Predicting the Stock Market One Quarter at a Time

*Note: Sub-titles are not captured in Xplore and should not be used

1st Cory Carter
Computer Science Master of Science Student
Oakland University
Auburn Hills, USA
email address or ORCID

2nd Tom Tisdall
Software Engineering and IT Master of Science Student
Oakland University
Auburn Hills, USA
ttisdall@oakland.edu

***Abstract*—Predicting the stock market is a very challenging task, with hundreds if not thousands of papers addressing this problem. This paper sets out to explore predictingstock quarterly trends by determining the best quarterly features to use and then leveraging them to determine the direction of the upcoming quarterly report. The thought was to eventually marry this will daily trends to improve accuracy of predicting stock market pricess by factoring in the quarterly direction. By determining the stocks that will go up or down for a quarter, we will predict whether buying or selling is the best course of action for a stock. Testing for accuracy, we will simuilate buying and selling in the stock market based on our predictions to determine if this project was a succsess or failure.**

****Add more details to this section as we further develop our strategy**

***Index Terms*—Index Terms - Stock Market, Prediction, Machine Learning**

I. Introduction

Determining the direction of a stock from day to day is both a challenging and frustrating problem. There are so many features to choose from and that factor into a stock changing price from day to day. It can be as simple as positive or negative news, changes in the overall market or industry, technological breakthroughs, production hurdles and many more, which don't even cover the technical or quantative numbers that typically are used to determine a stock rating and predicting its movement from day to day. This paper is goign to attempt a different tact on predicting stocks, at a more generic and higher level, using the companies quarterly returns. The thought is this information can be added to other predictors to better refine the accuracy of those models. In this project we have gathered twenty five Technology, Software and IT Services large and mega cap stocks with earnings history for their entire existence, typically between ten and thirty years. We started with the features used in the paper Stock Market Trends Prediction after Earning Release [1]. We reviewed and refined those quarterly features by using Mutual Info Regression over the datasets that we obtained, which provided us with the most important columns to leverage in our machine learnging algorithm. We decided to implement Support Vector Machine (SVM) to help us determine what stocks to buy or sell each quarter. Using our model, we then started with a set amount and bought/sold as recommended to see how we faired in the long run.

II. Review of existing techniques for predicting the stock market

For this project, we started by reviewing papers that were published in the realm of stock market analysis via machine learning. One of the first papers we found Stock Market Analysis: A Review and Taxonomy of Prediction Techniques [5] discusses how stock market analysis falls into four categories, statistical, pattern recognition, machine learning, and sentiment analysis. Since this paper focuses on machine learning, we focused on that part of their paper. They and other papers they referenced felt that the stock market can be predicted, the challenge is determine what information provides the most value in prediction the stock price. In the paper they discussed they suggest the agorithm space continues to grow with newer techniques like Random Forrest replacing naïve Bayes and Artificial Neural Networks (ANN), Deep Neural Networks (DNN), and Sentiment Analysis becoming more popular. They note that in cases where transaction fees are not considered that the traditional models perform better, but where transaction costs are included, the DNN models perform better. Bin, Ahmed and Fadel in their paper Stock Market One-Day Ahead Movement Prediction Using Disparate Data Sources [3] that using data from outside the stock market could help improve the prediction accuacy. In thier paper they proposed using Wikipedia traffic along with stock data helped improve their SVM model. Using sentiment analysis to predict stock price gains and losses seems to be gaining popularity and some success. The article Automated Stock Price Prediction Using Machine Learning [4] used a SVM with traffic visits to Wiki to determine Apples stock price over time at an 85% success rate. Unfortunately their prediction testing only used Apple, it woudl have been interesting to see how the model performed against other stocks. Intestingly enough they did find that the ordering of the features made a difference in their modeling, reminding us that there are many aspects to defining and developing an accurate model. Automated Stock Price Prediction Using Machine Learning [4] also applied sentiment analysis by using news articles to determine positive or negative trends. In their case they found

that using a SVM was the best algorithm. Seeing the benefits of sentiment analysis, this woudl be something we would like to include in future research.

Changing direction we found other articles that compared recent techniques such as Particle Swarm Optimization, Least Square Support Vector Machine (LS-SVM) and Artificial Neural Network (ANN) to determine which one was a better fit. A Machine Learning Model For Stock Market Prediction [6] found that using Particle Swarm to optimize Least Square SVM was better than just LS-SVM by itself and that LS-SVM was better than an Artificial Neural Network (ANN) which tended to run into an overfitting problem. Along the same lines An Emprical Study of Machine Learning Algorithms for Daily Trading Strategy [7] compared stocks from the US and Chinese markets across six Machine Learning Algorithms (Logistic Regression (LR), Support Vector Machine (SVM), Classification and Regression Tree (CART) , Random Forest (RF), naïve Bayes (BN) , eXtreme Gradient Boosting (XGB)) and six DNN algorithms (MultiLayer Perception (MLP), Deep Belief Network (DBN), Stacked Autoencoders (SAE), Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), Gated Recurrent Unit(GRU)).

They found that "some traditional ML algorithmshave a better performance than DNN algorithms in most ofthe directional evaluation indicators" and that the best DNN algorithms were not significantly better than the traditional ML algorithms without considering transaction costs. As transaction costs increase, they found that the ML algorithms performance decreased. The DNN models degraded when including transaction costs but at a much slower rate than the ML Algorithms.

Interestingly enough, along the same lines of differences in outside variables like transaction costs, Deep Architectures for Long-Term Stock Price Prediction with a Heuristic Based Strategy for Trading Simulations [8] found that when they created an application to simulate trading based on the predictions of their two models Long Short Term Memory (LSTM) and Convolutional Neural Network (CNN) that LSTM outperformed by gaining a higher total dollar amount, but that the CNN model had a higher number of days with gains. This once again shows that there are many influencers on the stock market that make it challenging to predict.

III. Formulation of our solution

IV. Applicable mathmatical theory

V. Experimental Results

VI. Summary and Conclusion

Citations

- Predicting the daily return direction of the stock market using hybrid machine learning [2]
- Stock Market One-Day Ahead Movement Prediction Using Disparate Data Sources [3]
- Automated Stock Price Prediction using Machine Learning [4]
- Stock Market Analysis: A Review and Taxonomy of Prediction Techniques [5]
- International Journal of Computer Science and Telecommunications 4.17-23 [6]
- An Empirical Study of Machine Learning Algorithms for Stock Daily Trading Strategy [7]
- Deep architectures for long-term stock price prediction with a heuristic-based strategy for trading simulations [8]
- How to use Machine Learning to become a millionaire predicting the stock market [9]
- I spent 20 minutes trying to predict the stock market with AI, these are my results [10]

References

[1] C. Qian, W. Zheng, and R. An, "Stock market trends prediction after earning release," *WebSite(Stanford)*, 2016.

[2] X. Zhong and D. Enke, "Predicting the daily return direction of the stock market using hybrid machine learning algorithms." *Financ Innov 5, 24*, 2019.

[3] B. Weng, M. Ahmed, and F. Megahed, "Stock market one-day ahead movement prediction using disparate data sources," *ResearchGate*, 2017.

[4] M. Moukalled, W. El-Hajj, and M. Jaber, "Automated stock price prediction using machine learning," *WebSite*, 2019.

[5] I. Shah and Zulkernine, "Stock market analysis: A review and taxonomy of prediction techniques," *International Journal of Financial Studies 7.2*, 2019.

[6] O. Hegazy, S. Omar S., and M. Abdul Salam, "A machine learning model for stock market prediction," *International Journal of Computer Science and Telecommunications 4.17-23*, 2013.

[7] L. DongDong, S. Yuan, M. Li, and Y. Xiang, "An empirical study of machine learning algorithms for stock daily trading strategy," *Mathematical Problems in Engineering 2019(7):1-30*, 2019.

[8] C. Stoean, W. Paja, R. Stoean, and A. Sandita, "Deep architectures for long-term stock price prediction with a heuristic-based strategy for trading simulations," *PlosOne*, 2019.

[9] J. Xu, "How to use machine learning to become a millionaire predicting the stock market," *Website*, 2019.

[10] A. Edell, "I spent 20 minutes trying to predict the stock market with ai, these are my results," *Website*, 2018.