**Stock Market Predictor**

**Ethical Consideration Worksheet**

A close-up of a graph

Description automatically generated with medium confidence

Team 6: The Razorbacks Programmers

Elayne Blancas: Team Manager

Gregory Perez: Programming Manager

Jonathan Zamudio: Resource Manager

19 July 2021

**Table of Contents**

Social Benefit – Page 3

Bias – Page 4

Safety – Page 4

Accountability – Page 5

Privacy – Page 5

Quality – Page 6

Principles – Page 6

Works Cited – Page 8

**Social Benefit**

The most important applications of our Stock Market Predictor are the following:

1. To predict whether an individual stock will either overperform or underperform.
2. To calculate the long and short gains of our portfolio’s stocks, to then obtain its average return and Sharpe Ratio. The Sharpe Ratio characterizes the portfolio’s risk-return and trend.

Any scale of investors gain the most benefit from the application of predicting the market’s behavior. This application can be used as a tool to help investors decide when and which stocks to purchase for the growth of their investment [2]. The nation, as a whole, also greatly benefits from a healthy stock market since the stock market affects the US economy in the following manners:

1. Stocks allow the individual investor to own part of a successful company. Without stock markets, only large private equity investors and financial institutions could profit from America's free market economy [2].
2. Stocks allow savers to overcome inflation. There is an approximately 7% annual increase on stock prices, after taking into account inflation, which generously compensates most investors for the additional risk of owning stocks rather than bonds or keeping the money in a savings account [2].
3. Growing, successful businesses need capital to fund growth and the stock market is a key source. In order to raise money this way, owners must sell part of the company, and to do so they "take the company public" through an initial public offering (IPO) of the company's shares. An IPO raises a lot of cash. It also signals that the firm is successful enough to afford the IPO process. The drawback is that the founders no longer own the company; the stockholders do. Founders can retain a controlling interest in the company if they own 51% of the shares [2].

Typically, the stock market and economic performance are aligned. Therefore, when the stock market is performing well, it is usually a sign of a growing economy, vice versa. The most prominent measurement of economic growth is the Gross Domestic Product (GDP) [5]. GDP is the total monetary or market value of all the finished goods and services produced within a country’s borders in a specific time period. As a broad measure of overall domestic production, the GDP functions as a comprehensive scorecard of a given country’s economic health [5]. As a result, when the GDP is increasing that is an indication that individual businesses are expanding. Expanding business activity usually increases valuations and leads to stock market gains [3]. A stock market predictor could assist with early detection of threatening trends allowing us to uncover hidden nuances about the economy and gain a greater understanding of such a volatile industry.

## **Bias**

A potential bias present in our model comes from the data set itself. The sharpe ratio requires normalized/standardized data which is difficult to obtain with stocks. Therefore, there is a possibility that our data is not normally distributed since they have a lower limit yet infinite upper bound. This may cause our results to be more skewed to the right and not normally distributed.

There is reporting and coverage bias specifically. Reporting bias occurs when the frequency of properties are not accurately reflecting the real-word frequency. This is especially present in our model since there are various different variables that affect the stock market that we can’t account for, such as terrorist attacks, pandemics, and political changes. Coverage bias can be seen in our selection of the stocks we are looking at. We chose to look at a selected few of the top S&P 500. These stocks will show different returns than a lower tier or smaller company’s stock price. As such, our model would not really reflect the real-world distribution, since we haven’t selected data in a representative fashion. Another form of bias present is automation bias. We were encouraged to use a preexisting repo on Github to assist us with this ambitious goal. Their model could have had some preexisting bias in their model that we are not proficient enough yet to rectify.

We can overall mitigate some of these biases by training our model over some more representative data by taking a random sampling of stocks from not just the S&P 500, but also including the NASDAQ or DOW. In addition to this, we can help reduce the reporting bias if we were able to do some sentiment or fundamental analysis as well as applying beautiful soup and web scraping. These analysis types would allow us to take into account public feedback/other external factors such as social media and news updates that can affect stock price [4].

## **Safety**

Both big companies and small investors can take advantage of the stock market predictor in an effort to obtain faster riches. It is also possible for either party to manipulate the market in a selfish manner, negatively impacting other investors. Another possibility is that people begin to blindly believe the predictor and make impulsive investments as a result. In the event that the predictor is incorrect, this can be catastrophic for the emotional and psychological well being of the investors, and lead to potential fears in the public.

On the other hand, a stock market predictor can also be used to combat manipulation. For example, some people who buy short stock are notorious for being media moguls, because those people then go on TV and scare normal people into selling their stock, so the short sellers can profit [6]. With a stock market predictor, the investors and public can make their own assumptions.

## **Accountability**

Our Stock Price Prediction model will analyze data pertaining to the opening and closing stock prices for a portfolio of S&P 500 companies from the year 2010 to present day. The input will include numerical data pertaining to the selected stocks. This takes the form of a single feature vector that contains consecutive sequences of 240 days worth of daily historical stock data for each stock. Our model will output a prediction about the performance of our portfolio.

To accomplish this goal, we obtain daily data from Yahoo Finance for the top 30 S&P 500 companies and use a decade worth of historical data for our analysis. Our labels include date, open, close. We use yfinance to collect the daily historical data. With the gathered data, we anticipate splitting it into a 60 to 20 to 20 percent ratio for training, testing, and validation sets. The data will be split sequentially, because we are using a RNN using our collected time-series data. Our data will go back as far as a decade to ensure that our model is as representative as possible.

Our predictor uses random forest as a baseline, and a LSTM model. LSTM’s are very powerful in sequence prediction problems because they're able to store past information. This is important in our case because the previous price of a stock is crucial in predicting its future price. Our model predicts the likelihood of whether a stock will under or over perform. We use calculated metrics related to the stock, such as its average return and the Sharpe ratio to determine its performance. By taking the stock’s historical data over the past decade, we can calculate these metrics and give a reasonably accurate prediction. This prediction can be faulty, given the fact we will not be able to take into account the real-world situations that affect stock prices everyday [3].

The users should know the risk of the stock market: it is very volatile. Our model does not account for anything external and should, therefore, not be used as a main investment tool. Additionally, we would include a disclaimer against using our model as the investor’s sole investment advice.

Although we are limited in our resources and knowledge, we created the model to the best of our ability. We will be more than happy to address/rectify any concerns/ issues that our users may encounter in all efforts to ensure that the outputs are as accurate as possible. For future work we would adjust our inputs to include sentimental analysis and web scraping for a greater performance in accuracy.

**Privacy**

Our model only uses public data on stocks that everyone can easily find online with a simple search. The user does not need to input any of their personal information or vital stock portfolio information to use our model. Our model does not use or require any potentially sensitive data or information from our users in any way. Therefore, there should not be a privacy risk.

**Quality**

LSTM models have proven to be 'good' for time series forecasting. Therefore, it is a commonly used and widely accepted model for stock market predictions, positively reviewed in literature. To prevent overfitting, we dropped 20 percent of our data, a common practice used in machine learning. Additionally, in an effort to obtain a higher accurate score, we use a decade of the stock’s history to predict its future trends. In making this model pipeline, we would use the 3 prior years as our training data. This pipeline would take the years of data and go over it in a series of sequences, each sequence containing 240 days. This sequence of data would then give us the value of the stock.

Our model performs decently, having an overall accuracy rating in the 60-70% range. However, we acknowledge that the quality of the model is not the best. This type of predictor is difficult to make generalizable, since the stock market remains very volatile due to impacts from the political, economic, and social changes of the world.

## **Principles**

Despite its critical role in the economy, the stock market is not the same as the economy. The stock market is driven by the emotions of investors [2].﻿ They can exhibit irrational exuberance [2]. It occurs during an asset bubble and the peak of the business cycle. They become overly optimistic even though there is no hard data to support it [2]. Our mission is to provide stability in the emotions of the investors by predicting if a stock is either over or under fit, and the stock market’s trends. Our goal is to be utilized as an additional tool/resource in their decision-making process.

This model fits our mission and moral compass by providing valuable additional insight on the stock market for potentially new and preexisting investors. The model itself is not malicious or dangerous, however, in the wrong hands it can be used for such intent. Regardless of the circumstance, the stock market is very volatile to begin with, making it that much more difficult for the predictor to be used to cause major changes in the market.

Possible objections that may exist include people blindly trusting the predictor for their investments, and investors using this as a tool for manipulating others to take advantage of a favorable outcome. It is for this reason that we use a disclaimer and do our best to be as informative and honest as possible with our users. This predictor should not be used for an investor sole advising tool nor believed to its entirety since it does not include important sentimental data, for example, that impact the stock market on a daily basis.

Works Cited

[1] Amadeo, Kimberly. “How Stock Investing Affects the Us Economy.” *The Balance*, 27 Nov. 2020, www.thebalance.com/how-do-stocks-and-stock-investing-affect-the-u-s-economy-3306179.

[2] Duncan, Garrett. “The Gamestop Stock Scandal Exposes the Corruption of the Rich.” *The Ridge*, arhsnewspaper.com/9219/showcase/the-gamestop-stock-scandal-exposes-the-corruption-of-the-rich/.

[3] Fernando, Jason. “Gross Domestic Product (Gdp).” *Investopedia*, Investopedia, 21 July 2021, www.investopedia.com/terms/g/gdp.asp.

[4] Obthong, Mehtabhorn, et al. “A Survey on Machine Learning for Stock Price Prediction: Algorithms and Techniques.”

[5] Team, The Investopedia. “How Does the Performance of the Stock Market Affect Individual Businesses?” *Investopedia*, Investopedia, 9 June 2021, www.investopedia.com/ask/answers/042215/how-does-performance-stock-market-affect-individual-businesses.asp.