Predicting Tesla's Stock Price
ECS171 Fall 2021

Alvin Agana, Caitlin Brown, Tsung Chieh Chen,
Aranza Cortes, Mckenzie Hochberg, Grant Koziol, Ryan Place,
Clare Tran, Dingxian Wang, Stacy Zhu, Leon Zhuang

December 5, 2021

Contents

1 Introduction

The development of machine learning methods has allowed for more accurate and efficient prediction of the stock price and forecasting a company's market value. Tesla, Inc, an American company founded in 2003, primarily focuses on the expansion of all-electric vehicles, clean energy generation and storage products [1]. Tesla went public in 2010 with an opening price of $17 a share on the Nasdaq stock market. Tesla's IPO alone raised over $226 million [6]. Through the application of machine learning methods, the goal is to construct a predictive model on Tesla's stock market price.

Since the stakes are large due to the financial crisis and scoring profits, it is significant that there must be a secure prediction of the values of the stocks [3]. There is a turnover of trading of approximately thousands of billions dollars which effectively provokes individuals in making a profit through this. A share is more valuable the more it is conducted while on the other hand, it loses its value if it is put into transaction in a low volume as it becomes less important for some traders [3]. The market can therefore result in profits or losses when one expects it to behave a certain way. That is where the problem of generating a reliable model in predicting stock prices comes in. When given a stock market history, the goal is to find out when the most appropriate time for buying or selling the share would be in order to generate profit. If it is possible to precisely predict the future trends of a stock, this would greatly reduce the financial risk.

Stock price prediction can be applied to the historical data of TESLA INC (TSLA) stock as an example for the following designed prediction algorithm. TSLA is a challenging stock to predict as the previous trends for the stock is not one of undeviating, rather it is constantly fluctuating due to external factors such as supply and demand. An example of this includes missed predictions in the second quarter of 2017. Due to this error, TSLA had lost over 5% of the company stock price value which resulted in a loss worth $12 billion [6]. In addition, Elon Musk, the CEO of Tesla, publicly posted on a popular social media application, Twitter, announcing that he had planned to make the company private when the stock price reaches $420. Consequently, this resulted in an elevation of 10% in TSLA's stock price as investors were motivated to buy more shares before the supposed privatization announced by Musk [6].

As stock market predictions have the ability to significantly impact the global economy, it is important to understand how to analyze the volatility of the stock market. However, it turns out to be very difficult to predict the stock market as the rapidly changing state of it is too large to be captured in a model [1]. Even though there are these known difficulties involved with forecasting the stock market, this does not mean that there is no desire to develop a model that is reliable in doing so. In fact, it is actually the opposite in which there is a continuous desire to build and employ new approaches to tackle this problem in Artificial Intelligence, specifically Machine Learning [7].

The objective of this project is to successfully apply machine learning methods to design and execute predictive models to prognosticate future stock prices of TSLA. This can be done by analyzing the daily stock prices of Tesla while also considering the attributes provided. Ideally, the results can be applied appropriately to maximize significant gains and assist financial institutions with market research.

2   Literature Review

Since the financial data of the real world is complicated, there were many limitations that early studies using statistical methods faced [2]. This was due to numerous statistical assumptions like normality and linearity. Therefore, some machine learning techniques that have the ability to capture nonlinearity and complexity have been applied to predicting the stock market. These include artificial neural networks (ANN) and support vector machines (SVM). There have also been attempts on utilizing deep learning techniques in which deep belief network (DBN), convolution neural network (CNN), and recurrent neural network (RNN) are commonly used methodologies [2].

Developed by RNN, long short-term memory (LSTM) neural networks help forecast financial time series and are able to avoid long-term dependence issues as a result of its special storage unit structure [4]. Although LSTM is commonly used in speech and image recognition due to its attention mechanism, it is rarely used in finance even though it proves to be appropriate in this area. LSTM is able to remember important events that occurred in the past, from many time steps ago, due to its memory [10]. As a result of this memory feature, RNN proves to be advantageous compared to the traditional ANN because of its ability to evoke temporal patterns in data.

Deep learning models have promising results as they can significantly improve the accuracy. Although it was explored that RNN models like LSTM could perform better than ANN, there have been many approaches using ANN to predict stock prices. One application of this model used Nifty stock index dataset and was able to achieve an accuracy of 96% with the average accuracy over all the cases as 88% [9]. In this ANN approach, there are customizable parameters and various activation functions are executed that have the option for cross validation sets as well.

There is also another proposed system that uses both Regression and Classification in order to predict stock prices. The system will predict the closing price of stock of a company in regression while with classification, the system will predict if the closing price of the stock will increase or decrease the following day [5]. Some regression models that are applied include: simple linear regression, polynomial regression, and support vector regression. The classification algorithms to determine whether the stock will go high or low the next day include: SVM, logistic regression, and Naive Bayes. The accuracies obtained from the models depend largely on the input dataset and does not indicate which algorithm or model provides the best results when used.

To conclude this section, there have been previous findings that go beyond the typical models for regression that are usually implemented in forecasting stock prices. In order to approach the problem of predicting stock prices, we will not only turn towards regression analysis to help us but also to methods like LSTM and SVM. This is due to their promising results in modeling complex systems and capturing nonlinearity. The results from previous research demonstrate how there are other machine learning techniques that one may not initially think fits the problem statement but actually could produce favorable results. Since the stock market is known for being volatile, nonlinear, and dynamic, it is important to look into analytic techniques being explored in order to detect these trends.

3   Dataset Description

The dataset provides information about Tesla stock data from June 29, 2010 to October 8, 2021. It consists of 2841 observations in which the data is available on a daily basis. The currency of the data provided is all in terms of USD.

Within the TSLA.csv data, there are seven attributes: Date, Open, High, Low, Close, Adj Close, and Volume. Date is formatted as yyyy-mm-dd and is provided on a daily level, taking into account that there are no regular trading hours for stocks on the weekends. The Open attribute is the Tesla stock price at the start of trading day. Similarly, the Close attribute is the stock price at the end of the trading day. The Adj Close attribute stands for "adjusted close" which is the closing price after adjustments to reflect the stock's value after taking corporate actions into consideration. Adj Close factors in corporate actions including stock splits, dividends, and right offerings. The High attribute tells us what the maximum price is on that trading day. Meanwhile, the Low attribute is the minimum price on that trading day. The Volume is the amount of Tesla stocks traded on that day and is represented by a positive integer number and the rest of the attributes are all decimal numbers in USD.

4   Methods

4.1   Linear Regression

4.2   Polynomial Regression

4.3   Support Vector Regression

4.4   Long Short-Term Memory

5   Results

5.1   Linear Regression

5.2   Polynomial Regression

5.3   Support Vector Regression

5.4   Long Short-Term Memory

6   Conclusion

References

1. Example. `https://google.com`

A   Appendix

A.1   Python Code

Here we can include code from our notebooks using either the minted or verbatim package, if we want.

A.2   GitHub

Here is the URL for our GitHub repository:

`https://github.com/Diean233/Diean233.github.io`