|  |
| --- |
|  |
| Forecast-Driven Strategic Planning: Predicting Tesla Stock Performance Using Time Series Forecasting Analysis |
| |  |  |  | | --- | --- | --- | | Raushan Abenova | 2024338 |  | |

A logo for college computing

Description automatically generated

**Assessment Cover Page**

|  |  |
| --- | --- |
| *Student Full Name* | Raushan Abenova |
| *Student Number* | 2024338 |
| *Module Title* | Strategic Thinking |
| *Assessment Title* | CA 3 – Capstone Project Proposal |
| *Assessment Due Date* | 18-03-2025 |
| *Date of Submission* |  |

**Declaration**

By submitting this assessment, I confirm that I have read the CCT policy on academic misconduct and understand the implications of submitting work that is not my own or does not appropriately reference material taken from a third party or other source.

I declare it to be my own work and that all material from third parties has been appropriately referenced.

I further confirm that this work has not previously been submitted for assessment by myself or someone else in CCT College Dublin or any other higher education institution.

Table of Contents

[3](#_Toc198048229)

[1. Proposal Project 4](#_Toc198048230)

[Introduction 4](#_Toc198048231)

[2. Business Description 4](#_Toc198048232)

[3. Project Hypothesis and Objectives 5](#_Toc198048233)

[Hypothesis: 5](#_Toc198048234)

[General Objective: 5](#_Toc198048235)

[Technical Objective 5](#_Toc198048236)

[4. Scope and Methodology 5](#_Toc198048237)

[Scope 5](#_Toc198048238)

[Methodology 6](#_Toc198048239)

[Ethical considerations 7](#_Toc198048240)

[4. Project Management and Planning 7](#_Toc198048241)

[Data Source Overview 7](#_Toc198048242)

[Data Exploratory Data Analysis (EDA) 7](#_Toc198048243)

[Data Dictionary 8](#_Toc198048244)

[5. Modelling 8](#_Toc198048245)

[Summary Statistics 8](#_Toc198048246)

[Bias Distribution and Outlier Detection 10](#_Toc198048247)

[Pre-processing and Data Cleaning 10](#_Toc198048248)

[Trend Analysis Prior to Stationarity Testing 11](#_Toc198048249)

[Trend and Seasonality Analysis – Seasonal Decomposition of Tesla Closing Price for 365 days. 11](#_Toc198048250)

[To better understand the underlying structure of Tesla’s stock behaviour, a seasonal decomposition was applied to the Close price series using additive components. This technique splits the time series into Trend, Seasonality, and Residuals, offering insights into the nature and drivers of price movements over time. 11](#_Toc198048251)

[Stationarity Testing and ARIMA Experimentation 13](#_Toc198048252)

[6. Machine Learning Algorithm 13](#_Toc198048253)

[6.1. Short-Term Seasonal Decomposition and Its Role in Modelling 13](#_Toc198048254)

[6.2. Stationarity Testing and Hypothesis Interpretation /ADF Test 14](#_Toc198048255)

[Stationarity Confirmation After Differencing 15](#_Toc198048256)

[ACF and PACF Analysis for ARIMA Model Selection 15](#_Toc198048257)

[ARIMA Modelling and Forecast for 2025-02-19 16](#_Toc198048258)

[SARIMA Forecasting – Experimental Prediction for February 19, 2025 18](#_Toc198048259)

[LSTM Forecasting – Experimental Prediction for February 19, 2025 19](#_Toc198048260)

[8. Modelling and Forecasting for April 15, 2026 20](#_Toc198048261)

[ARIMA Forecast for 2026-04-15 20](#_Toc198048262)

[SARIMA Forecast for 2026-04-15 21](#_Toc198048263)

[LSTM Forecast for 2026-04-15 22](#_Toc198048264)

[9. Evaluation and Comparison of Models for forecasting for 2026-04-15 22](#_Toc198048265)

[ARIMA, SARIMA, and LSTM Metrics 22](#_Toc198048266)

[10. Conclusion 24](#_Toc198048267)

[References: 25](#_Toc198048268)

# Figures

Figure 1………………………………………………………………………………….

Figure 2

Figure 2

Figure 2

Figure 2

Figure 2

Figure 2

Figure 2

Figure 2

Figure 2

Figure 2

Figure 2

Figure 2

Figure 2

Figure 2

Figure 2

# 1. Proposal Project

## Introduction

In the modern financial environment, the ability to anticipate market behaviour is a vital component of strategic decision-making. This project focuses on predictive analytics as a tool for forecasting stock performance, with a case study on Tesla Inc., a leading electric vehicle and technology company.

By applying time series forecasting models analysis to Tesla’s historical stock data, the goal is to develop insights, which how could support long-term planning, risk mitigation, and strategic investment decisions.

“Strategic analytics is not just about predicting the future, but about understanding the forces behind the patterns. Techniques like clustering give business leaders that necessary context.” T. Davenport, J. Harris (2007).

Given Tesla's volatility and its position in a high-growth sector, this analysis aims to explore how well different models like ARIMA, SARIMA and LSTM can capture market dynamics and project future stock behaviour, and probability of forecasting price Tesla's stock for April 15, 2026.

### 2. Business Description

Tesla Inc., founded in 2003, however publicly traded since 2010. It is a global leader in electric vehicles and clean energy. Its mission to accelerate the shift to sustainable energy has positioned it as both a technological innovator and a high-profile brand. Today Tesla offers electric vehicles, battery storage, solar products, and autonomous driving systems.

While the company has seen rapid growth, its stock is known for high volatility influenced by public sentiment, innovation news, and broader economic trends. Particularly, in April 2025 Tesla has been under financial pressure, reporting a 71% drop in quarterly net income and a 9% year-over-year revenue, that to cost nearly $19.3 billion. The factors, like weakening global EV demand, competition with Chinese firms, and very recently U.S. tariffs, it gives a tremendous affection on Company’s supply chain. Despite setbacks, Tesla is boldly pushing forward with innovation. In a short future a pilot robotaxi service in Austin will use self-driving Model Y vehicles, and the autonomous Cybercab is expected to enter production in next 2026. At the same time Bloomberg analysts express cautious optimism, noting that renewed focus and upcoming technologies may support a future rebound.

## **3. Project Hypothesis and Objectives**

### **Hypothesis:**

Despite Tesla’s well-known stock volatility, it is hypothesized that time series forecasting models—specifically ARIMA, SARIMA, and LSTM—can provide meaningful predictions for the company’s future stock price. The core testing approach of this project involves comparing

## these models using standard evaluation metrics such as R², MAE, MAPE, and RMSE to assess their real-world applicability in financial forecasting. **Hypothesis 1 (H₁)** proposes that these models, each with unique strengths ARIMA for trend modelling, SARIMA for capturing seasonality, and LSTM for detecting nonlinear temporal patterns can generate a reasonably accurate forecast for Tesla’s closing price on April 15, 2026. **Hypothesis 2 (H₀),** in contrast, assumes that reliable forecasting is not possible due to the highly speculative nature of Tesla’s stock. Factors such as sudden market sentiment shifts, innovation cycles, external economic shocks, and the absence of stable seasonality may violate the foundational assumptions of these models. As a result, forecasts may exhibit high uncertainty or fail to reflect actual market behaviour.

## General Objective:

This report aims to investigate how time series analysis can uncover hidden trends, seasonal patterns, and volatility in stock prices, enabling more informed predictions. Using Tesla Inc. (NASDAQ: TSLA) as a case study—a young and highly volatile company since its 2010 IPO my purpose is to demonstrate how advanced modelling techniques (ARIMA, SARIMA, and LSTM) can decode complex market behaviour and provide actionable insights for investors, even in erratic financial environments.

## Technical Objective

Data Exploration: Analyse Tesla’s historical stock price data (daily closing prices) to identify trends, seasonality, and volatility clusters. Model Application:

- ARIMA to capture non-seasonal trends and autocorrelations;

- Extend to SARIMA to account for seasonal effects (cycles).

- Implement LSTM (a deep learning approach) to handle nonlinear dependencies and long-term patterns.

Applying Validation Prediction helps to compare model performance using metrics RMSE and MAE.

# 4. Scope and Methodology

## Scope

The project focuses on**Tesla Inc. stock data**, covering**historical trends from the beginning of IPO 29th June 2010 until 15th April 2025.**While forecasting future price movement. The scope includes:

**Data Exploration**plays key role by Evaluating Tesla’s**historical stock price behaviour.**What component Data Exploration bears:

* **Feature Engineering.** Feature engineering is the process of transforming raw data into meaningful predictors (features) that improve model performance. In **time series forecasting,** it is especially crucial because **extracting hidden patterns** in stock prices like Tesla’s are influenced by multiple factors : trends, seasonality, external shocks. Feature engineering helps isolate these components for better predictions (Hyndman & Athanasopoulos, 2018).
* **Handling Non-Stationarity:** Most time series data, like stock prices, weather forecasting are **non-stationary, where** mean average could be variance and change over time.

Without proper feature engineering, even advanced models like LSTMs will fail to generalize in noisy financial datasets." – Goodfellow et al. (2016, Deep Learning).

## Methodology

This project follows the **CRISP-DM (Cross-Industry Standard Process for Data Mining)**framework to ensure a structured and repeatable data analysis pipeline, like implementation and evaluation. Additionally, the research aims to explore the relationship between Tesla’s stock price and other variables, including trading volume and broader market indices like the S&P 500. A key goal is also to understand why major institutional investors (“big fish”) consistently include Tesla in their portfolios and how this aligns with long-term strategies in the high-growth electric vehicle (EV) sector. Moreover, applying the CRISP-DM methodology through visualizations of trends and seasonality, along with the exploration of statistical properties, helps to uncover underlying structures and detect anomalies within the data.

Several steps were taken to ensure the dataset was clean and suitable for time series modelling:

* Date Alignment: When loading the CSV file, the trade dates were initially ordered from 2025 back to 2010. Using data preparation techniques, the 'Date' column was properly formatted and sorted in chronological order, ensuring compatibility with forecasting models (see Figure 1).
* Feature Cleaning: Irrelevant features such as 'Changes%', along with columns containing null values, were removed to improve model accuracy and reduce noise.

A screenshot of a graph

Description automatically generated

A screenshot of a computer

Description automatically generated

*Figure 1*

## Ethical considerations

This analysis used publicly available Tesla stock data, ensuring no privacy concerns. Models were rigorously validated to mitigate bias, with assumptions and limitations clearly documented for transparency. The study maintains neutrality—evaluating forecasting methods without endorsing Tesla’s financial decisions. All processes were designed for reproducibility, adhering to ethical standards for educational research in applied machine learning.

# 4. Project Management and Planning

## Data Source Overview

The project followed the CRISP-DM framework, with key milestones including data acquisition, EDA, model development, and evaluation. Public Tesla stock data (2010-2025) was sourced via Yahoo Finance API (yfinance) and pre-processed for analysis. Python tools (pandas, statsmodels, sklearn) ensured workflow efficiency, with full documentation for reproducibility.

## Data Exploratory Data Analysis (EDA)

To better understand the nature of Tesla’s stock over the long term (June 2010 – until present time April 2025), we examined basic descriptive statistical features, like:

* The dataset includes 3,722 trading days, covering Tesla’s entire post-IPO market behaviour.
* The average closing price is $86.05, while the median is only $18.54, suggesting a right-skewed distribution. This means that the stock spent much of its time at lower price levels, with large increases in recent years pulling the mean upward.
* The standard deviation of $111.88 confirms high volatility in price movements.
* Volume ranged from ~1.8 million to over 914 million shares per day, highlighting extreme variations in market attention and investor activity.

A screenshot of a graph

Description automatically generated

*Figure 2*

## Data Dictionary

Below is a data dictionary describing the key variables used throughout the analysis:

| Variable Name | Description |
| --- | --- |
| TSLA | Stock ticker symbol for Tesla Inc. Used to retrieve historical financial data from Yahoo Finance. |
| Close | The closing price of Tesla stock for each trading day. This was the primary target variable for forecasting. |
| df\_diff | The transformed dataset obtained after applying the first-order differencing to the Close variable. Used to ensure stationarity for models like ARIMA. |
| Date | The trading date corresponding to each record. Properly sorted in chronological order for time series modelling. |
| Volume | The number of shares traded on each day. Used for exploratory analysis, though not a target variable. |

# **5.** Modelling

## **Summary Statistics**

To understand the dataset structure, histograms were generated for key variables: Date, Close, Open, and High. The Date histogram shows a uniform distribution, confirming even time coverage from 2010 to 2025 and validating the dataset’s suitability for chronological modelling. The Close, Open, and High price variables are right-skewed, with most values clustered at the lower end and a long tail toward higher prices. This reflects Tesla’s real-world market behaviour—lower prices in early years, followed by a sharp increase in line with the company’s rapid growth. Such skewed, non-normal distributions are typical in financial time series and informed the use of transformations like differencing and scaling prior to applying models like ARIMA and LSTM. As Robert Nau (2024) notes, “Don’t worry too much about normality when using ARIMA models,” since these models focus on time-based dependencies rather than data distribution. This validates the approach taken in this project.

A screenshot of a computer screen

Description automatically generated

*Figure 3*

The histogram and density plot above illustrate the distribution of Tesla’s trading volume from 2010 to April 2025. The data is positively skewed, with most trading days concentrated in the lower volume range between 0 and 200 million shares. The highest frequency is observed around the 80–100 million range, after which the frequency gradually declines, forming a long right tail. This skewed distribution indicates that while high-volume trading days exist (often due to major news or market events), they are relatively rare. Most of the time, Tesla’s trading activity remains within a moderate volume range. This insight is crucial, as extreme volume days could disproportionately affect volatility and should be carefully considered in model training and interpretation. The presence of outliers and asymmetry justifies the use of techniques like scaling or log transformation as volume are to be used as a feature in predictive modelling. (Figure 4).

A graph of a number of people

Description automatically generated with medium confidence

*Figure 4*

## **Bias Distribution and Outlier Detection**

The boxplot provides a visual overview of the distribution and variability of key Tesla stock features: Close, Open, High, Low, and Volume. It effectively highlights **skewness and outliers,** which are crucial in identifying potential data bias.

**Price variables** (Close, Open, High, Low) show right-skewed distributions with outliers on the upper end—reflecting days of unusually high price movements due to news events, earnings, or investor sentiment. These are not errors but real market behaviours and were **retained to preserve volatility**.

**Volume**displays the most extreme outliers, representing trading spikes often aligned with announcements or speculative activity—key signals in market analysis.

These patterns confirm the **non-normal, high-variance nature** of financial data. Accordingly, **nonlinear models like LSTM** were prioritized, as they handle such irregularities more effectively than linear models.

As Mandelbrot and Hudson (2004) argue, **financial markets are inherently fractal and non-Gaussian,** challenging the assumptions of classical models like ARIMA, which depend on normality and stationarity.

## Pre-processing and Data Cleaning

In preparation for time series modelling, the dataset underwent several essential preprocessing steps to ensure consistency and usability:

* Data Type Inspection: Initial inspection revealed that some columns, including Date and Change %, were stored as object types. This is problematic for time-based modelling and numerical analysis, so appropriate transformations were applied.
* Column Renaming and Flattening: To streamline further analysis and model training, column names were cleaned by flattening any multi-index structures and removing trailing characters. This helped avoid errors in later processing stages.
* Datetime Conversion: The Date column was explicitly converted into a datetime format using pd.to\_datetime(). This is a critical step for any time series analysis, as it allows for proper indexing, resampling, and chronological ordering of data.
* Missing Data Check: A check for missing or improperly parsed dates was conducted using .isna(), and the result confirmed that no missing dates were present after conversion. This ensures that the time index is complete and the data can be safely used for modelling.
* Irrelevant Feature Handling: Features such as Change % were identified as non-numeric and not directly useful for modelling. These were flagged for removal or further transformation, depending on model requirements.

These cleaning steps ensured that the dataset was well-structured, free from parsing issues, and ready for accurate time series forecasting using models like ARIMA or LSTM.

## Trend Analysis Prior to Stationarity Testing

Following initial data cleaning and preprocessing, a time series plot of Tesla’s closing price was generated to visually inspect the overall trend and volatility from 2010 to 2025. This visualization serves as a critical intermediate step—providing an intuitive overview of the stock's behaviour before proceeding with formal statistical analysis, such as the ADF test and seasonal decomposition. (Figure 5).

A graph showing the price of a tesla

Description automatically generated

*Figure 5*

## Trend and Seasonality Analysis – Seasonal Decomposition of Tesla Closing Price for 365 days.

### To better understand the underlying structure of Tesla’s stock behaviour, a seasonal decomposition was applied to the Close price series using additive components. This technique splits the time series into **Trend, Seasonality**, and**Residuals**, offering insights into the nature and drivers of price movements over time.

* **Original Series:** The top panel shows the raw closing price, which exhibits long periods of stability followed by extreme growth and high volatility starting around **2019–2020**. This aligns with Tesla’s global expansion, increased investor interest, and broader adoption of electric vehicles.
* **Trend Component**: The trend line reveals a **long-term upward trajectory**, particularly steep between **2020 and 2022**, indicating a significant and sustained market valuation increase. The recent flattening and dip around 2023–2024 highlight the onset of market corrections or external macroeconomic effects.
* **Seasonality:** The seasonal component shows **repeating short-term patterns**, suggesting a cyclical effect in Tesla’s stock price. These could reflect quarterly earnings, investor cycles, or sector-based sentiment recurring annually.
* **Residuals**: The residuals component captures the **random noise** and unexplained variation not covered by trend or seasonality. The spikes in residuals after 2021 indicate periods of unexpected market reactions, possibly due to news events, regulatory changes, or global economic shocks.

This decomposition validates the presence of both **long-term growth trends** and **repeating seasonal effects**, justifying the use of time series models that can handle such structure—particularly **SARIMAX** and **LSTM**. (Figure 6)

*Figure 6*

A graph of different colored lines

Description automatically generated

## **Stationarity Testing and ARIMA Experimentation**

An initial experiment with the ARIMA model was conducted to forecast Tesla’s stock price for February 19, 2025. This test aimed to evaluate model performance on a known value before applying the same configuration to forecast April 15, 2026.

ARIMA assumes stationarity—constant mean and variance over time—so a **stationarity check**was essential. The **Augmented Dickey-Fuller (ADF) test** was used to assess this, where a p-value above 0.05 suggests non-stationarity. If confirmed, **differencing**is applied to stabilize the series.

Once stationarity was established, the model’s parameters (p, d, q) were determined and fitted to historical data. The validated configuration from February 2025 served as the foundation for the 2026 forecast.

This two-step experimentation helped confirm ARIMA’s structure suitability for Tesla’s data and established a **benchmark** for evaluating more advanced models like **SARIMA and LSTM,** which handle seasonality and nonlinearity more flexibly.

A close-up of a computer screen

Description automatically generated

*Figure 7*

## 6. Machine Learning Algorithm

## 6.1. Short-Term Seasonal Decomposition and Its Role in Modelling

As part of the modelling and evaluation process, a short-term seasonal decomposition was applied to Tesla’s closing price using an additive model with a 5-day period. This technique helps isolate the time series into three key components: **Trend**, **Seasonality**, and **Residuals**, providing a clearer view of what drives short-term fluctuations in the stock price (Hyndman and Athanasopoulos, 2021).

 Model Insight Before Forecasting:  
Decomposing the time series before fitting an ARIMA model offers a better understanding of recent patterns, especially when forecasting over short horizons. It visually confirms whether the most recent portion of the series follows a stable pattern or exhibits volatility that could impact forecast accuracy.

The reasons why perform short-term decomposition are:

 Focus on Weekly Patterns:  
A 5-day period was chosen because Tesla stock trades Monday to Friday, meaning this decomposition captures weekly cycles. These cycles might reflect investor behaviour (e.g., early-week rallies, end-of-week slowdowns), short-term earnings speculation, or news reaction lags.

 Residuals and Model Suitability:  
By analysing the residuals, we assess how much variability remains unexplained after accounting for trend and seasonality. If residuals show strong structure or patterns, it suggests that more complex models (like LSTM or SARIMAX with external variables) may be needed. If residuals are randomly distributed, simpler models like ARIMA can be considered sufficient.

A screenshot of a computer screen

Description automatically generated

*Figure 8*

## 6.2. Stationarity Testing and Hypothesis Interpretation /ADF Test

Before applying time series forecasting models like ARIMA, it is essential to assess whether the data is **stationary**, meaning it has a constant mean and variance over time. To evaluate this, the **Augmented Dickey-Fuller (ADF) test** was applied to the Tesla Close price series

(Figure 9). The ADF test is a statistical hypothesis test that determines whether a unit root is present in the time series.

The test follows this structure:

* **Null Hypothesis (H₀)**: The series has a unit root (non-stationary).
* **Alternative Hypothesis (H₁)**: The series is stationary.

The results of the ADF test showed:

* **ADF Statistic** = -1.407
* **p-value** = 0.5788

Since the p-value is **greater than 0.05,** we fail to reject the null hypothesis. This indicates that the Tesla stock price series is **non-stationary**, and therefore, differencing is required before applying the ARIMA model. This transformation helps stabilize the mean of the series and ensures the model assumptions are met for accurate forecasting.

This test is a foundational step, comparable to checking residuals or variable significance in linear regression models, and ensures the reliability of downstream time series predictions.

*Figure 9*

**A screenshot of a computer

Description automatically generated**

## Stationarity Confirmation After Differencing

After the initial Augmented Dickey-Fuller (ADF) test indicated that the original Tesla closing price series was non-stationary**, first-order differencing** was applied to the data to stabilize its mean and remove trends. This transformation helps meet a core assumption of ARIMA and SARIMA models**: stationarity (Figure 10).**

To verify the effect of differencing, the ADF test was re-applied to the transformed series (df\_diff['Close']). The test returned the following results:

* **ADF Statistic**: -11.07
* **p-value**: 4.61 × 10⁻²⁰

Since the p-value is **significantly below the 0.05 threshold**, we can **reject the null hypothesis** and conclude that the differenced series is now **stationary**. This confirms that the series is suitable for use in ARIMA and SARIMA modelling.

A white rectangular object with a white border

Description automatically generated

*Figure 10*

## ACF and PACF Analysis for ARIMA Model Selection

After confirming stationarity using the Augmented Dickey-Fuller test, the next step in developing the ARIMA model was to identify suitable values for the autoregressive (p) and moving average (q) components. This was done by generating Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots based on the differenced Tesla closing price series.

* The **ACF plot** (left, *Figure 11*) shows how the time series is correlated with its own lagged values. It is used to determine the **q parameter** in ARIMA (the number of lagged forecast errors included in the model).
* The **PACF plot** (right, *Figure 11* ) isolates the direct relationship between the series and its lagged versions, excluding indirect effects from intermediate lags. It is used to determine the **p parameter** (the number of lagged observations to include as predictors).

From the plots:

* The ACF shows a significant spike at lag 1 and quickly drops off, suggesting **q = 1.**
* The PACF also shows a clear drop after lag 1, indicating **p = 1**.

These observations support using an **ARIMA(1,1,0)** model, where d = 1 comes from the differencing step. Performing this analysis ensures the model is well-calibrated to the underlying structure of the data and not overfitted, leading to more accurate and stable forecasts.

A screenshot of a computer screen

Description automatically generated

*Figure 11*

## ARIMA Modelling and Forecast for 2025-02-19

To test ARIMA’s suitability for long-term forecasting, a validation was performed by predicting Tesla’s closing price for February 19, 2025. The Augmented Dickey-Fuller (ADF) test initially indicated non-stationarity (p > 0.05), so first-order differencing was applied. The differenced series passed the ADF test (p < 0.05), confirming stationarity.

Based on ACF and PACF plots, the ARIMA(1,1,0) model was selected and trained on data up to April 15, 2025. The in-sample forecast for February 19, 2025, was approximately $354.14, closely matching the actual price and indicating strong model performance.

This result suggests that ARIMA(1,1,0) effectively captures Tesla’s short- to medium-term temporal dynamics.

*Figure 12*

A screenshot of a computer program

Description automatically generated

To identify the most appropriate ARIMA model configuration for forecasting Tesla's stock price, it was employed **Akaike Information Criterion (AIC)**-based model selection (Figure 12). AIC is a widely accepted metric that balances model fit and complexity, helping to avoid overfitting. Lower AIC values indicate better model performance. With implemented a **grid search** over combinations of ARIMA model parameters:

* p (autoregressive term): 0 to 2,
* d (differencing): 0 to 1,
* q (moving average term): 0 to 2.

This nested loop evaluated each combination's AIC score using the differenced series (df\_diff['Close']), and stored the result. The model with the lowest AIC score was considered optimal. After evaluating all combinations, the **ARIMA(1,1,0)** model yielded the **smallest AIC** (23330.70), which makes it the best choice based on this criterion. This model includes:

* One lag of the autoregressive (AR) component,
* First-order differencing to achieve stationarity,
* No moving average (MA) component.

This suggests that the series is well explained by the first lag of past values, and the differencing step is crucial to eliminate non-stationarity. The absence of a moving average term simplifies the model while still maintaining forecasting accuracy, as validated in the short-term forecast to 2025-02-19.

According to Hyndman and Athanasopoulos (2021), selecting the optimal ARIMA model using the lowest Akaike Information Criterion (AIC) ensures a balance between goodness of fit and model simplicity. In this project, ARIMA(1,1,0) provided the lowest AIC score among tested combinations, making it the most statistically suitable configuration.

A graph showing a line

Description automatically generated with medium confidence

*Figure 13*

After identifying ARIMA(1,1,0) as the optimal configuration using the lowest AIC value, the model was trained on the entire historical Tesla dataset up to April 15, 2025. A one-step-ahead forecast was generated for February 19, 2025, yielding a predicted closing price of $354.14. This forecast was then compared with the actual closing price for that date.

Based on visual inspection (see Figure 13), the predicted value closely aligns with the actual price, with only a minor deviation, indicating high accuracy. The closeness of the forecast to the real value demonstrates that the ARIMA(1,1,0) model effectively captured the trend and short-term dynamics in Tesla’s price data.

Such accuracy is especially valuable for financial analysts and investors, as even small errors in short-term forecasts can affect trading and portfolio strategies. According to Hyndman and Athanasopoulos (2021), well-tuned ARIMA models can provide reliable short-term forecasts when data is properly pre-processed and stationarity is ensured.

## SARIMA Forecasting – Experimental Prediction for February 19, 2025

To evaluate SARIMA’s ability to model and forecast Tesla’s stock price, an experimental prediction was carried out for February 19, 2025, using the SARIMA(1,1,0)(1,1,0,5) model. This configuration includes a 5-day seasonal cycle, aligned with the typical weekly trading schedule.

The model was chosen based on seasonal decomposition results and prior evidence suggesting potential weekly patterns in stock market behaviour. Model fitting yielded key statistical outputs used to validate its forecasting capability:

* AIC: 24940.844
* BIC: 24959.505
* Log Likelihood: -12467.422

The model coefficients were statistically significant (p < 0.001), indicating a strong relationship between lagged terms and current price in both autoregressive and seasonal components. The forecasted price for February 19, 2025, was $345.77, closely matching the actual price of $354.14, with a forecast error of just 2.4%. This result highlights the model’s strength in short-term prediction. As noted by Box et al. (2015), SARIMA models are effective for seasonal time series, particularly when short-term cycles like weekly patterns are present.

Given this successful validation, the same SARIMA configuration was extended to forecast Tesla’s closing price for April 15, 2026, one year beyond the last observed data point.

A graph with blue line

Description automatically generated*Figure 14*

## LSTM Forecasting – Experimental Prediction for February 19, 2025

To evaluate the performance of deep learning techniques in forecasting Tesla’s stock price, a Long Short-Term Memory (LSTM) neural network was trained using the **90 previous trading days**before the target date, **February 19, 2025.** LSTM networks are particularly well-suited for time series forecasting because they can retain long-term dependencies and handle sequential data with complex, non-linear patterns.

The model was trained on scaled values of the stock's closing prices and validated on a one-step forecast. The **predicted closing price for 2025-02-19 was $354.17**, while the**actual price was $360.56,** yielding an error of just **1.17%**, demonstrating a **high degree of accuracy.**

A graph with orange and blue lines

Description automatically generated

*Figure 15*

The figure above shows how closely the LSTM’s prediction (red dot) aligned with the actual price (orange dot), even in a period of relatively high volatility. The input history used (blue line) and the forecast sequence (orange dashed line) highlight the model's ability to learn underlying trends.

According to Brownlee (2017), LSTM models are highly effective for financial time series because they are capable of modelling long memory processes and are not restricted by linear assumptions like classical models such as ARIMA.

Encouraged by this result, the same LSTM architecture can be extended to **forecast Tesla’s stock price for April 15, 2026** by feeding the most recent 90-day input sequence before that date. This allows a consistent framework to be applied for both validation and long-term forecasting.

# 8. Modelling and Forecasting for April 15, 2026

## ARIMA Forecast for 2026-04-15

Following the successful validation of the ARIMA(1,1,0) model on February 19, 2025, the same configuration was used to forecast Tesla’s closing price for April 15, 2026. This decision was based on the model’s strong alignment with the actual 2025 value, indicating its suitability for short-to-medium-term forecasting.

To prepare for the 2026 forecast, the model was retrained on the full dataset from June 2010 to April 15, 2025. Using the .get\_forecast() method, a 365-day projection was generated. As ARIMA operates on differenced data, the forecast was converted back to actual price values by applying a cumulative sum and adding the last known closing price.

This process resulted in a forecasted closing price of $430.41 for April 15, 2026. The result illustrates ARIMA’s strength in extrapolating from recent trends and offers a reasonable baseline for long-term planning.

By validating the model before full deployment, this approach aligns with best practices in financial forecasting, reinforcing confidence in the use of ARIMA for future investment decisions.

A graph showing a blue line

Description automatically generated

*Figure 16*

## SARIMA Forecast for 2026-04-15

To forecast Tesla’s stock price for April 15, 2026, the SARIMA(1,1,0)(1,1,0,5) model was applied. This configuration was selected based on its earlier performance when validated on February 19, 2025. A 5-day seasonal cycle was used to reflect weekly trading activity.

The model was trained on Tesla’s data from 2010 to April 15, 2025, and used to predict the next 365 days. Forecasts were generated using model\_sarima.get\_forecast, and 95% confidence intervals were added via forecast\_obj.conf\_int() to visualize the uncertainty range.

The model predicted a closing price of $189.60 for April 15, 2026. This is a notable drop from the 2025 value (~$354), but it falls within the model’s wide confidence interval, reflecting the forecast’s high uncertainty.

SARIMA’s linear and seasonal structure suits medium-term forecasting but struggles with long-term predictions, especially when data includes structural changes. For strategic investment planning, this result offers a baseline but should be combined with more adaptive models like LSTM or SARIMAX.

For investors, this forecast is not a fixed outcome but part of a broader scenario, highlighting potential volatility typical of high-growth tech stocks.

**A graph of a graph

Description automatically generated with medium confidence***Figure 17*

## LSTM Forecast for 2026-04-15

To evaluate long-term forecasting using deep learning, an LSTM (Long Short-Term Memory) model was implemented to predict Tesla’s stock price on **April 15, 2026.** Known for its ability to capture **temporal dependencies and nonlinear patterns**, LSTM is well-suited for financial time series (Hewamalage et al., 2021).

The model was trained on **90-day input sequences**, each predicting the following day's closing price. Data was scaled using **MinMaxScaler**and reshaped for LSTM’s required format. The architecture included **two LSTM layers (64 units)** and a dense output, trained over **50 epochs**with a batch size of 32 using the **Adam optimizer**and **MSE loss.**

The model forecasted a price of approximately **$257.51** for April 15, 2026. Visual analysis showed that the LSTM tracked historical trends well, confirming its strength in learning from sequence data. However, it remains sensitive to scaling and lacks interpretability compared to models like ARIMA.

For financial analysts, LSTM offers valuable **data-driven insights,** but given its “black-box” nature, forecasts should be used in combination with **macroeconomic and sector-specific context** to support investment decisions.

# **9. Evaluation and Comparison of Models for forecasting for 2026-04-15**

## **ARIMA, SARIMA, and LSTM Metrics**

To assess the forecasting performance of the three selected models—ARIMA (1,1,0), SARIMA (1,1,0)(1,1,0,5), and LSTM—we applied a set of standard evaluation metrics: **R², MAE, MAPE,** and **RMSE.** These metrics provide complementary insights into the accuracy and robustness of each model:

* **R² (coefficient of determination)** quantifies how well the predictions explain the variance of the actual data (closer to 1 is better).
* **MAE (mean absolute error)** reflects the average magnitude of prediction errors.
* **MAPE (mean absolute percentage error)** shows the average error as a percentage of actual values—values above 100% indicate poor accuracy.
* **RMSE (root mean squared error)** penalizes larger errors more strongly than MAE, offering a sensitive measure of forecast precision.

The following table and visualizations summarize the results for each model:

**A screenshot of a graph

Description automatically generated**

Figure 18

While all models produced negative R² values, indicating they underperformed compared to a simple mean predictor, LSTM consistently delivered better results than ARIMA and SARIMA across other metrics. It achieved the lowest MAPE and RMSE, suggesting a closer fit to actual price movements. In general, lower values of MAE, MAPE, and RMSE indicate fewer prediction errors and greater model accuracy.

ARIMA’s results, including a 100% MAPE and negative R², highlight its limitations in dealing with the volatility and structural changes typical of financial time series.

SARIMA performed the worst among the tested models, primarily due to a mismatch between its imposed 5-day seasonal cycle and the actual behavior of Tesla’s stock, which is highly volatile, non-stationary, and lacks regular short-term patterns.

Short seasonal cycles (e.g., weekly) can improve forecasts only when such seasonality truly exists. As Hyndman and Athanasopoulos (2018) note, “choosing the correct seasonal period is crucial; overfitting with short, artificial cycles may lead to explosive errors in absence of real periodicity.”

Speculative stocks like Tesla, driven by sentiment, innovation, and external shocks, rarely exhibit stable cycles. Tsay (2010) emphasizes that short-term seasonal effects are more relevant to low-volatility assets, not those prone to structural breaks and irregular movements.

In this case, SARIMA’s rigid structure led to overfitting and poor results—with RMSE of 22.41 and MAPE exceeding 1100%.

By contrast, LSTM, although not perfect, adapted better to Tesla’s volatility, achieving lower error metrics and more closely following observed trends.

# 10. Conclusion

This project explored how time series models—ARIMA, SARIMA, and LSTM—can be applied to real financial data, using Tesla’s stock as a case study. Beyond forecasting, the goal was to understand how **seasonality, cycles, and volatility** affect model accuracy.

The experiments revealed that **traditional models like ARIMA and SARIMA** struggle with highly volatile assets like Tesla due to assumptions of linearity and consistent seasonal patterns. As Ghosh and Chaudhuri (2015) noted, short-term seasonality often fails under **speculative or high-frequency market conditions**—a finding echoed in SARIMA’s poor performance using a 5-day cycle.

In contrast, the **LSTM model**, while not perfect, handled nonlinear dynamics better, outperforming in terms of **MAPE and RMSE.** However, its **black-box nature** limits transparency, which can be a concern in professional financial decision-making.

Overall, the project highlights that model choice must be **context-driven**. Even advanced tools can fail if misapplied. Effective forecasting requires not just model flexibility but also a deep understanding of **data structure and behavioural patterns**.

This could serve for future research should focus on identifying external influences—such as**structural breaks or technical indicators**—that affect model performance, forming a basis for improved **regime detection and adaptive forecasting**.

# References:

1. Thomas H. Davenport & Jeanne G. Harris, 2007. *Competing on Analytics: The New Science of Winning.*  Publisher: Harvard Business Review Press.

ISBN-13: 978-1422103326;

2. Robert Nau, Duke University, Fuqua School of Business., Available: <https://people.duke.edu/~rnau/411home.htm>;

3. Bloomberg.com, *Tesla Signals Caution*, Revisits 2025, Growth Outlook, April 22, 2025.

4. Reuters.com, *Tesla Doubles Down on Robotaxi Timeline*, April 24, 2025;

5. Business Insider, *Tesla's Profit Falls Sharply Amid EV Market Saturation*, April 2025;

6. Mandelbrot, B. and Hudson, R.L., 2004. The (Mis) Behaviour of Markets: A Fractal View of Risk, Ruin, and Reward. New York: Basic Books;

7. Fischer, T. and Krauss, C., 2018. Deep learning with long short-term memory networks for financial market predictions. European Journal of Operational Research, 270(2), pp.654–669;

8. Hyndman, R.J. and Athanasopoulos, G., 2021. *Forecasting: Principles and Practice*. 3rd ed. Melbourne, Australia: OTexts. Available at: https://otexts.com/fpp3/;

9. Hyndman, R.J. and Athanasopoulos, G., 2021. *Forecasting: Principles and Practice*. 3rd ed. Melbourne: OTexts. Available at: https://otexts.com/fpp3/;

10. Hyndman, R.J. and Athanasopoulos, G., 2021. Forecasting: Principles and Practice. 3rd ed. Melbourne: OTexts. Available at: <https://otexts.com/fpp3/>;

11. Box, G.E.P., Jenkins, G.M., Reinsel, G.C. and Ljung, G.M., 2015. *Time Series Analysis: Forecasting and Control.* 5th ed. Hoboken, NJ: Wiley;

12. Brownlee, J., 2017. *Deep Learning for Time Series Forecasting: Predict the Future with MLPs, CNNs and LSTMs in Python*. 1st ed. Machine Learning Mastery;

13. Hyndman, R.J. & Athanasopoulos, G. (2018). *Forecasting: Principles and Practice* (2nd ed.). OTexts. <https://otexts.com/fpp2/>.