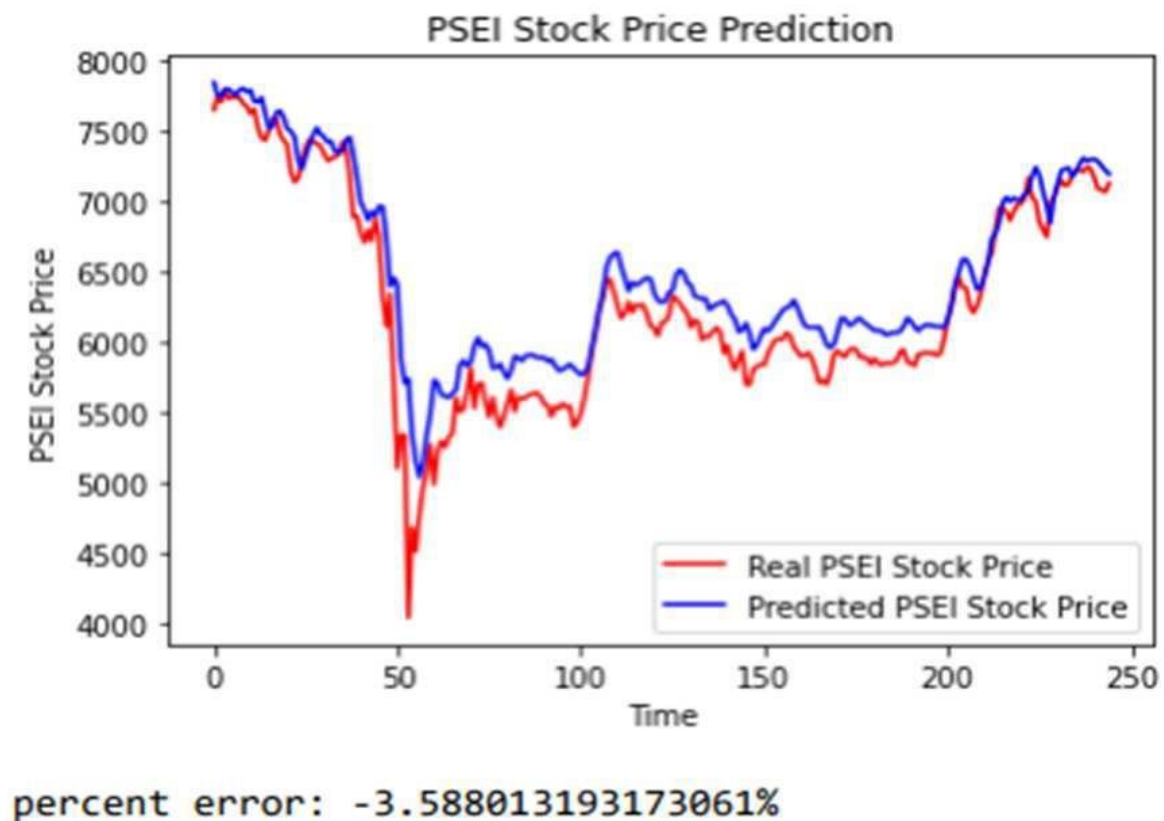

percent error: -3.588013193173061%

Fig: - Stock prediction

Chapter 4

Applications

The stock market prediction has attracted much attention from academia as well as business. Due to the non-linear, volatile and complex nature of the market, it is quite difficult to predict. As the stock markets grow bigger, more investors pay attention to develop a systematic approach to predict the stock market. Since the stock market is very sensitive to the external information, the performance of previous prediction systems is limited by merely considering the traditional stock data. New forms of collective intelligence have emerged with the rise of the Internet (e.g. Google Trends, Wikipedia, etc.). The changes on these platforms will significantly affect the stock market. In addition, both the financial news sentiment and volumes are believed to have an impact on the stock price. In this study, disparate data sources are used to generate a prediction model along with a comparison of different machine learning methods. Besides historical data directly from the stock market, numbers of external data sources are also considered as inputs to the model. The goal of this study is to develop and evaluate a decision-making system that could be used to predict stocks' short term movement, trend, and price. We took advantage of the open source API and public economic database which allow us to explore the hidden information among these platforms. The prediction models are compared and evaluated using machine learning techniques, such as neural network, support vector regression and boosted tree. Numbers of case studies are performed to evaluate the performance of the prediction system.

Stock Recognition, Handwritten stock Recognition, Stock Classification, Healthcare, Stock Verification and Analysis, Stock Forecasting and Predictions. Are the applications of Stock marketing prediction.

Chapter 5

Advantages and Disadvantages

- **Advantages**

  1. Removes the Investment Bias.
  2. Minimizes Your Losses.
  3. Assures Consistency.
  4. Gives a better idea about Entry and Exit points.
  5. Allows the smart way of making money.

- **Disadvantages**

  1. The model requires lots of computational power.
  2. It is hard to explain.
  3. It requires lots of data.
  4. Data preparation for neural network models need careful attention.
  5. Incorrect prediction leading to loss in market.

- **Future Scope**

Stock market prediction means determining the future scope of market. A system is essential to be built which will work with maximum accuracy and it should consider all important factors that could influence the result. Various researches have already been done to predict stock market prices.

Conclusion

We have successfully implemented machine learning algorithms on the dataset for predicting the stock market price. We applied data preprocessing and feature selection on the dataset. We applied four algorithms: KNN, SVM, Random Forest, Logistic Regression on the dataset. We analyzed the difference of the algorithms by calculating the performance metrics (accuracy, Recall, precision, f-score). We also found the advantages and disadvantages of the algorithms. We conclude that Random Forest is the best algorithm out of the four with an accuracy rate of 80.7%. Future scope of this paper would be adding more parameters that effect the stock market prediction. Adding more number of parameters will ensure better estimation.

References

1. Alvin Ken Steven C. Chua; Ellysa D. Pua; Elmer P. Dadios; Robert Kerwin C. Billones. AI To Predict Price Movements in the Stock Market. Available on https://ieeexplore.ieee.org/document/9731955

2. Stock market prediction through artificial intelligence, machine learning and neural networks (AIMV) 2021 International Conference on Artificial Intelligence and Machine Vision.|978-1-6654-4211-4/21/$31.00 ©2021 IEEE | DOI: 10.1109/AIMV53313.2021.9670919

3. International Stock Index Prediction Using Artificial Neural Network (ANN) and Python Programming, 2021 7th International Conference on Control, Instrumentation and Automation (ICCIA).|978-1-6654-0350-4/20/$31.00©2021IEEE | DOI: 10.1109/ICCIA52082.2021.9403580.

4. PSE Academy, "Types of STOCKS", PSE Academy, 17-Sep-2011. Accessed: 19-Jan-2021. [Online]. Available: https://www.pseacademy.com.ph/LM/investors~details/id-1316266249549/Types_of_STOCKS.html.

5. The Heritage Foundation, "The Importance of Trade", The Heritage Foundation, n.d.. Accessed on: January 19, 2021. [Online]. Available: https://www.heritage.org/trade/heritage-explains/the-importance-trade.