

# **Apple Stock Prediction**

**By: Grant Chiu Kwong Yeung**

# Introduction

The Stock Market, also known as the equity market, is a crucial aspect of modern finance and economics. It refers to a collection of exchanges, platforms and over-the-counter (OTC) marketplaces where stocks or shares of publicly held companies are bought and sold. The stock market provides a crucial source of funding for companies to grow and expand their operations. It also offers investors an opportunity to invest in companies they believe will grow and generate returns.

Formal exchanges, such as the New York Stock Exchange (NYSE) and the NASDAQ, are regulated by the government and follow strict rules and procedures for trading. The OTC market, on the other hand, operates in a decentralized manner, where transactions take place directly between buyers and sellers without intermediaries.

The primary objective of this data science project is to create a predictive model that can identify the trend of Apple Inc. stock price over time. The project will make use of time-series analysis, which involves studying historical data to understand patterns and make predictions about future trends. The end goal is to provide valuable insights and inform investment decisions.

The stock market plays a crucial role in the global economy, providing a platform for companies to raise capital and for investors to participate in the growth of these companies. Understanding the trend of stock prices, such as that of Apple, is crucial for making informed investment decisions and this data science project aims to contribute to this goal.

# Data

The dataset used for this analysis was obtained from Yahoo Finance, a trusted and reputable source for daily stock market statistics. The data encompasses a ten-year period, from 2013 to 2022. The following features are included in the dataset: closing price and adjusted closing price. The closing price represents the cost of shares at the end of the trading day, while the adjusted closing price takes into consideration various factors such as dividends, stock splits, and new stock offerings. As a result, the adjusted closing price is considered a more precise measure of the stock's value, given that it builds upon the closing price.

- High - The highest value of the stock achieved during the day's trading
- Low - The lowest value of the stock achieved during the day's trading
- Open - The value of the stock at the opening of the market for that day
- Close - The value of the stock at the closing of the market for that day
- Volume - The number of shares that were traded for that day
- Adj Close - The stock price of a publicly traded company after accounting for any corporate actions such as dividends, stock splits, and new stock offerings.

It is worth noting that the use of adjusted closing price in this analysis is crucial for providing a more accurate representation of the stock's value over time.

The raw dataset obtained from Yahoo Finance consists of 2518 rows and 7 variables. While the dataset was usable in its raw form, it was necessary to pre-process the data to reduce the model size without compromising its accuracy.

The first step was to check for missing values in the dataset. Once the missing values were addressed, the next crucial step was to convert the 'Date' column into a datetime object, which is essential for time-series analysis. This conversion enables the analysis to effectively

consider the time component and facilitate a more comprehensive analysis of the stock price trends over time. Additionally, the dataset was resampled to provide a more comprehensive understanding of the stock price trends over time. The resampling process aggregates the data into specified time intervals, such as weekly or monthly. This enables a more comprehensive analysis of the long-term trend of the stock price and eliminates the noise associated with daily fluctuations. Furthermore, to address any missing values within the resampled dataset, linear interpolation was performed. The 'time' method was chosen as the interpolation method due to its capability to consider the time component and generate predictions based on the trends observed in the historical data. This helps to ensure that the predictions are as accurate as possible and are in line with the trends observed in the data.

## Exploratory Data Analysis

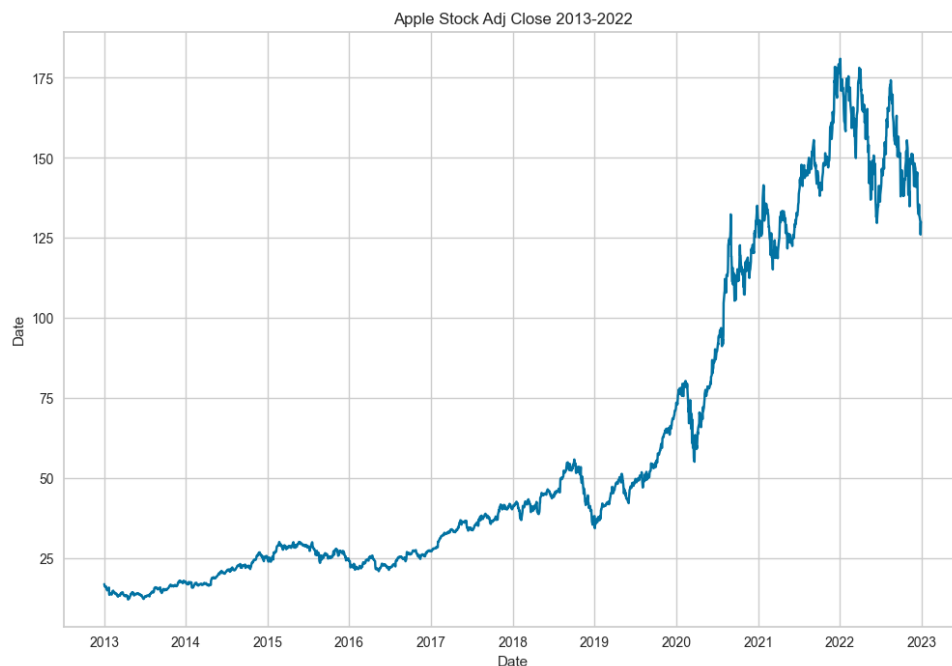


Figure 1- Apple Stock Adjusted Close Price

Apple stock price has experienced significant fluctuations from 2013 to 2022. In 2013, the stock price started at around \$55 and increased gradually to reach a peak of \$134 in February 2015. After that, the price experienced a downward trend and dropped to around \$90 by May 2016. From 2016 to 2018, the stock price showed some signs of recovery and reached a new high of \$233 in October 2018. However, the price then dropped again, reaching a low of \$140 in January 2019. From 2019 to 2020, the stock price was relatively stable, with a minor increase, reaching \$137 in September 2020. In 2021, the stock price experienced a significant increase, reaching an all-time high of \$157 in August 2021, due to the COVID-19 pandemic driving demand for technology products and services. Apple stock price has shown both upward and downward trends from 2013 to 2022, but overall has had a positive trend.

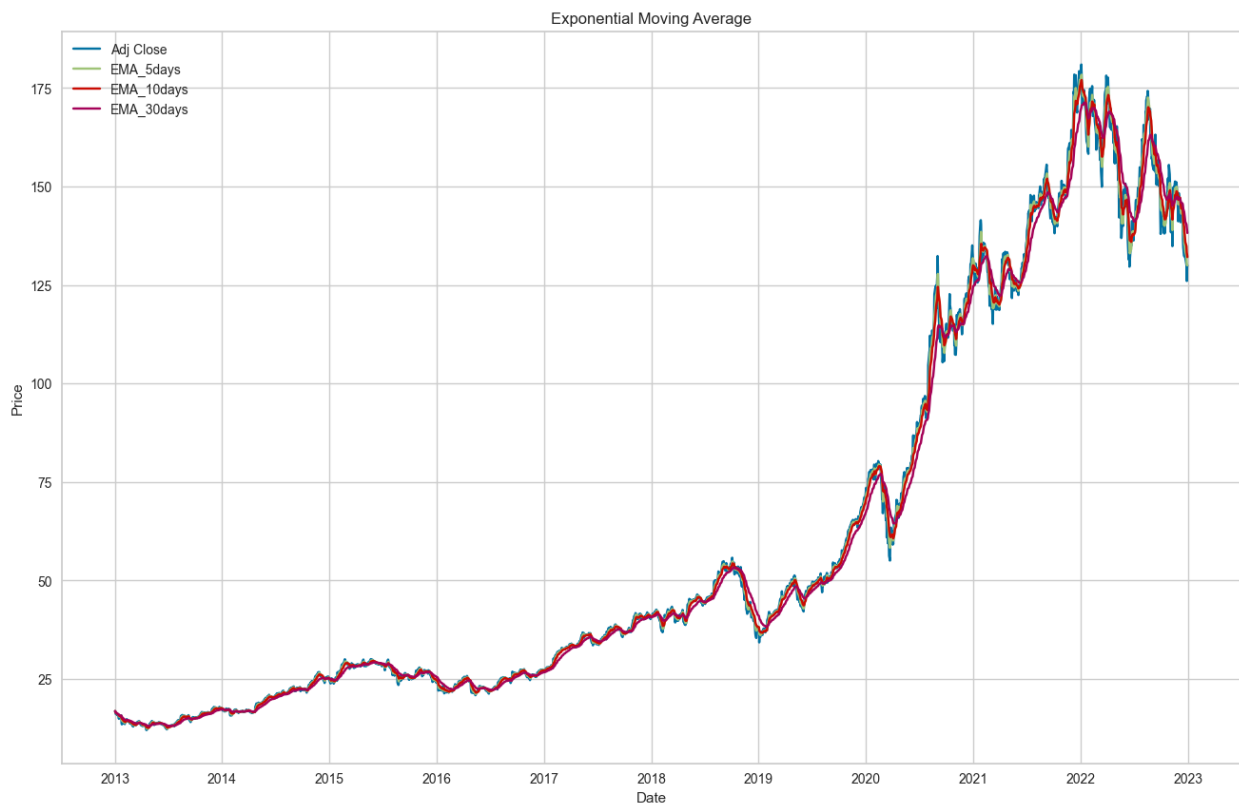


Figure 2 - Moving Averages for Apple Stock

Moving averages are important in stock price analysis as they help to identify trends, smooth data, generate signals, and confirm trend strength, providing valuable information for investment decisions. In general, during upward trends, the 5-day and 10-day moving averages would be above the 30-day moving average, indicating a short-term bullish sentiment. During downward trends, the 5-day and 10-day moving averages would be below the 30-day moving average, indicating a short-term bearish sentiment. During periods of stability or consolidation, the moving averages would be relatively flat and close to each other, indicating a lack of clear direction for the stock price. It is important to note that the moving averages can be subject to change based on the stock price fluctuations, so they should be monitored regularly and used in conjunction with other technical analysis tools to make informed investment decisions. This is shown in Figure 2.

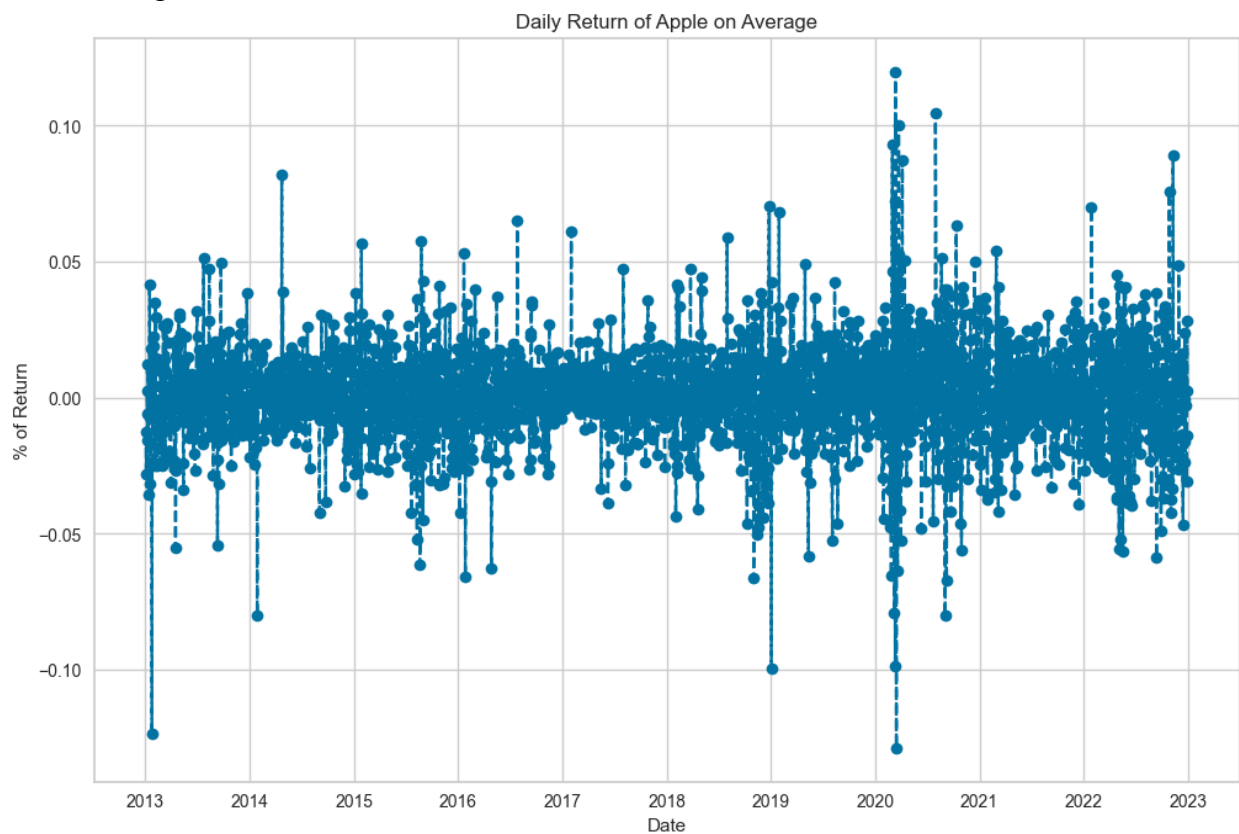


Figure 3- Daily Return of Apple Stock on Average

Looking at daily returns for a stock price can provide valuable information for investors and traders. Daily returns help measure the volatility of a stock and its ability to generate consistent returns. They can also be used to evaluate the performance of a stock compared to the overall market or to its peers in the same industry. Daily returns also play a crucial role in risk management and in making investment decisions in the short term. Observing daily returns can help identify trends in a stock's price movement, which can be useful in making investment decisions. However, it's important to note that daily returns may not provide a complete picture of a stock's performance and should be considered in the context of its longer-term performance and overall market conditions.

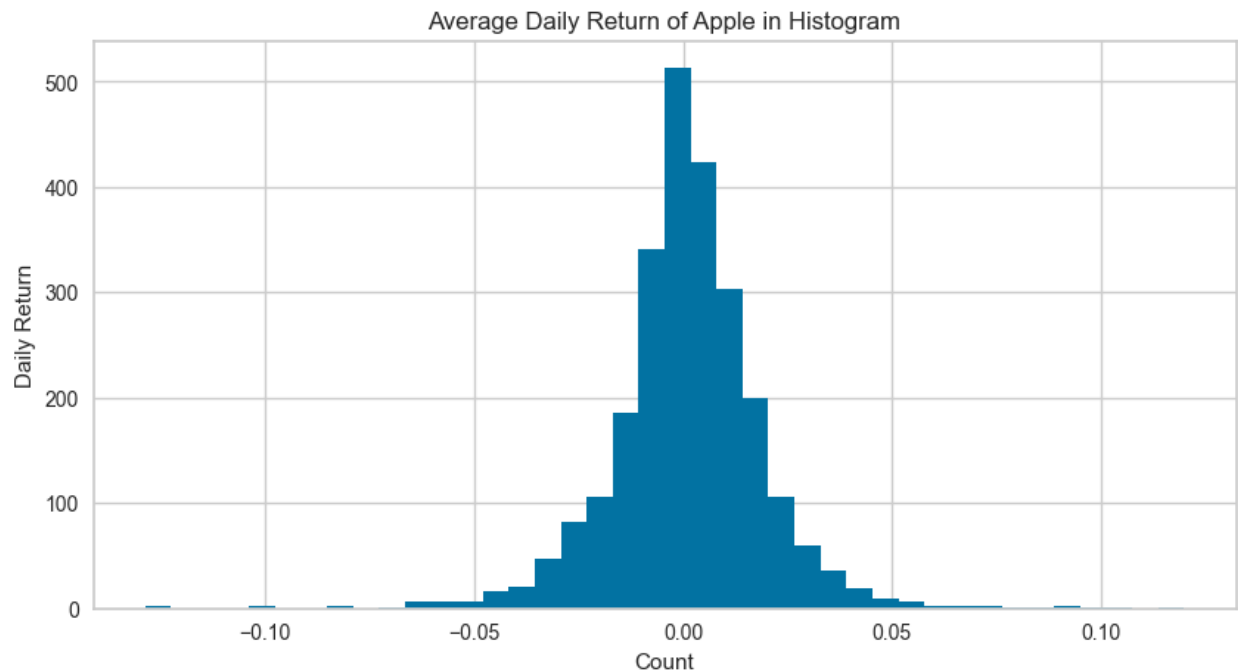


Figure 4 - Daily Return of Apple Stock on Average in Histogram

As shown in Figure 4, the histogram presents a visual representation of the distribution of daily returns for a specific stock over a certain period of time. The histogram in this case demonstrates a relatively symmetrical distribution, albeit with a slight tilt towards positive returns. This suggests that, over the long term, investors can expect a higher probability of experiencing

positive returns compared to negative ones. It is important to keep in mind, however, that past performance is not indicative of future results and that the stock market can be inherently volatile in the short term.

## Modeling

	Model	MASE	RMSSE	MAE	RMSE	MAPE	SMAPE	R2	TT (Sec)
dt_cds_dt	Decision Tree w/ Cond. Deseasonalize & Detrending	6.2534	5.9740	11.4229	14.7903	0.1390	0.1555	-0.2256	0.0600
gbr_cds_dt	Gradient Boosting w/ Cond. Deseasonalize & Detrending	6.4983	6.2770	11.8629	15.5669	0.1446	0.1638	-0.3720	0.0667
ets	ETS	7.0735	6.4980	12.7916	16.1037	0.1665	0.1887	-0.5984	0.0200

Figure 5 - Top 3 Models with lowest RMSE

In my analysis, I utilized PyCaret for the training of my time series model. I evaluated three different models, including Decision Tree, Gradient Boosting, and a blended model consisting of the top three models with the lowest Root Mean Squared Error (RMSE) score. The blended model approach was adopted with the intention of improving the RMSE score, as it combines the strengths of multiple models to make more accurate predictions.



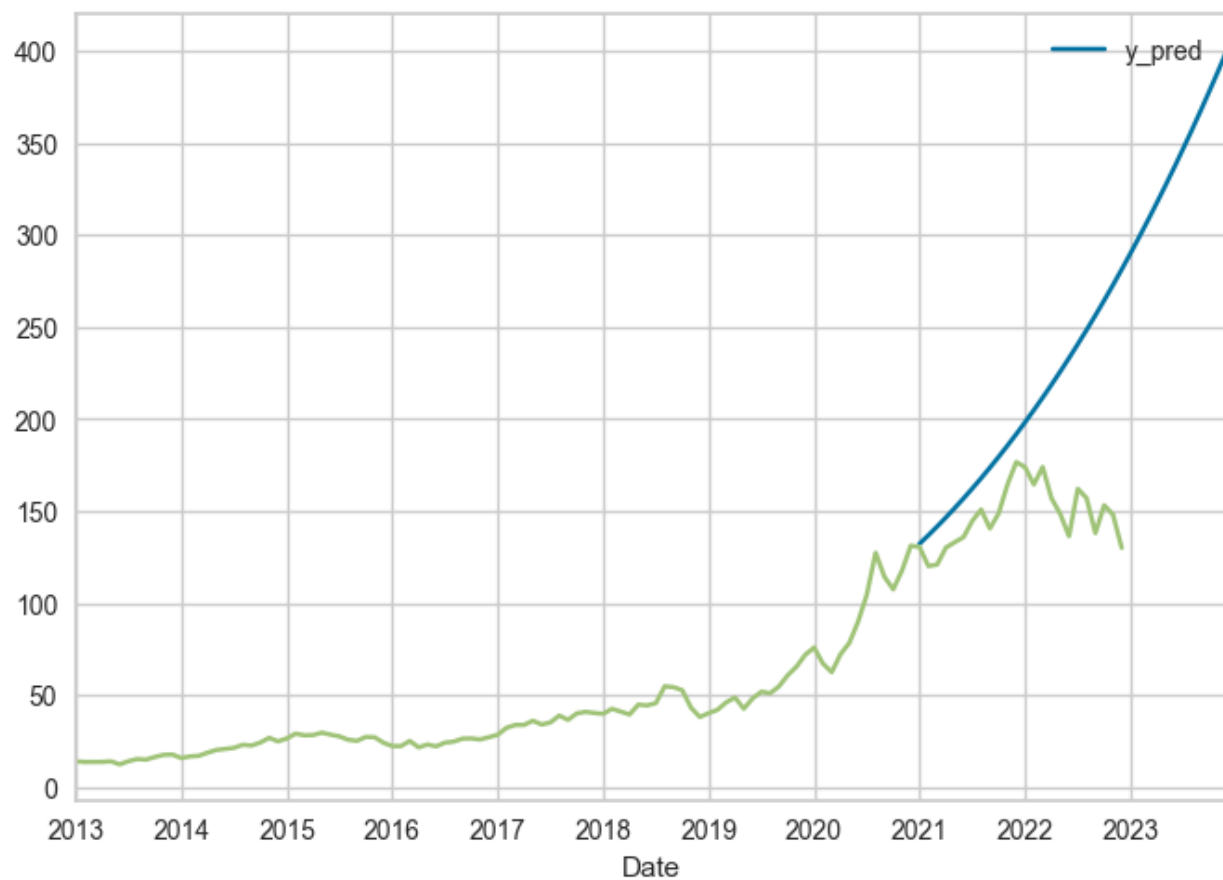


Figure 6 - Apple Stock price prediction Decision Tree Model (RMSE 92.23)

For the first model in Figure 6, we utilized a Decision Tree model with conditional deseasonalization and detrending. Our evaluation revealed that the model predicts a generally upward trend for the stock, but with a significant mismatch in relation to the actual price movements of Apple's stock. Thus far, the model has not proven to be effective, as the stock has experienced rapid gains since 2022. This highlights the limitations of the model and the need for further refinement or exploration of alternative models.

We used the Root Mean Squared Error (RMSE) as a metric to assess the effectiveness of our models. Among the models evaluated, the Decision Tree model performed the worst, with an RMSE score of 92.23. The use of RMSE as a performance metric allows us to quantify the

difference between the predicted values and the actual values, providing us with a clearer picture of the model's accuracy and effectiveness.

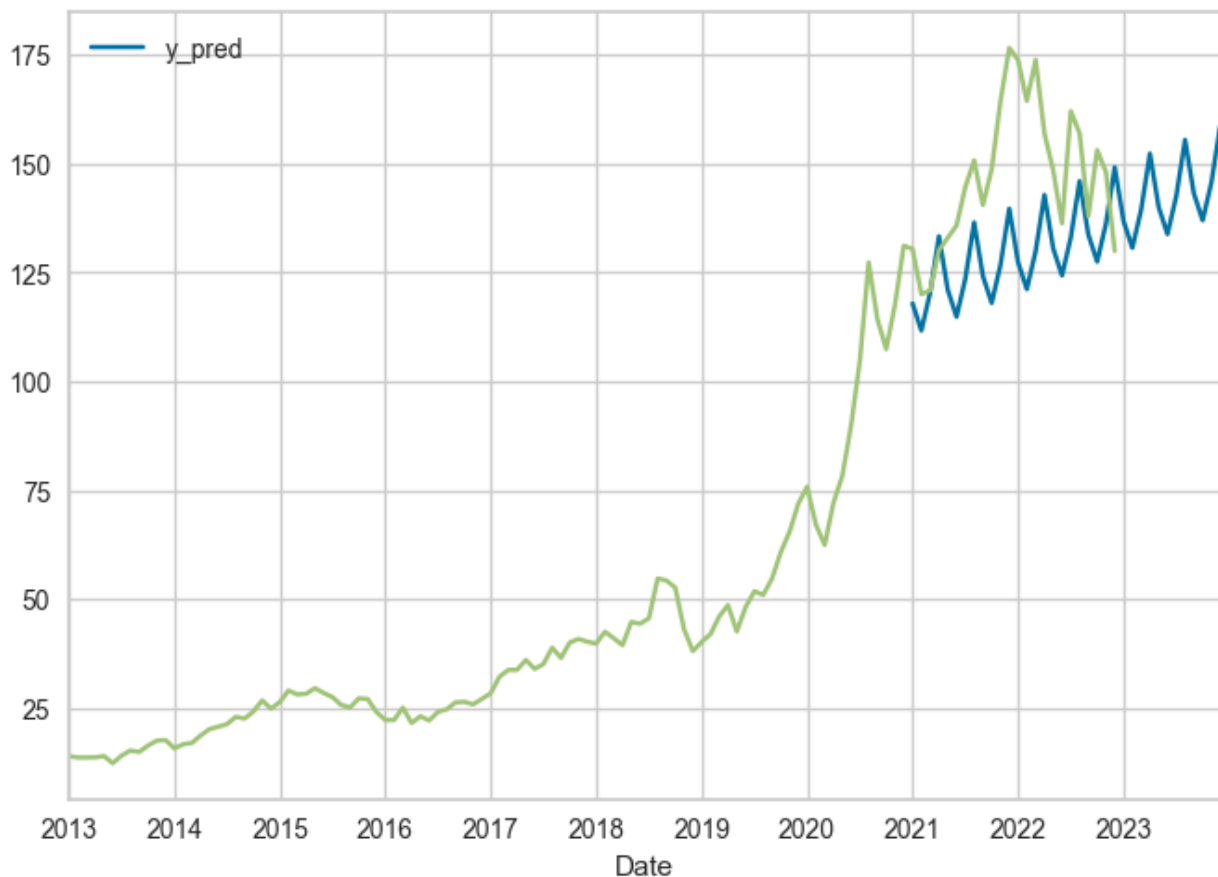


Figure 7 - Apple Stock price prediction Gradient Boosting Model (RMSE 27.02)

While the second model does display an improved prediction of the stock trend, it is important to approach these findings with caution. The RMSE score of 27.02 indicates a reduced discrepancy between the predicted values and the actual values, but it is not a guarantee of accuracy. There are many factors that can impact the stock market, including economic conditions, company performance, and investor sentiment. As such, it is not possible to say with confidence that the forecast generated by this model will remain accurate over time. It is important to continuously monitor and evaluate the performance of the model and consider additional factors that may impact the stock market.



Figure 7 - Apple Stock price prediction Blend Model (RMSE 34.57)

In the case of the blended model, we observed a slight upward trend that does not align well with the actual trend of the stock. The blended model approach can be beneficial as it combines the strengths of multiple models, potentially leading to a more accurate prediction. However, in this case, the blended model is not performing as well as expected, scoring a 34.57 RMSE. To improve the performance of the blended model, it may be beneficial to consider blending different models. By exploring alternative models and considering their unique strengths and weaknesses, it may be possible to create a more accurate blend that better predicts the trend of the stock.

# Conclusion

In conclusion, the time series analysis of Apple stock prices between 2013 and 2022 was performed using PyCaret and three machine learning models: Decision Tree, Gradient Boosting, and Blend Models. The Gradient Boosting model demonstrated the best performance with an RMSE score of 27.02, making it an effective tool for predicting the stock prices of Apple.

However, it is important to note that the results of this study should not be used as the sole basis for investment decisions. There are numerous factors that can impact stock prices, including economic conditions, company performance, and investor sentiment, and a machine learning model cannot account for all of them. Additionally, stock prices can be subject to sudden and significant fluctuations, making prediction an inherently uncertain process. As such, the results of this study should be used only as one of many sources of information when making investment decisions. This study highlights the importance of machine learning models in time series analysis and the potential for further improvement in stock price prediction, but it should not be relied upon as the sole basis for investment decisions.

In order to improve the project and make the predictions even more accurate, it would be beneficial to implement a news API. The integration of a news API in the time series analysis project can enhance the predictions by leveraging the power of Natural Language Processing (NLP) and machine learning. The API would allow for a systematic search and analysis of news articles relevant to the stock market and the company in question. This would provide a deeper understanding of what is being said about the company and its impact on the stock market. By incorporating a news API, the time series analysis would be able to identify any positive or negative news about the company in real-time and respond accordingly. Positive news can be

leveraged for promotional purposes, while negative news can be addressed proactively to prevent potential PR issues. This integration would provide a more comprehensive understanding of the stock market and the factors that drive stock prices, leading to improved predictions and valuable insights for investors. The implementation of a news API would optimize the time series analysis to deliver real-time, actionable information for better decision making.