

Market Abuse Surveillance TechSprint (July 2024) video Transcript.

Team 4. Spring Autumn (KCL).

Delegate 1

Ladies and gentlemen, we are proud to present our work about market abuse surveillance with AI identical in economic physics. So with the evolving electronic trading market, the traditional work based method may not adopt timely and identify the complex market abuse. And so we propose a method incorporating economic physics and machine logic model to enhance the existed surveillance tools or propose you to test whether the economic physics methods improve the performance of microabrios detection.

So the basic market measure include price volume because that volatility and however there are merely the plan description on the results of market dynamics and actually on a microscopic scale the macro dynamics are shaped by price jumps and they are driven by the older dynamics in the micro to microstructure and under the behaviour of the orders. So our research is based on the six months transaction reporting data of UK equities market selected from the company from the day the company has no release.

So where are you in? We're using a merge data set of transaction reporting data and the new speed. Yeah. So in our model, we concept the order of physical particle with the motion A1 dimensional price axis and and we define the price difference of a price jump and the time spent at the previous price and the distance and type of a micro movement. And we treat the other side as particle mass and with with the data we can further depend and derive the velocity, momentum and the kinetic energy of of every single orders. And then we can further aggregate them to to compute the the general market methods.

Delegate 2

So our purpose is to detect anomalies related to market abuse using a large unlabeled data set. And for this we chose the isolation forest algorithm. And on the left hand side diagram you can see the logic of this method. The method The model detects anomalies by recursively splitting the data into two sets based on their values until the maximum tree depth is reached.

The shorter the average past length, the more likely a point is an anomaly. And on the right hand side plot, you can see how we preposes the data based on news release time. We only select news release during the market opening hours from 9:30 to 5:00 PM and cite them as the origin in the timelines for

trades in each company each day. After modelling, we average the scores for each trades as there might be more than one news, even more news. Sorry, one news event per day per company.

Delegate 1

So in this plot, there are only a plot of price and the different physical methods of different companies. And you and you also study we Yeah. And you Please note that the pink colours interval by the pink colours are the equivalent to the time where there is a news release in this company. And the wider the interval, it means there are there are many news release at the same time. Yeah. And we found that the larger the stock market cap, the more likely we will observe a similar trend in the price and the accumulating momentum, which is the first plot and this the third plot.

However, there are also some companies that show a different paper in the past and ad the digital methods. There are also some other funny facts that we found during our study for example, but you know about this, this the news release, we observed the CFO's right before this new news release in the cumulative momentum and the price did drop actually news release and then the price recover and move up again along with the by force.

Delegate 2

So this plot compares model results showing the anomaly percentage detected by the models trained on the different parameters trained on the same parameter setting but with different features. So the blue bars represent the results of the model train and the original features like price, quantity and tree time from the given data site from FCA. The white bars represent the result of the model train, are a combination of original features and the physical measures which produced by our our physical model.

You can see that the white bars detect some subtle anomaly patterns that might be overlooked by the traditional method. And this is the two plots provided examples of trades for one company. The original model detects 100% of the 100% of this trade as normally as shown in the left hand side as the totally great block. Well, the physical combined model detect more reasonable anomaly which you can see from the right hand side part. There's only a few great block over there in the in the plots. This indicates that the physical measures reduce the false positive of the original model.

Delegate 1

So in general, our measure of the combined econo physics and the machine of the model providing innovative insights from the angle of micro to micro structure and order dynamics and enhance the understanding of micro dynamics and improve the micro array of detection.

At this stage, we we're only using the transaction reporting data and the newsability data without the order book data because it's the first time we got in touch with such data set.

We didn't, we didn't have time to the construct the order book, but in the future we may introduce the order book data with the information of order submission calculation and order book depth in order to develop the completed kinetic model for the linked order book. Slight malfunction, so I'll stick with the hand mic also. Thank you so much for that very interesting use concept of physics for this. Yeah, Questions anyone, Right at the back, well, two at the two towards the back.

Thank you.