Artificial Intelligence in the Stock Market: Predicting Prices

Abstract
This research project focuses on stock price prediction through A.I. models as well as machine learning algorithms to maximize profit potential, improve investments, and eliminate risk. Essentially, this project will demonstrate the modern implementation of A.I. in predicting the stock market. Potentially lucrative company stocks and shares have attracted investors as well as general interest in the stock market for decades[4], leading more people to try to predict the rise or fall of market prices. However, industry volatility and the seemingly unpredictable nature of the stock market have led many buyers to invest impulsively, sell their shares at the wrong time, or purchase stock from the wrong company.

To combat these problems, we trained and tested A.I. models on our collected, classified data in order to generate accurate predictions. These models achieved average prediction errors of 0.12% for the stock prices of Amazon, 0.13% for the stock prices of Google, and 0.07% for Microsoft's stock prices on the testing datasets.

Keywords — Stock, Model, Investors, Error

Introduction

Although it is more than 400 years old, the recent explosions of cryptocurrency, NFTs, and other such forms of digital assets have sparked a sudden interest in the stock market. The stock exchange's average daily volume, which has more than doubled since 2019 to total $38.3 million in contracts[8], attests to this renewed interest.

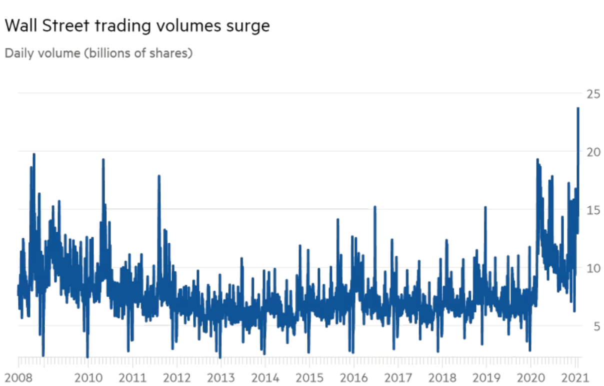

Figure 1 - FT Graph of the Daily Trading Volume on the Stock Market by Billions of Shares (2008-2021)

Nevertheless, the driving force behind any interest in the stock market is the potential to make huge profits relatively quickly. Subsequently, investors have constantly been trying to develop or modify different methods for predicting the stock prices of different companies. In the past, most buyers relied on calculating the moving average with Technical Analysis[2], Fundamental Analysis, or simply heuristics.

By calculating the moving average of a stock's price, investors would create a constantly updated average price that also mitigated random, short-term fluctuations for that stock price. Fundamental Analysis, on the other hand, allowed investors to evaluate a company's financial situation as well as current market and economic trends down to an intrinsic value.

However, this project applies a newer, more sophisticated and efficient prediction method known as Quantitative Technical Analysis. This is a mathematical approach to making accurate predictions for stock prices by measuring signals through statistical modeling. Therefore, the algorithm can generate multiple predictions for any one company as well as compare opportunity costs by contrasting investment opportunities.

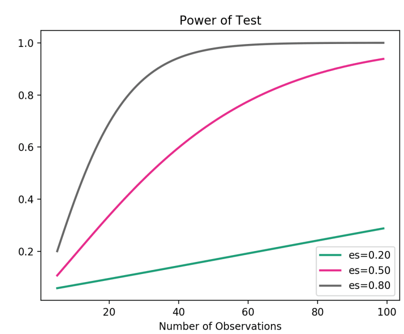

Figure 2 - Example of Statistical Modeling using Power Analysis and Classes to Calculate Power Curves

Naturally, stock market investors would be interested in an algorithm or model working to predict a certain company's stock price to determine if they should purchase shares of said company or not. Such an algorithm would also attract the attention of investors interested in short selling stocks of a company, shaping their ultimate financial outlooks. Perhaps most importantly, publicly traded companies would likely take an interest in these stock prediction algorithms to adapt their financial or monetary policies, especially with regards to their investors and shareholders.

Goal

Is it possible to program an algorithm and model through artificial intelligence that can accurately predict the stock price of any one company on a given day?

In this paper, we will demonstrate that it is possible to use technology and artificial intelligence to accurately predict the stock price of any on a given day, as well as provide an in-depth explanation of the actual programming behind the models.

Importance

The increasingly elusive economic environment, coupled with unnerving studies showing that around 80% of investors lose money in the stock market[5] as they do not understand how the market works, stress not only the dire situation of the market, but also the lack of financial responsibility amongst most investors. In other words, many investors are practically clueless about relevant monetary and industry information, in turn causing them to make poor investments or other such erroneous decisions in the market[6].

Fortunately, this project has developed a program with the use of A.I. models to assist buyers in making better investments through accurate, calculated predictions. Essentially, our models solve the issues of financial responsibility by guiding investors and minimizing risk, a necessity in navigating the illusive, volatile stock market.

Dataset

This project consists of numerical databases, including prediction history, stock prices, and measured error data. The A.I models are also trained to recognize market patterns with price fluctuations as well as compare between predicted outputs and actual outcomes. Additionally, the algorithm holds an up-to-date library that allows users to input a predefined set of dates on which to run the models.

To generate a prediction, the models must evaluate multiple different categories of the numerical data. First, the user must set the amount of time that the models will take stock price data from, whether the last five days or the last five years. As outlined in the programmed algorithm, the models then import this data from a Yahoo Finance library to analyze the stock prices at the open and close of the market for each day within the inputted timeframe. The models then split the data and given values into random sets of training and validation data or testing data.

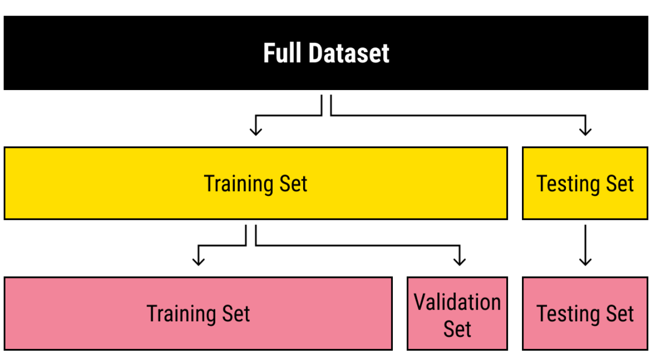

Figure 3 - Division of a Dataset into Training and Testing Sets of Data by an A.I. Model

Next, the models adapt the weights, or learnable parameters, of the training set of data to compare weight accuracies to the random values of the testing set of data at the end of the process. Both models subsequently experience their first learning curves with the data, optimizing parameters to improve accuracy. The models also apply the Multiplicative Weights Method, which is an algorithmic technique that assigns initial weights to each model base and adjusts these weights multiplicatively according to each model's respective performance. This multiplicative weights framework essentially solves the problem of prediction by optimizing the model bases, allowing the algorithm to iteratively decide which successful models to follow and combine.

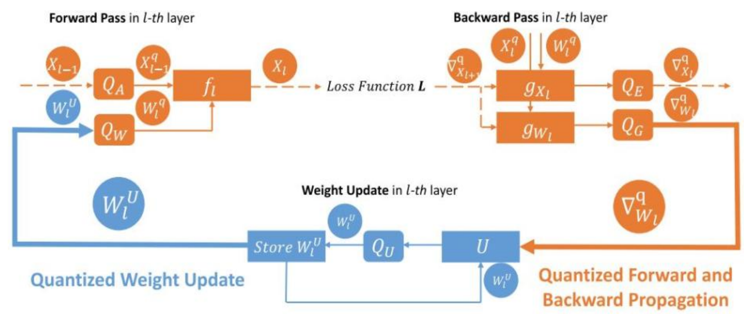

Figure 4 - Multiplicative Weights Method in Model Training with Logarithmic Number System

Now, the algorithm has established weights within both models, which allow the program to test each model on all of the days outlined in the given timeframe. As they did before, the models adapt based on their accuracies and inaccuracies. The models will then adapt to the other models' performances by evaluating weights based off of the other's precision. After an iteration, these weights will be used to combine the predictions of each model into one final prediction, also giving the percent error for each model.

At this point, the algorithm has analyzed stock price history as well as its own prediction history, compared its predictions with actual outcomes, identified market patterns, and weighed its models on performance to output a predicted stock price.

Methods

This project applies several different machine learning and data processing algorithms to generate an accurate stock price prediction. These machine learning methods include a Linear Regression Model, Neural Networks, Boosting Algorithms, and Game Theory.

The Linear Regression Model, one of the two models used in the program to output an prediction, finds a linear correlation based on patterns in the stock's price history. The Linear Model, although utilizing fewer parameters than the Neural Network Model, initializes the exact significance of each daily stock price, which then becomes the weight of that day. These weights are in turn determined by frequencies and irregularities in the stock price history, meaning an outlying stock price for a certain day is weighed less because it is an abnormality.

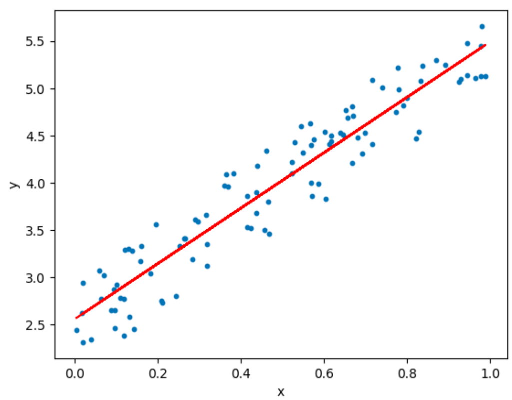

Figure 5 - Demonstration of a Linear Regression Model Evaluating a Dataset to Predict an Outcome

The Neural Network is another model used in the program to output a prediction with a weight system. However, Neural Networks differ from the Linear Model in that they have more parameters and are also able to find nonlinear patterns, meaning they are able to perform nonlinear fitting. Therefore, the Neural Network can detect frequencies in both accurate and inaccurate predictions as well as iterate through the dataset to measure the algorithm's accuracy on a given day. The Neural Network then analyzes for patterns between days where the model was relatively inaccurate and future dates, adapting the predictions to account for random fluctuations in the stock price. The model is also able to detect seasonal patterns and weekly changes in stock price history to ultimately learn and generate an increasingly accurate prediction.

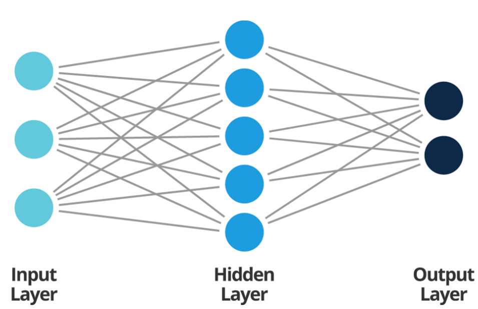

Figure 6 - Diagram of a Neural Network with Weighed, Hidden Layers and Complex Parameters

The Boosting Model is an algorithm that operates as a collaborative system to assess model performance and optimize the predictions of the Linear Model and Neural Network. The boosting algorithm works to reduce model errors in predictive data analysis by increasing the weights of samples with more significant errors. Weights are also assigned based on model performance, meaning that the models with the best predictions will have a greater influence over the final output. Additionally, the Linear Model and Neural Network can adapt to each other's analyses and predictions, essentially meaning that the Neural Network can access the specialized weighting system of the Linear Model while the Linear Model can evaluate and apply the Neural Network's results. Ultimately, the Boosting Model drives the optimization of the generated prediction by refining model accuracy and performance.

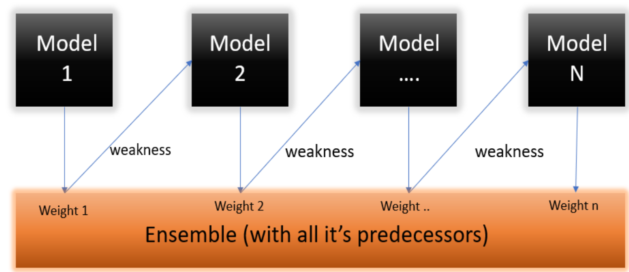

Figure 7 - Diagram of a Boosting Model and an Adaptive Ensemble Technique

Finally, the program will implement a game theory framework to optimize the weights of each model. Game theory is a branch of mathematics that analyzes how models apply reinforcement learning algorithms to solve issues of conflict and cooperation[7]. This method details assessing the impacts of each model's performance on the other[9] and working with the boosting algorithm to merge each model, generating the most accurate prediction.

Advantages

While traditional machine learning models have struggled with evaluating dynamic scenarios and generating accurate long-term stock price predictions[3], the Neural Network and Linear Model act interdependently to ensure reliability and adapt to current market trends. The flexibility of the Neural Network as well as the simple weights of the Linear Model work in unison to create an optimized prediction. Similarly, the boosting method and game theory improve model performances and predictions by reducing bias as well as reaching optimal output in the strategic setting. In particular, the boosting method reduces bias and variance in the machine learning ensemble while the game theory framework allows the Linear Model and Neural Network to evaluate the accuracies of different predictions.

Results

Below are the model-generated stock price predictions of Microsoft, Amazon, and Google over two separate, random dates by both the Linear Regression Model and the Neural Network. For comparison, the actual stock prices for the specified dates are listed next to the model predictions. Further down is the average percent error for the predictions of each company's stock prices by both models.

27 February, 2023

	LRM	NN	Actual Price
Microsoft	$253.50	$253.31	$252.46
Amazon	$95.13	$94.91	$94.28
Google	$91.99	$91.45	$90.09

***LRM - Linear Regression Model Prediction**
***NN - Neural Network Prediction**

29 February, 2023

	LRM	NN	Actual Price
Microsoft	$250.10	$250.76	$250.76
Amazon	$93.76	$93.80	$93.87
Google	$89.75	$89.85	$90.16

***LRM - Linear Regression Model Prediction**
***NN - Neural Network Prediction**

	LRM Error	NN Error
Microsoft	0.34%	0.17%
Amazon	0.51%	0.37%
Google	1.28%	0.89%

***LRM Error - Linear Model Average Percent Error**
***NN - Neural Network Average Percent Error**

Conclusion

This project draws its success from the innovative weighting techniques, adaptive learning structures, and complex model systems that we implemented. To verify the accuracies of the Linear Model and Neural Network, the A.I. models were tested on other companies as well, producing a surprisingly low average prediction error of 0.21% for Tesla's stock prices and 0.18% for Apple's stock prices. As machine learning algorithms and A.I. models advance, these systems can also be updated or optimized to account for political, cultural, and socioeconomic trends in certain industries to generate predictions at an even higher accuracy. The almost-limitless applications of modern A.I. prediction methods in helping investors reduce risk while maximizing potential profits highlight the need for such technologies in the unpredictable financial market.

Open Questions

In the future, I intend to modify the program as well as adapt the dataset to be able to factor relevant socioeconomic or sociopolitical news and events into the models when they are used to generate predictions. The increasing influence of social media on the stock market[1], as seen with the GameStop short squeeze in 2021, signifies the societal shifts that have given rise to, for example, cryptocurrency stocks. The new program would essentially be able to evaluate the impacts of social media posts and news articles by quantifying such data and assigning values to that data. Thus, the A.I. models would be able to take into account relevant news and online publications to generate an even more accurate prediction, which would also be a significant breakthrough in stock prediction technologies.

References

[1] Bishop, J. (2021, December 2). *From GameStop To Tesla – How Social Media is Driving the Stock Market*. Maddyness. https://www.maddyness.com/uk/2021/12/02/from-gamestop-to-tesla-how-social-media-is-driving-the-stock-market/

[2] Fernando, J. (2023, January 9). *Moving Average (MA): Purpose, Uses, Formula, and Examples*. Investopedia. https://www.investopedia.com/terms/m/movingaverage.asp

[3] Gülen, K. (2023, January 11). *Applying Machine Learning in Financial Markets: A Review of State-of-the-art Methods*. DATACONOMY MEDIA. https://dataconomy.com/2023/01/stock-prediction-machine-learning/

[4] Malkiel, B. (1973). *A Random Walk Down Wall Street: Including a Life-Cycle Guide to Personal Investing* (pp. 11-39). W. W. Norton & Company, Inc.

[5] Plummer, S. (n.d.). *Why is the Stock Market Down, And Why Is It Crashing?* The Annuity Expert. https://www.annuityexpertadvice.com/why-is-the-stock-market-down-crashing/

[6] *Reasons Why & How People Lose Money in Stocks*. (2022, December 1). Angel One. https://www.angelone.in/blog/6-reasons-why-some-people-lose-money-in-stock-markets

[7] Silva, M. (2019, October 26). *Game Theory to Interpret Machine Learning Models and Predictions - The Ultimate Guide*. Towards Data Science. https://towardsdatascience.com/the-ultimate-guide-using-game-theory-to-interpret-machine-learning-c384cbb6929

[8] *Trading & Data.* (n.d.). New York Stock Exchange. https://www.nyse.com/trading-data

[9] Vaidya, D. (n.d.). *Game Theory*. WallStreetMojo. https://www.wallstreetmojo.com/game-theory-2/