Timeframe Trading Algorithms

Student Name: A.L. Gillies

Supervisor Name: M. R. Gadouleau

Submitted as part of the degree of MEng Computer Science to the

Board of Examiners in the School of Engineering and Computing Sciences, Durham University

***Abstract* —** The abstract must be a Structured Abstract with the headings **Context/Background**, **Aims**, **Method**, **Results**, and **Conclusions**. This section should not be longer than half of a page, and having no more than one or two sentences under each heading is advised.

Context/Background - Algorithmic trading is characterised by an entirely hands off approach to stock market trading. All data manipulation, mathematical inference, machine learning and trade execution is done autonomously. With this approach, how much of an improvement can be gained over a standard interest rate provided by a high street bank, in the time frame given?

Aims - Using the average interest rate calculated from British banks, the aim of this paper is to show, through implementation of statistical and machine learning techniques that algorithmic trading can improve the annual return on investment over a given time frame.

Method - This paper will consider two possibilities for implementation of the system, a purely statistical method, relying on known practices and techniques, and a hybrid system incorporating both statistical reasoning and machine learning. The known statistical practices are mostly used by human traders to allow for data insight and are well vetted. The machine learning techniques are widely used in other contexts, with limited academic papers being available for this area.

Results -

Conclusions -

***Keywords* —** Algorithmic, Machine Learning, Statistics, R, Trading, Stocks

I INTRODUCTION

The stock market has been an early adopter of technology since its inception, with companies wanting to get an edge over their fellows and thus earning the most money. The first computer usage in the stock market was in the early 1970s with the New York Stock Exchange introducing the DOT system or the Designated Order Turnaround system, this allowed for bypassing of brokers and routed an order for specific securities to a specialist on the trading floor. Since this point the use of machines to allow for increase throughput and speed has been pandemic. From this point it was inevitable that computers would be used to aid in the decision making process of what to buy or sell and when. This was shown to be very effective and got significant traction in the financial market in 2001 with the showcase of IBMs MGD and Hewlett-Packard's

ZIP, these two algorithmic strategies were shown to consistently outperform their human counterparts. These were both based on academic papers from 1996 so the academic conception of algorithmic applications in financial markets has been present for several decades. Whilst in the current day over one billion shares are traded every day, this would not be possible without computerised assistance.

Unique ID	Deliverable	Description
DL1	Simulate the financial market	Have data for at least 10 companies for at least a year, with data for each minute where data is available.
DL2	Allow buying and selling of stocks	Have a functional buying and selling mechanism, with the data collected for each transaction processed.
DL3	Implement statistical methods	Implement as many statistical methods as are beneficial to allow for the insight into the data for each stock.
DL4	Implement a purely statistic strategy	Using just the statistical methods implemented in DL3, create a strategy that will buy and sell stocks to maximise profit made over the time frame given.
DL5	Create a hybrid strategy	Implement a machine learning trading strategy that uses the stock data as well as any statistical methods that are helpful to maximise profit made over the time frame given.
DL6	Implement tracking systems	Implement graphical and table outputs for the results of the computer logic and trading performance.
DL7	Create a testing criteria	Create a method with which to test the strategy so as to avoid over fitting.

Table 1: Deliverables

II RELATED WORK

III SOLUTION

How the simulation was set up.

The logic behind the simulation was to allow trading at a fine grain level, this needed to be permitted by the data chosen to perform the algorithms on. A drawback was that 'tick data' or data for each trade executed was found to be too costly for the project so 'minute data' was used. Minute data provides an open, high, low, and close price for a stock in the given minute, this gave plenty of data points per day of trading. A simplification was created at this point, trading was only done on the price that opened the minute. This was done to reduce the complexity of calculating any inter-minute values. The running of the simulation was based on the progression

through each date of available trading and for each day iterate through every minute of the trading day, starting at 09:30 until 16:00. Each trading day has 390 points at which trading is possible, if the data permits it. This is available for X days in the 18 months of data that was used.

Talk through what data was used.
Talk through why that data was used.
Why was the simulation set up that way?
What statistical methods were used first?
How do these methods work?

IV RESULTS

V EVALUATION

VI CONCLUSIONS