Chapter 7
Stock Market Analysis and Prediction Using ARIMA, Facebook Prophet, and Stacked Long Short-Term Memory Recurrent Neural Network

Parvathi R
Vellore Institute of Technology, Chennai, India

Xiaohui Yuan
University of North Texas, USA

ABSTRACT

Stock analysis involves comparing a company's current financial statement to its financial statements in previous years to give an investor a sense of whether the company is growing, stable, or deteriorating. Stock market analysis helps in getting insights into a company's stock and to make better decisions in buying or selling shares in the stock market. This chapter proposes a method to analyze and predict stock market prices based on historical data of 4 MNCs namely, Amazon, Apple, Google, and Microsoft. The prediction is implemented using three models; namely, ARIMA model, Facebook's Prophet model, and lastly a self-constructed, stacked LSTM model. The results of the three models are compared and analyzed. Mean absolute error is used to analyze the performance of the models on real-time test data. The minimum loss achieved by Facebook Prophet Model is 2.445, by ARIMA Model is 10.782, and the Stacked LSTM Model achieved a minimum loss of 6.552.

DOI: 10.4018/978-1-6684-9804-0.ch007

INTRODUCTION

Forecasting is the process of predicting a future event with the help of historical data. It is widely used in many applications such as business, industries, finance, etc. Forecasting problems involves analysis of time, and are classified into three types,

- Short term forecasting, where the prediction is done for the duration of few seconds, minutes, days, weeks or months.
- Medium term forecasting, where the duration of prediction ranges from one to two years.
- Long term forecasting, where the duration of prediction is beyond two years.

A time series data consists of sequential observations or information for a particular variable. For our research paper, the variable taken into consideration is the stock price which is either univariate or multivariate. Univariate data comprises of information about a single stock, whereas multivariate data includes stock prices of multiple companies at various time instances. Time series data analysis aids in identifying different trends, patterns, and cycles or periods that is present in the data. Knowledge about the bullish or bearish mode in case of stock market data, helps in wise investment of money. Additionally, analyzing the patterns helps in determining the good performing companies for a certain time period. As a result, time series analysis and forecasting are extremely important areas of research. The existing approaches for predicting stock prices are classified as technical analysis, fundamental analysis, and time series forecasting (Devadoss & Ligori, 2013).

Fundamental analysis is defined as a type of investment analysis specifically used for long-term forecasting, helps in determining the company's share value by assessing the profits, sales, earnings, and other economic factors. In technical analysis, the historical stock data is used for prediction of future prices. One of the popular technical analysis methods is the moving average which is defined as the unweighted mean of the last n data points. and is best suitable for short term predictions. It can be considered as the unweighted mean of past n data points. This method is suitable for short term predictions. The third method is the analysis of time series data and involves two classes of algorithms, namely linear model and non-linear models.

The different linear models are AR, ARMA, ARIMA and using predefined equations to fit a mathematical model to a univariate time series. The main disadvantage is that the model cannot identify inter dependencies as they consider only univariate time series data, and also does not account for the latent dynamics existing in the data. Hence resulting in a model, that is not capable of identifying dynamics or patterns in the whole data. Non-linear models involve methods like