THE CURIOSITY CUP 2025

A Global SAS® Student Competition

Evaluating the Impact of the AI Revolution on Stock Market Performance: A Comparative ARIMA Analysis

Aleksandra Czuszek, Szymon Czuszek, Mateusz Medoń, Mateusz Zastawny
University Of Economics In Katowice (LoveForStonks)

ABSTRACT

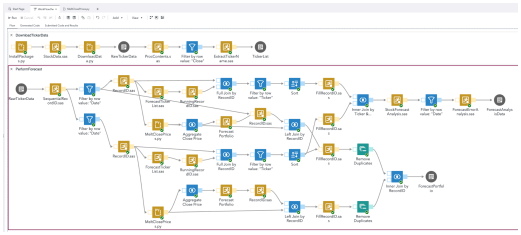
The rapid advancements in artificial intelligence, exemplified by technologies like ChatGPT, have disrupted multiple sectors, including technology, semiconductors, cybersecurity and healthcare (Gill et al., 2022). These shifts are expected to significantly influence stock market performance, presenting opportunities and challenges for investors.

This project evaluates the impact of the AI revolution on stock price predictions by comparing ARIMA-based forecasts for 18 companies across diverse industries, using December 2022 as the cutoff date. Statistical tests and forecast error comparisons will be conducted to uncover patterns of stock price disruption associated with the rise of AI-driven technologies.

The primary objective is to provide valuable insights into how AI advancements shape financial markets, particularly their effect on predictive accuracy and market dynamics (Ohana, Ohana, Benhamou, Saltiel, & Guez, 2021). Leveraging the automation-friendly capabilities of SAS® Studio, this project is designed to be reusable for analyzing different periods, sectors, or Black Swan events—whether those events occur in the future or have already shaped market behavior in the past (Phadnis, Joshi, & Sharma, 2021).

INTRODUCTION

The recent surge in artificial intelligence, spearheaded by the release of ChatGPT, has disrupted industries worldwide.



Display 1. SAS® Viya Analytic Process

As AI-driven technologies continue to emerge, their impact on the financial sector—particularly the stock market—remains a critical question (Golić, 2019). This paper investigates the influence of AI on stock predictions by comparing ARIMA-based forecasts for 18 leading companies before and after the AI revolution. Leveraging historical data and statistical tests, the study evaluates whether AI has caused significant shifts in stock behavior and forecast accuracy. Identical models were trained for both periods to ensure comparability, with the pre-AI model benefiting from a larger dataset. However, the extended out-of-sample period for the pre-AI model poses a challenge, as rapidly evolving market dynamics post-AI may have introduced forecasting inaccuracies. The forecast period for both models spans the post-ChatGPT era, concluding at the end of 2024, enabling a direct,

observation-by-observation comparison of performance. The analysis was conducted using a robust analytical process in SAS® Studio. This analytical process seamlessly integrates traditional SAS® languages like SAS® 4GL and SAS® Macro with Python and low-code tools available exclusively through the SAS® Studio. Figure 1 illustrates the end-to-end process designed for analyzing ARIMA forecasts before and after the recent AI boom (Haigh, 2024). The use of SAS® analytic process allowed the team to leverage the strengths of multiple programming languages in one, end-to-end flow. Python was used for establishing an API connection with the Yahoo Finance stock database, while SAS® was used for statistical analysis and transformations.

Loading data into SAS® and data preprocessing

In order to load stock data into the SAS® environment in a reproducible manner, a Yahoo finance library using the API capability was employed through a Python tool. The first step was installing the Python packages needed, with an executable command using pip. Further on, a SAS® datalines command was used to create a portfolio of AI-related stocks, which were displayed with a proc print procedure, as portrayed in Figure 2.

The portfolio can be manually modified, allowing analysts to alter the analysis. Once selected, it is processed through a Python tool that fetches daily stock data from Yahoo Finance, generating 1000 records. More selected tickers result in more columns. The data, including date, Close, Low, High, Open, and Volume, is stored for later queries.

A PROC CONTENTS step filters columns containing "Close," and a custom SAS® script extracts tickers for further analysis. In the second flow stage, PerformForecast, records are labeled, and the dataset is split at December 2022 to train ARIMA models on Before AI Revolution and After AI Revolution periods, each with 500 observations.

SAS® macros automate forecasts for both individual companies and the entire portfolio. The final dataset is structured for easier analysis, with key columns like MainRecordID, ForecastBeforeRevolution, Date, ClosePrice, and ForecastAfterRevolution. Individual company forecasts generate more rows, while portfolio-level forecasts remain compact.

Obs	Ticker	Company	Sector
1	NVDA	NVIDIA	Semiconductors
2	MSFT	Microsoft	Technology
3	GOOGL	Alphabet_(Class_A)	Technology
4	META	Meta_Platforms	Technology
5	AMZN	Amazon	Technology
6	AMD	AMD	Semiconductors
7	INTC	Intel	Semiconductors
8	TSM	Taiwan_Semiconductor	Semiconductors
9	IBM	IBM	Enterprise_AI
10	ORCL	Oracle	Enterprise_AI
11	SNOW	Snowflake	Enterprise_AI
12	TSLA	Tesla	Automation_&_Robotics
13	PATH	UiPath	Automation_&_Robotics
14	PLTR	Palantir	Cybersecurity_&_AI
15	CRWD	CrowdStrike	Cybersecurity_&_AI
16	MRNA	Moderna	Healthcare_&_AI
17	ILMN	Illumina	Healthcare_&_AI
18	PFE	Pfizer	Healthcare_&_AI

Display 2. Chosen portfolio

Forecast results of ARIMA

The pre-AI and post-AI revolution ARIMA(1,0,1) analyses for the "Close" variable use datasets of 500 observations. Both exhibit significant autocorrelation ($p < 0.0001$ across all tested lags) and show strong AR(1) parameters (pre: 0.9971, post: 1.0000, both $p < 0.0001$), while MA(1) remains insignificant in both cases. This comparison is crucial as it establishes that the underlying time series dynamics remained consistent across both periods, allowing us to isolate and assess the impact of the AI revolution on forecast deviations rather than changes in inherent market behavior.

For the pre-AI period, the model mean is 2549.13 (SD: 480.14), with MU value estimated at 2494.8. The variance estimate from the model's residuals is 2574.46, AIC 5348.63, and SBC 5361.27. Post-AI, the mean rises to 2624.07 (SD: 484.05), MU drops to 1739.0, The residual variance estimate declines to 1490.85, AIC to 5077.48, and SBC to 5088.12. These changes reflect a potential market disruption, emphasizing the significance of ChatGPT's emergence as potentially indicative of a Black Swan event, which may have altered stock price dynamics.

Both models pass residual diagnostics, confirming their adequacy and suitability for effective forecasting. The AR(1) dominance highlights strong autocorrelation, further supporting the use of these models for forecasting the 'Close' variable.

Data Analysis and Results

After assessing the goodness of fit of the ARIMA models in both periods, we ensured that the results could be compared to themselves (Contreras, Espinola, Nogales, & Conejo, 2003). In this section, we will go over the results of the code used to analyze the portfolio forecasts, individually for every company. To visualize the data, a combination of SAS® visualization techniques was employed. The results gave us great insight into the differences between each of the companies and forecast periods.

Correlation & t-test Analysis

The first great insight into the difference between the forecast period for the entire portfolio was a t-test performed based on the value of both forecasts, as seen in Output 1.

The TTEST Procedure

Difference: ForecastBeforeRevolution - ForecastAfterRevolution

N	Mean	Std Dev	Std Err	Minimum	Maximum
14400	-51.0513	106.9	0.8907	-488.0	152.5

Mean	95% CL Mean		Std Dev	95% CL Std Dev	
-51.0513	-52.7973	-49.3054	106.9	105.7	108.1

DF	t Value	Pr > |t|
14399	-57.31	<.0001

Pearson Correlation Coefficients, N = 14400
Prob > |r| under H0: Rho=0

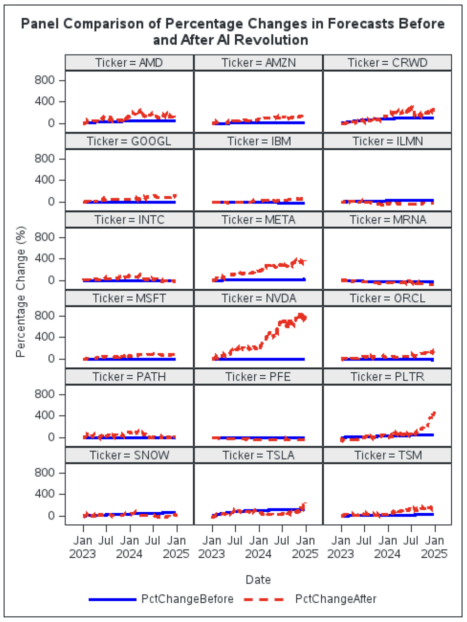
	ForecastBeforeRevolution	ForecastAfterRevolution
ForecastBeforeRevolution	1.00000	0.60244
ForecastBeforeRevolution		<.0001
ForecastAfterRevolution	0.60244	1.00000
ForecastAfterRevolution	<.0001	

Output 1. Proc t-test & CORR

As seen in the t-test, the both forecast differences are statistically significant, which means that the two groups of forecasts differ enough that this is likely not a coincidence. This tells us that there is tangible statistical proof that the emergence of ChatGPT, and the AI Revolution that followed, had a substantial impact on the stock exchange, most notably on the companies like NVIDIA, which benefited greatly from the sale of GPUs, needed to train large language models. The correlation analysis on the other hand, showed increased volatility (higher std dev) suggesting greater uncertainty or stronger fluctuations in stock performance post-revolution. While past forecasts still influence post-revolution forecasts (r = 0.6024), they are not perfectly aligned, indicating a shift in market behavior. This suggests that AI's emergence strongly impacted, and altered stock market expectations, especially in AI related companies.

Close price forecast analysis

Output 2 presents a side-by-side comparison of stock price percentage change in forecasts before and after the AI revolution for multiple companies. Each panel corresponds to a different stock ticker, showing forecasted stock prices over time. As can be observed in the output, the post-revolution forecasts are higher and more volatile. The red dashed lines (ForecastAfterRevolution) generally show higher predicted values compared to the blue solid lines (ForecastBeforeRevolution). This aligns with the statistical analysis where the mean forecast increased from 113.84 to 164.89. Stocks such as META and NVDA show significant increases, indicating a strong impact of AI-related developments. Many stocks exhibit a steeper upward trajectory post-revolution. Companies like CRWD, PFE, NVDA, and PLTR have an increasing spread between the pre-and post-revolution forecasts over time, suggesting an acceleration of expected growth. This could imply that AI-driven innovations have significantly changed investor expectations. Some stocks, such as ORCL (Oracle) and SNOW (Snowflake), show little to no deviation between pre- and post-revolution forecasts. This suggests that the AI revolution had less impact on certain industries like pharmaceuticals. Some stocks (e.g., META, NVDA, ILMN) show greater fluctuations in post-revolution forecasts. This aligns

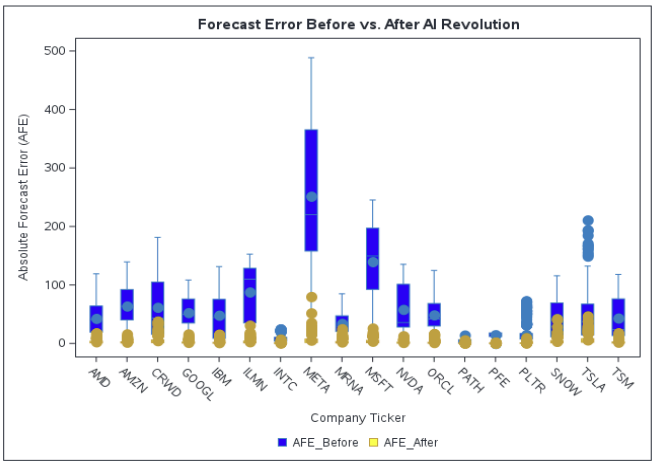


Output 2. Percentage change in stock price - Panel Comparison

with the increased standard deviation (133.91 vs. 81.79) found in the correlation analysis. It suggests that AI-related market conditions introduced more uncertainty in some sectors.

Error analysis

Output 3 shows a comparison of the absolute forecast error before (blue) and after (yellow) AI advancements, sparked by ChatGPT.



Output 3. The forecast error box plot

The absolute forecast error before is significantly larger for most companies, suggesting that forecasts made before AI didn't account for the black swan event, as the one caused by the deployment of the ChatGPT model. The absolute forecast error after values are much smaller, implying that forecasts made after the AI revolution are more precise and reliable. Even though an event as unforeseen as this happened, the trend created by it was stronger, and in turn, easier to capture by the ARIMA model. This trend is visible across multiple companies, with notable improvements for META, MSFT, NVDA, and TSLA, suggesting that for those companies, the impact of the AI revolution was quite significant. Some stocks remained unphased by the AI revolution sparked by ChatGPT, while others completely skyrocketed, leaving them unrecognizable from what was forecasted with data up until December 2022, further proving the black swan theory.

CONCLUSION

The findings of this study provide strong evidence that the emergence of ChatGPT and the subsequent AI revolution created a significant structural break in stock market behavior. The observed discrepancies between forecasts generated before and after December 2022 indicate that traditional forecasting models, which relied on pre-AI market trends, failed to anticipate the rapid growth of AI-driven companies. The substantial reduction in absolute forecast error (AFE) in post-AI forecasts suggests that the impact of AI advancements was largely unpredictable using historical data alone. Notably, companies such as NVIDIA, Meta, and Microsoft experienced the most pronounced deviations, reinforcing the hypothesis that AI-centric firms benefited disproportionately from this market shift. These results support the classification of the AI revolution as a black swan event, as it introduced an unforeseen and transformative change in financial markets.

REFERENCES

Contreras, J., Espinola, R., Nogales, F. J., & Conejo, A. J. 2003. "ARIMA models to predict next-day electricity prices." *IEEE transactions on power systems*, 18(3), 1014-1020.

Gill, S. S., Xu, M., Ottaviani, C., Patros, P., Bahsoon, R., Shaghaghi, A., ... & Uhlig, S. 2022. "AI for next generation computing: Emerging trends and future directions." *Internet of Things*, 19, 100514.

Golić, Z. 2019. "Finance and artificial intelligence: The fifth industrial revolution and its impact on the financial sector." *Zbornik radova Ekonomskog fakulteta u Istočnom Sarajevu*, (19), 67-81.

Haigh, T. 2024. "How the AI Boom Went Bust." *Communications of the ACM*, 67(2), 22-26.

Ohana, J. J., Ohana, S., Benhamou, E., Saltiel, D., & Guez, B. 2021. "Explainable AI (XAI) models applied to the multi-agent environment of financial markets." *In Explainable and*

Transparent AI and Multi-Agent Systems: Third International Workshop, EXTRAAMAS 2021, Virtual Event, May 3–7, 2021, Revised Selected Papers 3 (pp. 189-207). Springer International Publishing.

Phadnis, C., Joshi, S., & Sharma, D. 2021. "A study of the effect of Black Swan events on stock markets–and developing a model for predicting and responding to them." *Australasian Accounting, Business and Finance Journal*, 15(1), 113-140.

ACKNOWLEDGMENTS

This paper would not have come to life if it weren't for our advisor's suggestion to band together and enroll in this competition, so we wish to express our profound gratitude to PhD Mariusz Żytniewski, a university lecturer at the University of Economics in Katowice.

RECOMMENDED READING

- *Base SAS® Procedures Guide*
- *SAS® For Dummies®*

CONTACT INFORMATION

Aleksandra Czuszek

Your comments and questions are valued and encouraged. Contact the author at:

Aleksandra Czuszek

University of Economics in Katowice

aleksandra.rozanska@edu.uekat.pl

Szymon Czuszek

Your comments and questions are valued and encouraged. Contact the author at:

Szymon Czuszek

University of Economics in Katowice

szymon.czuszek@edu.uekat.pl

Mateusz Medoń

Your comments and questions are valued and encouraged. Contact the author at:

Mateusz Medoń

University of Economics in Katowice

mateusz.medon@edu.uekat.pl

Mateusz Zastawny

Your comments and questions are valued and encouraged. Contact the author at:

Mateusz Zastawny

University of Economics in Katowice

mateusz.zastawny@edu.uekat.pl