

---

# Machine Learning Project - Stock Fluctuations

---

Mary Letey

Morgan Allen

Colton Williams

Aniq Shahid

## Abstract

Research project investigating relationships between media concerning companies and the aforementioned company's stock price. Theoretically, the stock price of a company is based in part on public perception of their market. This perception can easily be affected by media sources. The project team hypothesizes that social media sources and news sources may be used to predict daily stock price fluctuations within the technology industry. **TODO :: One sentence on results.**

## 1 Introduction

Stock prices generally appear unpredictable; however, *theoretically* they are quite simple. Considering the most basic supply-demand model in economics, stock prices are functions of the supply and demand for ownership in a company. Because the "consumers" in this supply and demand model are viewing a stock as an investment, demand represents confidence in the performance of the company. Furthermore, an increasing stock price represents increased confidence in the potential of the company. Thus the stock price may be affected by a change in the public's understanding of a company's perceived growth and risk. Forms of media, such as news articles, may have a huge impact in these perceptions.



Figure 1: Facebook's stock price over the last year

For example, when examining Facebook's stock price over the last year, one of the most prominent features in Figure 1 is the dramatic drop around the beginning of April. Recall, Mark Zuckerberg testified before congress on April 8th. This hearing caused concern about Facebook's future, where traders expected Facebook's stock price to continue declining around a window of April 8th, according to Business Insider [1].

This project will examine the extent to which media sources affect stock prices, and how well we can use media data to predict stock price changes.

## 2 Data and Methods

The focus of this project is to investigate the effects of media on stock price changes inside the technology industry. A single industry was chosen in order to maintain consistent results. The companies in this project representing the tech industry are:

- Hewlett Packard
- IBM
- Seagate
- Western Digital
- **TODO :: replacement for Dell here**

The focus of the media data is incorporating textual data from news articles into our model. The articles were gathered from reputable news sources such as the Wall Street Journal.

Furthermore, data from Google Trends is also included to obtain correlation between certain terms to model public interest in the company.

Numerical data measuring the closing price and volatility of shares was retrieved from Bloomberg. Other numerical data included macro-economic factors, such as the S&P500, crude oil, and gold.

## 3 Results and Discussion

The Google Trends data provided an excellent starting point for this analysis. As one would expect, given that they measure public interest in a company, the Google Trends data combined with previous stock data predicted the closing price of the company's shares.

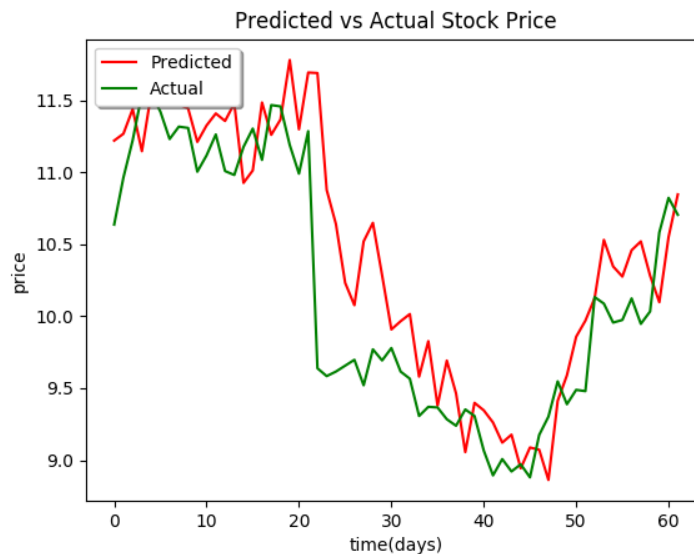


Figure 2: Using Google Trends to Predict Closing Price

For example, Figure 2 shows almost exact results using a standard regression, where the Predicted and Actual curves are almost completely correlated.

## 4 Conclusions

## 5 References

- [1] Ciolli, Joe (2018, April 10) *Facebook traders are bracing for the worst ahead of Mark Zuckerberg's hearing.* Retrieved from <http://www.businessinsider.com/facebook-stock-price-hedging-ahead-of-mark-zuckerberg-congressional-testimony-hearing-2018-4>