Forecasting Tesla stock prices using ARIMA models

MSc Research Project

Masters in Data Analytics

Manchu Leela Prakash

Student ID: X23214210

School of Computing

National College of Ireland

Supervisor: EAMON NOLAN

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	Manchu Leela Prakash
Student ID:	X23214210
Programme:	Masters in Data Analytics **Year:** 2024
Module:	MSC RESEARCH PROJECT
Supervisor:	EAMON NOLAN
Submission Due Date:	12 Dec
Project Title:	Forecasting Tesla stock prices using ARIMA models
Page Count:	20 pages

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: MANCHU LEELA PRAKASH

Date: 12 Dec

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Forecasting Tesla stock prices using ARIMA models

MANCHU LEELA PRAKASH

X23214210

Abstract

The project's findings include further exploited by financial institutions, including such entities as hedge funds and portfolio managers. This group can benefit from the unveiling of reliable forecasting models provided by the project since such deliverables can be used to manage one's portfolio better. In general, due to improved asset allocation, investment funds' performance can be increased (Mei *et al.*, 2018). Moreover, those working in the area of financial analytics and being researchers, or data scientists, can make use of the study's methodological contributions.

1 Introduction

1.1 Background

Tesla, Inc. is one of the leading companies in the automotive and industrial sectors. But with its leadership in new electric vehicles and energy solutions, as well as high market valuations, Tesla's stock price can be somewhat volatile and difficult to predict (Afeef *et al.,* 2018). This issue is particularly complicated for investors and traders trying to use timely information to understand potential gains or losses.

The ARIMA model has gained a significant acknowledgement in the domain of stock market forecasting. Time series forecasting in various cases is beneficial for the given model that reflects a range of trends associated with a particular stock or asset. Therefore, ARIMA stands for AutoRegressive Integrated Moving Average and is frequently characterized as the model series with aft in the general class and specific types. Discussing the applicability of ARIMA models for time series forecasting, one should mention the benefits of integrating the model in this process.

1.2 Problem Statement

For anyone seeking to maximize the return on investment, understanding the future direction of a stock price is the key to achieving a successful trading strategy. Yet it is not that simple with the stock market. In the case of Tesla, short-term price predictions are mostly impossible for a number of reasons (Wijesinghe and Rathnayaka, 2020). The first is that the stock market is highly unpredictable and can be affected by a single comment, invention, or improvement by someone or a firm.

Commonly used forecasting models often fail to accommodate such dynamism and produce inaccurate and outdated results. ARIMA models, which are among the most popular time series forecasting tools, might be able to provide more accurate predictions on stock price behaviour. The research problem to be solved is to ensure that the ARIMA model can provide accurate short-term forecasts of a major security volatility such as Tesla, as well as to identify the limitations of the model Ensures support the necessary will be provided.

1.3 Aim and Objectives

1.3.1 Aim

To forecast Tesla's stock prices using ARIMA models, providing accurate short-term predictions to support investment and trading decisions.

1.3.2 Objectives

- To explore the application of ARIMA models in forecasting Tesla stock prices.
- To analyse the relationship between historical stock price data and predictive accuracy using ARIMA techniques.
- To identify the limitations and challenges of using ARIMA models for forecasting in volatile stock markets like Tesla.
- To propose and implement countermeasures for improving forecasting accuracy and model performance when predicting Tesla stock prices using ARIMA models.

1.4 Research Question

How effective are ARIMA models in forecasting Tesla stock prices compared to other time-series forecasting methods?

1.5 Motivation for the Project

This project's motivation arises with the growth of demand for reliable and efficient forecasting tools for prediction of stock price trend in real-time. ARIMA models are strongly recognized as the one of the most powerful and efficient time series forecasting approach that can be used for the analysis of the historical stock price data and making prediction for the future. However, the current limitations of ARIMA model due to its inappropriateness towards highly volatile markets raise the questions about its accuracy and applicability for stocks such as Tesla.

1.6 Project beneficiaries

Investors, traders, and financial analysts are the primary interested parties regarding this proposal. Their main interest is to make reliable stock price predictions. It is problematic to predict and invest in Tesla stocks. The high volatility of the stocks makes volatility an issue for large regular investors as well as for smaller retail investors (Ma, 2020). The study will develop an effective model to forecast stock prices, namely, improve the ARIMA model to make short-term prediction more accurate.

The project's findings can be further exploited by financial institutions, this including such entities as hedge funds and portfolio managers. This group can benefit from the unveiling of reliable forecasting models provided by the project since such deliverables can be used to manage one's portfolio better. In general, due to improved asset allocation, investment funds' performance can be increased (Mei *et al.,* 2018).

1.7 Structure of the Study

The structure of the study contains several basic components, which suggest what the investigation is about. The layout of the report commences from the Abstract, which presents a brief overview of the study's content. The Introduction includes the Problem statement, aims and objectives, motivation of the study. In the Literature Survey, researchers look at previous studies and define the gap which the current project should cover. The Research Methodology section determines the steps, methods, models, and technologies that would be used during the investigation. The Design and Implementation Specifications prepare an assignment for the further solution of the problem.

2 Related Work

2.1 Introduction

The review evaluates the existing theories and methodologies in the context of the previous research, highlighting the trends, debates, and inconsistencies. Moreover, this section assists in developing the rationale for the relevance of the objective and the need for the further support of the research in the field (Mauerand Venecek, 2022). This literature review aimed to present a variety of approaches, theoretical backgrounds and empirical results in the domain of artificial intelligence, machine learning, and data science. The related areas have been developing fast, and continuous research is delving into new applications and opportunities as well as challenges.

2.2 To explore the application of ARIMA models in forecasting Tesla stock prices.

As per Weng *et al.,* (2022), Autoregressive integrated moving average model has gained much recognition for its applicability to time series forecasting. In particular, it has found application in the stock price. As the price is subject to constant fluctuations, it is particularly essential to explore, which way the price is going. For one thing, a model offers the flexibility required to process the non-stationary stock price data. On the other hand, by stabilizing the variance of time series through differencing, this modelling approach proves useful in ascertaining which way stock price changes. Functionally, this model comprises autoregression, moving averages, and integration.

According to Subakkar *et al.,* (2023), the ARIMA application in stock price forecasting has been extremely popular as the stock market is barely predictable. Instead of sinking into theoretical discussions, one can consider the example of Tesla's stock. Tesla's stock is hardly predictable, due to the company's extraordinary activity and innovation. ARIMA can be effectively used to predict the short-term peaks and falls of the price. Significant aspect is that the model will adjust to new trends and peculiarities of the stock movement changes as time goes by.

As per Yang (2024), the effectiveness of the use of ARIMA models to forecast the stock prices of such electric car manufacturing organizations as Tesla; such works generally refer to building ARIMA models through identifying a set of the best-fitting parameters, namely, p, d, and q, autoregressive, differencing, and moving average, respectively. Some scholars concluded that ARIMA models prove highly valuable assets to predicting short-term trends, yet they might be less efficient in terms of long forecasting. Thus, significant drops in stock prices might be the result of sudden market shocks or other events surrounding the company in question, such as innovating electric vehicles, quarterly earnings' announcements, or regulatory issues, and these might require additional adjustments made to enhance the effectiveness of ARIMA model application.

2.3 To analyse the relationship between historical stock price data and predictive accuracy using ARIMA techniques.

As per Khan and Alghulaiakh, (2020), the primary advantage of working with historical stock price data is that it is possible to predict stock prices within a broad time range. To a great degree, the effectiveness of Autoregressive Integrated Moving Average models in discussing historical stock data and the accuracy of predictions is based on such a premise. As a statistical and quantitative technique, ARIMA is widely recognized for its predictive accuracy. As a general rule, historical stock prices are used to predict future ones. In fact, the historical stock price data is the best foundation for forecasting because it di9scusses past trends, cyclical behaviours, and other aspects.

According to Reddy (2019), the accuracy of ARIMA models to predict stock prices mainly relies on the quality and properties of historical data. In general, stock prices are a non-stationary series of data when varying over time because of different market conditions. While ARIMA takes this into consideration by applying differencing, it appears only effective when the model can turn the stock data into stationary series, where mean and variance are consistent over time. In one way or another, the correlation between the data's linear components is captured by ARIMA, which seems essential to use in creating forecasts. As for shortcomings, however, the historical data must be rich enough and not contain any noise or anomalies.

According to the viewpoints of Ashok and Prathibhamol, (2021), historical stock prices positively influence the accuracy of prediction when the ARIMA model is applied. Such a claim is especially relevant in periods where no substantial external or internal shocks affect the market. Stock movement patterns are, in general, influenced by multiple factors, including movements within the exchange, and incorporation of past data is beneficial. However, on the other hand, the rate of the above movement has to be non-fluctuant, as the major impact of ARIMA on stocks is to provide the same forecast its historical data is based. Thus, for high-volatility stocks, such as the shares of technology companies like Tesla, the accuracy of ARIMA forecasts may lack, as unexpected movements among the investors or sporadic market movements may disrupt the price patterns which used to be set in the history.

2.4 To identify the limitations and challenges of using ARIMA models for forecasting in volatile stock markets like Tesla.

As per Hemkar Goswami, (2021), ARIMA (Automatic Regressive Index Moving Average) models are popular forecasting approaches because they have been observed to handle linear patterns and trends well in historical data patterns. Moreover, the present work also established that applying ARIMA in volatile stock markets such as Tesla bears a number of limitations and challenges that should not be ignored. In more specific concerns, a key issue relates to the fact that stock price data is non-stationary and non-linear by nature, something that ARIMA models do not well address. Furthermore, stock fluctuation due to expectations, planned and unplanned events, and market sentiment, news, any macro factors, and so on, are not new to investors especially when it comes to trading in Such highly unpredictable and volatile stocks as Tesla. These factors result to great fluctuations in prices that are specific and difficult to predict hence may not be effectively addressed by the ARIMA models, which are mostly based strongly on past patterns of data.

According to the viewpoints of Rubio *et al.*, (2023) another drawback of ARIMA models is that they are conditionally stationary, and that means that their parameters, including the mean and variance, must also remain fixed throughout the period of analysis. Generally, stock price tends to be non-stationary, because of the high variability in the stock market values, which does not enable efficient predictions by the means of ARIMA. Still, with data transformations to obtain stationarity (for example, differencing), the model may potentially not account for sharp spikes or drops characteristic of Tesla's stock prices. However, the whichever use of differencing might help to reduce trends, will only add to losing sufficient volatility in the future.

According to Silva *et al.*, (2018), ARIMA models work with univariate time series, and this assumes that no other factors can greatly affect stock prices; factors such as macroeconomic variables, earnings calls, industry news, among others. In the case of tesla, various forces exogenous to the organization such as technological forces, political forces and production constraints can lead to big swings in price. ARIMA model does not take such external variables and therefore it is not well suitable in occasions whereby short-term prices prove very sensitive to such factors.

2.5 To propose and implement countermeasures for improving forecasting accuracy and model performance when predicting Tesla stock prices using ARIMA models.

Hybrid Modeling Approaches: A possible solution to this is construction of hybrid models by combining ARIMA with machine learning approaches like Artificial Neural Networks (ANNs) or Support Vector Machine (SVM). ARIMA captures the linear patterns and a machine learning component captures the non-linear relationship. This method is a good combination of both frameworks, helping the model to handle the undetermined surrounding price action for Tesla (Khashei and Hajirahimi, 2019). A possible example is a hybrid ARIMA-ANN model to capture the short-term trends of ARIMA and ANNs can use complex, non-linear relationships associated with varying external influences.

Incorporation of Exogenous Variables: ARIMA depends only on past price data, we can improve this model with exogenous variables (ARIMAX). This model can add external factors, such as Tesla quarterly earnings, production announcement dates, global indices (SP500, Nasdaq) interest rates & regulatory changes. Such variables add a larger background and enhance the model's capability to anticipate how other events can impact Tesla prices (BELLO, 2021). Accordingly, adding in those accounts for decreased predictions accuracy during surprise market movements.

Dynamic Data Transformation Techniques: Volatile stock data often violate one of the most fundamental requirements of ARIMA, stationarity. Dynamic data transformations like seasonal decomposition of time series (STL decomposition) or wavelet transformation can be used to deal with this (Dhyani *et al.*, 2020). They are useful for filtering out noise, stabilizing variance and minimizing the influence of less significant trends or cyclical movements in Tesla stock price. This allows the ARIMA model to concentrate on actual trends and make more reliable short- and long-term forecasts.

Regularization to Prevent Overfitting: When ARIMA is applied to most, or for that matter any, of the highly volatile markets, overfitting becomes serious issue. To combat this, regularization can be used (Lasso or Ridge regression), which will correct the excessively complex models. This is to avoid ARIMA to over-fit history data and focus on noise or outliers what is very common in Tesla stock price movements (Kulaglic and Ustundag, 2021). Synchronization of a new predictive model, which reduced overfit and improved generalization to new data points allows for greater assurance of prediction.

Adaptive ARIMA Models: A different approach is to utilize adaptive ARIMA type models including Auto-ARIMA, or time-varying ARIMA. Such variants automate parameter selection with real time data to allow the model to be dynamic and adaptive to Tesla's rapidly changing market conditions (Hossain *et al.*, 2021). It can automatically pick optimal parameters and orders thus reducing the feature extraction process required and enhance the prediction power of that model additionally with time as it adapts to underlying data.

2.6 Literature Gap

Finding an ARIMA model to Predict Stock price is valuable for most the data. However, this leaves plenty of gaps unaddressed. First, nearly all the studies assume linearity and stationarity of data for ARIMA to be effective but stocks are usually non-linear and have a tendency to not remain stationary due to high volatility. More studies are required to broaden the scope by incorporating external variables and hybridization approaches into ARIMA paradigms to improve its forecasting ability further.

Moreover, ARIMA is well-known for short-term forecasting, are its long-term predictions also accurate in unpredictable markets. Additionally, very limited research develops complete countermeasures for model overfitting or tackles ARIMA application computation challenges concerning high-frequency stock data. Further, little attention has been afforded the real-time estimation of adaptive ARIMA models in stock price prediction through simulation studies which warrant some further research opportunities to shape the utility of such models in dynamic market settings.

2.7 Summary

This chapter started with the importance of literature reviews that allows one to acquire information about research gaps and review available knowledge. The presentation went through the advantages and disadvantages of ARIMA models for predicting stock prices such as Tesla where behaviour may be considered volatile. Different papers focused on the advantages of this model for short-term forecasting but also pointed problems like stationarity assumption, inability to capture non-linearities in data and high sensitivity to exogenous influences.

3 Research Methodology

3.1 CRISP-DM

This method gives a systematic approach to addressing data-related problems. In terms of predicting Tesla's stock prices using ARIMA models and data from Yahoo Finance, the approach is defined as, In the first phase, which is Business Understanding, the main goal is to forecast stock prices of Tesla for better funding resolution. Precise predictions can help stakeholders, involving analysts and investors, in maintaining risk and making sophisticated investment choices (Amirzadeh *et al.*, 2022). In the next phase, which is Data Understanding, the dataset is collected using the yfinance library in Python for Tesla stock prices, this library provides past data, involving low, high, close, and adjusted close prices. This data extends a specific time which is 5 years, and the exploratory analysis includes showing the trends in stock price, analyzing trends like seasonality, and identifying inconsistency or missing values in the data. The fundamental ingredients of data analysis are prepared well the drive into ARIMA modeling. Measures include dealing with missing responses, date column transformation into the datetime format, and, respectively, verification of the data sequence order. Also, the time series is made stationary by using techniques such as differencing where necessary to make the series stationary. The prediction is done by using the ARIMA (Auto-Regressive Integrated Moving Average) model. In model selection, one identifies the values to use for p, d, and q and there are also numerous methodologies to use including the ACF and PACF plots (Sable *et al.*, 2023). The model is first trained on the statistical history data, and detected by several indices such as Mean Absolute Error (MAE) and Root Square Error (RMSE). Further tuning takes place sequentially to improve the model. The accuracy of the model is assessed by differentiating the forecasted stock prices as opposed to the actual costs in the test dataset. The examination involves evaluating whether the forecasting aligns with the esteemed trends and computing error metrics to assess performance. If the outcome is not satisfying, the model or data variables might be re-evaluated for enhancement (Amirzadeh *et al.*, 2022).

Once the model is deployed and found reliable the ARIMA model is then used to predict the future stock prices of Tesla.

3.1.1 Business Understanding

The goal of business understanding is to predict Tesla's stock prices across a specific period to help analysts, financial institutions, and investors in making data-centric decisions. Prediction of the stock price makes it difficult to analyze trends in the market, optimize investment strategies, and maintain risks. It is also important to forecast its stock because Tesla is one of the most volatile stocks in the market with inherent risks affecting the financial industry. The objective is to use historical stock price data from Yahoo Finance, and then use ARIMA, a time series analysis, to predict the stocks (Amirzadeh *et al.*, 2022).

3.1.2 Data Understanding

The dataset consists of past data on Tesla's stock prices from the last five years, it is downloaded from Yahoo Finance by using the yfinance library. It involves primary financial metrics such as close, open, low, high, trading volume, and adjusted close prices, along with the data furnishing as the time index. Data understanding includes assessing volatility, trends, and seasonality in Tesla's stock prices across time (Zhao *et al.*, 2023). The first step is always to examine the magnitude of price changes by making a graphical display of the price data for any abnormalities or gaps. Statistical summaries are useful to determine the oscillation of the prices and the variations in the volume. This understanding however provides the basis of preprocessing and modeling to arrive at a set of data that can be aptly used in ARIMA-based forecasting.

EDA

1. **Tesla Stock Price Over Time**

The plots are for the closing stock prices of Tesla with the help of Matplotlib. The line graph is plotted using plt.plot for the Close price column of Tesla data which represents the changes in stock price. The figure size is set to 12x6 inches which gives it an easier appearance. This plot is named Tesla Stock Price Over Time to brandish the intention of the visualization. The former axis indicates the date, and the latter axis reveals the closing stock value in U.S. dollars. A legend is incorporated to label the line as the close price of Tesla. Such visualization makes it easier to observe the cyclicality and possible volatility of the price of Tesla stock as compared to a benchmark stock.

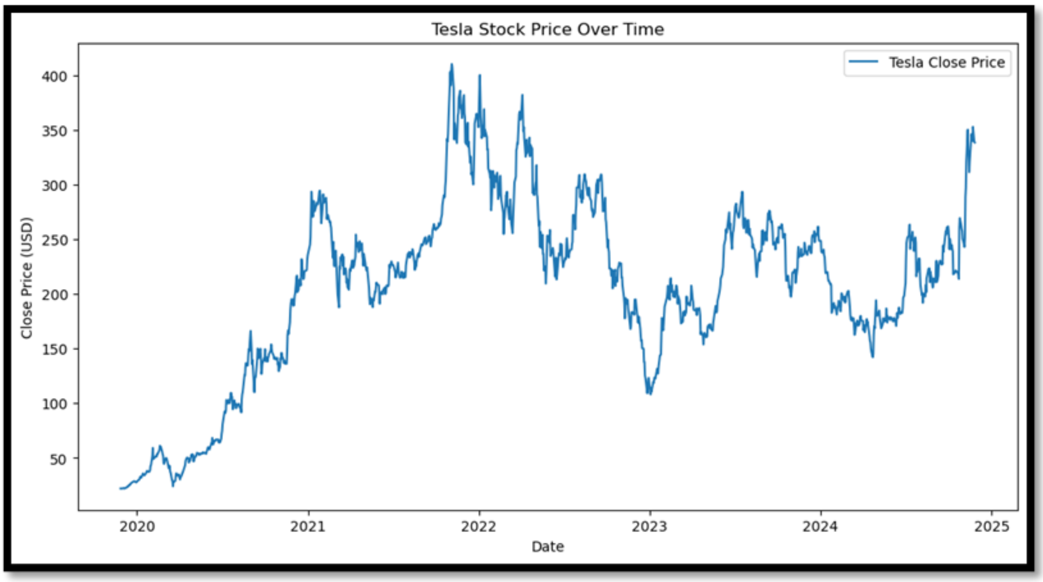

Figure 1: Tesla Stock Price Over Time

(Source: Self-Developed)

2. **Tesla Stock Price and Moving Averages**

This plot represents the Tesla stock price with its 7-day and 30-day moving averages. The plot represents the closing price of the stock which is represented by the Close column with the two moving averages one is for a 7-day moving average and the other is a 30-day moving average. By advancing these averages, any cyclical distortions that exist in the data are removed and the more significant trends become apparent as moving averages of over specified windows of stock prices. The 7-day MA responds to short-run price fluctuations better than the 30-day MA which follows longer-term trends. On the x-axis there is the dates while the y-axis portraying the stock prices in USD. This legend assists in differentiating between the closing price and two moving averages and assists in understanding the prices movements and analysis.

Figure 2: Tesla Stock Price and Moving Averages

(Source: Self-Developed)

3.1.3 Data Preprocessing

This phase makes sure that the dataset must be clean and ready for ARIMA modeling. In this phase missing values are checked by using tesla.isnull().sum(), and any gaps are solved via insertion or forward-filling. Feature selection is also done to improve the dataset, in it, the moving average is calculated in a 7-day and 30-day window to get short and medium trends, which helps in acknowledging price movements. And in last, the dataset is split into testing and training sets for modeling (Zhao *et al.*, 2023).

3.1.4 Modeling

This phase includes forecasting frameworks to predict the stock prices of Tesla. The ARIMA (Auto-Regressive Integrated Moving Average) model catches linear trends, autocorrelation in time series data, and seasonality, absolute for non-seasonal data. SARIMA enhances ARIMA by adjoining seasonal components to manage periodic trends, making it applicable for data along with continuing trends (Sonkavde *et al.*, 2023). The deep learning model, LSTM (Long Short-Term Memory) outperforms traditional techniques to capture highly nonlinear relationships as well as long-term dependencies in the sequentially organized data of highly volatile stocks. Both models are built from historical data and checked in terms of metrics like RMSE and selected according to the accuracy of forecasts and characteristics of the sets.

3.1.5 Evaluation

This phase compares the three models ARIMA, SARIMA, and LSTM, which depend on predicting preciseness using metrics such as R-squared (R^2), and MSE (Mean Squared Error). ARIMA perceived an MSE of 72.70 and an R^2 of 0.989, depicting robust linear pattern catches but having limited accessibility to non-linear patterns.

SARIMA enhanced seasonality handling, having an MSE of 0.0005 and R^2 of 0.9894, depicting better forecasting for periodic data. LSTM, while catching crucial aspects, has an MSE of 295.07 and R^2 of 0.95, emphasizing challenges along with volatility. SARIMA represents the best performance, maintaining precision, and adaptability for predicting the stock price of Tesla.

Later, the Augmented Dickey-Duller (ADF) test on the closing stock price of Tesla to check for stability, a critical presumption for time series prediction models like ARIMA. The result shows the involvement of the ADF statistic, the p-value, and the typical values at distinct consequence levels (1%, 5%, and 10%).

```
ADF Statistic: -2.369513369639033
p-value: 0.15052652863242327
Critical Values: {'1%': -3.4356048614183443, '5%': -2.8638605461891617, '10%': -2.5680054872544145}
```

Figure 3:Result of Augmented Dickey-Fuller test on 'Close' column

(Source: Self-Developed)

If the value of p is more than the consideration level which is usually 0.05 then the null hypothesis cannot be declined, this shows that the time series is not stable and this is correct for the stock of Tesla. To this end, the code uses first-order differencing by applying the diff() method which reduces trends and seasonality by subtracting each value from the previous one. The differenced series is then plotted to represent the fluctuations in Tesla's closing price after excluding the persisting pattern (Kurek, 2024).

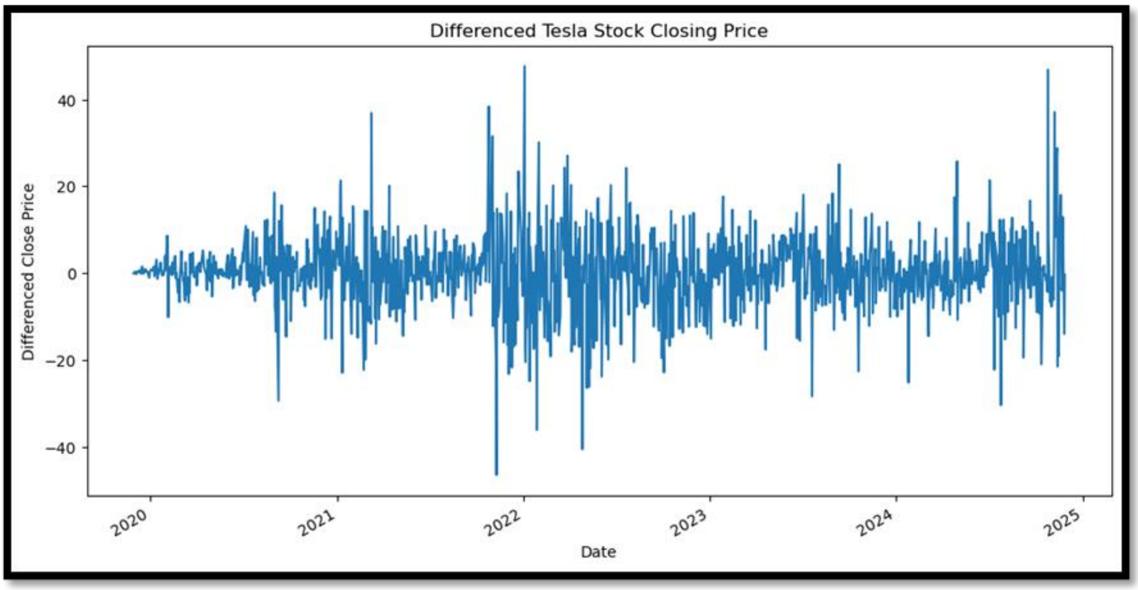

Figure 4:Differenced Tesla Stock Closing Price

(Source: Self-Developed)

3.1.6 Deployment

Deployment requires incorporating SARIMA into an easy-to-use application or system to forecast the real-time Tesla stock price. The model is serialized using libraries such as pickle or joblib for more convenient real use and then the whole process takes place on the server or cloud platform. A frontline interface in the form of an online dashboard or API receives parameters (e.g., forecast horizon) and provides results accompanied by graphs. Periodic inputs into the model with current stock data help to keep the model current (Sonkavde *et al.*, 2023). Also, real-time or time-sensitive information, for example, high volatility rates may be added, giving indications of growth or decline in share rates that investors and financial analysts may consider.

4 Design Specification

Some of the features are run by using the framework and it is apparent that use they are used in the methodology for numerous application software, libraries, and programming languages which help analyze modeling and data. Mainly the language used in this is Python, and it is well known for their libraries as well as for frameworks that are essential for machine learning and data science (Han and Kwak, 2023).

- **Pandas**: This is mainly because the environment for the analysis and modification of the pandas library is convenient data it is also capable of changing the provided data set after it brings out its analysis on it.
- **Matplotlib and Seaborn**: These visualization libraries were used in the determination of certain of the plots as well as charts that we employed when discussing various aspects of data distributions, relation concerning the present state of affairs of the prediction models, together with their future performances (Lavanya *et al.*, 2023).
- **Scikit-learn**: Some of them are available in scikit-learn and this includes logistic regression, and other techniques such as random forest, decision tree, k nearest neighbors, and so forth upon which the present study has used.
- **NumPy:** It is a numeric processing library, which includes array computations for effective data analysis especially when combined with Pandas (Lavanya et al., 2023).
- **yfinance**: yfinance is a great Python library that is often used to pull data from Yahoo Finance. It contains historical stock price information including dividends, splits, and other financial statistics.

5 Implementation

In addition, the review examined how historical stock price data relates to predictive accuracy. In addition, the section also discovered constraints including overfitting and the need for extensive training to completely realize its effectiveness, calling for hybrid techniques and incorporating external elements to further bolster forecasting ability. The chapter ended with some countermeasures to enhance the performance of the ARIMA model in dynamic, volatile markets such as that of Tesla.

6 Evaluation

6.1 ARIMA

The model output gives typical observations into the time-series behavior of Tesla's stock closing prices. The summary of the model involves feature estimates, goodness-of-fit measures, and statistical importance (Tiwary and Mishra, 2022). Consequential p-values for MA and AR coefficients indicate that they are important to the model.

```
                               SARIMAX Results
==============================================================================
Dep. Variable:                   TSLA   No. Observations:                 1258
Model:                 ARIMA(1, 1, 1)   Log Likelihood               -4477.593
Date:                Wed, 27 Nov 2024   AIC                           8961.186
Time:                        14:52:09   BIC                           8976.595
Sample:                             0   HQIC                          8966.977
                               - 1258
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1         -0.5894      0.515     -1.145      0.252      -1.598       0.420
ma.L1          0.5641      0.527      1.070      0.285      -0.469       1.598
sigma2        72.7019      1.662     43.752      0.000      69.445      75.959
===================================================================================
Ljung-Box (L1) (Q):                   0.02   Jarque-Bera (JB):               904.01
Prob(Q):                              0.90   Prob(JB):                         0.00
Heteroskedasticity (H):               1.59   Skew:                             0.06
Prob(H) (two-sided):                  0.00   Kurtosis:                         7.15
===================================================================================
```

Figure 5: SARIMAX Result

(Source: Self-Developed)

The remaining plot gives a visual estimation of the model's suitability. Mainly, unused should look like white noise, showing no left autocorrelation or structured patterns in the data. If the leftovers show randomness, the ARIMA model efficiently catches the basic form of the series. Contrarily, trends or patterns in fragments may show misspecification or underfitting.

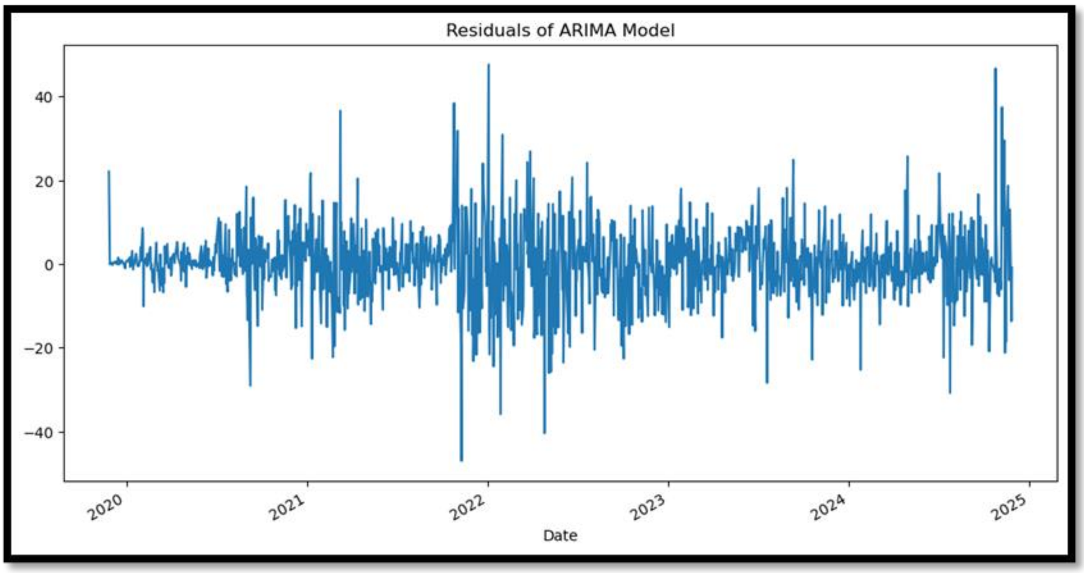

Figure 6: Residuals of the ARIMA model

(Source: Self-Developed)

On the other hand, MSE gives an accurate measure of error, and R^2 gives a comparative measure of model fit. As one, these metrics aid in evaluating the ARIMA model's reliability in prediction. If the MSE is comparably low but R^2 is not satisfactory, it might show that the time series have complexities or patterns which is not fully apprehended by the ARIMA (1,1,1) model. Additional improvements are still desirable, for example, selecting other appropriate values for ARIMA orders or adding more exogenous variables could give even better results (Subakkar *et al.*, 2023).

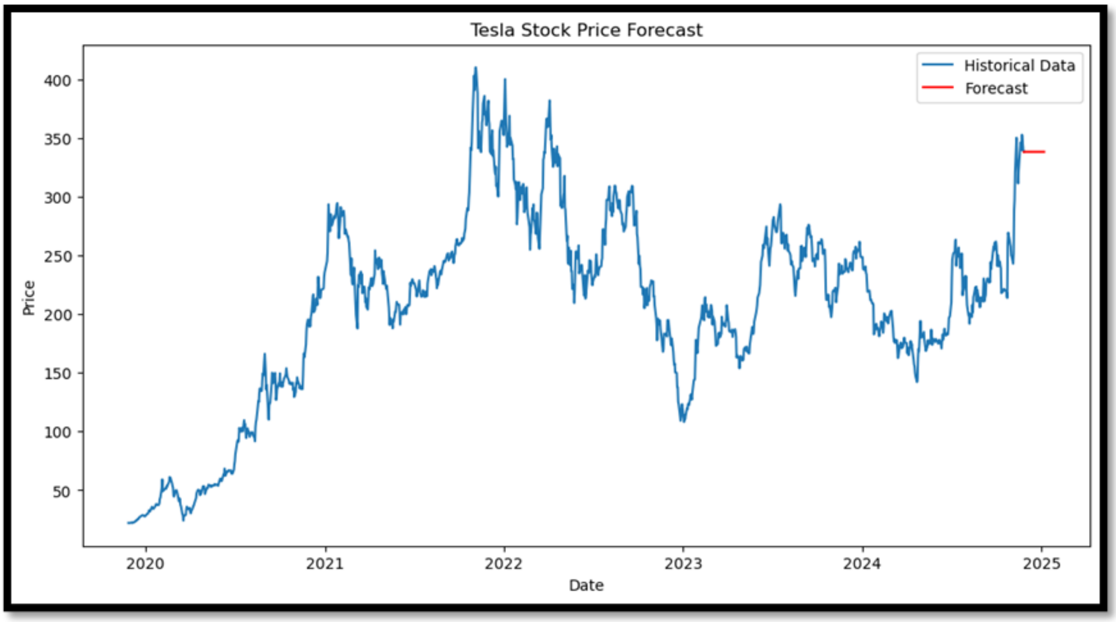

Figure 7: Tesla Stock Price Forecast

(Source: Self-Developed)

6.2 SARIMA

The seasonal=True shows that the data reveal seasonality, along with m=12, indicating a periodicity of 12-time pace.

```
Performing stepwise search to minimize aic
 ARIMA(2,1,2)(1,0,1)[12] intercept   : AIC=-6019.368, Time=3.52 sec
 ARIMA(0,1,0)(0,0,0)[12] intercept   : AIC=-6026.478, Time=0.19 sec
 ARIMA(1,1,0)(1,0,0)[12] intercept   : AIC=-6023.706, Time=0.33 sec
 ARIMA(0,1,1)(0,0,1)[12] intercept   : AIC=-6023.570, Time=0.49 sec
 ARIMA(0,1,0)(0,0,0)[12]             : AIC=-6027.385, Time=0.08 sec
 ARIMA(0,1,0)(1,0,0)[12] intercept   : AIC=-6025.072, Time=0.23 sec
 ARIMA(0,1,0)(0,0,1)[12] intercept   : AIC=-6024.959, Time=0.30 sec
 ARIMA(0,1,0)(1,0,1)[12] intercept   : AIC=-6024.317, Time=0.62 sec
 ARIMA(1,1,0)(0,0,0)[12] intercept   : AIC=-6025.109, Time=0.21 sec
 ARIMA(0,1,1)(0,0,0)[12] intercept   : AIC=-6025.089, Time=0.18 sec
 ARIMA(1,1,1)(0,0,0)[12] intercept   : AIC=-6023.714, Time=0.32 sec

Best model:  ARIMA(0,1,0)(0,0,0)[12]
Total fit time: 6.476 seconds
```

Figure 8: Stepwise performance

(Source: Self-Developed)

The function repetitively examines combinations of seasonal orders (P, D, Q, m) and ARIMA orders (p, d, q) by using norms like AIC (Akaike Information Criteria) to manage model complexity and fit. The trace = True option records the assessing process, and the summary() result gives observations into the chosen model's parameters, diagnostic metrics, and coefficients. It is an automated technique that streamlines the identification of prime SARIMA parameters, improving predicting accuracy for seasonal data (Lin, 2023).

```
                               SARIMAX Results
==============================================================================
Dep. Variable:                      y   No. Observations:                 1258
Model:               SARIMAX(0, 1, 0)   Log Likelihood                3014.692
Date:                Wed, 27 Nov 2024   AIC                          -6027.385
Time:                        14:52:26   BIC                          -6022.248
Sample:                             0   HQIC                         -6025.454
                               - 1258
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
sigma2         0.0005    1.1e-05     43.840      0.000       0.000       0.001
===================================================================================
Ljung-Box (L1) (Q):                   0.63   Jarque-Bera (JB):               885.95
Prob(Q):                              0.43   Prob(JB):                         0.00
Heteroskedasticity (H):               1.59   Skew:                             0.05
Prob(H) (two-sided):                  0.00   Kurtosis:                         7.11
===================================================================================
```

Figure 9: SARIMAX Result

(Source: Self-Developed)

The estimated mean values are shown in the red line since they depict the model's predicted stock prices. The vertical bars to the right of the forecasted mean represent the confidence intervals, which define the less certain range of values for models falling within the region of confidence intervals because of the existing uncertainty level (Subakkar *et al.*, 2023). The pink color used in the figure can be interpreted as the range of forecast confidence if the confidence is at 95%.

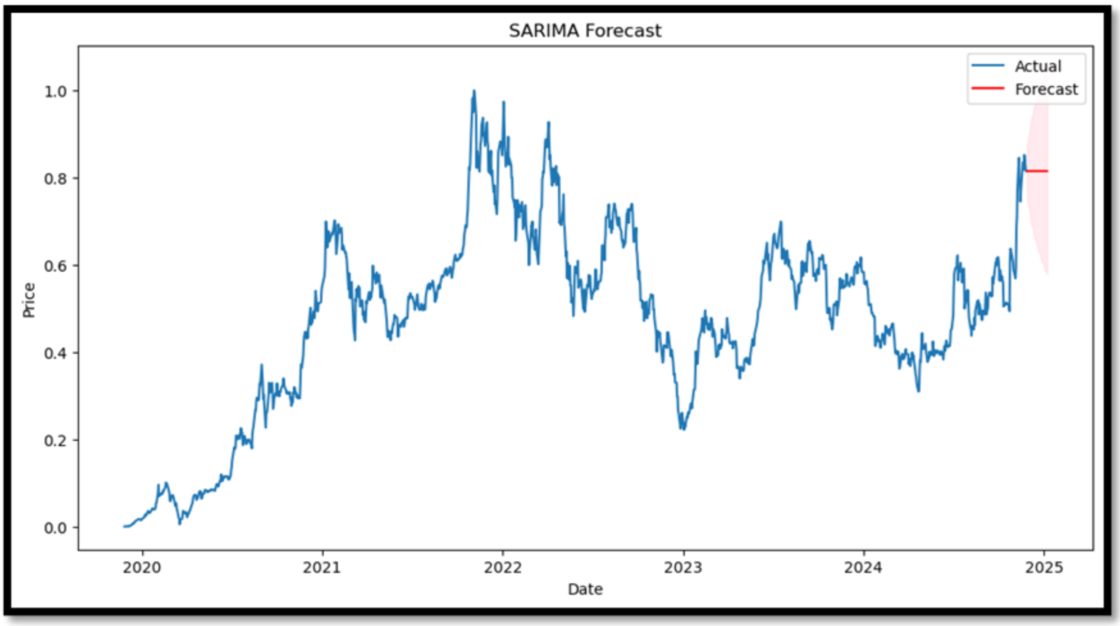

Figure 10: SARIMA forecast

(Source: Self-Developed)

This plot permits us to make a direct comparison of historical prices with the forecast ones so that we can get some idea of how the model believes that the stock will develop itself. The confidence interval also demonstrates the level of risk associated with the forecast and the more the intervals are spread out the more risk involved. Generally, the visualization aids in assessing both the forecasted direction and the possible risk of Tesla's future stock prices, giving important information for decision-making.

The result analyses two main metrics, R-squared (R^2) and Mean Squared Error (MSE), to examine the performance of the SARIMA model on the in-sample data which is past data. MSE assessed the mean squared difference among the predicted and actual values (fitted values) deriving out the SARIMA model. A lower MSE shows that the forecasting models are nearest to the true values, showing better model preciseness. If the value of MSE is high, it advises that the SARIMA model is not catching the fundamental patterns of the time series effectively, deriving huge forecasting flaws (Tiwary and Mishra, 2022).

6.3 LSTM

This model is composed of an input layer of LSTM (50 units), a hidden layer of LSTM (50 units), and two dropout layers to avoid overfitting. The last layer of the model is a Dense layer which gives the predicted closing price. The model is built with the Adam optimizer and mean squared error function provided for regression problems.

The model is trained for 20 epochs in the case of training data and uses a batch size of 32 and finally, the accuracy is computed by the test set. In the following section, the results from the model are rescaled back to the correct scale of the stocks using the scaler.inverse_transform() function. This step is important because the data used to train the classifiers was normalized before training to enhance the accuracy of the classifiers. The results are assessed by differentiating the forecasted and actual costs on the test set. The progress of the LSTM model can be identified by metrics like R-squared or Mean Squared Error, showing how suitably the model catches trends in Tesla stock prices (Sk and Javvadi, 2023). If the forecasting nearly follows the true value, it could propound that the LSTM model is efficient in predicting future prices, via extreme tuning, like expanding training epochs or adjusting hyperparameters, which would enhance its preciseness.

Later, the plot is plotted to compare the predicted and actual price of Tesla stock. By using matplotlib the figure size is set to 12x6 inches for better description. Then after two sets of data are plotted, the true price is shown in blue and the predicted price is shown in red color. It also provides an easy means of ensuring the easiest forms of relativity between the two datasets. The plot is labeled as "Tesla Stock Price Prediction" which gives the identity of the graph showing the trends of the stock price of Tesla. To the lower horizontal axis, the label sequence is "Time," and this is the period where the stock prices are observed in intervals of day or months, etc. On the y-axis, there is a legend with information that indicates that it is normalized to the price of Tesla's stocks. A legend is included to ensure that it becomes easy for the viewers to identify what dataset corresponds to which price line between the actual and the predicted prices (Alkhatib *et al.*, 2022). Lastly, the plot is shown by using plt.show() to allow to see the trend of stock price prediction of Tesla and thus indicates how effectively the model can predict the prices of Tesla stocks.

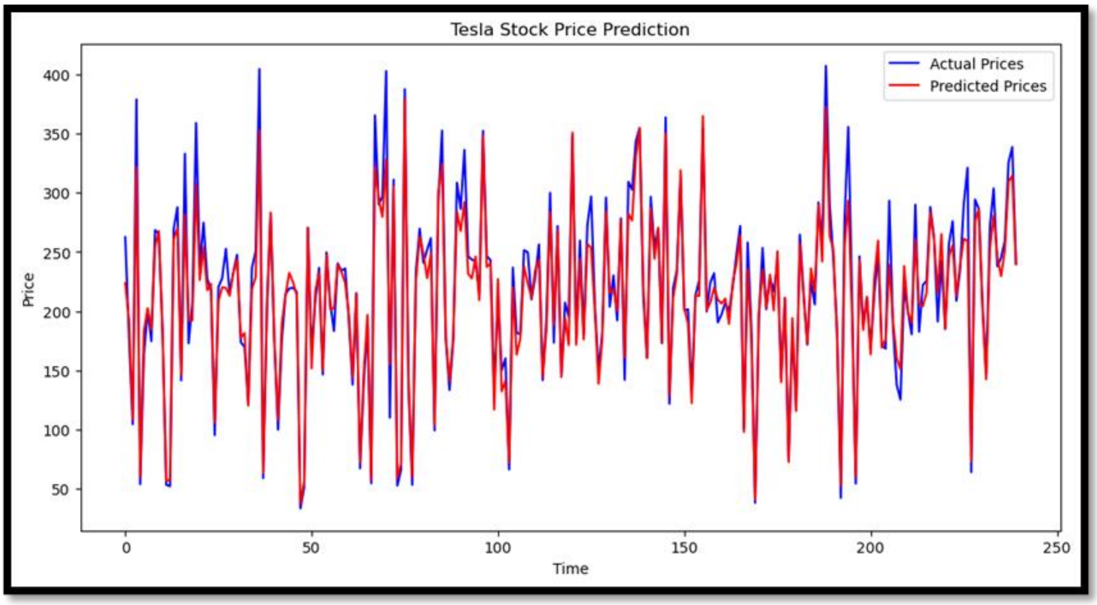

Figure 11: Tesla Stock Prediction

(Source: Self-Developed)

6.4 Discussion

These three models give a different approach to predicting time series by their limitations and strengths. ARIMA works suitably whenever the data represent a linear pattern and could be distinct to separate trends or seasonality (Subakkar *et al.*, 2023). It is majorly easy to work well, implement, and estimate effectiveness by smaller datasets. However, one of the greatest weaknesses of ARIMA is the inability to model non-linear relations or compounding seasonality patterns. Stock prices are said to be financial data that are known to have fluctuations, non-stationarities, and shifts that are not easy to model by ARIMA methods. In Tesla's stock price prediction, SARIMA (Seasonal ARIMA) enlarges ARIMA by appending seasonal components to manage time series data along with the periodic seasonality or trends, making it stronger than ARIMA whenever seasonality is here. It involves additional variables for seasonal autoregression (P), seasonal moving averages (Q), seasonal differencing (D), and seasonal periods (m). This is much more effective for datasets such as Tesla, stock prices, which might show trends that repeat periodically (like yearly or monthly cycles). SARIMA is a more general approach than the basic one that has been just described because it deals more directly with the seasonal parts of the series and can handle changes that repeat at appropriate intervals (Ma, 2020). LSTM is a type of Recurrent Neural Network (RNN) that is specifically designed to catch long-term relativity in successive data. It can determine non-linear relations and model typical patterns, which makes them extremely effective for predicting an eruptive economic market. As long, as ARIMA and SARIMA models need to work along with the data that have been formed stable (Ma, 2020).

7 Conclusion and Future Work

This study focused on forecasting Tesla's stock prices using three distinct models: ARIMA, SARIMA, and LSTM. As mentioned in the introduction part, accurate prediction of the future stock prices is crucial for investments portfolio choices and risk management. ARIMA is the most fundamental time-series model used in this study because of its capacity to detect linear features. ARIMA was shown to be well suited for modeling the Tesla stock price but as seen it could not capture non-linear trends and volatility.

To overcome these limitations, a SARIMA model has been used where seasonality has been added to predict data having the periodic nature. SARIMA was found to be a better model as compared to its basic form ARIMA, in terms of Mean squared error (MSE) and R squared (R^2), that described it ability to estimate better for the fluctuating SES of the Tesla stock prices. The study also used an LSTM model which although it is designed for capturing non-linear properties and can have some difficulty in futures stock prices during periods of high volatility. However, it was helpful in gaining greater understanding of the main trends in price changes.

Further, statistical tests such as the ADF test were applied, which also evidenced that the Tesla stock price data is stationary and thus effectively utilisable for time series model. Lastly, we found out that SARIMA was the most effective model for forecasting Tesla stock price with the effectiveness and flexibility of the model in contrast to ARIMA and LSTM models, where there is some level of effectiveness based on the particular kind of model being used.

References

Afeef, M., Ihsan, A. and Zada, H., 2018. Forecasting stock prices through univariate ARIMA modeling. *NUML International Journal of Business & Management*, *13*(2), pp.130-143.

Alkhatib, K., Khazaleh, H., Alkhazaleh, H.A., Alsoud, A.R. and Abualigah, L., 2022. A new stock price forecasting method using active deep learning approach. *Journal of Open Innovation: Technology, Market, and Complexity*, *8*(2), p.96.

Amirzadeh, R., Nazari, A. and Thiruvady, D., 2022. Applying artificial intelligence in cryptocurrency markets: A survey. *Algorithms*, *15*(11), p.428.

Ashok, A. and Prathibhamol, C.P., 2021, January. Improved analysis of stock market prediction:(ARIMA-LSTM-SMP). In *2021 4th Biennial International Conference on Nascent Technologies in Engineering (ICNTE)* (pp. 1-5). IEEE.

BELLO, A.O., 2021. FITTING AUTOREGRESSIVE INTEGRATED MOVING AVERAGE WITH EXOGENOUS VARIABLES MODEL ASSUMING LOGNORMAL ERROR TERM (Doctoral dissertation).

Choi, H.K., 2018. Stock price correlation coefficient prediction with ARIMA-LSTM hybrid model. *arXiv preprint arXiv:1808.01560*.

Dhyani, B., Kumar, M., Verma, P. and Jain, A., 2020. Stock market forecasting technique using arima model. *International Journal of Recent Technology and Engineering*, *8*(6), pp.2694-2697.

Doshi, A., Issa, A., Sachdeva, P., Rafati, S. and Rakshit, S., 2020. Deep stock predictions. *arXiv preprint arXiv:2006.04992*.

Ghous, H., Malik, M.H., Mahrukh, A. and Zaffar, A.M., 2023. Exchange stock price prediction using time series data: A survey. *Pakistan Journal of Humanities and Social Sciences*, *11*(2), pp.1110-1124.

Han, S. and Kwak, I.Y., 2023. Mastering data visualization with Python: practical tips for researchers. *Journal of Minimally Invasive Surgery*, *26*(4), p.167.

Hemkar Goswami, A.K., 2021. Stock Market Prediction-A Comparative Analysis.

Hossain, M.R., Ismail, M.T. and Karim, S.A.B.A., 2021. Improving stock price prediction using combining forecasts methods. *IEEE Access*, *9*, pp.132319-132328.

Khan, S. and Alghulaiakh, H., 2020. ARIMA model for accurate time series stocks forecasting. *International Journal of Advanced Computer Science and Applications*, *11*(7).

Khanderwal, S. and Mohanty, D., 2021. Stock price prediction using ARIMA model. *International Journal of Marketing & Human Resource Research*, *2*(2), pp.98-107.

Khashei, M. and Hajirahimi, Z., 2019. A comparative study of series arima/mlp hybrid models for stock price forecasting. Communications in Statistics-Simulation and Computation, 48(9), pp.2625-2640.

Kulaglic, A. and Ustundag, B.B., 2021. Stock Price Prediction Using Predictive Error Compensation Wavelet Neural Networks. *Computers, Materials & Continua*, *68*(3).

Kurek, P., 2024. Algorithmic Trading with Long Short-Term Memory Recurrent Neural Network and Stationary Time Series.

Lavanya, A., Sindhuja, S., Gaurav, L. and Ali, W., 2023. A comprehensive review of data visualization tools: features, strengths, and weaknesses. *Int. J. Comput. Eng. Res. Trends*, *10*(01), pp.10-20.

Lin, J., 2023. The Fed's Interest Rate Policy and Changes in Tesla's Share Price: Evidence from ARIMA Model. *Highlights in Business, Economics and Management*, *20*, pp.185-192.

Ma, Q., 2020. Comparison of ARIMA, ANN and LSTM for stock price prediction. In *E3S Web of Conferences* (Vol. 218, p. 01026). EDP Sciences.

Mauer, B. and Venecek, J., 2022. Writing the Literature Review. *Strategies for Conducting Literary Research, 2e*.

Mei, W., Xu, P., Liu, R. and Liu, J., 2018. Stock price prediction based on arima-svm model. In *International Conference on Big Data and Artificial Intelligence* (p. 4).

Panchal, S.A., Ferdouse, L. and Sultana, A., 2024, July. Comparative Analysis of ARIMA and LSTM Models for Stock Price Prediction. In *2024 IEEE/ACIS 27th International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD)* (pp. 240-244). IEEE.

Reddy, C.V., 2019. Predicting the stock market index using stochastic time series ARIMA modelling: The sample of BSE and NSE. *Indian Journal of Finance, 13*(8), pp.7-25.

Rubio, L., Palacio Pinedo, A., Mejía Castaño, A. and Ramos, F., 2023. Forecasting volatility by using wavelet transform, ARIMA and GARCH models. Eurasian Economic Review, 13(3), pp.803-830.

Sable, R., Goel, S. and Chatterjee, P., 2023. Techniques for Stock Market Prediction: A Review.

Silva, E.G., Júunior, D.S.D.O., Cavalcanti, G.D. and de Mattos Neto, P.S., 2018, July. Improving the accuracy of intelligent forecasting models using the perturbation theory. In 2018 International Joint Conference on Neural Networks (IJCNN) (pp. 1-7). IEEE.

Sk, K.B. and Javvadi, S., 2023. Predictions of Tesla Stock Price Based on Machine Learning Model. https://www.researchgate.net/publication/379908383_Predictive_Analysis_of_Tesla_Inc_Stock_with_Machine_Learning

Sonkavde, G., Dharrao, D.S., Bongale, A.M., Deokate, S.T., Doreswamy, D. and Bhat, S.K., 2023. Forecasting stock market prices using machine learning and deep learning models: A systematic review, performance analysis and discussion of implications. *International Journal of Financial Studies*, *11*(3), p.94.

Subakkar, A., Graceline Jasmine, S., Jani Anbarasi, L., Ganesh, J. and Yuktha Sri, C.M., 2023, January. An Analysis on Tesla's Stock Price Forecasting Using ARIMA Model. In *Proceedings of the International Conference on Cognitive and Intelligent Computing: ICCIC 2021, Volume 2* (pp. 83-89). Singapore: Springer Nature Singapore.

Tacchini, S., 2024. *Forecasting stock market behavior using artificial intelligence* (Master's thesis, Universitat Politècnica de Catalunya).

Tang, Y., 2024. The impact of brake failure rights protection event on Tesla Motors: Stock prediction based on ARIMA model. In *SHS Web of Conferences* (Vol. 188, p. 01011). EDP Sciences.

Tiwary, S. and Mishra, P.K., 2022, August. Time-series Forecasting of Stock Prices using ARIMA: A Case Study of TESLA and NIO. In *Proceedings of the 2nd Indian International Conference on Industrial Engineering and Operations Management Warangal* (pp. 16-18).

Vuong, P.H., Phu, L.H., Van Nguyen, T.H., Duy, L.N., Bao, P.T. and Trinh, T.D., 2024. A bibliometric literature review of stock price forecasting: from statistical model to deep learning approach. *Science Progress*, *107*(1), p.00368504241236557.

Weng, Q., Liu, R. and Tao, Z., 2022. Forecasting Tesla's Stock Price Using the ARIMA Model. *Proceedings of Business and Economic Studies*, *5*(5), pp.38-45.

Wijesinghe, G.W.R.I. and Rathnayaka, R.M.K.T., 2020, December. Stock Market Price Forecasting using ARIMA vs ANN; A Case study from CSE. In *2020 2nd International Conference on Advancements in Computing (ICAC)* (Vol. 1, pp. 269-274). IEEE.

Yang, K., 2024. Forecasting EV Stock Trends Based on ARIMA Model Represented by Tesla and BYD. *Highlights in Business, Economics and Management*, *30*, pp.135-141.

Zhao, C., Wu, M., Liu, J., Duan, Z., Shen, L., Shangguan, X., Liu, D. and Wang, Y., 2023. Progress and prospects of data-driven stock price forecasting research. *International Journal of Cognitive Computing in Engineering*, *4*, pp.100-108.