Jiteshwar Anjale
DSC680, Weeks 1-4
23-September-2023

Project 1: Milestone 3

Topic:

Effective use of data science methodologies in financial advisory companies so that it can efficiently advise their customers and manage funds with more accuracy.

Business Problem:

Everyone wants to know if they're getting a great deal and what to expect from their investments in the future. If you could know exactly how much your stock investment would be worth in a year, two years, or 10 years, you'd want to know, wouldn't you?

The stock market or the price of individual stocks is traditionally based on some fundamentals that an investor or investment advisor evaluates to make a judgment as to whether to invest in those stocks. Market behavior is affected by other factors too such as geopolitical events, natural disasters, and the like. Unfortunately, there is no reliable way to predict the future expected stock price for a company. This is a challenge to the investors and their consultants.

Background/History

It is hard to predict the performance of a company stock or Fund based on company's traditional fundamentals. Fundamentals are nothing but metrics such as cash flow, return on assets, debt, R&D investment, earning per share, company estimates, their "Generally Acceptable Accounting Practices", and many more. However, stocks have shown themselves to have healthy books, but have not traded well in the market and their stock price fluctuation doesn't reflect the health of the company. Often time, other stimuli affect the stock price. Most of the stimuli are not predictable but the variation in the stock price over time may reveal more insight into the stock behavior.

Commonly used technical indicators:
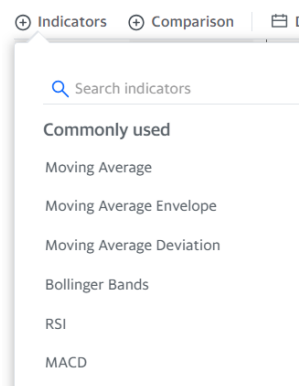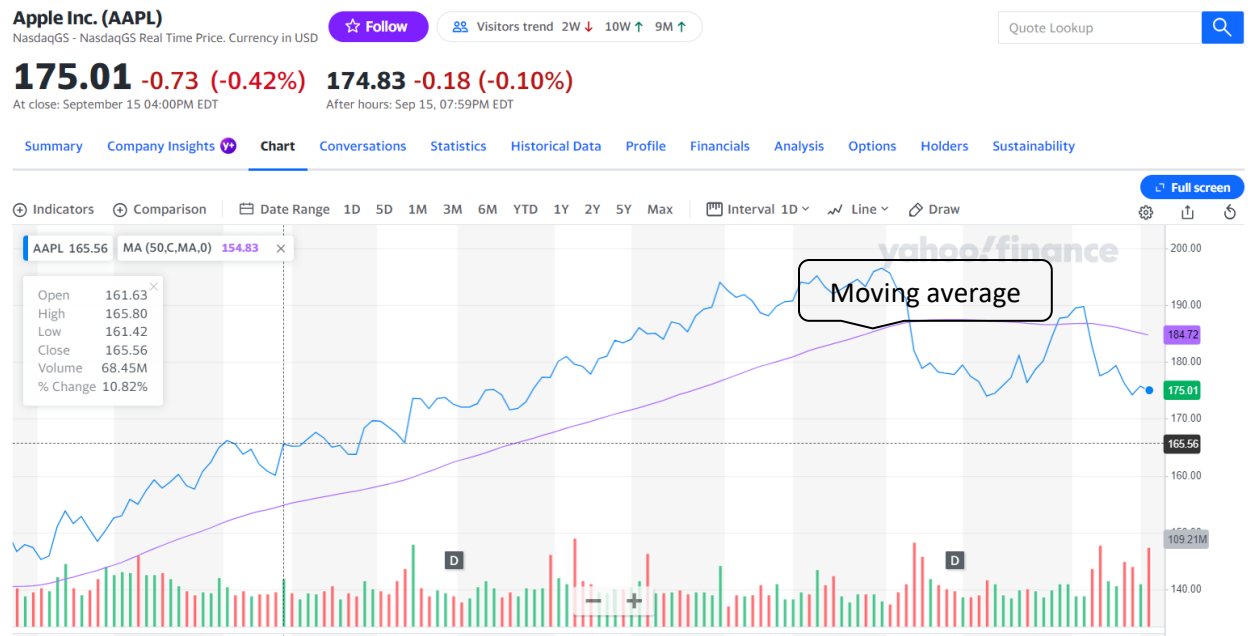
The goal of this project is to provide tools to "visualize" stock behavior over time, predict stock price along with metrics to grade the prediction.

Data Explanation

For this phase of the project, we will use traditional stock parameters. This dataset will have.

Datasets

To establish an initial proof of concept for stock analysis, I will use the historical data for individual stocks. I have collected around 7500 different stock data and 1300 EFT data from Kaggle. In addition to static data, I have time series data from different portfolios for Amazon, Dominos, Bitcoin, Netflix.

The dataset from yfinance will have values for Date, Open, High, Low, Close, Adjusted Close, and Volume.

Data Preparation

Both Datasets will be augmented with a column for label. Label is extracted from the filename. For example, file name mcd.us.txt will be converted to a table with column Label having the value of 'MCD'. We will then add columns for the following technical indicators.

Bollinger Bands

Relative Strength Index (RSI)

Aroon Oscillators

Acceleration Bands

Data Management

Methods

We will use Python as the primary programming language. We will use Matplotlib, Seaborn, Keras, Pandas, yfinance and other necessary libraries. We will show technical indicators for individual stocks. We apply an AI model in Neural networks to predict the price of a given stock and show graphs and metrics to show confidence and accuracy of the prediction.

Analysis

Technical Indicators

Technical indicators are intrinsically pattern-based because they are based on time. They are signals that are produced by the price, volume, and other parameters. Shown graphically, these signals will show spike, declines, gradual/sharp decrease, or gradual/sharp increase, etc. These visuals will help technical analysist predict future price movements.

We created 4 technical indicators. The Bollinger band indicator is a popular one. Shown below, the black line is the 30-day average closing price of the stock, and the gray area is the band range. The upper part of this gray area is the 30-day average plus 3 standard deviations, and low part of the gray area in the average minus 3 standard deviation. The wider the band, the more volatile the stock is. There are
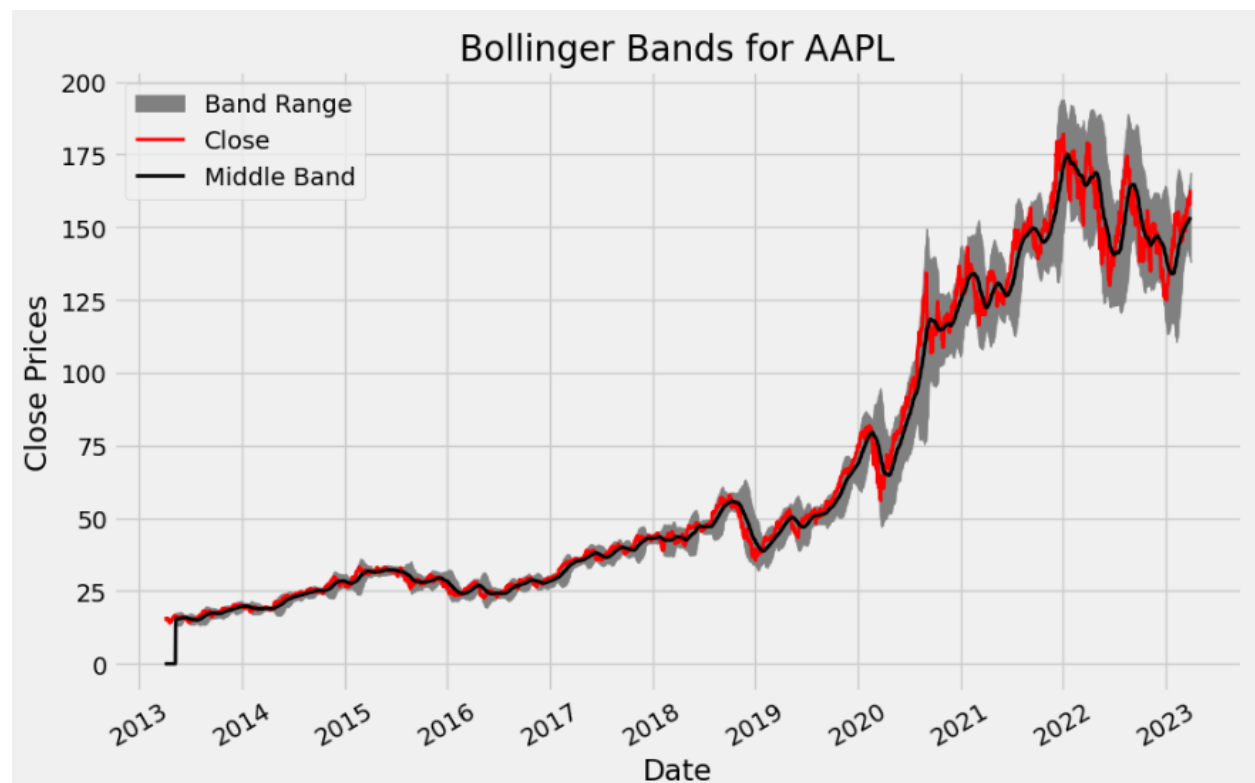
Relative Strength Index or RSI, measures speed and magnitude of a stock's most recent price changes (up or down). This will help evaluate whether a stock is overvalued (don't buy) or undervalued(buy). Shown below, the stock for AAPL was overvalued in 2016,2017 and undervalued since 2021-22.
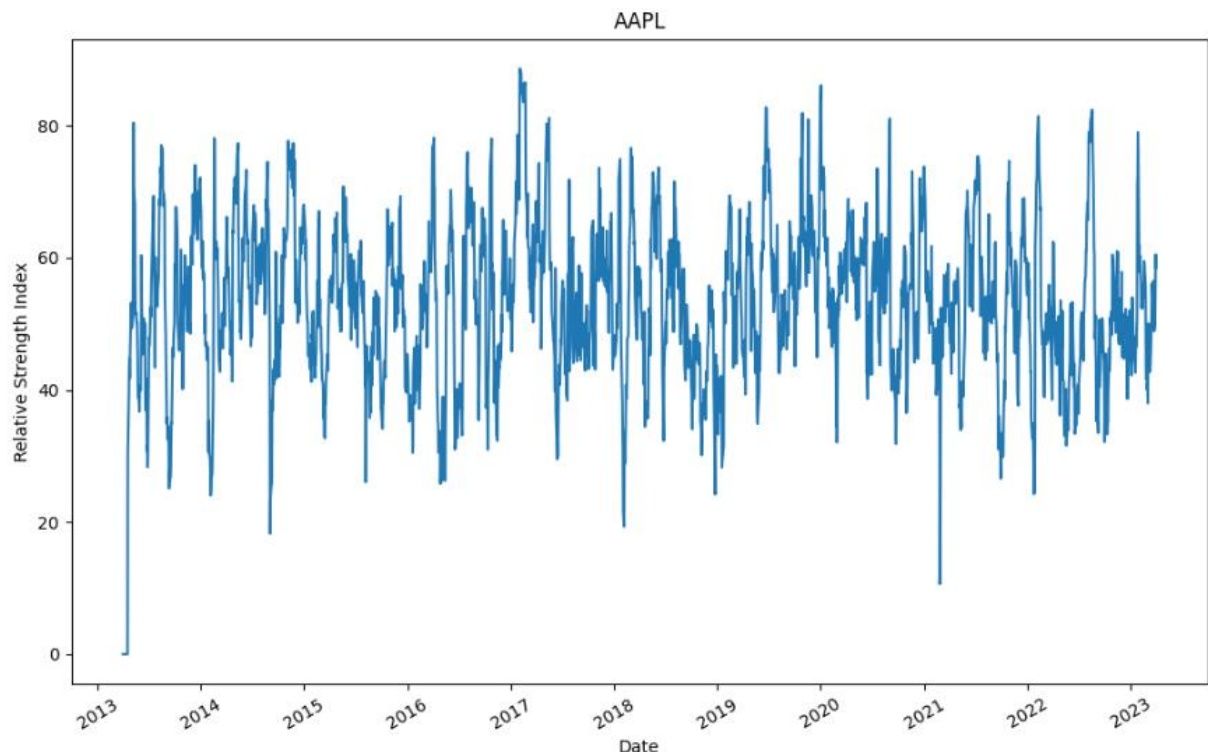
Aroon Oscillators are technical indicators that are used to identify the strength and direction of a trend. The indicator consists of two lines: the Aroon-Up and the Aroon-Down. The Aroon-Up measures the number of periods since the highest high, while the Aroon-Down measures the number of periods since the lowest low. When the Aroon-Up is above the Aroon-Down, it indicates an uptrend, while when the Aroon-Down is above the Aroon-Up, it indicates a downtrend. Shown below, APPL is trending downward.

Aroon Oscillator for AAPL
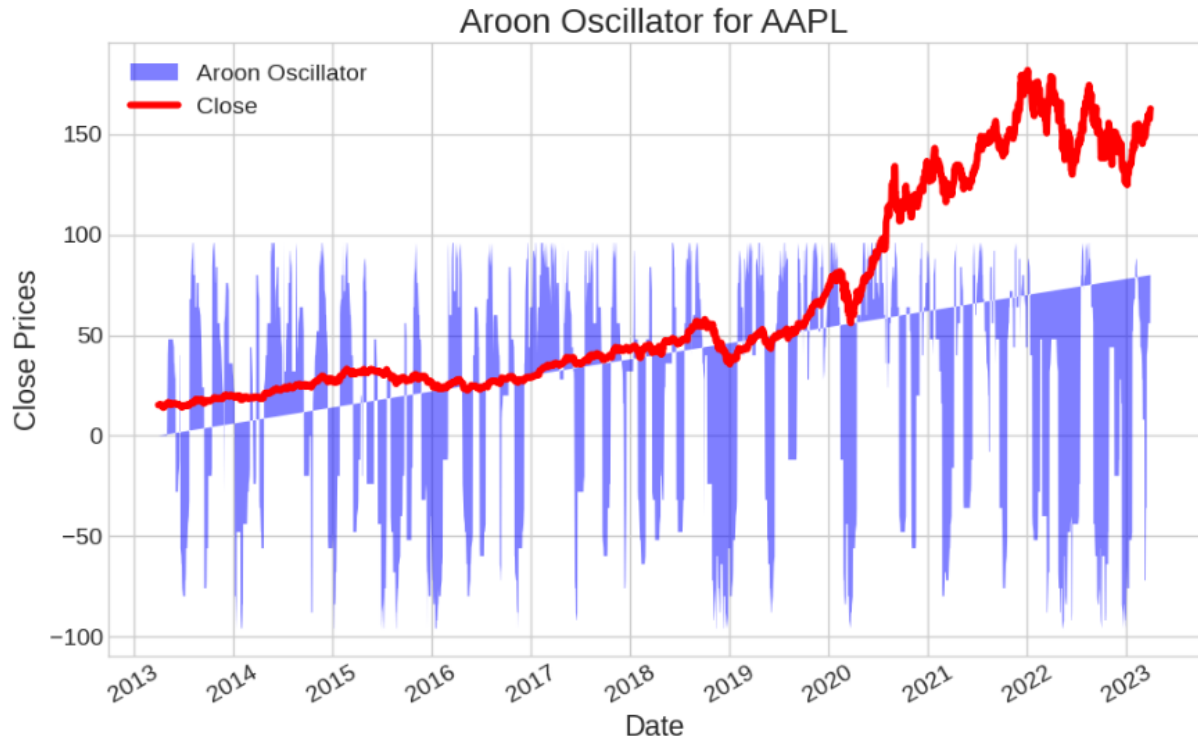
Acceleration Bands are a technical indicator that uses volatility to identify potential price trends in the market. The indicator consists of three lines: the upper band, the lower band, and the centerline. The upper and lower bands are based on the standard deviation of the price from the centerline, while the centerline is typically a moving average. Shown below, when the price touches or crosses the upper band it indicates that the trend is losing its momentum and due for a reversal.

Acceleration Bands for AAPL
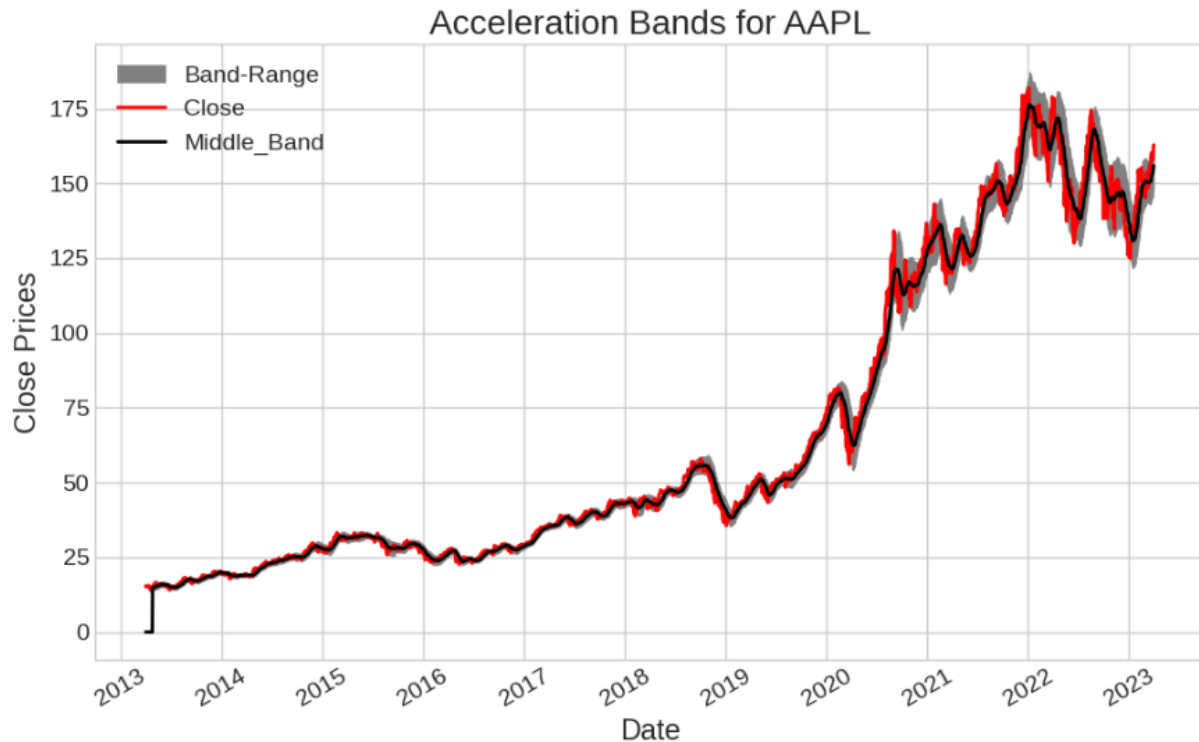
RNN Model

We used a RNN model and used two LSTM layers(shown below). LSTM is suitable for time series dataset. We used the closing price in the training data set to predict closing price in a test dataset.
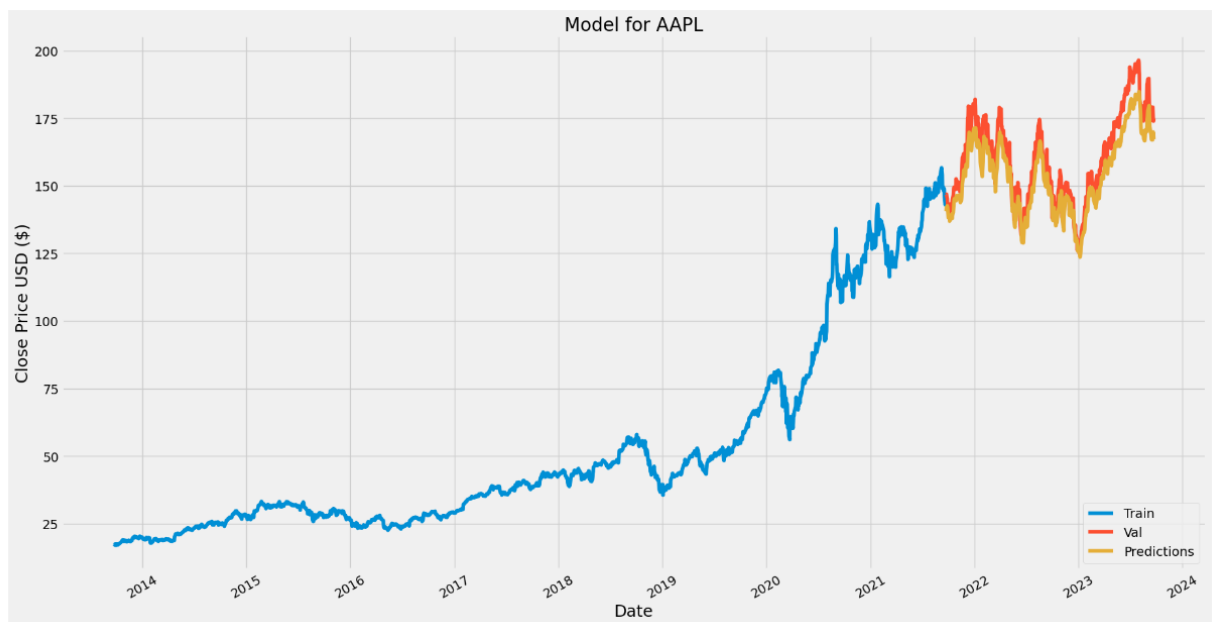
The graph below shows the result. The disparity in the green and orange line shows the accuracy of the model.
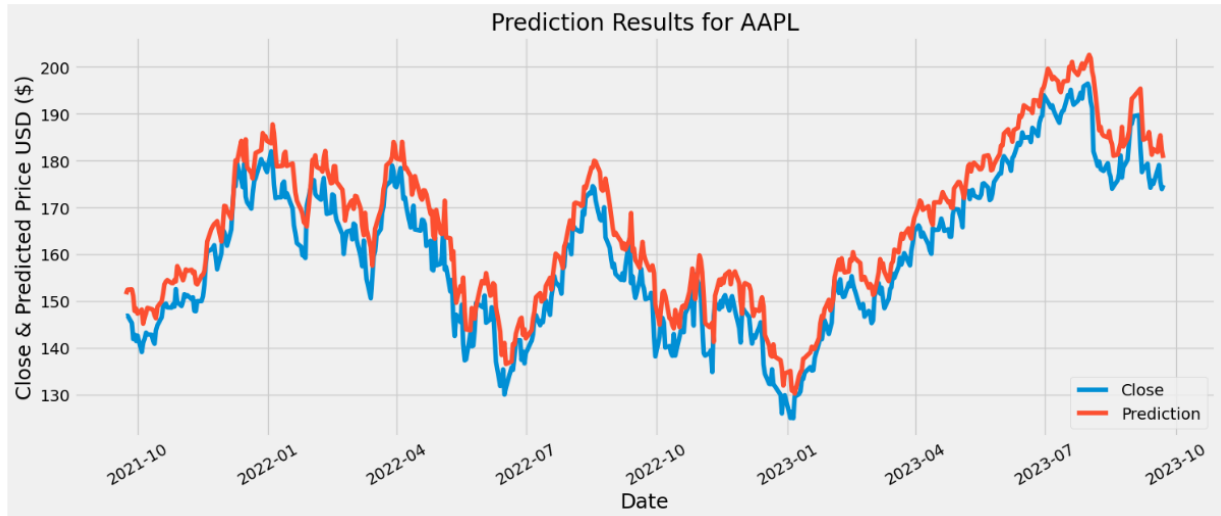

Model for AAPL
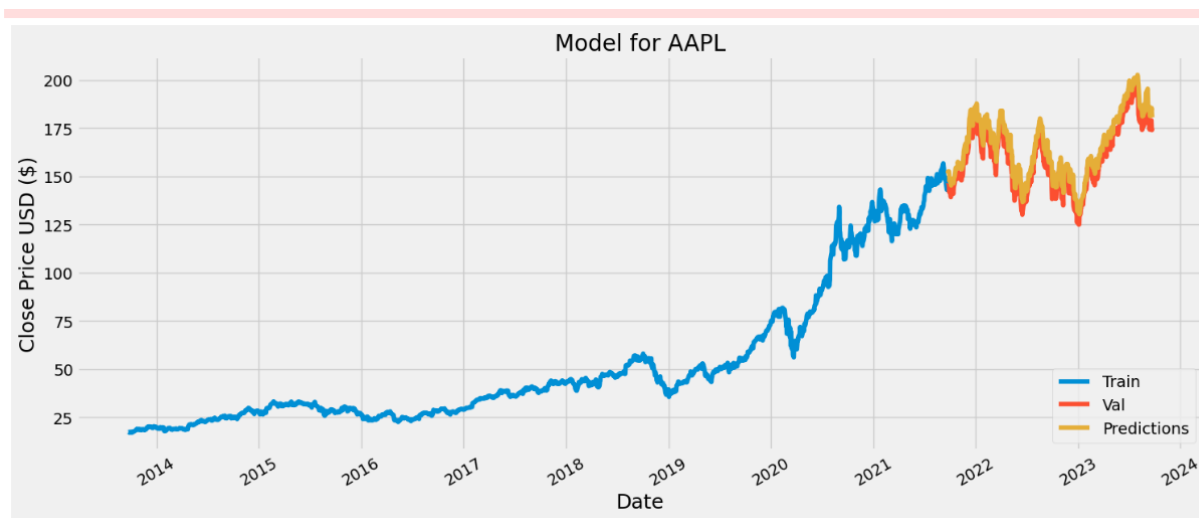
Shown below, the disparity seems quite wide and persistent.
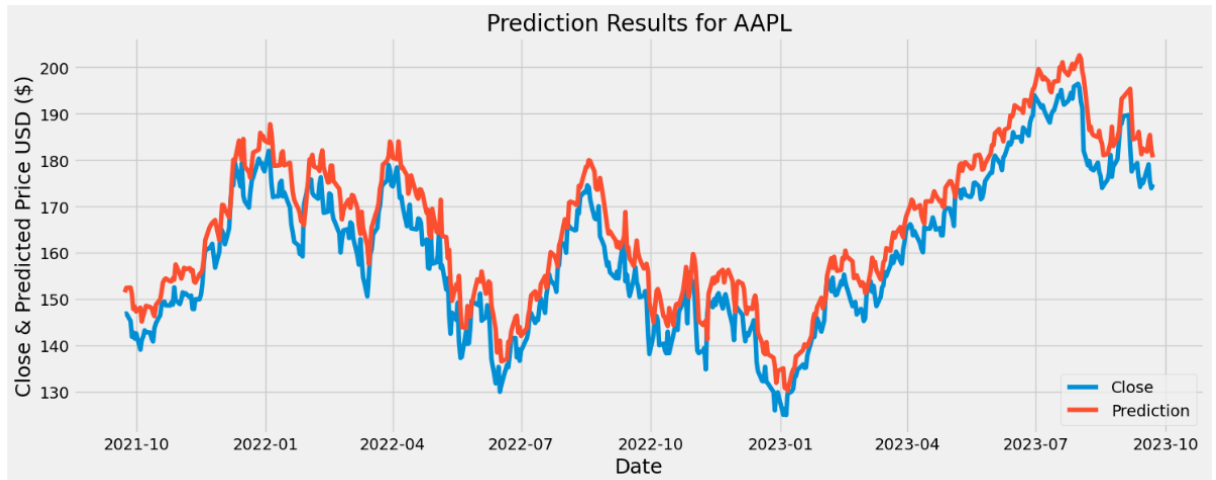
Using a different approach, we came up with this result:

No noticeable difference was found.

Prediction Results for AAPL

Conclusion

We showed that technical indicators help investors make more accurate decisions about buying or selling stocks. They also help with the timing of the transaction, and the price target. Combined with fundamentals of the stock, and other factors such as interest rate, consumer confidence, overall health of the economy (both nationally and globally) and other external factors, investors of today can make educated and low risk decisions.

Assumptions

We must assume that the investors are knowledgeable about the interpretation of the technical indicators, but not necessarily about machine learning. The machine learning aspect of this project is mainly the responsibility of the developers and data scientists.

Limitations

One major limiting factor to this project is data quality. We must ensure that the quality of data is monitored and always maintained. Other limiting factor that we must be aware of is time to market constraint. Model and calculation accuracy is also a limitation that cannot be overlooked.

Ethical Considerations

There are some traditional ethical considerations in a typical service provider relationship.

Bias and fairness: Ensuring that the algorithms and models used in the project do not perpetuate or amplify biases and discrimination against any group.

Data privacy: Ensuring that the data collected is obtained in a legal and ethical manner and that personal information is not misused or disclosed without consent.

Transparency: Ensuring that the data sources, analysis methods, and findings are transparent and easily understandable to all stakeholders.

Security: Ensuring that the data is secure and protected against unauthorized access or theft.

Informed consent: Ensuring that individuals whose data is being used in the project are fully informed about the project's purpose, risks, and benefits and have given their informed consent to participate.

Impact on society: Ensuring that the project's results do not have a negative impact on society or vulnerable populations.

Future Uses & Product Roadmap

The following improvements will be put on the roadmap.

Additional technical indicators

These technical indicators can be added to the portfolio:

1. Stochastic Oscillator(%K and %D)
2. Chaikin Money Flow
3. Parabolic SAR
4. Price Rate of Change
5. Volume Weighted Average Price

New Model

Research and create classification model to predict up or down. Other RNN models and more advanced LSTM model can also help improve our portfolio,

Graphical User Interface

A graphical user interface to allow users to select stocks, select graphs of different types (i.e., RSI, PVT, BBs, etc.) and be able to view the data daily, monthly, yearly, etc. The GUI will have to be web based and accessible by authorized personnel.

Other algorithms.

Research and Implement new ML algorithms.

Dataset

Augment data set with additional parameters to gauge the market fluctuation better.

Recommendations

Market research at any level is an important asset. We recommend that the results of all the analysis done be used as a basis for future market analysis. This will also help in the improvement of the process of stock analysis.

Implementation Plan

To implement this project, we will perform the following main tasks:

1. Gather and prepare dataset.
2. Prepare Design Document.
3. Code, and test
4. Present
5. Deploy

Challenges/Issues

As of now it is very early to document all the challenges and issues that I could face during the course of the project. But things like data quality of the datasets used in the project, algorithm or approach to use for the analysis and prediction, accuracy of the algorithm can be used as primary challenges in this project. I am confident that I rely on the previous course notes and professor's guidance to mitigate all these challenges.

Questions and Answers

1. What is technical analysis?
2. Is technical analysis better than fundamentals?
3. What was the change in price of the stock overtime?
4. Have you investigated using feature engineering to generate new features?
5. Have you considered other measurements that may be useful as features?

6. Are there other models or algorithms that can be used
7. Are there other technical indicators?
8. What are recurrent neural networks?
9. Explain LSTM algorithm.
10. Can use of ML adversely affect the stock market trading?
11. Do you think much improvement could be made using an ensemble?
12. Do you think you would see any improvement by using other transformations on different feature variables?

References:

Kaggle Datasets https://www.kaggle.com/datasets/borismarjanovic/price-volume-data-for-all-us-stocks-etfs

Stock Market Basics https://www.nerdwallet.com/article/investing/stock-market-basics-everything-beginner-investors-know

Stock Market Prediction
https://encyclopedia.pub/entry/16165#:~:text=Technical%20analysis%20is%20the%20study%20of%20stock%20prices,data%20using%20technical%20indicators%20to%20forecast%20stock%20prices.

4 Ways to Predict Market Performance
https://www.investopedia.com/articles/07/mean_reversion_martingale.asp

7 Technical Indicators to Build a Trading Toolkit https://www.investopedia.com/top-7-technical-analysis-tools-4773275

Using Technical Indicators to Develop Trading Strategies
https://www.investopedia.com/articles/trading/11/indicators-and-strategies-explained.asp

Technical Indicator: Definition, Analyst Uses, Types and Examples
https://www.investopedia.com/terms/t/technicalindicator.asp