An Applied Approach to Financial Market Analysis: Forecasting NASDAQ and NYSE Stock Price with Residual Network 2D CNN Models

A Study of Image-Based Prediction Techniques for Enhanced Financial Forecasting

Tucker S.E.

2491092

Supervisor Dr Ali Salimian

A dissertation presented for the degree of
MSc Applied Artificial Intelligence

London
South Bank
University

Computer Science and Informatics
London South Bank University
United Kingdom

December 2024

Abstract

This dissertation investigates the application of residual network-based 2D convolutional neural networks (CNNs) for predicting NASDAQ and NYSE stock prices. The study focuses on transforming traditional numerical stock data into an image, enabling the model to capture intricate spatial and temporal patterns within the data. Unlike traditional methods that depend heavily on quantitative analysis, this approach leverages visual data processing to uncover insights that conventional techniques might miss. Through three experiments on one model, the study demonstrates that image-based prediction models provide significant advantages in addressing challenges such as non-linearity, which are common in the stock market. Experiment 3, which had historical stock prices, indices, and economic data, achieved the highest predictive accuracy, with an R^2 score of 0.99. However, it performed the worst when forecasting future horizons and smaller unseen data. Experiment 1, which only had historical stock prices, had a Mean Squared Error (MSE) of 299.776 when forecasting future horizons, with some stocks having an MSE as low as 0.004 despite having an R^2 score of 0.97 during standard testing, indicating the value of feature engineering and the simplicity of the data. While the stock market is inherently complex, the residual network lays the foundations for future work and improvement despite the limitations of the model, the novel approach, which involves how the data was transformed bridges the gap between traditional numerical methods and the current approaches. Future work could explore re-training the model or incorporating contextual data to enhance predictions further.

Acknowledgements

This dissertation would not have been possible without the invaluable support and guidance of many individuals. The author thanks Dr. Ali Salimian and Francis Babayemi, whose insights and feedback were instrumental in shaping the course of this work. To Dr. Bugra Alkan for teaching the module Industrial Cyber-Physical Systems, the skills gained, especially regarding writing the literature review, were valuable and made the process of doing it at a larger scale easier. Gratitude is also extended to the School of Engineering's Computer Science and Informatics department for providing essential resources, allowing for a thorough exploration of the research questions. Also, a special thanks goes to Vijay Lee from the Business School for imparting foundational knowledge in finance, which was crucial in understanding the topic explored here. Thank you all for your significant contributions to this project.

Contents

List of Figures

List of Tables

1 Introduction

1.1 Problem Statement

In March 2024, the New York Stock Exchange (NYSE) had 2,260 listed companies, with a total market capitalization of 28.4 billion USD, making it the largest stock exchange in the world (Statista 2024). However, despite its vast scale and importance, the NYSE has a long history of significant volatility. Black Monday in 1987 saw one of the most dramatic market crashes. The aftermath of 9/11 led to another steep decline (Jones and Dean 2020 and Elliott et al. 2024), and the 2008 financial crisis, fuelled by the collapse of the subprime mortgage, caused a further drop, with the Dow Jones losing 33% of its value over six months (Kosakowski 2024 and Alvarez-Ramirez, Rodriguez, and Ibarra-Valdez 2020). In the years following, events like Brexit caused disruptions in market liquidity for non-U.S. stocks (J. C. Kim, Mazumder, and Su 2024, Shahzad et al. 2018 and Arshad, Rizvi, and Haroon 2019). At the same time, the COVID-19 pandemic resulted in unexpected gains for companies such as Zoom, Amazon, and AstraZeneca (Klebnikov 2021, Imbert 2020, and D. Lee, Waters, and Mancini 2020). These historical events demonstrate the ongoing challenge of stock market volatility. For investors, policymakers, and analysts, such unpredictability poses a significant hurdle to making informed decisions. Traditional forecasting models often need help anticipating these sudden changes, mainly driven by events outside typical market patterns. This study seeks to address this gap by investigating whether artificial intelligence, specifically a 2D convolutional neural network (CNN), can offer improved accuracy in predicting stock market trends. By transforming numerical data into an image for CNN to interpret, this approach aims to develop a model that is better suited to respond to abrupt market changes.

Figure 1: The NYSE floor the day after Pearl Harbor in 1941. (Kassel 2016)

1.2 Statement of Purpose

This dissertation investigates whether using a residual network 2D CNN to forecast the share price of NASDAQ and NYSE stocks by transforming numerical data into images can improve prediction accuracy for unseen data compared to traditional methods. Transforming stock data into images also provides a new way of interpreting stock market trends. It will explore the possibility of uncovering new patterns that could be overlooked with the use of standard methods.

1.3 Research Questions

1. How do image-based techniques using a 2D CNN compare to traditional numerical prediction methods in capturing stock market patterns?

2. What unique patterns or insights, if any, are revealed through image-based representations of stock market data that are not readily apparent in conventional numerical data representations?

3. To what extent does transforming numerical stock market data into images improve the generalisation of the model across different market conditions?

4. What pre-processing or transformation techniques enhance the effectiveness of the 2D CNN model in predicting stock market indices?

1.4 Overview of Methodology

This study uses a quantitative approach, applying a residual network 2D CNN by transforming historical data into images. The CNN will be trained on historical daily stock prices, which are preprocessed into visual representations to capture potential patterns in the data. After training, the model's accuracy is tested on new data, using metrics like Mean Absolute Error (MAE) to gauge its predictive performance. Chapter 3 will go into more detail on the methodology, including how the data is prepared, how the model is trained, and how its accuracy is measured.

1.5 Rationale and Significance

Stock market forecasting plays a crucial role in helping companies make informed decisions and in some instances, the uncertainty of mergers and acquisitions impacts the share prices. Take Disney's acquisition of Pixar in 2006. This move brought Pixar's creative strength to Disney's portfolio, increasing both Disney's stock performance and revenue growth (Barnes 2008). The opposite occurred with the merger between AOL and Time Warner in 2000, which resulted in a significant decline in price as expected returns never materialised, known as the biggest merger mistake (Nickerson and Rarick 2010). Investors plan and devise strategies to hedge their portfolios to compensate for unavoidable downturns, helping investors feel secure and in control. Investors need to know when to use long to buy or short to sell stocks and securities to maximise returns and minimise losses. Diversifying portfolios is another strategy to spread investments across various assets or industries to reduce risk. If an investment incurs losses, other investments can compensate for the loss to maintain the portfolio's profitability (Elsayed, Nasreen, and Tiwari 2020).

The market serves as a benchmark to measure performance by comparing competitors within the same industry by providing a standard for evaluating financial performance, operational efficiency, and growth potential (Mishkin and Eakins 2016). Organisations can identify their strengths and weaknesses by looking at key performance indicators (KPIs) against industry averages to set goals and make informed decisions to attract new investments (Grybauskas, Stefanini, and Ghobakhloo 2022). Market trend analysis can influence policy makers by offering economic indicators such as interest rates and inflation (Zakhidov 2024). For example, during the 2008 crisis, the stock market signalled extreme economic instability, prompting central banks to cut interest rates and implement quantitative easing to stimulate recovery. Similarly, market drops reflected recession fears during the COVID-19 pandemic, leading to massive fiscal stimulus and monetary easing (Wright 2021). In 2016, after the Brexit referendum, sharp declines in stock markets indicated economic uncertainty, leading the Bank of England to lower interest rates and introduce stimulus measures to support the UK economy (Kenourgios, Dadinakis, and Tsakalos 2020). Managing risk leads to investor confidence, especially when predictions are consistent and reliable. As COVID-19 spread and fears of a major economic downturn grew, investors worried about their portfolios and began to panic sell. Large indices like the S&P 500 dropped by more than 30% in just weeks. The quick drop created urgency for many investors to sell off stocks to limit losses. The uncertainty about how the pandemic would affect the economy really added to the market's volatility leading to a drop in investor confidence, which then drove stock prices down even further (Shu, R. Song, and Zhu 2021).

Accurate predictions are essential for market efficiency, which helps drive a more rational allocation of capital. This empowers investors to make informed decisions and pushes stock prices closer to their intrinsic values. As a result, capital is directed toward companies with strong financial health and promising growth potential. When capital is allocated wisely, it can stimulate economic growth by directing resources toward profitable ventures, which in turn helps shield investors from potential losses. A stark reminder of the pitfalls of poor allocation is the 2008 crisis, where many banks poured money into subprime mortgage-backed securities without a proper grasp of the associated risks (S. Zhang et al. 2021). As housing prices declined, many borrowers defaulted on their loans, causing securities to lose value rapidly, leading to enormous losses for investors and a severe recession (Antoniades 2016).

1.6 What Is the Stock Market

The stock market contains public companies with stock listed on the exchange to be bought and sold on a primary and secondary market. The primary market is where securities like bonds and stocks are issued and sold for the first time, allowing companies and governments to raise capital directly from investors. Unlike the secondary market, where securities are traded between investors, the proceeds from primary market transactions go directly to the issuing company to gain capital (Navyatha and Gaddam 2022). Securities that have previously been issued are traded between investors on the secondary market, facilitating the exchange of securities among investors, with platforms like NYSE and NASDAQ, the main focus of this report (Arseneau, Rappoport, and Vardoulakis 2020 and Maurin 2022). Being a complex system influenced by various factors, the stock market is inherently multivariate. Economic factors such as inflation, interest rates, GDP and employment figures significantly shape investor expectations and stock price movements. Movements in interest rates can affect the cost of borrowing and profitability, while inflation can impact companies' margins (Woode, Owusu Junior, and Adam 2024 and Alzoubi 2022).

Major organisations influence stock market trends through their performance reports and strategic decisions. Positive earnings reports or dividend announcements give investors insight into a company's health (Trinh, Haddad, and Tran 2022). When these indicators surpass expectations, stock prices can surge and the positive sentiment spreads across the sector. However, the opposite may trigger sell-offs, especially when unexpected, affecting individual corporations and related industries (King and Fogarty 2022). The performance of influential companies like Apple can drive trends in the whole market, moving index values and affecting the entire sector sentiment (Kirsch 2024). Volatility is the rapid fluctuations in the price of stocks or market indices over time. Many factors, including investor sentiment and geopolitical factors, influence it.

Figure 2: A photo of the NASDAQ board in 2004 when Google first appeared on the market. (Photo taken by Chris Hondros as cited by Hale 2021)

These factors create extra complexity, uniquely impacting different industries and regions, resulting in price swings. This kind of volatility can mean quick profit opportunities for short-term traders, but for long-term investors, increases risk. When volatility is high, uncertainty follows, which can lead to swift losses or gains. Periods of economic uncertainty often push volatility even higher, making the market unpredictable. (Y. Zhang et al. 2023). Liquidity is another key part of the stock market and speaks to how easily assets like stocks can be traded without making waves in their price. High liquidity means there's a good balance of buyers and sellers, allowing trades to flow smoothly with minimal price impact. Factors influencing liquidity include the number of market participants, trading volumes and market size. Liquidity contributes towards maintaining price stability and ensuring an efficient market. In a liquid market, prices reflect supply and demand more accurately (Abudy 2020).

Risk and return are fundamental to investment decisions. Return refers to the profit investors earn and stocks generally offer higher long-term returns, but the returns fluctuate more widely in the short term due to the inherent risks in the market. In market terms, risk is the uncertainty regarding the difference between the actual and expected returns with the possibility that money can be lost. There are different types of risk, including systematic risk, which affects the entire market, such as economic downturns or interest rate changes, and unsystematic risk, which is industry-specific to individual companies or sectors, volatility increases risk (Mazumdar, D. Zhang, and Y. Guo 2020).

The risk-return trade-off is the relationship between risk and return in the stock market, higher potential returns tend to come with higher risk. Market sentiment, the overall attitude or feeling of investors toward the stock market or a particular stock at a given time, connects with the risk-return trade-off because it influences how investors perceive and respond to risk in pursuit of returns. When sentiment is bullish, investors may become more willing to take on risk, often pushing up prices for riskier stocks, such as those in emerging markets or high-growth industries, expecting substantial rewards. In contrast, bearish sentiment is when investors avoid riskier investments due to fears of losing money and invest in lower-risk stocks like the established companies offering consistent but lower returns. Negative market sentiment increases risk aversion and as a result, riskier assets may experience declines in value as investors prioritise capital preservation over high returns. Sentiment indicators, such as the fear and greed index or put/call ratios, help gauge whether the market is leaning toward extreme optimism or fear, thereby assisting investors in anticipating potential shifts.

Global events, including political changes, natural disasters and economic reports, can immediately influence the stock market, affecting investor sentiment and causing fluctuations in stock prices. Global markets are highly interconnected, and significant events in one area often create ripple effects worldwide. This requires investors to monitor global developments closely and adjust their portfolios to manage risk and capitalise on opportunities. This interconnectedness results in both volatility and opportunities across international stock markets.

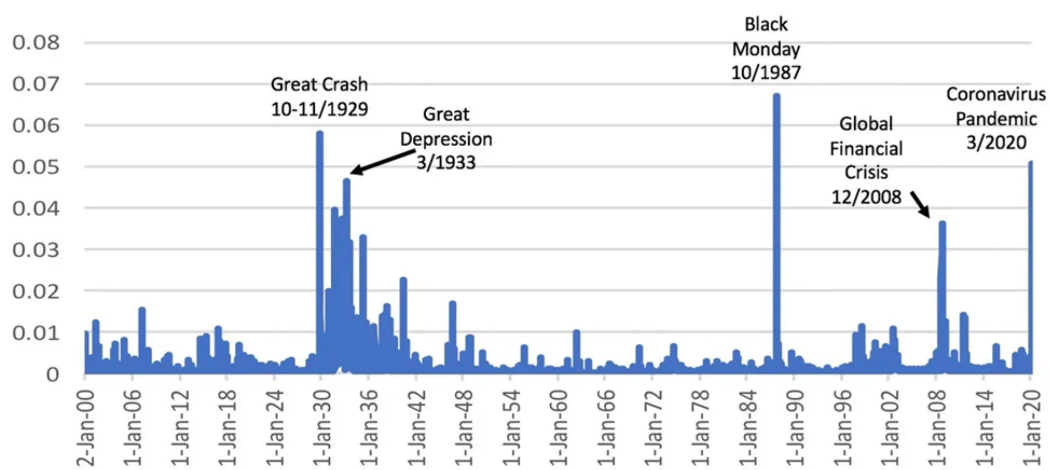


Figure 3: The volatility of the US stock market over 10-day trading periods illustrates the market jumps during some of the famous market shocks. (Baker et al. 2020)

1.7 Stock Market Forecasting Challenges

Stock market forecasting presents significant challenges due to many factors. The Efficient Market Hypothesis (EMH) suggests that all available information is already reflected in stock prices, making beating the market complex. An effective market also ensures the market players are following the rules because there is no advantage within the information to generate abnormal profits (Kelikume, Olaniyi & Iyohab, 2020 as cited by Ehedu and Obi 2022). Stocks are sold at their fair market value on exchanges, which means purchasing at a discount or selling at an inflationary price is difficult making it problematic to surpass the market's performance (Downey and Scott, 2020 as cited by Ehedu and Obi 2022). It has also been said that the efficiency of the market is reliant on the knowledge available, so the EMH has been categorised into three forms: strong, semi-strong, and weak. Strong means that the information available is such that investors can confidently analyse and trade, making quick decisions from new information. The weak form is when prior price information reflects current prices accurately making historical prices insufficient to forecast. Furthermore, the competitive nature of the investor leads to information being kept secret. Random Walk, another market theory, suggests that stock prices move randomly and are not influenced by past price movements or patterns, suggesting that predicting future prices based on historical data is impossible because price changes are random and independent. This randomness makes it difficult for investors to accurately predict the movements of the stock market (Maloumian 2022). Chaos theory states that stock market forecasting is difficult due to the dynamic and complex nature of the market. It suggests that even small, seemingly insignificant changes can lead to many different outcomes, a concept known as the butterfly effect (Debnath 2022). As a non-linear system influenced by several correlating factors, the behaviour of the stock market is unpredictable and can produce outcomes that are hard to predict using traditional methods. To create a model that can accurately predict future stock prices means that the EMH, Random Walk and Chaos theory are false. Overfitting is a common issue, where models are too closely aligned with historical data, capturing noise rather than the right patterns and failing to generalise (Jerez and Kristjanpoller 2020). Since the stock market is influenced by numerous external factors, as previously discussed, these unpredictable characteristics make accurate forecasting efforts a challenge.

1.8 Historical and Current Forecasting Approaches

Predicting stock prices has always been a key focus for forecasters with early methods, including technical and fundamental analysis. Technical analysis relies on the study of historical price and volume patterns, using tools like moving averages and relative strength indices to anticipate potential stock movements. On the other hand, the fundamental analysis evaluates a company's financial health and overall economic conditions to estimate its intrinsic value, aligning with the Intrinsic Value Hypothesis that stock prices eventually reflect a company's true worth (Nti, Adekoya, and Weyori 2019). Over time, statistical approaches such as ARIMA (Autoregressive Integrated Moving Average) and GARCH (Generalised Autoregressive Conditional Heteroskedasticity) models became popular for analysing time series data. ARIMA combines components like autoregression, integration, and moving averages to predict stock price trends (G. Kumar, S. Jain, and U. P. Singh 2021), while GARCH focuses on volatility clustering, making it valuable for managing market risk (He et al. 2021). More recently, advancements in machine learning and deep learning have enabled models like Random Forests, CNNs (Convolutional Neural Networks), and LSTMs (Long Short-Term Memory) to detect intricate patterns and trends, often outperforming traditional techniques (S. W. Lee and H. Y. Kim 2020; Varadharajan et al. 2024). Hybrid approaches, like CNN-LSTM models, improve forecasting by combining methods that identify patterns in both space and time (Mehtab and Sen 2022). The move from traditional methods to these advanced techniques shows how stock price forecasting has evolved to tackle the increasing complexity of financial data.

1.9 Role of the Researcher

The researcher plays a central role in the research journey. This includes regular contact with the supervisor, reflecting back on the research questions, crafting the methodology, and managing the collection of the stock market data and analysis. The Gantt chart will be used to keep the project on track, which outlines timelines and milestones for each phase. Conducting a risk assessment is also essential, as it identifies potential challenges and prepares strategies to address them. Staying objective throughout the process is crucial to ensure that the findings are solid and insightful. Connecting with existing literature is also crucial for providing context and gathering insights that can refine the approach. Interpreting the results and discussing the implications for financial forecasting is a significant responsibility. Throughout the research process, being clear about any biases or perspectives that could influence the study is important to maintain the integrity of the work.

1.10 Researcher Assumptions

1. The research assumes a certain level of market efficiency, implying that all available information is reflected in stock prices, though recognising that inefficiencies and anomalies can occur, which distorts the pricing, such as geopolitical events and sudden economic crises impacting the model's ability to perform on unseen data, especially when the anomalies fall outside the patterns captured by the reliance on historical data.

2. There is an acknowledgement of the impact of behavioural biases, such as bullish and bearish behaviour, on stock market movements, which leads to irrational stock movements. While image-based representations may capture some patterns associated with these behaviours, the model might struggle to differentiate between genuine market trends and noise arising from extreme

emotional reactions or speculative bubbles.

3. It is assumed that the historical data used for analysis are accurate and representative of market conditions.

4. Recognition that no single model can perfectly predict stock market behaviour due to its inherent complexity and the influence of unpredictable external factors.

5. While striving for objectivity, it is acknowledged that the researcher's background and previous knowledge may introduce subtle biases in interpreting the results. These biases could influence how results are evaluated, especially when comparing the CNN-based model to other approaches.

1.11 Organisation of the Dissertation

The organisation of this dissertation begins with a comprehensive Literature Review comprising an introduction, review of literature, conceptual framework, discussion, and summary. Following this, the methodology section details the research approach, setting and context, sample and data sources, data collection methods, data analysis methods, and trustworthiness issues. It will cover the limitations and delimitations followed by the experimental setup, introducing how it will be done and summarising the experimental procedures. The Results section will present the findings, followed by the Discussion that highlights the results, providing a thorough analysis along with considerations for future work. Finally, the dissertation concludes and reflects on learning, summarising possible key insights and the overall learning process throughout the dissertation.

2 Literature Review

2.1 Introduction

The stock market forecasting field has undergone considerable evolution driven by the increasing accessibility of large datasets and the rising complexity of modern financial systems. Within this change, deep learning models such as CNNs and RNNs have gained attention due to their capacity to model complex temporal relationships. Specifically, stock market prediction has seen frequent applications of architectures like LSTM networks, GRU, and CNN-RNN hybrids. A common objective in recent studies is to improve prediction accuracy and model adaptability through varied approaches to data pre-processing, transformation, and evaluation. One of the most prominent challenges in financial forecasting is financial data's noisy, non-stationary nature. However, there is a reassuring solution: the critical role of pre-processing steps. These steps, which range from scaling methods, such as Min-Max normalisation, to more advanced transformations that support chaos theory and data decomposition techniques like wavelet transformation, play a role in preparing time series data for models. They enable the model to handle temporal complexities better to improve their reliability. Articles in this review demonstrate various transformation strategies to optimise the data, including methods such as phase-space reconstruction and the use of Lyapunov exponents for chaos detection.

Evaluation metrics are another focal point, as diverse error measures and statistical indices provide a comprehensive assessment of model effectiveness. Metrics such as Mean Squared Error (MSE), Root Mean Square Error (RMSE), and Mean Absolute Error (MAE) are widely used to quantify prediction accuracy. In contrast, other metrics, like the Directional Change Statistic (Dstat) and Theil's Inequality Coefficient (Theil's U), offer insights into model robustness and directional accuracy. Each evaluation metric is evidence of the model's effectiveness in capturing the nuances in the trends within the stock market and minimising predictive errors, ensuring a thorough evaluation process. Furthermore, the articles shows several novel model architectures to improve prediction outcomes through hybridised methods. For instance, the combination of CNN and RNN layers is popular for its ability to process spatial and temporal data aspects, which is especially relevant in multidimensional datasets such as those used for NASDAQ and NYSE predictions. Hybrid approaches often aim to capture both the short-term fluctuations and long-term trends within the financial market for a more accurate and nuanced prediction. By reviewing these foundational concepts, this review will detail the key advancements and challenges in the use of deep learning for stock market forecasting, how current research methodologies align with and address the core difficulties inherent to financial data resulting in identifying critical gaps and future research directions.

2.2 Review of Literature

2.2.1 Prediction Models and Data Pre-processing

The foundation of stock market forecasting has significantly shifted from traditional statistical approaches to ML and deep learning (DL) methods. Early stock market forecasting models primarily relied on statistical techniques such as ARIMA and regression analysis suited for stable time series data. However, traditional approaches like the ARIMA model operate under certain stability assumptions that limit its capacity to address the complex, non-linear characteristics of the stock markets. Building on traditional statistical approaches, Daryl, Kumara, and Suhartono 2021 examined the use of SARIMA for stock market forecasting, which builds on ARIMA by adding seasonal components to address periodic patterns in data. While SARIMA can handle seasonality better than ARIMA, the authors found that accuracy diminishes in highly volatile or non-linear data. This highlights the need for more adaptive methods, such as ML and DL approaches. In this context, Sharma, R., and Deepa 2022 extended these statistical approaches by incorporating exogenous variables into the SARIMAX model, further adapting SARIMA for forecasting stock indices like the S&P BSE SENSEX and S&P 500. Their study showed that the Prophet model, designed for handling seasonality and holidays, outperformed SARIMAX in predictive accuracy, illustrating the advantages of using additional variables in modelling. Durairaj and Mohan 2022 focuses on enhancing these models by integrating neural network components, specifically with ResNet-ARIMA architectures, to manage stable and dynamic patterns more effectively. This integration seeks to bridge the gap in traditional statistical models by allowing for a more adaptive response to market fluctuations while leveraging the foundational strengths of ARIMA in trend analysis.

ML models gained popularity with increased computational power and data availability. Techniques such as SVMs and Random Forests (RFs) became alternatives to traditional statistical models, showing enhanced adaptability to the non-linear aspects of stock data. However, as noted by Y. Wu et al. 2023, achieving optimal performance in ML models, particularly those that integrate convolutional and transformer networks, often requires extensive tuning to handle the complexities of multivariate time series (MTS) data. Their study proposed an MTS classification of the Convolutional Transformer Network (MCT-Net) by combining convolutional layers for local feature extraction with transformer encoders to capture global dependencies, which has proven effective in reducing computational load and improving accuracy across multiple datasets. Similarly, J. Wu et al. 2021 used structural information from price graphs after converting the data, demonstrating the benefits of feature engineering and data transformations like using graph-based embeddings like struc2vec where the model encodes relationships between data points in the series to mitigate data volatility.

In their work on hybrid models, Ali et al. 2023 introduced a hybrid approach that leverages Empirical Mode Decomposition (EMD) combined with LSTM networks to improve the accuracy of financial time series forecasting. EMD is a signal processing technique adept at handling non-linear and non-stationary data by decomposing complex time-series signals into a set of intrinsic mode functions (IMFs), each representing distinct oscillatory components. Decomposition aids the LSTM model by isolating meaningful trends and fluctuations while reducing the impact of noise, advantageous for managing the volatile characteristics of the stock market data. Lv et al. 2022 proposed a hybrid leveraging Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) combined with ARMA and LSTM (CAL), which demonstrated better accuracy across several metrics compared to traditional and DL models. For example, in predicting the DAX index, the CAL model achieved an MAE of 72.3340, substantially lower than the LSTM model (MAE of 167.0816) and the ARIMA model (MAE of 136.0422). This method allows the model to address stable and volatile market conditions by decomposing the stock index time series into linear and non-linear components. ARMA handled stationary data and LSTM focused on non-linear patterns. Jeong et al. 2021 emphasises the importance of data normalisation, specifically Min-Max scaling, to enhance model performance in DL applications. They observed normalising the data as being crucial for models sensitive to feature scales to promote stable convergence during training and improve accuracy. Although their experiments focused on temperature prediction, they demonstrated that consistent normalisation enhances model robustness, particularly for datasets with wide fluctuations in values.

Barra et al. 2020 explored a CNN-based approach by encoding time series data into Gramian Angular Field (GAF) images, allowing the CNN to identify complex patterns. By structuring raw stock data as image representations, they used CNN's strengths in spatial feature extraction without excessive pre-processing, ensuring critical information was preserved. In contrast, Bhandari et al. 2022 utilised LSTM models to predict the S&P 500 index, using structured pre-processing methods, such as Min-Max scaling for data normalisation and denoising using Haar wavelets to address noise in stock data. Their study highlighted that carefully structured and transformed data, feature selection and targeted pre-processing can improve LSTM's ability to capture non-linear patterns in volatile markets. In their work on fuzzy time series models, Carvalho Tavares, Ferreira, and Mendes 2022 explored an innovative approach using a red-black tree structure to partition the universe of discourse to forecast the stock indices. This method employs red-black trees' self-balancing properties to ensure data clusters are statistically well-distributed and to support improved forecasting accuracy. The red-black tree framework aids in effectively managing the non-linear and volatile characteristics of stock market data, which often exhibit complex patterns that traditional models may struggle to capture. Their method highlights how non-traditional models like fuzzy logic-based systems can capture complex patterns without relying solely on DL methods. In their study, Rokhsatyazdi et al. 2020 proposed an optimised approach to stock price forecasting by enhancing an LSTM model's performance through hyperparameter tuning with Differential Evolution (DE). By fine-tuning key parameters, such as batch size, window size, and the number of LSTM layers, the LSTM achieved lower RMSE than traditional statistical models, showing the importance of hyper-tuning to enhance prediction accuracy.

Tripathy et al. 2023 highlighted the importance of careful pre-processing to enhance model accuracy, especially in balancing noise reduction with the retention of critical market signals. In their study, feature selection techniques like LASSO and Principal Component Analysis (PCA) were used to refine the input data by focusing on the most relevant features, reducing dimensionality, and addressing collinearity issues. These steps ensured that the GRU model captured relevant patterns in the data without excessive noise to enhance its predictive accuracy. While smoothing can help manage volatility, excessive application risks losing important details, showing the need for thoughtful pre-processing to achieve reliable forecasting results. Similarly, Yang, Chen, and Liu 2021 explored optimisation techniques to improve stock price prediction by using Particle Swarm Optimization (PSO) to fine-tune the parameters in an RNN. Their approach enhanced the model's ability to capture complex market patterns by improving global and local search capacities, highlighting how PSO-driven parameter tuning can boost model performance in dynamic financial contexts. In addition, Jarrah and Derbali 2023 examined pre-processing strategies tailored to stock data's volatility, utilising exponential smoothing to filter noise effectively, thereby enhancing model stability in forecasting applications. Tang and Shi 2021 contributed to this body of research by analysing how balanced pre-processing can enhance model accuracy during volatile market periods. While smoothing can reduce data volatility, their work emphasises that excessive transformation risks missing subtle but crucial signals necessary for accurate predictions. The need for balanced pre-processing was further reinforced by H. Song and Choi 2023, who found that excessive transformations like aggressive smoothing or dimensionality reduction can simplify the data too much, resulting in the loss of essential information. They observed that the model accuracy could be negatively affected when critical signals were removed during pre-processing. These examples highlight the importance of balancing noise reduction while retaining meaningful patterns for better model performance, demonstrating the challenges of balancing the need for clean and scaled data to avoid losing vital information. Understanding which transformations optimise model performance without sacrificing critical market signals remains a core area of research for finding methods that reduce noise while preserving key data characteristics.

Lu and X. Xu 2023 made some advances in stock price prediction with a Time series Recurrent Neural Network (TRNN) model, tailored for both predictive accuracy and computational efficiency. Their research found that compared to traditional RNN, LSTM, ARIMA, and GARCH models, the TRNN model reduced error rates, with an MAE at approximately 9.23% lower in most scenarios. The TRNN model's effectiveness comes from optimised data pre-processing and compression techniques, allowing it to efficiently handle time series stock data without sacrificing accuracy. Gülmez 2023 further contributed by evaluating an optimised deep LSTM network for stock price prediction, focusing on the impact of hyperparameter tuning through the Artificial Rabbits Optimisation (ARO) algorithm. Their approach consistently outperformed traditional LSTM and ANN models; for example, the LSTM-ARO model achieved an MAE of 3.804 for AXP, significantly lower than the 5.159 MAE of the standard LSTM. Similarly, the LSTM-ARO model achieved an MAE of 3.318 for AMGN. In contrast, the ANN model had a higher MAE of 6.193, showing the benefits of optimised LSTM configurations in handling complex stock data.

2.2.2 CNN Architectures and Performance Evaluation

The transition from traditional models to DL architectures, particularly for CNNs, marks a significant advancement in the ability to forecast the stock market. CNNs were initially designed for image recognition but have shown considerable promise in time series forecasting, especially when the data is transformed to leverage CNNs' pattern recognition strengths. Ramanathan and McDermott 2021 emphasises the potential of CNN-based models for capturing complex, non-linear dependencies, especially in high-frequency trading environments. The ResNet architecture builds on the strengths of traditional CNNs by incorporating residual connections, which allow certain layers to be bypassed, addressing the vanishing gradient issue that can arise in deep networks. While the deep layers of ResNet are beneficial for extracting features, overfitting remains a potential risk, especially in cases with limited or sparse data. Meanwhile, Zhao et al. 2023 highlighted the importance of integrating relational information into time series models to predict the stock market. They proposed a Time Series Relational Model (TSRM) that combines time series data with stock relationship information derived from clustering techniques. Similarly, S. W. Lee and H. Y. Kim 2020 addressed forecasting challenges with super-high-dimensional time series data by developing a novel framework using Convolutional LSTM networks, trend sampling, and specialised data augmentation techniques. Their proposed model, NuNet(DA, T), demonstrated significant improvements in forecasting accuracy compared to baseline models. For example, on the S&P500 index, NuNet(DA, T) achieved an MSE of 3,658.5, while on the KOSPI200 and FTSE100 indexes, it achieved MSEs of 1,326.0 and 2,648.7, respectively. The model's performance was enhanced by two key data augmentation methods: column-wise random shuffling, which reduced overfitting by randomising the order of companies in each mini-batch, and trend sampling, which prioritised recent data to capture dynamic market trends better. These techniques enabled the model to learn high-level features from large, complex datasets, improving forecasting accuracy. Both Zhao et al. 2023 and S. W. Lee and H. Y. Kim 2020 emphasised the necessity for adaptive and generalisable models in stock market forecasting that can maintain performance across diverse and fluctuating market conditions as conventional models often struggle with the dynamic nature of stock data, particularly during periods of heightened volatility.

In addition to standard metrics like MSE and MAE, Dstat and Theil's U have emerged as valuable performance metrics in financial forecasting. Dstat, which evaluates a model's ability to predict the correct direction of price changes, provides insight into a model's performance on directional trends rather than just magnitude. This means that Dstat assesses whether the model can accurately capture the general movement—upward or downward—of price changes over time, not just the size or extent of individual price fluctuations, which is very relevant in financial applications where capturing the correct direction is often more important than the exact value. Studies by Ruan, W. Wu, and Luo 2020 indicate that CNN-based models tend to achieve higher Dstat values, demonstrating their ability to capture directional trends in a volatile market as opposed to traditional models. Theil's U is another metric that provides a relative measure of forecast accuracy by comparing the model's performance to a naive or random-walk forecast. Lower values of Theil's U indicate better predictive accuracy relative to a baseline, with values below 1 implying the model is more accurate than a naive prediction. On the S&P500, the NuNet(DA, T) model achieved an MSE of 3,313.7, which was notably lower than the values achieved by traditional models.

Hybrid architectures like CNN-LSTM models have gained significant attention due to their ability to combine the strengths of CNNs and LSTMs in financial forecasting. Semenoglou, Spiliotis, and Assimakopoulos 2022 highlighted the computational demands of CNN and hybrid CNN-LSTM models. These hybrids require substantial processing power and memory, making real-time predictions challenging when applied to high-frequency trading scenarios. Furthermore, Anand and Nayak 2020 pointed out that DL models require large volumes of historical data to achieve reliable accuracy, which can be a limiting factor in emerging markets with shorter financial histories. Cai et al. 2024 introduced an explainable dual-mode CNN (XDM-CNN) framework specifically designed for MTS classification (MTSC), leveraging both 1D and 2D convolutional layers to capture complex patterns. The focal point of this framework is its 2D CNN component, which operates on time-frequency graphs created from the original time series data using Ensemble Empirical Mode Decomposition (EEMD) and the Hilbert-Huang Transform (HHT). By transforming the data into the time-frequency domain, the 2D CNN can capture nuanced high-level spatial and frequency-based features often challenging to detect in raw time-domain signals alone. This dual-mode approach allows the XDM-CNN to achieve exceptional classification accuracy, with performance metrics showing comparable or better results than other leading models in MTSC. The XDM-CNN achieved an average accuracy of 73.47% and ranked highly across multiple benchmark datasets, demonstrating robust performance. Moreover, the model's interpretability is enhanced through

Gradient-weighted Class Activation Mapping (Grad-CAM) visual explanations as a heatmap and numerical importance calculations based on information entropy, measuring the importance of the features to reduce uncertainty. Unfortunately, the paper did not disclose the other metrics, such as MAE or MSE, making it difficult to compare with other methods. Another example of a successful hybrid architecture that combines CNN with other techniques is the CNN-BiLSTM model.

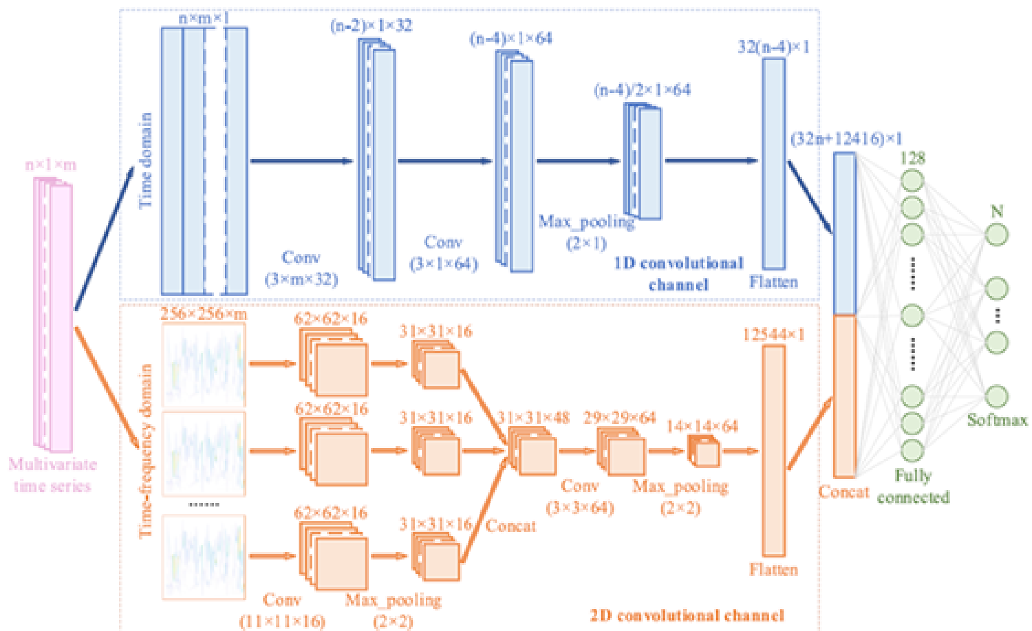


Figure 4: Cai et al. 2024 model architecture using both images and numerical data to make a prediction.

In their study, Jeong et al. 2021 applied this CNN-BiLSTM setup to forecast temperature instead of stock prices, taking advantage of the complementary strengths of both CNN and Bi-LSTM layers. Here, the CNN handled the feature extraction from observed and forecasted temperature data, while the Bidirectional LSTM (Bi-LSTM) layer captured sequential time dependencies. This approach allowed the model to track temperature patterns effectively, showing how well it can handle complex time dependencies in environmental data. Although the study focused on temperature forecasting, it highlights the potential of CNN and recurrent network hybrids to adapt to sequential patterns in time-sensitive applications. In addition to advancements in CNN architectures, enhancing model interpretability has become crucial in financial forecasting, where understanding the drivers of predictions is as important as accuracy.

Boniol et al. (2022) introduced the Dimension-wise Class Activation Map (dCAM), an interpretability method specifically for multivariate time series. This technique allows CNNs to highlight critical features and dimensions that contribute most significantly to the model's decisions, providing insight into which variables or time steps influence predictions. Such interpretability can be invaluable in financial contexts, helping analysts understand how different variables impact stock forecasts, especially when high-dimensional data is involved. In addition to advancements in CNN architectures, enhancing model interpretability has become crucial in financial forecasting, where understanding the drivers of predictions is as important as accuracy. Boniol et al. 2022 introduced the Dimension-wise Class Activation Map (dCAM), an interpretability method specifically for MTS. This technique lets CNNs pinpoint the key features and dimensions that have the biggest impact on the model's decisions, offering a clearer picture of which variables or time steps drive predictions. In financial contexts it helps analysts see how different variables shape stock forecasts, especially in complex, high-dimensional datasets.

2.2.3 Hybrid Models in Stock Market Forecasting

Hybrid models combining different ML and DL architectures have emerged as a significant area of research in financial forecasting. They leverage the strengths of various model components to overcome the limitations of standalone methods in terms of accuracy, adaptability, and resilience to market volatility. Gourav Kumar, Uday Pratap Singh, and Sanjeev Jain 2021 developed a hybrid model that combines PCA for dimensionality reduction, PSO to optimise initial weights, and the Levenberg–Marquardt (LM) algorithm for training a feed-forward neural network (FFNN). This approach addresses challenges like data complexity and local minima issues in training. The PSO-LM-FFNN model outperformed traditional FFNNs and alternative hybrid models in predicting the closing prices of indices like Nifty 50, Sensex, and S&P 500. The author stated higher accuracy and efficiency, such as the S&P 500 with an RMSE of 9.34E-04, Mean Absolute Percentage Error (MAPE) of 1.16E-04 and Theil's U of 1.87E-04. showing it could better capture complex stock price patterns than other methods tested in the study.

Y. Wang and Yuankai Guo 2020 investigates the adaptability of hybrid models for stock market volatility prediction but focuses on a DWT-ARIMA-GSXGB hybrid model, using Discrete Wavelet Transform (DWT) to separate data into components handled by ARIMA for low-frequency data and an optimised XGBoost model for high-frequency data. This hybrid aims to improve predictive performance and reduce error by utilising linear and non-linear model strengths, such as an accuracy rate 87.88% and an MAE as low as 0.00183 for specific datasets. On the other hand, Yeh et al. 2023 explores the effectiveness of multitask learning (MTL) in time-series classification, particularly using a 2D CNN architecture. The model enhances feature extraction and improves accuracy across multiple tasks by leveraging shared parameters and an attention mechanism to focus on critical data features. Beyond CNN-RNN and CNN-GRU hybrids, some researchers have explored combining CNNs with Attention Mechanisms. Semenoglou, Spiliotis, and Assimakopoulos 2022 presented a CNN-Attention-LSTM hybrid model where an attention layer was inserted between the CNN and LSTM layers. This attention layer helped the model focus on the most relevant parts of the data, giving weight to specific time points or features critical to predicting future stock prices. By directing the model's focus to important patterns and reducing the influence of irrelevant ones, the attention mechanism improved the performance across various market conditions. This hybrid showed a substantial increase in model accuracy during market anomalies and rapid shifts.

Hybrid models are not limited to DL combinations; some researchers have integrated traditional ML models with DL architectures to enhance model interpretability and performance. For example, Sharma, R., and Deepa 2022 introduced a stacked hybrid model combining CNN, LSTM, and traditional ARIMA techniques. Their model first applied CNNs and LSTMs to capture complex non-linear dependencies within the data, followed by an ARIMA model to refine the predictions based on linear relationships. In a separate study, Varadharajan et al. 2024 focused on the performance of LSTM-RNN models in stock price prediction, demonstrating LSTM's strengths in capturing non-linear patterns within volatile data. The study did not directly compare CNNs with other DL models. However, it emphasised LSTM's predictive accuracy over traditional approaches like ARIMA, achieving lower errors across metrics such as RMSE (2.51) and MAPE (1.84%). Each hybrid model illustrates different ways of combining spatial, temporal, and even interpretative layers to enhance prediction accuracy and robustness in stock market forecasting. The diversity of architectures reflects the evolving needs of stock market forecasting and the ongoing efforts to balance accuracy, computational efficiency, and model interpretability. Khanarsa, Luangsodsa, and Sinapiromsaran 2020 examined advancements in forecasting models. Highlighting the shift toward DL techniques using ResNet for time series order identification within the ARIMA framework. The study reports ResNet's potential to outperform the traditional methods in handling complex data, suggesting that the model is well-suited for capturing intricate dependencies.

2.2.4 Image-Based Models from Other Domains and Their Application to Stock Market Prediction

The transition towards image-based models in predictive analytics has roots in fields where pattern recognition and spatial relationships are paramount. Initially designed for visual data, CNNs have shown exceptional capabilities in medical imaging, climate science, and even text analysis when data is transformed into image-like formats. In medical imaging, CNNs have achieved remarkable accuracy in tasks like tumour detection, organ segmentation, and disease diagnosis. The domain has widely adopted techniques like transfer learning and multi-resolution analysis, allowing models to learn from vast datasets and adapt them to smaller, domain-specific datasets. Transfer learning's application to financial data transforms historical price movements and volatility patterns into image formats akin to multi-resolution MRI scans that reveal underlying structures at various scales. For instance, using a GAF or Recurrence Plot (RP) to convert time series data into images enables CNNs to capture the intricate relationships between past and present prices, similar to how MRI scans capture tissue patterns at different resolutions. Sayed, Himeur, and Bensaali 2023 explored an innovative application of 2D CNNs in building occupancy prediction, focusing on transforming time series environmental sensor data into 2D images. By converting data like temperature, humidity, and CO2 levels into visual representations, the model captured complex patterns within the time series that traditional approaches might miss. The 2D CNN analysed the spatial relationships between sensor readings over time to enhance classification accuracy for occupancy states. The model achieved a high accuracy rate of 95.6%, significantly outperforming conventional models using raw time series data. Moon, Hossain, and Chon 2021 used synthetically generated time series data representing AR and ARMA processes to evaluate the effectiveness of CNNs in accurately determining model orders, a task relevant in fields like signal processing and physiological data analysis. Using both the original time series and its time-frequency representation, their CNN model—based on ImageNet architectures such as Inception-v3 and ResNet50-v2—outperformed traditional model selection criteria, such as Akaike and Bayesian Information Criteria, in accuracy. Transforming time-series stock data into images can enhance CNNs' capacity to identify complex interdependencies and recurring patterns.

In climate science, CNNs have been used to model and predict weather patterns by analysing spatial-temporal data across large geographic regions. Techniques like spatiotemporal convolutions allow CNNs to capture spatial correlations and temporal dependencies in data, making them well-suited for forecasting tasks involving changes over time. The spatial dependencies within financial markets—where individual stocks or indices influence each other like weather cells in climate systems—can similarly be mapped and analysed through image-based techniques. Although Jeong et al. 2021 applied this approach to temperature data, similar spatiotemporal modelling techniques could theoretically be adapted to financial data. By structuring stock data as matrices where rows and columns represent different features and time intervals, the CNNs could capture cross-sectional influences between features over time, providing insights into market trends. In a study of physiological data analysis, J. Li and Q. Wang 2022 showed how converting time series data into 2D images can make a real difference in classification accuracy, especially for datasets with individual variability. By transforming ECG and EEG signals into images, they found that CNNs could generalise better across individual differences than traditional feature-based time series approaches. While their focus was on physiological data, this method has a clear crossover to financial forecasting. Much like physiological signals, stock market data often contains hidden patterns that shift over time, and converting these time series into 2D images allows models to pick up on complex dependencies and relationships that go unnoticed.

2.2.5 Applying Image-Based Techniques to Financial Time-Series Data

Converting financial data into image formats and leveraging CNNs' image recognition capabilities has opened new pathways for enhancing prediction accuracy. Techniques such as GAF and RP, widely explored in image-based time series research, have shown potential in financial forecasting by transforming numerical time series data into images. These transformations preserve temporal dependencies, allowing CNNs to detect nuanced patterns that raw time series data might not reveal as effectively. The GAF technique encodes each time series data point as an

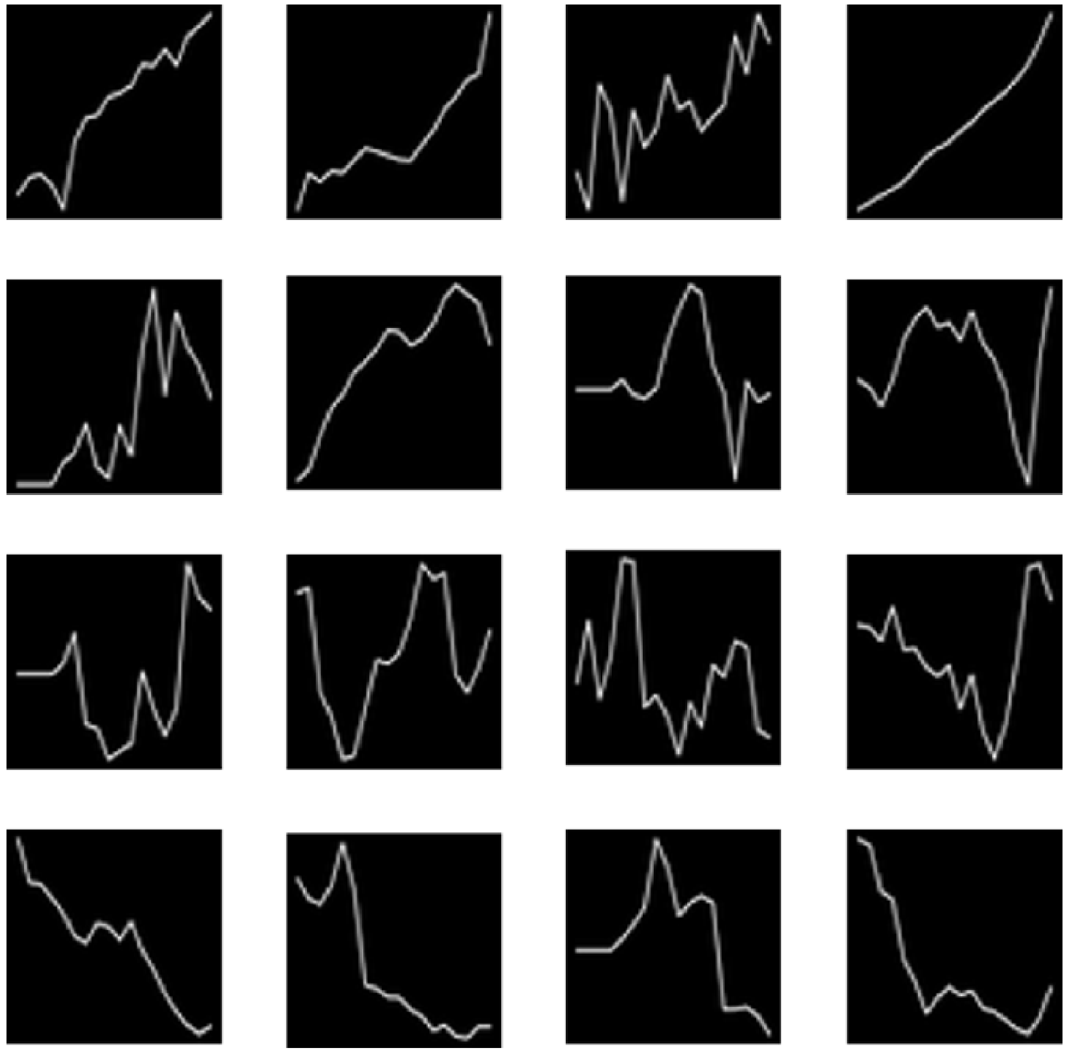


Figure 5: Semenoglou, Spiliotis, and Assimakopoulos 2022 created a line plot from the time series data using matplotlib and thickened the line. A single channel 8-bit integer was created to be fed into the CNN.

angle relative to the series' global maximum, highlighting overall trends and seasonal patterns in financial data. Similarly, recurrence plots represent instances when a time series returns to a prior state, allowing CNNs to identify recurring structures that may correlate with market trends. While these transformations are often discussed in the broader literature, they show the power of image-based methods for extracting complex patterns in time series data, a critical advantage in volatile financial markets. RP was initially used to visualise non-linear dynamics; RPs represent instances when a time series returns to a previous state, offering insights into recurring patterns. In financial forecasting, RPs have been employed to detect cyclical market behaviours that may aid trend prediction. Sayed, Himeur, and Bensaali 2023 2D CNN can be adapted to stock market data by transforming stock prices, trading volume, volatility, and other relevant indicators into 2D images. For instance, one could use heatmaps, recurrence plots, or candlestick patterns over specified time intervals to visualise market behaviours. A 2D CNN could then capture short-term dependencies in the data as spatial patterns and reveal trends or anomalies more effectively than standard numerical models. This method allows the model to recognise small shifts and correlations that indicate trend reversals for better prediction. By combining various financial indicators in a visual format, this approach could leverage the strengths of 2D CNNs in spatial feature extraction for financial forecasting applications.

2.2.6 Comparative Analysis and Key Trends in Image-Based Financial Models

Comparing traditional and image-based approaches highlights the advantages of image-based models in capturing complex financial patterns. For instance, Anand and Nayak 2020 demonstrates that CNNs, which can process image-like representations of data, often outperform traditional ML models, particularly in identifying patterns related to market cycles, seasonality, and momentum. Transforming data with methods like GAF and RP increases its dimensionality, enhancing feature extraction and making training slower and more computationally demanding. Incorporating PCA can streamline high-dimensional inputs by retaining only the most essential components, helping manage computational load while mitigating overfitting risks. To capture both spatial and temporal dependencies, hybrid models like CNN-RNN are recommended, as they can balance the strengths of each architecture, particularly in volatile financial markets.

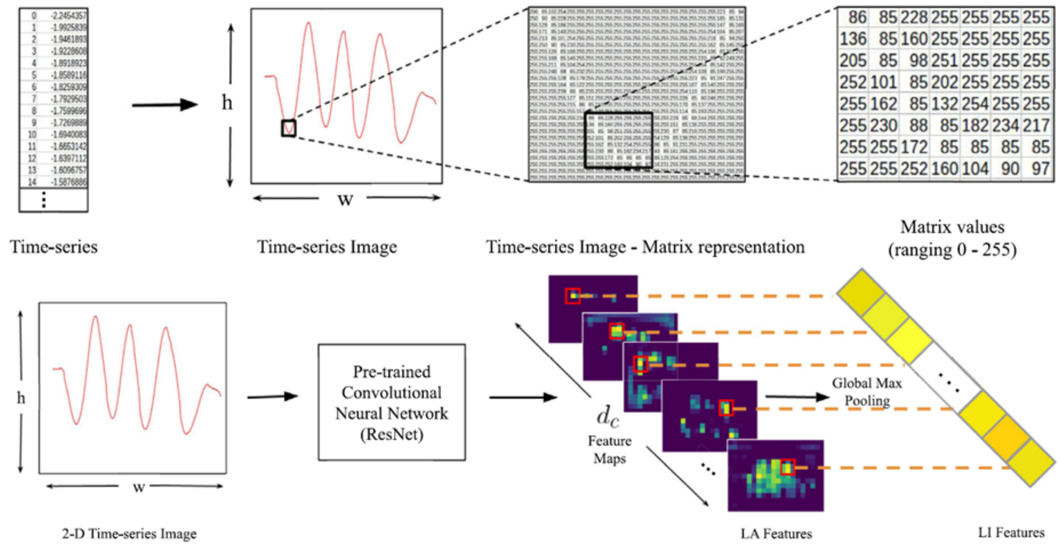


Figure 6: Anand and Nayak 2020 uses DeLTa method to transform time series data into 2D grids, capturing local patterns and temporal relationships and visualising them as images.

2.2.7 Justification for Using Image-Based Models in Stock Market Prediction

Image-based models, especially CNNs, bring clear advantages to stock market prediction by capturing spatial and temporal patterns that are tough to spot in raw time series data. Financial markets are complex systems with intricate dependencies, periodic cycles, and patterns that frequently vary in scale and intensity. Turning financial data into image-like representations, like GAF or RP, lets CNNs pick up on complex relationships that traditional time series analysis might miss. One big reason for using image-based models is CNNs' ability to handle high-dimensional data. By structuring stock data as images, these models can spot correlations and market signals like recurring trends, volatility clusters and momentum shifts. Techniques like PCA can help reduce the dimensionality of the images while preserving essential information. Another key advantage of image-based models is their flexibility in capturing both short- and long-term dependencies, which is especially valuable in the fast-moving world of the stock market. Studies, such as those by Zaheer et al. 2023, indicate that while CNNs alone may sometimes struggle in volatile conditions, combining CNNs with RNNs or LSTMs can enhance their robustness, enabling the model to capture temporal dependencies and providing a more comprehensive view of market trends. Homenda et al. 2024 reinforce the argument for image-based models in stock market prediction by showcasing the effectiveness of time series-to-image transformations for classification. Their study demonstrates that CNNs applied to image-transformed time series can outperform traditional methods, specifically when MTS are converted to multi-dimensional images. They also show that different image transformations—such as using original values and first and second differences—enable CNNs to capture intricate patterns within the data. This supports the argument that CNN-based image models can improve pattern recognition and capture the complex dependencies of stock market data

2.3 Conclusion

The literature review has detailed the potential and challenges of using image-based techniques, specifically 2D CNNs, for stock market prediction. Traditional numerical prediction methods, including ARIMA and ML approaches, have effectively captured linear and non-linear relationships within stock market data. However, these models often hit a limit due to their reliance on carefully pre-processed data, and they need help adapting to complex, non-linear patterns. CNN and hybrid models do have an edge over traditional methods in terms of accuracy, adaptability, and pattern recognition. However, these approaches need help regarding generalisability and robustness across varying market conditions. In addressing the first research question, the literature highlights that image-based techniques, particularly those using CNNs, can detect complex spatial and temporal patterns. Studies like Anand and Nayak 2020 have demonstrated CNNs' strength in capturing intricate relationships such as market cycles and seasonality, which traditional models like ARIMA struggle with due to their reliance on stability assumptions and limited handling of non-linear data. By transforming stock data into images, CNNs leverage spatial pattern recognition, providing a distinct advantage over numerical methods in identifying underlying structures within volatile financial markets. For question two, image-based representations, as discussed by Homenda et al. 2024, reveal complex interdependencies and patterns that might remain hidden in traditional time-series data. GAF and RP allow CNNs to detect recurring structures and volatility clusters. They also demonstrated how using different transformations enhances the CNNs' ability to capture intricate details within the data that add a predictive edge. For question three, transforming numerical data into images can help models generalise better across varying market conditions by capturing broader temporal and spatial dependencies. The literature suggests that CNNs, particularly when used with hybrid approaches like CNN-LSTM, can adapt more effectively to fluctuations in market volatility, as discussed by Zaheer et al. 2023. However, some studies, including those by Homenda et al. 2024, suggest that while CNNs trained on image data can capture complex dependencies, their generalisability may still be limited in highly volatile markets without supplementary methods like PCA to manage dimensionality and reduce overfitting. The final question regarding the pre-processing techniques, effective pre-processing and transformation techniques are crucial for optimising CNN performance. Studies by Jeong et al. 2021 and Barra et al. 2020 indicate that methods like Min-Max scaling, data normalisation, and dimensionality reduction improve model convergence and accuracy by reducing noise and managing scale fluctuations. Additionally, converting data into image formats like GAF or RP enhances CNNs' spatial feature extraction abilities, enabling them to capture meaningful patterns within the data. However, studies like those by Barra et al. 2020 emphasise the risk of excessive pre-processing, suggesting that over-simplification during transformation can strip away essential details that models need to detect critical patterns. This reinforces the need for a balanced approach in data preparation by simplifying data enough to aid model training while preserving the intricate signals necessary for accurate market predictions.

2.4 Assumptions and Future Directions

The literature points to a few core assumptions in stock market prediction research. The idea is that stock prices primarily reflect all available information, aligning with the concept of market efficiency. Behavioural biases, like bullish and bearish sentiments, shape trends, and while CNNs can capture some of these patterns, they may not fully account for unpredictable shifts in investor sentiment. These assumptions highlight the balancing act needed in stock market forecasting—getting the data handling right, designing effective models, and being realistic about the limits when trying to predict such a complex system. In this context, 2D CNN models, especially those built on the residual network configuration, really stand out for financial forecasting. With its depth and skip connections, a residual network can work through complex stock data with much better accuracy. By turning time series data into images, residual networks can capture spatial dependencies and pick up on intricate patterns that traditional methods can easily miss. The residual layers also help control overfitting, which makes the residual network more generalisable and robust, especially for high-frequency trading and volatile market conditions. Looking ahead, refining pre-processing pipelines, trying different transformation techniques, and developing models that can adapt to various market conditions are of real value. This project will use a novel residual network CNN model to boost prediction accuracy, improve generalisability, and explore new ways to interpret stock market data through image-based methods.

3 Methodology

3.1 Introduction

Predicting the stock market is a challenging task due to the time-dependent nature of the data. A combination of historical trends, economic events and shifts in market sentiment shapes the stock prices. The patterns within the time series often carry valuable information that could help forecast future movements. By treating this as a time series problem, sequential dependencies and relationships within the data can help the model better capture the behaviours that influence stock indices. Financial markets are dynamic, with trends, seasonality, and autocorrelation playing a key role with properties that are often non-stationary, meaning their statistical properties evolve, requiring transformations like differencing or logarithmic adjustments to extract meaningful patterns. The inherent noise and unpredictability of the market, driven by trends and random events, add to the complexity of making an accurate prediction. This section will outline the chosen model architecture, a residual network-based 2D CNN specifically designed to handle transformed financial data in image formats to leverage these temporal patterns to address the stock indices' dynamic, stochastic behaviour and provide some predictive insights. The model design incorporates several advanced features that enhance its performance and generalisability.

3.2 Research Approach and Rational

The transformation of numerical data into an image-based format is the prime step to enabling the residual network-based 2D CNN model. This approach is a stepping stone between raw numerical data sets and the spatial learning capabilities of CNNs by utilising a structured transformation process. Numerical stock and economic data are not traditionally suitable for 2D CNNs, as they lack the spatial relationships associated with image data. However, the numerical data are reshaped into grid-like structures that capture temporal patterns and feature interrelationships through pre-processing and reformatting, allowing the CNN to identify complex dependencies within the data. The data contains weekly stock prices, other indices and economic indicators for analysis. The data preparation starts with merging the stock data with other economic variables, ensuring the data is comprehensive enough to capture both market trends and external macroeconomic influences. The columns, excluding the dates, are shuffled to introduce fluctuations and reduce potential bias. This step maintains the temporal integrity of the dates while randomising the arrangement of features to aid with the model adapting to various patterns. Given that the data are recorded weekly, the temporal sequencing is preserved by creating overlapping slices of sequential data. Each slice corresponds to a rolling window of weekly observations, offset by one week. For instance, one slice captures a specific week's data, way. To be able to model the data using the CNN, the data are normalised using Min-Max scaling, transforming all features to a range between 0 and 1. This step ensures the model processes each feature equally to minimise the risk of dominant variables skewing the results, as it is crucial for ensuring the model does not disproportionately favour features with larger magnitudes, a common challenge in financial data sets where scales can vary significantly depending on the stock within the index. Normalisation reduces the risk of feature dominance and aids in stable convergence during training, a process highlighted in studies such as Jeong et al. 2021 and Barra et al. 2020. Once normalised, the data are reshaped into 2D arrays where rows represent temporal units, such as the weekly observations and columns correspond to features, such as stock indices and economic indicators.

The reshaped data are designed to align with the input requirements of the residual network 2D CNN architecture. The resulting grids for a column length of 3,287 measure 19×173, where 19 corresponds to the weekly observations in each slice, and 173 is the total number of features. These grids are then extended into a three-dimensional tensor similar to grayscale images in traditional CNN models by adding a single-channel dimension, yielding an input shape of $(\text{batch\_size}, 19, 173, 1)$. The shape is fed into the network, where the rows capture temporal dependencies and the columns reflect feature interrelationships. The structure allows the CNN to learn temporal progressions and feature correlations to uncovering patterns that may not be evident in numerical data.

Steps were taken during data preparation to define and isolate the targets to ensure the model predicted only the target variables and not all the input features. Variables such as adjusted closing prices for the NASDAQ and the NYSE were identified as the target data and explicitly separated from the features used as inputs to the model. This was achieved by dropping the target columns from the input dataset, so that it was excluded from the features the model would analyse. The target values were then stored in a separate array, carefully aligned with the corresponding input slices by date, to maintain the temporal structure of the data. The relevant target values were extracted for each time window, iterating through the data sets and creating a structured array ready for use in prediction. The model's output layer was designed to match the dimensions of this target array, ensuring that the number of predicted values matched the defined targets. This process was essential for maintaining the integrity of the predictive framework, ensuring the model focused on the needed targets without any overlap or leakage between inputs and outputs.

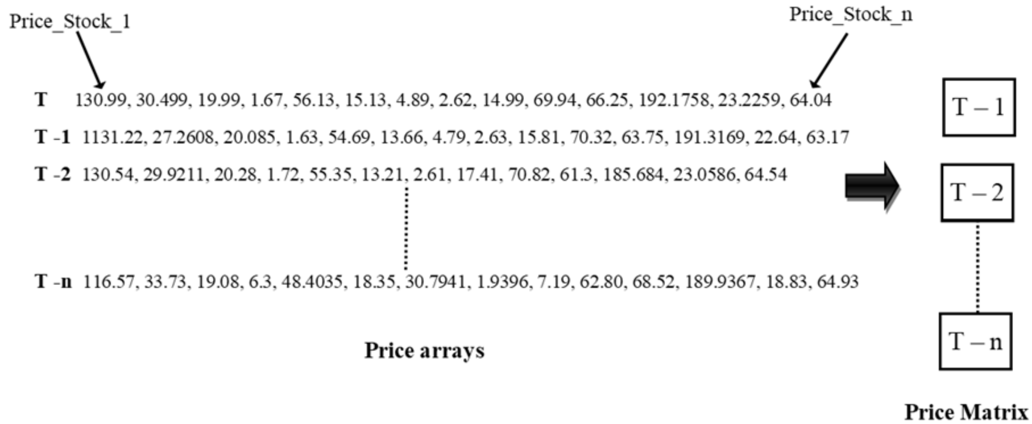


Figure 7: Transformation of price arrays into a structured price matrix for modelling, illustrating sequential stock prices (T, $T-1$, $T-2$, ... $T-n$) organised into rows for input processing.

Techniques such as Gramian Angular Fields (GAF) and Recurrence Plots (RP) mentioned by Sayed, Himeur, and Bensaali 2023, Barra et al. 2020 and Homenda et al. 2024 are often used to convert data into images for CNNs; however, this study adopts a custom transformation tailored to the nuances of financial data termed Grid-Based Temporal-Feature Structuring (GTFS), by structuring grids to capture temporal sequences and feature interrelationships, the methodology ensures that the CNN can extract both local and global patterns within the data. The initial convolutional layer focuses on capturing local dependencies within each grid, while deeper layers with residual connections enable the network to learn hierarchical and global patterns.

This model shares similarities with the approach outlined by M. Kirisci and E. Yolcu 2023, who implemented a CNN-based method for forecasting financial time series. Like their model, this one uses convolutional layers to extract both spatial and temporal features, with residual connections ensuring the preservation of information across deeper layers and preventing performance degradation. Dropout is applied in the dense layer of 300 units at 0.1 (10%) to address overfitting, a technique Kirisci and Yolcu emphasise as crucial for generalising models to the unpredictable nature of financial data. However, unlike their single-branch design, this model introduces a multi-branch structure, with three independent inputs capturing distinct temporal data slices. This adaptation allows for a richer representation of temporal dynamics and builds on their approach while tailoring it to the complexities of this research.

Each branch undergoes a series of convolutional and pooling operations to extract meaningful patterns from the input data. Convolutional layers apply filters, such as 3×3 or 19×19, with stride 1×1, to identify spatial and temporal dependencies in the grid-like structures. These filters generate feature maps, where activations highlight specific characteristics, such as shifts in stock prices or correlations between economic indicators. Batch Normalisation is applied after each convolutional layer to maintain consistent feature scaling and aid convergence. As the data flows through the layers, the model moves from identifying straightforward patterns, such as temporal edges, to capturing more abstract, high-level dependencies. Residual connections help retain crucial information and support the learning of intricate interdependencies within the data. Max Pooling layers reduce the dimensionality of the feature maps by downsampling, allowing the model to prioritise the most important activations while simplifying the data structure (Mehmet Kirisci and O. C. Yolcu 2022). This enables the model to focus on key features and adapt more effectively to different market conditions. The ReLU activation is applied after each convolution introduces non-linearity, which is essential for learning the complex, non-linear relationships found in financial datasets. This hierarchical extraction of features, progressing from simple patterns to complex structures, allows the CNN to uncover the dependencies that inform its predictions.

The fully connected layers that follow the convolutional and pooling stages consolidate the features extracted by the CNN branches. These dense layers connect all neurons from the previous layer to all neurons in the next, enabling the model to learn complex, high-level representations. The first dense layer consists of 15 units with ReLU activation, introducing non-linearity, followed by a second dense layer with 300 units, also using ReLU activation. The final output layer uses a linear activation function, which is tailored for regression tasks, with neurons corresponding to the target variables in the dataset, making it ideal for forecasting stock prices or indices seen in Figure 8.

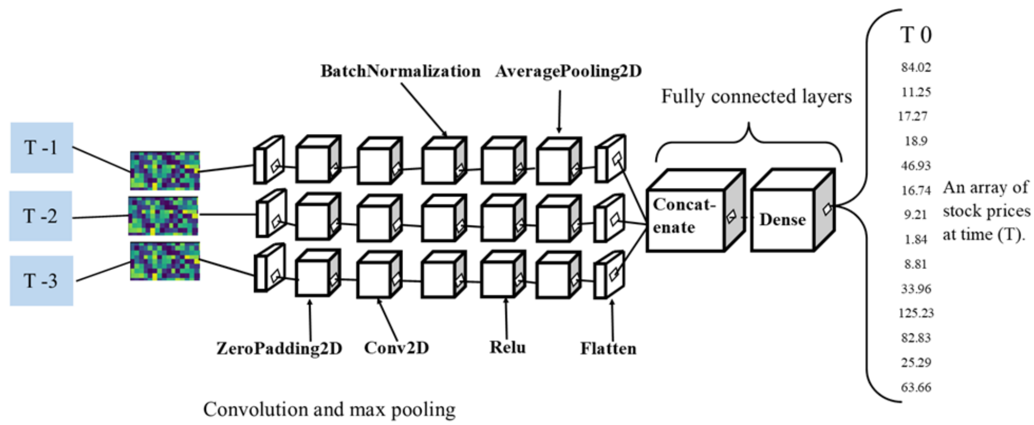


Figure 8: Illustration of the residual network architecture showing the transformation of temporal inputs (T -1 to T -3) through convolutional and max pooling layers, followed by fully connected layers, to predict an array of stock prices at time (T0).

To optimise training, the model utilises the Adam optimiser chosen for its adaptive learning capabilities, which is particularly suited for noisy financial data. An initial learning rate of 0.001 was subsequently reduced to 0.0001 and 0.00001 in successive phases to improve convergence and stability.

The training data is processed into 3D tensors with dimensions corresponding to the input shape required by the CNN branches. Each tensor slice represents a transformed segment of the data, reshaped into a grid-like image format to maximise the CNN's feature extraction capabilities. The preparation chosen aligns with the literature's findings that CNNs perform optimally when trained on structured image data.

3.3 Data and Research Context

The NASDAQ and NYSE stock data set consists of 195 columns of stock indices and 3,280 rows of weekly prices in a float format, dated from 12/06/2020 to 01/03/2024. All columns contain complete data with no missing values, ensuring consistency for analysis. The mean and median values are closely aligned across most columns, suggesting that many variables show some symmetrical distributions, although certain columns show larger gaps between these measures, indicating possible skewness in some of the stock data. Fluctuations across the dataset vary significantly as reflected in the standard deviation, with some stocks like `NLTX`, `SMH` and `VSAT` displaying considerable fluctuations, while others, such as `AAAU` and `SOHOO`, demonstrate greater stability. The range of values further highlights these differences with stocks like `A` and `NLTX` spanning wide minimum-to-maximum intervals, whereas `AAAU` has a much narrower spread. Certain columns, including `NLTX` and `SMH`, show extreme maximum values relative to their 75th percentiles, suggesting potential outliers as stock `ASMB` has the most outliers illustrated in Figure 9. This mix of predictable and highly volatile patterns across the dataset required careful preprocessing like Min-Max scaling to ensure diverse stock behaviours do not disproportionately influence model predictions. More charts and tables relating to this data are in Appendix A and B.

The exploration and merging of data sets for this project involved gathering and refining a substantial amount of historical stock market and economic data. One data set included core indices such as the Dow Jones Industrial Average (DJI), S&P 500 (GSPC), Russell 2000 (RUT), FTSE 100 (FTSE), and the German DAX (GDAXI). These indices were analysed for their adjusted closing prices over a specified period. The second, more comprehensive dataset incorporated additional macroeconomic factors, including consumer price index (CPI), industrial production (IP), and monetary aggregates like money supply (WM2NS). The additional data made it possible to evaluate the influence of external economic indicators on model performance, offering a basis for comparing the impact of enhanced input features.

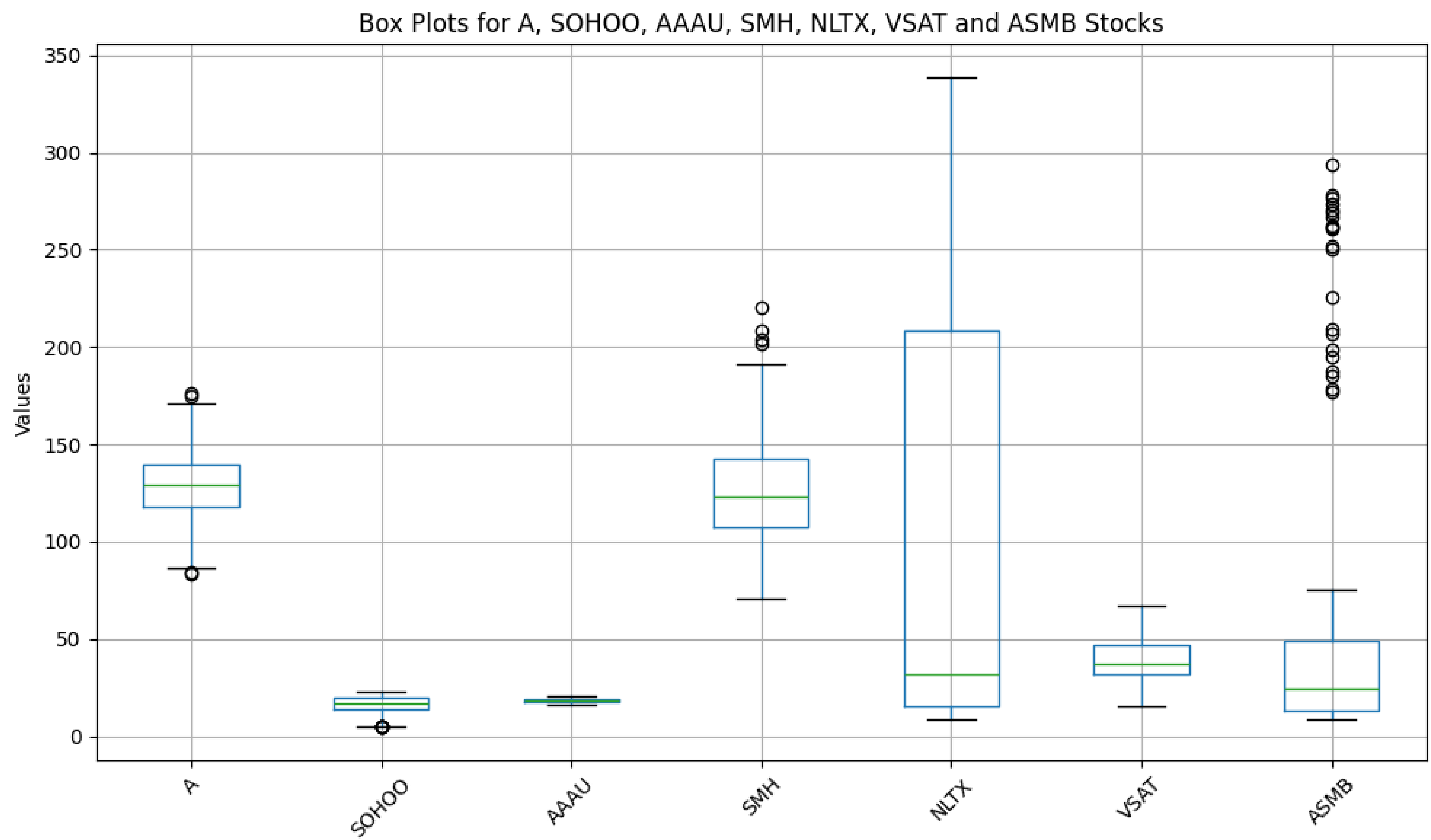


Figure 9: Box Plot of Stock Prices Across Selected Features.

Pearson's Correlation was a key statistical method used to assess the linear relationships among the variables in the data sets. The correlation heatmaps in Appendix B, created through this technique, offered insights into how strongly or weakly different indices and economic indicators were interrelated. This analysis was critical for feature selection, as it was informative, offering a clear understanding of the relationships. Indices showing a strong correlation at either 0.7 and above for positive correlations or -0.7 and below for negative correlations were prioritised for being included in the final data set. At the same time, those with weaker or redundant relationships, including near-linear dependencies, were excluded to minimise noise and optimise the model's learning efficiency.

The importance of Pearson's Correlation in feature selection is highlighted by its ability to quantify linear relationships between variables, which is important in predictive modelling. Recent studies, such as that by G. Li et al. 2022, published the efficacy of correlation-based methods in improving feature selection processes for financial time series analysis. These approaches are crucial for avoiding overfitting and ensuring the model generalises well to unseen data, allowing it to focus on key relationships, such as the influence of CPI on index performance or industrial production on small-cap indices like the Russell 2000. The analysis revealed significant positive correlations between adjusted closing prices from other indices and key economic factors such as CPI and industrial production (IP), suggesting their potential relevance for inclusion in the final data set.

The economic data needed to be factorised as the values for IP were small in comparison to the other data. One option was to remove it, but as it had a strong correlation with the FTSE, which was heavily correlated with the NASDAQ and NYSE, the decision was made to keep it and make the other data more equal so that the model would not treat the IP column as insignificant. More detail on this can be found in Appendix B.

Using two distinct data sets with one having the core stock indices and another enriched with macroeconomic indicators, allowed for a comparative assessment of the model's predictive ability. The additional economic data provided the model with a broader range of contextual information to enable it to detect underlying drivers of market trends that are not directly seen in stock indices alone. The dual data set approach facilitated a nuanced evaluation of whether extra information improves the generalisability and accuracy of the stock market predictions, providing a clear basis for selecting variables that align intending to uncover meaningful patterns for forecasting.

3.4 Data Collection Methods

Historical data for major stock indices, including the S&P 500, Russell 2000, FTSE, and GDAXI, were retrieved using Yahoo Finance via the `yfinance` library. Detailed stock data for the primary indices, NASDAQ and NYSE, were obtained from Alpha Vantage, which provided historical prices as a reliable source. Additional macroeconomic data, including the CPI, IP, and WM2N, were retrieved from the Federal Reserve Economic Data (FRED) system. These data sets span from June 2020 to March 2024, capturing a range of market conditions and economic cycles.

3.5 Limitations and Delimitations

The methodology of transforming numerical data into grid-like structures for CNN analysis introduces inherent limitations. For instance, while the process captures temporal patterns and feature interrelationships, it does not account for non-linear relationships or irregularities that may exist beyond the temporal and spatial domains, as highlighted by Barra, 2020. Additionally, the data's granularity regarding weekly observations rather than daily limits the model's ability to analyse intra-week fluctuations, which could be important for short-term forecasting.

Table 1: A sample of the combined data set of economic data and indices for **Experiment 3** where the first three rows show the before being factorised. **Experiment 2** did not have the last 3 columns as part of the input.

Date	**Adj Close DJI**	**Adj Close FTSE**	**Adj Close GDAXI**	**Adj Close GSPC**	**Adj Close RUT**	**CPI**	**WM2NS**	**IP**
2020-06-12	25605.5	6105.2	11949.2	3041.3	1387.6	257.0	18129.2	91.67
2020-06-15	25763.1	6064.7	11911.4	3066.6	1419.6	257.0	18129.2	91.67
2020-06-16	26289.9	6242.8	12315.7	3124.7	1452.3	257.0	18129.2	91.67
2020-06-12	67.746	78.869	69.756	61.294	58.815	85.22	85.103	91.67
2020-06-15	68.163	78.346	69.535	61.804	60.168	85.22	85.103	91.67
2020-06-16	69.557	80.647	71.895	62.975	61.552	85.22	85.103	91.67

Not all of the economic data could be included as it was unavailable for free download. As a result, the data input is incomplete as it does not capture the optimum number of factors influencing the NASDAQ and the NYSE in order to make an accurate forecast. Despite this, the information that is easily accessible should allow the model to forecast the trend of each stock, as it is assumed that investors are only concerned with making a significant loss or gain.

The reliance on historical data for training assumes that past patterns and relationships will persist in the future, an assumption that may not hold during unprecedented economic events according to the random walk hypothesis (Maloumian 2022). Moreover, while the model uses Min-Max scaling to normalise features, the technique could unintentionally suppress the importance of smaller-scaled data fluctuations that may contain relevant information.

Lastly, the model architecture itself, while sophisticated, is computationally intensive regarding time, limiting the ability to scale the analysis to larger datasets. These constraints highlight the need for further exploration and refinement to ensure the model's applicability.

3.6 Summary

This study reshapes numerical stock and economic data into grid-like image structures, allowing a residual network-based 2D CNN to analyse and detect intricate financial patterns. By organising the data to highlight temporal sequences and feature relationships, the approach connects the unique demands of numerical data with the spatial learning strengths of CNNs. Weekly stock prices and key economic indicators were merged and normalised using Min-Max scaling to ensure balanced feature contributions. Temporal integrity was maintained through overlapping slices of weekly observations and Pearson's Correlation guided the selection of features by prioritising variables with strong relationships while excluding redundant ones. The model's architecture incorporates convolutional layers, residual connections, and max pooling, allowing detailed patterns. This is especially true with the additional macroeconomic data, which enriches the analysis and enables the evaluation of external factors affecting stock trends. Despite innovative techniques, limitations include the inability to analyse intra-week fluctuations and reliance on accessible data, highlighting areas for future improvement.

4 Experimental Setup

4.1 Introduction

The experimental setup was designed to evaluate the effectiveness of a residual network-based 2D CNN in forecasting stock market trends using structured financial and economic data. The aim was to assess the model's ability to extract meaningful patterns from temporal and feature relationships in the data while incorporating macroeconomic variables to enhance predictive accuracy. This section outlines the experimental process, including data preparation, model structure, and evaluation metrics.

4.2 Experimental Process

The experimental process followed a structured approach to ensure the model was trained and evaluated close to conditions representative of real-world scenarios. Google Colab was used to run the model using Python on an Intel(R) Core(TM) i7-1065G7 CPU @ 1.30GHz, 1.50 GHz and 16GB RAM laptop. This was because the laptop was limited in handling the computational power required to train the model, which can significantly affect its performance. Furthermore, hardware-induced variations, such as those caused by differences in GPUs, have been shown to influence not only model accuracy fairness and learning efficiency, highlighting the importance of the computational platform in machine learning tasks (Hooker et al. 2023).

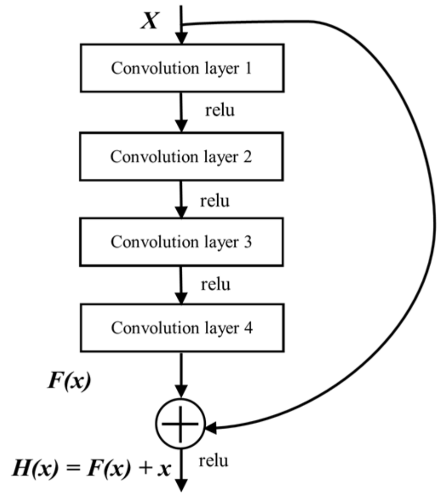


Figure 10: Residual blocks with skip connections are prominent components in residual networks. These blocks allow the input X to bypass intermediate layers via a skip connection and be directly added to the output $F(X)$ of those layers. This design facilitates the training of deeper networks by mitigating issues such as the vanishing gradient problem, thereby enhancing performance and convergence G. Xu et al. 2024.

Three model variations were explored:

- **Experiment 1 (E1):** Utilised only historical stock data to assess the predictive power of market trends alone.

- **Experiment 2 (E2):** Incorporated additional indices to investigate whether these features enriched the model's understanding of interdependencies in financial data.

- **Experiment 3 (E3):** Expanded the dataset to include other macroeconomic indicators, such as industrial production and money supply, to evaluate the influence of broader economic conditions on forecasting accuracy.

The data preparation pipeline followed the methodology outlined earlier, ensuring consistent and comparable transformations across all three models. The data was

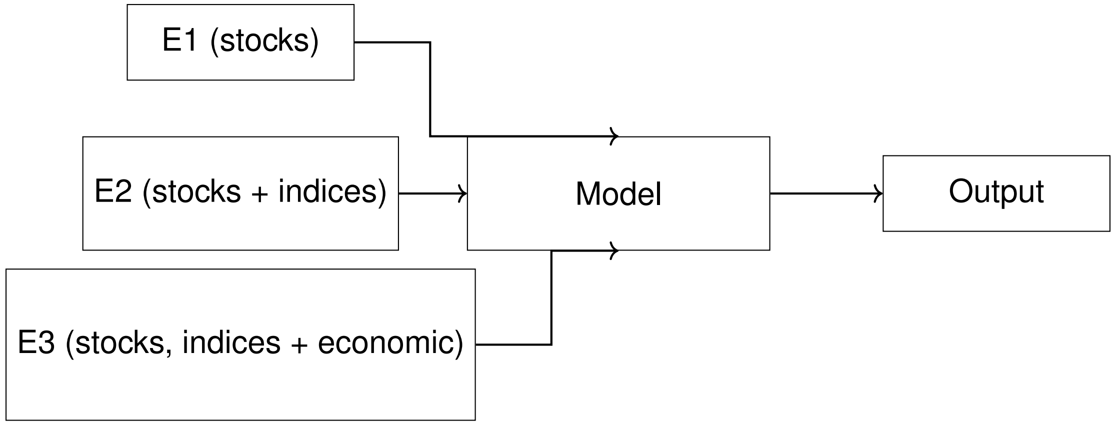


Figure 11: Diagram illustrating the integration of three experiments into a single model leading to an output.

split into training and validation sets, with 80% allocated for training and 20% reserved for validation. This division allowed the models to learn from a sufficiently large dataset while retaining enough samples to assess their performance on unseen data.

4.2.1 Training and Validation

The training process for all three models utilised the Adam optimiser, known for its adaptability to noisy financial data. The training needed to be aggressive as it was realised that the `ReduceLROnPlateau` wasn't enough to train the model. Although early stopping helps to prevent overfitting, it is not enough. The models were trained on 100 epochs, then 40 and finally another 40 for fine-tuning at a lower learning rate. While **E1** did not use dropout, **E2** and **E3** incorporated a 10% dropout rate in the fully connected layers to mitigate the risk of overfitting since both had extra data to learn from. Each model was specifically trained to predict the adjusted closing prices for the NASDAQ and NYSE indices. The primary evaluation metric was MSE, selected for its ability to highlight large deviations, making it particularly suited for assessing forecasting performance. Validation loss was carefully monitored across epochs by plotting the curve to ensure consistent learning progress and the models' generalisability to unseen data.

4.2.2 Experimental Objectives

The first objective was to evaluate the predictive performance of the baseline model (**E1**), which relied solely on historical stock data to assess the utility of market trends in isolation. The second objective centred on **E2**, which incorporated additional indices to determine whether these features provided complementary insights and enhanced forecasting capabilities. Finally, the third objective focused on **E3**, which introduced broader macroeconomic indicators, such as industrial production and money supply, to examine the influence of external economic variables on stock market trends. The experiments will also be tested on unseen data and its accuracy at predicting the future horizon beyond the data model has seen.

The comparative analysis of these three experiments provided a detailed understanding of the trade-offs and advantages of different feature sets. By systematically exploring these variations, the study aimed to identify the most effective combination of inputs for accurate and insightful stock market forecasting.

4.3 Evaluation Metrics

The performance of the model was assessed using a range of evaluation metrics to provide a comprehensive understanding of its capabilities: MSE was chosen as the primary metric as it has been highlighted in the literature as a common and sensitive metric for measuring forecasting accuracy, especially in financial contexts where large deviations can significantly impact decisions. For example, studies by Zhao et al. 2023 and S. W. Lee and H. Y. Kim 2020 used MSE effectively to compare model predictions against observed values, ensuring that large prediction errors are penalized more heavily.

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

Where y_i represents the actual values, $\hat{y}_i$ represents the predicted values, and n is the total number of observations.

Realised volatility was calculated using rolling standard deviations, which is relevant for evaluating models in dynamic markets. Financial forecasting research often incorporates volatility measures to understand model performance under fluctuating conditions. For instance, studies like those by Sayed, Himeur, and Bensaali 2023 explore fluctuations and their predictive relevance using derived metrics from time series analysis.

$$\text{Volatility} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(r_i - \bar{r})^2}$$

Where r_i represents the returns for time i, $\bar{r}$ is the mean return, and n is the number of observations in the rolling window.

The literature highlights the importance of correlation metrics in ensuring that models include variables that are both relevant and non-redundant. Pearson's correlation is particularly effective for this purpose, as it identifies relationships between features that can improve the model's learning process. In this case, it was not only used for feature selection but also to examine the relationship between volatility and error, helping to understand how fluctuations in stock prices may affect forecasting inaccuracies. Jeong et al. 2021 emphasises Pearson's correlation as a practical approach for enhancing feature relevance and evaluating model performance. It is calculated as:

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

Where x_i and y_i are the data points for the two variables, and $\bar{x}$ and $\bar{y}$ are their respective means.

The accuracy of the model's predictions was evaluated by comparing its forecasts to actual stock prices across different timeframes. MAPE is commonly referenced in financial forecasting for its interpretability and practical significance in measuring relative prediction errors. This aligns with methodologies discussed in studies that highlight balancing precision with interpretability in real-world financial contexts.

$$\text{MAPE} = \frac{100}{n} \sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right|$$

Where y_i represents the actual values, $\hat{y}_i$ represents the predicted values, and n is the total number of observations.

RMSE provides an interpretable measure of error because it is expressed in the same units as the target variable. It is calculated as the square root of MSE, offering a balance between penalising larger errors and maintaining interpretability.

$$\text{RMSE} = \sqrt{\text{MSE}}$$

This metric complements MSE by emphasising prediction accuracy while accounting for more significant deviations in forecasting.

MAE evaluates the average magnitude of errors without considering their direction, providing an intuitive and straightforward measure of prediction accuracy.

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i|$$

Unlike MSE and RMSE, MAE does not penalise larger deviations as heavily, making it a helpful metric for scenarios where all errors are treated equally.

4.4 Summary

The experimental setup demonstrates the potential of using a residual network-based 2D CNN for financial forecasting by simulating realistic conditions and testing the model with different types of data inputs. Transforming numerical financial data into grid-like image formats enabled the model to use spatial learning techniques to identify complex temporal patterns and feature relationships. Three data sets, ranging from historical stock data to data sets enriched with other indices and macroeconomic indicators, were used and the study highlights the model's adaptability to different data complexities. MSE was used to highlight large prediction errors, which carry significant weight in financial decision-making, with the addition of realised volatility to measure how effectively the model responds to shifts in market conditions. Both metrics are commonly used in financial forecasting, ensuring the model's evaluation is practical and aligned with real-world needs.

Furthermore, incorporating macroeconomic data, such as CPI and industrial production, reflects the broader economic influences on market behaviour. By ensuring the experimental setup addresses actual forecasting needs, the study has the potential to highlight the residual network-based CNN as a reliable tool to predict a near-accurate forecast.

5 Results

5.1 Introduction

This section presents the findings derived from the experiments to address the research questions outlined. The results reflect the outcomes of the residual network model applied to transformed stock data for the NASDAQ and NYSE indices. Quantitative metrics such as MSE and R^2 were used to assess the performance of three models. The results will explain each model's accuracy trends, comparative performance across stocks, and the impact of volatility on prediction errors. The model's sensitivity to high-volatility stocks, the relationship between volatility and error, and the effect of extremely low volatility on prediction accuracy are examined. Additionally, this analysis sheds light on the strengths and limitations of each model, as summarised in Tables 2 and 18.

5.2 Experiment 1

5.2.1 Standard Testing

The model can reasonably predict stock prices, as reflected by the MSE error 7.900 for the stocks in Table 2. Among the selected stocks, predictions for ADUS displayed the lowest error at 2.995, whereas predictions for ESNT had a higher margin of error, reaching 8.362. The model achieved an impressive R^2 score of 0.97, considering it was only trained on historical stock prices. However, significant outliers were observed, particularly for AAMC, where the model's prediction deviated substantially, resulting in an error of 562.252, this was also clear in Figure 47. The correlation coefficient between volatility and error was calculated at 0.505, indicating a moderate positive relationship, where higher volatility results in increased prediction errors.

5.2.2 Future Horizon Forecasting

Figure 12 to 19, depicting the actual versus predicted values for BEP, ACP, ADUS, AWP, ASG, CASY, ESNT and FCO over the 15 weeks from 8th March to 14th June 2024 in 15-day intervals, highlights the varying accuracy of predictions during this horizon testing. ESNT, with relatively low volatility at 0.0272, did not predict the price well. ADUS also had low volatility at 0.0451 but has a high MSE at 525.808, reflected in Figure 21 where the gap is wide. AWP had the lowest MSE of 0.004 and exhibited low volatility; however, the error (= Real − Predicted) was not as low as BEP and ACP as AWP had an error of 0.635 compared to 0.050 and 0.056 consecutively.

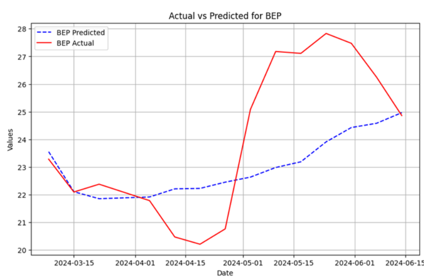


Figure 12: A line graph of prediction vs actual for BEP, which had the lowest error.

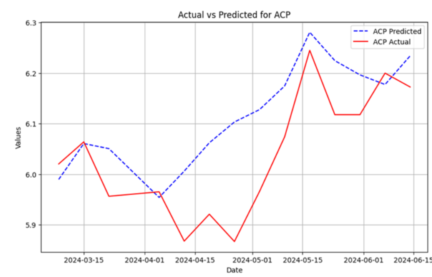


Figure 13: A line graph of prediction vs actual for ACP, which had the second lowest error.

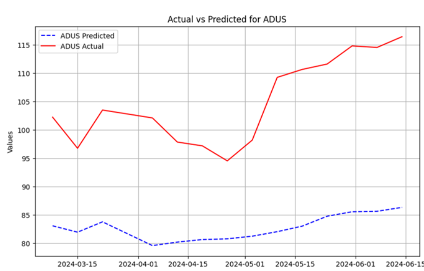


Figure 14: A line graph of prediction vs actual for ADUS, which had an error of 15.292.

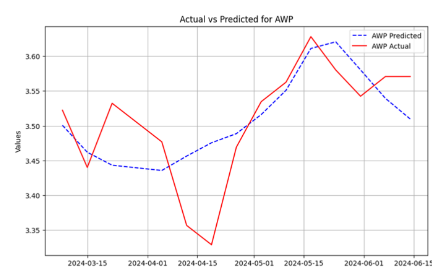


Figure 15: A line graph of prediction vs actual for AWP, where the prediction line seems to capture the moving average rather than the stock price.

The Figure 20 shows E1's overall performance. It shows a dense concentration of errors near 0, indicating consistent near accuracy for the majority of stocks. However, the distribution also highlights significant outliers with errors exceeding

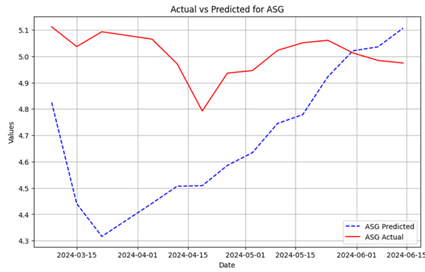


Figure 16: A line graph of prediction vs actual for ASG, showing an accurate prediction at roughly 9 weeks before deviating again.

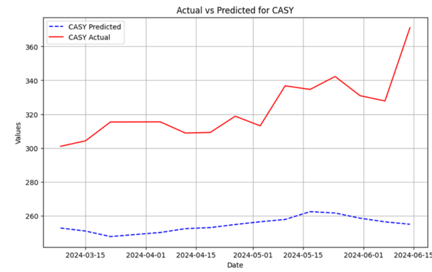


Figure 17: A line graph of prediction vs actual for CASY, which was poorly predicted.

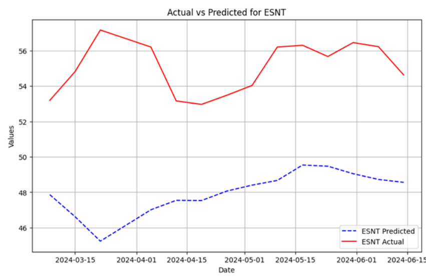


Figure 18: A line graph of prediction vs actual for ESNT. It had an error of 14.949.

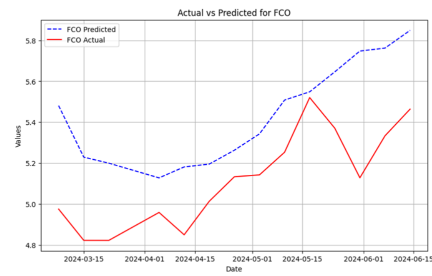


Figure 19: A line graph of prediction vs actual for FCO, which failed to capture the overall pattern in the stock price.

±50, which inflated the MSE. Performance metrics reported include an MAE of 8.12, an RMSE of 17.314, and an MSE of 299.776. These values reflect a balance of generally minor errors offset by substantial deviations for a few stocks. Prominent outliers include IYH, which recorded the highest MSE at 43,497.021, a clear contributor to the upper tail of the error distribution, although EQC had the highest error, which similarly skews the overall results. In contrast, stocks like BEP (error: 0.050) and ACP (error: 0.056) are at the zero peak of the histogram with high predictive accuracy recorded.

5.2.3 Smaller Unseen Data Testing

The model was tested on an unseen dataset with a smaller shape to evaluate its generalisation. There was a variation in accuracy across different stocks, with errors ranging from 0.071 to 16800.2. Stocks like KNCT, with a volatility of 0.0321, dis-

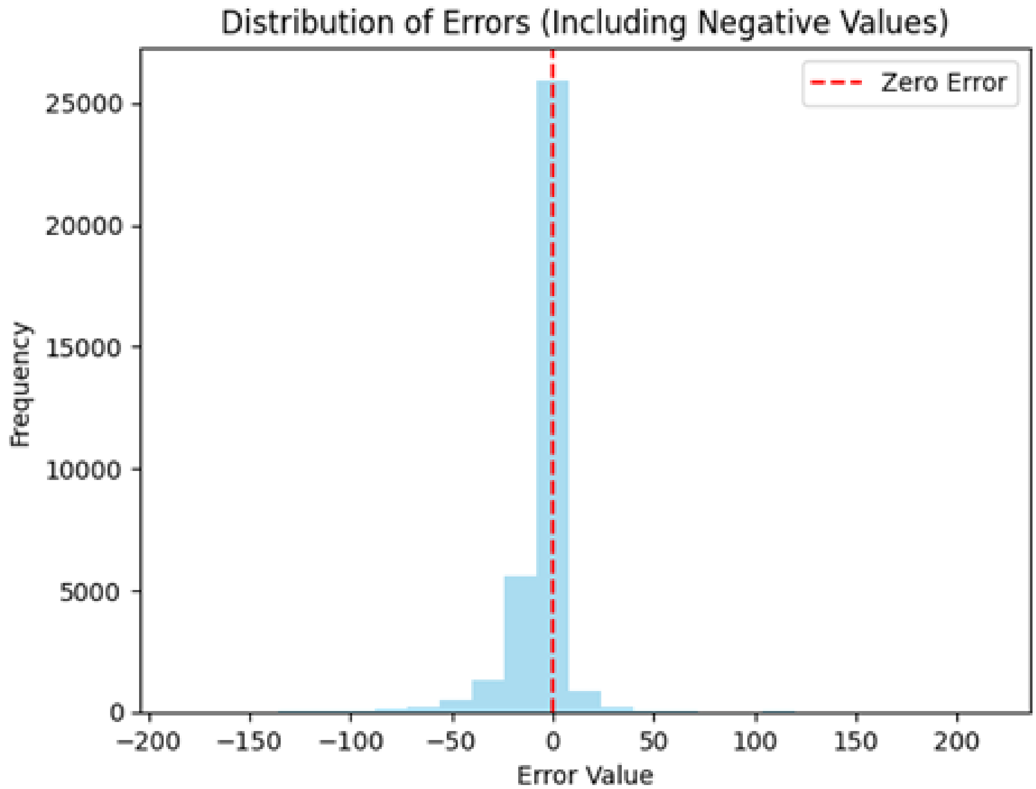


Figure 20: A histogram showing a high concentration of errors at zero for E1 when tested to forecast the horizon.

played the lowest prediction errors at 0.071, a closer alignment between actual and predicted values. In contrast, HEPA, with a high volatility (0.1484) and AMED, with a lower volatility (0.0578), recorded high errors. A correlation of 0.38 between volatility and prediction errors was observed, suggesting that volatility levels influence performance. While the model demonstrated potential in handling unseen data, its effectiveness varied by stock, pointing to areas that could benefit further refinement.

5.3 Experiment 2

5.3.1 Standard Testing

A balanced performance was achieved with an R^2 score of 0.981 (Figure 49). Individual results for stocks such as KAR and HERD highlight this performance, with errors of 14.624 and 9.922, respectively and a general consistency in error distribution. However, stocks like ADUS and IG presented challenges, particularly ADUS, which had a significantly high error of 97.194 despite its low volatility of 0.05479 (Table 2). With a volatility of 0.010411, IG had the most significant error of 241.240, suggesting that extremely low volatility might contribute to unpredictable predictions. The average error across a selection of stock predictions was 26.719, showing that while the model generally performed well, certain stocks deviated from their actual values. The correlation coefficient of 0.0043 between volatility and error suggests that the relationship between the two is weak. Perhaps there are other factors contributing to prediction accuracy. The scatter plot of volatility versus error further visualises this trend, showing a clustering of errors in the lower volatility range with occasional outliers (Figure 50).

5.3.2 Future Horizon Forecasting

In Figure **??**, which captures the Actual vs Predicted values for KAR, HERD, ESNT, ADUS, and FTDS, the 15-week horizon reveals varying prediction accuracies among these stocks. KAR showed consistent over-prediction, particularly during mid-April, with sharp discrepancies contributing to its MSE of 10,623.77. In contrast, with a low volatility of 0.0626, HERD displayed stable trends between actual and predicted values, resulting in a much smaller MSE of 125.881. ESNT, with an MSE of 1,245.136, demonstrated over-prediction from mid-April onwards, as predictions consistently exceeded actual values. In contrast, FTDS has an MSE of 23.994 with a closer alignment between actual and predicted, which experiences even less volatility (0.0261). The histogram in Appendix B illustrates a broader distribution of error values compared to Experiment 1. Although most errors cluster around 0, there is an increase in higher errors, showing greater variability in the model's performance. The performance metrics confirm this trend, with an MAE of 46.25, RMSE of 72.35, and MSE of 5235.29.

Outliers such as CACI (MSE of 141,503.83) and HUBB (125,617.13) contribute heavily to the overall high MSE. Meanwhile, stocks like ASG (MSE of 0.0378) and

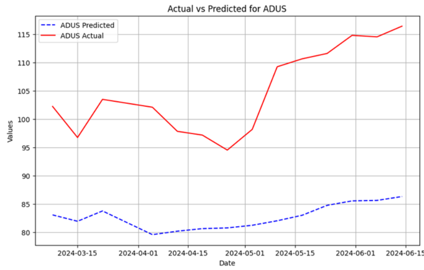


Figure 21: A line graph of prediction vs actual for ADUS, which had an error of 15.292.

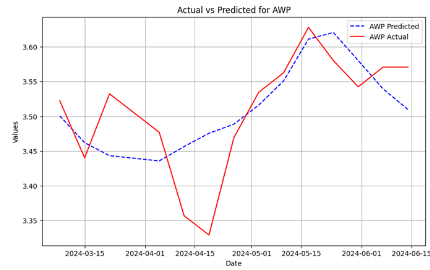


Figure 22: A line graph of prediction vs actual for AWP, where the prediction line seems to capture the moving average rather than the stock price.

MBI (MSE of 0.1452) represent the lower-error cluster, highlighting the model's capability to provide highly accurate predictions for specific stocks.

5.3.3 Smaller Unseen Data Testing

The results showed a positive correlation between volatility and prediction error, with a coefficient of 0.368. Stocks such as LYRA, with higher volatility (0.099), displayed significant prediction errors (14,537.58), while DORM, with lower volatility (0.044), achieved minimal errors (0.025). The stocks in the Top 50 Highest Volatility category showed larger errors, highlighting the model's difficulty with volatility. Figure 23 provides additional context to the error distribution during the evaluation, highlighted as a dense cluster of errors around zero; however, the spread of errors, including notable deviations beyond ±100, indicates challenges with certain stocks. This distribution aligns with the more significant errors observed for stocks such as LYRA while capturing minimal errors for more stable stocks like DORM.

5.4 Experiment 3

5.4.1 Standard Testing

The training and validation loss curves show a consistent decline, with both losses converging smoothly over time, indicating that the model was well-trained and ef-

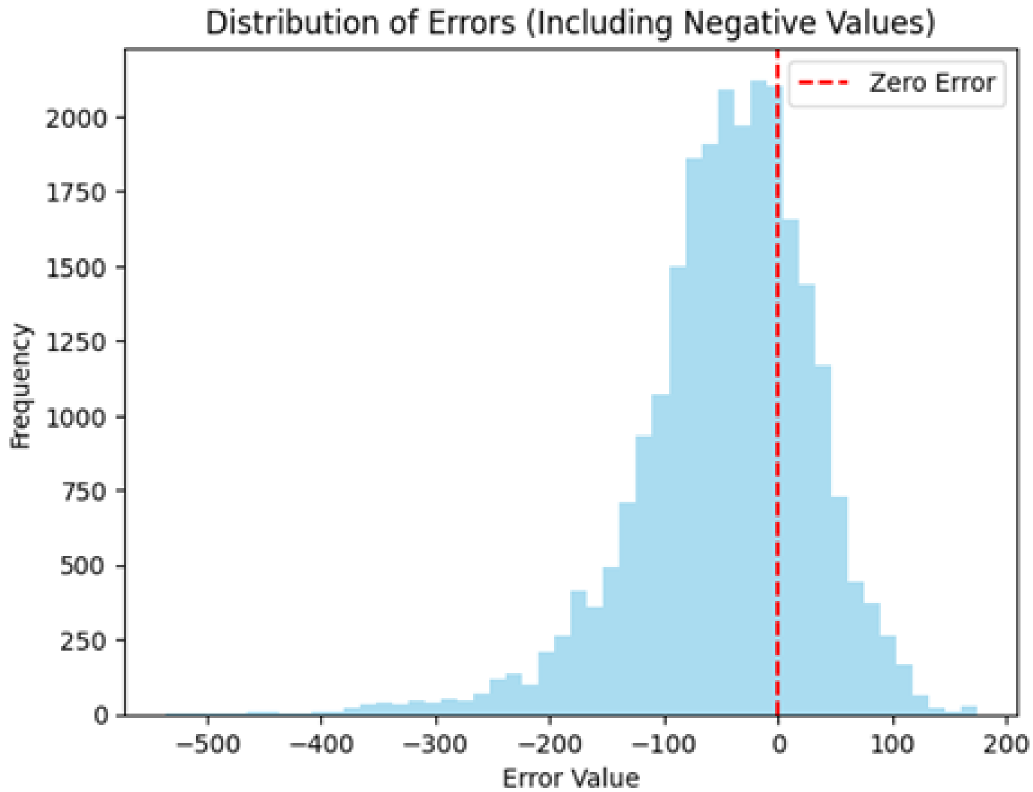


Figure 23: A histogram showing a wide spread of errors.

fectively generalised for validation data. Numerically, the model's average error was the second lowest among the three at 25.68 (Table 2). Experiment 3 achieved the highest R^2 score among the three models, at 0.99 (Figure 24), demonstrating its capacity to predict stock prices well. The data points align very closely with the diagonal line, indicating high predictive accuracy. While most predictions are precise, minor deviations at higher price levels suggest slight inaccuracies in specific instances. Notably, for row five, the R^2 score was 0.9909. Although the individual stock prediction for "IBRX" had a significant error, with the actual stock price being 1.1182 as it predicted 4.9149, resulting in a high error of 339.5399. The scatter plot shows a tight clustering of predicted and actual values along the diagonal, confirming the model's effectiveness. The correlation coefficient between volatility and error was low at 0.0126, suggesting a limited direct impact of volatility on prediction error. However, the error distribution relative to volatility shows that while the model performs well under low-volatility conditions, higher-volatility scenarios still pose challenges, as evidenced by the results for "IBRX" stock and the scatter plot (Figure 52).

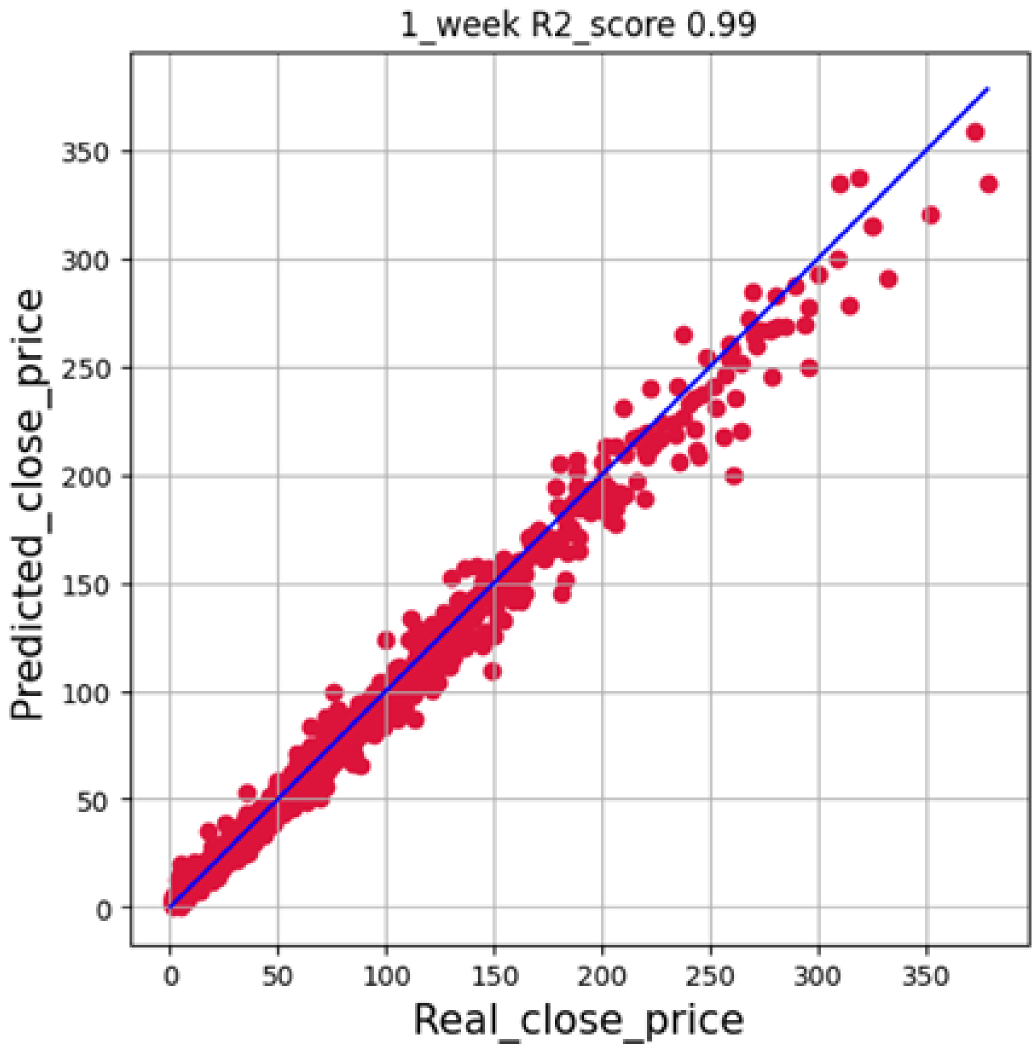


Figure 24: The scatter plot for Experiment 3, achieving an impressive R^2 score of 0.99.

Stock	**Volatility (E1)**	**Error (E1)**	**Volatility (E2, E3)**	**Error (E2)**	**Error (E3)**
KAR	0.062628	3.093479	0.063323	14.623855	12.213029
HERD	0.024999	4.979856	0.025425	9.922360	4.081338
ESNT	0.037666	8.362080	0.037999	3.584887	1.481312
ADUS	0.054161	2.994905	0.054790	97.193541	108.135527
FTDS	0.026198	6.048334	0.026660	8.269088	2.506808
Average Error	**-**	**5.09573**	**-**	**26.7187**	**25.6836**

Table 2: Comparison of Experiment 1, Experiment 2, and Experiment 3 Volatility and Errors

5.4.2 Future Horizon Forecasting

Figure 25 highlights the model's varying performance across stocks during the 15 weeks. ASG demonstrated exceptional predictive alignment, with both actual and predicted values remaining stable throughout, consistent with its low table error of 0.0378. Similarly, MAV maintained close alignment across the horizon, supported by a low error of 0.2277 in the table. FCO displayed minor discrepancies, particularly in the later weeks, yet its error of 0.1454 remained relatively low. In contrast, CASY showed significant divergence, especially in June, with actual values consistently falling below predictions, contributing to a much higher error of 101,779.30. The histogram (Figure 26) reveals a concentrated distribution of errors near 0, reflecting consistent performance for many stocks. However, a notable spread of larger deviations signals the occasional misprediction. The performance metrics reflect this, with an MAE of 46.41, RMSE of 72.61, and MSE of 5272.15, comparable to Experiment 2, suggesting similar overall predictive accuracy. Outliers, such as CACI (141,328.96) and HUBB (125,589.31), heavily influence the edges of the histogram, indicating areas where predictions deviated significantly. However, stocks like KFFB (0.017) and GOVX (0.120) are clustered near the peak,

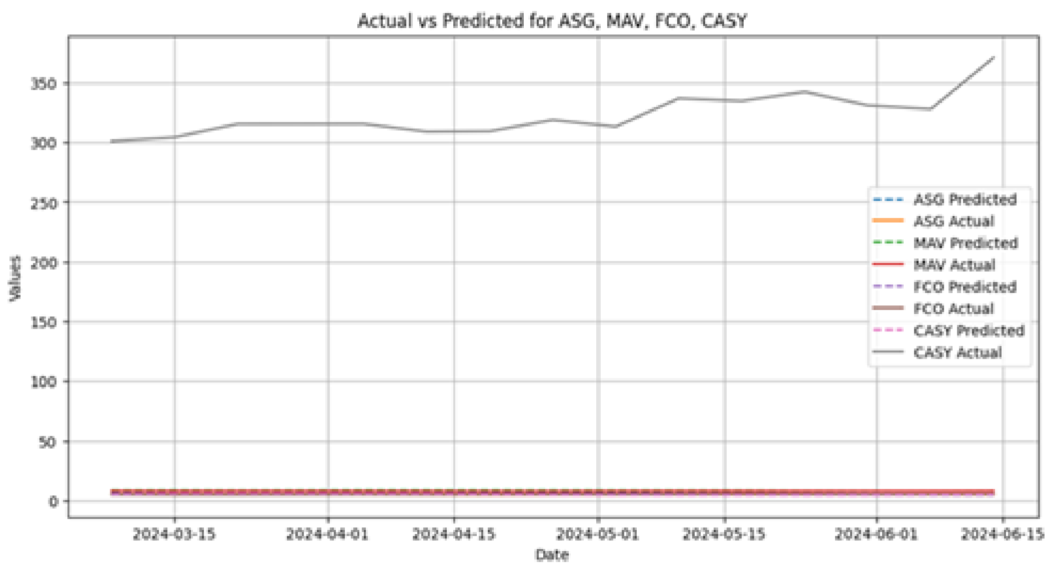


Figure 25: Actual vs Predicted values for ASG, MAV, FCO, and CASY

highlighting the model's capability to deliver highly accurate predictions for some stocks.

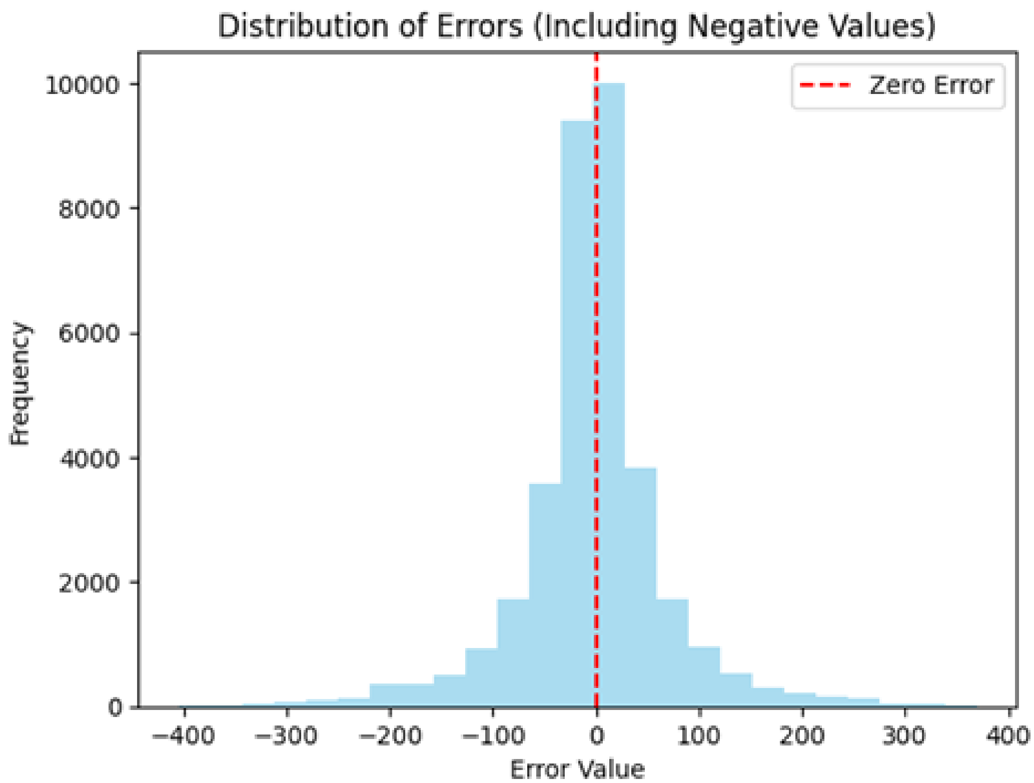


Figure 26: A histogram showing a wider distribution of error.

5.4.3 Smaller Unseen Data Testing

A correlation coefficient of 0.452 was recorded between volatility and prediction error, showing a slightly higher positive relationship for volatile stocks with high errors evidenced by MBIO which had the highest error (23,232.14) and high volatility (0.1174). Similar trends were seen in highly volatile stocks like BNKD at 0.627 and an error of 615.22 and GOVX with a volatility of 0.489 and an error of 7,999.74. In contrast, stocks such as BG (0.402) and BDJ (0.798) from the Top 50 Lowest Error Stocks demonstrated the model's accuracy in predicting more stable and less volatile stocks.

Table 3: A Sample to the Top 50 Stocks with Highest Volatility

Stock	**Volatility**	**Real**	**Predicted**	**Error**
BNKD	0.626591	21.55	-111.03	615.22
DZSI	0.57247	1.37	-34.52	2619.87
MBIO	0.548532	0.21	-48.58	23232.14
APM	0.542266	4.59	62.80	1268.16
GOVX	0.488739	1.74	-137.46	7999.74

Table 4: A Sample of the Top 50 Stocks with Lowest Error

Stock	**Volatility**	**Real**	**Predicted**	**Error**
BG	0.042176	102.27	102.68	0.40
BDJ	0.016122	7.93	7.87	0.80
IDOG	0.017308	30.26	30.51	0.83
BAUG	0.012195	39.54	40.01	1.18
AEE	0.022810	69.60	70.44	1.21

6 Discussion

6.1 Introduction

This section evaluates the model's performance in addressing the research objectives, including its strengths and limitations in predicting stock prices. Key insights are discussed, including trends observed in the results and challenges related to volatility and accuracy. The findings provide a foundation for exploring enhancements, particularly in improving predictions for high-volatility stocks and refining the techniques for future work.

6.2 Discussion

This study explores the potential of the residual network in stock market forecasting, focusing on its performance during standard testing, future horizon forecasting, and smaller unseen data testing. As detailed in Chapter 5, the residual network outperformed baseline numerical models in predictive accuracy during standard testing, as demonstrated by metrics such as MSE and R^2 score. For instance, E3 achieved an R^2 score of 0.99, showing its capacity to forecast stock prices accurately. This aligns with previous research, such as Ramanathan and McDermott 2021, which highlighted the effectiveness of residual networks in learning spatial hierarchies and extracting high-level features. Similarly, S. Zhang et al. 2021 reported the strengths of CNNs in identifying intricate relationships and handling noisy data, attributes needed for predicting volatile and complex stock markets. The pre-processing approach adopted in this study, including column shuffling, Min-Max scaling, and data transformation into arrays and images, preserved essential numerical structures while reducing noise. Literature by Zhao et al. 2023 supports such strategies, which aim to balance the retention of critical data patterns with minimised noise introduction. This approach facilitated compatibility with GTFS, enabling spatial structuring of temporal data. GTFS proved effective, as shown in E2, which achieved the lowest MSE among the top 50 high-error cases. Figures 47 and 52 demonstrate how GTFS enhanced the detection of cyclical market behaviours and anomalies, particularly during volatile periods. However, limitations emerged in high volatility scenarios, as depicted in Figures 47 and 50, where a wider error spread prompts the need to refine the model.

Across all the experiments, the relationship between volatility and prediction errors was a recurring theme. E1's scatter plot (Figures 47) revealed that while the model performed well under low-volatility conditions, prediction errors increased significantly with rising volatility. This trend was repeated across the experiments, with E3 achieving the highest R^2 score (0.99) but still encountering challenges with high-volatility stocks. E2, which incorporated indices correlated with the NASDAQ and NYSE, demonstrated improved performance metrics compared to the other experiments, particularly for high-volatility stocks. The reduction in average errors among the top 50 stocks with the highest prediction errors (Tables 14 and 16) suggests that leveraging correlated features can enhance generalisation across diverse market conditions. Future horizon forecasting results emphasised the importance of adapting the model's learning strategy for extended temporal predictions. Models tested across varying forecast horizons revealed consistent generalisation for shorter horizons, as illustrated in Figures **??** and 55, while longer horizons introduced greater prediction errors. E1 outperformed E2 and E3 in this area; perhaps the simplicity of the data offered better adaptability, as extra-economic data may expose the model to more noise. Nonetheless, the limitations of long-term forecasting in E3 highlight the need for adjustments in learning dynamics to improve accuracy over extended periods.

Testing with smaller unseen data sets further exposed the model's inability to generalise well when input shapes differed from those used during training. The data had smaller dimensions, requiring adjustments to the saved model to accommodate the new input and output shapes. It has been assumed that this modification impacted the predictions, leading to significantly higher errors in comparison to the future horizon forecasting. The need to retrain the model on data sets with differing shapes to ensure improved performance and generalisation has been realised. Another contributing factor is the stock market's dynamic nature, as it behaves like a living organism, adapting to various environmental changes and stressors. Stocks are constantly added or removed, as evidenced by the reduction in the stock pool from 3,280 to 2,517 within nine months. This evolving cycle presents an additional challenge for the model, which must generalise across historical data and adapt to the constantly shifting market composition. This fluidity requires a flexible model architecture capable of learning and re-learning patterns as new stocks enter the market and existing ones leave.

The integration of correlated indices in E2 demonstrated the potential of leveraging auxiliary features to improve performance. However, persistent challenges with extreme volatility, the dynamic stock environment and differences in dataset shapes highlight the need for architectural enhancements. The plateauing of validation losses across all experiments further indicates residual overfitting, particularly in E3, where validation losses remained consistently higher than training losses (Figures 45 and 48). The residual network demonstrates significant promise for stock market forecasting, particularly for stable, low-volatility stocks. However, its limitations in managing high-volatility conditions, unseen data with different shapes, and the dynamic nature of the stock market require targeted refinements. Improvements in pre-processing, feature engineering, model architecture, and re-training using data sets with varying shapes and better dropout configurations could improve the model. E2's results suggest incorporating correlated indices and optimising learning strategies can strengthen the model's generalisation. However, E1, with limited data, performed the best overall; there needs to be a balance regarding the added features so as not to introduce more noise, as some stocks exhibiting low volatility also had high errors.

6.3 Future Work

This section will address possible ways to enhance model performance, particularly under volatile market conditions and further steps the research could take. Optimising the model for real-time use and adapting it to the dynamic stock market are critical avenues for future research. Retraining periodically to account for changing input shapes and addressing issues with unseen data would improve generalisation. Investigating the inclusion of daily stock data instead could help the model capture intra-week fluctuations to improve accuracy. Furthermore, identifying additional factors influencing stock market behaviour beyond volatility may address high errors in low-volatility stocks. The study also highlights the importance of refining training processes to reduce overfitting. Adjustments to dropout placements and rates and experiments with learning schedules could improve model stability. Expanding the dataset to include broader indices and macroeconomic variables could enhance the model's ability to account for external factors driving volatility, but a balanced approach needs to be taken as it could result in more noise. While forecasting remains inherently uncertain, this research demonstrates that residual network-based 2D CNNs provide a promising foundation. Future studies can build upon this foundation by integrating adaptive mechanisms and refining feature engineering techniques to develop accurate and adaptable models to the complexities of real-world financial markets.

7 Conclusions and Reflection on Learning

The primary aim of this dissertation was to evaluate the effectiveness of residual network-based 2D CNN models in forecasting NASDAQ and NYSE stock prices by transforming numerical data into an image. Through the implementation and testing of three experimental setups, the research has achieved many of its original objectives but has also revealed areas for improvement. The results demonstrate that the residual network approach offers significant potential for extracting spatial and temporal patterns from stock data, particularly in stable, low-volatility conditions, as evidenced by Experiment 3, which achieved an impressive R^2 score of 0.99. The integration of additional indices in Experiment 2 further highlighted the capability of extra features to improve performance, especially for high-volatility stocks, aligning with the original objective of exploring the model's ability to generalise across diverse market conditions. However, the study also exposed some limitations; mainly, the model's sensitivity to extreme volatility remained challenging, where high volatility led to significant prediction errors. Furthermore, some stocks with low volatility also had high errors, which was a surprise. While unsurprising, the inability to generalise well to smaller unseen datasets with varying input shapes demonstrates a limitation in the model's performance, indicating that the objective of creating a highly robust framework was not entirely achieved. Surprisingly, the experiment that performed best overall had the lowest R^2 score because the model performance improved with more data during standard testing. More time was needed to retrain the model and experiment with different layers to see whether the model improved as well as feature engineering, as it was believed that some of the economic data may have introduced more noise to the model. Several areas could have been explored further, as getting lost in the data and the training process to improve the model was very easy, and the desire to keep pushing for a better model performance took away much time. Comparing the residual network with a 1D CNN would have been interesting to test how it compares to a traditional DL model that only accepts numerical data and ascertain whether the image approach is superior at finding unique patterns. Although it was advised not to bother, the decision to test the model under various conditions to emulate the real world was then taken. Testing the model with daily data instead of weekly might have also revealed finer missed temporal patterns. It was realised through hypertuning that the model did not perform well when the kernel was changed from 1,1 to 2,2 and it was difficult to get the right training schedule for the model training and validation loss to be a smooth curve. Despite the model's limitations, this dissertation is believed to contribute to the industry by presenting a novel approach to forecasting the stock market.

Glossary

Residual Network (ResNet) A deep neural network architecture that uses skip connections to mitigate vanishing gradient problems and improve training for deeper networks.

Grid-Based Temporal-Feature Structuring (GTFS) A method of transforming numerical financial data into grid-like image structures, capturing temporal sequences and feature interrelationships for analysis by convolutional neural networks.

Volatility A statistical measure of the dispersion of returns for a given stock or market index, often used as an indicator of risk.

Convolutional Neural Network (CNN) A type of deep learning model designed to process structured data like images, capable of recognising spatial and temporal patterns.

Mean Squared Error (MSE) A metric that measures the average squared difference between actual and predicted values, penalising larger errors more heavily.

Mean Absolute Error (MAE) A metric that measures the average absolute difference between actual and predicted values, treating all errors equally.

Industrial Production (IP) An economic indicator that measures the output of the industrial sector, including manufacturing, mining, and utilities.

Money Supply (WM2NS) A measure of the total amount of monetary assets available in an economy at a specific time.

Adjusted Closing Price A stock's closing price after adjustments for corporate actions such as dividends and stock splits.

Dropout A regularisation technique in neural networks that randomly omits certain neurons during training to prevent overfitting.

Min-Max Scaling A normalisation method that scales data to a specific range, often between 0 and 1, to ensure uniformity across features.

Pearson's Correlation A statistical measure of the linear relationship between two variables, ranging from -1 (perfect negative correlation) to +1 (perfect positive correlation).

Rolling Window A method of analysing time-series data by calculating metrics over a sliding window of observations.

ReLU (Rectified Linear Unit) An activation function in neural networks that outputs the input directly if positive, otherwise outputs zero.

Skip Connections Connections in residual networks that bypass one or more layers, allowing gradients to flow directly to earlier layers and improving training.

Table of Abbreviations

Abbreviation	Full Form
ARIMA	Autoregressive Integrated Moving Average
SARIMA	Seasonal Autoregressive Integrated Moving Average
SARIMAX	Seasonal Autoregressive Integrated Moving Average with Exogenous Variables
GARCH	Generalized Autoregressive Conditional Heteroskedasticity
EMD	Empirical Mode Decomposition
LSTM	Long Short-Term Memory
CNN	Convolutional Neural Network
GTFS	Grid-Based Temporal-Feature Structuring
MSE	Mean Squared Error
MAE	Mean Absolute Error
FTSE	Financial Times Stock Exchange
NYSE	New York Stock Exchange
NASDAQ	National Association of Securities Dealers Automated Quotations
DJI	Dow Jones Industrial Average
GDAXI	German DAX Index
IP	Industrial Production
WM2NS	Money Supply

Table 5: Table of Abbreviations

References

Abudy, M. M. (2020). “Retail investors’ trading and stock market liquidity”. In: *The North American Journal of Economics and Finance* 54. DOI: 10.1016/j.najef.2020.101281.

Ali, Muhammad et al. (2023). “Prediction of Complex Stock Market Data Using an Improved Hybrid EMD-LSTM Model”. In: *Applied Sciences* 13.1429, pp. 1–24. DOI: 10.3390/app13031429.

Alvarez-Ramirez, J., E. Rodriguez, and C. Ibarra-Valdez (2020). “Medium-term cycles in the dynamics of the Dow Jones Index for the period 1985–2019”. In: *Physica A* 546, pp. 1–11. DOI: 10.1016/j.physa.2019.124017.

Alzoubi, M. (2022). “Stock market performance: Reaction to interest rates and inflation rates”. In: *Banks and Bank Systems* 17.2, pp. 189–198. DOI: 10.21511/bbs.17(2).2022.16.

Anand, Gaurangi and Richi Nayak (2020). “DeLTa: Deep local pattern representation for time-series clustering and classification using visual perception”. In: *Knowledge-Based Systems* 212, pp. 1–16. DOI: 10.1016/j.knosys.2020.106551.

Antoniades, A. (2016). “Liquidity Risk and the Credit Crunch of 2007–2008: Evidence from Micro-Level Data on Mortgage Loan Applications”. In: *Journal of Financial and Quantitative Analysis* 51.6, pp. 1795–1822. DOI: 10.1017/S0022109016000740.

Arseneau, D. M., W. Rappoport, and A. P. Vardoulakis (2020). “Private and public liquidity provision in over-the-counter markets”. In: *Theoretical Economics* 15.4, pp. 1669–1712. DOI: 10.3982/TE3419.

Arshad, S., S. A. Rizvi, and O. Haroon (2019). “Impact of Brexit vote on the London stock exchange: A sectorial analysis of its volatility and efficiency”. In: *Finance Research Letters* 34, pp. 1–4. DOI: 10.1016/j.frl.2019.07.013.

Baker, Scott R. et al. (2020). *The Unprecedented Stock-Market Reaction to COVID-19*. Available at: https://insight.kellogg.northwestern.edu/article/what-explains-the-unprecedented-stock-market-reaction-to-covid-19 [Accessed 04 December 2024].

Barnes, B. (2008). *Disney and Pixar: The Power of the Prenup*. https://www.nytimes.com/2008/06/01/business/media/01pixar.html. Accessed: 19 October 2024.

Barra, S. et al. (2020). "Deep Learning and Time Series-to-Image Encoding for Financial Forecasting". In: *CAA Journal of Automatica Sinica* 7.3, pp. 683–692. DOI: 10.1109/jas.2020.1003132.

Bhandari, Hum Nath et al. (2022). "Predicting stock market index using LSTM". In: *Machine Learning with Applications* 9, pp. 1–15. DOI: 10.1016/j.mlwa.2022.100320.

Boniol, Paul et al. (2022). "dCAM: Dimension-wise Class Activation Map for Explaining Multivariate Data Series Classification". In: pp. 1175–1189. DOI: 10.1145/3514221.3526183.

Cai, Wei et al. (2024). "An explainable dual-mode convolutional neural network for multivariate time series classification". In: *Knowledge-Based Systems* 299, pp. 1–15. DOI: 10.1016/j.knosys.2024.112015.

Carvalho Tavares, Thiago Henrique Barbosa de, Bruno Pérez Ferreira, and Eduardo Mazoni Andrade Marçal Mendes (2022). "Fuzzy time series model based on red–black trees for stock index forecasting". In: *Applied Soft Computing* 127, pp. 1–14. DOI: 10.1016/j.asoc.2022.109323.

Daryl, Aditya Winata, Sena Kumara, and Derwin Suhartono (2021). "Predicting Stock Market Prices using Time Series SARIMA". In: pp. 92–99. DOI: 10.1109/ICCSAI53272.2021.9609720.

Debnath, T. K. (2022). "The Effect of Chaos Theory in the Field of Business: A Review". In: *International Journal of Progressive Research in Science and Engineering* 3.5, pp. 88–92. URL: https://www.researchgate.net/

publication/363354624_The_Effect_of_Chaos_Theory_in_the_Field_of_Business_A_Review.

Durairaj, M. and B. H. Krishna Mohan (2022). “A convolutional neural network based approach to financial time series prediction”. In: *Neural Computing and Applications* 34, pp. 13319–13337. DOI: 10.1007/s00521-022-07143-2.

Ehedu, V. C. and K. Obi (2022). “Efficient Market Hypothesis (EMH) and the Nigerian Stock Exchange In The Midst Of Global Financial Crisis”. In: *International Journal of Academic Management Science Research* 6.8, pp. 263–273. URL: https://www.researchgate.net/profile/Callistar-Obi/publication/364304726_Efficient_Market_Hypothesis_EMH_and_the_Nigerian_Stock_Exchange_In_The_Midst_Of_Global_Financial_Crisis/links/634459dcff870c55ce164c58/Efficient-Market-Hypothesis-EMH-and-the-Nigerian-Stock-Exchange-In-The-Midst-Of-Global-Financial-Crisis.pdf.

Elliott, T. et al. (2024). “The Impact of Terrorism on Liquidity and Information Asymmetry: Evidence from Non-U.S. Stocks Listed on the NYSE”. In: *Defence and Peace Economics* 1, pp. 1–7. DOI: 10.1080/10242694.2024.2332168.

Elsayed, A. H., S. Nasreen, and A. K. Tiwari (2020). “Time-varying co-movements between energy market and global financial markets: Implication for portfolio diversification and hedging strategies”. In: *Energy Economics* 90. DOI: 10.1016/j.eneco.2020.104847.

Grybauskas, A., A. Stefanini, and M. Ghobakhloo (2022). “Social sustainability in the age of digitalization: A systematic literature review on the social implications of industry 4.0”. In: *Technology in Society* 70. DOI: 10.1016/j.techsoc.2022.101997.

Gülmez, Burak (2023). "Stock price prediction with optimized deep LSTM network with artificial rabbits optimization algorithm". In: *Expert Systems With Applications* 227, pp. 1–16. DOI: `10.1016/j.eswa.2023.120346`.

Hale, Kori (2021). *Nasdaq Secures SEC Approval To Increase Board Diversity*. Forbes. Available at: `https://www.forbes.com/sites/korihale/2021/08/18/nasdaq-secures-sec-approval-to-increase-board-diversity/` [Accessed 28 November 2024].

He, Houtian et al. (2021). "A seasonal-trend decomposition-based dendritic neuron model for financial time series prediction". In: *Applied Soft Computing* 108, pp. 1568–4946. DOI: `https://doi.org/10.1016/j.asoc.2021.107488`.

Homenda, Władysław et al. (2024). "Time series classification with their image representation". In: *Neurocomputing* 573, pp. 1–15. DOI: `10.1016/j.neucom.2023.127214`.

Hooker, Sara et al. (2023). *The Hardware Lottery Revisited: Fairness in Machine Learning and Computational Inefficiencies*. DOI: `https://doi.org/10.48550/arXiv.2312.03886`.

Imbert, F. (2020). *Nasdaq jumps more than 2% led by Amazon, S&P 500 goes positive for 2020*. `https://www.cnbc.com/2020/07/19/stock-market-futures-open-to-close-news.html`. Accessed: 19 October 2024.

Jarrah, Mutasem and Morched Derbali (2023). "Predicting Saudi Stock Market Index by Using Multivariate Time Series Based on Deep Learning". In: *Applied Sciences* 13.8356, pp. 1–18. DOI: `10.3390/app13148356`.

Jeong, Seongyoep et al. (2021). "Temperature Prediction Based on Bidirectional Long Short-Term Memory and Convolutional Neural Network Combining Observed and Numerical Forecast Data". In: *Sensors* 21.941, pp. 1–20. DOI: `10.3390/s21030941`.

Jerez, T. and W. Kristjanpoller (2020). "Effects of the validation set on stock returns forecasting". In: *Expert Systems with Applications* 150. DOI: 10.1016/j.eswa.2020.113271.

Jones, C. and T. H. J. Dean (2020). "Worst day for markets since Black Monday crash of 1987; Worst crash since 1987". In: *The Times*, pp. 1–2.

Kassel, Matthew (2016). "Once-Popular NYSE Has No Plan to Reopen to the Public". In: *The Wall Street Journal*. URL: https://www.wsj.com/articles/once-popular-nyse-has-no-plan-to-reopen-to-the-public-1470621840.

Kenourgios, D., E. Dadinakis, and I. Tsakalos (2020). "Brexit referendum and European stock markets: a sector analysis". In: *Managerial Finance* 46.7, pp. 913–933. DOI: 10.1108/MF-07-2019-0366.

Khanarsa, Paisit, Arthorn Luangsodsa, and Krung Sinapiromsaran (2020). "Self-Identification ResNet-ARIMA Forecasting Model". In: *WSEAS Transactions on Systems and Control* 15, pp. 196–211. DOI: 10.37394/23203.2020.15.21.

Kim, J. C., S. Mazumder, and Q. Su (2024). "Brexit's ripple: Probing the impact on stock market liquidity". In: *Finance Research Letters* 61, pp. 1–10. DOI: 10.1016/j.frl.2024.105030.

King, T. A. and T. J. Fogarty (2022). "The dramaturgy of earnings guidance: an institutional analysis of a soft landing". In: *Accounting, Auditing & Accountability Journal* 35.4, pp. 1093–1120. DOI: 10.1108/AAAJ-06-2017-2983.

Kirisci, M. and E. Yolcu (2023). "A CNN-Based Model for Financial Time Series Forecasting". In: *Neural Computing and Applications* 35.5, pp. 4231–4245. DOI: 10.1007/s11063-022-10767-z.

Kirisci, Mehmet and O. Cagcag Yolcu (2022). "A New CNN-Based Model for Financial Time Series: TAIEX and FTSE Stocks Forecasting". In: *Neural Processing Letters* 54, pp. 3357–3374. DOI: 10.1007/s11063-022-10767-z.

Kirsch, J. (2024). *What Are The Magnificent Seven Stocks?* https://www.forbes.com/sites/investor-hub/article/what-are-magnificent-seven-stocks/. Accessed 19 October 2024.

Klebnikov, S. (2021). *Explosive Growth During The Coronavirus Pandemic*. https://www.forbes.com/sites/sergeiklebnikov/2020/07/23/5-big-numbers-that-show-amazons-explosive-growth-during-the-coronavirus-pandemic/. Accessed 19 October 2024.

Kosakowski, P. (2024). *The Fall of the Market in the Fall of 2008*. https://www.investopedia.com/articles/economics/09/subprime-market-2008.asp. Accessed 16 October 2024.

Kumar, G., S. Jain, and U. P. Singh (2021). "Stock Market Forecasting Using Computational Intelligence: A Survey". In: *Archives of Computational Methods in Engineering* 28, pp. 1069–1101. DOI: 10.1007/s11831-020-09413-5.

Kumar, Gourav, Uday Pratap Singh, and Sanjeev Jain (2021). "Hybrid evolutionary intelligent system and hybrid timeseries econometric model for stock price forecasting". In: *International Journal of Intelligent Systems* 36, pp. 4902–4935. DOI: 10.1002/int.22495.

Lee, D., R. Waters, and D. P. Mancini (2020). *Prospering in the pandemic: the top 100 companies*. [Accessed 19 October 2024]. URL: https://www.ft.com/content/844ed28c-8074-4856-bde0-20f3bf4cd8f0.

Lee, Si Woon and Ha Young Kim (2020). "Stock market forecasting with super-high dimensional time-series data using ConvLSTM, trend sampling, and specialized data augmentation". In: *Expert Systems with Applications* 161, pp. 1–20. DOI: 10.1016/j.eswa.2020.113704.

Li, Guanzhi et al. (2022). "Pearson Correlation Coefficient-Based Performance Enhancement of Broad Learning System for Stock Price Prediction". In: *IEEE Transactions on Circuits and Systems II: Express Briefs* 69.5, pp. 2413–2417. DOI: 10.1109/TCSII.2022.3160266.

Li, Jingjing and Qiang Wang (2022). “Comparison of the representational ability in individual difference analysis using 2-D time-series image and time-series feature patterns”. In: *Expert Systems With Applications* 215, pp. 1–25. DOI: 10.1016/j.eswa.2022.119429.

Lu, Minrong and Xuerong Xu (2023). “TRNN: An efficient time-series recurrent neural network for stock price prediction”. In: *Information Sciences* 657, pp. 1–18. DOI: 10.1016/j.ins.2023.119951.

Lv, P. et al. (2022). “Stock Index Prediction Based on Time Series Decomposition and Hybrid Model”. In: *Entropy* 24.146, pp. 1–18. DOI: 10.3390/e24020146.

Maloumian, N. (2022). “Unaccounted forms of complexity: A path away from the efficient market hypothesis paradigm”. In: *Social Sciences Humanities Open* 5.1. DOI: 10.1016/j.ssaho.2021.100244.

Maurin, V. (2022). “Liquidity Fluctuations in Over-the-Counter Markets”. In: *The Journal of Finance* 77.2, pp. 1325–1369. DOI: 10.1111/jofi.13106.

Mazumdar, K., D. Zhang, and Y. Guo (2020). “Portfolio selection and unsystematic risk optimisation using swarm intelligence”. In: *Journal of Banking and Finance Technology* 4.1, pp. 1–14. DOI: 10.1007/s42786-019-00013-x.

Mehtab, S. and J. Sen (2022). “Analysis and Forecasting of Financial Time Series Using CNN and LSTM-Based Deep Learning Models”. In: vol. 302. DOI: 10.1007/978-981-16-4807-6_39.

Mishkin, F. S. and S. G. Eakins (2016). *Financial markets and institutions*. 8th. Boston: Pearson.

Moon, Jihye, Md Billal Hossain, and Ki H. Chon (2021). “AR and ARMA model order selection for time-series modeling with ImageNet classification”. In: *Signal Processing* 183, pp. 1–11. DOI: 10.1016/j.sigpro.2021.108026.

Navyatha, K. and C. Gaddam (2022). “Review Of Literature On Ipos- A Study”. In: *International Journal of Research and Analytical Reviews* 9.2.

Nickerson, I. and C. Rarick (2010). “The Merger of AOL and Time Warner: A case study”. In: *Journal of the International Academy for Case Studies* 16.7, pp. 103–110. URL: https://www.researchgate.net/publication/291711379_The_merger_of_AOL_and_Time_Warner_A_case_study.

Nti, I. K., A. F. Adekoya, and B. A. Weyori (2019). “A systematic review of fundamental and technical analysis of stock market predictions”. In: *Artif Intell Rev* 53, pp. 3007–3057. DOI: 10.1007/s10462-019-09754-z.

Ramanathan, Amrutha and James McDermott (2021). “Fall detection with accelerometer data using Residual Networks adapted to multi-variate time series classification”. In: pp. 1–8. DOI: 10.1109/IJCNN52387.2021.9533670.

Rokhsatyazdi, Ehsan et al. (2020). “Optimizing LSTM Based Network For Forecasting Stock Market”. In: pp. 1–7. DOI: 10.1109/CEC48606.2020.9185545.

Ruan, Jinlong, Wei Wu, and Jiebo Luo (2020). “Stock Price Prediction Under Anomalous Circumstances”. In: pp. 4787–4794. DOI: 10.1109/BigData50022.2020.9378030.

Sayed, Aya Nabil, Yassine Himeur, and Faycal Bensaali (2023). “From time-series to 2D images for building occupancy prediction using deep transfer learning”. In: *Engineering Applications of Artificial Intelligence* 119, pp. 1–11. DOI: 10.1016/j.engappai.2022.105786.

Semenoglou, Artemios-Anargyros, Evangelos Spiliotis, and Vassilios Assimakopoulos (2022). “Image-based time series forecasting: A deep convolutional neural network approach”. In: *Neural Networks* 157, pp. 39–53. DOI: 10.1016/j.neunet.2022.10.006.

Shahzad, K. et al. (2018). “UK’s stock market reaction to Brexit process: A tale of two halves”. In: *Economic Modelling* 80, pp. 275–283. DOI: 10.1016/j.econmod.2018.11.013.

Sharma, Suraj Prakash, Jeyanthi R., and K. Deepa (2022). "Forecasting India SP BSE SENSEX and USA SP-500 Benchmark Indices Using SARIMAX and Facebook Prophet Library". In: pp. 1523–1530. DOI: 10.1109/ICICCS53718.2022.9788127.

Shu, M., R. Song, and W. Zhu (2021). "The 'COVID' crash of the 2020 U.S. Stock market". In: *The North American Journal of Economics and Finance* 58. DOI: 10.1016/j.najef.2021.101497.

Song, Hyunsun and Hyunjun Choi (2023). "Forecasting Stock Market Indices Using the Recurrent Neural Network Based Hybrid Models: CNN-LSTM, GRU-CNN, and Ensemble Models". In: *Applied Sciences* 13.4644, pp. 1–26. DOI: 10.3390/app13074644.

Statista (2024). *New York Stock Exchange (NYSE)*. [Accessed 15 October 2024]. URL: https://www-statista-com.lsbu.idm.oclc.org/study/117133/new-york-stock-exchange-nyse/.

Tang, Chen and Yanlin Shi (2021). "Forecasting High-Dimensional Financial Functional Time Series: An Application to Constituent Stocks in Dow Jones Index". In: *Journal of Risk and Financial Management* 14.343, pp. 1–13. DOI: 10.3390/jrfm14080343.

Trinh, Q. D., C. Haddad, and K. T. Tran (2022). "Financial reporting quality and dividend policy: New evidence from an international level". In: *International Review of Financial Analysis* 80. DOI: 10.1016/j.irfa.2022.102026.

Tripathy, Nrusingha et al. (2023). "Deep Learning-based Gated Recurrent Unit Approach to Stock Market Forecasting: An Analysis of Intel's Stock Data". In: *International Journal of Smart Sensor and Adhoc Network* 4.1, pp. 19–25. DOI: 10.47893/IJSSAN.2023.1234.

Varadharajan, Vivek et al. (2024). "Stock Closing Price and Trend Prediction with LSTM-RNN". In: *Journal of Artificial Intelligence and Big Data* 4.877, pp. 1–13. DOI: 10.31586/jaibd.2024.877.

Wang, Yan and Yuankai Guo (2020). “Forecasting Method of Stock Market Volatility in Time Series Data Based on Mixed Model of ARIMA and XGBoost”. In: *IEEE* 17.3, pp. 205–221. DOI: 10.23919/JCC.2020.03.017.

Woode, J. K., P. Owusu Junior, and A. M. Adam (2024). “Dynamic interdependence structure of industrial metals and the African stock market”. In: *Resources Policy* 88. DOI: 10.1016/j.resourpol.2023.104455.

Wright, Nicholas D. (2021). “Neither Triumph nor Disaster: United Kingdom Responses to COVID-19 and the Future of National Security”. In: *PRISM* 9.4, pp. 134–154. ISSN: 21570663, 21570671. URL: https://www.jstor.org/stable/48697196 (visited on 10/19/2024).

Wu, Junran et al. (2021). “Price graphs: Utilizing the structural information of financial time series for stock prediction”. In: *Information Sciences* 588, pp. 405–424. DOI: 10.1016/j.ins.2021.12.089.

Wu, Yupeng et al. (2023). “An Aggregated Convolutional Transformer Based on Slices and Channels for Multivariate Time Series Classification”. In: *IEEE* 7.3, pp. 768–779. DOI: 10.1109/TETCI.2022.3210992.

Xu, G. et al. (2024). “Development of Skip Connection in Deep Neural Networks for Computer Vision and Medical Image Analysis: A Survey”. In: *arXiv preprint*. URL: https://arxiv.org/abs/2405.01725.

Yang, Fuwei, Jingjing Chen, and Yicen Liu (2021). “Improved and optimized recurrent neural network based on PSO and its application in stock price prediction”. In: *Soft Computing* 27, pp. 3461–3476. DOI: 10.1007/s00500-021-06113-.

Yeh, Chin-Chia Michael et al. (2023). “Multitask Learning for Time Series Data with 2D Convolution”. In: pp. 9–16. DOI: 10.1109/icmla58977.2023.00010.

Zaheer, Shahzad et al. (2023). “A Multi Parameter Forecasting for Stock Time Series Data Using LSTM and Deep Learning Model”. In: *Mathematics* 11.590, pp. 1–24. DOI: 10.3390/math11030590.

Zakhidov, G. (2024). “Economic Indicators: Tools for Analyzing Market Trends and Predicting Future Performance”. In: *International Multidisciplinary Journal of Universal Scientific Prospectives* 2.3, pp. 23–29.

Zhang, S. et al. (2021). “Measurement of capital allocation efficiency in emerging economies: evidence from China”. In: *Technological Forecasting and Social Change* 171. DOI: 10.1016/j.techfore.2021.120954.

Zhang, Y. et al. (2023). “Geopolitical risk and stock market volatility: A global perspective”. In: *Finance Research Letters* 53. DOI: 10.1016/j.frl.2022.103620.

Zhao, Cheng et al. (2023). “Stock Market Analysis Using Time Series Relational Models for Stock Price Prediction”. In: *Mathematics* 11, pp. 1–14. DOI: 10.3390/math11051130.

A Extra Tables

Table 6: Summary Statistics for Selected Stock Data

	A	AA	AAAU	AADI	AADR	AAL	AAMC	AAME	AAOI	AAON
count	195	195	195	195	195	195	195	195	195	195
mean	128.80	37.99	18.50	15.95	54.07	16.09	13.67	2.78	7.17	48.72
std	18.80	16.80	0.98	8.47	5.83	3.29	10.78	0.86	4.73	11.79
min	83.82	10.66	16.31	1.63	43.09	10.92	2.69	1.60	1.56	32.12
25%	117.92	27.10	17.76	8.08	48.87	13.63	8.31	2.10	2.71	38.87
50%	128.86	36.16	18.41	15.35	53.70	15.46	10.58	2.52	7.24	45.56
75%	139.76	47.47	19.29	22.28	59.27	18.14	13.81	3.16	9.64	56.84
max	176.38	90.02	20.63	37.50	65.52	24.97	57.49	5.59	21.73	84.30

	NEWTL	NLTX	OPOF	RNLC	SMH	SOHON	SOHOO	VSAT	WCLD	WDC
count	195	195	195	195	195	195	195	195	195	195
mean	22.55	100.89	20.04	29.79	124.04	16.53	16.36	38.40	39.68	49.36
std	1.21	100.98	3.65	3.33	27.16	4.94	4.96	10.26	11.54	11.59
min	19.55	8.40	13.52	20.74	70.82	4.60	4.59	15.53	23.55	30.59
25%	21.75	15.41	16.97	28.46	107.30	14.69	14.10	32.04	29.68	39.58
50%	22.68	31.70	20.60	30.59	122.89	17.40	17.07	37.24	35.80	47.09
75%	23.40	208.40	23.03	31.80	142.68	20.15	20.05	46.74	49.51	57.64
max	24.93	338.80	27.08	36.80	220.54	23.51	22.87	67.00	64.21	77.17

Table 7: Top 10 Stocks with Most Outliers

Column Name	**Count**
ATEST	14
CCLDP	12
CRBP	11
ASMB	10
KEQU	10
BLBX	8
GLPG	8
ABIO	8
CYTK	7
ENG	7

Table 8: Bottom 10 Stocks with Fewest Outliers

Column Name	**Count**
BTOG	1
CDXC	1
CARV	1
GLOB	1
CARG	1
BZQ	1
BYD	1
GROW	1
GRPN	1
DUST	1

Table 9: Financial Data (Part 1)

Index	**Date**	**Adj_Close_DJI**	**Adj_Close_FTSE**	**Adj_Close _GDAXI**
0	2020-06-12	25605.539	6105.200	11949.280
1	2020-06-15	25763.160	6064.700	11911.350
2	2020-06-16	26289.980	6242.800	12315.660
3	2020-06-17	26119.609	6253.300	12382.140
4	2020-06-18	26080.100	6224.100	12281.530

Table 10: Financial Data (Part 2)

Adj_Close_GSPC	**Adj_Close_RUT**	**CPI**	**WM2NS**	**IP**
3041.310	1387.680	257.004	18129.2	91.6745
3066.590	1419.610	257.004	18129.2	91.6745
3124.740	1452.260	257.004	18129.2	91.6745
3113.490	1426.530	257.004	18129.2	91.6745
3115.340	1427.330	257.004	18129.2	91.6745

Stock	**Volatility (E1)**	**Real (E1)**	**Predicted (E1)**	**Error (E1)**	**Volatility (E2)**	**Real (E2)**	**Predicted (E2)**	**Error (E2)**	**Volatility (E3)**	**Real (E3)**	**Predicted (E3)**	**Error (E3)**
KAR	0.063	15.140	15.608	3.093	0.063	136.55	116.58	14.624	0.063	136.55	119.87	12.21
HERD	0.025	36.808	34.975	4.980	0.025	30.987	27.913	9.922	0.025	30.987	29.723	4.081
ESNT	0.038	53.200	48.751	8.362	0.038	19.985	19.269	3.585	0.038	19.985	20.281	1.481
ADUS	0.054	96.350	93.464	2.995	0.055	2.340	4.614	97.194	0.055	2.340	4.870	108.14
FTDS	0.026	46.190	43.396	6.048	0.027	54.943	50.400	8.269	0.027	54.943	53.566	2.507
Average Error	-	-	-	**5.096**	-	-	-	**26.719**	-	-	-	**25.684**

Table 11: Detailed Comparison of Volatility, Real and Predicted Prices, and Errors for Experiment 1, 2, and 3

Stock	Volatility	Real	Predicted	Error	SE
ALLK	0.515203	2.65	3.178916	19.959093	0.279752
CLEU	0.445323	2.385	13.661012	472.788750	127.148447
KOD	0.389507	3.04	5.375727	76.833115	5.455621
IMTE	0.374204	2.61	6.206541	137.798490	12.935107
CRBP	0.322030	5.5	6.225388	13.188874	0.526188
IMUX	0.303182	1.44	1.605500	11.493030	0.027390
LTRY	0.273063	2.8	4.796227	71.293838	3.984922
ATXS	0.264678	6.79	9.265766	36.461946	6.129417
AAMC	0.230315	4.35	28.807968	562.252141	598.192199
EIGR	0.225331	6.189	10.114609	63.428805	15.410406
... Remaining rows omitted for brevity ...					
Average	0.211204	6.67058	8.665292	95.942954	48.586632

Table 12: Top 50 stocks with the highest volatility, along with their real, predicted values, errors, and standard errors (SE) for **Experiment 1**.

Stock	Volatility	Real	Predicted	Error	SE
AAMC	0.230315	4.35	28.807968	562.252141	598.1921987
ATXG	0.119233	1.13	6.772195	499.309322	31.83436442
CAMP	0.114378	5.3061	31.150562	487.070773	667.9362161
CLEU	0.445323	2.385	13.661012	472.788750	127.1484466
DOGZ	0.184245	3.37	18.18301	439.555196	219.4252653
FAMI	0.203164	1.25	6.211551	396.924057	24.61698833
CUTR	0.130528	3.03	13.611702	349.231088	111.9724172
BTAI	0.187015	3.03	13.433203	343.340025	108.2266327
EGRX	0.08664	4.96	21.111881	325.642767	260.8832598
BW	0.146026	1.44	5.800993	302.846767	19.01825995
... Remaining rows omitted for brevity ...					
Average	0.142502	4.286122	12.777432	222.590589	126.1603671

Table 13: Top 50 stocks with the highest prediction errors, along with their real, predicted values, errors, and standard errors (SE) for **Experiment 1**.

Stock	**Volatility**	**Real**	**Predicted**	**Error**	**SE**
ALLK	0.520555	41.3615	39.352409	4.857393	4.036446646
CLEU	0.45078	39.3848	35.983761	8.635411	11.56706628
KOD	0.393908	33.17	28.404119	14.368045	22.71362171
IMTE	0.334842	6.11	6.068199	0.684146	0.001747324
CRBP	0.325413	23.8386	22.636244	5.043737	1.445659951
IMUX	0.305878	2.1366	1.95224	8.628646	0.03398861
LTRY	0.275614	37.6253	34.294376	8.852883	11.09505469
ATXS	0.267562	19.14	17.004473	11.157405	4.560475568
AAMC	0.232731	33.3833	30.869539	7.529995	6.318994365
EIGR	0.22772	99.2	96.770424	2.449169	5.90283954
... Remaining rows omitted for brevity ...					
Average	0.21253844	38.83137	36.6408813	9.853995	32.92852535

Table 14: Top 50 stocks with the highest volatility, including real and predicted values, error, and standard error (SE) for **Experiment 2**.

Stock	**Volatility**	**Real**	**Predicted**	**Error**	**SE**
IG	0.010411	4.763	16.253256	241.239888	132.025982
AGCO	0.050136	4.63	12.579053	171.685808	63.187443
EDN	0.076645	1.275	3.092165	142.522763	3.30208864
BILL	0.103582	2.4585	5.869095	138.726656	11.6321583
AGNG	0.023521	3.3344	7.826067	134.706917	20.1750724
HIE	0.029804	2.9862	6.834128	128.857022	14.8065499
IXC	0.040349	6.627	14.366182	116.782591	59.8949380
CQP	0.047735	5.27	10.92454	107.296766	31.9738226
MBSD	0.006197	2.83	5.621246	98.630613	7.79105423
ADUS	0.05479	2.34	4.614329	97.193541	5.1725724
... Remaining rows omitted for brevity ...					
Average	0.05344762	6.354626	8.82565254	72.22480062	28.0351174

Table 15: Top 50 stocks with the highest error, including real and predicted values, error, and standard error (SE) for **Experiment 2**.

Stock	Volatility	Real	Predicted	Error	SE
ALLK	0.520555	258.3713	260.751923	0.921396	5.667366
CLEU	0.45078	39.3848	37.6367	4.438515	3.055854
KOD	0.393908	26.67	29.439024	10.382542	7.667494
IMTE	0.334842	8.37	8.215385	1.847247	0.023906
CRBP	0.325413	23.8386	24.097578	1.086381	0.067070
IMUX	0.305878	6.228	6.369201	2.267192	0.019938
LTRY	0.275614	62.5342	62.515221	0.03035	0.000360
ATXS	0.267562	19.14	17.455164	8.802696	2.838672
AAMC	0.232731	65.9	74.045578	12.360513	66.35044
EIGR	0.22772	99.2	99.711525	0.51565	0.261658
... Remaining rows omitted for brevity ...					
Average	0.21253844	40.978368	40.53604336	13.38527662	13.720

Table 16: Top 50 stocks with the highest volatility, including real and predicted values, error, and standard error (SE) for **Experiment 3**.

Stock	Volatility	Real	Predicted	Error	SE
IBRX	0.179746	1.1182	4.914936	339.539938	14.41520425
MAIN	0.03034	5.27	20.306091	285.314826	226.0840326
AGNG	0.023521	3.3344	12.736378	281.969101	88.39719031
EDN	0.076645	1.275	4.866414	281.679535	12.89825452
IG	0.010411	4.763	16.627392	249.094936	140.7637975
CRT	0.07397	1.4	3.937448	181.246287	6.438642353
CTBB	0.052057	1.871	5.135989	174.505009	10.66015317
AGCO	0.050136	4.63	12.616401	172.492456	63.78260093
BILL	0.103582	2.4585	6.6264	169.530201	17.37139041
HIE	0.029804	2.9862	7.506611	151.376695	20.43411561
HOMZ	0.030288	2.85	7.073066	148.177763	17.83428644
IXC	0.040349	6.627	16.371964	147.049396	94.96432336
CGO	0.032283	2.44	5.578728	128.636382	9.851613458
CCOI	0.04203	1.42	3.022259	112.835124	2.567233903
ADUS	0.05479	2.34	4.870371	108.135527	6.402777398
FITBP	0.020787	6.25	12.906516	106.504257	44.30920526
FTI	0.074051	17.12	35.037384	104.657617	321.0326494
MBSD	0.006197	2.83	5.73643	102.700696	8.447335345
FWRD	0.059188	3.93	7.906392	101.18046	15.81169334
ARMP	0.128178	3.5775	7.184474	100.823858	13.01026144
DDLS	0.018384	1.79	3.58745	100.416203	3.230826503
LIVN	0.058857	4.67	0.354201	92.415397	18.62612101
... Remaining rows omitted for brevity ...					
Average	0.054045	4.86229	8.996673	108.402857	37.976335

Table 17: Top 50 stocks with the highest error, including real and predicted values, error, and standard error (SE) for **Experiment 3**.

	High Volatility	**High Error**
E1	48.587	126.160
E2	32.929	28.035
E3	13.720	37.976

Table 18: MSE for the top 50 stocks with high volatility and high prediction error.

B Additional Images

Figure 27: A line graph for stock A. The data shows a significant upward trend peaking in mid-2021, followed by fluctuating declines and recoveries.

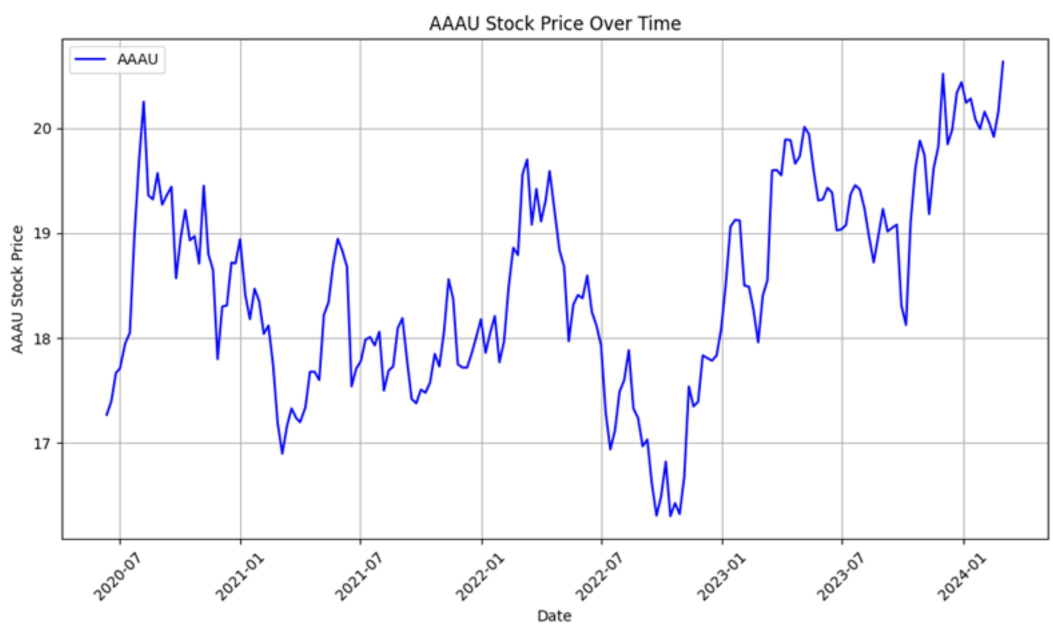


Figure 28: A line graph for stock AAAU. The overall trajectory suggests long-term growth despite fluctuations in the short term.

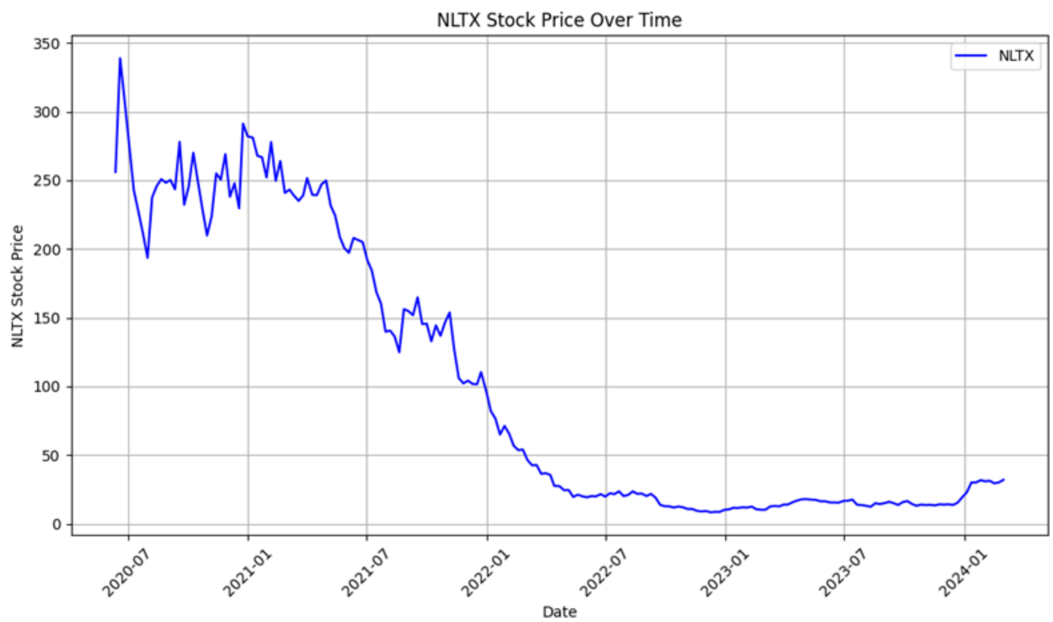


Figure 29: A line graph for stock NLTX. The data reveals a sharp initial decline from a peak of over 300, followed by intermittent fluctuations throughout 2021. From 2022 onward, the stock experiences a consistent downward trajectory.

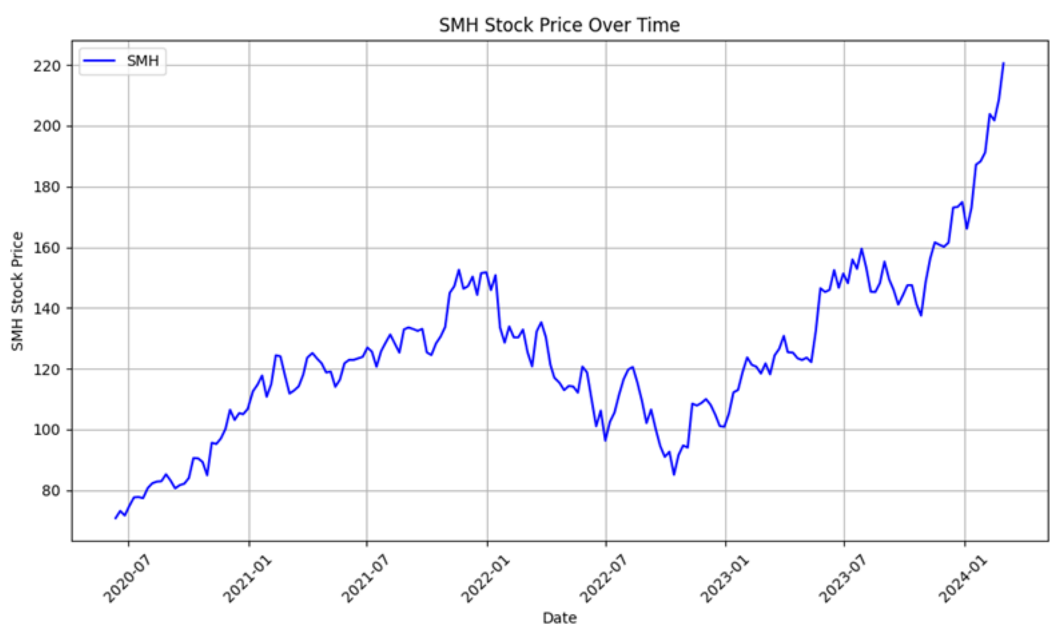


Figure 30: A line graph for stock SMH. The data reveals a steady upward trend, with periods of consolidation and slight declines, particularly in mid-2022. However, from mid-2023 onward, the price demonstrates a sharp and sustained increase, indicating strong growth momentum in recent months.

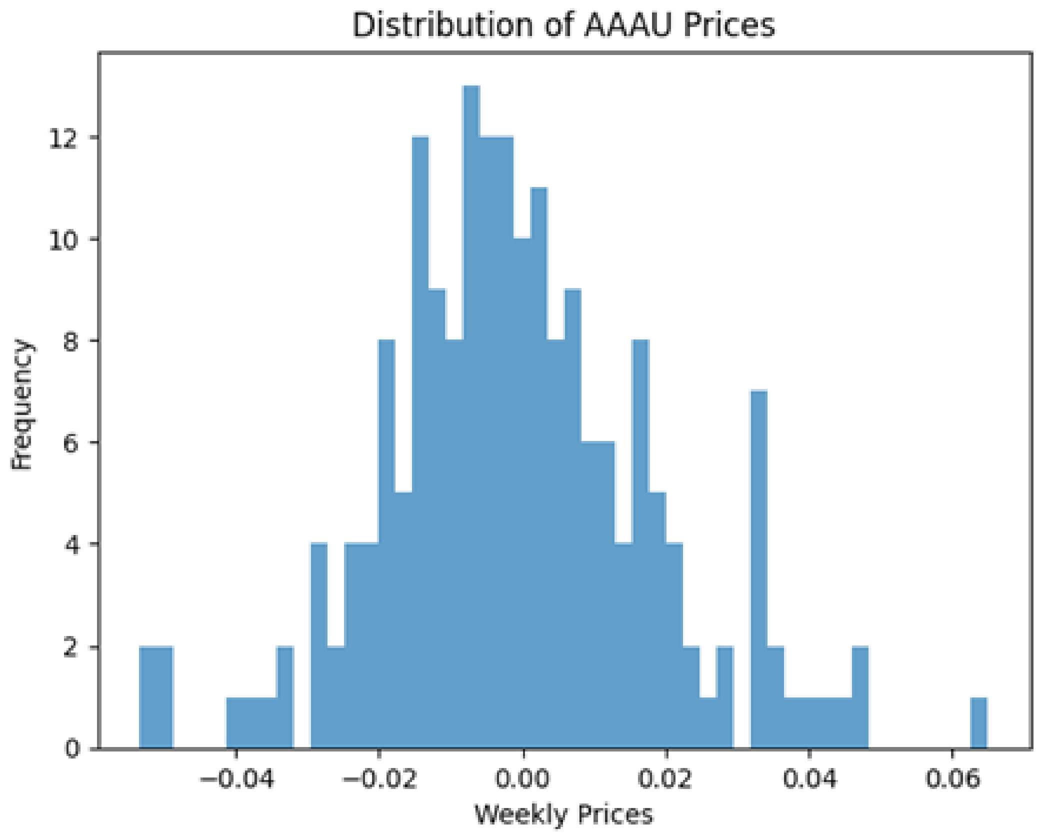


Figure 31: The graph shows the distribution of AAAU weekly prices, following a bell-shaped pattern centred around 0. The graph suggests a relatively normal distribution, with most price changes clustered near the mean and gradually tapering toward the extremes.

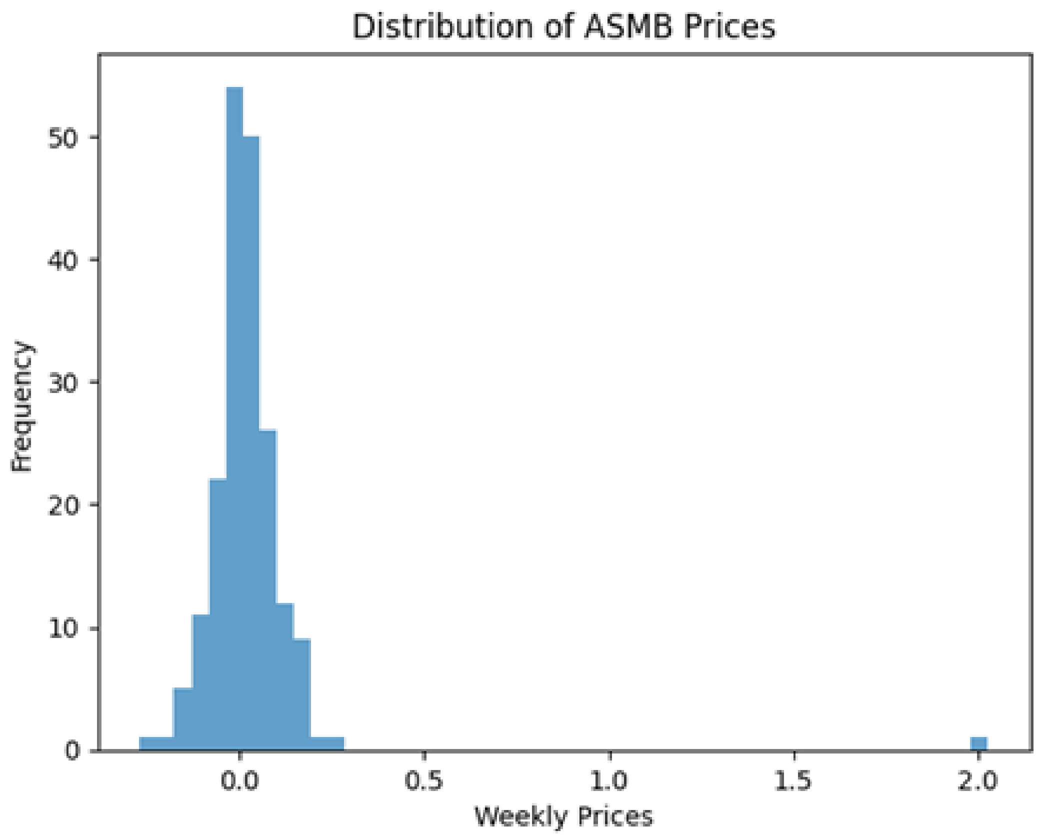


Figure 32: The graph shows the distribution of ASMB weekly prices, revealing a concentration of values near zero and a single notable outlier around 2, indicating most prices were low with rare instances of significantly higher values..

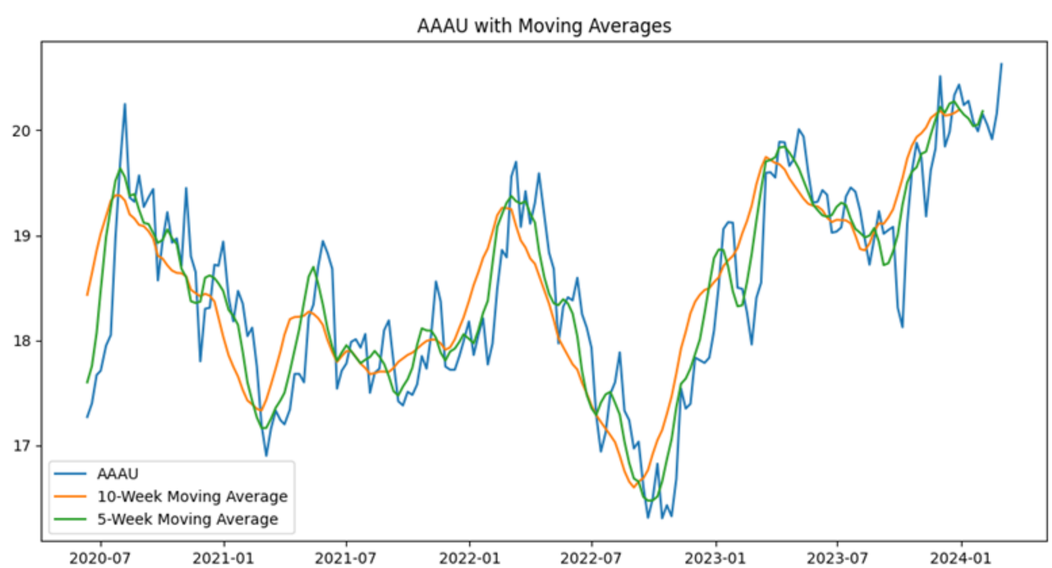


Figure 33: The graph illustrates the price movement of the AAAU stock from mid-2020 to early 2024 alongside its 5-week and 10-week moving averages. The moving averages smooth out short-term fluctuations, highlighting the underlying trends and aiding in identifying potential momentum shifts over time.

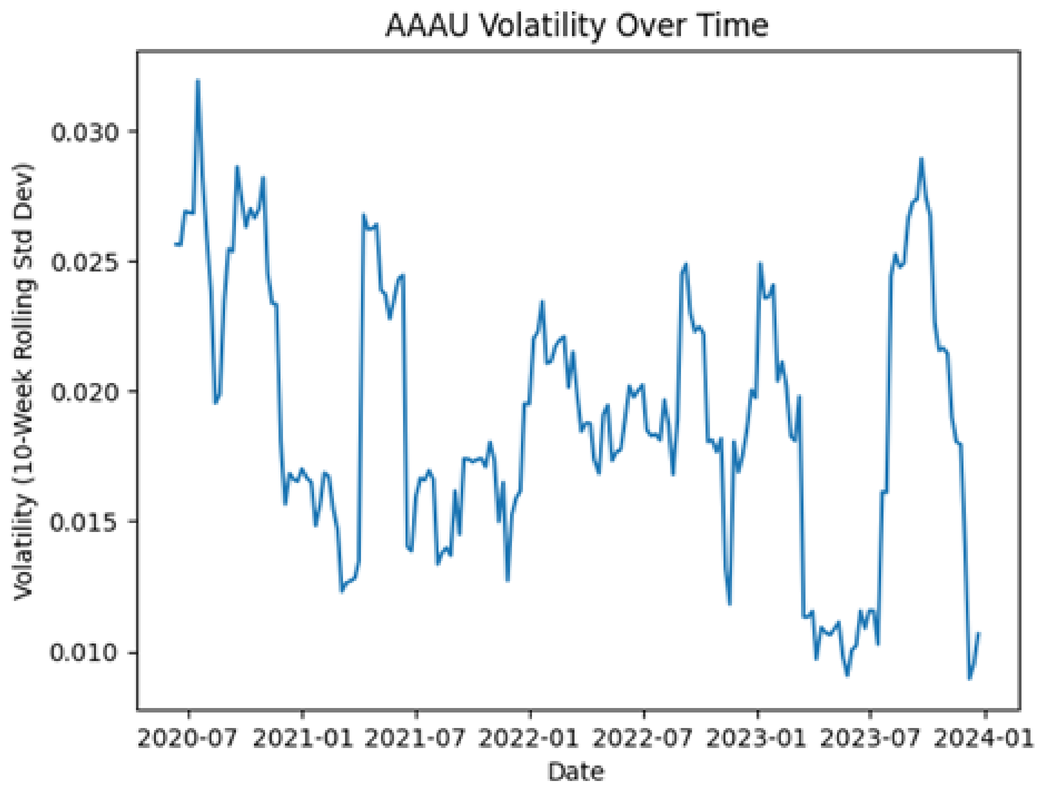


Figure 34: The graph depicts the volatility of AAAU over time, measured as a 10-week rolling standard deviation, showing periods of increased and decreased price fluctuations between mid-2020 and early 2024.

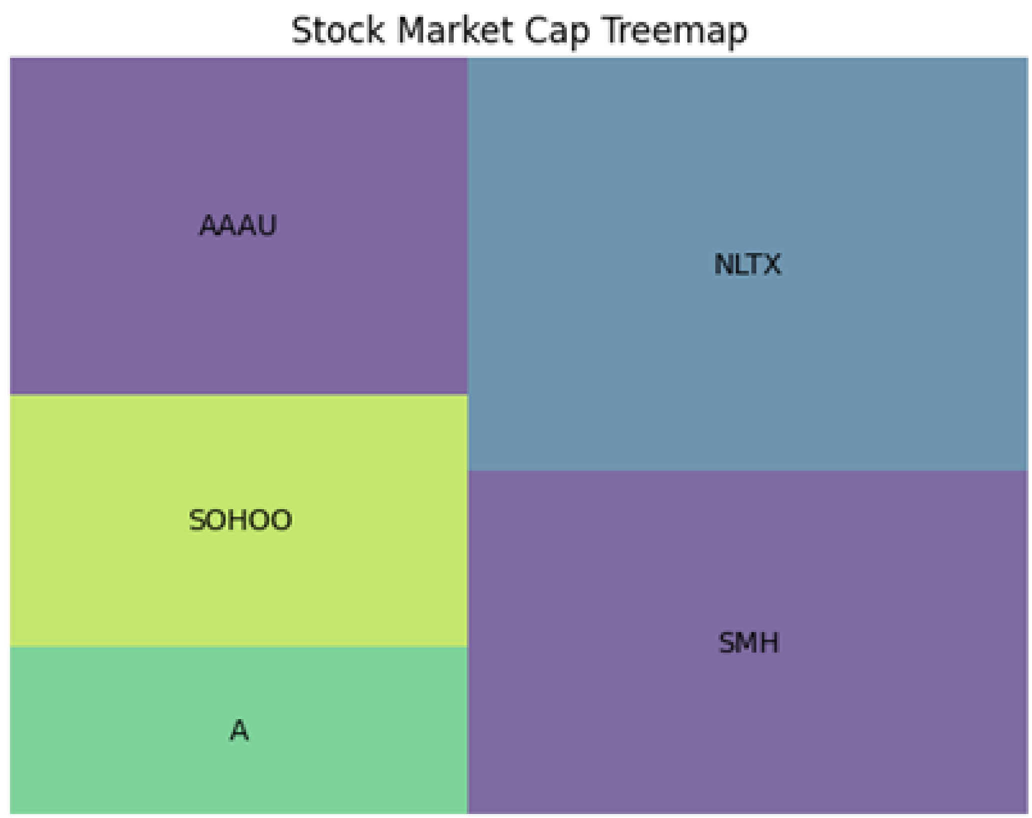


Figure 35: This chart shows the relative market capitalisation of several stock tickers (AAAU, SOHOO, A, NLTX, SMH). The size of each rectangle corresponds to the market cap of the respective ticker, with larger rectangles (e.g., NLTX and SMH) indicating higher market caps. In comparison, smaller rectangles (e.g., AAAU and A) represent lower market caps. .

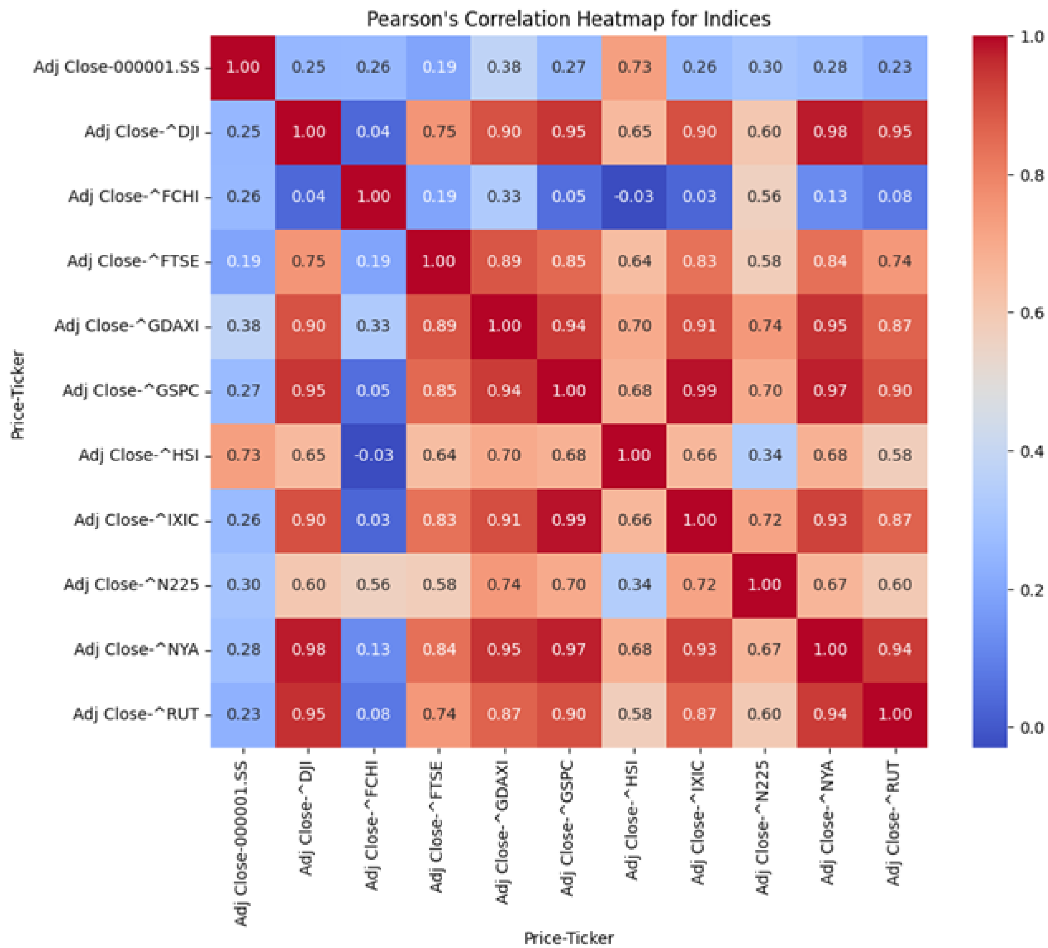


Figure 36: The NASDAQ and the S&P are heavily correlated. It also has a high correlation with the Dow Jones, DAX, Russell 2000, NYSE and the FTSE 100. The NASDAQ is mildly correlated with the Nikkei and the Hang Seng Index. The NYSE is heavily correlated with the DAX, Dow Jones, Russell 2000, S&P 500 and NASDAQ. It is strongly correlated with the FTSE 100.

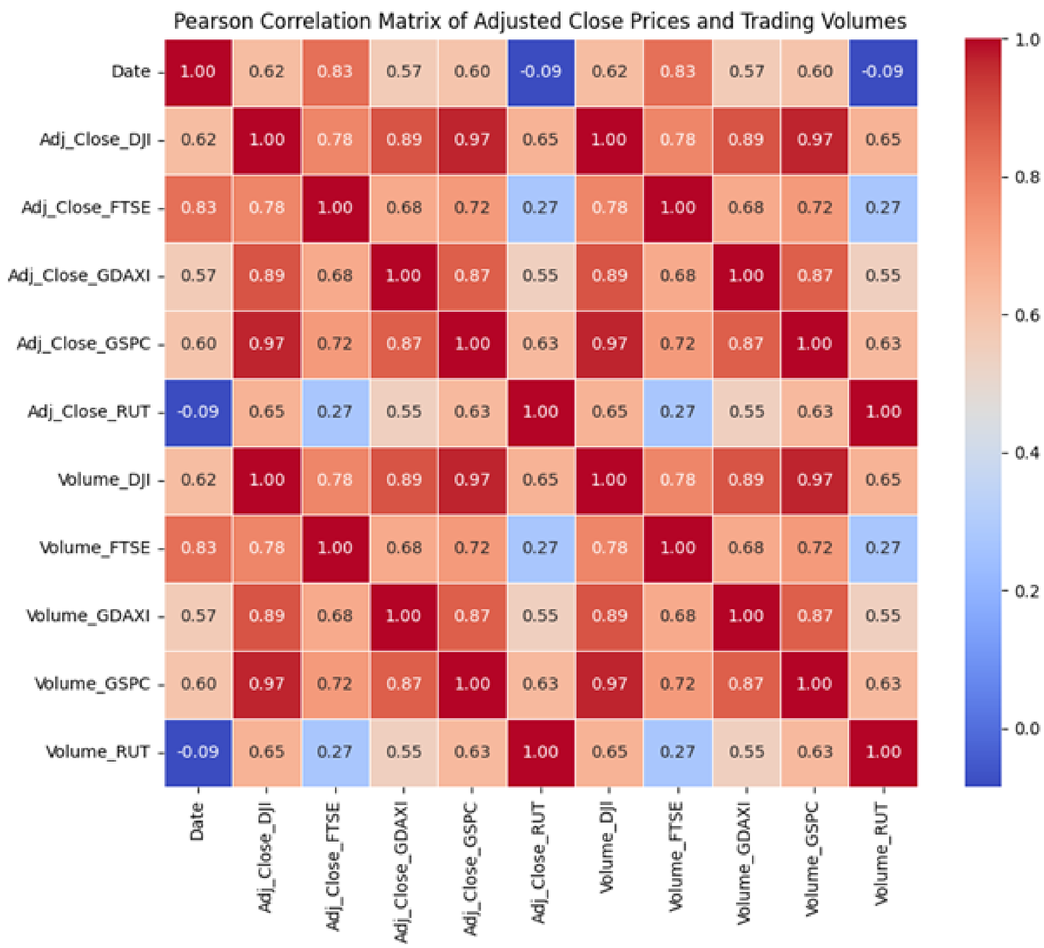


Figure 37: Unsurprisingly, the volume is heavily correlated with the indices it is associated with.

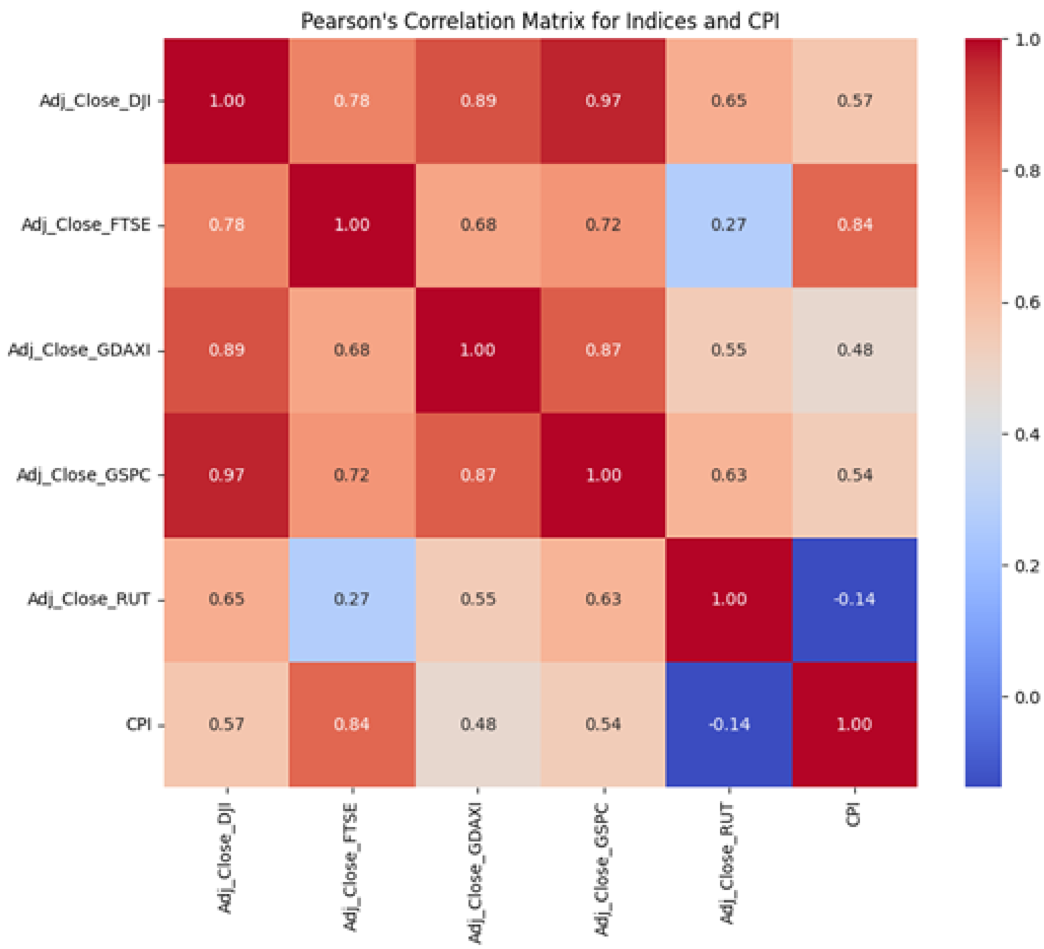


Figure 38: The FTSE is correlated with the US inflation rate (CPI) and so it is highly likely that NASDAQ and NYSE also correlate with inflation.

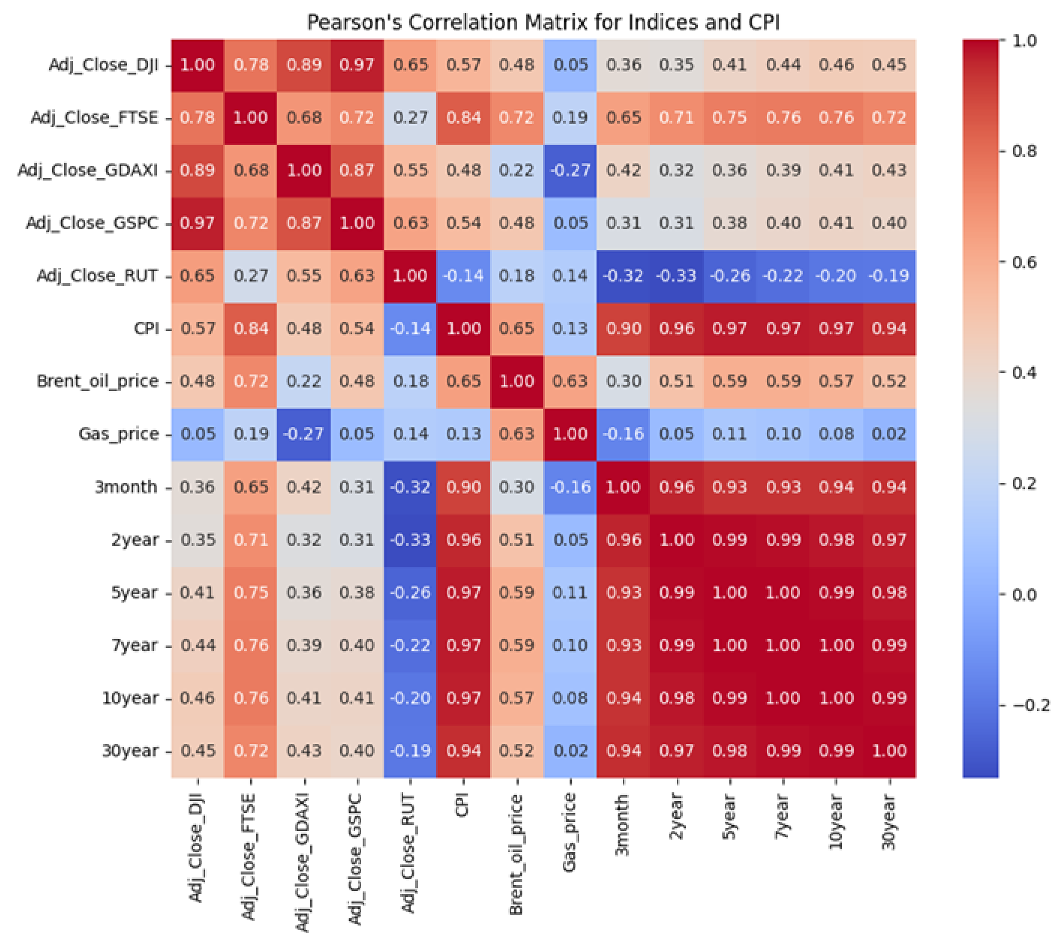


Figure 39: The Treasury yield is heavily correlated with Inflation and mildly correlated with the FTSE. This suggests that the inflation data is enough to help with the prediction. The treasury yield will be excluded. Oil and gas prices are also insignificant.

Figure 40: Comparison of indices' real returns and annualised realised volatility. While real returns across indices display significant short-term fluctuations, the annualised realised volatility varies over time, indicating periods of heightened or reduced market stability across different indices.

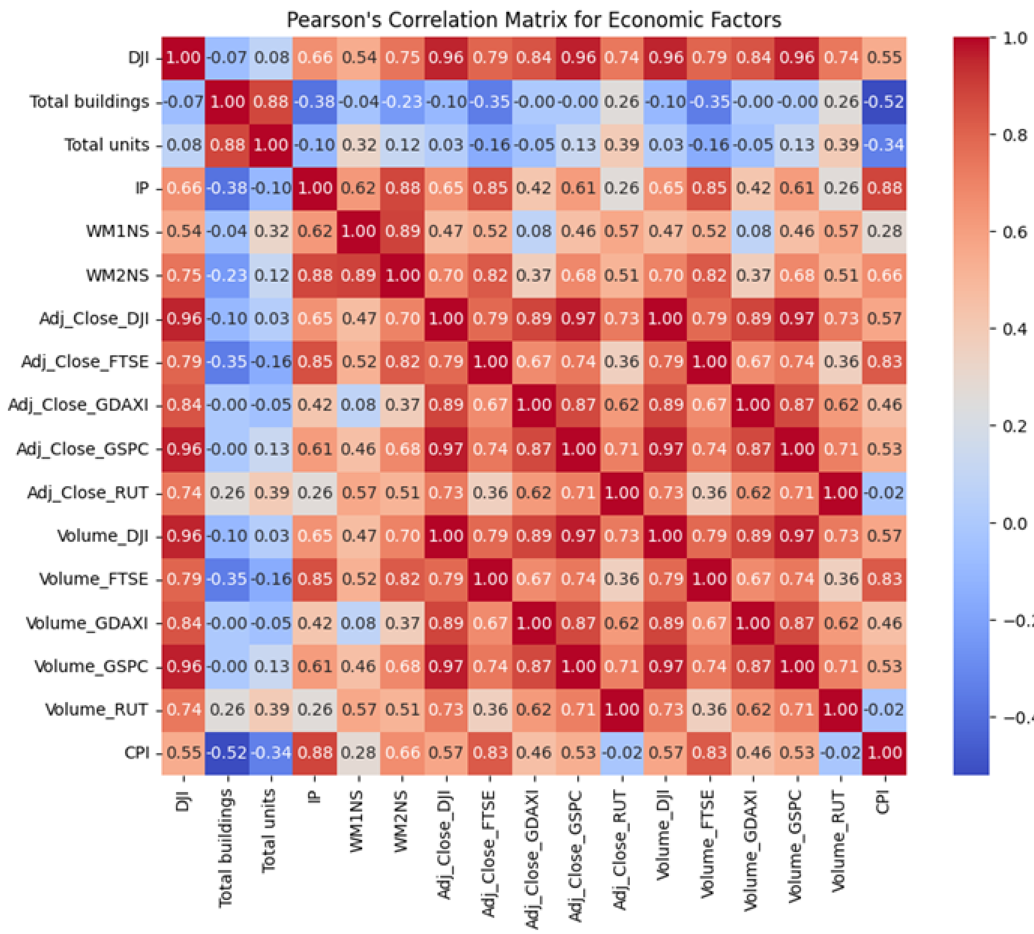


Figure 41: It appears that industrial production (IP) is heavily correlated with the FTSE and the money supply (WM2NS).

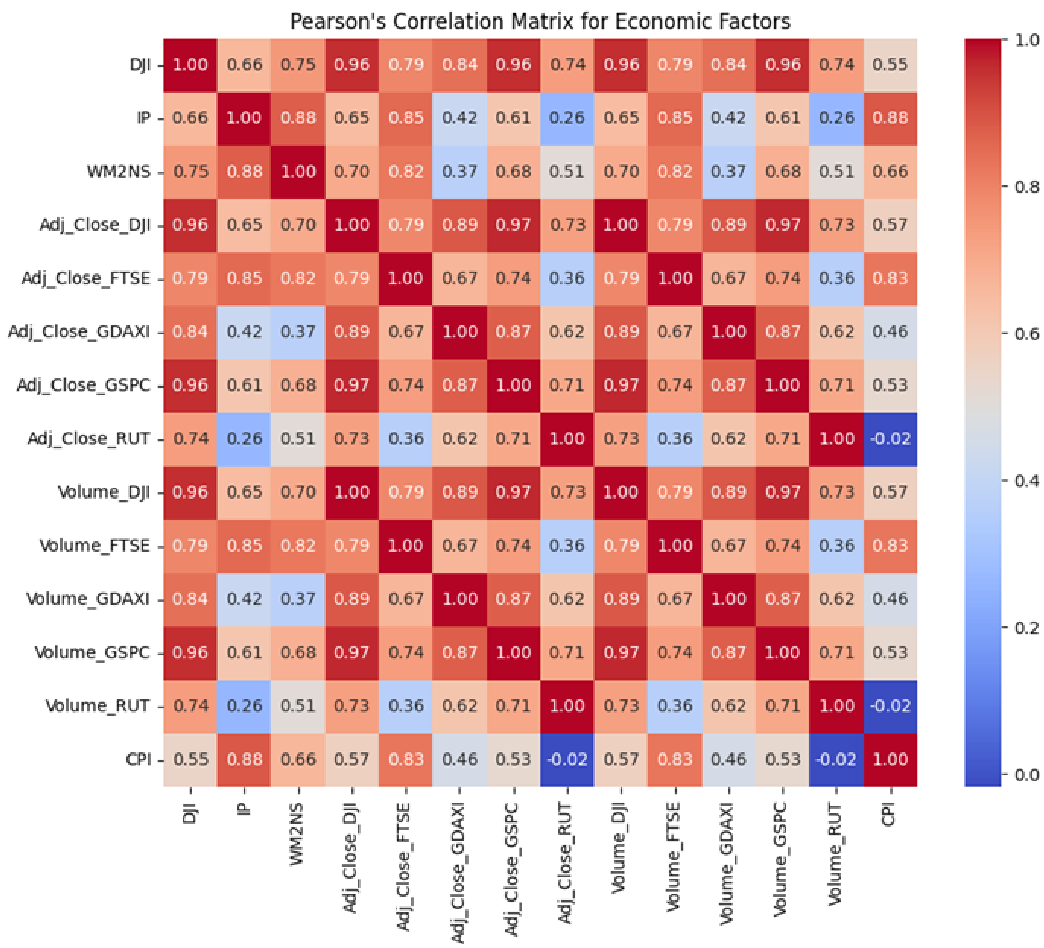


Figure 42: A reduced version of the Pearson's Correlation heat map for readability.

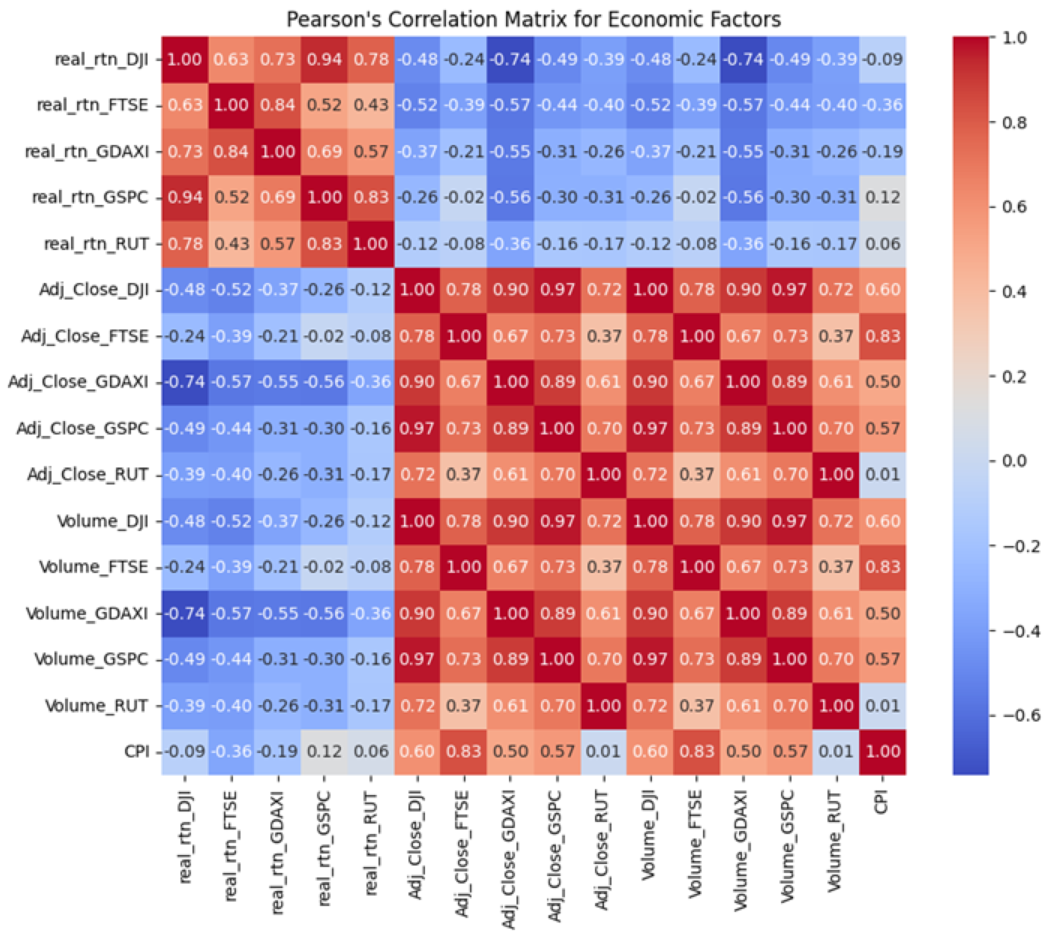
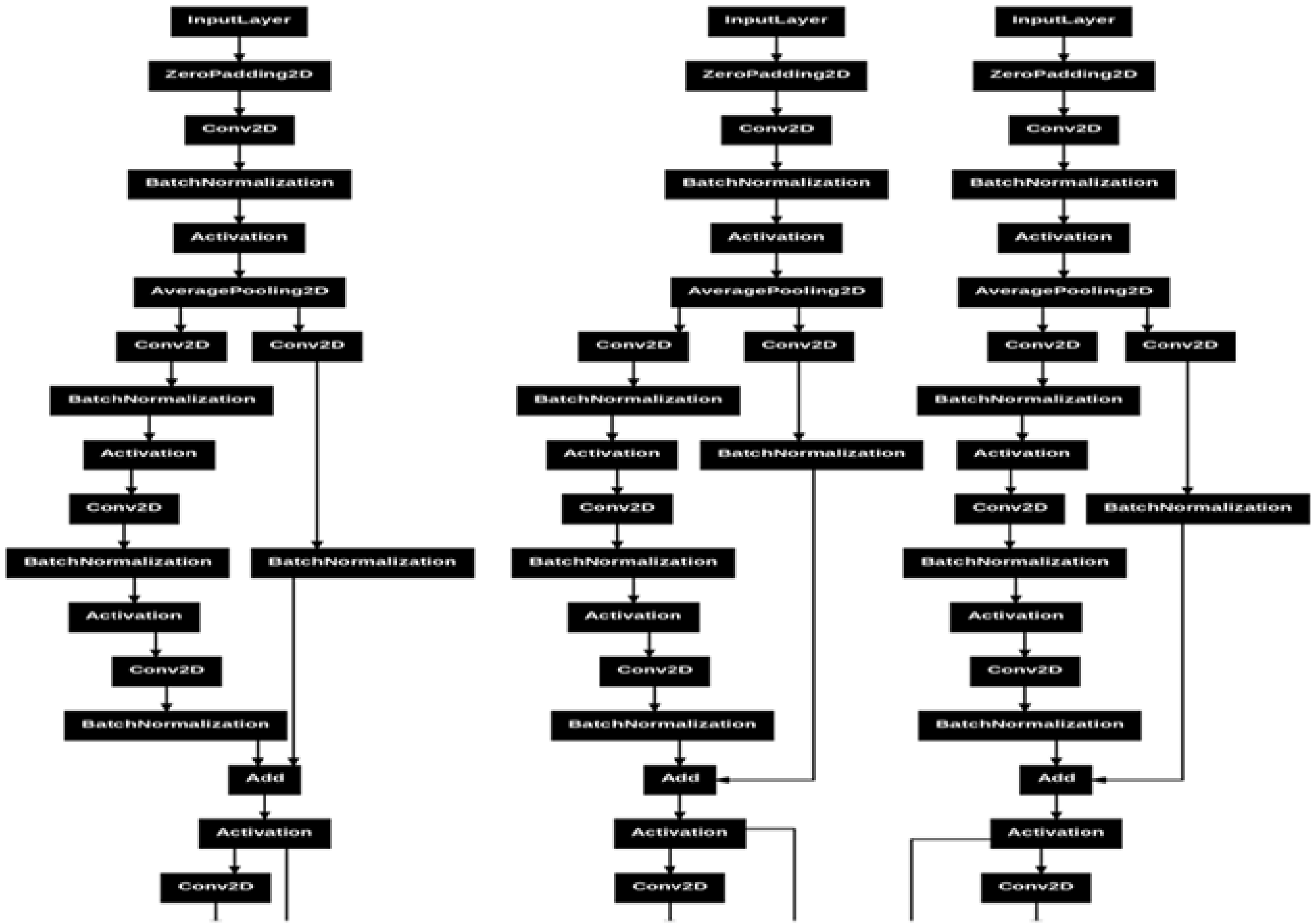
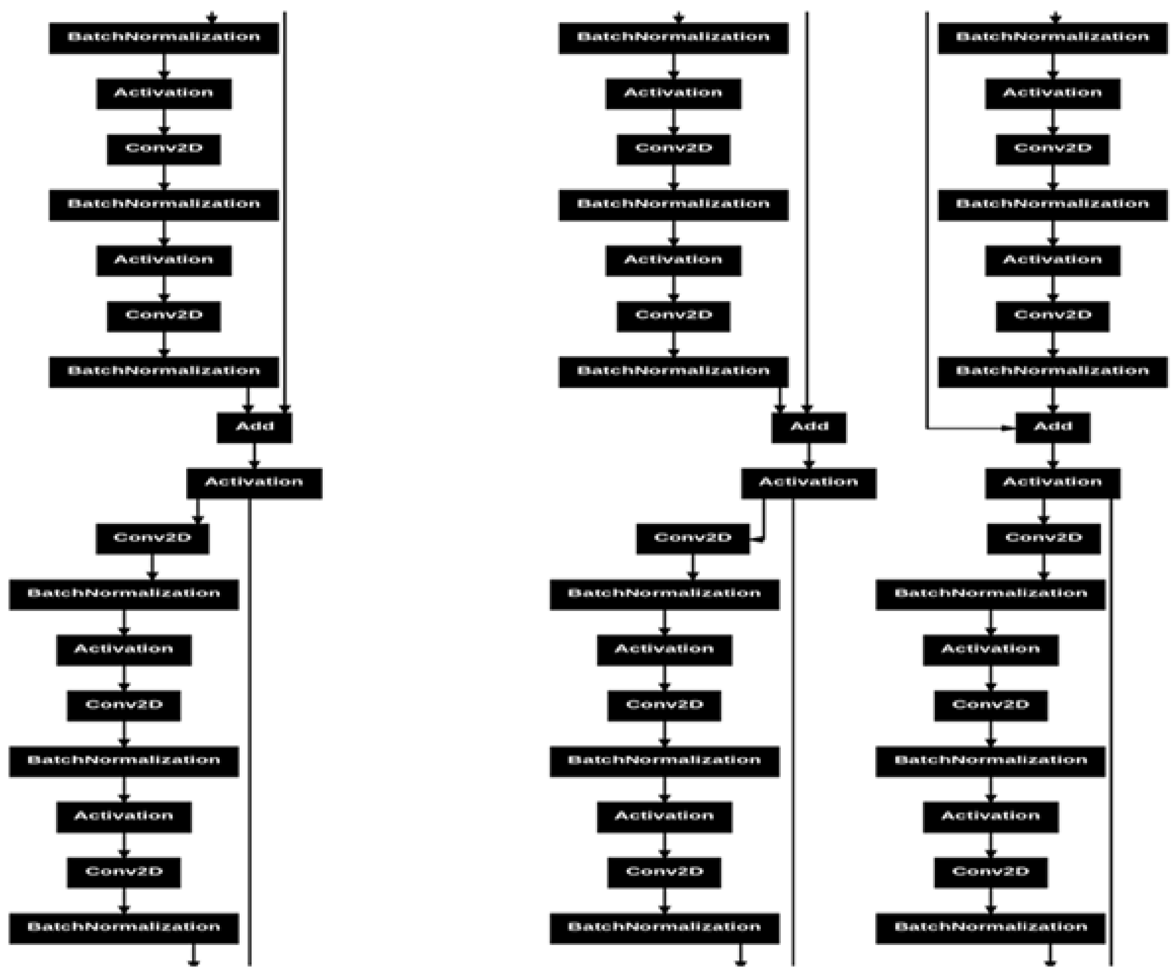
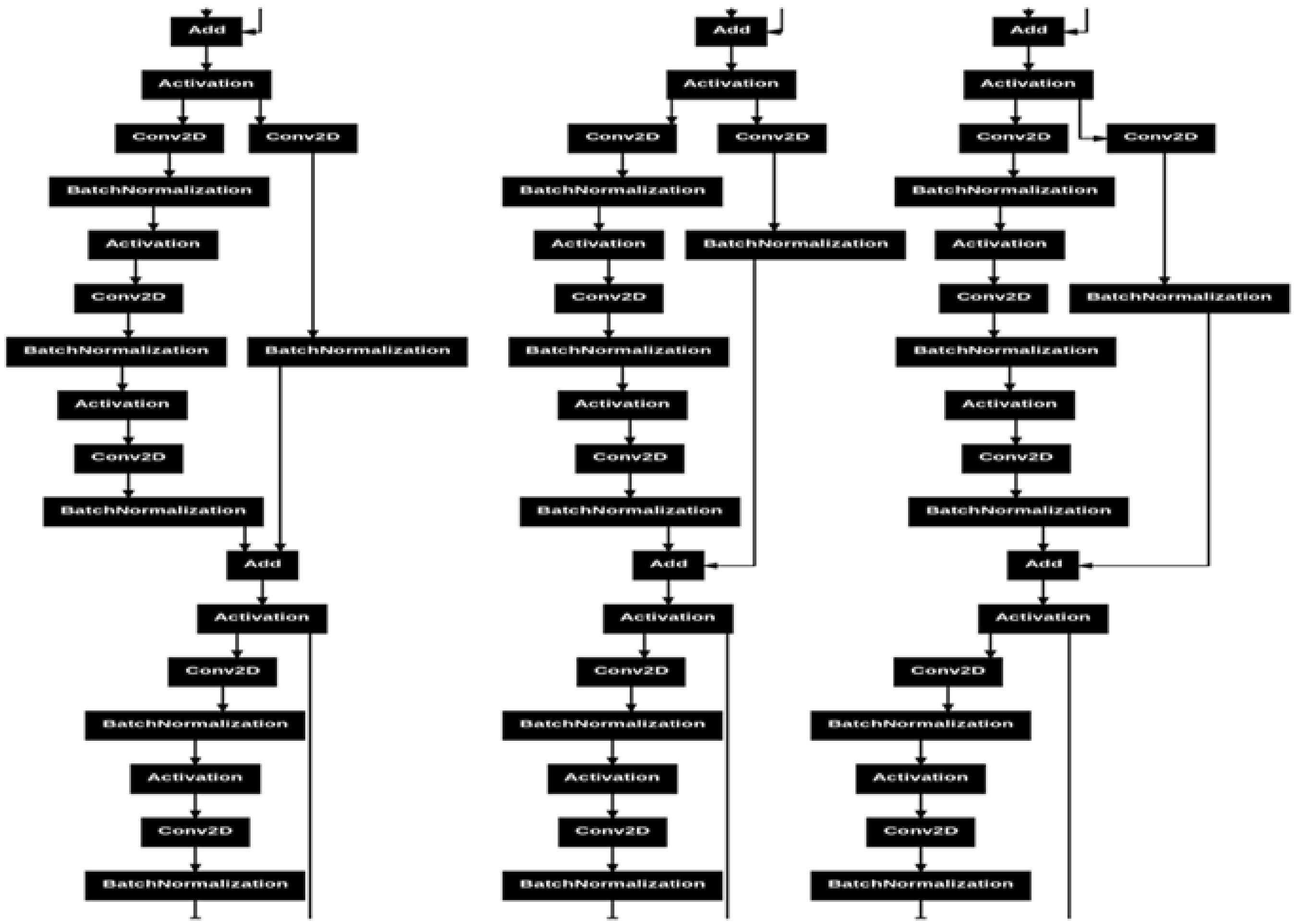
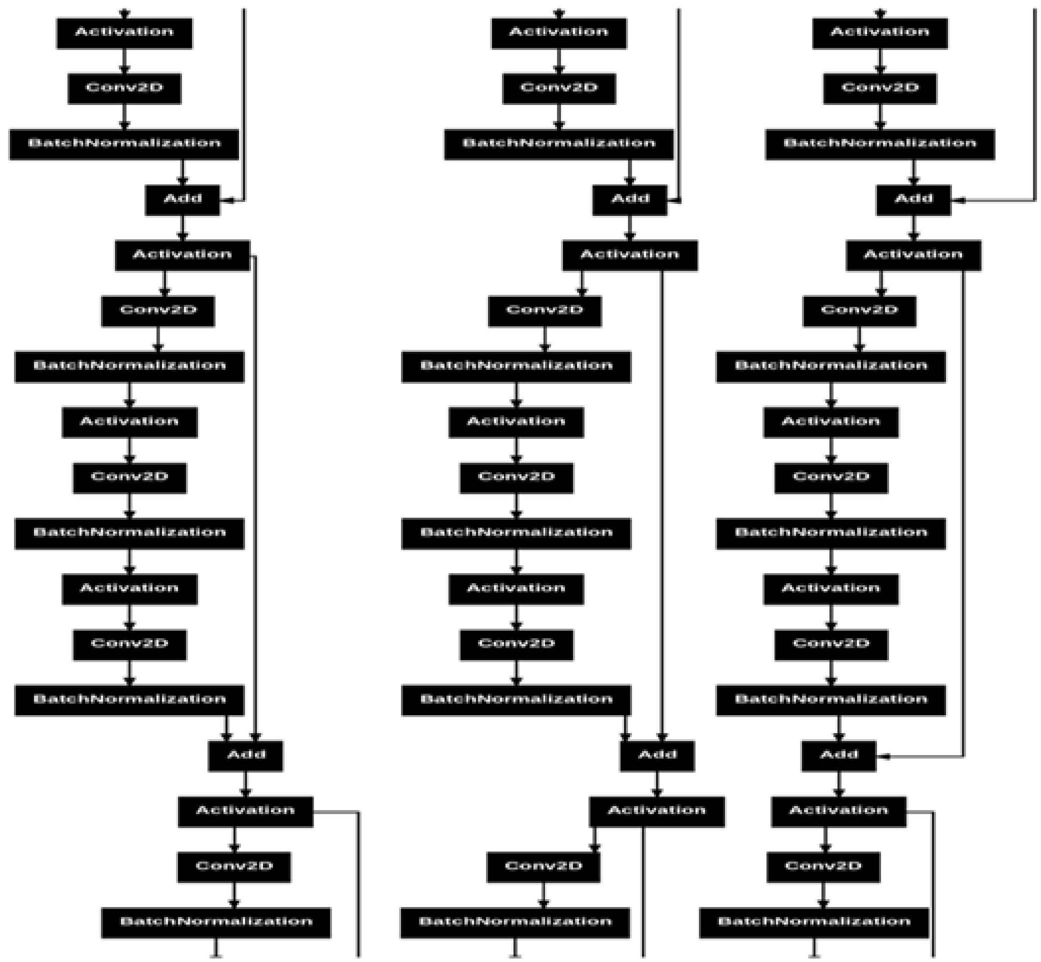


Figure 43: The real returns were added to the heat map and although the real returns for the Dow Jones (DJI) have a negative correlation of -0.74 with the German Dax (GDAXI), it appeared rather insignificant to include in the final features.


InputLayer
ZeroPadding2D
Conv2D
BatchNormalization
Activation
AveragePooling2D
Conv2D
Conv2D
BatchNormalization
Activation
Conv2D
BatchNormalization
BatchNormalization
Activation
Conv2D
BatchNormalization
Add
Activation
Conv2D
InputLayer
ZeroPadding2D
Conv2D
BatchNormalization
Activation
AveragePooling2D
Conv2D
Conv2D
BatchNormalization
Activation
BatchNormalization
Conv2D
BatchNormalization
Activation
Conv2D
BatchNormalization
Add
Activation
Conv2D
InputLayer
ZeroPadding2D
Conv2D
BatchNormalization
Activation
AveragePooling2D
Conv2D
Conv2D
BatchNormalization
Activation
Conv2D
BatchNormalization
BatchNormalization
Activation
Conv2D
BatchNormalization
Add
Activation
Conv2D


BatchNormalization
Activation
Conv2D
BatchNormalization
Activation
Conv2D
BatchNormalization
Add
Activation
Conv2D
BatchNormalization
Activation
Conv2D
BatchNormalization
Activation
Conv2D
BatchNormalization
BatchNormalization
Activation
Conv2D
BatchNormalization
Activation
Conv2D
BatchNormalization
Add
Activation
Conv2D
BatchNormalization
Activation
Conv2D
BatchNormalization
Activation
Conv2D
BatchNormalization
BatchNormalization
Activation
Conv2D
BatchNormalization
Activation
Conv2D
BatchNormalization
Add
Activation
Conv2D
BatchNormalization
Activation
Conv2D
BatchNormalization
Activation
Conv2D
BatchNormalization


Add
Activation
Conv2D
Conv2D
BatchNormalization
Activation
Conv2D
BatchNormalization
BatchNormalization
Activation
Conv2D
BatchNormalization
Add
Activation
Conv2D
BatchNormalization
Activation
Conv2D
BatchNormalization
Add
Activation
Conv2D
Conv2D
BatchNormalization
Activation
BatchNormalization
Conv2D
BatchNormalization
Activation
Conv2D
BatchNormalization
Add
Activation
Conv2D
BatchNormalization
Activation
Conv2D
BatchNormalization
Add
Activation
Conv2D
Conv2D
BatchNormalization
Activation
Conv2D
BatchNormalization
BatchNormalization
Activation
Conv2D
BatchNormalization
Add
Activation
Conv2D
BatchNormalization
Activation
Conv2D
BatchNormalization


Activation
Conv2D
BatchNormalization
Add
Activation
Conv2D
BatchNormalization
Activation
Conv2D
BatchNormalization
Activation
Conv2D
BatchNormalization
Add
Activation
Conv2D
BatchNormalization
Activation
Conv2D
BatchNormalization
Add
Activation
Conv2D
BatchNormalization
Activation
Conv2D
BatchNormalization
Activation
Conv2D
BatchNormalization
Add
Activation
Conv2D
BatchNormalization
Activation
Conv2D
BatchNormalization
Add
Activation
Conv2D
BatchNormalization
Activation
Conv2D
BatchNormalization
Activation
Conv2D
BatchNormalization
Add
Activation
Conv2D
BatchNormalization

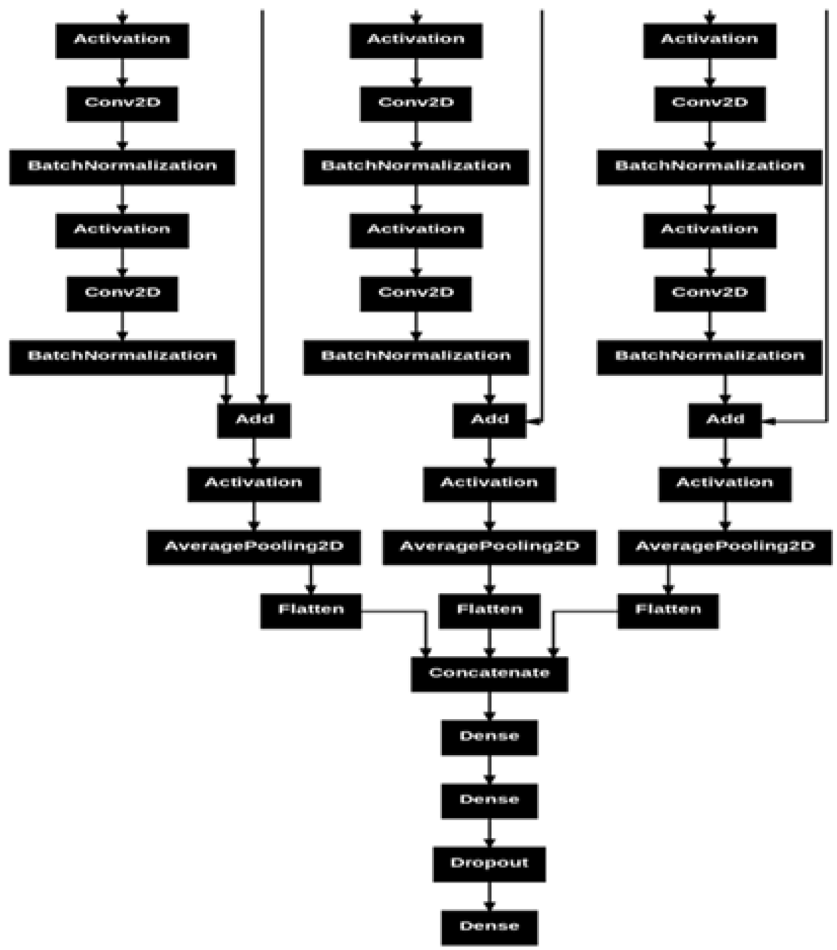


Figure 44: The model architecture of the residual network.

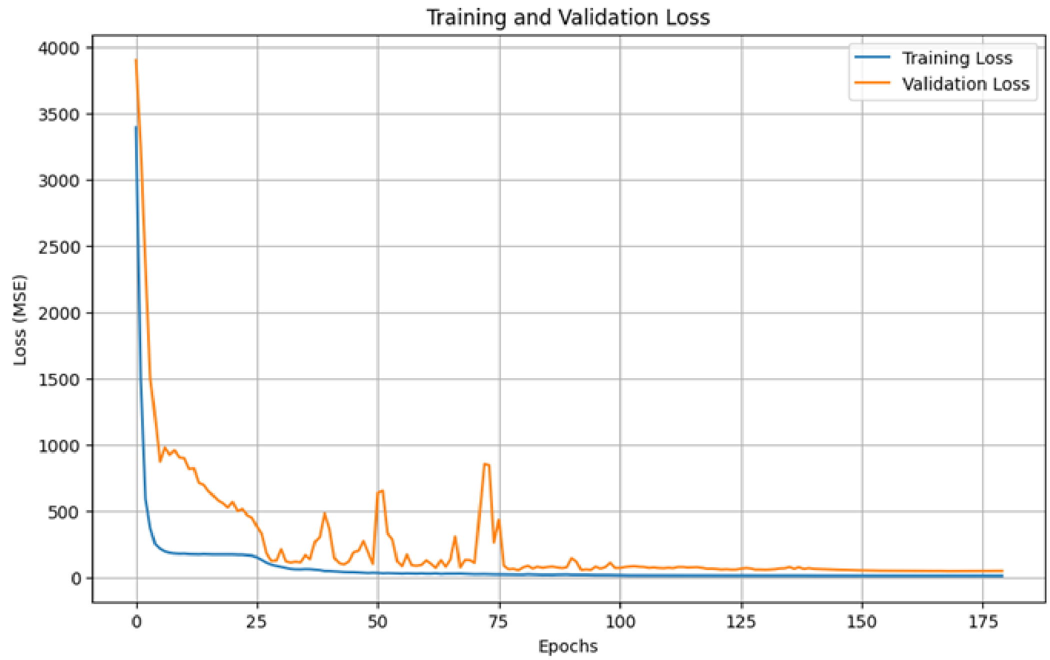


Figure 45: (Experiment 1) The chart shows the training and validation loss for the model across 175 epochs, measured in terms of MSE. Initially, both losses exhibit a sharp decline, indicating that the model quickly learns the data patterns during the early training stages. The training loss continues to decrease steadily, eventually approaching zero, which reflects the model's ability to fit the training data effectively. In contrast, the validation loss shows fluctuations, particularly after 25 epochs, suggesting the presence of instability in the generalisation of unseen data. These fluctuations gradually reduce over time, with the validation loss eventually stabilising, though at a higher level than the training loss. This disparity points to a mild overfitting of the model, where performance on the training data is superior to that on the validation set. Despite this, the overall downward trend in both curves suggests effective learning and adaptation throughout the training process. There is a plateau at approximately 100 epochs. This plateau indicates that the model has reached its learning capacity, with minimal improvement in loss despite continued training.

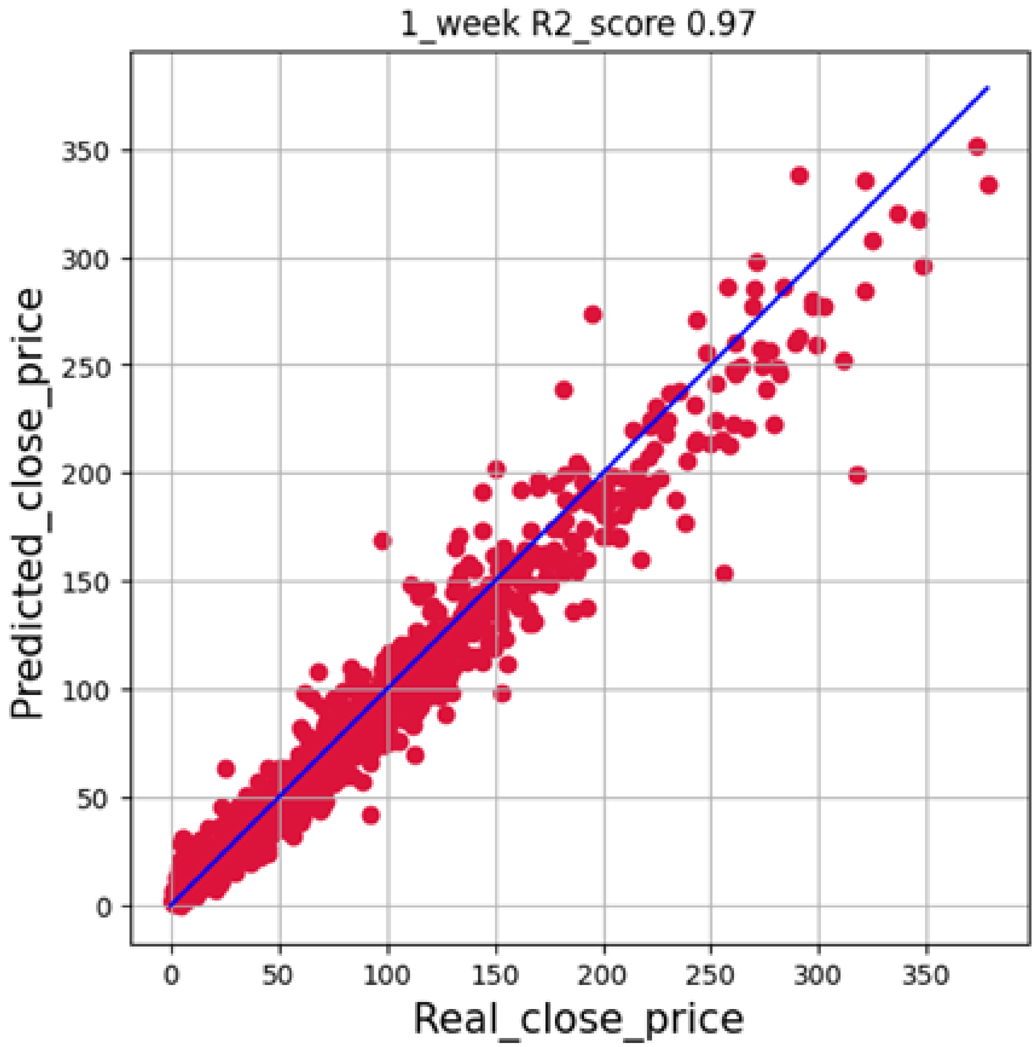


Figure 46: (Experiment 1) The scatter plot shows the relationship between the actual closing prices (x-axis) and the predicted closing prices (y-axis) for the model, with an R^2 score of 0.97. The points are tightly clustered around the diagonal line, representing perfect prediction accuracy and indicating a strong correlation between the predicted and actual values. While most data points align closely with the diagonal, a few deviations are observed, particularly at higher price ranges, suggesting minor inaccuracies in some predictions. The high R^2 score indicates the model's ability to effectively capture the underlying patterns in the data and make accurate predictions for the majority of cases.

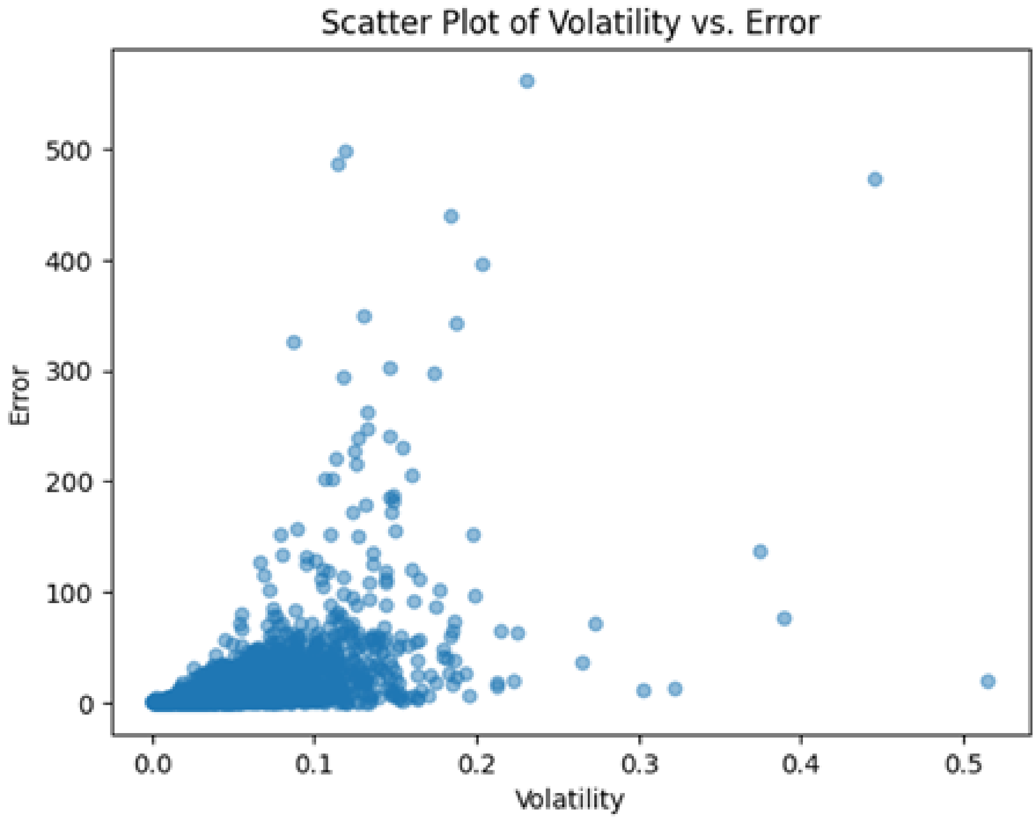


Figure 47: (Experiment 1) The scatter plot presents the relationship between volatility (x-axis) and prediction error (y-axis) for the model. Most data points are concentrated in the lower volatility range (below 0.1), with corresponding errors typically remaining low. As volatility increases, the spread of errors becomes more pronounced, with a small number of extreme cases exceeding 500 in error magnitude. This pattern suggests that the model performs well in stable, low-volatility conditions but needs help maintaining accuracy as market volatility increases. These observations highlight the model's sensitivity to volatility and potential limitations in handling more volatile scenarios.

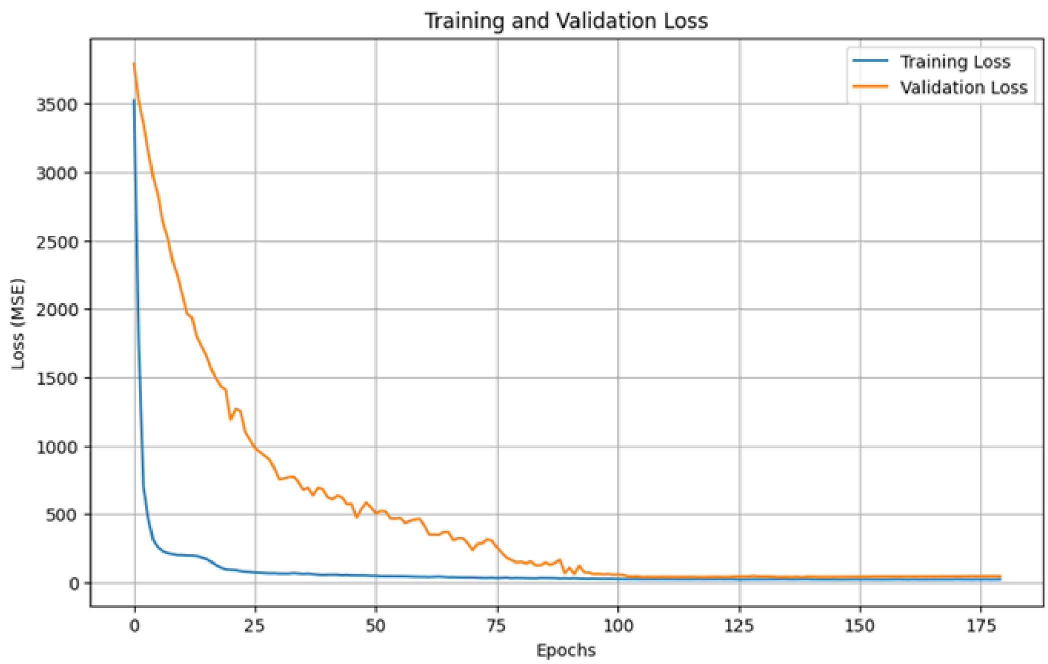


Figure 48: (Experiment 2)The training and validation loss curves decline consistently as the epochs progress, with both losses plateauing towards the later stages of training. The training loss approaches near zero, while the validation loss converges to a slightly higher value. This pattern suggests that the model generalises effectively to unseen data, although the persistent gap between the two curves may indicate minor overfitting.

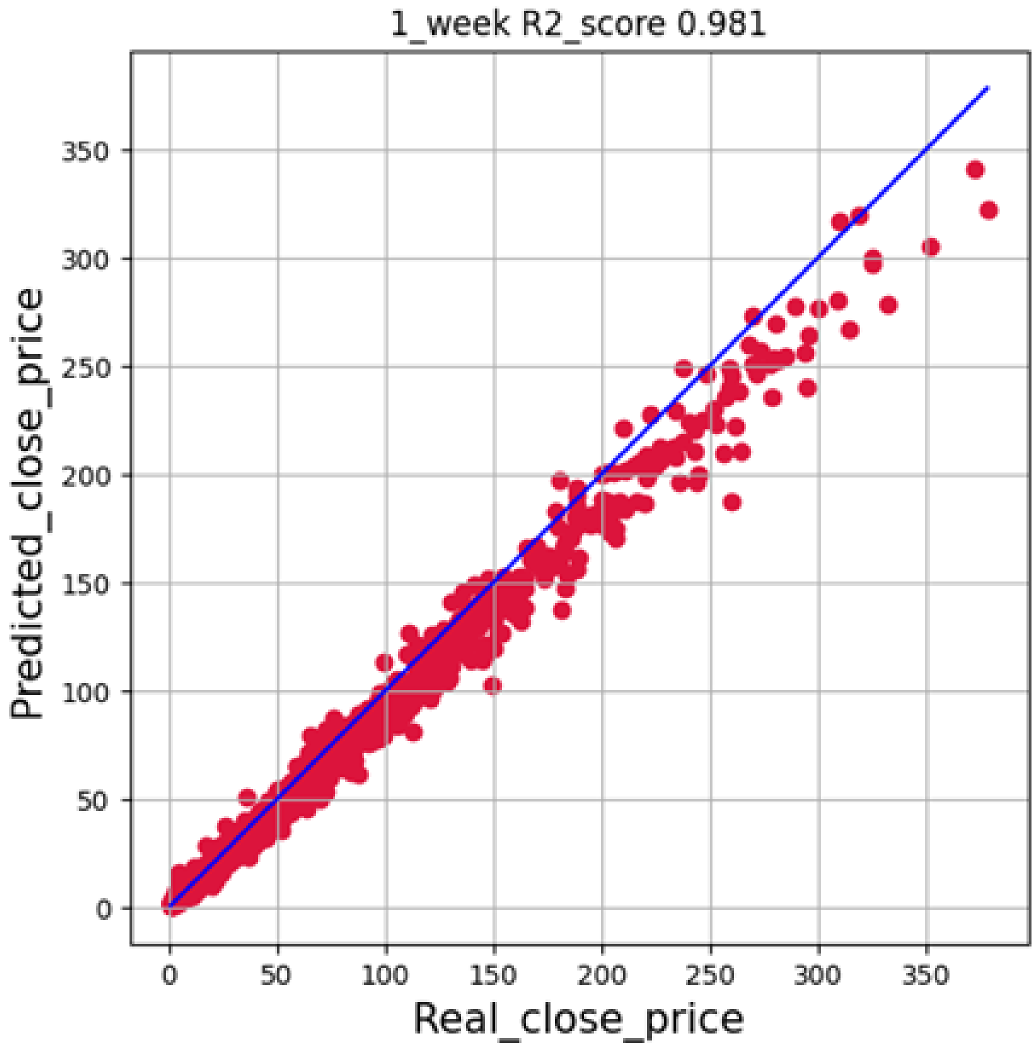


Figure 49: (Experiment 2)The scatter plot shows the relationship between actual and predicted stock prices for Experiment 2, with an R^2 score of 0.981, indicating a really good predictive accuracy. The points closely cluster around the diagonal line, representing a near-perfect prediction. While most data points adhere tightly to this line, a few deviations occur, particularly at higher price ranges, reflecting slightly reduced precision.

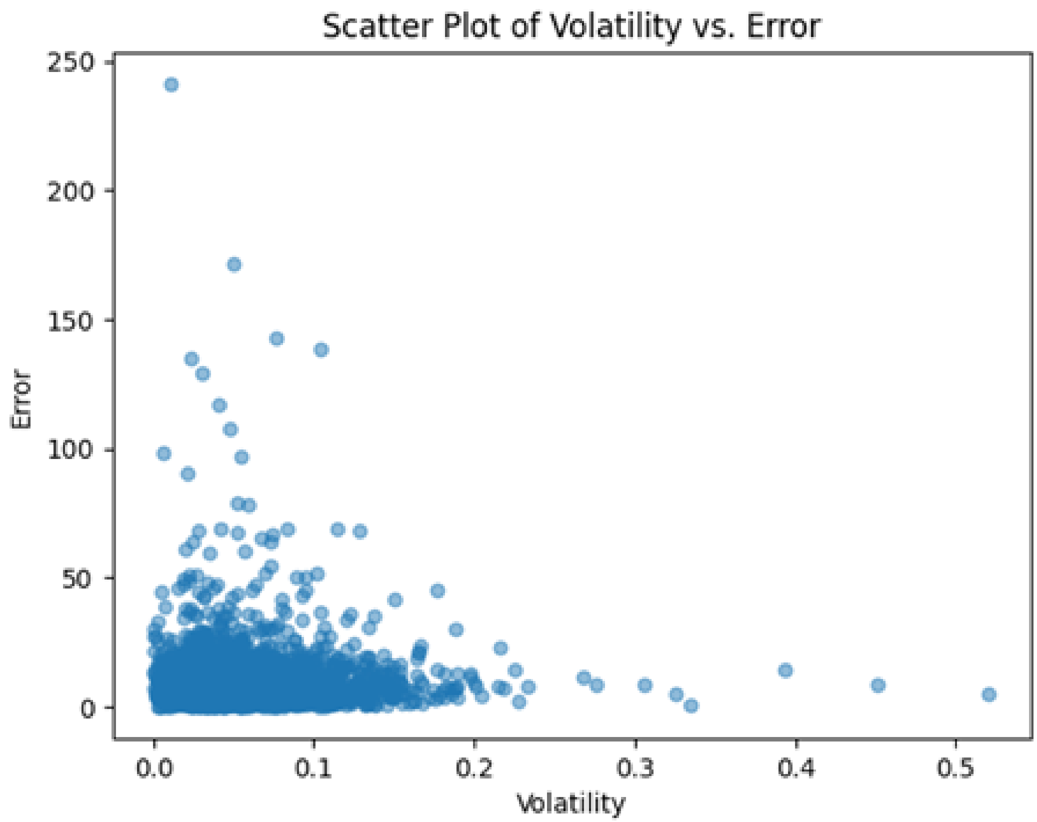


Figure 50: (Experiment 2)The scatter plot illustrates the relationship between market volatility and prediction error. Most points are concentrated at lower volatility levels, with minor errors, suggesting the model performs well under stable conditions. However, errors are more pronounced as volatility increases, with a few significant outliers, indicating that while the model handles low-volatility scenarios well, its accuracy diminishes in more volatile market conditions.

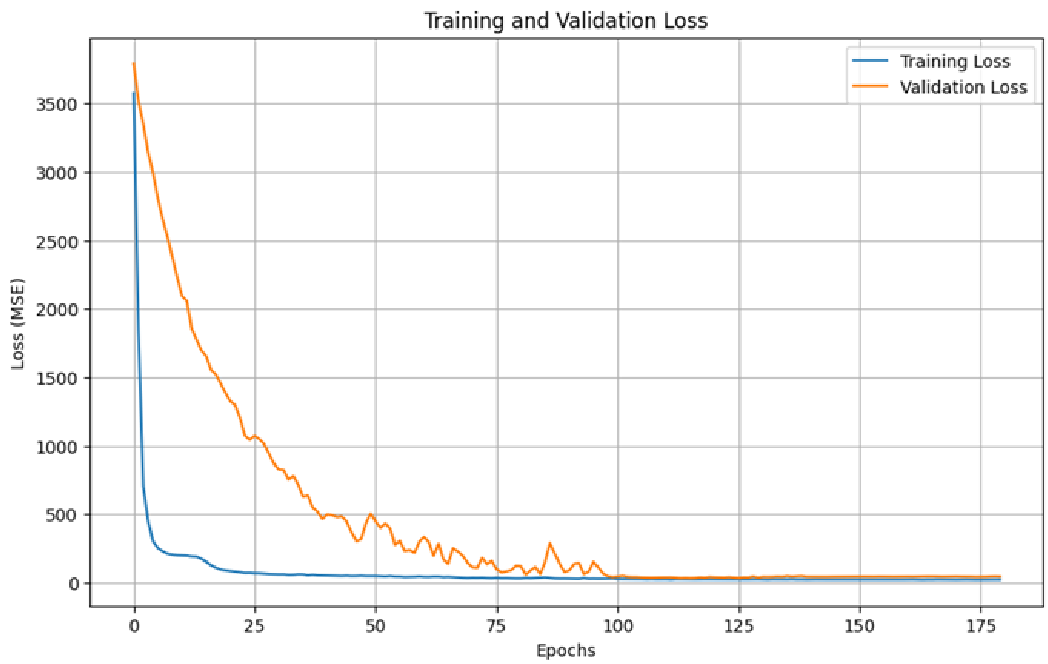


Figure 51: (Experiment 3)The training and validation loss curves shows a steady decline in loss throughout 175 epochs. The training loss decreases rapidly, approaching near zero, while the validation loss stabilises at a higher value, reflecting the model's generalisation to unseen data. Although there is a small gap between the two curves, suggesting minor overfitting, the overall trend shows that the model has trained well.

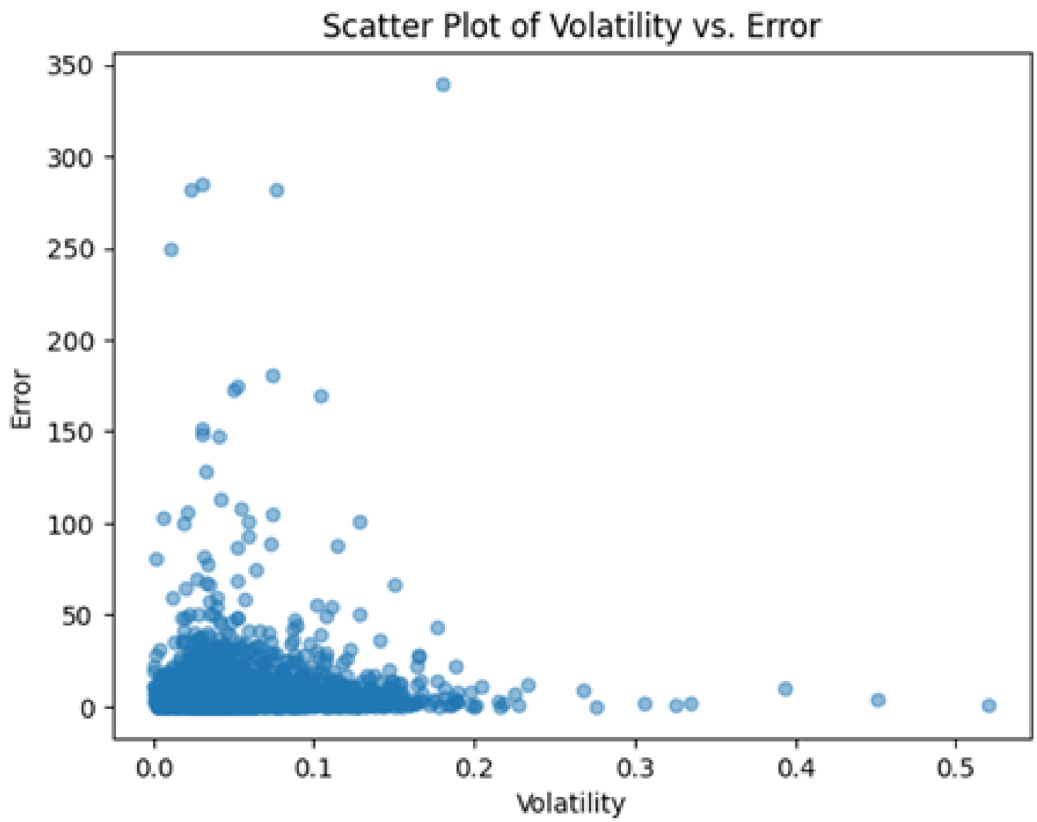


Figure 52: (Experiment 3)This scatter plot explores the relationship between market volatility and prediction error. Most data points are concentrated in the lower volatility, with minimal prediction errors. A wider spread of errors emerges as volatility increases, with a few scattered outliers. These results highlight that while Experiment 3 performs exceptionally well in low-volatility conditions, it encounters challenges maintaining accuracy when volatility is high.

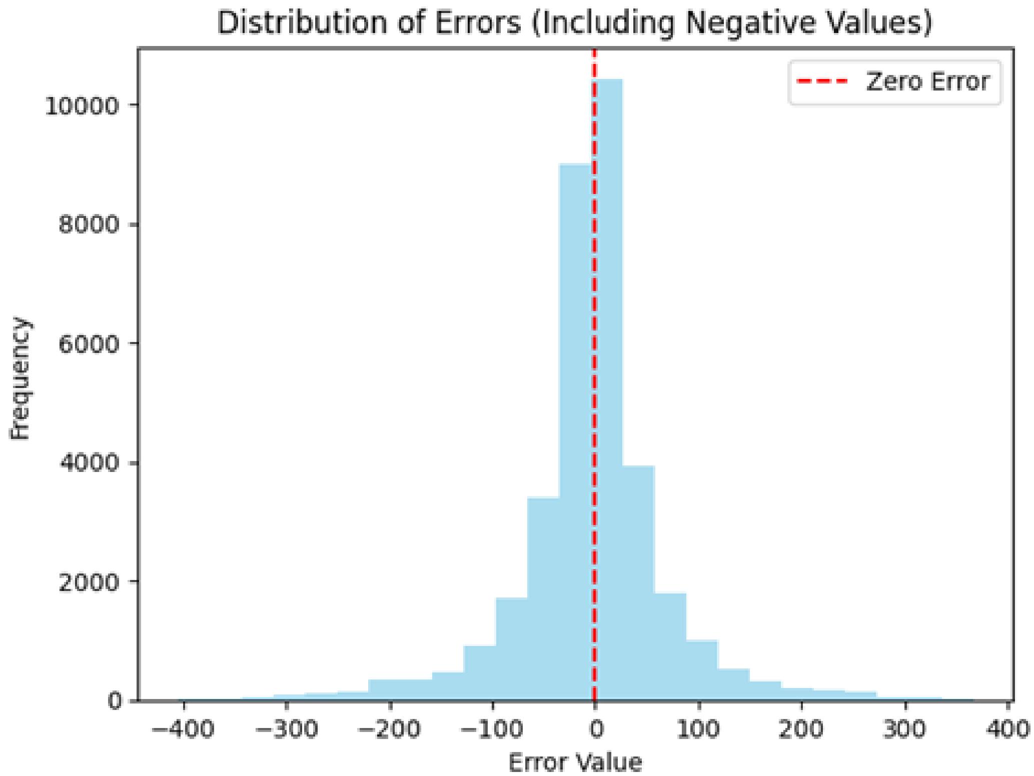


Figure 53: A histogram of E2 predicting the horizon.

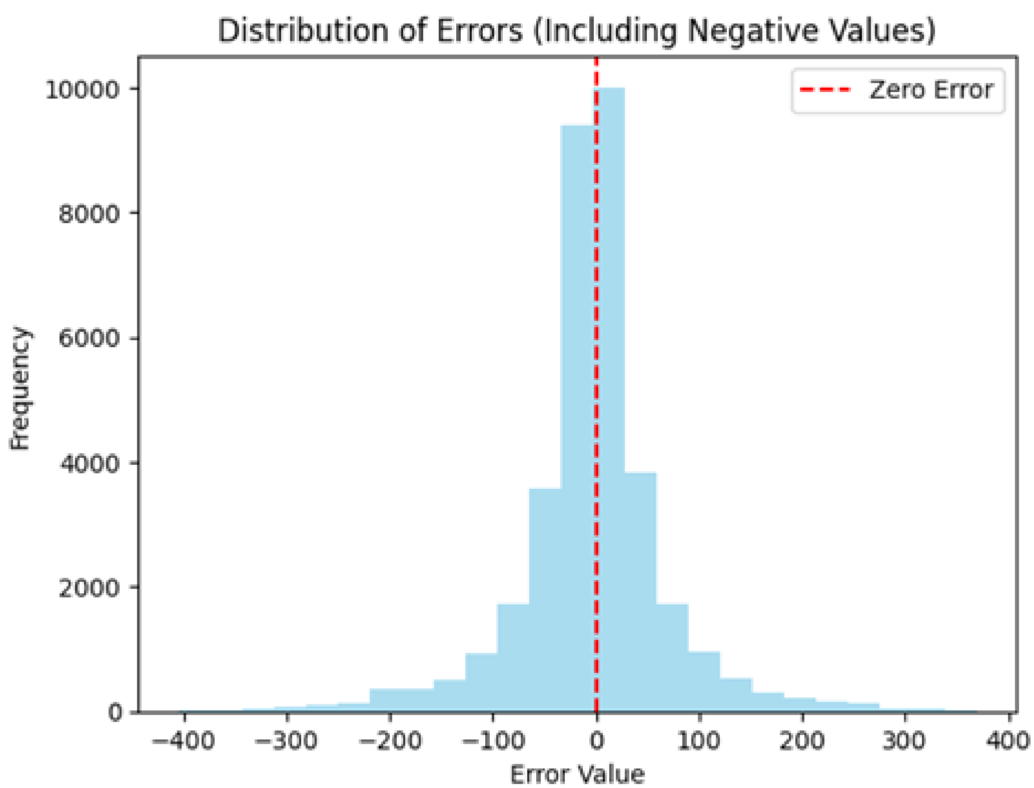


Figure 54: A histogram of E3 predicting the horizon.

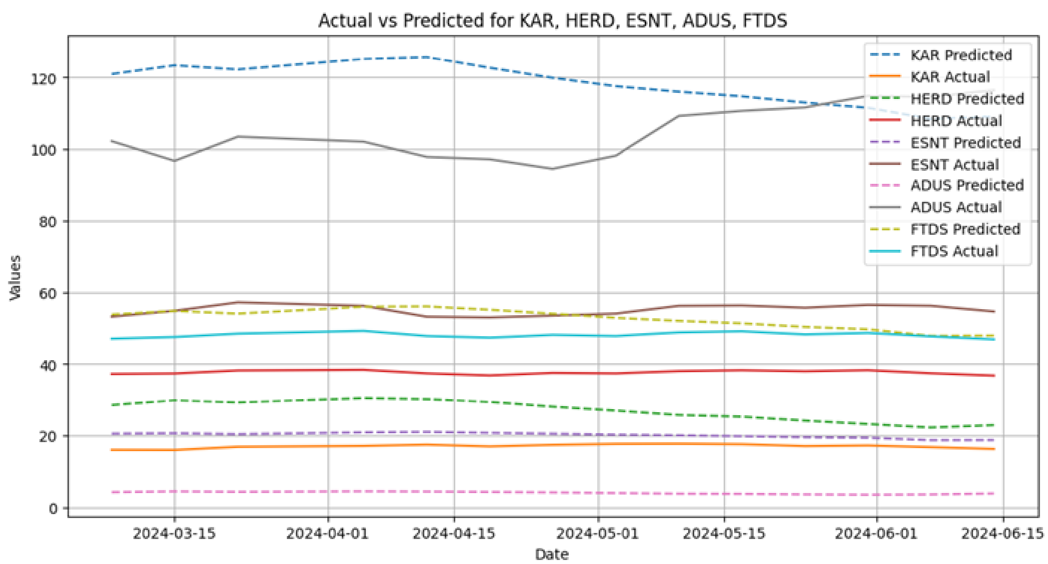


Figure 55: The line graph shows the Actual vs Predicted for some stocks, many of which have been poorly predicted.

C Code