Title

University of New Haven

DSCI: 6002 -1 - Intro to Data Science – Fall 2023

Time Series Analysis For Stock Prediction

Team – 13

- **Ramya**, Boggala Ramesh
- **Krishnaveni,** Peesapati
- **KavyaSree,** Mokatala
- **Amer,** Sohail Shaik

Contents

Contents:

Objective

Business challenge

There are several challenges that need to be addressed to make accurate stock price predictions:

- Data quality
- Non-linearity
- Overfitting
- Dynamic nature of markets
- Lack of transparency
- Data volume
- Limited predictability

Solutions

- To address data quality issues, it's essential to ensure that the data used for stock price prediction is accurate, complete, and up to date. This can be achieved by using data from reputable sources, performing data cleaning and validation, and implementing quality control measures to identify and correct errors.
- Models can become too complex and overfit to the training data, making them less accurate when applied to new data. It's essential to balance the model's complexity with its ability to generalize to new data.

Technical Report

Executive Summary:

The executive summary includes about the studying of how stock market, prices change over time. We aimed to create a smart model to tell the difference between what really happens in the stock market and what we predict. We used a special model called ARIMA for this. The results showed that prices vary over time. The ups and downs of the stock are kind of what we expected. That wasn't a clear repeating pattern. The differences between what we predicted and what really happened mostly same but there exist very few changes. So, using the ARIMA model seems like a good way to understand and predict stock market changes accurately.

Highlights of Project:

- Time-series analysis is the study of data sets over time. It's commonly used in finance and business to understand how the price of stock changes over time and how different strategies affect performance.
- For this project we have used the models called ARIMA Model and Linear Regression.
- Linear regression is a data analysis technique that predicts the value of unknown data by using another related and known data value. It mathematically models the unknown or dependent variable and the known or independent variable as a linear equation.
- ARIMA is a statistical model utilized for analyzing and predicting time series data.

Abstract

Abstract:

- Investing in the stock market has always been a popular way to grow wealth, but predicting short-term stock prices is challenging. This paper focuses on using a forecasting model called Autoregressive Integrated Moving Average (ARIMA) to predict the daily stock prices of companies listed on the Bombay Stock Exchange (BSE) and National Stock Exchange (NSE). The ARIMA model is effective for short-term and daily stock predictions.

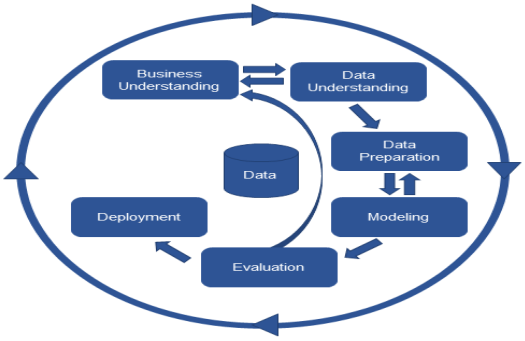
- The research extends to electric vehicle companies like Tesla and NIO, whose stock prices have experienced fluctuations. Researchers commonly use various models, including econometrics and deep learning techniques like Long Short- Term Memory (LSTM) and Recurrent Neural Networks (RNN), for stock price predictions.

- The main goal of this paper is to predict the stock prices, including closing prices, of Tesla and NIO. The ARIMA (p, d, q) model is applied, utilizing differencing techniques to make the data suitable for forecasting. Additionally, machine learning algorithms are employed, with training and testing data used in fixed proportions. The study aims to extrapolate the data into the future to predict stock prices accurately. The managerial implications include aiding companies in better planning and execution based on these predictions.

Methodology

Methodology

We have used CRISP-DM Methodology, The CRISP-DM (Cross-Industry Standard Process for Data Mining) methodology serves as a comprehensive framework for guiding data mining and machine learning projects. In its iterative process, it begins with understanding the business objectives, emphasizing the importance of aligning data analysis with overarching business goals. Below are the various steps which we have taken while building the Model.



- Collect the appropriate dataset for Stock Prediction
- Explore and understand the data.
- Clean the data and Perform engineering and preprocessing to prepare for the modeling step.
- Select the suitable predictive model and train it with the data.
- Deploy the model on a live server and integrate it into a web application to predict the pricing in real time.
- Monitor the model in production and iterate.

Methodology

Data collection:

- To understand how things change over time, we looked at different sites for information. We used sources like google dataset search, world bank open data, GitHub, climate data online and Kaggle. These places have lots of different datasets, and we picked one from there for our study.

Data analysis:

- Application of time series analysis to identify patterns and trends. Implementation of the regression models and Arima model to predict the differences.

Modeling and evaluation:

Arima model:

The ARIMA model is like a super tool used in math to predict future values in a sequence of numbers. It is especially great for things like predicting stock prices or weather patterns. This tool has three important parts:

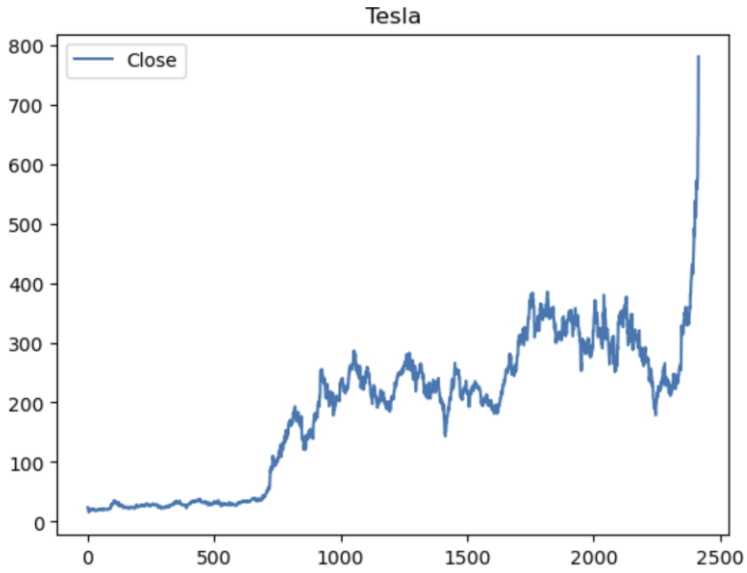
- Autoregressive part
- Integrated part.
- Moving average (ma) part:
- We write the ARIMA model as ARIMA (p, d, q)

Linear regression model:

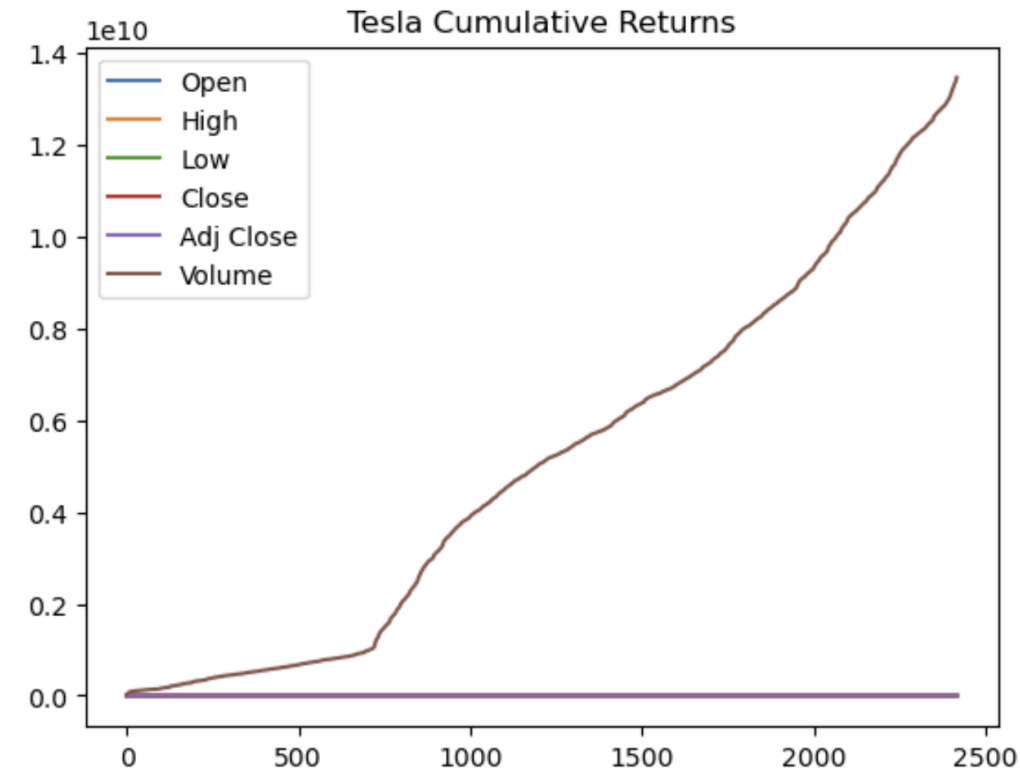
Result Section

Results Section:

Tesla Data:

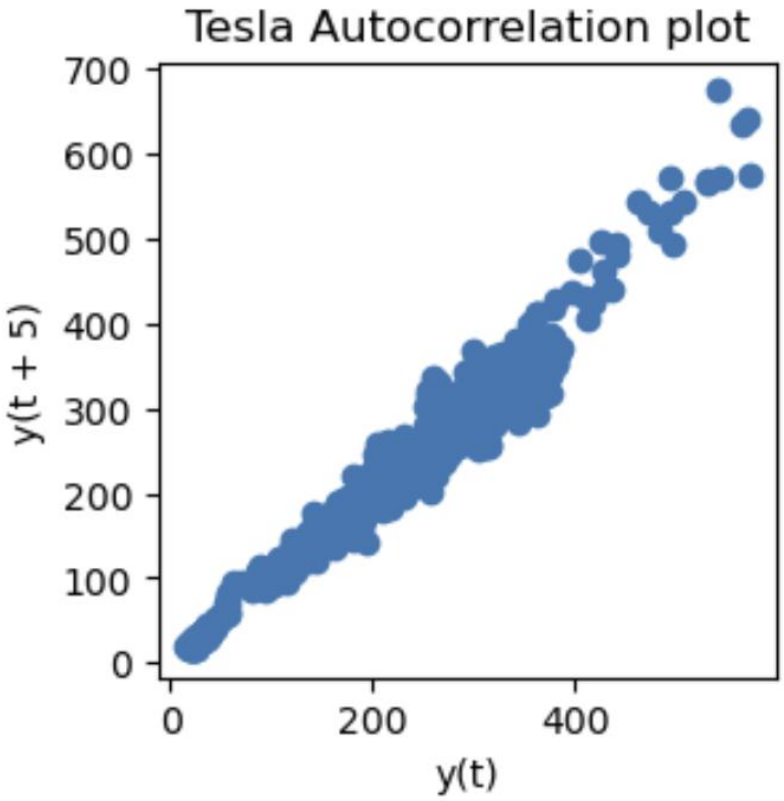


Tesla Cumulative Returns:



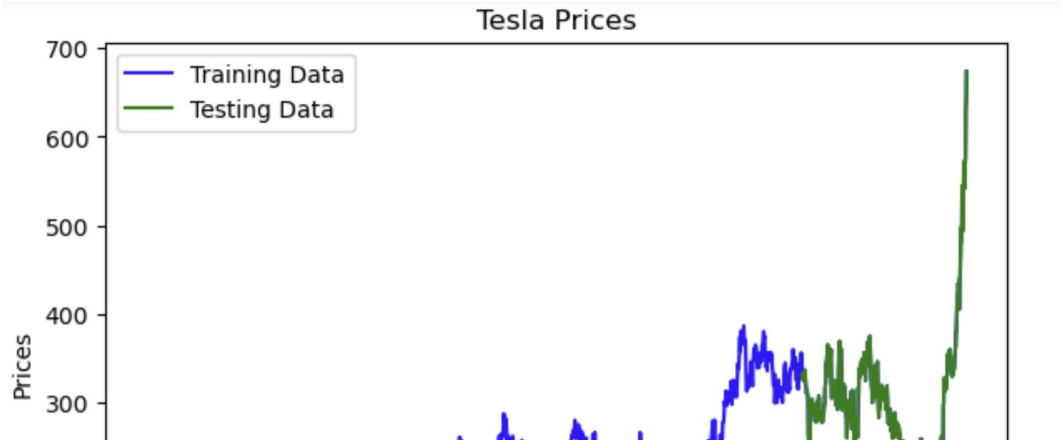
Result Section

Tesla Auto-Correlation Plot:



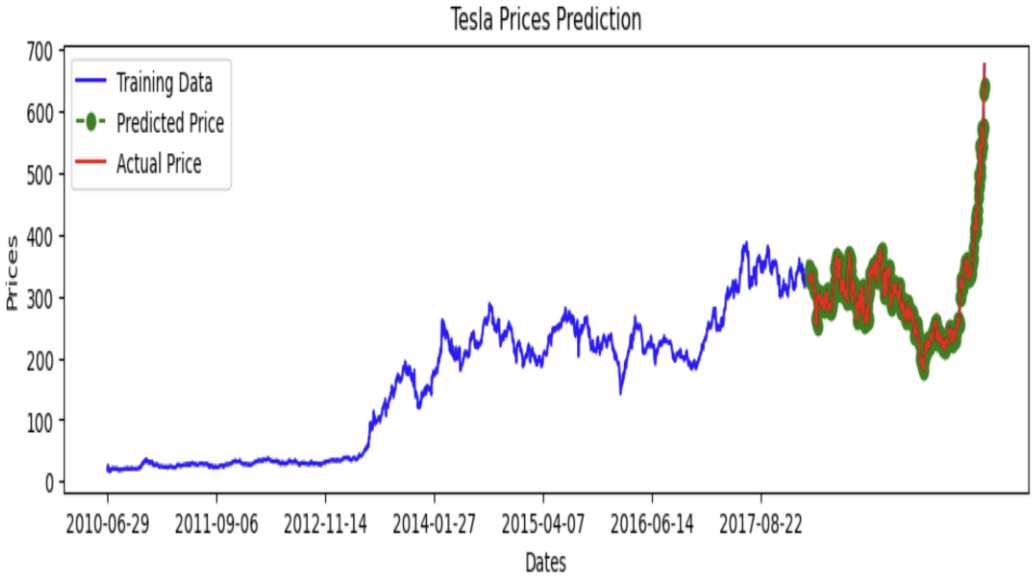
ARIMA (Autoregressive Integrated Moving Average) for Time Series Prediction:

Actual:



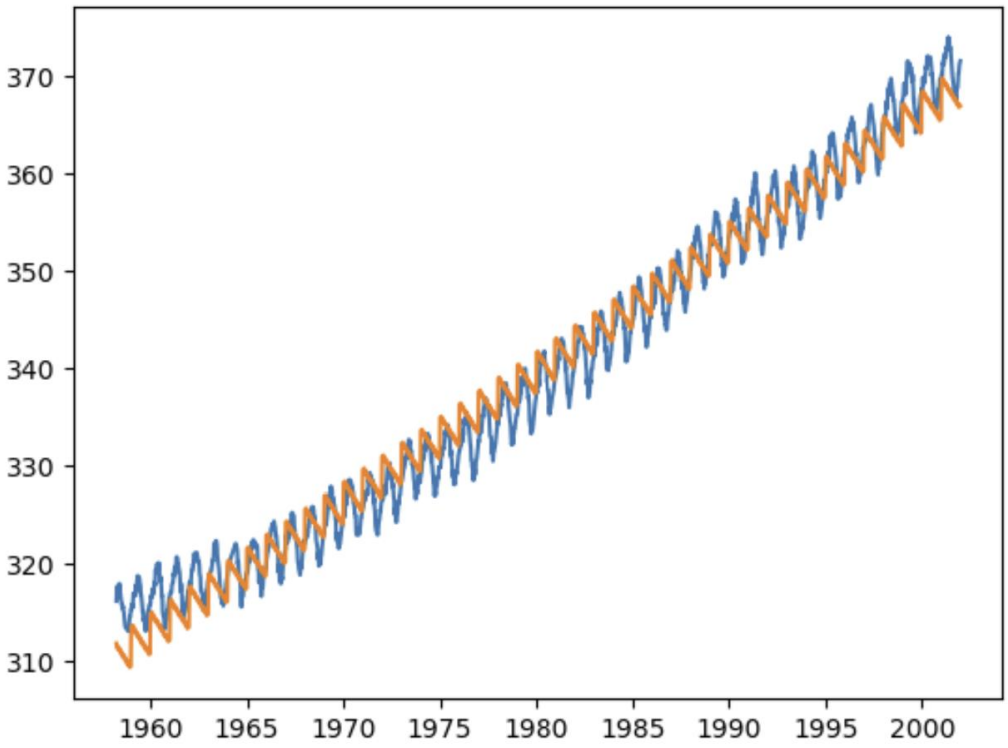
Result Section

Tesla Prices Prediction:



Tesla Linear Regression Model:

```
[<matplotlib.lines.Line2D at 0x1be2aaca750>]
```



Conclusion

Conclusion:

- In summary, time series analysis offers a useful approach for predicting the stock market.
- By utilizing statistical, machine learning, and deep learning models, along with appropriate features and indicators, businesses can enhance their forecasting capabilities. However, it's important to recognize that stock market prediction is complex and uncertain.
- Factors such as economic conditions and investor sentiment contribute to volatility, making accurate predictions challenging.
- Continuous model evaluation, monitoring, and adaptation is necessary for optimal performance. While time series analysis can improve accuracy, it's crucial to manage expectations and understand the inherent limitations of stock market prediction.

References:

GitHubLink:
https://github.com/RamyaBoggalaRamesh/Final_Project