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

Evelyn Yemurai Zhou

College of Economic and Management Sciences
University of South Africa
Preller St, Muckleneuk Ridge,
Pretoria, 0002
South Africa
62330985@mylife.unisa.ac.za

Abstract

The Covid 19 pandemic exacerbated global economic and equity market uncertainty globally with as much as a 30 percent decline recorded in Europe and the United States. Despite the overreach of the COVID-19 pandemic, businesses and industries were not adversely affected similarly. The dissimilarity of COVID-19 impact was also in evidence in national stock markets. As such, this research proposes a machine learning method – an artificial neural network (ANN) – to model the prediction of stock returns for large firms trading on the Johannesburg stock exchange (JSE) during the COVID-19 regime. Until the advent of COVID-19, sustainability was a global business agenda. In a novel approach, the current research study was augmented to include environmental (E), social (S), and governmental (G) policy features to capture sustainability value. Accordingly, to efficiently predict stock returns from firms traded on the Johannesburg Stock Exchange (JSE), the current research augmented traditional efforts to include a COVID-19 feature and separate E, S, and G factors. The study also proposes a novel extension to the literature with the implementation of 'explainable' artificial intelligence (XAI) to assess the trustworthiness of estimated model results. Overall this study revealed that COVID-19 had a mixed impact on the performance of selected top 40 shares trading at JSE. The importance of the separated ESG factors was also shown. For example, Aspen, a leading African pharmaceutical company, experienced negative elasticities due to COVID-19 and varying impacts attributable to ESG factors.

Keywords: Return Prediction, COVID-19, Sustainability, *Explainable Artificial Intelligence (XAI)*

1. Introduction

Epidemics and pandemics usually threaten the ongoing vitality of affected communities as they impose different costs on different entities and populations. Inter alia, this includes medical and productivity costs [14]. In late 2019 the world woke up to the news of a novel disease – COVID-19. In a short space, the disease morphed from being an epidemic to one of the worst pandemics in recent memory. Like previous global crises, the pandemic led to massive economic uncertainty and caused severe global challenges [7,12]. The pandemic further threatened the attainment of the United Nations (UN) Sustainable Development Goals (SDGs), primarily relating to goals 1, 3, and 8; goals focused on poverty reduction, good health, and inclusive economic growth, respectively [1,13].

Global stock markets likewise experienced a significant decline during the early stages of the COVID-19 pandemic [4,10,14]. This research investigates the impact of COVID-19 on the return-generating process of the largest South African companies trading on the JSE using machine learning methods. Specifically, the study implicates an artificial neural network (ANN) to model stock valuation basics. Predicting stock prices and their associated returns is essential for decision-makers who want to understand a stock's market efficiency and intrinsic value to allocate invested wealth efficiently. But, it is also well understood that alternative AI prediction models are not equally trustworthy. Hence, it is now commonplace to explain any bias that could dampen the trustworthiness of the model's prediction. Model trustworthiness and fairness are addressed by applying an 'explainable AI' (XAI) technique. By harnessing the trust precepts of XAI, this study extends previous studies that examined the pandemic's effect on JSE stock returns [8,9,11]. The remainder of the paper is organized as follows, section 2 details the methodology adopted, section 3 presents the findings and discussions of the results, and section 4 concludes the study.

2. Research Methodology

Following the methodology of previous studies [2,3,5], the study gathered secondary data from the JSE stock market. The automated mining function of the WinORS-AI software platform [15] (aka WinORS) was used to acquire time-series data and necessary data and statistical analysis. The data consists of the monthly closing prices for the largest 40 JSE-listed firms from Jan 2017 to September 2022. As per Equation 1, the following equity pricing model provides the method for the study:

$$P_{AGL,t} = \beta_1[P_{JTOPI,t}] + \beta_2[P_{SA-VIX,t}] + \beta_3[P_{COVID,t}] + \beta_4[P_{E,t}] + \beta_5[Pr_{S,t}] + \beta_6[P_{G,t}], \quad (1)$$

where, JTOPI is the market index, SA-VIX is the volatility index that provides an indication of the market sentiment on the JSE, and COVID is a global pandemic proxy computed as a ratio of reported COVID-19 deaths to COVID-19 confirmed diagnoses. Model sustainability effects are expressed through the pricing factors for the E, S, and G global indices. Returns were computed by taking the log difference of prices, $r_{j,t} = ln(P_{j,t+1}) - (P_{j,t})$. Log returns are used for two fundamental properties: time additivity (consistency) and normality. The specific WinORS-hosted RANN selected for this study was the K4-RANN. Empirical estimates for both OLS and the RANN were obtained by rewriting equation 1 to express the valuation function in the returns of the data as expressed in equation 2:

$$r_{AGL,t} = \beta_1[r_{JTOPI,t}] + \beta_2[r_{SA-VIX,t}] + \beta_3[r_{COVID,t}] + \beta_4[r_{E,t}] + \beta_5[r_{S,t}] + \beta_6[r_{G,t}] \quad (2)$$

The study continued with fitting equation 2 for each of the JSE top 40 companies using both OLS and the ANN methods. Before executing the K4-RANN models, the data were transformed by the de-correlating zero-phase component analysis (ZCA). After scaling, the total observations were split into two sets – training (33%) and validation (67%). The radius parameter was inclusively set to 1.0 for Gaussian activation functions. Generalized cross-validation (GCV) error minimization and Akaike Information Criterion (AIC) statistic guide the solutions. Three of the 40 models solved are presented next to provide a comparative analysis of results and predictions.

3. Results

3.1 Estimated Model Effects

The table 1 and 2 below show the model results from OLS and the K4-RANN techniques. Given the computed MSE, the Aspen Pharmacare (APN), Sasol Ltd. (SOL), and Mondi PLC (MNP) models were selected as focus firms for the following analysis. Aspen Pharmacare Holdings Limited is the largest South African multinational pharmaceutical company. Sasol Limited is an integrated energy and chemical company based in Sandton, South Africa. Lastly, Mondi Plc is an international packaging and paper group with about 100 production sites across more than 30 countries, predominantly in Europe, North America, Russia, and South Africa.

Table 1: Multivariate Regression Analysis Output

	Model Features/Variables						**Parameters**	
	JTOPI	**SA-VIX**	**COVID-19**	**E**	**S**	**G**	**MSE**	**R^2**
APN	0.51	-0.75**	-0.75	0.95	-0.55	-0.85	0.10	0.57
SOL	2.55***	-0.88***	-30.04***	1.24	-1.87	1.47	1.63	0.25
MNP	0.20**	-0.09*	5.08***	0.10	-0.56	-0.56	0.05	0.37

*The model intercept was omitted as the results are obtained for comparability with the machine learning technique. NOTE: ***significant at 1% level of significance, ** significant at 5% level of significance, *significant at 10% level of significance.*

Reference is made to table 1. The results show that only the South African Volatility Index had a significant negative impact (at a 5% significance level) on the stock returns of Aspen. Importantly, the OLS results find the other drivers, including COVID-19, had no significant effect on stock returns. These results contrast with the other two companies, which were significantly impacted at a 1% significance level by the COVID-19 feature. Interestingly, the impact was negative for Sasol and positive for Mondi, an indicator of the pandemic's varying impact on companies operating in different industries. In table 2 below, the results of applying the K4-RANN are presented.

Table 2: K4-RANN Analysis Output

	Model Features/ Variables						Parameters	
	JTOPI	**SA-VIX**	**COVID-19**	**E**	**S**	**G**	**MSE**	**R^2**
APN	-0.05	-0.10	-0.13	-0.12	-0.12	-0.13	5.65e-04	0.99
SOL	0.32	0.32	0.30	0.31	0.29	0.29	6.16e-03	0.99
MNP	0.11	-0.02	-0.08	0.02	-0.06	-0.09	1.34e-05	0.99

Based on comparative MSE (fitness) and *R-squared* values, model performance decidedly favors the nonlinear mapping of the ANN. K4-RANN mapping of the Covid factor identifies Aspen as the most negatively affected (-0.13) firm. The Covid factor elasticity implies that a 10 percent increase in the COVID-19 index will result in a decline of 1.3 percent in the company's expected return. The negative coefficients for the pharmaceutical company that was Africa’s sole producer of vaccines are not surprising, considering Aspen Ltd had a business footprint covering more than 150 countries, each with a unique sovereign response to the pandemic [6]. By contrast, based on model weights, Sasol is the only focus firm to experience beneficial stock return effects from the Covid factor (0.30). With a 10 percent increase in the COVID-19 factor, the model predicts a 3.1 percent increase in the expected return of the company's share price. Interestingly, Sasol experienced roughly equal and positive effects among the three focus firms across all features.

The RANN predictive ability chart (figure 1) visually compares actual and predicted stock returns over the training and validation data sets. The network feature map (figure 2) provides a view of model mapping across features and neurons to obtain the predicted output.

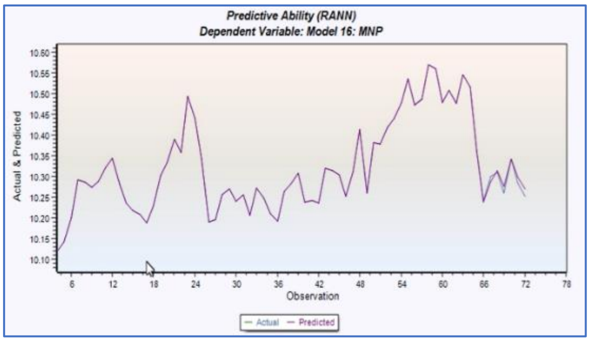

Figure 1: Predictibility for MNP

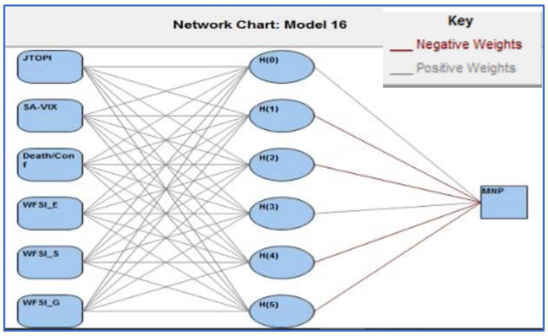

Figure 1: MNP Feature Map

A review of model parameters confirms the visual interpretation. Table 2 presents the overall model performance metrics. Each model shows a very low error and high (pseudo) *R-Square value*. Because it best demonstrates the role of positive and negative predictor effects, the feature chart for the Mondi model is displayed in figure 2. On this chart, the negative weights display in red color and the positive ones in gray color.

3.2 The Predicted Returns Model: AI Explainability

Interpretable AI in WinORS is based on the computation of SHapley Additive exPlanations (SHAP values). The literature recognizes the technique for its fair approach to allocating value from an ML estimation. The XAI results in this study highlight the ranked explanatory effect attributed to each of the six features with an event date set to the 10^{th} month of the COVID-19 pandemic (the period following lockdown measures in South Africa). The XAI analysis for Mondi and Sasol is not shown for governed brevity. The following SHAP analysis features provide a detailed review of XAI findings for Aspen Phamacare.

The SHAP values summary plot (figure 3) and heatmap (figure 4) for Aspen Phamacare paint an informative picture of the model features. Surprisingly, the covid factor ranks below all other features in this model. The summary plot depicts the dichotomous structure of this feature. High SHAP values are associated with decreased functional, *f(x)*, value. The associated heatmap confirms the opposite. The functional, *f(x),* increases when the Covid effect is low to absent (blue). The SA-VIX index was the dominant feature of the AI model. For this feature, high SA-VIX SHAP values produce low expected returns.

The heatmap provides a confirmatory analysis. The black vertical bar for SA-VIX is the highest, and the function value, *f(x)*, decreases when the heatmap is red. The G factor is likewise explanatory. As the second-ranked feature, the plot demonstrates how high G SHAP values counteract the positive effect induced by SA-VIX. In fact, high negative SHAP values have a decidedly negative impact on return generation. At the same time, the S and E indexes also show high negative SHAP values, which are associated with a reduction in the firm's predicted expected returns.

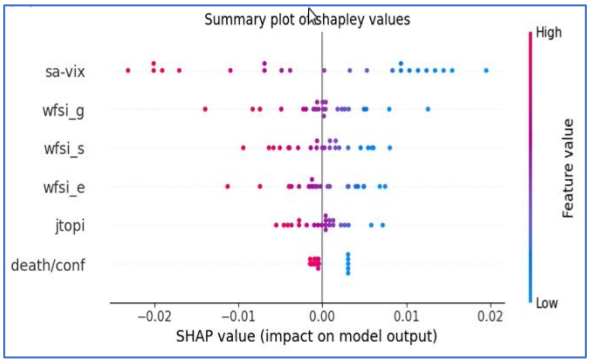

Figure 3: Aspen Pharmacare Summary Plot

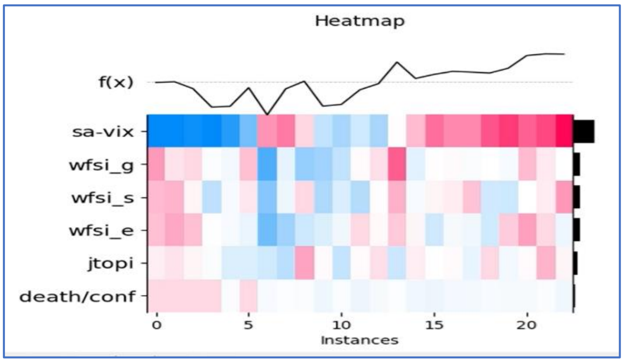

Figure 4: Aspen Pharmacare Heatmap

4. Conclusion

The primary purpose of this study was to use an ML technique to model the stock returns of large companies trading on the JSE over the COVID-19 pandemic period. A secondary purpose included invoking XAI methods to investigate the trustworthiness of all estimated models. Based on K4-RANN estimated model weights obtained by solving equation 2 above, the ML study contributed new evidence on how the existence of a Covid factor can affect firms differently based on their exposure to multinational forces. The ML models also produced solutions with decidedly lower error measures. In addition to the observed accuracy of the estimated solution, XAI feature ranking provided evidence of trustworthy results for ranked features. SHAP value plots identified the Covid factor as a detrimental influence on the ability of firm's to generate equity returns. Finally, the findings help to clarify the importance of conducting an XAI analysis on AI results intended for the burgeoning fintech community.

Acknowledgments

The methods and algorithms described and developed in this paper were guided by Professor Gordon Dash, a professor in the Computational Finance and Decision Sciences Faculty at the University of Rhode Island USA, and Professor Nina Kajiji, a Principal of the NKD Group Inc and an adjunct associate professor in the Computer Science and Statistics Department at the University of Rhode Island USA.

REFERENCES

[1] Dash, G. H. and Kajiji, N. 2021. Behavioral Portfolio Management with Layered ESG Goals and Ai Estimation of Asset Returns. *Available at SSRN 3953440*, Article ID.

[2] Davis, S. J., Liu, D. and Sheng, X. S. 2022. Stock prices and economic activity in the time of coronavirus. *IMF Economic Review*, 70 (1): 32-67.

[3] Dogan, O., Tiwari, S., Jabbar, M. and Guggari, S. 2021. A systematic review on AI/ML approaches against COVID-19 outbreak. *Complex & Intelligent Systems*, 7 (5): 2655-2678.

[4] Grima, S., Özdemir, L., Özen, E. and Romānova, I. 2021. The interactions between covid-19 cases in the usa, the vix index and major stock markets. *International Journal of Financial Studies*, 9 (2):1-19

[5] Gu, S., Kelly, B. T. and Xiu, D. 2019. Empirical asset pricing via machine learning. *Chicago Booth Research Paper*, Article ID(18-04): 2018-2009.

[6] Kew, J. 2020. Aspen's Covid-19 flop bodes ill for Africa vaccine making drive news24

[7] Khan, K., Zhao, H., Zhang, H., Yang, H., Shah, M. H. and Jahanger, A. 2020. The impact of COVID-19 pandemic on stock markets: An empirical analysis of world major stock indices. *The Journal of Asian Finance, Economics and Business*, 7 (7): 463-474.

[8] Makatjane, K. and Moroke, N. 2022. Examining stylized facts and trends of FTSE/JSE TOP40: a parametric and Non-Parametric approach. *Data Science in Finance and Economics*, 2 (3): 317-344.

[9] Mbatha, V. M. and Alovokpinhou, S. A. 2022. The structure of the South African stock market network during COVID-19 hard lockdown. *Physica A: Statistical Mechanics and its Applications*,1-19

[10] Mokoena, S. and Nomlala, B. 2022. Analysis of the relationship between COVID-19 and the stock market performance in South Africa. *International Journal of Finance & Banking Studies*, 11 (2): 25-33.

[11] Thangamuthu, M., Maheshwari, S. and Naik, D. R. 2022. Volatility Spillover Effects during Pre-and-Post COVID-19 Outbreak on Indian Market from the USA, China, Japan, Germany, and Australia. *Journal of Risk and Financial Management*, 15 (9): 378, 1-15

[12] Tiwari, A. K., Abakah, E. J. A., Karikari, N. K. and Gil-Alana, L. A. 2022. The outbreak of COVID-19 and stock market liquidity: Evidence from emerging and developed equity markets. *The North American Journal of Economics and Finance*, 62 (1):1-25

[13] United Nations Department of Economic and Social Affairs Sustainable Development (UNDESA). 2022. THE 17 GOALS | Sustainable Development (un.org)

[14] Vengesai, E. 2022a. COVID-19 and Stock Market Volatility in South Africa: A Cross-Sector Analysis. *Asian Economic and Financial Review*, 12 (7): 473-493.

[15] WinORSe-AI 2021 (Version V21). The NKD Group, Inc., Providence, RI: https://www.nkd-group.global.