Predicting Stock Returns – Chaos Theory and Machine Learning

Michele Orlandi ISYE6767 Fall 2022

1. Introduction

In this project we attempt to predict whether future stock returns will be positive or negative, using two Machine Learning models – a Multilayer Perceptron Classifier and a Support Vector Classifier – trained and tested on daily data of 110 NASDAQ and NYSE stocks between January 1st 2000 and November 12th 2021. The predictor variables are technical indicators attempting to quantitatively identify the formation of Elliott Waves within a certain time frame.

Elliott Waves are the core of trading Chaos Theory. In his pioneering research Dr. Bill Williams, a trader and author of 'Trading Chaos'[1], explains a fundamental issue with current quantitative approaches to trading. Price paths and returns are almost always modeled through linear mathematics and parametric statistics; Brownian Motions, Black-Scholes models, Fourier Transforms, they all try to explain movements that are random to make sense of a chaotic market and "bring stability" to our reasoning.
What if chaos wasn't unexplainable? What if instead of fearing randomness, we instead embraced it? Here is when Elliot Waves come into play, or more specifically, fractals.

Nature's complexity is extremely difficult to explain through linear mathematics, financial markets, as a reflection of human nature and thus nature itself, are also hard to make sense of through classical Newtonian/Euclidian math and physics. Fractals however, as a measure of irregularity and scale, can be very helpful in identifying timely entry and exit for trades. Elliot Waves, built on fractal geometry, show the geometric representation of a random pattern that remains constant at different scales of time. The ability to spot the formation of these waves is crucial to accurately time trades to maximize profits. The basic rhythm of an Elliot Wave is composed of 5 waves. Here is a diagram of a wave in an upward trending market.

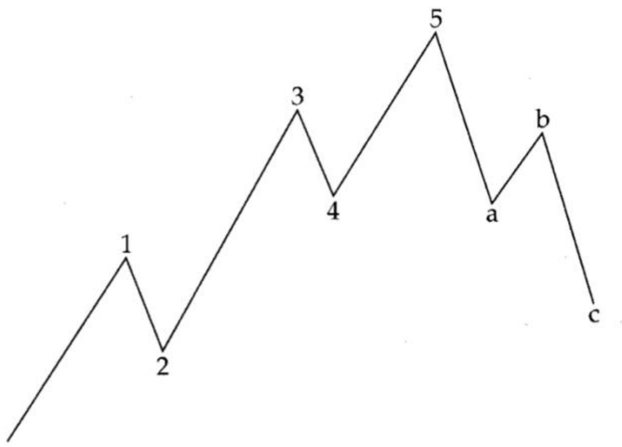

Figure 1. Basic Elliot Wave Rhythm

[1] Williams, Bill. *Trading Chaos: Applying Expert Techniques to Maximize Your Profits*. Wiley, 1995.

In this project we will use several technical indicators used to identify these Elliott Waves to predict whether future returns will be positive or negative.

Notation used in the report:

- MACD: Moving Average Convergence Divergence
- TWR: Alligator Indicator
- MFI: Market Facilitator Index
- EMA: Exponential Moving Average
- MA: simple Moving Average
- NN: Neural Network
- MLP: Multilayer Perceptron
- SVC: Support Vector Classifier

2. Data Cleaning

We were given various sets of tickers traded on both the New York Stock Exchange and the NASDAQ from which to retrieve daily stock data. We used the Yahoo Finance API to firstly retrieve data for 10 firms between 2000 and 2021, and to also get the data for the top 100 NASDAQ firms by market capitalization for the same period.
As some of the data might be incomplete or have holes, we decided to forward fill. In some cases, not enough data was available for the twenty years due to the age of certain firms so, should the data be not enough to provide relevant results we ignore it.

3. Feature Engineering

After cleaning the data, we engineered our necessary features to spot the formation of Elliott Waves. These can be summed up into 4 technical indicators:

- 5/34/5 Moving Average Convergence Divergence (MACD)
- 5/13/34 Alligator Indicator (a.k.a. Tide-Wave-Ripple)
- Bottom or Top Fractals
- Squats

These indicators will produce 9 time series of continuous data that will act as our predictor variables. Here we provide a brief explanation of the meaning of each indicator:

- MACD = (5-period EMA of close prices) – (34-period EMA of close prices)
- MACD signal = 5-period EMA of MACD oscillator
- Tide = 5-period MA of close prices
- Wave = 13-period MA of close prices
- Ripple = 34-period MA of close prices
- Top Fractal = Highest High spotted at the center of 9 consecutive candles
- Bottom Fractal = Lowest Low spotted at the center of 9 consecutive candles
- Volume Change = percentage change in trading volume
- MFI = (High – Low) / Volume

How do these indicators work in relationship to Elliot Waves? To a trader, they provide binary signals when they meet certain conditions. Ideally, when these signals appear either at the same time or within few time intervals of each other, they would indicate the end of a wave and the concomitant start of the next one. We interpret the indicator as follows:

- Buy signals:
  - Bottom fractal forms
  - MACD > MACD signal (positive trend)
  - Tide > Wave & Wave > Ripple (positive momentum)
  - Squat = Volume change > 0 & MFI change < 0 (market movement has lost directional momentum and is ready for a reversion)
- Sell signals:
  - Top fractal forms
  - MACD < MACD signal (negative trend)
  - Tide < Wave & Wave < Ripple (negative momentum)
  - Squat = Volume change > 0 & MFI change < 0 (market movement has lost directional momentum and is ready for a reversion)

Once we calculate the data it is necessary to standardize it as the indicators provide significantly different absolute values.

Our prediction variable will be a series of binary signals indicating whether the n-period return is positive (1) or negative (0). Since Elliott Waves tend to form in a range between 120- and 140-time intervals we chose a 24-day holding period return as our prediction. We then shift the prediction data back by the holding period.

We thus have preprocessed and engineered our features and are ready to apply our classifiers to each data set.

4. Models

We chose two Machine Learning Models to classify our data:

- Multilayer Perceptron Classifier (Neural Network)
- Support Vector Classifier

We chose the Neural Network as our first classifier because of its relatively higher performance when applied to non-linear problems and speed in its predictions. While we chose to also use a SVC because of the strength of its kernel function, its scalability to high-dimensional data and the relatively less risk to incur into overfitting.

After a Brute force search for the optimal hyperparameters for the NN we settled on the following because they provided the best results:

- Hidden Layer Size = (50, 50, 50)
- Activation Function = tanh
- Solver = lbfgs
- Alpha = 0.0001
- Learning Rate = adaptive

For the Support Vector Classifier, we used a Grid Search Convergence to find the optimal parameter for each dataset. We didn’t run a GridSearchCV for the Neural Network as warnings

would have cluttered the console, we decided to thus use brute force to find the overall best parameters.
Despite the many advantages of these models, they don't come without certain drawbacks, the main one being very slow training time.

5. Results

After fitting each data set to both models, we retrieve each one's accuracy and precision scores and plotted them against two given benchmarks, 50.14% accuracy and 51.41% precision.

Firstly, let's analyze the NN's results on the initial set of 10 firms. The output was impressive as the average accuracy was around 65% and precision was about 68%, well above the benchmarks.

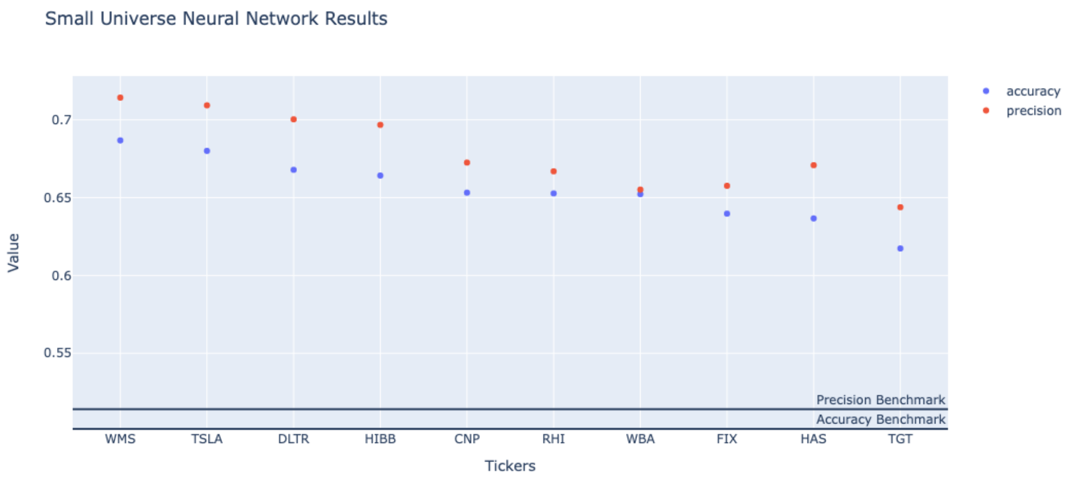

Figure 2. Small Stock Universe MLP Classifier Results

The same remarks can be made after running the MLP Classifier on a larger universe of stocks. Below the top 20 performers are shown. As we can see the results are even better with an average 70% accuracy and precision, with the top performer accurately classifying results 83% of the time.

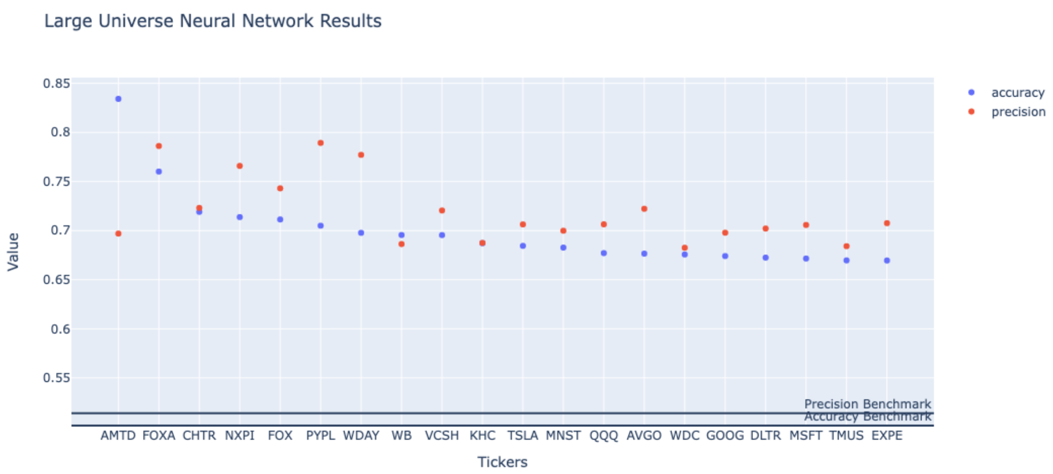

Figure 3. Large Stock Universe MLP Classifier Results

Next, we analyze the results of the SVC on the small set of stock data. As shown below, the classifier once again shatters the metrics benchmarks but provides slightly worse results than the Neural Network. It is interesting to notice how the top performers through a Support Vector Classifier are almost the same as the NN.

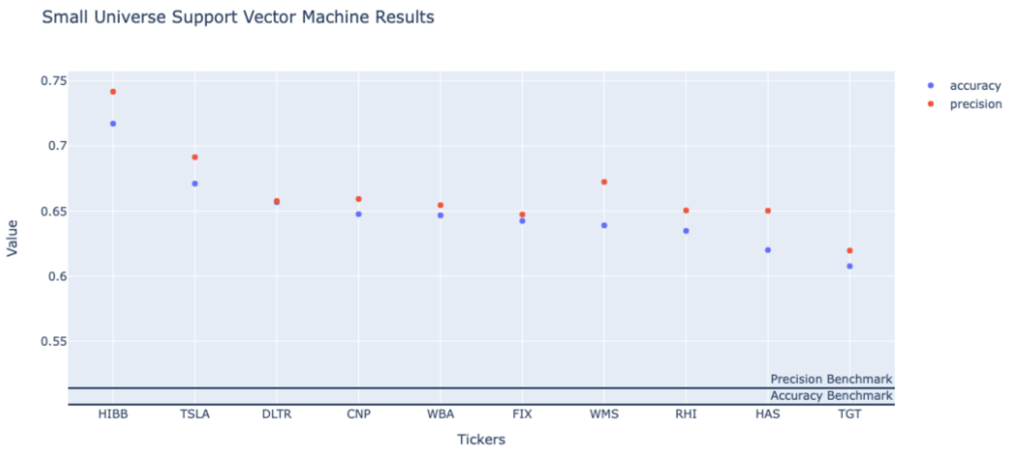

Figure 4. Small Stock Universe Support Vector Classifier Results

When looking at the larger universe of stocks the results are also promising. They once again show an average of 70% accuracy and precision, with all the top 20 performers surpassing the benchmarks, and the top performer scoring the highest accuracy of both models at 86%.

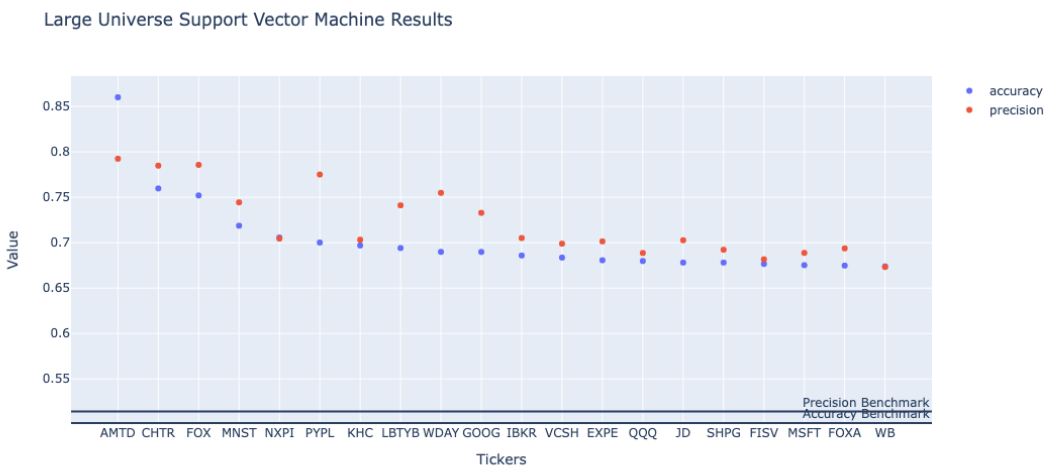

Figure 5. Large Stock Universe Support Vector Classifier Results

6. Conclusions

In conclusion, both Machine Learning models provide a fairly accurate prediction of whether the future 24-day return will be positive or negative.

It provides great insight on the effectiveness of these quantitative indicators and their ability to spot optimal entry and exit points for a trader. Nonetheless further research is needed to increase accuracy scores. The formation of Elliott Waves is more evident in erratic markets and lower time frames, as a matter of fact most of Bill Williams's research was focused on the highly volatile commodities markets and intra-day trading. Therefore, our obvious next steps are to test the validity of these models on lower timeframes and on more volatile securities. We didn't seem to suffer from any overfitting problems either as our out-of-sample errors don't increase as our degrees of freedom increase.

Hopefully, through this project and its results, we won't be afraid of chaos anymore and we can see it for what it really is, a higher, natural form of order.