

# CS 254 Machine Learning

## Project Proposal

Nikhil Sai Reddy Choppa, Tyler Scott, Varshney Gentela  
February 14, 2023

Instructor: Prof. Safwan R. Wshah

### 1 Introduction

The goal of this project is to develop an Automated Trading Bot with Machine Learning in Python and also if time permits we will deploy our ML model in AWS. This project seeks to address the major problem with trading stock, which can lead to preventing large financial losses for individuals. Our machine learning model aims to make trading more profitable for individuals participating in stock trading by improving the time taken and accuracy of future results in trading to produce great profits.

### 2 Problem Definition and Algorithm

People who are active in trade on the stock market are prone to losses of wealth. Through our machine learning model, we would like to increase the probability of a higher hit ratio (percentage of successful and profitable trades) for individuals.

#### **The Algorithm: Predictive Modeling**

- The training model is built on historical data to make predictions on new, unseen stock trends.
- We will use this historical data in conjunction with Non-Linear Regression and an Elastic Net Regularization term to see possible future trends of stocks such as the S&P 500 and other Stocks.
- We will also automate trading strategies, such as technical indicators, and execute trades on predefined sets of trading rules to the machine learning algorithm.

**Tech Stack:** Python, AWS, OOPs, Numpy, Pandas, Matplotlib, scikit-learn, Keras, Tensorflow, Broker APIs, and many more...

### 3 Dataset

We will be using datasets of the last 50 years of Open price, Closed price, High, Low, Volumes, and Change% of everyday data for 5 companies (Microsoft, Google, Apple, Amazon, and Tesla) from investing.com. The datasets are labeled, and we don't need any special hardware to process our data.

<https://www.investing.com/equities/microsoft-corp-historical-data>

<https://www.investing.com/equities/google-inc-historical-data>

<https://www.investing.com/equities/apple-computer-inc-historical-data>

<https://www.investing.com/equities/amazon-com-inc-historical-data>

<https://www.investing.com/equities/tesla-motors-historical-data>

### 4 Related Work

To build upon the work of other researchers, we have reviewed and compared our approach to two related studies.

#### **Journal 1:**

*Logistic Regression Model For Predicting Performance of S&P BSE30 Company Using IBM SPSS*

As the title implies, this study is an attempt at predicting the future behavior of companies that are in the S&P BSE30. The goal of this study was to answer these questions:

1. Can the return on stocks be explained with the help of different financial ratios?
2. Can we analyze stock return using a logistic regression model?

The model used applies Logistic Regression with a Maximum Likelihood Function that extends the Ordinary Least Squares method of Linear Regression. This is then combined with the Newton-Raphson Equation, a method for solving systems of nonlinear equations. The companies are separated into “Good” and “Poor” classifications; “Good” being a high-profit yield and “Poor” being a lower profit yield. It is then analyzed to predict the probability of either increasing or decreasing profits over time. This will help with providing a solid predictor for future investment opportunities.

### **How Ours Differs:**

Our approach to this problem would be different by the implementation of an Elastic Net Regularization. The choice of Elastic Net Regularization should fit our model as the multiple sets of data we use may have different quantities of features we can use. This will allow us to shape our model to more closely model Ridge Regularization if we have many features and also leaves open the ability to shift it towards Lasso Regularization if the data has fewer features present. We feel that this method would help us keep the model relatively modular and applicable to multiple different sets of data, as opposed to the single dataset used in the Journal.

### **Journal 2:**

*Classify Stock Market Movement Based on Technical Analysis Indicators Using Logistic Regression*

The goal of this study is to “identify the significant technical indicators to predict stock market movement by using logistic regression.” The study focuses on only one specific company out of Bursa, Malaysia; the company in question is not named. It utilizes these 7 technical indicators in its model: Moving Average, Exponential Moving Average, Relative Strength Index, Moving Average Convergence Divergence, Rate of Change, Stochastic Oscillator, and Volume Trading.

### **How Ours Differs:**

Our project is different in that we are using multiple sets of data with the combination of automatic trading strategies with different technical indicators to execute trades on a predefined set of trading rules to forecast future market trends.

## **5 Bibliography**

Smita, Mrinalini. “Logistic Regression Model for Predicting Performance of S&P Bse30 Company Using IBM SPSS.” *International Journal of Mathematics Trends and Technology*, vol. 67, no. 7, 2021, pp. 118–134.,  
<https://doi.org/10.14445/22315373/ijmtt-v67i7p515>.

Zaini, Bahtiar Jamili, et al. “Classify Stock Market Movement Based on Technical Analysis Indicators Using Logistic Regression.” *Beh, Hui*, no. Sang, 19 Mar. 2019.  
[https://www.akademiabaru.com/doc/ARBMSV14\\_N1\\_P35\\_41.pdf](https://www.akademiabaru.com/doc/ARBMSV14_N1_P35_41.pdf)