COVID Information Commons (CIC) Research Lightning Talk

Transcript of a Presentation by Evelyn Yemurai Zhou (University of South Africa), July 26, 2023

Title: Advances in Machine Learning Explainability to Contextualize Equity Market Sustainability in South Africa During the COVID-19 Era

YouTube Recording with Slides

Summer 2023 CIC Webinar Information

Transcript Editor: Lauren Close

---

Transcript

*Slide 1*

Evelyn Yemurai Zhou: Thank you so much, Lauren, and greetings to you all. As Lauren said, my name is Evelyn Zhou, and I am a financial management student from the University of South Africa. The title of my research is "Advances in Machine Learning Explainability to contextualize Equity Market Sustainability in South Africa During the COVID-19 Era."

*Slide 2*

For this presentation, I'll start by looking at the research objectives, methodology, results analysis, then go on to conclusions and future research, then end by acknowledgments.

*Slide 3*

As a financial management student, I was so fascinated in trying to understand and assess the impact of COVID-19. The easiest way to do this was to look at the stock markets. Therefore, I looked at the Johannesburg stock market, which is the largest stock market in the African continent. We do have a number of - we do have 353 companies listed at the Johannesburg Stock Exchange, but for someone to quickly have a glimpse of what really transpires in the JSE - one method is to look at what we call the Johannesburg Top 40 Listed companies. By just assessing the the top 40 listed companies, they constitute more than 80 percent of the market capitalization. Therefore it would be easier to really see what really happens at the Johannesburg Stock Exchange. The first objective was for this research was to employ [inaudible] in machine learning algorithm to assess sustainability effects attributable to the impact of COVID-19 on the stock returns of the JSE top 40 listed companies, which, for this year, 2023, was coming up from the eight industries although we do have 10 Industries in general.

Another objective was to examine the interpretability and robustness of machine learning in trying to estimate - by implicating explainable artificial intelligence.

*Slide 4*

For data collection, I managed to have to find the closing stock prices for the JSE top 40 listed companies and the data was from January 2017 to September 2022. After gathering data, I ran 40 OLS regression models and analyzed the results. Then I ran also 40 machine learning modules then did post-work analysis using SHAPELY values and for global interpretability I did use SHAPELY heat maps. Then I reviewed and I explained the results. I had quite a number of independent variables - that is Johannesburg  Top 40 index, South African volatility index, COVID index, sustainability indices (that is environmental social and governance). The parameter settings where data transformation was STD-1. Training was 33% of observations. Validation: 67% of the observations. Radius was set at 1.0. Then, the error minimization rule was the GCV.

*Slide 5*

Using the Akaike Information Criterion AIC) in mean squared error the best three models were for the three sectors: basic materials, financial services, and consumer services industries. These companies are Anglogold Ashanti (ANG), which is the third largest gold mining company in the world measured by production. I reselect Financial Services industry - there was there is a Standard Bank which is one of the largest financial services groups in South Africa. It is Africa's leading lender in terms of assets. And it selects Clicks for the consumer services industry. This is the leading pharmacy organization and so part of the healthcare group, which is part of the consumer services industry. Looking at my results, OLS and RANN results produce identical parameter settings except for Standard Bank, which is a financial services company. For Standard Bank, during COVID - this can be attributed to this negative sign parameter for COVID - or Standard Bank, it can be attributed to the fact that during the pandemic, COVID pandemic, not necessarily by a Standard Bank alone but most of the financial services companies were - offered - started offering relief services. That is, in terms of ATM cash withdrawal fees, monthly fees to their clients. And also they did waivers in terms of relief in terms of home loans, student loans, and other financial portfolios.

*Slide 6*

For Anglogold Ashanti and Clicks, they were influenced positively by COVID-19. This can be attributed to the fact that during times of crisis and pandemics like COVID-19, most of the basic materials companies, since they trade with the precious metals, people tend to prefer and buy shares in companies which deal with precious materials most of the time. For Clicks, it is a consumer services provider which was declared to be an essential service provider. So they were not really impacted by COVID-19, the companies in that industry, because they were operating full-time. People always buy consumer services, goods, so they were not really closed. For markets, that is for JTOPI, another variable there - the OLS results implicated the negative  - a positive result relationship.  What is the machine learning modules negative relationship result the negative relationship with the stock returns of these particular companies. For South African

Volatility Index - normally the financial theory postulates an inverse equity price volatility relationship. Yet an exception was for Standard Bank, a financial services provider. For sustainability results (ESG), OLS and RANN results had mixed parameters. Most of the companies from the basic materials industry were impacted severely by governance, by the governance in this because most of those companies were not allowed to operate because they were not declared to be essential service providers. Whereas for financial services industry - companies they were allowed to operate so they were impacted by the social indices - unlike the basic materials companies. For Clicks and other consumer services, it was environmental indices which had quite significant impact on the stock returns of this particular companies.

*Slide 7*

These graphs are for Clicks, a pharmaceutical company. This is to just give you a clear picture of the OLS and the difference between OLS regression and K4-RANN. That is, machine learning modules. As you can see from these graphs, the precision of the K4-RANN machine learning model is quite precise. It's amazing compared to that of OLS. To the right, I do have a network chart with the red line showing the features which had negative impacts on the stock returns of Clicks. Then, the gray lines are showing the features which had positive impacts on the stock returns of this particular company.

*Slide 8*

Moving on, this is SHAP heat maps. For this, I selected three companies - the best performing three companies in three different Industries. Anglogold Ashanti, Standard Bank, and Clicks. The blue line is showing the start of COVID-19 and the black bars to the end of each graph are showing the overall contribution of that feature to the stock returns of this particular companies. For Anglogold Ashanti, as you can see, COVID, precious metal index, and JTOPI are the variables - are the features which had quite significant impact on the stock returns of this particular company. For Standard Bank, a financial services provider, South African Volatility Index, social and environmental indices are the features which had a significant impact on the stock returns of this particular company. For Clicks, the the company at the far right - the precious metal ratio, South African Volatility Index, and COVID are the features which had a significant impact on the stock returns of this particular company. They were not really affected the same, but it was different across the industries.

*Slide 9*

So what I can say is the conclusion from my research is that the performance of complex of complex stock markets modules based on the traditional parametric statistical method is very weak when applied to developing economies like South Africa, and Africa in general. So it will be best for researchers and policymakers to use novel learning - novel machine learning models in trying to develop and produce reliable explainable results when applied to market instruments. For future research, it should be really crucial to use advanced machine learning techniques like deep learning and gradient boosting to compare and contrast results to identify the best performing algorithm for stock market predictability and also to increase the accuracy of the

expected rate of return predictions for traded instruments and stock indices that relies on machine learning will be really encouraged.

*Slide 10*

I just want to end by acknowledging and giving a special thanks to my academic mentors in Professor Gordon Dash and Professor Nina Kajiji, both from the University of Rhode Island in the United States. And I also want to extend my deepest appreciation to the NKD-Group which provided me with access to WinORS software package which I used in conducting this research.

*Slide 11*

I just want to end - these are some of the references which I used for this particular presentation

*Slide 12*

And with this I thank you!