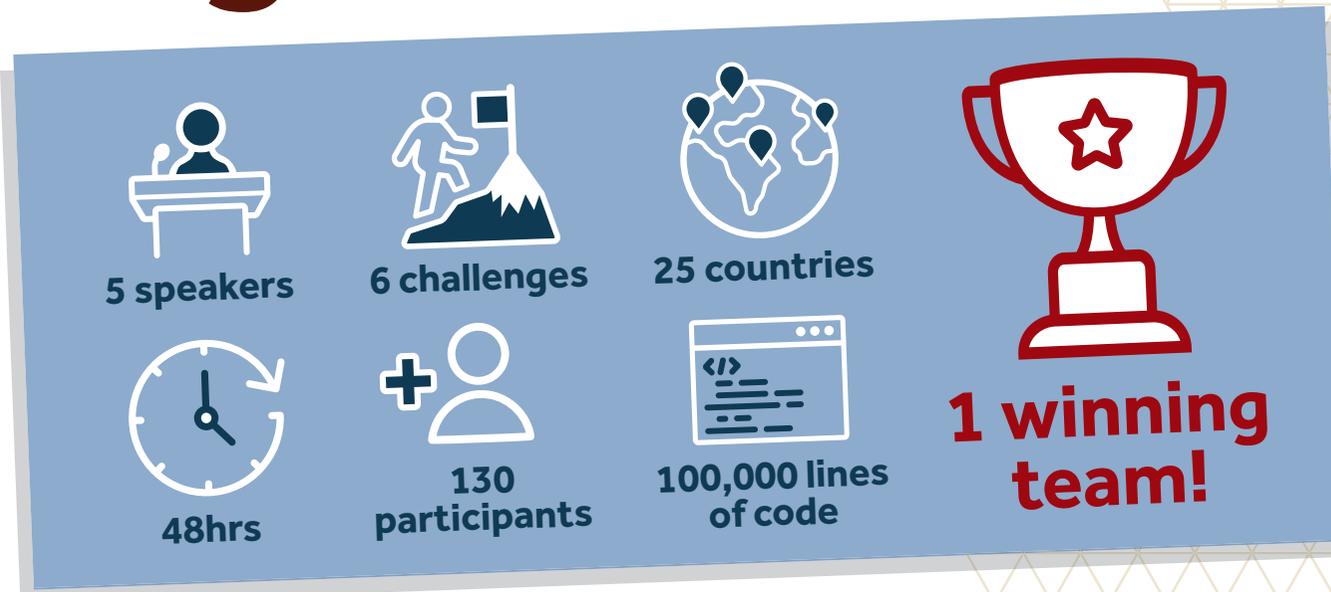


Algothon Virtual '21



5 speakers

6 challenges

25 countries

48hrs

130 participants

100,000 lines of code

1 winning team!

The infographic features a blue background with white icons: a speaker at a podium, a person climbing a mountain, a globe with location pins, a clock, a person with a plus sign, and a code editor window. A red trophy with a star is positioned to the right of the statistics.



Aspect was delighted to sponsor the Algothon 2021 run by the Imperial Algorithmic Trading Society.

In its third year, the event brought together some of the brightest student and graduate minds for 48 hours of coding challenges. It is the largest financial machine learning hackathon in Europe.

The event also included some excellent talks from leading researchers in the field and from Aspect staff on topics including the importance of data visualisation and trading algorithms in execution research.

Here, we shed light on the solutions and techniques used during the event and draw comparisons between the challenges and our processes at Aspect.



**Imperial College
London**



The Challenge

Dataset cleaning

The Tools

A corrupted time series data set

The Criteria

Mean squared error of the cleaned dataset and corrupted data

Data Cleaning Challenge: Winning Solution

Whilst the data cleaning challenge was a simple one to familiarise participants with the competition, the skills required to complete the challenge are essential to perform well both in the competition and in systematic investing. The participants received a distorted dataset, and their job was to remove the anomalous data points.

Most participants noticed that there were several large data points and removed these by filtering large absolute values. One of the techniques the better submissions applied was to transform the data into returns and filter large returns, then convert back to the time domain; this removed the more subtle noise and vastly improved their estimate of a clean dataset.



Word from Aspect: Hieu Nguyen, Lead Data Analyst

As a quantitative investment manager, data is central to all of our trading decisions and having clean data, free of outliers, is imperative. With numerous asset classes and instruments in our coverage, we require a robust, reliable and performant solution to meet the many challenges in this area.

For example, a highly volatile equity will have very different price movements to stable interest rates which rarely change. Additionally, intraday data used in live trading would need to be cleaned in real time and on a constant basis. These are the types of challenges we face everyday.

We also employ many tools to “health-check” our data and the processes involved, so we can be proactive in case of outages or delays from vendors. This helps minimise the impact from such incidents and allows us to continue trading smoothly.



The Challenge

Use machine learning techniques to enhance the portfolio returns of a trading strategy

The Tools

- List of trades including market / time-on / time-off and the profit & loss
- Universe of liquid futures on stock indices, FX, fixed income and commodities

The Criteria

Portfolio returns calculated as the sum of returns according to the actions recommended by the model built, and scored using the Sharpe Ratio of daily returns over a test period of 2 years

Machine Learning Challenge: Winning Solution

Participants were given anonymised intraday futures trade data taken from a live strategy run by Aspect. The goal was to provide an 'action' for each trade: -1 to invert the trade, 0 to cancel the trade, and 1 to allow the trade, in order to maximise PnL/Sharpe Ratio. Participants were free to use any external data they liked to construct their features, applying machine learning techniques to explore cross-sectional predictability to overlay and modulate portfolio allocations. We had a variety of submissions using gradient boosting and neural network approaches. The winning team used a transfer-learning technique which allowed a neural network to be trained on an orthogonal dataset and then applied to the trades data. This reduced overfitting and allowed the model to stay relatively simple, whilst improving returns.



Word from Aspect: Darius Horban, Quantitative Developer

Machine learning is gaining increasing attention across the finance industry. With the hunt for alpha so competitive, it is crucial that we explore these cutting-edge techniques to stay ahead of our peers. With such a high barrier to entry and enormous degrees of freedom, machine learning models allow us to construct unique, low-beta strategies that maximise the predictive power of our data.

Here at Aspect, we are often faced with deeply complex, higher dimensional and non-linearly constrained optimisation problems for which supervised learning models are ideally suited. When developing these technologies, a particular focus is applied to interpretability and risk management so we can maintain high levels of transparency and accountability opposite our clients.

The Challenge

Assign a weighting to a long-only portfolio of stocks for each trading day, with the portfolio being restricted to stocks with all three E, S and G components having a score of >45

The Tools

Encrypted data for US stocks.

Stock features included:

- Percentage change
- Float market capitalisation
- ESG features for each stock
- S&P500 target daily percentage change

The Criteria

Mean of absolute difference in daily percentage returns of the portfolio against an S&P500 Benchmark

3

ESG Challenge: Winning Solution

Financial understanding of how the S&P500 Index is constructed was the key to this challenge. A good solution should build upon a naive solution of assigning weights proportional to floating market capitalisation to stocks fulfilling the ESG criteria. Improvements on this could include clustering the stocks based on ESG criteria, their sector/industry or past returns, then replacing stocks failing ESG screens with similar stocks.

Most solutions were built upon regression or machine learning methods on the returns data, which in fact under-performed the naive solution. This highlighted the importance of financial understanding of the problem before applying statistical techniques to solving it. Only three teams submitted solutions that outperformed a naive solution and achieved joint 1st place.



Word from Aspect: Rosie Reynolds, Chief Commercial Officer

Responsible investment remains at the forefront of our clients' minds and we need to ensure we meet their evolving expectations in this area. We believe that there is no one-size-fits-all approach to ESG integration, and our approach varies depending on the investment strategy adopted and asset classes traded. We seek to incorporate ESG into the investment process where appropriate to do so; in other cases, we need to demonstrate that we have been thoughtful about the reasons it may not be appropriate.

Nor does Responsible Investment begin and end with ESG; other elements – including being a responsible market participant and acting as a responsible corporate citizen – are also part of the story, and we dedicate a lot of effort to these areas.



The Challenge

- Develop a long-short strategy in: tech, energy or gold
- Remain market neutral and provide a time series of weights of long and short stocks

The Tools

Five stocks from one of the three sectors provided

The Criteria

- Performance of models by calculating cumulative returns – 0.5* max drawdown
- Output needed to be a number between -1 and 1

4

Long-Short Sector Strategy Challenge: Winning Solution

The winning team managed to develop a robust architecture in their neural network, with relu and tanh as their activation function, with the hidden layer twice the size of the input layer to optimise out-of-sample performance.

From this solid model, they were able to gauge which of the three sectors yielded best results.

As intended, the sectors chosen gave opportunities for a long-short strategy given their mean-reversion type behaviour.



Word from Aspect: Graham Baird, Researcher

By taking long and short positions such that the overall position nets to zero, the objective is to hedge out exposure to the sector or the market as a whole.

Provided the stocks in which long positions are held outperform relative to the stocks selected as shorts, then the overall portfolio stands to gain irrespective of the performance of the sector or wider market.

Within Aspect, such market neutral relative value strategies have been applied across multiple asset classes, by a number of different programmes.



The Challenge

- Build a visual dashboard to best interpret a dataset
- Both the visualisation tool and the analysis derived had to be meaningful

The Tools

- A website hosted using Heroku
- Data from the data cleaning challenge

The Criteria

Functionality, insights derived and user experience were equally scored

5

Data Visualisation Challenge: Winning Solution

In this challenge teams could stand out in one or more of the three criteria (UI, functionality, insights derived from the data). Some built a very good user interface but their data analysis tools were not sophisticated. Other teams submitted plain HTML pages with amazing insights. The winning solution excelled in the insights and functionality of the dashboard. The design of the page was at the level of other submissions, but they displayed all the main financial indicators and did all sorts of distribution analysis of the given data. They analysed the data and plotted it in many ways, allowing for in-depth analysis and interpretation. These creative visualisations of the data showed things no other group had seen.



Word from Aspect: Hakeem Gbadebo, Investment Solutions

The challenge with large datasets is being able to answer questions rapidly as well as to derive actionable insights. In the case of investment management, this transparency is key for research and innovation as well as building partnerships with clients.

A good visualisation tool needs to:

- Connect to data from many types of sources (e.g. databases, files, websites, maps, APIs and codebases)
- Automatically detect relationships between datasets and incorporate time intelligence for time series datasets
- Clean, transform, model and optimise data efficiently without slowing down the machine
- Create unique and original visuals to captivate audiences
- Facilitate sharing of results in many formats (e.g. pdf, ppt, HTML)

Data visualisation is evolving with the emergence of AI tools such as natural language processing, which allows a user to ask the tool questions like we would to a person.



The Challenge

- Predict whether a single ticker's price will go up or down in the next sample
- Timeout of 10 seconds to classify all the test data
- Any algorithm and language was permitted

The Tools

- Raspberry Pi (4GB)
- An ARM device running Raspbian OS
- Pandas and SKLearn

The Criteria

- Teams' classifiers had to achieve over 50% accuracy on unseen test data to be considered
- Teams then ranked based on the time taken to classify

6

Low Latency Challenge: Winning Solution

The winning solution, with the minimal response time, was a very neat and simple C++ program which always guesses that the ticker's value will increase. This team's earlier submissions attempted a more complicated predictor, written in Python and using logic regression with SKLearn. The final solution was audacious, and gambled on the chance that the majority of ticks would continue to increase in value in the unseen test data, as they did in the training data.

The winning solution was quite similar to one of the example solutions we provided. Even a relatively basic attempt can be successful. Is this strategy applicable in the real world? Absolutely not! It serves as a salient reminder of the importance of optimised execution.



Word from Aspect: Dr. Ilija Zovko, Senior Researcher

While Aspect does not trade high frequency strategies, understanding latencies in execution algorithms is nonetheless important. This helps in limiting costs and being taken advantage of by faster market participants.

For certain programmes, efficient execution significantly contributes to their success. This is achieved both by employing algos based on latest market microstructure research and by receiving relevant market information in a timely, low latency fashion.

At Aspect, we measure trading costs at transaction level. While trade cost reduction is a never-ending endeavour, understanding the cost details helps us make rational decisions when choosing the area of execution improvement we want to tackle next.

