Leveraging Machine Learning for Financial Forecasting: A Dual Approach to Meme Stock Price and GDP Prediction

Master of Philosophy

By

Piyumi Perera

Strathclyde University

2024

Leveraging Machine Learning for Financial Forecasting: A Dual Approach to Meme Stock Price and GDP Prediction

By

Student Name: Piyumi Perera

Student No: 202260979

For the Master of Philosophy

Under the Supervision of Prof. Anil Fernando

2024

Table of Contents

Abstract

This dissertation explores the adoption of machine learning in financial analysis through a systematic investigation of stock price prediction and GDP forecasting. The first study examines the relationship between online sentiment and meme stock price movements. This is done by conducting sentiment analysis on Reddit's WallStreetBets discussion thread and identifying daily popular stock tickers and their sentiment scores. The study then uses several recurrent neural networks, including variations of Long Short-Term Memory (LSTM) models such as single-layered LSTM, regular stacked LSTM, bidirectional LSTM model and Gated Recurrent Unit (GRU) model. These models were trained on five years of historical and technical data, highlighting the impact of online sentiment on Meme stock price fluctuations and the potential of AI-driven models to capture these dynamics. Models were tested for real-time applications for three consecutive days. Results demonstrated that the single-layered LSTM model outperformed other models with low error rates. For example, NVDA with average RMSE: 4.64, MAE: 3.38, MAPE: 0.035, and similar performance observed for ASTS with average RMSE 2.03, MAE: 0.92, MAPE 0.091 and LUNR (RMSE: 0.41, MAE 0.28, MAPE 0.066). However, by the third day regular stacked LSTM model slightly outperformed for NVDA, while single-layered LSTM dominated with better predictive power for other stocks (SMCI – RMSE 112.726, MAE: 86.946, MAPE: 0.111; AI – RMSE: 1.6, MAE: 1.197, MAPE: 0.036).

The second study extends the use of AI in macroeconomic forecasting, focusing on the prediction of the GDP of the United Kingdom (UK) using vital macroeconomic variables, including energy prices, unemployment rate, inflation, net migration and Real Effective Exchange Rate (REER) from 1990-2018. For the prediction, several machine learning models, such as Support Vector Regression (SVR), Random Forest (RF) and Gradient Boosting Machines (GBM), were implemented and compared together with Shapley Additive exPlanantions (SHAP). The models were assessed using evaluation metrics such as Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and R^2 score. The findings underscored the significant role of macroeconomic variables in economic forecasting and illustrated the potential of AI-driven models to provide valuable insight into financial markets and economic indicators. Among them scaled SVR

model achieved best performance with RMSE: 83,492. 048, MAE: 77, 219.274, MAPE: 4.2% and R^2 score of 0.042.

Together, these studies demonstrate the adoptability and potential of machine learning in addressing complex financial and economic prediction tasks and underline practical implications. The integration of sentiment analysis for stock price prediction and macroeconomic modelling for GDP forecasting showcases machine learning's ability to handle diverse data types, from unstructured textual data in online platforms to structured economic indicators. By combining these approaches, the research highlights how AI can uncover hidden patterns and relationships that traditional financial models might overlook, providing a more nuanced understanding of market behaviors and economic trends.

Acknowledgement

I, Piyumi Perera, express my heartfelt appreciation to my supervisor, Prof Anil Fernando, for his continuous guidance, expert insight and patience across this research journey. His mentorship played a vital role in directing and ensuring the successful completion of this work. I am truly grateful for his encouragement and constant support.

Table of Figures

Table of Tables

Definition of Key Terminologies

- Sentiment Analysis – The process of determining the emotional tone expressed through a body of text (Bogle & Potter, 2015).
- Meme Stocks – Stocks that have gained popularity among investors due to viral online discussions (Aloosh, et al., 2021).
- WallStreetBets (WSB) – A subreddit on Reddit online forum where retail Investors discuss trading strategies, often prompting meme stocks (Nobanee, et al., 2023).
- Stock Price prediction – The process of forecasting the future price of a stock using Historical Data and models (Hossain, et al., 2018).
- Recurrent Neural Networks (RNNs) – A type of artificial neural network designed to recognize patterns in sequential data (Pawar, et al., 2019).
- LSTM (Long Short-term Memory) – A specialized form of RNN capable of learning long-term data dependencies (Sandhya, et al., 2022).
- GRU (Gated Recurrent Unit) – A type of RNN models that are designed to solve vanishing gradient problem in sequential data (Pirani, et al., 2022).
- VADER (Valence Aware Dictionary and Sentiment Reasoner) – A lexicon and rule-based sentiment analysis tool specifically designed for social media platforms (Hutto & Gilbert, 2014).
- Technical Indicators – Calculations based on historical stock price data that help identify the trends and patterns in stock market data (Oriani & Coelho, 2016).
- Gross Domestic Product (GDP) – The total market value of goods and services produced within a country for a specific period. It is a key indicator that gauges the health of an economy (Ruiz, et al., 2011).
- Linear Regression (LR) – A basic Statistical model that estimates the relationship between dependent and one or more independent variables using a linear equation (Trivedi, et al., 2021).
- Macroeconomic Variables – factors that describe the state of an economy at a large scale (Ali Shah, et al., 2013).
- Support Vector Regression (SVR) – SVR is a type of Support Vector Machine (SVM) used for regression tasks. It finds the hyperplane that best fits the data in a high dimensional space while minimizing the prediction error (Xu, et al., 2014).

- Random Forest (RF) – RF is an ensemble learning method that utilize multiple decision tree for training and output the average prediction of the regression trees (Yoon, 2021).
- Gradient Boosting Machines (GBM) - this is an ensemble learning technique that builds models sequentially, where each models try to correct the error of the previous models, leading to better predictive performance (Yoon, 2021).
- SHAP (Shapley Additive Explanation) – SHAP is a method that explains individual predictions made by a machine learning model attributing to the predictions to each feature based on SHAPLEY values from cooperative game theory (Kim, et al., 2024).

1 CHAPTER ONE

1.1 Introduction

The rise of social media has triggered a multitude of unexpected changes worldwide, mainly in how people consume and interact with news and information. Unlike in the past, news and information are consumed in a rapid, on-demand manner, with headlines instantly reaching digital devices through a number of social media platforms. The already growing digital consumption was accelerated by the COVID-19 pandemic in 2020 (McNamara, et al., 2020) Consequently, financial markets and related communication methods have also shifted towards Internet platforms. Hundreds and thousands of information relating to financial securities and the equity market are constantly shared through popular social media platforms by institutions and individuals. As a result, retail investment has gained growing popularity among the general public that is driven by the swift growth of social media platforms and the availability of low commission, easily accessible trading platforms.

Meme stocks, known for their extreme volatility, have recently gained significant global attention due to their responsiveness to social media-driven, irrational herding behaviours. This phenomenon is highly influenced by platforms like Twitter and Reddit, especially by the subreddit r/wallstreetbets community (Umar, et al., 2021). The influence of meme culture on investing, particularly in the case of GameStop and the subsequent short squeeze, has demonstrated the power of collective sentiment in driving stock price movements.

The first study explores the relationship between online sentiment and meme stock price changes by performing sentiment analysis on r/wallstreetbets' discussion threads to identify the most popular stock tickers and their sentiment score. Subsequently, it is aimed to predict the next day closing price of identified popular stocks individually using several recurrent neural network models such as Single-Layered Long Short-Term Memory (LSTM), Regular Stacked LSTM, Bidirectional LSTM, Gated Recurrent Unit (GRU) that are trained on historical and technical data spanning nearly a decade.

Gross Domestic Product (GDP) is a critical measure of economic health that denotes the total value of goods and services produced by a country over a specific period (Landefeld, et al., 2008). Accurate GDP prediction is an essential component of informed government planning, policymaking, and business decision-making since it aids in anticipating economic trends and effective resource allocation (Landefeld, et al., 2008). Traditional econometric models such as General Equilibrium (DSGE), Ordinary Least Squares (OLS) regression (Pérez-Pons, et al., 2021) and Autoregressive Moving Average (ARMA) (Choi, 2012) have been dominating economic forecasting for an extended period of time. However, these traditional methods struggle to interpret the non-linear interdependent relationships among macroeconomic variables that are growingly important components of modern economies.

Accurate GDP forecasting requires more advanced methods that are capable of capturing these complexities, considering the substantial fluctuations in the global economies, including the UK, that affect the energy markets, trade conditions and labour dynamics (DeJong, et al., 2005). In recent times, machine learning models have emerged as powerful alternatives to traditional methods in capturing these interconnections more accurately (DeJong, et al., 2005). This study aims to predict the GDP of the UK by employing several machine learning models, including Linear Regression (LR), Support Vector Regression (SVR), Gradient Boosting Machines (GBM), and Random Forest (RF). Additionally, SHAP (Shapley Additive exPlanations) is used to interpret the influence of each macroeconomic variable on the GDP predictions to enhance transparency and understanding of the model outputs while providing valuable insight into the impact of macroeconomic variables on the economic direction of the UK.

The second study aims to address the limitations of traditional models while providing more accurate and interpretable predictions that ultimately contribute to better economic decisions and policy formulation in the UK.

This thesis is aimed at making a considerable contribution to the current body of literature, as the following research papers have been produced from both studies - Impact of Energy Prices and Macroeconomic Variables on GDP Prediction UK: Machine Learning Approach (Perera & Fernando, 2024) in the Journal of Business and Management Studies address the gap in literature by exploring the predictive impact of important macroeconomic variables and Stock Price Prediction with PCA-Stacked-LSTM, GRU and Voting Regressor Ensemble: A SHAP Interpretation, currently under review in the Review of Business and Economics Studies addresses gaps in current literature by implementing RNN models in stock market domain and addresses the explainability issue in next day price predictions through SHAP. The first research paper provides summaries of the fifth and sixth chapters of the thesis, while the second paper contributes to the third and fourth chapters. These research papers aim to extend the academic discourse in macroeconomic forecasting and financial market predictions.

1.2 Motivation and Research Scope

The remarkable emergence of meme stocks has exposed the need for understanding the dynamic relationship between social media sentiment and stock market behaviour. Traditional financial prediction models often overlook the collective influence of individual investor sentiment that emerges through online platforms. This novel trend in financial markets has presented unique opportunities to develop models that can harness sentiment data to predict stock price movements.

The research scope of this study consists of the collection and analysis of Reddit, r/wallstreetbets' data, identification of the most popular stock tickers based on the compound sentiment scores, and the development of a predictive model with the aid of historical prices and technical data.

As for the UK GDP prediction study, accurate GDP prediction is a crucial requirement for understanding the relevant economic direction and performance in making informed policies and business decisions (Maccarrone, et al., 2021). Traditional econometric models have provided a valuable foundation for econometric forecasting. Although traditional models fall short of capturing the complex, non-linear, and interdependent nature of macroeconomic relationships (DeJong, et al., 2005). The growing importance of energy price changes, labour market changes, and trade dynamics in shaping the UK economy has created a need to adopt more advanced predictive models that can effectively manage and understand these complexities.

Machine learning techniques have emerged as a powerful tool in economic forecasting due to their ability to handle large datasets and their capability to identify intricate relationships (Maccarrone, et al., 2021) among variables. This study is motivated by the need to enhance GDP prediction accuracy by leveraging machine learning models with crucial macroeconomic variables and thereby overcoming the shortcomings of conventional approaches. Moreover, the application of SHAP values acts as an added interpretability aspect, which makes the predictions more transparent and actionable to policymakers, businesses and other relevant parties. Further, it produces clear insights into the impact of each macroeconomic variable on GDP specific to the UK.

The UK's economic landscape is characterised by fluctuations in energy markets and significant labour and trade dynamics (Barnett, et al., 2014), which present unique challenges and opportunities for GDP prediction. While most literature is focused on the US economy, this research aims to provide insight specific to the UK, filling the gap in the literature and providing practical information for UK decision-makers.

The scope of this research is focused on GDP prediction of the UK by utilising macroeconomic variables as predictors such as energy prices, unemployment rate, inflation, net migration and Real Effective Exchange Rate (REER) from 1990 -2018. This involves the application of a variety of machine learning models, such as LR, SVR, RF, and GBM, to evaluate the effectiveness of GDP prediction. This study also assesses the impact of data processing techniques, such as scaling and hyperparameter tuning, on model performance. Additionally, SHAP was integrated into the models to enhance understanding and transparency by gaining an in-depth understanding of each variable's influence on the GDP.

1.3 Background of Research

Meme stocks have become a viral phenomenon in recent times (Nobanee, et al., 2023). These stocks are heavily influenced by online retailer communities and social media news. Out of all the social media sites and communities, 'r/ WallStreetBets' (WSB) has gained the most popularity due to the discussions of this forum that hypes up unconventional stocks and surge the stock prices to unprecedented heights. Specifically, stocks such as GameStop (GME) and American Multi-Cinema (AMC) have become household names for this novel intersection in stock markets known as 'meme stocks' (Umar, et al., 2021).

Meme stocks differ from conventional stocks like FTSE 100 or S&P500, which are typically driven by company performance, financial reports and broader economic indicators. Instead, meme stocks are characterised by high volatility, short-term speculative behaviour and large-scale price changes that often disconnect from the underlying value of the company. This novel category of stocks is dominated by retail investors who have less experience and knowledge of stock trading but are highly influenced by discussions on social media platforms, where sentiment and collective enthusiasm can override traditional market signals (Nobanee, et al., 2023).

The ability to predict meme stock prices poses unique challenges because conventional prediction models that perform well on blue-chip stocks often fail to detect and capture the erratic nature of meme stocks (Nobanee, et al., 2023). Therefore, this project aims to comprehensively analyze the daily popularity and sentiment of the top five most-mentioned meme stocks on WallStreetBets utilizing advanced sentiment analysis techniques, and out of them, the top three stocks will be utilized in predicting next-day closing price using RNNs which have shown to excel at predicting time series data, particularly in financial markets and stock prediction field (Rout, et al., 2017). Particularly, this thesis focuses on comparing RNN architectures, including single-layer LSTM, Regular Stacked LSTM, Bi-LSTM and GRU, to determine which model best predicts meme stock prices.

This study contributes to the growing body of literature exploring the behaviour of retail investors influenced by social media discussions. Previous studies in the stock market domain have mostly concentrated on traditional stocks and fundamental factors. Therefore, this study expects to enhance the current understanding of the stock market by focusing on the novel and most likely impactful variable in the future of stock markets.

When considering the second study, GDP is a significant indicator of the economic health of a country that provides valuable insights into the overall worth of products and services generated by economic activities within a certain period (DeJong, et al., 2005). Classical econometric models have traditionally been a major source of GDP projections (Pérez-Pons, et al., 2021). Furthermore, these conventional models have been a crucial component of macroeconomic forecasting as they offer a foundational framework for assessing economic growth and detecting sudden changes in the economy, which is essential in planning and implementing effective government policies and making accurate business decisions (Pérez-Pons, et al., 2021). Out of the conventional economic forecasting models OLS, Generalised Spatial General Equilibrium (GSGE), Autoregressive Moving Average (ARMA) and Autoregressive Integrated Moving Average (ARIMA) have been cornerstones of economic forecasting by producing insightful forecasts based on historical connections between economic variables (Stevenson, 2007).

Even though traditional econometric models have produced valuable insights into the dynamics of the global and regional economies, they often struggle with unconventional shifts in markets due to their limitations. These limitations have invalidated most of the traditional models in recent times since globalisation and technological advancements have created complex interdependencies among economic variables that can not be captured by pre-determined equations or models that assume linearity (Pérez-Pons, et al., 2021). The global financial crisis in 2008 (Edey, 2009) can be taken as a prime example for this, since many traditional models failed to capture the sudden shock to the global economy that generated worldwide spirals to many sectors. Similarly, Brexit introduced considerable uncertainties and volatilities in the UK (Adam, 2019), which further demonstrated how the traditional econometric models are compromised in reflecting the unconventional changes in the economies. Therefore, these real-life situations underscore the need for more sophisticated predictive approaches that are capable of handling non-linear patterns and interactions among macroeconomic variables.

Energy prices, unemployment rate, and inflation are key macroeconomic variables considered in this study. These variables are often tightly related to one another by creating complex direct and indirect effects. For instance, growing energy prices have the potential to raise the cost of production, which results in inflationary pressures that can have a negative influence on GDP growth and consumer spending patterns (Khan & Chhapra, 2016). Similarly, unemployment is a major factor that affects household income, which in turn can generate harsh impacts on saving rates and demand for goods and services, which eventually reflects through the GDP (Chowdhury, et al., 2019). Therefore, more sophisticated methods are crucial in recognizing these non-linear and frequent interdependencies among economic variables. Because conventional models are likely to oversimplify representations of economic dynamics through predefined formulas (DeJong, et al., 2005).

In recent times, machine learning models have become strong substitutes to overcome the shortcomings of traditional models. Machine learning models are highly valuable and well suited to capturing subtle changes in contemporary economic systems because they have the ability to process large datasets and discover intricate, direct and indirect non-linear patterns among economic variables (Pérez-Pons, et al., 2021). These models have gained growing popularity in various fields, such as finance and healthcare (Hossain, et al., 2021) due to their ability to provide accurate and flexible solutions in the face of complexities. Specifically, in economics, machine learning offers the opportunity to enhance GDP predictions by overcoming linearity constraints and incorporating high-dimensional data (Hossain, et al., 2021). Recent studies have shown that machine learning techniques such as SVR (Ülker & Ülker, 2019) and RF (Muchisha, et al., 2021) have outperformed traditional statistical approaches like time series analysis, offering improved accuracy in GDP forecasting.

The economic landscape of the UK presents unique challenges that justify the application of advanced machine learning models for GDP prediction. Between 1990-2018, the UK faced significant economic fluctuations driven by events such as the Gulf War (1990 -1991) that caused a sharp rise in oil prices due to instability in the Middle East (Wilson, 1994), the dissolution of the Soviet Union in 1991 that forced centralized economies to shift to market-oriented systems, which created a global impact on trade and investment flows (Flemming, 2010), the global financial crisis (2007-2008), which resulted in global recessions and massive unemployment (McKibbin & Stoeckel, 2010). Most importantly, the Brexit referendum in 2016 brought significant uncertainties to the UK economy by affecting trade, investment and labor flows (Hodson & Mabbett, 2009).

The above-mentioned global phenomenon contributed to the increased volatilities and complexities in the UK's economy during the study period (1990 – 2018). These incidents made it difficult for traditional models to accurately capture the full scope of economic dynamics. Therefore, by applying machine learning techniques together with SHAP interpretation, this study aims to address these complexities and provide more actionable insights into the factors that are driving the GDP of the UK.

1.4 Problem Statement

The study of meme stocks and their unconventional price fluctuations driven by social media sentiments is an emerging field, mainly influenced by the threads on social media platforms like Reddit WallStreetBets (WSB). Although social sentiment plays a substantial role in identifying popular meme stocks, traditional predictive models often fail to identify the unique patterns and behaviours associated with these stocks. This research aims to address two of the most crucial challenges in the field of meme stock price prediction and offer practical insights for financial analysts and investors.

1. Identification of popular meme stocks utilising social sentiment: Regardless of the current impact of social media on meme stock popularity, there is a lack of systematic methods to identify the most trending meme stocks using sentiment analysis. There is a crucial gap in effectively leveraging social media data to extract the stocks that gain the most traction among retail investors. Therefore, there is a vital need to address this gap since the discovery of popular meme stocks at the early stages of their popularity

is crucial in making more trustworthy and accurate predictions and investment decisions (Shiller & Robert, 2020).

2. Enhanced prediction accuracy with RNN models utilising technical indicators and financial data as predictors: traditional stock prediction techniques that utilize technical and historical financial data function well with ordinary stocks. However, they tend to frequently fail to capture the extreme volatilities of meme stocks. Unlike traditional machine learning models, novel neural networks have shown promising results in capturing temporal correlations and complicated data patterns (Pirani, et al., 2022). Therefore, there is a growing need for the application of neural networks for meme stocks while using their technical and historical financial data as predictors.

This study aims to accurately predict the next day closing price of the most popular meme stocks discovered through the sentiment analysis of discussions on the WSB community. For that, RNN models has been constructed, utilizing technical indicators and financial data as predictors. The main objective of this study is to create a solid framework that will assist financial analysis and independent investors in navigating the quickly changing domain of meme stocks.

Regarding the second research, conventional economic models have been beneficial for GDP prediction. However, they frequently fail to take non-linear and interdependent characteristics of macroeconomic variables into consideration (DeJong, et al., 2005). Due to these restrictions, there is a crucial need for more sophisticated forecasting models that are accustomed to different economies. In light of the significant volatilities in the UK's energy markets, labor trends and trade dynamics, it is imperative to develop models that are capable of predicting GDP specifically for the UK. This study addresses this gap by utilizing machine learning techniques that can capture complex patterns in macroeconomic variables together with SHAP to interpret the impact of each macroeconomic variable on GDP fluctuations, allowing for more accurate and transparent GDP prediction.

1.5 Aims and Objectives

The meme stock price prediction project aims to develop a predictive framework to forecast the next-day stock price of meme stocks, identified through daily sentiment analysis of WallStreetBets posts and comments. The price prediction model utilizes variations of LSTM and GRU with financial data and technical indicators, with the aim of producing a comprehensive tool for investors and analysts to navigate through meme stocks' volatility effectively without relying only on social media trends.

Objectives –

1. Develop a method to identify daily popular meme stocks: Design a systematic approach to analyzing the sentiment of posts and comments on WallStreetBets daily to identify the most popular meme stocks.
2. Implement RNN models: Develop and implement RNN models that utilize historical financial data and technical indicators to predict the next day closing price of the identified meme stocks.
3. Evaluate model performance: Assess the prediction accuracy of the RNN models using metrics like Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE). Then, assess the stock price prediction against actual values for three consecutive days.
4. Provide recommendations to investors and future studies in the field: Based on the findings, practical guidelines and strategies shall be provided for retail investors and financial analysts to improve their decision-making process when handling meme stocks along with recommendations to enhance accuracy and improve prediction models specifically for future studies in the meme stock field.

The prime objective of the second study is to predict the GDP of the United Kingdom by utilizing several macroeconomic variables from 1990 to 2018. These variables are utilized as predictors on machine learning models, and their prediction performance is compared. Specifically, this study aims to:

1. Analyze the impact of key macroeconomic variables on GDP prediction with more weight on energy price influence.

2. Application of machine learning models – Random Forest (RF), Support Vector Regression (SVR), Gradient Boosting Machines (GBM) and a statistical model - Linear Regression (LR) as a baseline model to predict GDP.

3. Application of SHAP to interpret each macroeconomic variable's contribution to each model's predictions.

4. Compare the performance of these models using evaluation metrics such as Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and R^2 score.

5. Identify the model that offers the highest prediction accuracy for GDP.

1.6 Research questions

The studies address the following questions:

1. How can sentiment analysis of WallStreetBets be effectively utilised to identify the most discussed meme stocks on a daily basis?
2. What key financial indicators can be utilised to increase the prediction accuracy of the RNN models?
3. How is the prediction accuracy of RNN models compared to one another?
4. How can the insights gained through sentiment-driven price predictions be translated into practical investment strategies for retail investors and financial analysts?

5. How do macroeconomic variables, particularly energy prices, influence GDP prediction in the UK?

6. What machine model demonstrates the highest accuracy in predicting GDP using macroeconomic data?

7. How do different models perform in terms of accuracy and prediction reliability?

8. How can SHAP be used to enhance the interpretability of each model?

1.7 Novelty and Contribution

The meme stock price prediction study contributes to the existing literature by offering a unique perspective on the performance of predictive models in the context of meme stocks that differ significantly from conventional stocks like FTSE100 or S&P500. Since, in traditional stock markets, price movements tend to be driven by long-term fundamentals and established trends, and complex models. Models such as Bi-LSTM, Extreme Gradient Boosting (XGBoost), Deep Neural Networks (DNN) and Auto Regressive Integrated Moving average (ARIMA) are highly utilized and have shown to be effective on conventional stocks due to their complexity, need for the vast amount of data and long-term trend dependencies.

However, this study reveals that, for meme stocks where retail investors' herding behaviour and short-term trends play a major role, simpler models like single-layered LSTM display better performance. As the results present, single-layer LSTM strikes a balance between complexity and the ability to adapt to the highly volatile and sentiment-driven nature of meme stocks. Hence, these findings challenge the conventional assumption presented by prior literature that expresses that more complex models always outperform simpler models. As this study demonstrated, model performance can vary depending on the type of stock being analysed.

The insights gained from this study can help retail investors and financial analysts better understand how meme stocks behave, offering a predictive edge in short-term trading. Additionally, the findings underscored the potential of using sentiment analysis from platforms like WallStreetBets as a filtering mechanism to focus on popular stocks. A combination of the sentiment indication together with another layer of market analysis could drive informed decision-making. This could be particularly useful in determining when to enter or exit trades in top-rated meme stocks while minimising the risk of catching up with the hype cycles in the wrong time which can lead to sudden price drops.

The UK GDP prediction study contributes several ways to the current literature on machine learning for economic forecasting. One of the key contributions is the comprehensive comparison of multiple machine learning models, such as SVR, GBM, and RF, along with the baseline statistical model, LR. By applying and evaluating several machine learning models, this study provided an in-depth analysis of the strengths and weaknesses of each approach for GDP prediction in the context of the UK. The detailed analysis helped illustrate how each model handled the complexities of macroeconomic data, such as non-linearity and interdependencies among features. Such comparative analysis offers valuable insight into the suitability of different machine learning techniques for macroeconomic forecasting while highlighting the importance of selecting the type of model depending on the inherent relationship between predictors and target variables. This can aid future researchers and practitioners in selecting appropriate models for similar economic variables-related tasks.

This research makes a significant contribution by evaluating the impact of feature scaling and hyperparameter tuning on the performance of machine learning models. This was done by systematically applying and testing with both scaled and unscaled data and tuning the hyperparameters. This study clearly illustrated how hyperparameter tuning affects kernel-based models and tree-based models. As an example, hyperparameter tuning extensively improved the SVR model's predictability but some models, such as RF and GBM, experienced no obvious difference in predictions with parameter tuning. Similarly, scaling improved the SVR model's performance, although it contributed to deteriorated predictions in LR and RF models. This analysis provided insight into the sensitivity of different models to preprocessing methods. Therefore, this study contributed valuable insight into designing and optimising machine learning-based models in economic forecasting.

One of the study's major contributions is the application of SHAP to provide transparency and interpretability to the prediction models. By incorporating SHAP into the prediction models, this study bridged the gap between the 'black box' nature of machine learning models and the need for transparency in economic decision-making. SHAP analysis allowed for the relative importance of each macroeconomic feature in determining GDP, which ensured the explainability of the model's predictions to the stakeholders and policymakers, enhancing the study's relevance and applicability.

As expressed through the literature review, most of the past studies have focused the GDP predictions on the US or other larger economies. This study focused on UK-specific macroeconomic data from 1990 to 2018, contributing a novel perspective to the GDP prediction with relevance to the UK. This provided insight into how different macroeconomic variables have impacted GDP in the UK over an extended period while highlighting trends and influences that are specific to the region. This context-specific research produced essential input to economic forecasting as it accounts for unique economic conditions, policies and structural characteristics specifically related to the UK, making the findings highly relevant to regional policymakers and economists.

By employing a range of models, from simple linear regression to more complex ensemble methods like GBM, this study contributed to the understanding of the trade-off between model complexity and generalizability. While more complex models like GBM and RF captured the non-linear relationships more effectively with their default parameters, the baseline LR model provided reasonable performance with its simpler structure. This contribution is valuable for practitioners who are seeking to strike a balance between model complexity, interpretability, and prediction accuracy, particularly in settings where computational efficiency and understandability are important considerations.

Therefore, these contributions make this study unique in its thorough application of multiple machine learning techniques, SHAP detailed feature analysis and the UK-specific focus. Hence, by providing comprehensive insight into model performance, interpretability and the impact of different macroeconomic variables on GDP, this research not only enhances the current understanding of economic forecasting but also provides a practical understanding of the interdependencies between economic variables of the UK for future studies and policymakers.

1.8 Dissertation Outline

The structure of this dissertation is organized as follows:

Chapter 1: Introduction

This Chapter sets the foundation for the research by providing an overview of the key objectives, direction, and significance of the study. The introduction also sets the foundation for the study by highlighting the sudden rise of meme stocks influenced by online communities such as WallStreetBets.

Additionally, this chapter established foundation to the second study by providing background information by providing information of brief historical development of GDP prediction and its value to the economic analysis and direction. It presents the problem statement, emphasizing the limitations of traditional econometric models in capturing non-linear relationships between macroeconomic variables.

Furthermore, research questions and objectives of both studies are presented to guide the flow of both the research. The overall structure of this thesis is also summarized through this chapter.

Chapter 2: Literature Review

The literature review synthesizes the body of research on stock price prediction, sentiment analysis in price forecasting, and the use of machine learning models in financial forecasting. It provides an in-depth discussion on the emergence of meme stocks, prior studies on meme stocks, and the power of social media sentiment in shaping stock market trends. Further, it discusses the strengths and limitations of machine learning models like LSTM, GRU, and traditional econometric approaches that are applied in financial forecasting.

Moreover, this chapter addresses GDP and the impact of macroeconomic variables on the economy through fundamental economic theories such as Neoclassical, Keynesian and endogenous growth theory. Then, it discusses the traditional methods of GDP prediction. This chapter also highlights the emergence of machine learning in economic forecasting. Finally, identifies gaps in existing literature while explaining how this study aims to address them. This chapter establishes the theoretical framework and produces a comprehensive view of the current study.

Chapter 3: Sentiment Driven Meme Stock Price Prediction: Analysing WallStreetBets discussion to Forecast Stock Prices with Recurrent Neural Networks.

This Chapter addresses the research design and methodology as well as the results of the first study.

Research Design and Methodology: This first part of the third chapter describes the research design, outlining the methods and tools used to conduct the analysis. It covers the data collection process, which includes data scraping from Reddit's WSB using PRAW and historical data extraction from Yahoo Finance, as well as technical indicator generation using the extracted historical data. Further, this chapter details the machine learning models (Single-layered LSTM, Regular stacked LSTM, Bi-LSTM and GRU) employed to predict the next-day stock prices, explains the evaluation metrics and real-time next-day price prediction utilised to assess the model performance.

Results: Secondly, this chapter presents the results of the study, highlighting the performance of the different machine learning models across the selected meme stocks. The chapter includes visualisations, such as plots and loss curves, as well as a comparative analysis of evaluation metrics (RMSE, MAE, MAPE) for each model. Comparative analysis is conducted to identify which model performed best across different stocks and why certain models outperformed others based on characteristics and volatility.

Chapter 4: Impact of Energy Prices and Macroeconomic Variables on GDP Prediction UK: Machine Learning Approach with SHAP Interpretability.

This Chapter addresses the research design, methodology, and results of the second study.

Research Design and Methodology: The first section of the chapter outlined the design of the study that utilised machine learning models to forecast the UK GDP prices using macroeconomic factors as predictors. It describes the methods used in gathering data from different sources. The preprocessing techniques, the evaluation methods and the evaluation metrics used for model performance evaluation are also explained in this chapter. Further, the application of SHAP to gain an in-depth understanding of the underlying factors of predictions is also interpreted through this chapter.

Results: The latter part of this chapter presents the results of each model and their performance according to various performance analysis metrics. Additionally, a comparative analysis is provided on each model's performance in order to highlight the best-performing model in the GDP prediction of the UK. A thorough SHAP analysis was also included to show how each variable affects the predictions.

Chapter 5: Discussion and Analysis

This chapter interprets the findings from the previous chapter, connecting the results to the literature review. It analyses the strengths and weaknesses of each study's best and worst-performing models and the potential reasons behind their performance.

The former part of the discussion chapter analyses the meme stock price prediction models. It expresses the reasons for the Single Layered LSTM model's exceptional performance in terms of meme stocks. Further, this chapter reflects on the role of sentiment analysis in filtering relevant stocks for prediction and how the characteristics of meme stocks differ from traditional stock market assets.

Subsequently, this chapter presents interpretations of the key findings of the second study in the context of the UK economy, focusing mainly on the impact of energy prices and other macroeconomic factors on GDP prediction. By comparing the findings of this study with prior work relating to this area, the results are validated. Further, this chapter highlights the practical consequences of the findings that can provide novel and valuable perspectives to UK policymakers and explores the theoretical implications of the findings in relation to economic theories.

Chapter 6: Conclusion and Recommendations

The main conclusions of both research studies are included in this chapter. This chapter also emphasises how crucial it is to use machine learning models that are suited to the relevant study's characteristics. It evaluates how the study accomplished relevant objectives while emphasising on the originality and additions to the current work in the field of economic forecasting. Additionally, the study's limitations are emphasized while providing suggestions for future work are given to further improve the current understanding of financial forecasting.

2 CHAPTER TWO

Literature Review

2.1 Introduction

This chapter presents a comprehensive review of the existing literature on the application of machine learning in financial forecasting with a focus on two distinct but interconnected domains, namely, meme stock price prediction and GDP forecasting. As financial markets evolve in complexity and volatility, traditional econometric models have displayed limitations in capturing the nonlinear, high-frequency and sentiment-driven dynamics of modern financial behaviour. In response, machine learning methods have gained traction due to their ability to process large volumes of data, identify hidden patterns and improve predictive accuracy. The first part of the review explores and compares recent developments in stock market prediction with traditional methods. As well as the rise of highly volatile, socially driven assets known as meme stocks. The second part focuses on GDP prediction, a crucial task in macroeconomic planning and policy making. This review explores the evolution of macroeconomic forecasting from statistical models to machine learning. By synthesising recent advances in both domains, this chapter highlights the potential and versatility of machine learning in addressing diverse forecasting challenges.

2.2 The COVID-19 Pandemic Influence on Economic Indicators and the Stock Market

The COVID-19 pandemic is an unprecedented event that has significantly impacted the global economy and resulted in notable drawbacks in many fields (McNamara, et al., 2020). This critical event demonstrated severe uncertainties that were highly visible through macroeconomic variables (Altig, et al., 2020). Especially, the surge in unemployment rates helps illustrate the magnitude of the crisis. In the UK, the unemployment rate increased from 4% to 5% by the end of the year 2020. As a result of that, benefit claims increased by 1.43 million pounds between March and May 2020, further increasing to the highest level of 2.70 million by August 2020 (UK Parliament). Meanwhile, the USA's unemployment rate soared from 3.5% to an all-time high of 14.7% (Suomi, et al., 2020). GDP is another significant indicator that highlights the economic consequences of the pandemic. As an example, the UK experienced its largest monthly GDP drop of 20.4% in April 2020; correspondingly, the USA GDP crashed by 11.2% from Q4 2019 to Q2 2020 (Jena, et al., 2021).

The economic response to COVID-19 was unpredictable, mainly due to the constant and extensive economic downturns (Altig, et al., 2020). This led to the emergence of different measures of uncertainty. Stock market volatility is one of the most dominant measures, while the volatility index (VIX) is the most prominent near-term measure of stock market volatility (Malkiel, et al., 2018) (Baker, et al., 2016). The one-month VIX index measures the implied 30-day market volatility procured from options on the S&P 500 stock market index. Typically, the rise of the VIX index implies the poor performance of the stock market. In early 2020, before the rise of COVID-19, the 1-month VIX index was valued at a steady range between 10 and 15. However, in March 2020, the daily VIX reached 82.7, which surpassed the highest VIX value recorded in history, which was 80.9 during the 2008 Global Financial Crisis.

Other effective measures of economic uncertainty include the Twitter Economic Uncertainty (TEU) index (Altig, et al., 2020) that tracks variants of economic uncertainty-related words in tweets starting from 2010. Macro forecaster disagreement is another effective measure that calculates the standard deviation across forecasts one year ahead of the annual real GDP growth rate. The Economic Policy Uncertainty (EPU) Index (Baker, et al., 2016), which tracks the frequency of economic uncertainty-related words in around 2000 US newspapers; the index uses the average uncertainty measured from 1985 to 2010 as a baseline value of 100.

Table 2.1 COVID-19 Impact on Different Measures

Measure	Value in Jan 2020	% Increase at the peak
VIX 1-month implied volatility, US	13.3	497
Twitter economic uncertainty, US	139.8	594
News, economic policy uncertainty, US	110.1	683
Macro forecaster disagreement, UK	0.5	1960

The stock market reacted strongly to the alarming signals of uncertainty indicated by the macroeconomic indicators. Although the COVID-19 pandemic showed similarities with the Spanish flu, the impact on the stock market was unmatched by any other infectious outbreak (Baker, et al., 2020). Even though the Spanish flu caused more mortality rates than COVID-19, it had minimal recorded stock movements that did not exceed 2.5% (Angel, et al., 2021). In contrast, COVID-19 caused over 12 significant changes in the US stock market (Baker, et al., 2020). As for an example, during only a period of one month (Feb- March 2020) S&P 500 went through three temporary trading halts due to drops exceeding 7%, resulting in a loss of over third of its value.

During the first wave of COVID-19, governments worldwide implemented stringent measures to reduce the virus spread, which caused the closure of essential businesses, extreme restrictions on flights and mandatory stay-at-home orders. These government-enforced restrictions, coupled with social distancing, significantly toppled the modern service-driven economy (Baker, et al., 2020), which resulted in a substantial contradiction in the economy and the stock markets.

Although stock markets globally suffered substantial losses due to the COVID-19 pandemic, most of them gained a fast and rapid recovery. As an example, NASDAQ, mainly composed of giants in the tech industry, gained fast recovery within a short period of time. While the Dow Jones Industrial Average, S&P 500, DAX(Germany), and Shanghai Composite Index (China) also made a swift recovery by the end of 2020 (Seven, et al., 2021). These recoveries were influenced by a combination of several factors, including government stimulus packages, monetary policy interventions, investor optimism, and an accelerated performance of sectors like healthcare, communication, and technology.

Out of these reasons, investor behavior change has the most impact on the stock market recovery. During the lockdown period, many people tended to explore the stock market through online trading platforms. Consequently, Robinhood's free trading service gained the highest popularity, tripling the trades on the app (Bellofatto, et al., 2018). Unlike institutional investors, these new retail investors showed rapid reactions to news data and tend to draw investments towards risky and controversial investments (Boehmer, et al., 2021).

2.3 Retail Investors and Investing Platforms

Retail investors are individuals who invest their own money in financial securities without professional training in investing or trading (Boehmer, et al., 2021). During the pandemic, retail investors grew primarily due to low-cost, easily accessible investing offered by online investing applications (Chaudhry, et al., 2021). Trading platforms such as Robinhood, Trading 212, Moneybox, and Freetrade offer investors the opportunity to trade stocks, options and other securities without any formal trading and minimum deposit. These platforms provide this service to the public, expecting to provide democratised, easily accessible investment opportunities to everyone (Chaudhry, et al., 2021). However, due to the lack of basic knowledge in investing, retail investors tend to make emotion-driven, illogical investment decisions.

The easy accessibility, free trading options and incentives used by investing platforms often encourage retail investors to behave irrationally (Chaudhry, et al., 2021). For example, investment platforms like Robinhood, Trading 212 and Weibull provide new users with free stocks for opening an account, while other platforms like Stash, M1 Finance and Acorns offer both cash and stock bonuses through referral programs. These techniques and gamified investment processes offered by online trading platforms potentially further drove retail investors' reckless investment decisions triggered by the excitement of quick gains.

Most of the retail investors consist of millennials who are more tech-savvy and highly influenced by social media and financial information online (Pagano, et al., 2021), (Li & Jasper, 2023). This demographic shift has brought a novel dynamic to the stock market, where social media and online communities play a vital role in influencing investment decisions. Especially platforms like Reddit, particularly WallStreetBets, have become the prime source of strategies and knowledge sharing for high-risk investments.

2.4 Memes and Meme Culture

Meme is a term introduced by evolutionary biologist Richard Dawkins in his book The Selfish Gene (Dawkins, 2016). The book used the term to represent a simple idea, style, or behaviour that spreads among people in a culture. In the internet culture context, memes usually take the form of an image, video, or text that is rapidly shared through social media and other forms of communication (Shifman, 2013). Typically, they are humorous in nature.

Meme culture refers to the process of creation, sharing and interpretation of memes. It has now become an essential aspect of the internet culture and plays a significant role in how people communicate and express opinions or facts online (Shifman, 2013), (Olena, et al., 2020). In the current social media-driven world, memes have created a strong communication channel where complex ideas and emotions are encapsulated in a more concise and shareable format, although ones who are unfamiliar with memes may find it hard to comprehend.

2.5 r/Wallstreetbets, GameStop and Short Squeeze

Reddit is a centralised community platform founded in 2005 that allows people to share their thoughts, knowledge and other related facts on various interests. Reddit currently has over 100,000 active communities with over 16 billion posts and over 82 million daily active users (Reddit, 2024). r/WallStreetBets (WSB) is one of the top communities on Reddit, with over 16 million members. It is known for the investment strategies and experiences shared by the users as they mostly present more unorthodox approaches to investing, often characterised by bold humour, absurdity and disregard for traditional financial advice. It is known for the investment strategies and experiences shared by the users as they mostly present more unorthodox approaches to investing, often characterised by bold humour, absurdity and disregard for traditional financial advice.

The GameStop Phenomenon was one of the pivotal events that catapulted WSB into the global spotlight. GameStop was a struggling American video game, consumer electronics and gaming merchandise company that had been heavily shorted by hedge funds, which had been speculated for the stock prices to fall. In 2019, Redditor Keith Gill, also known as 'DeppFuckingValue' on Reddit, started buying GameStop shares and call options while posting his trades and analysis on WallStreetBets. By January 2021, many other users took notice of Gills's posts and the potential for a short-squeeze due to the large number of short positions against GameStop. By the end of January 2021, GameStop shares skyrocketed from around \$20 to \$483 as retail investors continued to invest in them, eventually forcing short sellers to cover their position by buying back shares, creating a feedback loop that intensified the price increase (Umar, et al., 2021).

Due to the short squeeze, some hedge funds, like Melvin Capital, which has significant short positions, suffered great losses, leading to a call for additional capital or closure of short positions at a loss. This unexpected event caused noteworthy volatility in the market, leading to brokerage restrictions on heavily shorted stocks by brokerage firms, including Robinhood. This was a controversial action, and the brokerage firms were accused of market manipulation. Further, in the long run, this event attracted the attention of lawmakers, resulting in a congressional hearing to investigate the circumstances and regulatory requirements for a short squeeze (Oxford Analytica, 2021).

Therefore, the Wall Street Bets' GameStop short squeeze remains a landmark event in financial history (Figure 2.1), as it demonstrated the collective power of individual retail investors and the growing influence of social media on financial markets (Umar, et al., 2021).

Figure 2.1 GME Closing Price Increase

2.6 Meme Stocks

Meme stocks refer to the shares of companies that gain popularity among retail investors through social media platforms such as Reddit, Twitter, LinkedIn and TikTok (Yahya, et al., 2021). These stocks mostly attract the attention of retail investors due to unconventional factors such as viral trends, humour and coordinated buying efforts rather than fundamental financial performance, often leading to significant price volatility. GameStop, AMC Entertainment, and Blackberry are the most recent meme stocks that have experienced dramatic price movements due to online investor activities. These price movements were mostly fueled by online encouragement, potential short squeezes and retail investors' fear of missing out (FOMO), and WallStreetBets played a vital role in encouraging retail investors to take such high risks for potentially high returns. Some describe this novel phenomenon as a rebellious act by retail investors against professional investors (McCamy & Christal, 2021).

2.7 Evaluation of Stock Market Performance

A stock is considered a meme stock if it has gained distinct popularity among social media platforms, leading to highly inflated valuations (Yahya, et al., 2021). Therefore, most market participants view these stocks as overvalued, as their prices are mostly driven by social media hype rather than the fundamental value of the relevant company (Li & Jasper, 2023). To assess whether these meme stocks are overvalued, traditional stock evaluation methods play a crucial role. Below are some of the key evaluation metrics used to evaluate stock performance. These traditional methods provide the foundation for understanding whether the stock's current price reflects its intrinsic value or if it has been highly inflated through speculative trading.

2.7.1 P/E ratio

The Price-to-Earnings (P/E) ratio (2.1) is one of the most widely used measures of stock valuation and anticipated performance (Ogaluzor & Ebom, 2023). It is traditionally calculated by dividing the stock's current market price by its earnings per share (EPS) over the past year. Several forms of P/E ratio can be derived from using different timeframes and future earnings expectations. These variations of the P/E ratio can provide a wider scope for analysis. A higher P/E ratio usually indicates investors' confidence in the growth of the company, and the stock may be overpriced if those expectations are not met.

$$\frac{P}{E} = \frac{Share\ Price}{Earnings\ per\ share} \qquad 2.1$$

Regardless of the popularity of the P/E ratio as a valuation tool, its interpretation can vary significantly due to several influential factors (Ogaluzor & Ebom, 2023). First, the specific year in which the P/E ratio is calculated can affect its interpretation, as the average target P/E ratios can annually shift based on the investor sentiment and the economic condition of that year. Secondly, the relevant organisation's industry or sector plays a vital role. Meaning a highly advancing sector like technology often has a higher P/E ratio due to optimistic expectations. In contrast, organisations in more mature, slow-growth sectors like healthcare and real estate typically have lower P/E ratios. Thirdly, the organisation's market capitalisation has a direct influence on the P/E ratio, meaning larger companies have stable earnings, resulting in a higher P/E ratio. Also, company-specific factors, positive and negative news, breakthroughs or major contracts directly or indirectly influence the P/E ratio.

Since meme stocks are often driven by social media hype and retail investor sentiment, the P/E ratio can be extremely high or even negative due to the stock price deviating from its fundamental value. Hence, significant deviations from historical P/E averages or sector benchmarks can signal the overvaluation or undervaluation of a stock.

2.7.2 EV/EBITDA

This is another popular measure for assessing a company's valuation that provides a more comprehensive view (Persson & Ståhlberg, 2007)The Enterprise Value (EV) represent the total value of the company calculated by adding the firm's market capitalisation (total equity) to its total debts and subtracting any cash or equivalents to cash. This accounts for both the equity and debt of a company, giving a clear idea of a company's total value in the market. The Earnings Before Interest, Taxes, Depreciation and Amortisation (EBITDA*)* *(*2.2*)* serve as a proxy for operating cash flow by removing the effects of financial and accounting decisions, which makes it a valuable indicator of a company's operating performance (Persson & Ståhlberg, 2007), (Mauboussin, 2018).

$$\frac{EV}{EBITDA} = \frac{Equity + Debt}{EBITDA} \qquad 2.2$$

This ratio is usually expressed as a multiple (e.g., 8x - which indicates that the enterprise value of a company is eight times its annual EBITDA). This measurement provides a standardised value that can be compared with various other companies regardless of their size or capital structure, as it focuses on the company's core profitability irrespective of the way they are financed. This is more effective in comparing companies in the same industry since it neutralises the impact of different levels of debt and depreciation policies and aligns with the industry-specific norms and growth prospects.

For meme stocks, the EV/EBITDA ratio can reveal whether the stock is trading at a high multiple compared to its peer companies. Similar to the P/E ratio, a high EV/EBITDA can indicate overvaluation. By comparing a stock's value with that of a traditional peer company, investors can identify whether the price movement is grounded in fundamentals or purely driven by social media sentiment.

2.7.3 DCF Model

The Discounted Cash Flow (DCF) model (2.3) is a fundamental method for valuing a firm by estimating future cash flows and adjusting them to their present value. Unlike other valuation methods, this model considers the firm's expected future cash flows derived from historical performance data, which provides a futuristic perspective to its valuation (Bilych, 2013). This model estimates the present value of a company by calculating the net present value (NPV) of the future cash flow and subtracting it from its debts. This results in the firm's equity value, which can be divided by the number of outstanding shares to estimate the intrinsic value per share.

The DFC model's cash flow consists of two categories: Free Cash Flow to the Firm (FCFF) and Free Cash Flow to Equity (FCFE). FCFF refers to the available cash flow after covering all operational expenses before debt payments, which provides an overall view of the firm's cash generation capabilities. FCFE stands for the cash flow available to equity shareholders after accounting for debt servicing, expenses and necessary reinvestments, proving more realistic valuation that is more preferable by equity holders as it directly relates to the accruing value of equity holders (Bilych, 2013).

$$DCF = \frac{CF_1}{(1+r)^1} + \frac{CF_2}{(1+r)^2} + \cdots + \frac{CF_n}{(1+r)^n} \qquad 2.3$$

DCF – Discounted cash flow, CF_i - Cash flow period I, r– Interest rate, n – time in years before The future cash flow occurs

The discount rate is one of the most crucial components of the DCF model that converts future cash flows to their current value by factoring in the time value of money. After applying to the firm's value, the Weighted Average Cost of Capital (WACC) can be derived (2.4), representing the average return rate required by all security holders. WACC presents a combined cost of both cost and equity, reflecting a company's total cost of capital. While most components of WACC, such as the market value of equity and debt, can be derived from the company's financial statements, the cost of equity and debt calculation requires additional steps (Rehman & Raoof, 2010).

$$WACC = \frac{\sum_{i=1}^{N} r_i \ . MV_i}{\sum_{i=1}^{N} MV_i} \quad 2.4$$

WACC – Weighted average cost of capital, N – number of sources of capital r_i *– required rate of return for security, i- security* MV_i *– market value for all outstanding securities*

The cost of equity represents the return required by equity investors for the risk they endure, which is often calculated using the Capital Asset Pricing Model (CAPM), which considers the risk-free-rate (typically the return on 10-year U.S treasury bond, the stock's beta (relative volatility compared to the market), and the expected market return which is often based on historical averages such as the S&P 500 performance. On the contrary, the cost of debt represents the actual rate a company incurs on its borrowed funds after factoring in the tax advantages of interest payments since they are generally tax deductible. This adjustment is made by multiplying the cost of debt by (1-tax rate).

Although the DCF model is a popular and powerful traditional tool for measuring a company's intrinsic value, it has limitations as it mainly relies on assumptions about future cashflows, market returns, discount rates and company-specific risks. Even though minor changes in these speculated inputs can significantly change the outcome, it remains as a comprehensive framework that offers insight into the current share price alignment with the fundamental value (Bilych, 2013).

While the DCF model is more theoretical and long-term focused, it helps derive how meme stocks' market valuation is disconnected from their actual cash-generation potential. Therefore, a stock trading significantly above its DFC value clearly indicates that the stock's value is driven mainly by market sentiment, herding behaviour, and speculative trading.

2.8 RNN Models and Their Variants

Recurrent Neural Networks (RNNs) (Figure 2.2) are a class of artificial neural network that has been specifically designed to process sequential data, where every input is influenced by the prior input in sequence (Pawar, et al., 2019). This makes RNNs particularly effective in natural language processing, time series data analysis or any other task that requires the inputs to be in order (Patel, et al., 2021).

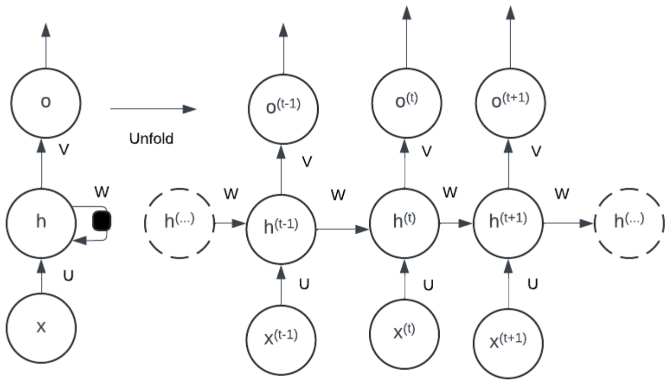


Figure 2.2 RNN Model Structure

There are several key components in RNNs –

- Sequential Inputs: RNNs process data sequences by taking each sequence element one at a time. (e.g.- stock price on day T is predicted by considering day 1 price, day 2 price …. Up to day T-1 price)
- Recurrent Structure: The 'memory' aspect of RNNs that allows them to capture temporal dependencies of data. This means the network output at each step is also influenced by the input at that step.
- Hidden Strate: the hidden state, denoted by h(t), serves as the memory of the network, which is a function of the previous state h(t-1) as well as the current input x(t). This facilitates the RNN to retain information from the prior inputs, which is critical in making accurate predictions in sequential tasks. And *f* stands for the function that updates the hidden state based on the current input and the previous inputs, which is an essential component in sequence-based tasks.

2.9 Machine Learning in Stock Price Prediction

To make well-informed decisions, investors need to conduct technical, sentimental and fundamental analysis (Sandhya, et al., 2022). This is a highly challenging and stagnant task because of the continuous flow of information in today's digital world. However, machine learning poses the ability to fuse both the tedious task of data crunching and human decision-making capabilities to aid in making well-informed decisions.

The oldest study relating to the use of machine learning in stock market predictions dates back to 1996, where they proposed a mixed-integer linear model for portfolio optimisation. This model utilised transaction cost, minimum transaction and minimum holdings, which are mostly disregarded by current optimisation models (Speranza & Grazia, 1996). The earliest study of neural networks in financial modelling utilised a neural network to estimate the nonlinear dynamics of the Kalman filter. This filter was designed to refine different types of noises that are commonly found in financial data (Bolland, et al., 1997). Most of the early studies have utilised Genetic Algorithms (GA) to solve the problem of financial time series forecasting (Lee, et al., 2001), (Ma, 2020), (Araújo, et al., 2007).

The application of machine learning in stock price prediction has gained significant attention in recent years, leading to the development of various models that are aimed at enhancing forecasting accuracy. Out of these machine learning models, the deep learning approaches. In particular, LSTM networks have been extensively employed for stock price forecasting. Studies such as (Nguyen-Trang, et al., 2024) utilized LSTM algorithms to predict stock price trends in Vietnam and were able to achieve a high accuracy rate of 93%. Whereas the research by (Vuletić, et al., 2024) reviewed the advancements of deep learning models for financial forecasting, where they highlight the improved accuracy and effectiveness of models such as transformers and Generative Adversarial Networks (GANs) in capturing complex market patterns.

Similarly, a combination of multiple ML techniques has reflected promising results in improving prediction accuracy and performance. A multi-stage approach integrating several models was proposed to enhance forecasting accuracy, which emphasised the potential of ensemble methods in stock price prediction (Xu, et al., 2024). Furthermore, the integration of convolutional filters with LSTM, combined with a versatile sliding window technique, has been utilised in more recent studies to refine prediction results and training performance (Wu, et al., 2023). Innovative neural networks are also currently being explored to further advance stock price predictions, such as the Hidformer Model (Liu, et al., 2024), a Transformer-style neural network (Mozaffari & Zhang, 2024) has been investigated for its potential in financial time series forecasting, which offers valuable insights in the practical application of transformer architectures in the stock price prediction domain.

2.10 Historical and Technical Data in Stock Price Prediction

Most studies on the subject matter rely on historical market prices and technical analysis to forecast the following day's market price or the direction of price movement (bullish or bearish). Market data consists of trading features that occur in a stock market, such as closing price, opening price, highs, lows and more. traders have historically used open-high-low (OHLC) candlestick charts to recognise patterns in the stock market (Velay, et al., 2018). Traders tend to perform technical analysis on stock data, which involves visual analysis of charts created over time to identify variations in price, volume and price momentum (Nesbitt, et al., 2004). Patterns such as gaps, spikes, wedges, head-and-shoulder, and triangles are derived from these visual data charts (Parracho, et al., 2010). These patterns have the ability to signal investors about the future development of the stocks. (Kim, et al., 2018) has developed a pattern matching system (PMTS) based on the Dynamics Time Wrapping (DTW) algorithm to predict the future price of KOSPI 200. The study used the DTW to grasp a matching pattern in the time series data with a known pattern. According to their results, the PMTS generates good, annualised returns. Also, the study highlights that most stocks are more profitable closer to the clearing time. (Çelik & Erkan, 2020) employ a two-stage stacking ensemble model along with empirical model decomposing (EMD) to forecast stock market direction based on the daily closing prices of six market indices. This approach discards 20% of predictions where the model underperforms, as decomposition proves to be more effective in achieving reliable predictions compared to enhancing model accuracy through feature engineering. (Zhang, et al., 2020) Train a random forest model with historical data from the Shenzhen Growth Enterprise Market (China) to predict the stock price movement (up, down, flat, unknown) and the market's growth or decline rate interval within a specific prediction duration. As per the results, the proposed model was able to deal with market volatility and outperform existing models with respect to return per trade and accuracy.

Studies that are conducted employing numeric data such as (Ou, et al., 2010), (Xu, et al., 2018), (Srijiranon, et al., 2022) prove that Artificial Neural Networks (ANN), Support Vector Machine (SVM) and Least Square SVM (LS-SVM) outperform other forecasting models such as Particle swarm optimization (PSO) and traditional neural networks. (Yang, et al., 2023) evaluate the performance of different ML algorithms on stock indexes such as S&P 500, CSI 300. This study separately executes the algorithms on stocks under transaction cost and no transaction cost. According to their results, traditional machine learning algorithms such as SVM, Naïve

Bayes, Random Forest, extreme gradient boosting and classification and regression tree (CART) perform better on directional evaluation indicators of the stocks with the transaction cost. Meanwhile, DNN architectures such as RNN, LSTM, GRU, multilayer perceptron (MLP), and deep belief network (DBN) generate better results for stock prices without the transaction cost. (Hossain, et al., 2018) Also employed two DNN architectures: LSTM and GRU. This model was trained using a considerably large time series data set of S&P 500, spanning up to 66 years (1950 – 2016). In this model, the input data was processed through the LSTM model to generate a first-level prediction, and the outcome of the LSTM model was passed through the GRU to generate the final prediction. This model was able to achieve a mean squared error (MSE) of 0.00098.

More recent studies have highlighted the importance of including external economic indicators to enhance predictive models. (Haque, et al., 2023) introduced a novel approach which incorporates anticipated macroeconomic policy changes together with historical stock prices, which demonstrated improved prediction accuracy. This emphasises the significance of considering broader economic contexts in stock price forecasting.

2.11 Sentiment Data in Stock Price Prediction

Much like other sectors, social media and news articles that reflect public opinion, rumours, speculations and insights from analysts regarding various companies and other markets highly influence the stock markets' direction (Khan, et al., 2022). Even though collecting and processing sentiment data present considerable challenges, sentiment factors such as positive, negative and neutral sentiment derived from sentiment analysis play a crucial role in the decision-making process (Jing, et al., 2021). In terms of sentiment analysis-based stock price prediction, the earliest work dates back to 1998, when a system was developed in an attempt to forecast the daily closing price of three prominent stock market indices in Asia, America and Europe by analyzing leading financial newspapers using data mining techniques, advanced keyword tuple counting and transformation. The forecasted prices were made available in real time through a website (Wuthrich, et al., 1998).

The sentiment classification is mainly done by utilising either a machine learning approach or a lexicon-based approach, further segregated into dictionary-based and corpus-based approaches (Agrawal & Bhardwaj, 2021). Although both approaches contain different strengths and weaknesses, many researchers utilised machine learning in their models since

ML is more flexible and can learn patterns more accurately than the lexicon-based approach. Factors such as prediction objectives, number of predefined classes, and time horizon for prediction influence the way news data is labelled (Nikfarjam, et al., 2010). News articles are represented by a set of features with each feature assigned a specific weight, and these weighted feature vectors are used as inputs for a classifier (Xing, et al., 2018). Support Vector Machines (SVM) is the most frequently used classification method in several studies (Fung & Glenn, 2003), (Soni & Narendrakumar, 2019). While others utilise decision trees (Das & Thoudam, 2023), where they apply two different models that include a neural network- a token-based approach and BERT (Bidirectional Encoder Representation from Transformers), which is a character-based model to forecast the next day return ratio. These studies generally indicated that models present more effective and accurate performance on news headlines compared to analysing the whole article.

Twitter is one of the primary social media platforms used for sentiment analysis due to its brevity. (Tayal & Satya, 2009). A study that implemented Stanford sentiment analysis on both Twitter and financial news headlines identified that most stock markets are highly influenced by social media and news data, which results in increased prediction accuracy. (Khan, et al., 2022). While another study used logistic regression classifiers and SVM on financial news headlines from NASDAQ and breaking news on Twitter related to 30 DIJ listed companies to predict the hourly stock price direction (Alostad & Hasan, 2015). This study also confirmed that analysis on social media news articles can influence the hourly stock price direction. (Pagolu, et al., 2016) Also employed N-gram and Word2Vec to analyse the polarity of sentiments behind the tweets. This model has gained an accuracy of around 70%, and they have found a correlation between tweet sentiment and price data of up to 71.82%. Furthermore, the study concluded that more sentiment data could further increase the accuracy of the model. Although Twitter serves as an available source of sentiment analysis, like any other social media platform, it also contains a large volume of spam content that can mislead the sentiment analysis. Therefore, scholars like (Liu, et al., 2017) Employ a blacklist dictionary to enhance real time spam tweets detection. Additionally, (Khan, et al., 2022) apply MNB classifiers to categorise tweets as spam or low-quality content.

2.12 Stock Market Prediction with Numeric and Sentiment Data

A hybrid model that combines empirical mode decomposition (EMD), principal component analysis (PCA), and long short-term memory (LSTM) to predict one step forward of the closing price of stocks in the Thailand stock market demonstrated that hybrid models produce high accuracy predictions on numeric data (Srijiranon, et al., 2022). LSTM is a type of RNN that can solve long-range dependence. Currently, RNNs have gained attention due to machine learning's constant growth in time series prediction. EMD has the ability to decompose time series data by extracting the trends of data while eliminating the nonlinear and non-stationary trends. According to the researchers (Srijiranon, et al., 2022), (Nava, et al., 2018) The application of EMD to raw data outperforms the other models. However, EMD contains the issue of mode splitting and mixing (Stallone, et al., 2020) to overcome this issue (Wu & Norden, 2009), (Yang, et al., 2023) propose ensemble empirical mode decomposition (EEMD) where white Gaussian noise is added to the original signal; While, (Hu, 2021) propose complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) where a different noise is added to the residue of the current stage, which avoids the numerical errors of EEMD. These hybrid approaches incorporate sentimental, historical, and technical data, demonstrating improved results compared to traditional methods that rely on a single data type.

In the study by (Minh, et al., 2018) propose a framework named Stock2Vec and a two-stream gated recurrent unit (TGRU network to predict the stock price direction of S&P500. This study applies sentiment analysis to financial data on Reuters financial news, employing the Harvard IV-4 sentiment dictionary. And the results indicated that both Stock2Vec and TGRU networks yield promising outcomes in analysing financial and sentiment data. Further, a study on the CSI 300 index utilising GRU along with a recurrent neural network (RNN) coupled with news data on the Chinese stock market on Sina Weibo demonstrated that the model performed well in forecasting market volatility. (Ding, et al., 2015) Propose a hybrid model that combines neural network models with sentiment analysis on the S&P 500 index. This study considers over 10 million events that occurred within 7 years. These news events were represented as vectors, and a deep convolutional network was trained to predict the temporary and lasting impact those events have on the index. Moreover, they were able to achieve an accuracy of 64.21% on the S&P 500 index. Similarly, a study on Nifty50 based on gathered data from websites, news headlines and price movement by employing LSTM together with LIME demonstrated considerable performance, together with explanation on the predictions. (Sandhya, et al., 2022).

Predicting outcomes using high-dimensional financial, technical and sentiment data can adversely affect the learning model and increase the computational time (Cao, et al., 2016). Therefore, to address this issue, researchers have applied feature extraction methods such as swarm intelligence (SI) algorithms (Kumar, et al., 2022), genetic algorithms (Abraham, et al., 2022), slap swarm algorithm (Ibrahim, et al., 2019) and principal component analysis (PCA) (Srijiranon, et al., 2022) to derive the most impactful features.

2.13 Current Direction of Sentiment Analysis in Finance

Sentiment analysis has become an increasingly prominent tool in finance, especially with the escalation of the individual investor’s involvement with the stock markets, investor sentiment is recognised as a key driver of market behaviour. Most recent studies have focused on leveraging social media platforms Twitter, Google Trends, and Reddit (Khan, et al., 2022) To gauge public sentiment to gain public sentiment for stock price predictions. However, even before the widespread use of the internet, an extensive study done on news headlines' impact on stock market dynamics revealed that world events and news data have a significant impact on stock movements. (Coelho, et al., 2019) (Pawar, et al., 2019).

More recently, a study done by analysing tweets from February – December 2008 by extracting the public mood state through Opinion Finder and the Profile of Mood States (POMS) and evaluating them through fuzzy neural networks found that the sentiments were able to predict the closing price of DIJA stocks with 87% accuracy. (Lachanski, 2016). Building on the approach, another study used a keyword classification algorithm together with Support Vector Machines (SVM), Naïve Bayes and C45 classification algorithms to predict the closing price of NASDAQ and DIJA using Twitter, where that achieved 63% accuracy. Another study on twitter data took a different approach where they downgraded tweets containing advertisements or non user contents using Stanford’s NLP processor and manually curated datasets were used to improve sentiment accuracy the accumulated data were analyzed for sentiment with n-gram and word2vec algorithms to measure 5 minutes stock changes, the results displayed that use manually labeled data produced domain specific advantage, although the it limited the study scale to 1000 tweets. (Dickinson, et al., 2015).

Further, recent studies exploring the correlation between social media sentiment and financial news and their impact on stock returns have applied machine learning and have found better accuracy. (Khan, et al., 2022) tested different machine learning methods on social media and financial news sentiment, and found that random forest produced the best results for market value predictions with 83.22% accuracy. The study on the correlation between StockTwists sentiment and stock price changes also highlighted that domain-specific sentiment analysers could significantly improve prediction accuracy (Coelho, et al., 2019). A study by (Sousa, et al., 2019) also agreed with improved predictions by fine-BERT model, which achieved a 69% correlation between financial news and DIJ movement. Several other recent studies that have applied machine learning models such as LSTM, SVM and GRU together with BERT NLP analysis have demonstrated promising results on popular stock indices (Sun, et al., 2023).

During recent times financial sector is increasingly leveraging large language models such as BERT. (Weng, et al., 2022), FinBERT and GPT-based model (Shobayo, et al., 2024) for sentiment analysis. These models have demonstrated superior performance in interpreting complicated financial texts, which resulted in improved market predictions. For example, researchers have indicated that the GPT-3-based OPT model achieved a prediction accuracy of 74.4% for stock market returns, which is a huge improvement compared to models like BERT and FinBERT (Kirtac & Germano, 2024).

In order to address the unique challenges of financial linguistics, there is a growing focus on fine-tuning LLM models with domain-specific data (Jeong, 2024). this approach enhances the model's ability to understand and analyse domain specific terminologies, leading to more accurate sentiment assessments. Studies have shown that fine-tuned models significantly outperform general-purpose LLMs and outperform their non-specialised counterparts in financial sentiment tasks. moreover, due to the scarcity of labelled financial data, researchers are currently exploring the use of synthetic data to fine-tune sentiment analysis models (Shobayo, et al., 2024). This innovative approach aims to enhance model performance by supplementing real data with AI-generated datasets, thereby addressing data limitations and improving robustness in sentiment predictions. (Lin & João Alexandre, 2024).

While sentiment has proven to be a valuable component of financial prediction models, there is a notable gap in understanding how it can add value to highly volatile meme stocks. Much of the existing literature centres around well established, traditional stock indices such as S&P 500, FTSE 100 and DIJ, where sentimental data from conventional news sources and social media sentiments re-analysed alongside price and volume indicators. Additionally, these studies do not consider the impact of sentiments on individual stocks inside the indices. Therefore, there is a noticeable gap in comprehensive studies focusing on the predictive power of sentiment analysis in retail investor-driven stocks characterised by extreme social media influence. Moreover, limited research has explored the extent to which integrating technical indicators into the prediction of such highly volatile stocks. This gap emphasises the growing requirement for a more tailored approach to utilise sentiment data for meme stocks, where retail investor sentiment plays a major role.

This study addresses this gap by using sentiment analysis on Reddit WSB and a filtering mechanism to identify daily most popular meme stocks rather than directly incorporating sentiment scores in to prediction models. This approach was followed due to practical constraints in accessing extensive Reddit data and the rapid shifts in sentiments associated with meme stocks which could have introduced noise. Therefore, this study utilized sentiment scores to identify daily most popular meme stocks and their polarity. These stock price data together with technical indicators were analyzed through variations of LSTM along with GRU model to generate accurate next day prices to provide the investors and analysts a practical insight to the short term behavior of meme stocks.

2.14 Gross Domestic Product

GDP *(2.5)* is a comprehensive measure that quantifies the total value of goods and services produced in an economy. It is considered a representation of the wealth of an economy by investors, regulators and academics, which they utilise as a tool in the decision-making process (Provost & Fawcett, 2013). The GDP of an economy can be measured based on the market value using the following equation. Where the price of product n is P and the quantity purchased is represented in Q. Here, it is assumed that only one year's worth of production is considered. This measurement is known as nominal GDP (Bhandari & Frankel, 2017).

$$GDP = (P1 * Q1) + (P2 * Q2) + \ldots + (Pn + 1) + (Pn\,Qn) \tag{2.5}$$

But changes in nominal GDP do not always represent the actual state of the economy (Ulrich, 2004) Therefore, real GDP *(2.6)* is calculated after adjusting the inflation and comparing the current year with baseline GDP. The real GDP is represented by the equation below, where

‘PB,n’ is the base year price, product ‘n’ and ‘Qc,n’ for the amount of n goods purchased in a given year. (Sa'adah & Wibowo, 2020).

$$Real\,GDP = (PB,1 * QC,1) + (PB,2 * QC,2) + \ldots + (PB,n + 1) + (PB,n * QC,n) \tag{2.6}$$

GDP is a strong indicator of the growth or decline of an economy (Provost & Fawcett, 2013). Therefore, countries tend to maximise GDP during fiscal planning to achieve better economic growth (Divya & Rama Devi, 2014). Predictions of the economic status of a country are highly valuable for the performance of organisations and policymakers, as well as for public welfare and quality of life. However, it is a complex task to predict GDP as it involves complicated calculations, while the official data are primarily released to the public after a one-quarter delay (Yoon, 2021). Hence, there is a substantial need for proper GDP prediction with currently available data.

The GDP of the United Kingdom is a critical measure of the nation’s economic activities that reflects the total value of goods and services produced within its territory over a specific period of time. As the sixth largest economy in the world, the UK’s economy highly influences both domestic and international economic policies and directions. (Hodson & Mabbett, 2009). The UK’s economy mainly consists of the highly developed service sectors that include tourism, manufacturing, finance, retail and health care, which provide significant contributions to the GDP (Hodson & Mabbett, 2009). Historically, the UK experienced fluctuations in its GDP growth due to various factors such as global market shifts, government policy changes and external shocks like the 2008 financial crisis. (McKibbin & Stoeckel, 2010), the Brexit referendum (Adam, 2019) and the most recent COVID-19 pandemic (Keogh-Brown, et al., 2020). The economic variables, such as energy prices and trade dynamics, play a substantial role in shaping the UK’s economic performance, especially due to its reliance on imports of most energy components and raw materials. Therefore, accurate prediction of the UK’s GDP is an essential requirement for policymakers and businesses to make informed decisions about resource allocation, investments and economic planning in a global interconnected market. (Hodson & Mabbett, 2009).

2.15 Macroeconomic variables and GDP

2.15.1 Energy Price Impact on GDP

Macroeconomic variables play a critical role in influencing a country's GDP by affecting production, consumption and overall economic performance. (Chaudhry, et al., 2021). In terms of energy prices, the earliest study that addresses the impact of energy prices on GDP dates back to 1983 (Hamilton, 1983), which showed a significant negative impact of oil prices on GDP. Subsequently, studies such as (Burbidge & Harrison, 1984) and (Gisser & Goodwin, 1986) Confirmed these findings. Later studies presented a decline in this impact. (Baumeister & Peersman, 2013), (Kilian, 2008) attributing this declined impact to diminishing oil expenditure shares, improved monetary policy responses, and changes in economic structure.

(Kilian, 2008) Introduced different drivers of oil price shocks, including crude oil supply shocks, aggregated demand shocks, and oil market-specific demand shocks. The analysis found that aggregated demand shocks have a more significant and longer-term effect on GDP than supply shocks. This overturned the earlier conclusion that assumed oil prices were primarily driven by supply disruptions. Other studies (Edelstein & Kilian, 2007) added to this hypothesis by showing that tax reforms and demand factors, including speculation, play a significant role in determining oil prices and their effect on GDP.

Several studies across different countries have found that energy price shocks have outsized effects on GDP relative to the share of energy expenditure. (Mork, 1989), (Wang, 2022). Whereas, some studies have highlighted the role of price volatility in reducing GDP, even for oil-exporting countries, by causing uncertainty and investment delays. (Maeda, 2007), (Nyangarika, et al., 2018). This evolving understanding of the oil-GDP relationship has led to more sophisticated methods of analysing the effect of energy prices, such as the price innovations approach. (Khiabani & Naderian, 2017), which identifies the sources of price change.

The main components of this model are:

- Crude oil supply shocks – the unexpected changes caused by changes in physical oil supply due to exogenous events such as geopolitical tensions, wars, production cuts and natural disasters. (e.g., the OPEC oil embargo in the 1970s significantly reduced oil supply. Which caused a more immediate but short-term impact on oil prices and GDP, affecting the economy primarily in the first year after the shock (Fattouh & Mahadeva, 2013)

- Aggregated demand—This fluctuation arises when global demand for goods and services increases and directly affects demand for oil. (For example, Sudden economic booms can increase demand for energy as more manufacturing and transportation take place. This is more significant and disruptive for GDP than a supply shock, which has a long-term impact.
- Oil-market-specific demand shocks—these are also known as precautionary demands and speculative shocks, where the demand for oil increases influenced to the fear of future supply shortages. These demand shocks have delayed yet significant impacts on both oil prices and GDP.

Therefore, energy prices impact GDP both negatively and positively, depending on the cause of the price movement and the broader economic environment. The balance between these effects determines the overall impact.

- The (Table 2.2) presents past literature that has studied the impact of energy prices and demand on the UK economy and each sector of the economy. As presented by these studies, energy prices mostly have an inverted U-shaped pattern relationship. As presented by most studies, UK economy appears to maintain a U-shaped relationship with energy prices, where price increases positively impact the economy up to a certain point.

Table 2.2 Relationship Between Energy and Economy

Papers	Impact on the Economy
(Lee, 2006)	Monotonic relationship
(Allan, et al., 2007)	Inverted U-shaped pattern
(Warr, et al., 2010)	Inverted U-shaped relationship
(van de Ven & Fouquet, 2017)	Inverted U-shaped relationship
(Rashid & Kocaaslan, 2013)	non-pronounced U-shaped pattern

In light of recent global events that led to significant fluctuations in energy costs, several studies have underscored the invaluable effects of rising energy prices on economic performance. According to the (OECD, 2023) 5% increase in energy prices could lead to a 0.4% reduction in the productivity of an economy within a year, This immediate negative impact is based on operational costs of business and reduced consumer spending power. Another (Ayhan & Elal, 2023) The analysed economies of OECD countries indicated that a 10% increase in energy prices has the potential to suppress economic growth by approximately 0.15%. This effect varies across economies based on their dependencies on energy and economic structure. In terms of the sector impact of energy price fluctuations, according to the (European Central Bank, 2023) escalating energy prices have considerably constrained consumer spending by affecting real incomes. This situation has led to a shift in consumption patterns with households reducing expenditure on non-essential goods and services. In summary, recent studies illustrate the complex dynamics between energy prices and GDP. Although energy price fluctuations can pose challenges to economic growth, strategic investments in clean energy and improvements in energy efficiency contribute to pathways of mitigating these effects.

2.15.2 Unemployment Rate

The unemployment rate is one of the most visible indicators of economic health which plays a critical role in determining GDP (Ülker & Ülker, 2019). Since high unemployment rates typically lead to lower household incomes, reduced consumer spending and decreased aggregate demand, all of which negatively impact GDP growth. (Rahman, 2013). In the UK, unemployment peaked during the 2007-2008 global financial crisis, resulting in the unemployment rate reaching an all-time high of 8.5% by the year 2011. This period demonstrated significant contractions in consumer spending, especially in retail and service sectors that are highly dependent on household consumption (Edey, 2009). On the other hand, in 2019, the UK's unemployment rate reached a historically low level of 3.8%, which supported strong GDP growth powered by higher consumer confidence and spending. (Simionescu, et al., 2020). However, this growth was disrupted by the unexpected COVID-19 Pandemic, with all economic sectors being disproportionately affected. (Keogh-Brown, et al., 2020). Therefore, the unemployment rate and GDP performance appear to have an inverse relationship.

Okun's Law (Okun, 1962) clearly describes this empirical relationship between the unemployment rate and economic growth. The law suggests that when an economy experiences faster GDP growth, the unemployment rate tends to decline. And when the GDP growth is slow and negative, the unemployment rate increases. Specifically, this study explains that every 1% increase in the unemployment rate roughly decreases GDP by 2-3%. This theory suggests that economic growth drives job creation, while economic contraction leads to job losses. Although in practice, when an economy reaches its potential, the unemployment rate tends to remain stable (Grainca, 2022). However, rapid growth is essential for the reduction of the unemployment rate. While the relationship between GDP and unemployment has evolved over the years due to factors such as labour market changes and productivity improvement, Okun's Law remains a fundamental concept of macroeconomic analysis.

2.15.3 Real Effective Exchange Rate (REER)

REER *(*2.7*)* is a measure of a currency value against a weighted average of several foreign currencies adjusted for inflation differences between the home country and its trading partners (Chinn, 2006). This shows the competitiveness of a domestic currency against its major international trading currencies. An increase in REER indicates a decline in competitiveness in international trade, where exports become expensive and imports become cheaper.

$$REER = NEER \times \frac{CPI^*}{CPI} \quad 2.7$$

REER – Real Effective Exchange Rate, NEER – Nominal Effective Exchange Rate (A weighted average of the bilateral exchange rates of the home country relative to those of its trading partners), CPI – Consumer Price Index of the home country (used as a proxy for domestic inflation), CPI^* - Weighted average of the consumer price index of the home country's trading partners (used as a proxy for foreign inflation)

$$NEER = \prod_{i=1}^{n} \left(\frac{E_i}{E_{i0}}\right) w_i \quad 2.8$$

NEER – Nominal Effective Exchange Rate, E_i – Exchange rate of the domestic currency relative to trading partner I (expressed as units of the foreign currency per unit of the domestic currency), E_{i0} – base period exchange rate, W_i – Trade weight of country i, n – Number of trading partners considered

according to the prior literature on REER impact on GDP growth, some studies (Comunale, 2017), (McLeod & Mileva, 2011) Suggested that a weaker REER can positively influence the GDP growth by boosting the traded goods sector. Since a weak REER supports 'learning-by-doing-effects' where it helps compensate for institutional weaknesses and market failures in developing countries by promoting investments in tradable sectors (Rodrik, 2008). According to the 'Bertrand Scale', REER depreciation can have a long-lasting positive effect on GDP; according to this scale, firms compete based on price and increasing return to scale, where the production process becomes more efficient as output increases. (Di Nino, et al., 2011)which results in increased exports and economic growth. This was practically demonstrated through Italy's economic growth from 1861-2011 (Di Nino, et al., 2011). Several studies agree that monetary undervaluation of the exchange rate supports economic growth, especially for developing countries. (Hausmann, et al., 2005), (Rodrik, 2008), (Ribeiro, et al., 2020).

On the contrary, most recent studies argue that undervaluation doesn't always promote economic growth, especially for developing countries, particularly when episodes of overvaluation are excluded. (Kappler, et al., 2011), (Nouira & Sekkat, 2012) Agree that large currency appreciations have a limited negative impact on both developing and advanced economies, especially when the appreciation depends on factors such as productivity and capital inflow. (Bussière, et al., 2015).

When considering the REER of the UK, it significantly impacts the UK's Trade balance, which is a key component of GDP (Ditta, et al., 2021) A strong REER has the potential to make exports more expensive for the international markets while making imports cheaper, potentially leading to a trade deficit. As an example, after the Brexit referendum in 2016, the value of the British pound fell sharply, which resulted in a lower REER. While this temporarily caused a negative impact by increasing import costs, it made UK exports more competitive in the global markets, especially in sectors like manufacturing and financial services (Wheeler, et al., 2008). Conversely, when REER strengthened during the early 2000s, UK exports faced increased challenges as the goods became more expensive to overseas consumers, which resulted in reduced export volume and a consequent drop in GDP.

2.15.4 Inflation

Inflation is known as the sustained increase in the general price levels of goods and services in an economy over a certain period, where each unit of currency reduces consumers' purchasing power. (Wei & Xie, 2022). Inflation is commonly measured through indices such as the consumer price index (CPI) or the producer price index (PPI). These indices track changes in the prices of a basket of goods and services over time. (Wei & Xie, 2022).

Since 1990, a significant number of studies have been conducted to identify the impact of inflation on a country's GDP using cross-country data and time series analysis, revealing disparities between nations. (Arai, et al., 2004). From a theoretical perspective, economists argue that moderate inflation can benefit economic growth. (Kirshner, 2001). However, monetarists claim that it has detrimental effects. (Cooley, et al., 1991). Therefore, the exact nature of the inflation-GDP relationship is still unclear. To identify this relationship, various statistical methods have been applied, such as the unit root test. (Lee & Chang, 2008), causality test (Korkmaz, 2019) and cointegration analysis (Manning & Andrianacos, 1993). However, recent studies suggest that there is a threshold inflation rate, and beyond this level, the relationship between inflation and GDP varies. Below the threshold, inflation most likely has a neutral or even positive impact on GDP growth, while above the threshold, inflation begins to negatively impact the economy. (Arai, et al., 2004). This threshold may vary depending on the country; for example, in Pakistan, it has been found to be 9% (Mubarik & Riazuddin, 2005), in Jordan, 2% (Sweidan, 2004) and in Taiwan and Japan, it is 7.25% and 2.52%, respectively. (Lee & Wong, 2005).

In the UK, the Bank of England closely monitors inflation to ensure it remains at the target level, usually around 2% (Joseph, et al., 2024). When inflation rises beyond this limit, it erodes consumer purchasing power, increases business costs and generally slows down the economy. As an example, during the 1970s oil crisis, UK inflation surged over 25%, resulting in a significant drop in economic growth and requiring drastic policy measures (Nelson & Nikolov, 2003). And more recently, due to the COVID-19 pandemic in 2022, the UK experienced major inflation levels exceeding 10%, which was driven by the post-pandemic supply chain issues and energy price spikes. This inflation rise caused lower consumption, reduced real income and a rise in interest rates (Victor, et al., 2021). Therefore, uncontrolled high inflation has proven to reduce investor and consumer confidence, which negatively impacts GDP growth (Mankiw, 2019).

2.15.5 Net Migration

Net migration is defined as the difference between the number of people entering a country (immigrants) and the number of people exiting a country (emigrants) during a specific period, which impacts the population size and composition of a country. (United Nations Department of Economic and Social Affairs (UN DESA), 2019). According to the (United Nations Development Program (UNDP), 2020), even though migrants only make up to 3.5% of the global population, they contribute to 9% of the world's GDP. Studies conducted by the (IMF, 2024), (World Bank Group, 2024) Have projected that increasing immigration in developed countries by 3% could boost global GDP by USD 365 billion by 2025. Several studies have been carried out on the impact of migration on GDP and other macroeconomic indicators. As an example, the study done by (Kudaeva & Redozubov, 2021) States that a 1% increase in migration can lead to a 0.01% increase in real GDP. According to prior literature, there are three main categories of migration-

- Labour migration – impacts industries such as manufacturing.
- Educational migration – influences scientific and technological sectors.
- Refugee migration inversely impacts the economy due to social support needs for refugees.

There is a complex, multifactorial relationship between migration and the labour market, consumer demand, and public services, which eventually influences the GDP (Rayevnyeva, et al., 2023). Positive net migration can increase the workforce, boost consumer demand and fill skill gaps in key industries such as healthcare, construction and services (Rayevnyeva, et al., 2023). In the UK, net migration played a significant role during the economic growth in the early 2000s, as an influx of workers from the EU contributed to a larger, flexible labour market. However, following Brexit in 2016, strict immigration policies were introduced, reducing migration from the EU. This resulted in a labour shortage in critical sectors, which put upward pressure on wages and limited productivity. Therefore, Brexit, in turn, slowed the economic output (Outhwaite & Menjívar, 2019). Hence, net migration is a crucial component in maintaining a proper workforce and sustainable growth. This is specifically important to the UK with an ageing population, since the labour shortage can constrain GDP expansion (Borjas, 2014).

2.16 Traditional GDP Prediction Methods

GDP forecasting methods can be categorised as quantitative forecasting and qualitative forecasting. Qualitative forecasting mainly focuses on experts' judgements and opinions. It is mainly used when historical data is unavailable. (Maccarrone, et al., 2021). Business and consumer survey data (Claveria, 2021), Delphi method (Abreu & Mesias, 2020) Panel consensus forecasting (Anesti, et al., 2021) These are some of the most commonly used qualitative methods in qualitative economic forecasting.

In terms of quantitative methods, traditionally, statistical and econometric models have been used for GDP predictions. These models typically depend on fitting data to a predetermined relationship between input variables (indicators) and the output (target) variable. These models implicitly assume a stochastic process underlying the relationship between the target and the indicators. The most commonly used models are ARIMA. (Ho & M. Xie, 2021), Vector Autoregression (VAR) (Roush, et al., 2017) Structural Equation Modelling (SEM) (Chien & Hu, 2008), and Partial Least Squares (PLS) (Ahmmed, et al., 2021).

ARIMA model (Box & Jenkins, 1976) predicts GDP by analysing past data trends through a combination of autoregressive (AR) terms, Differencing (I) to achieve stationarity, and Moving Average (MA) to represent past forecast error. The model is denoted as ARIMA (p, d, q) *(2.9),* where p is the number of autoregressive terms, d is the number of differences needed for stationarity, and q is the number of lagged forecast errors in the prediction. This method has become a standardised approach for time series forecasting, guiding researchers through model identification, parameter estimation and diagnostics checks. The earliest study in the application of the ARIMA model in economic and financial forecasting dates back to 1991 (Pindyck & Rubinfeld, 1991), where they illustrate how the model can effectively capture inherent patterns in economic data and provide short-term forecasting for macroeconomic variables like GDP, Inflation and unemployment. Since then, several studies have utilised the ARIMA model for GDP forecasting (Yang, et al., 2016), (Ho & M. Xie, 2021) However, the ARIMA model contains certain constraints, such as the model's assumption of linearity, which can restrict the non-linear patterns in data, and the models struggle with long-term forecasts, especially when the data exhibit volatility and seasonality (Atif, 2024).

ARIMA model for time series data-

$$y_t = c + \emptyset_1 y_{t-1} + \emptyset_2 y_{t-2} + \ldots + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \ldots + \theta_q e_{t-q} + e_t \tag{2.9}$$

y_t - Value at time t, $\emptyset_i$ – Autoregressive Parameters, θ_i – Moving Average Parameter, e_t – Error term (white noise) at time t, c - Constant

The Vector Autoregression (VAR) model (Sims & A, 1980) It is another statistical model that measures the linear interdependencies among several time series variables to produce comprehensive macroeconomic forecasting. (Enders & Jones, 2016). This model treats every variable in the system as endogenous, meaning each variable is a function of its past values and the past values of all other variables in the system. Moreover, the VAR use the variables' lags to predict the target variable; as an example, in two-variable systems with GDP and inflation, each variable is explained by its past values as well as the past values of the other variable. (Forni, et al., 2023). This method is known for its simplicity and flexibility in capturing dynamic relationships among several features that are more accurate in short-term predictions with a substantial dataset. (J. Roush & Hu, 2017).

Structural Equation Modelling (SEM) is another popular traditional method that provides insight into the complex structural dynamics of GDP. The SEM model combines factor analysis and multiple regression. (Alam, et al., 2019) To generate predictions. This statistical model differs from prior models as it can simultaneously identify complex direct and indirect relationships between observed variables and target variables through multiple interrelated regressions. This model has displayed strengths in providing insight into the causal pathways and relationships between variables. Therefore, this model stands out from others as it allows for rich interpretations in complex economic structures. However, due to its high complexity, the model is susceptible to extreme specifications that make it nearly impossible to interpret without deep expertise. (Spulbar, et al., 2021).

Partial Least Squares (PLS) is another statistical model (Assunção & Fernandes, 2024) that simultaneously decompose the predictor variables and response variables to find linear connections, also known as latent variables, that can heighten the covariance between the predictors and the response, allowing the derived variables to be relevant for predicting the response variable. This model is highly utilised when traditional regression methods fail due to multicollinearity or high data dimensionality. PLS helps identify key latent variables that drive the response variable, in this case, the GDP growth, through dimensionality reduction.

Other than the above methods, many prior studies have utilised different versions of statistical modelling for GDP prediction, such as Seasonal ARIMA (SARIMA) (Abonazel & Abd-Elftah, 2019) Bayesian Vector Autoregression (BVAR) (Karagöz & Keskin, 2016) and Exponential Smoothing model (ETS) (Jafari-Samimi, et al., 2007). Although these traditional methods produce a robust framework for GDP forecasting, they tend to portray limitations in handling large, complex datasets and non-linear relationships. For example, a study (Bahelka & Weerd, 2024) Evaluated ARIMA along with VAR and Mixed Data Sampling (MIDAS) models for forecasting the GDP of Madagascar, which highlighted that ARIMA has better capabilities in capturing temporal dependencies. Whereas the study (Oancea, 2024) compared SARIMA with LSTM networks to forecast Romania's GDP, and according to the findings, LSTM models offered superior predictive accuracy.

2.17 Machine Learning in Economic Predictions

In recent times, a considerable number of studies have leaned towards machine learning methods for macro and micro econometric forecasting. (Granger & Engle, 1987) Since machine learning algorithms take a different approach to forecasting, they make minimal assumptions about the fundamental relationship between variables. Instead, they depend on algorithms to discover a function that best explains the relationship between input and target variables. While machine learning has been used to a certain extent for economic forecasting in the past (Swanson & White, 1997) (Teräsvirta, et al., 2005), (Stock, et al., 2002) It is only recently that a substantial number of studies have applied these methods for macroeconomic forecasting.

Machine learning models such as decision trees and random forests excel at capturing complex, non-linear relationships in large datasets, which makes them highly effective in forecasting economic variables like stock price prediction and GDP growth. (Maccarrone, et al., 2021), and studies such as (Li & Chen, 2014) has revealed that use of methods like lasso regression and Elastic Net allows for better feature selection and interpretability, making them highly effective in economic forecasting by reducing the black box nature of the machine learning models.

Random forests (Yoon, 2021) and Decision Trees (Muchisha, et al., 2021) They are among the most widely used machine learning methods for economic forecasting due to their ability to model non-linear relationships through various data. Hence, models like random forests have shown exceptional performance. (Pérez-Pons, et al., 2021) Beyond traditional models like ARIMA or VAR, when forecasting high-frequency financial data. These models reduce variance in data and improve prediction accuracy by creating an ensemble of decision trees that are especially useful for dimensional data with many predictors. (Maccarrone, et al., 2021).

Support Vector Machines (SVM) is another machine learning model that was created for classification problems but has been adopted by several studies for regression tasks in economic forecasting. (Cao & Tay, 2021).Moreover, many studies have demonstrated that SVM can outperform linear models in capturing non-linear patterns in economic data (Hu & Wang, 2022). SVMs are particularly useful when data has a complex structure, as they create hyperplanes to maximise the margin between different data points to enhance prediction accuracy. (Tian, et al., 2012).

Artificial Neural Networks (ANNs) are also increasingly common in econometric forecasting as they are able to model complex non-linear relationships by following the human brain's neural structure. (Zhang & Yu, 1998). As an example (Shams, et al., 2024) Use the PC-LSTM-RNN model, which combines Principal Component Analysis (PCA), Long Short-term Memory (LSTM), and Recurrent Neural Networks (RNNs) to capture the long-term dependencies among variables. Another widely used model is Bayesian Vector Autoregression (BVAR)

(Yoo, 2023), (Sílvia, et al., 2019), which enhances traditional VAR models by incorporating Bayesian priors to address overfitting and improve forecast accuracy. Additionally, the XGboost algorithm also portrayed promising results in prior studies, as it can handle large datasets and improve prediction accuracy through advanced regularisation techniques and scalable implementation of gradient boosting. (Chu & Qureshi, 2023), (Bharathi & Navaprakash, 2024). However, even if neural networks can be powerful and accurate in identifying non-linear relationships among various variables, they require careful tuning and validation to avoid overfitting, especially when used with smaller datasets.

Further, machine learning models face significant criticism for being 'Black Boxes', meaning their lack of interpretability makes it challenging to understand the underlying factors that are driving their predictions (Buckmann, et al., 2021). Hence, even though machine learning models have proven to outperform traditional models, traditional models such as ARIMA and VAR still dominate the economic forecasting domain in practice.

These recent studies that compare machine learning methods with conventional methods for GDP prediction have revealed notable differences in performance and stability. (Martin, 2019). According to their findings, ANNs produce better performance in capturing the nonlinear relationships, especially when they are combined with other models like Support Vector Machines and decision trees. (Jena, et al., 2021), even though they may overfit to smaller datasets. Regardless, LSTM has demonstrated superior accuracy in time series GDP forecasting, outperforming traditional SARIMA through effective modelling of long-term dependencies. As an example, a hybrid-RNN-LSTM model that utilised transfer learning indicated outstanding performance in the Gambia with an R^2 score of a staggering 91.28% (Jallow, et al., 2025). Deep learning models in multi-country settings also appeared to outperform linear regression when economic indicators were included, making the models ideal for cross-country analysis. (Subian, et al., 2024). In terms of real-time GDP nowcasting, machine learning models such as RF and elastic nets exceeded traditional forecasting methods in predictive accuracy, offering strong promise for immediate policy use (Richardson, et al., 2021) Even though challenges remain in data consistency, overall, machine learning model almost always seem to outperform conventional methods.

This study addresses notable gaps in previous studies that employ machine learning for GDP prediction by incorporating a substantial set of macroeconomic variables concerning the UK. These variables are often analysed in isolation without considering their combined impact on the GDP (Ali Shah, et al., 2013). Additionally, prior studies have frequently focused on a limited set of input variables or utilised traditional econometric models that have faced difficulties in capturing the complex, non-linear relationships inherent to econometric data. (Stock, et al., 2002), (Andreas, et al., 2008), (Teräsvirta, 2005) (Afrouzi & Bhattarai, 2023). This study leverages the abilities of advanced machine learning models such as SVR, RF and GBM together with Linear regression as a baseline model to uncover intricate patterns among the variables while providing comparative analysis on the models' performance.

By utilising UK-specific data, it is expected to address the contextual gaps in existing generalised studies. (Li, et al., 2021) These generalised studies often failed to account for the UK's unique economic conditions and peculiar traits. In contrast, this study explored the dynamic relationships between several critical macroeconomic indicators, focused only on the UK. The indicators consist of variables such as energy prices, inflation, unemployment rates, and REER. Additionally, this study evaluates interpretability by providing insight into the factors that hold more predictive power for GDP predictions.

To further enhance the interpretability of the machine learning models, SHAP was applied to each version of the models. This application helped determine the factors that influenced the predictions across different models. Further, it aided in breaking down the contribution of each variable to the predictions (Kim, et al., 2024). Thus, this study ensured transparency and a detailed understanding of each variable's contribution to the UK's GDP.

Thus, by incorporating machine learning techniques that are capable of capturing both linear and non-linear relationships, this study contributes to the growing body of literature with robust, data-driven insights into the interconnectedness of macroeconomic indicators and their impact on GDP. Moreover, it advances the field by offering an in-depth comparison of model performance and interpretability that not only enhances prediction accuracy but also provides practical implications for policymakers, economists and businesses that are looking to harness machine learning for economic forecasting.

3 CHAPTER THREE

Sentiment Driven Meme Stock Price Prediction: Analysing WallStreetBets Discussion to Forecast Stock Prices with Recurrent Neural Networks

3.1 Introduction

This chapter explores the use of sentiment analysis and recurrent neural networks (RNNs) for forecasting meme stock prices with a focus on the influence of online discussions of Reddit's WallStreetBets (WSB) subreddit. Meme stocks, known for their volatility and susceptibility to retail investor sentiment, require models that account for both technical and behavioural market signals. To capture this unique behaviour of meme stocks, sentiment data was retrieved using the Python Reddit API Wrapper (PRAW) and analysed using the VADER lexicon to generate a daily sentiment score to identify the most discussed stocks. Stock prices and technical indicators of these popular meme stocks were used to train three variants of LSTM and GRU. This chapter evaluates the predictive power of each model and assesses how effectively sentiment-driven inputs can enhance next-day price forecasting for socially influenced stocks.

3.2 Methodology Overview

This methodology (Figure 3.1) aims to produce a robust model for investors and analysts to easily identify popular meme stocks and navigate through their volatility effectively after proper consideration without solely relying on social media trends.

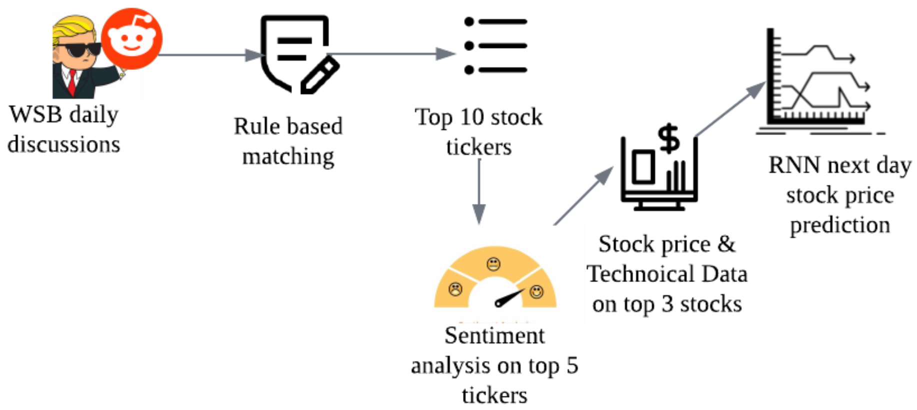


Figure 3.1 Methodology Meme Stock Price prediction

As depicted by the Figure 3.1Sentimental data was retrieved from WallStreetBets subreddit discussion threads using direct Python Reddit API Wrapper (PRAW), followed by rule-based matching, detecting and filtering the most discussed stock tickers based on predefined rules and criteria. Secondly, five of the most popular meme stock tickers were identified based on the number of mentions. A natural language processing tool known as VADER Lexicon was applied to analyse the sentiment of the currently available discussions and assign a sentiment score to the detected top five stocks. Subsequently, historical stock price data were extracted from Yahoo Finance, and technical indicators were generated based on the historical data for the top three stocks. Conclusively, historical and technical stock data were utilised to develop next-day closing price prediction machine learning models.

For model training top three meme stocks were selected. The dataset was split into 80% training and 20% testing, and all features were scaled using MinMaxScaler. Three versions of the LSTM RNN architecture and GRU were tested to observe how the predictions differ with model complexity. The models were implemented using Keras and TensorFlow. Experimental setup – epochs: 50, Batch size: 32, Loss function: Mean Squared Error (MSE), Optimisation: Adam. The validation strategy was to implement early stopping on validation loss with a sequence lookback window of 10 days. This experimental setup ensured a fair comparison of different versions of LSTM and GRU architectures in predicting the next day closing price of meme stocks using both sentiment and technical data.

Sentiment analysis played a crucial role in the first part of this study, which focused on identifying the most popular meme stocks on the Reddit WSB platform. The decision to incorporate sentiment analysis purely stemmed from the unique and unconventional behaviour of meme stocks that are heavily influenced by public sentiment. Unlike traditional stocks, which primarily respond to financial indicators and macroeconomic events. As explained in the previous chapter, meme stocks tend to exhibit significant price fluctuations solely based on viral discussions, hype and collective retail investor behaviour.

Due to resource constraints and smaller datasets of meme stocks that are mostly novel to the stock market, the study focused on four RNN architectures, namely single-layered LSTM, regular stacked LSTM, Bidirectional LSTM and GRU. These models were selected to evaluate how different levels of depth and directional memory affected predictive accuracy while maintaining computational feasibility.

Although sentiment analysis was emphasised in the context of meme stocks, the proposed forecasting framework is adaptable to conventional stocks as well. For traditional stocks, the model can operate solely on historical price and technical data without relying on the sentiment feature. In such a scenario, the sentiment feature can be omitted, but the models can still generate meaningful predictions based on market trends.

This flexibility in application indicates that this forecasting architecture is not limited to meme stocks. Any stock, regardless of its popularity among online communities or volatility, can be analysed using the same architecture. Therefore, the inclusion of sentiment analysis serves as an additional filtration layer of insight for sentiment-driven assets, but not a compulsory requirement for the core function of the prediction models.

3.3 Research Paradigm

This study is grounded on the positivist research paradigm, which insists on the existence of objective reality and that it can be measured and quantified through scientific and empirical methods (Crook & Garratt, 2005). The ontological stance of this study is realism since this study utilises real stock prices and sentiments expressed on the WallStreetBets forum that are caused by actual phenomena that exist independently of the observer. Meaning that the input variables of this study can be measured and analysed objectively. As for the epistemological stance, this study has adopted objectivism since the information about the relationship between posts and comments on WSB and stock price data was obtained through empirical observations and analysis. The outcome of the study relies on objective data collection and analysis methods to discover patterns and relationships. Further, this study identifies social media sentiments and stock market data as quantifiable variables that can depict insights into meme stock behaviours.

3.4 Theoretical Framework

This study is primarily based on behavioural finance theory, which explains that market prices are not entirely dependent on fundamental factors such as dividends, interest rates and earnings but are also significantly influenced by the sentiments and behaviours of the investors (Bellofatto, et al., 2018). Unlike traditional finance, behavioural finance identifies that investors are not always rational and make decisions based on all available information. Instead, it recognises that investors are humans who make decisions influenced by social influences, emotions and cognitive biases, which can lead to irrational financial decisions that result in market anomalies (Coelho, et al., 2019). This theory is reflected through meme stocks, where collective behaviours and communications of individual investors on social media platforms, particularly WSB, can lead to significant price fluctuations that demystify conventional financial logic.

Additionally, this study combines behavioural finance theory with machine learning frameworks to produce a price prediction model that is able to analyse and forecast stock prices based on historical and technical data of the relevant meme stock. Recurrent Neural Networks (RNNs) are considered a powerful tool for modelling and predicting complex time series data. (Pawar, et al., 2019). Therefore, a combination of behavioural finance with machine learning concepts enables the identification of temporal dependencies in sequential stock data, facilitating ideal analysis of meme stock price movement and temporal patterns in social media sentiment.

3.5 Data Mining and Extraction

3.5.1 WSB Data Extraction

This study utilised two primary data sources, namely Reddit WSB sentiment data and historical financial values of the stocks, to create a link between stock price data and social media sentiments while identifying the effect of sentiments on meme stocks.

The sentiment data of Reddit WSB were extracted using Python Reddit API Wrapper (PRAW), which provided a convenient platform to interact with Reddit's Application Programming Interface (API). Although PRAW provide convenient access to Reddit data, it has certain limitations, such as 60 requests per minute, and it does not provide access to extensive historical data. Moreover, WSB is a highly active subreddit with thousands of daily posts, especially threads like 'Daily Discussion' where users discuss trending market trends and share personal stories and memes (Nobanee, et al., 2023). Therefore, posts and comments were extracted, focusing on the likelihood of containing important sentimental information relating to meme stock trading. The main post flairs targeted were 'Daily Discussion', 'Weekend Discussion' and 'Discussion', which are known to attract considerable user engagement. To maintain the quality and relevance, extraction parameters were set as a minimum upvote ratio of 0.70 and a minimum number of upvotes with 20 upvotes for a post to be considered. Further, the extraction was set to automatically include submissions of a set of trusted authors according to Reddit guidelines, using the 'AutoModerator' variable. The extraction process includes navigating through 'hot' posts, collecting post titles and unique comments and compiling the extracted information into a structured format.

3.5.2 Historical Stock Data Extraction

Historical data for the most discussed meme stocks were extracted from Yahoo Finance using the 'yfinance' Python package. The data were extracted based on the day the relevant stock was issued for public trading, since most meme stocks are fairly recent. For the stocks that have been in the market for a longer period, the data were extracted for a period of 10 years spanning from 2014 to 2024, as it provides a comprehensive timeline that includes various market cycles, economic conditions and important events that impact the direction of stock prices. Further, this period encapsulates the rise of retail investing, assisted by meme stocks driven through social media and online trading platforms like Robinhood. Further, for specific stocks, this period includes the impact of the COVID-19 pandemic on the stock market. Hence, this period addresses the trends, volatility and impact of external factors on meme stock performance. The dataset consists of the 'Date', 'Open', 'High', 'Close', 'Adj Close' and 'Volume' data for each meme stock on separate occasions.

3.6 Data Preparation

3.6.1 R/WallStreetBets Data

The extracted data from WSB, focusing on popular flairs with high upvote ratios, were preprocessed to extract the collective sentiment expressed by individual investors. The initially collected posts and comments were filtered based on keywords and currently available 7054 stock ticker symbols on the NASDAQ stock screener (e.g., GME, AMC, AAPL).

The row text data extracted from WSB posts did not require substantial pre-processing to convert it into a suitable form for sentiment analysis, since the sentiment analysis was carried out using Valence Aware Dictionary and Sentiment Reasoning (VADER), a case-sensitive sentiment analysis tool that can differentiate between uppercase and lowercase text. VADER can interpret uppercase words as having more intensity and emphasis; for example, the word 'good' is identified as a positive sentiment while 'GOOD' could be interpreted as strongly positive due to the emphasis implied by the uppercasing. (Elbagir & Yang, 2019).

First, the posts were sorted based on post 'flair'. In Reddit, 'flair' is a tag that users or moderators can attach to posts to categorise the posts and provide additional context to the content. Some of the most popular flairs used in financial stocks-related subreddits are – 'Daily Discussion', 'Discussion', 'News', 'Analysis' and 'Meme'. It can be identified as a way of highlighting the importance of the posts. Therefore, for this study, posts flared as 'Discussion', 'Daily Discussion' and 'Weekend Discussion' were given priority.

- Daily Discussion—This is most likely part of the regular and continuous discussion that focuses on current events and daily stock market activities.
- Weekend Discussion – These posts cover broader discussion that reflects on the discussions during the week. Possibly, including discussions on investment strategies and week-in-review.
- Discussion – posts flaired as discussion cover topics related to stocks that are currently under analysis.

This filtering helps ensure the avoidance of 'Meme' flaired posts that mostly represent casual chat without relevant serious analysis. However, the filtration was set to also consider posts with no flairs, as not all posts get flaired. Therefore, it ensures that potentially important discussions are not missed out simply due to a lack of flair.

The posts were further filtered based on their popularity within the WSB subreddit using the 'hot' posts tag. They were filled to consider posts with a 70% or above upvote ratio. The upvote ratio (3.1) is the percentage of upvotes compared to the total number of both upvotes and downvotes a post receives.

$$Upvote\ Ratio = \frac{Number\ of\ Upvotes}{Number\ of\ upvotes\ +\ Number\ of\ Downvotes} \tag{3.1}$$

This iteration was included as it provides insight into the community's perception and interest in a post. A higher upvote ratio indicates that the post is more positively received and that most of the individual investors in the community share a similar interest in the relevant matter discussed in the post.

Further, the posts were filtered to consider only the ones that receive 20 ups or above. This filtration was included as it is a direct measure of how many Redditors identify the content as valuable and agreeable. As for an example, a post with high upvote ratio (e.g.- 80%) and high upvotes (e.g.- 10,000 ups) is most likely a popular topic and has little controversy. This provides a clear sense of the community's perception of the post.

From the above filtered posts, comments were extracted and sorted based on their newness. Just as with any other social media platform, comments on Reddit posts can potentially include hundreds and thousands of comments. For efficiency, Reddit only displays a small portion of the comments, and the rest of the comments are represented with 'MoreComments' as a placeholder for deeper comments. Therefore, the replace_more ()' method was used to replace the 'MoreComments' object holder with actual comments. For this study, the 'MoreComments' object was slightly expanded by expanding the limit by 10 to include 10 more objects, but not fully expanding the thread. Moreover, these comments were filtered by a minimum upvote score of two, and the comments below this threshold were disregarded from the analysis. Further, the comments were checked for the author's name, if the name is available, in order to filter out the deleted accounts and ensure unique comments by unique authors.

The next step in preprocessing was to identify stock ticker symbols accurately. On WSB, most users prefix stock ticker symbols with a '$' sign to clarify to other users that they are referring to a stock. Therefore, to standardise the textual data, the '$' sign was disregarded from words to avoid the sentiment analysis tool wrongfully identifying the '$' prefixed ticker symbol and the core ticker symbol as different stocks (e.g., '$GME' and 'GME' are the same). This helped standardise stock ticker symbols, which ensured consistency and accuracy.

Further, comments containing the relevant ticker symbol were associated with the ticker to help with the sentiment analysis. The author of the comments was tracked using a separate dictionary named 'cmt_auth'. Unique author tracking was enabled using the 'uniquecmt' label. This ensured that redundant data was avoided, as this label permitted the model to only consider the first comment of each author about a single stock ticker. As the final step, the acquired data were organised into separate dictionaries. These dictionaries consisted of important variables for the analysis, such as tickers, post titles, and comments.

3.6.2 Historical Price Data and Technical Data

Historical data were sourced through the Yahoo Finance website via Python, the Yfinance library for each top three popular meme stocks identified during the duration of the study. Particularly for older stocks, historical price data for the last ten years were collected since Meme stocks gained popularity during this period, especially in 2020 (Nobanee, et al., 2023).

Therefore, this time period contains unique characteristics of this phenomenon. Further, this period captures the before, during and after effects of the COVID-19 pandemic (Aloosh, et al., 2021), which provided valuable insight into the external shocks and recovery of meme stocks. Also, due to the lockdown, retail investors surged into the market through online trading platforms during this period (Aloosh, et al., 2021). Most importantly, during this significant period, several high-volatility events occurred, including the meme stock frenzy in 2021, where GameStop (GME) and AMC experienced an unprecedented surge in stock price (Aloosh, et al., 2021). Therefore, training the model with data from this period closely aligns with the current dynamics of meme stocks, ensuring that the predictions are relevant to the current market conditions.

The dataset gathered from Yahoo Finance consisted of 'Date', 'Open', 'High', 'Close', 'Adjusted Close' and 'Volume'. The 'Adjusted Close' price represents the closing price adjusted to corporate actions, such as stock splits or dividend distributions, to present a more accurate reflection of the true value of the stock. However, for this study, the 'Close' price was selected as the target variable for the prediction since it represents the last transaction price of the stock during the trading session, without historical market adjustments, presenting the real-time market value (Hossain, et al., 2018). Consequently, the 'Adjusted close' price was excluded from the predictors.

3.6.2.1 Feature Engineering

To enhance the meme stock price prediction model, technical indicators were derived from the historical data since the inherent features of the original stock price data were insufficient to construct a robust model, especially in the highly volatile meme stock environment (Yahya, et al., 2021). Technical indicators can provide valuable insight into stock data in terms of market trends, volatility, and market momentum that can guide the decision-making process, improving prediction accuracy (Vargas, et al., 2018).

Based on the historical price data, 13 additional technical indicators were created to enhance the dataset. The technical indicators consist of Moving Average Convergence/Divergence (MACD), Price Change Percentage Rate, and Exponential Moving Average (EMA).

EMA's are the averages of stock prices over a specific period, with more weight to recent years, making the indicator more responsive to recent price changes (Bandi, et al., 2021). MACD was generated based on the EMA lines by measuring the difference between the long-term EMA and from short-term EMA. This study used the default standard setting to set the period as a longer-term EMA (26 days) and a short-term EMA (12 days). The distance between these two EMA lines was identified as the MACD line. The MACD line helps identify the stock market trend (Oriani & Coelho, 2016). As an example, the stock is gaining upward momentum if the MACD line is above zero and if the MACD line is below 0, the stock is facing a downward trend.

The 9-day EMA line plotted on top of the MACD line is known as the signal line, as it helps identify buy or sell signals (Vargas, et al., 2018). If the MACD line crosses above the signal line, it is typically identified as a bullish or a sell signal, while the MACD line crossing below the signal line is identified as a bearish or a sell signal (Vargas, et al., 2018). Therefore, these technical indicators help the model spot changes in the direction, trend and strength of the meme stock prices, aiding informed decisions (Oriani & Coelho, 2016).

3.6.2.2 Feature Selection

With the combination of stock price data and technical indicators, the entire data set consisted of – Open, High, Low, Close, Adjusted Close, Volume, Percentage change, Short EMA, Long EMA, MACD, Signal Line, Log Return, Volatility, Standard Deviation, Upper Band, Middle Band Lower Band and Next day Price. To reduce the negative impact of redundant features that drive models to overfitting, where the model performs well with training data but fails to generalise with new data.

To avoid redundancy, some of the features were removed from the dataset. The removed features are-

- Close and Adjusted close price—They were removed while keeping the next-day closing price feature in the dataset. The next-day price serves as the dataset's target variable, and the Close and adjusted close price closely relate to the next-day closing price. Using the closing price of the current day can lead to data leakage in the system.

- Volume – Although volume provides insight into market activities, it is highly correlated with other features like volatility and standard deviation. It was less effective in the prediction. Further, technical indicators like MACD as a feature reduced the need for direct volume data, as the trend and volume are highly correlated.

- Short EMA and Long EMA – These components were removed from the features since they were used to assist in calculating the MACD. Therefore, these features were avoided due to multicollinearity, which can impact the model's performance.

Hence, the chosen features were Open, High, Low, Volatility, MACD, Signal Line, Pct_change, Middle Band, Standard Deviation, Upper Band, Lower Band, and Next Day. These features were derived from the dataset, especially to provide balance in price data, trend, momentum, and volatility to capture the complex dynamics of meme stock prices.

3.7 Sentiment Analysis Modelling

For this study, Valence Aware Dictionary and Sentiment Reasoner (VADER) was utilised to perform sentiment analysis on the gathered posts and comments on WSB. VADER utilises a rule-based lexicon approach to sentiment analysis, specifically designed to analyse sentiments on social media platforms. Unlike other machine learning models, this tool has been developed through a process of human validation and expert judgements (Elbagir & Yang, 2019). The core of the VADER sentiment lexicon is a combination of different sources, such as the general inquirer lexicon, human annotations and gold standard datasets such as tweets, movie reviews and Amazon reviews (Hutto & Gilbert, 2014).

The VADER sentiment analysis tool calculates a compound score based on each word's valence scores within a post or a comment (Hutto & Gilbert, 2014). This score was further enhanced by applying pre-defined rules such as word order sensitivity, presence of emphasis (e.g., capitalisation, punctuation) and degree of modifiers. In the current context, the compound score was used to represent the overall sentiment of each post or comment for the meme stocks being analysed in order to produce a quantifiable sentiment measure.

Unlike most pre-trained natural language processing models, VADER has the ability to refine the predetermined scoring dictionary with custom words (Elbagir & Yang, 2019). Therefore, a custom lexicon dictionary was incorporated with the existing VADER word list. Moreover, the custom dictionary consists of frequently used words and slangs by the WSB community. The WSB lingo was sourced through online sources e.g.- (Mathis, 2023), (r/WallStreetBets, 2022), (CLEMEN, 2024). Each term was assigned a sentiment score ranging from strongly negative (-4.0) to strongly positive (4.0) with relevance to the context in which it is typically used. This approach ensured that the sentiment analysis aligns more closely with the unique wordings used in the WSB community.

E.g., ‘citron’ and ‘Hindenburg’ are often associated with short selling. Therefore, they were assigned with strong negative score (-4.0). On the contrary, ‘moon’, ‘calls’ and ‘bull’ represent positive context hence assigned a strong positive score (4.0). whereas words like ‘stonk’, and ‘tendies’ are in between words; hence, they were assigned a sentiment score of 1.9 and 2.0, respectively.

For the sentiment analysis pipeline, the previously accumulated stock tickers dictionary was sorted in descending order based on the number of times each ticker was mentioned in the posts and comments; out of the sorted tickers, the top ten most mentioned stock tickers were selected. Another list of tickers consisting of the highest, most mentioned five stocks was created for the sentiment analysis. And a different dictionary named ‘scores’ was initialised to store sentiment scores for each stock ticker. Comments and post titles associated with each selected stock ticker were stripped of emojis, punctuations, stop words and numbers to clean the text for the sentiment analysis. Further, the cleaned text was tokenised into individual words. The remaining words were lemmatised to reduce the words to its base form to maintain consistency. For each word in the cleaned and tokenised comment, sentiment scores of negative, positive, neutral (ranging from 0 to 1) and compound scores (ranging from -1 to 1) were accumulated using VADER. Furthermore, for the recognised stock ticker symbols within the comments, it was configured to automatically assign a positive sentiment. Since in the Reddit WSB, if a stock is constantly discussed and analysed, there is a high probability of a positive outlook on the stock, as the WSB forum is mainly focused on buying opportunities (Aloosh, et al., 2021). Further, it simplifies the sentiment analysis process.

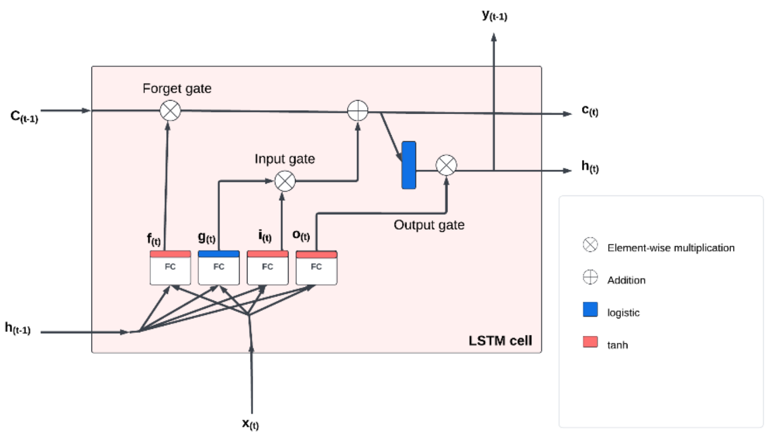
To obtain a comparable, normalised sentiment value, the accumulated scores for the comments were averaged by the number of words. The normalised scores were added to the total sentiment score of the relevant stock ticker symbol in the ‘scores’ dictionary. Finally, the total sentiment value of each ticker was calculated by averaging the sentiment scores by the number of times each ticker was mentioned, which presents an overall mood towards a particular stock mentioned in the comments.

3.8 RNN Modelling

The next step was to predict the individual stock price data using RNN models (single-layered LSTM, regular stacked LSTM, BI-LSTM, GRU) on historical and technical data. Unlike traditional feed-forward neural networks, RNNs assume that all inputs are dependent on each other, and they are specifically designed to process sequential data. Therefore, RNNs are particularly ideal for meme stock price prediction as previous days' stock price data influence the current stock price (Zdravković, et al., 2023). Out of the available RNN models, LSTM and GRU were explicitly chosen for this study as they are advanced types of RNNs that address the issue of vanishing gradients. This allows both LSTM and GRU models to learn long-term dependencies more effectively (Pawar, et al., 2019). This was a crucial requirement in identifying complex patterns in meme stock price data.

3.8.1 LSTM

Long Short-Term Memory (LSTM) network (Figure 3.2) has demonstrated better performance in various domains that involve sequential data, such as time series prediction, sentiment analysis and speech recognition (Pawar, et al., 2019). This high effectiveness in learning from sequential tasks makes the LSTM model highly suitable for tasks such as anomaly detection and financial forecasting.



$h_{(t-1)}$ & $h_{(t)}$: Short-term state at timestep t-1 and t

$c_{(t-1)}$& $c_{(t)}$: Long-term state at timestep t-1 and t

$f_{(t)}$: Forget gate vector.

$i_{(t)}$: Input gate vector

$g_{(t)}$: Candidate vector

$o_{(t)}$: Output gate vector

Figure 3.2 LSTM Model Structure

As depicted by the (Figure 3.2), the core advantage of LSTM lies in its unique architecture that learns which information to retain, discard or use as input for future predictions (Patel, et al., 2021). This process is facilitated by the gates that regulate the flow of information. When the LSTM model was fed with 3-dimensional arrays of individual meme stock data, structured with 10 timesteps, the LSTM model processed the data sequentially. At each time step, the long-term state (c(t)) is updated by combining the previous state (c(t-1)) with novel information as detected by the input gate (i(t)), forget gate (f(t)) and candidate vector (g(t)). After updating the cell state, the short-term state (h(t)) is calculated, reflecting the LSTM cell output for the current timestep. This output is influenced by both the updated long-term state (c(t)) and the output gate (o(t)), which determines which part of the cell state should be output as h(t). As the LSTM processes each of the 10 timesteps, it continuously updates the internal states based on the input data and gating mechanisms. Finally, the last output after processing the 10th timestep is passed through subsequent layers, known as the dense layers; this produces the final, next-day closing stock price prediction.

3.8.2 Single-Layered LSTM

This is the simplest form of the LSTM model (Figure 3.3). It was implemented to serve as a baseline to evaluate the performance of a more complex architecture. Since it is easier to interpret and debug, it has provided a foundational understanding of the LSTM process for handling meme stock data and identifying immediate and more straightforward relationships in the data.

This model was initiated with Keras, and the features were input to each timestep through the sequential model input layer. The LSTM layer consists of 128 LSTM neurons, meaning the model has the capacity to learn 128 different patterns in the data. Moreover, the 'return_sequence' was set to 'False' to ensure that the output is a single vector after processing the entire input sequence. Following the LSTM layer, a dense layer was created with 25 neurons that connect with all 128 neurons in the LSTM layer. And the output layer was initiated with a single neuron that is fully connected to all 25 neurons in the previous dense layer to predict the output. For this study, Mean Squared Error was chosen as the loss function to average the squared difference between stock prices and actual prices, which aided in heavily penalising larger errors. Due to the efficiency and ability to handle sparse gradients, the Adam optimiser was used to adjust the model's weight during the training to minimise the loss.

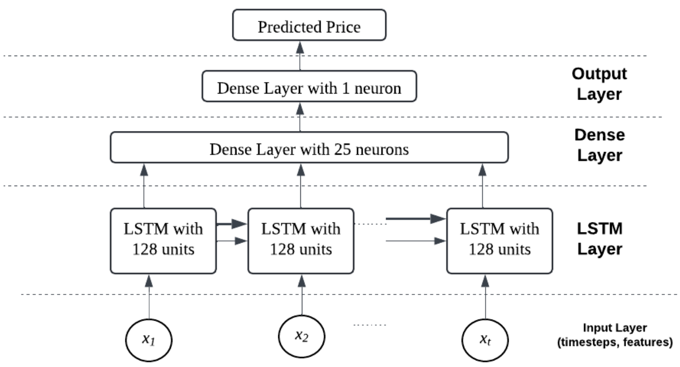


Figure 3.3 Single Layered LSTM Model Structure

The model was trained for 50 epochs with the batch size set to 32. This means that the dataset was divided into samples of 32 chunks that were processed at once in each epoch. A validation dataset was utilised to tune the hyperparameters of the model while helping prevent overfitting and data leakage. The early stopping was set to 10 to stop the training process so that the model would cease to show improvement.

3.8.3 Regular stacked LSTM

This model is slightly more complex than a single-layered LSTM as it stacks multiple LSTM layers to learn more complex patterns and hierarchical representations from data (Figure 3.4). Hence, it can potentially capture more intricate relationships within sequential data while uncovering more complex temporal relationships.

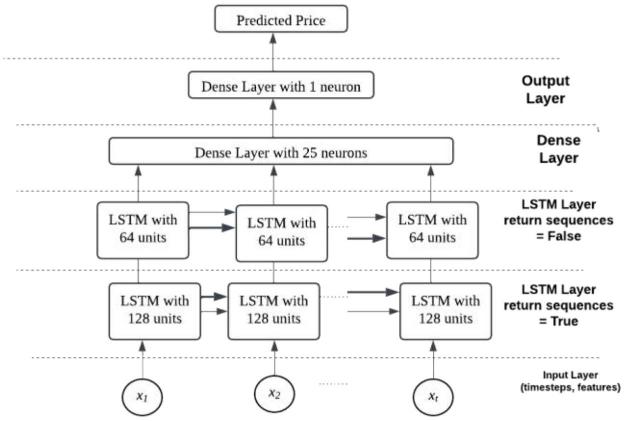


Figure 3.4 Regular Stacked LSTM model Structure

Similar to the single-layered LSTM model, the regular stacked LSTM model was constructed using the Keras sequential API. And the model was designed with two LSTM layers: the first layer consisting of 128 units that return sequences, and the second layer with 64 units that process the entire sequence and pass a single output to the dense layers. The first dense layer consisted of 25 units, followed by the next dense layer with a single unit that produced the output. The model was compiled with Adam optimiser and mean squared error loss function. The training was conducted over 50 epochs with a batch size of 32. And the early stopping was initiated at 10 epochs.

3.8.4 BI-LSTM model

This is a complex version of the LSTM model; this model processes the sequence in both forward and backwards directions, which allows the model to capture dependencies that might have been missed by unidirectional models. Particularly in meme stock price prediction, this can be identified as a useful feature of the model since past events influence future prices and

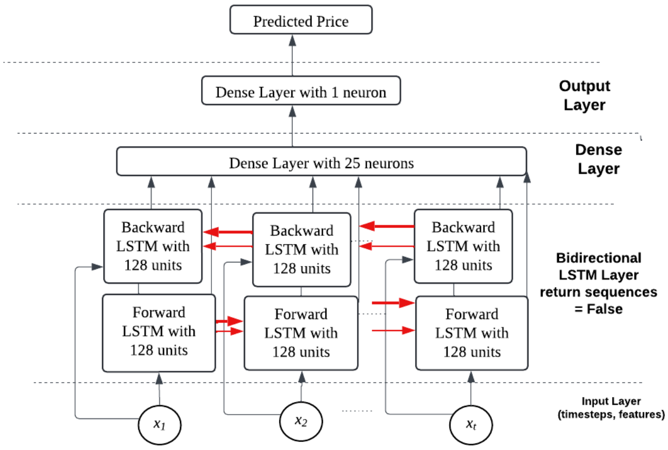


Figure 3.5 Bi-LSTM Model Structure

considering data in both directions can further enhance prediction accuracy (Figure 3.5).

The model was developed using similar epochs, batch sizes, and optimisers using the bidirectional library of Keras. The model processed the input data in both forward and backwards directions, providing a more comprehensive understanding of the data's sequential dependencies.

3.8.5 GRU

GRU is another variant of RNNs designed to address the vanishing gradient problem and improve the efficiency of sequential data processing (Figure 3.6). Unlike LSTM, which has separate memory cells and gates to control the flow of information, GRU combines both forget and input gates into a single update gate, simplifying the architecture and computational resources to speed up the training process (Gao, et al., 2021).

While both LSTM and GRU models are designed to handle long-term dependencies and mitigate the vanishing gradient problem that is common in traditional RNNs, GRU differs from LSTM in architectural simplicity and computational efficiency (Li & Qian, 2022). Unlike LSTM, which uses separate memory cells along with input, output and forget gates, the GRU combines all these mechanisms into a single update and reset gate. This streamlined structure allows GRU to train faster with the aid of fewer parameters, which can be advantageous for smaller datasets.

For this study, the GRU model was built with four layers, with each layer containing 64 units and using the 'tanh' activation function. Dropout regulation with a rate of 0.3 was applied to each GRU layer to prevent overfitting. The model was compiled with a Stochastic Gradient Descent (SGD) optimiser to gradually reduce the learning rate as the training progresses. Similar to LSTM Mean Squared Error (MSE), the loss function was used, and the early stopping callback was set to 10 epochs to halt training if no improvement in validation performance occurred. Further, the model was trained through 50 epochs, and the dataset was segmented into 32 batches to allow pattern learning through training data and generalising to the testing data.

In terms of predictive performance, the GDRU model, in fact, did not outperform the LSTM models. This suggests that the complex gating mechanism of LSTM may allow the model to capture nuanced patterns in highly volatile meme stock prices, particularly where subtle sequential relationships matter. However, GRU still demonstrated acceptable performance. Additionally, its lower training time may make it a viable option in a real-time or resource-constrained environment. These results underscore the importance of selecting models based not only on theoretical strength but also empirical performance aligned with the task and dataset characteristics.

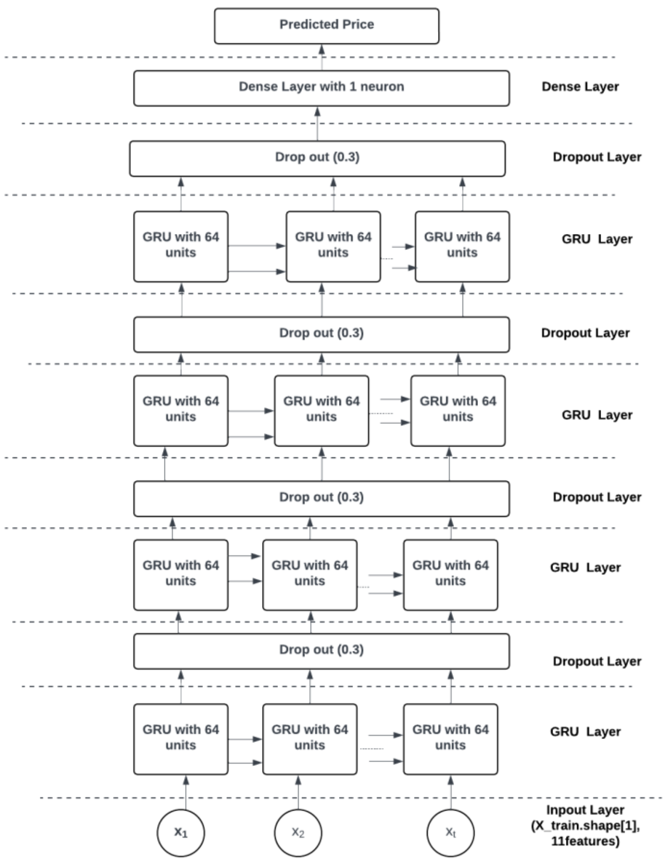


Figure 3.6 GRU Model Structure

3.9 Model Evaluation

3.9.1 Evaluation Metrics

The performance of the models was evaluated using popular error metrics as follows –

Root Mean Squared Error (RMSE)

3.2) is a widely used machine learning model evaluation metric in regression tests, which measures the average magnitude of prediction errors. It is calculated by acquiring the square root of the mean of the squared error differences between predicted values and actual values. This gives higher weight to larger errors, which makes it particularly sensitive to outliers. In the meme stock price prediction models, RMSE provides a quantitative representation of how well the models capture sudden price spikes and drops, which are highly common in meme stocks. A lower RMSE indicates that the model has higher accuracy and a closer fit to the actual price trend.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \qquad 3.2$$

y_i - Actual value of the target meme stock price for observation I, $\hat{y}$ - Predicted stock price for observation I, n – total number of observations,

Mean Absolute Error (MAE)

3.3*)* is another metric used to measure the average magnitude of errors in predictors. The measure is calculated as the average of the absolute difference between predicted values and actual values. Unlike RMSE, MAE assigns equal weights to all errors, making it less sensitive to outliers and more interpretable for understanding the average size prediction errors. In meme stock price prediction models, MAE was applied to assess the overall prediction accuracy of the models without the influence of extreme price fluctuations. A lower MAE suggests that the model consistently makes predictions that are closer to the actual values.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \qquad 3.3$$

y_i – Actual meme stock price, $\hat{y}_i$ - Predicted value of observation i, n – Total number of observations

Mean Absolute Percentage Error (MAPE)

3.4*)* represents the average percentage difference between predicted and actual values. This error metric provides the relative accuracy of the predictions. MAPE is necessarily useful when comparing models across stocks with different price ranges since it normalises the errors as a percentage of the actual values. In the current study, MAPE allows for a better understanding of the models' performance of all three top stocks across the study period. This makes it easier to evaluate the model's performance despite varying prices.

$$MAPE = \frac{100\%}{n} \sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \qquad 3.4$$

y_i - Actual closing price for observation i, $\hat{y}_i$ – Predicted price for observation I, n – number of observations, 100% - Multiplied to convert to percentage

3.9.2 Real-time evaluation

Other than the error metrics, from the collected data across all three days, the study was conducted (26 August – 28 August 2024), during the critical time period from 4 p.m. to 9 a.m. EST, real closing price analysis was performed on the model predictions. This offered a real-world assessment of the effectiveness of the meme stock price prediction models. The specific time period was selected for Reddit sentiment analysis because this time period encompasses both the previous day's post-market discussions and key speculative discussions about the next day's trading strategies and produces insight into the latest sentiment and dominant opinions that can influence the upcoming trading day. Further, this time period was specifically selected as it encompasses the latter days of the week, including Wednesday, Thursday and Friday, which include daily discussions that analyse daily trading patterns, weekend discussions that analyse the trading strategies of the entire week and also speculative discussions for the coming week.

Results

3.10 Reddit WSB Sentiment Analysis

According to Reddit's PRAW API documentation (Reddit, n.d.), access to Reddit content is limited to 60 requests per minute and up to 1000 items per request. To gain optimum access to valuable data within these constraints, daily comments and post titles were acquired from

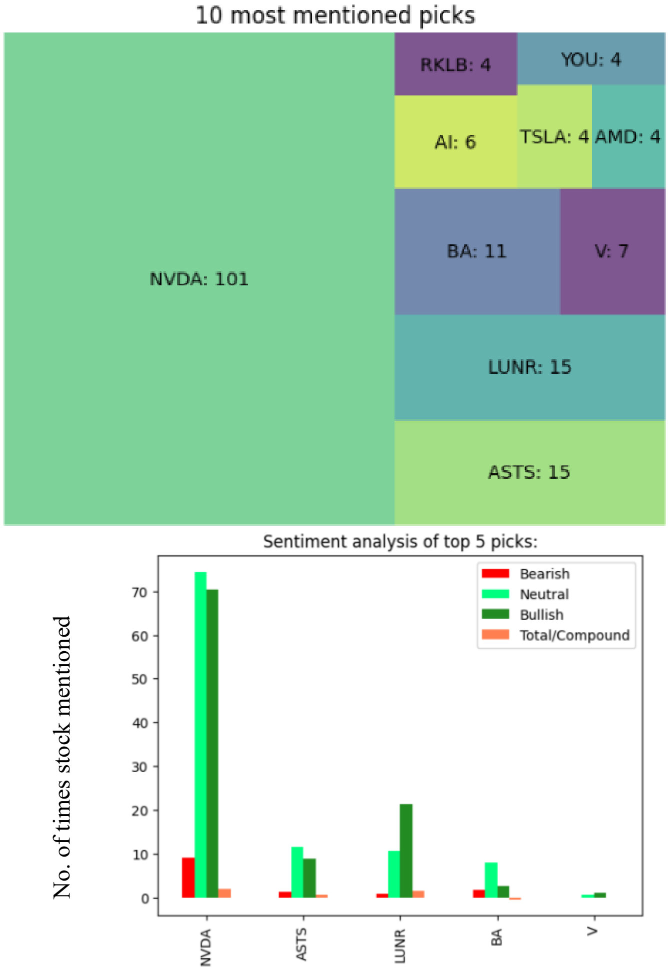
WallStreetBets using 'hot()' sorting criteria, which retrieve 100 posts per request to capture the most popular daily discussions. The acquired posts were further filtered to include posts that are fliared as 'Daily Discussions', 'Weekend Discussions' and 'Discussion' and posts without any other flair. Additionally, posts that meet or exceed a 0.70 upvote ratio with more than 20 upvotes were considered to ensure focus on highly relevant and engaging content.

From these acquired posts, comments were considered for the analysis if they had at least two upvotes and were unique comments by distinct authors, adhering to the criteria for meaningful engagement and avoidance of repetitive contributions from the same Redditor. This data collection was run during the time window between 4:00 PM to 9:00 AM EST, which included the prior day's post-market hours and the pre-market hours of the trading day. This period is crucial for speculators as it captures discussions on next-day trading strategies and reflects on the latest sentiments and dominant opinions before the market opens.

The code was executed consecutively from August 26, 2024, to August 28, 2024, to ensure comprehensive sentiment analysis and observe changes over time. This time frame also included end-of-weekend discussions, which often feature speculative trading strategies for the upcoming week and analysis of prior-week strategies. By capturing these dynamics and influential discussions, the sentimental data provided valuable insight into the most popular meme stocks among retail investors, which were then utilized for price prediction models.

3.10.1 August 26

On August 26th, after analyzing 3825 comments over 21 posts, the top ten most discussed stock tickers were detected as NVIDIA (NVDA) with 101 mentions, followed by AST SpaceMobile Inc (ASTS), Intuitive Machines, Inc (LUNR), each with 15 mentions and Boeing Co (BA) with 11 mentions and several other tickers such as Visa, Inc (V), C3.AI, Inc (AI), Rocket Lab USA, Inc (RKLB), Tesla, Inc (TSLA), Advanced Micro Devices, Inc (AMD), Clear Secure, Inc (YOU) with less mentions (Figure 3.7).



Sentiment analysis of top 5 picks:

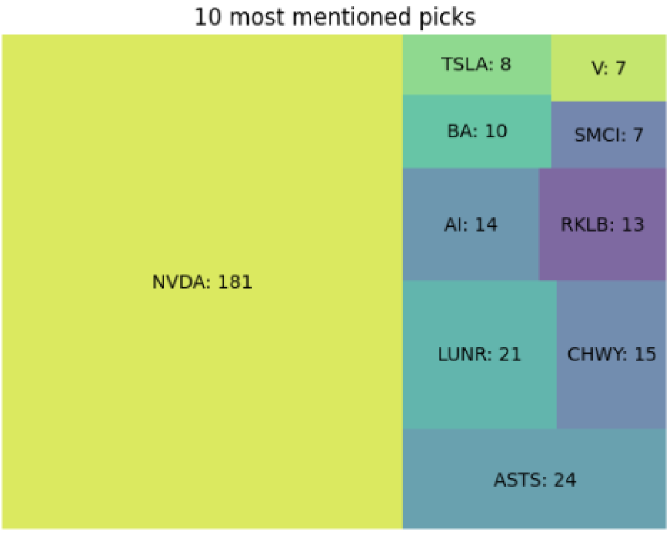
	Bearish	Neutral	Bullish	Total/Compound
NVDA	9.134467	74.481543	70.359574	2.053253
ASTS	1.193476	11.470873	8.83225	0.519902
LUNR	0.75	10.633333	21.416667	1.457397
BA	1.759947	7.969551	2.696262	-0.395901
V	0.036	0.643	1.083	-0.001

Figure 3.7 August 26 Sentiment Analysis – 10 most mentioned picks (top left), Total number of times stocks mentioned on posts and comments (bottom left) VADER sentiment scores of top 5 (right)

Vader sentiment score calculation: negative, neutral and positive scores – average sentiment per stock, calculated by averaging sentiment scores of each related comment and post title, total/ compound score – overall sentiment polarity (aggregated sentiment divided by the number of times stock was mentioned in posts or comments)

From the identified stock tickers, sentiment analysis was carried out on the top five stock tickers and categorised them into bullish, bearish and neutral sentiments, along with compound scores representing the overall sentiment. Apparently, NVDA showcases the highest level of engagement with a balanced mix of sentiments as 9.91 bearish, 74.48 neutral and 70.36 bullish comments, which resulted in a positive compound score of 2.05, which showcases WSB Redditors' positive sentiment towards NVDA. As depicted by the analysis, LUNR also presents a positive sentiment about the future of the stock, while ASTS obtained a lower yet positive sentiment. However, BA and V sentiment scores leaned towards bearish sentiment, with sentiment scores nearly zero.

3.10.2 August 27



Sentiment analysis of top 5 picks:

	Bearish	Neutral	Bullish	Total/Compound
NVDA	17.111468	135.575811	109.585905	3.650097
ASTS	2.463073	16.941813	19.12083	1.018879
LUNR	1.486111	13.875397	24.376984	2.110761
CHWY	1.552632	9.584928	12.287081	1.258101
AI	0.080	0.842	0.227	0.003

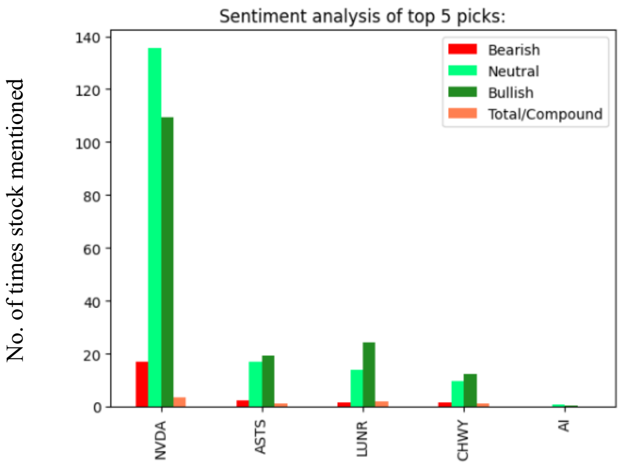
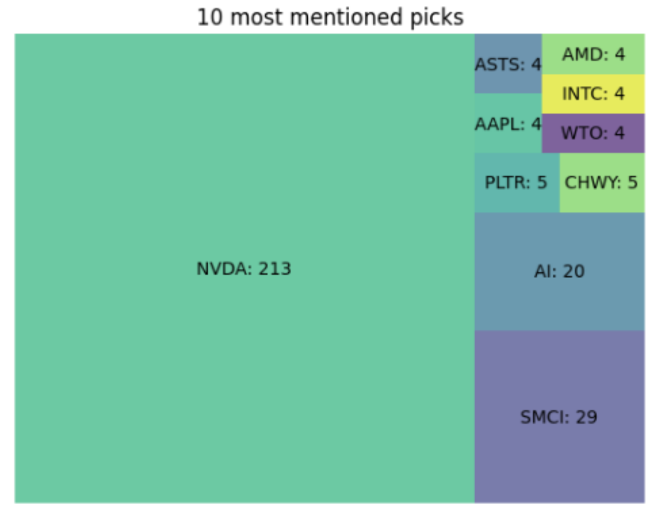


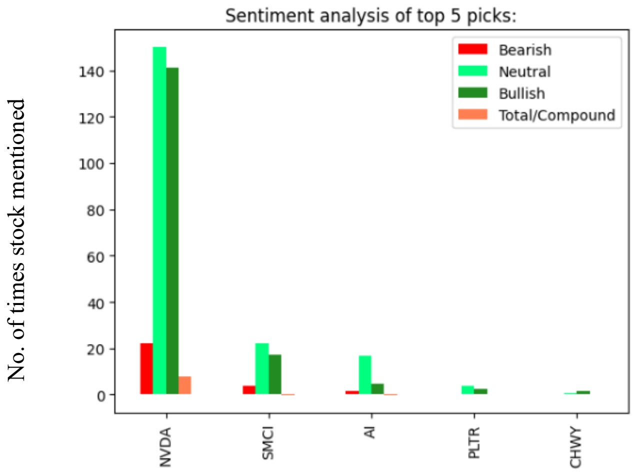
Figure 3.8 August 27 Sentiment Analysis 10 10 most mentioned picks (top left), Total number of times stocks mentioned on posts and comments (bottom left) VADER sentiment scores of top 5 (right)

Similar to the previous day's results, NVDA remained the highest-mentioned stock ticker with 181 mentions, followed by ASTS and LUNR consecutively with 24 and 21 mentions, which was slightly higher than the previous day. Although the following most popular meme stock, BA, was taken over by Chewy, Inc. (CHWY), an online retailer of pet products company with 15 mentions, AI stock by 14 mentions and Rocket Lab USA Inc (RKLB) with 13 mentions. TSLA and V remained at the lower level, with fewer mentions, together with Super Micro Computer Inc (SMCI) gaining some attention (Figure 3.8).

Compared to 26th August, NVDA became even more bullish (109.58) with a compound score of 3.65, reflecting growing investor confidence and positive news impacting the stock. Similarly, ASTS and LUNR showed higher bullish scores compared to the previous day. Notably, CHWY gained retail investor attention with a significant bullish sentiment of 12.29 and a compound score of 1.26. Although the sentiment of AI was minimal, with a negligible overall sentiment score.

3.10.3 August 28





Sentiment analysis of top 5 picks:

	Bearish	Neutral	Bullish	Total/Compound
NVDA	22.193496	150.050511	141.315809	7.881146
SMCI	3.911598	22.172936	17.193171	-0.334098
AI	1.719082	16.595327	4.58434	-0.065189
PLTR	0.285714	3.845238	2.42381	0.239371
CHWY	0.040	0.680	1.420	0.038

Figure 3.9 August 28 Sentiment Analysis 10 most mentioned picks (top left), Total number of times stocks mentioned on posts and comments (bottom left) VADER sentiment scores of top 5 (right)

On August 28, 2024, NVDA mentions further increased up to 213, confirming its position as the most popular stock among redditors in WSB. Followed by SMCI and AI with 29 and 20 mentions, respectively. Unlike previous days, the rest of the stocks gained fewer mentions.

With the increase of mentions, NVDA's compound sentiment score increased significantly up to 7.88, reflecting a bullish outlook on the stock. However, despite having a higher number of mentions and higher bullish and neutral scores, SMCI and AI gained slightly negative compound scores of -0.33 and -0.07, respectively, indicating a shift in the WSB community sentiment. Sentiment for PLTR and CHWY remained marginally positive with minimum overall mentions, reflecting less community interest and engagement towards these stocks compared to the leading stocks (Figure 3.9).

Across all three days, NVDA remained the most mentioned and discussed stock on Reddit WSB, with its compound sentiment score growing in a bullish manner from 2.05 to 7.88 in just two days. This highlighted the strong positive sentiment of the community towards NVDA. Although ASTS and LUNR maintained their position as the next most discussed stocks on WSB with a positive sentiment, the interest faded in these stocks by the third day. Instead, stocks like SMCI and AI gained the interest of the community, and the sentiment reflected a mixed to slightly bearish sentiment, suggesting a slight divergence in the community outlook.

3.11 Next-day Price Prediction with RNN

The RNN models were trained and tested on the historical prices and technical indicators of the top three most mentioned stock tickers on WSB. For each stock, the date range for the analysis was selected from the day it started trading publicly, as most of the stocks mentioned on WSB, like ASTS, LUNR and AI, went public very recently. However, some of the stocks have been in the market for years, like NVDA, TSLA and AMD, and have recently gained popularity. For those stocks, the last 10 years of historical price data were considered. Technical indicators were generated on the relevant datasets to enhance the predictive power of the models. This stock-specific time frame approach was considered to ensure that the models were trained in comprehensive and relevant data, reflecting both historical and current market trends.

3.11.1 Single-Layered LSTM

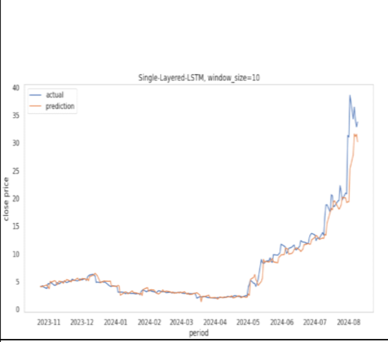
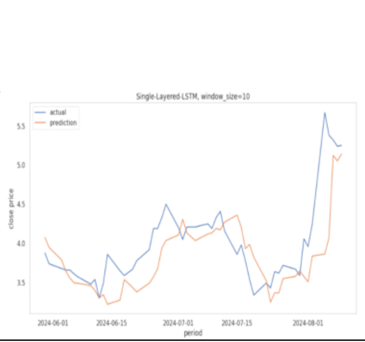
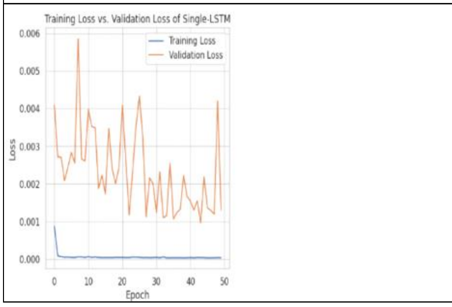
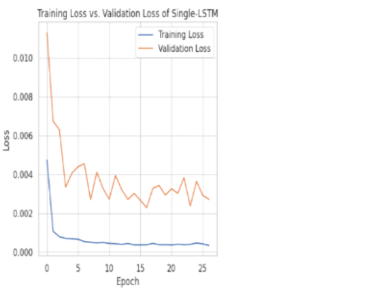
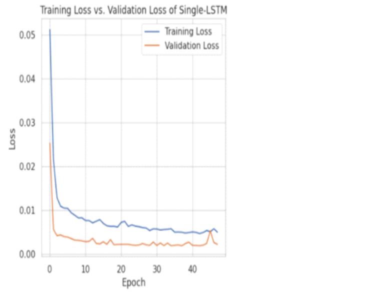
3.11.2 August 26

The single-layered LSTM model demonstrated strong predictive performance on NVDA stock, utilising historical price data from 26/08//2014 to 26/08/2024. As depicted by the graph, the model effectively captured the stock price's upward trend and volatility. The graph displayed a close alignment between the actual and the predicted prices over the testing period. The training and validation loss curves indicated a stable convergence, with training loss consistently appearing lower than the validation loss. This visual representation suggested that the model managed to generalise well without significant overfitting. This performance highlighted the ability of the model to leverage long-term historical data to predict the next day's closing prices accurately. This verified the robustness of the model in predicting the price of established stocks like NVDA with extensive data history.

For ASTS, the model was trained on data starting from 07/04/2021, as this stock was publicly listed from this date. The model's performance showed that it could generally track the overall direction of the stock price movement. However, it showed some lag in adjusting to rapid price changes. This is acceptable for the model, given the typical characteristics of high volatility and speculative trading activities. The training and validation losses showed more variability than NVDA, indicating challenges in identifying consistent patterns in short-term volatile data periods. However, the model's predictions were more reasonably aligned with the actual prices, indicating that the model provides useful insight into more recent and volatile stocks compared to well-established stocks.

LUNR is a stock that started public trading very recently on 13/02/2023. Therefore, the historical data was utilised from that day onwards. The predictive performance of this model showed difficulty in capturing the volatility and rapid price fluctuations of the newly listed stock. The actual and the predicted prices showed considerable divergence, especially during periods of sharp price movements, indicating the model's limited ability to adjust swiftly to sharp price changes and dynamic conditions. The training and validation loss plots showcased more unstable behaviour, with validation loss occasionally spiking, suggesting that the model struggled with overfitting issues and issues in finding a stable pattern in limited data. The overfitting issue occurred due to the limited historical data; therefore, early stopping was initiated with a patience value of 10 epochs, and dropout layers were introduced with a dropout rate of 0.2 to neutralise the model from becoming overly dependent on specific features. Despite these regularisation efforts, the results displayed only a marginal improvement. The validation loss curve remained unstable. This suggests that the model struggled to learn consistent patterns from the limited, noisy dataset. This reinforces the difficulty of applying deep learning models to newly listed stocks with limited price history and high volatility.

Table 3.1Single-Layered LSTM Results August 26 stock price prediction curves (top row), training loss vs validation loss (bottom row)

NVDA	ASTS	LUNR
		
		

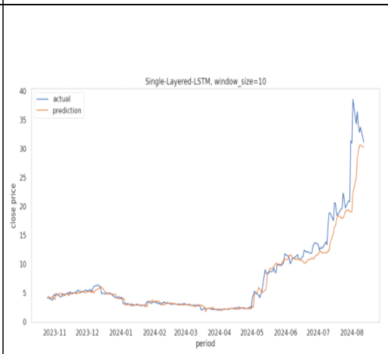
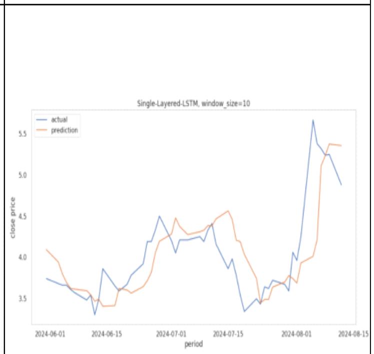
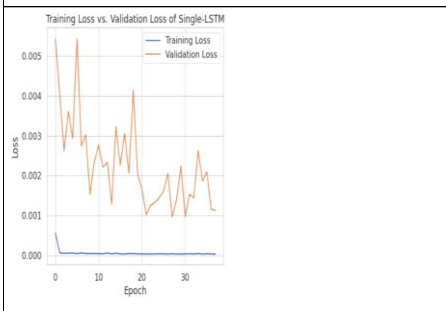
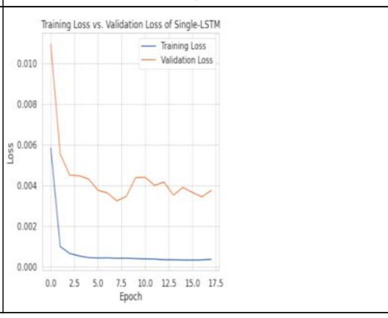
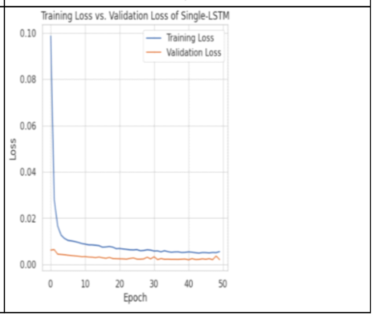
3.11.3 August 27

Similar to the previous day's performance, the model maintained a high level of alignment between the predicted and actual stock price for NVDA. The trend lines closely followed the actual price movements, including the most recent price fluctuations, indicating the model's effectiveness in capturing the long-term price direction. The training loss graph showed relatively low and stable behavior, indicating the model's ability to learn from the training data. Although the validation loss showed significant fluctuations, especially in the early epochs, suggesting the model struggled in generalizability to unseen data, and the high variability in validation loss compared to stable training loss pointed out the issues of overfitting and instability. Regardless, the validation loss followed a downward trend. To address this this, multiple regularization techniques were employed, including dropout layers increased from 0.2 to 0.4 early stopping with reduced epochs from 10 to 5 and the addition of L2 regularization. Although some improvement in validation stability was observed, the high volatility of stock data remained a limiting factor.

For ASTS, the model's predictive abilities were moderately successful. According to the graph, the prediction was able to follow the general upward trend of the stock price. Although, the prediction showed some occasional lag and some deviations from the actual price, especially during 06/2024 – 08/2024, when the price rapidly escalated. The loss graphs indicate relatively stable training loss with a volatile validation loss stipulating challenges in generalizing from training data to unseen test data that reflected on the stock's volatility and limited data availability due to its recent public listing in 2021.

The performance of the LUNR model persistently depicted difficulty in dealing with the highly volatile nature of the recently listed stock. Although the prediction line captured some of the general upward and downward trends, it lagged behind the actual price, especially during sharp price movements. The training loss showed consistency which demonstrated that the model is learning quickly from the training data. The validation loss also dropped initially but remained relatively stable and consistently lower than the training loss throughout the epochs, suggesting the model’s generalizability to validation data without overfitting.

Table 3.2 Single-Layered LSTM Results August 27 stock price prediction curves (top row), training loss vs validation loss (bottom row)

NVDA	ASTS	LUNR
		
		

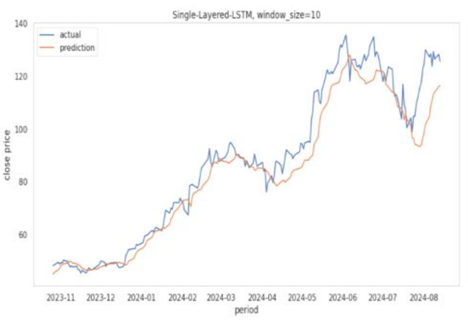
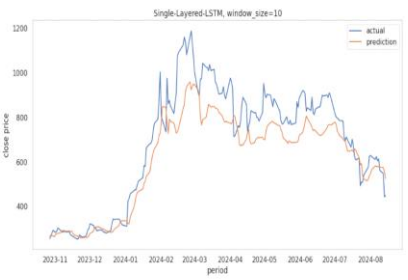
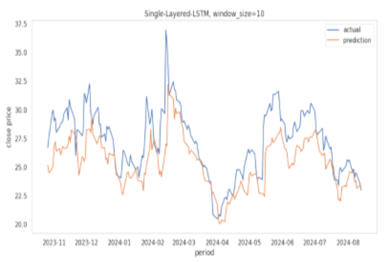
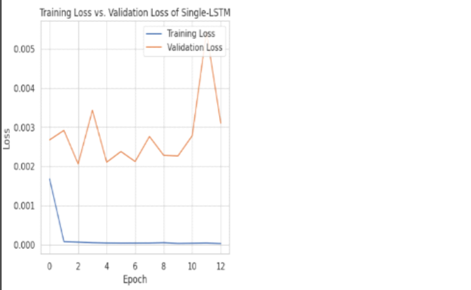
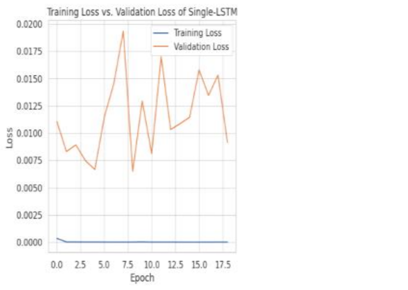
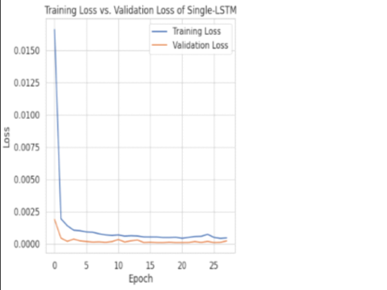
3.11.4 August 28

The model trained on NVDA data for 28/8 demonstrated strong predictive performance in accurately capturing the stock’s upward trend over the long term. The protection line closely followed the actual stock prices, showing the model’s ability to learn and replicate NVDA price movement effectively. However, some deviations could be observed, especially during times of rapid price changes, where the model occasionally lagged behind stock performance. In terms of the loss graphs, the training loss rapidly decreased and stabilized at a lower value in the early stage of the training process. Although validation loss exhibited considerable fluctuations throughout the epochs, which indicates the model's incapability to apply the learned patterns to the validation set as a result of overfitting where the model fits too closely to the training data. Therefore, the performance indicated that the prediction of this model is less reliable.

For SMCI, the model effectively captured the overall price trends. although particular challenges could be observed during periods of rapid price fluctuations. The predicted values are generally aligned with the actual stock prices. However, certain struggles could be observed in reflecting sudden drops in stock prices. Meaning the lags and smoother predictions depicted the models' incapability to capture the stock's volatility.

The model's performance for AI stock prediction demonstrated a decent alignment between the actual and the predicted values. The model was able to capture the overall trends and fluctuations of the stock's closing price. However, some variations between the actual and the prediction could be noticed during highly volatile periods. This indicated that while the model effectively recognizes general patterns, it struggles to detect unprecedented changes. As for the training and validation loss curves, both show a sharp decrease in early stages followed by stabilization at relatively low values, suggesting that the model learned the underlying pattern well without overfitting, as the validation loss did not deviate drastically from the training loss.

Table 3.3 Single-Layered LSTM Results August 28 stock price prediction curves (top row), training loss vs validation loss (bottom row)

NVDA	SMCI	AI
		
		

3.12 Regular Stacked LSTM

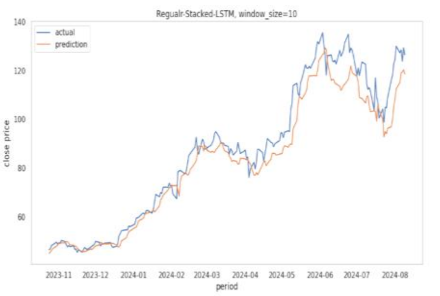
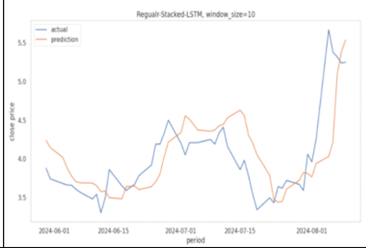
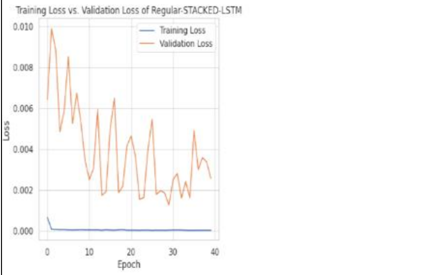
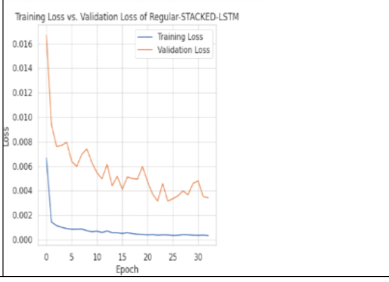
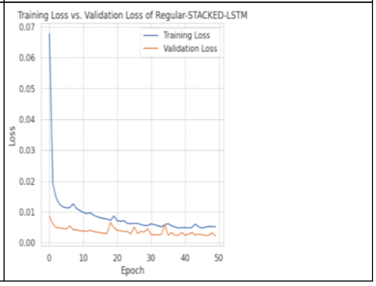
3.12.1 August 26

The model appeared to perform well in predicting the NVDA stock prices, as the predicted prices closely track the actual stock prices, capturing the general upward trend and smaller stock price fluctuations. This indicates that the model effectively learned the long-term patterns and short-term variations. The training and validation loss curves displayed a consistent learning process, with both losses decreasing consistently during the early epochs. However, the validation loss showed more variability compared to the training loss, indicating some degree of volatility and overfitting. Regardless of the fact, the validation loss remained at a lower level throughout the process, signifying the model's ability to generalize to unseen data at a satisfactory level.

As depicted by the graph, The model on ASTS stock demonstrated substantial accuracy in predicting the next day's price movement since the predictions were well aligned with the actual stock prices, closely tracking upward trends and daily fluctuations effectively. In terms of the training and validation loss, both losses showed a persistent decrease as the epochs increased, indicating rapid learning from the training data. However, validation loss slightly increased towards the last epochs, indicating the model struggled to generalize to the most recent trends.

As for the LUNR stock, the model was able to follow the general stock price trends. However, there were noticeable deviations, especially during the 15/6/2024 – 15/07/2024 period, with rapid price movement. Also, it struggles during periods of high volatility, particularly in early August 2024, when the model underestimated the magnitude of the price spike. The training and validation loss graph indicated that the model effectively reduced both losses throughout each epoch, stabilizing at a low value by the end of the training period. The validation loss remained slightly higher than the training loss, suggesting a reasonable generalization of the model without much overfitting with mild fluctuations. Hence, the model showed commendable performance in capturing essential trends, given the volatility and short public trading history.

Table 3.4 Regular Stacked LSTM Results August 26 stock price prediction curves (top row), training loss vs validation loss (bottom row)

NVDA	ASTS	LUNR
		
		

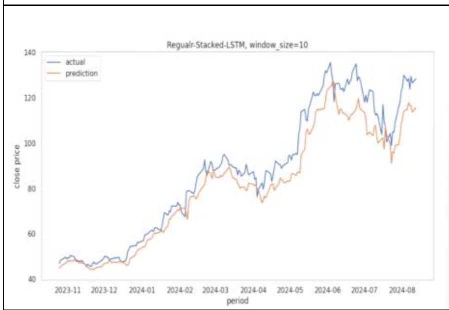
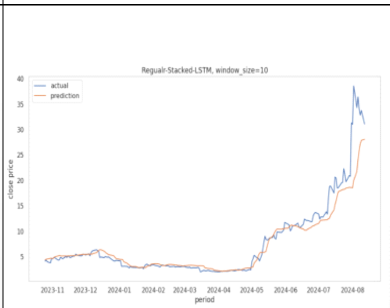
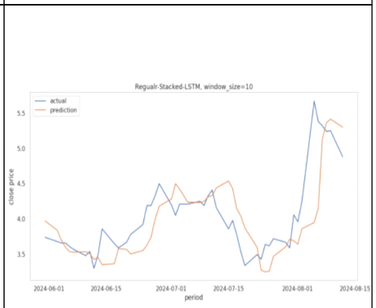
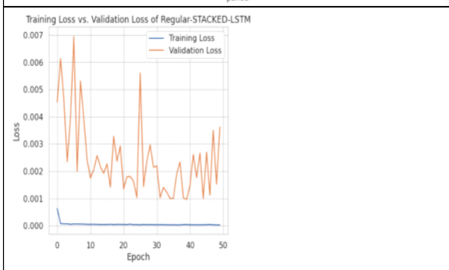
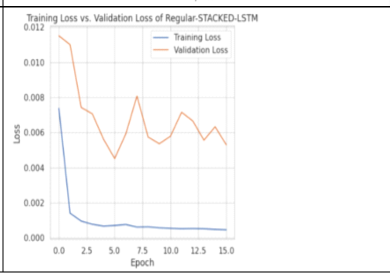
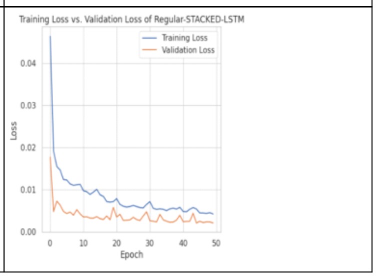
3.12.2 August 27

The regular stacked model's performance on NVDA confirmed that it can generally capture the overall upward trend and daily fluctuations of the stock price. However, the model showed mild incapability in capturing sharp price changes as there appeared to be instances of the prediction lagging or the model not fully capturing the magnitude of the sharp price changes. The training and validation loss graphs revealed a similar idea as the training loss remained consistent and low through the epochs while the validation loss fluctuated significantly, indicating certain overfitting of the model, as it seemed to perform well on the training data but less consistent on unseen data. The spikes indicated that the model may not be entirely reliable during periods of high volatility.

As for ASTS, the model demonstrated a generally effective prediction of the overall upward trajectory. The predictions aligned well with the actual price movements, especially during the gradual rise and sharp spike in the latter part of the observations. Meaning that the model is capable of capturing significant price trends and responds well to dynamic price changes. Although some lag could be identified during the period of rapid price increase, indicating that the model struggled with timing the abrupt market shifts that were more pronounced in volatile meme stocks like ASTS. As per the loss graph, both losses showed a notable decrease in the initial epochs, with the training loss stabilising at a lower level. However, the validation loss remained higher with more variability, implying the model’s struggle in generalising to unseen data.

Confirming the previous day’s results, LUNR stock price demonstrated moderate alignment between predicted and actual values. The model captured the general price movement of LUNR, including the highly contrasting peaks and troughs observed, especially from June to August 2024. However, there were notable discrepancies, particularly in the timing and extent of price fluctuations, where the model predicted more gradual changes compared to sharper actual movements of the market, especially towards the end of the observed timeframe, where there appeared to be a notable divergence as the actual price spiked more aggressively than prediction suggesting the model struggled with less mature stocks. The training and validation losses showed notable improvement in the early epochs as they experienced a steep ascent. However, validation loss remained above training loss, most likely due to high noise in the LUNR's recent market behaviour.

Table 3.5 Regular Stacked LSTM Results August 27 stock price prediction curves (top row), training loss vs validation loss (bottom row)

<table>
<tr><th>NVDA</th><th>ASTS</th><th>LUNR</th></tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</table>

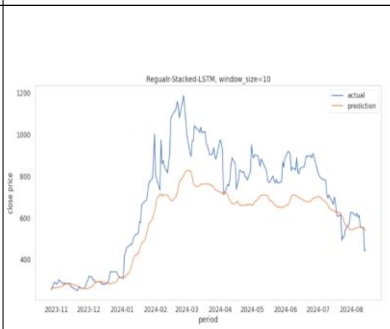
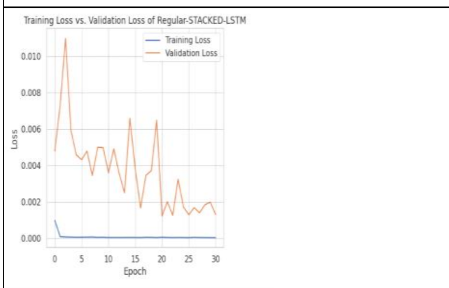
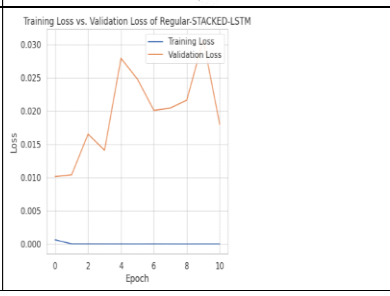
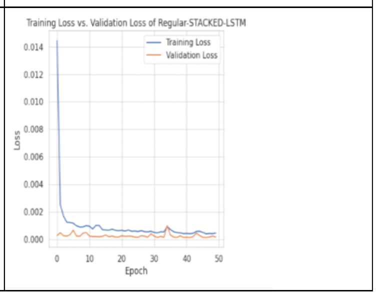
3.12.3 August 28

NVDA stock prices showed a strong correlation between actual and predicted values for the third consecutive day. The model effectively captured the consistent upward trends and occasional stock price corrections from late 2023 to mid-2024. The training and validation losses showed a pattern of rapid decline, suggesting the model quickly learned the underlying trends of the training data. However, validation loss shows a relatively greater variability and remains higher, with clearly visible spikes indicating potential overfitting in generalizing beyond the training dataset.

In terms of SMCI, the model showed limitations in capturing the volatility of the stock price, especially during early 2024. The model struggled to represent the rapid price increase. The training and validation graphs underscored this discrepancy as the validation loss remained significantly at a higher rate than training loss indicating the model's struggle in fully capturing the higher volatility and rapid fluctuation characteristic of the complex stock behavior.

In terms of AI stock, the model performed reasonably well in predicting stock prices as the predictions closely aligned with actual values despite discrepancies in a sharp increase in prices, especially during the beginning of 2024. The training and validation loss represented the model's ability to learn from the data set as the training loss remained flat in its learning with validation loss closely aligning with the training loss with occasional spikes. Which indicates the model's ability to generalize reasonably well to new data.

Table 3.6 Regular Stacked LSTM Results August 28 stock price prediction curves (top row), training loss vs validation loss (bottom row)

<table>
<tr><th>NVDA</th><th>SMCI</th><th>AI</th></tr>
<tr>
<td>
</td>
<td>
</td>
<td>
</td>
</tr>
<tr>
<td>
</td>
<td>
</td>
<td>
</td>
</tr>
</table>

3.13 BI-LSTM

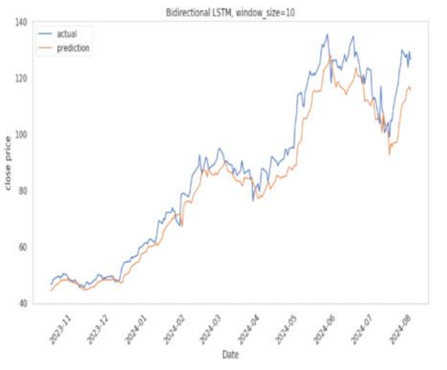
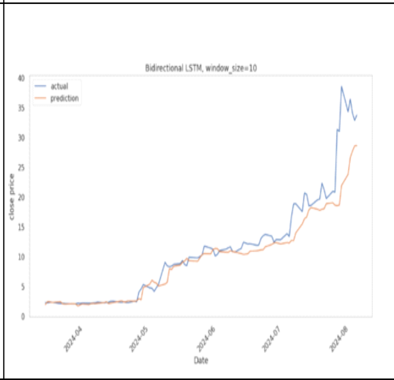
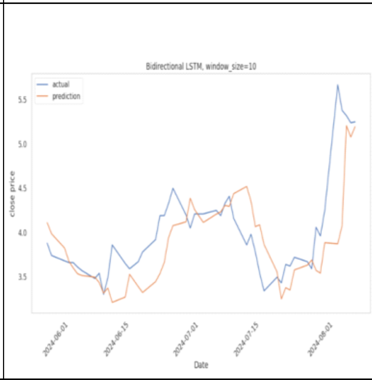
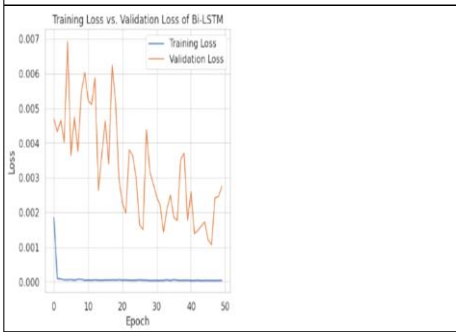
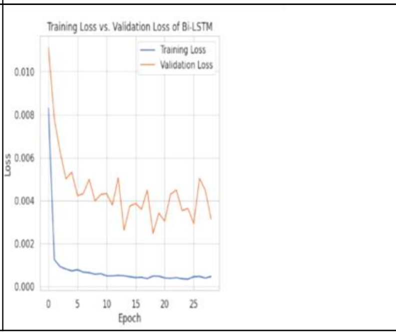
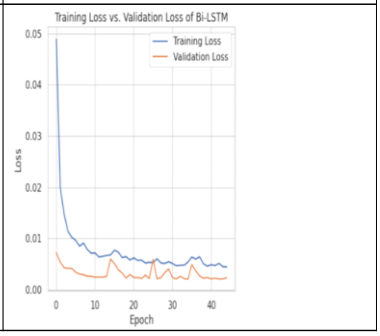
3.13.1 August 26

Like other LSTM models, the BI-LSTM model also showed persistence in capturing the trend of NVDA well by closely aligning with the actual price. However, the price prediction seemed to lag slightly from late 2023 to early 2024, when the market experienced sharp movements. The training loss remained plateau starting from the first few epochs, but the validation loss showcased clearly visible fluctuations indicating the model's difficulty in fitting with unseen data. Although the validation loss gradually decreased by the last epochs, suggesting that the model can predict unseen values without significant overfitting.

Similarly, the model showcased generally good alignment with predicted and actual values, demonstrating a good performance, especially during the initial growth period. However, the sharp increase in price by October 2024 was predicted by the model with a visible lag, indicating the model's ability to follow general trends, although it faced hardships in depicting sharp and unexpected changes in price. The training and validation loss curves represented a similar indication as both losses rapidly declined in the early epochs. However, the validation loss indicated occasional fluctuations with the deviations in the model's performance on unseen data.

As for LUNR, the model showcased moderate success in predicting the values but displayed better performance than ASTS in terms of predicting sharp price increases. However, the prediction line overshot the actual values during the middle of July 2024, which indicated a limitation in the model's ability to emulate sharp and rapid changes in the stock price. The training and validation loss curves illustrated a sharp decline in the first few epochs, indicating the model quickly grasped the critical patterns of the training data, with validation loss displaying mild fluctuations, indicating minor instabilities of the model due to the volatile nature of this latest stock.

Table 3.7 BI-LSTM Results August 26 stock price prediction curves (top row), training loss vs validation loss (bottom row)

NVDA	ASTS	LUNR
		
		

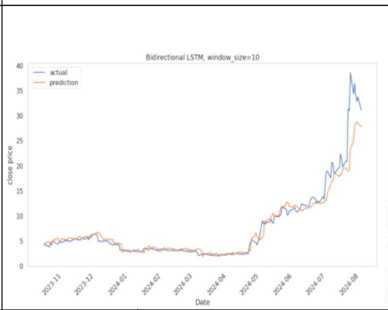
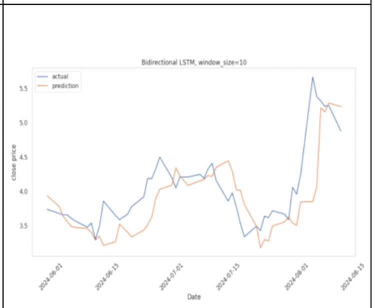
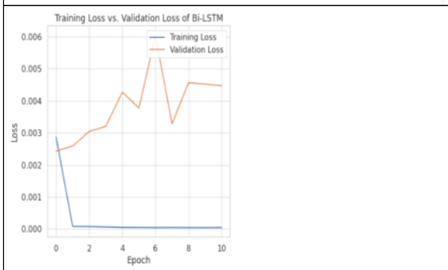
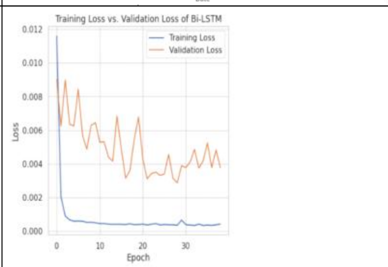
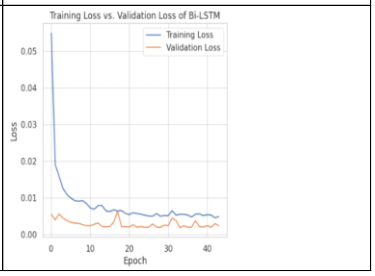
3.13.2 August 27

Similar to the previous day, the prediction line indicated a strong alignment to the actual price over the observed period. The model clearly depicted the upward trend of the stock price during early 2024, reflecting the strength in learning long-term patterns. However, slight deviation could be observed during the more volatile periods, such as mid-2024, where the prediction lagged behind. The training and validation loss curves also confirmed this trend as both curves were reduced at the beginning of the curve. However, validation loss showed upward fluctuation towards later epochs, suggesting clear variability in the model's performance on recent unseen data.

As for the ASTS, the model showed clear improvement compared to the previous day's results, as the model's predictions aligned more closely with actual stock prices. Especially from May to July 2024, when prices followed a clear upward trend. Although the model seemed to struggle with the steepest climb by the end of the period, the overall trajectory showed much compliance with the actual value. The loss curves also reflected this improvement, although certain spikes could be observed in the validation loss. The reduced gap between training and validation loss and the downward trend showcased a clear improvement of the model.

Similarly, LUNR stock price prediction indicated clear improvement, particularly during the sharp increase observed by the end of the prediction period. Which indicated that the model has likely identified the upward momentum and volatility of the stock price. This can be depicted as an improvement of the model since the predictions more visibly lagged behind the actual values and rapid changes. The loss curves also supported this improvement as both curves showed a decrease and stability at a lower value as the epochs increased, although minor discrepancies could be visible in the validation loss curve.

Table 3.8 BI-LSTM Results August 27 stock price prediction curves (top row), training loss vs validation loss (bottom row)

NVDA	ASTS	LUNR
		
		

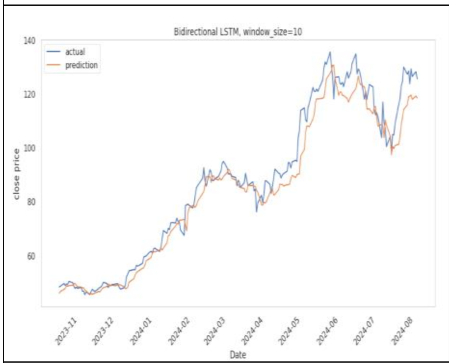
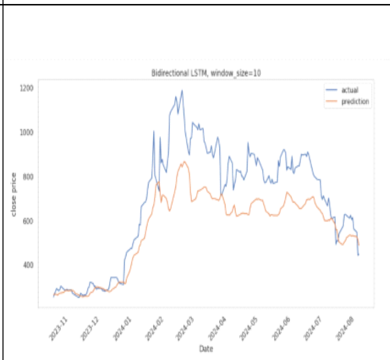
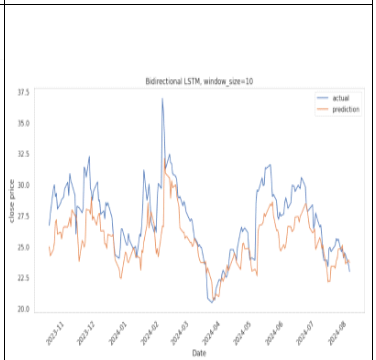
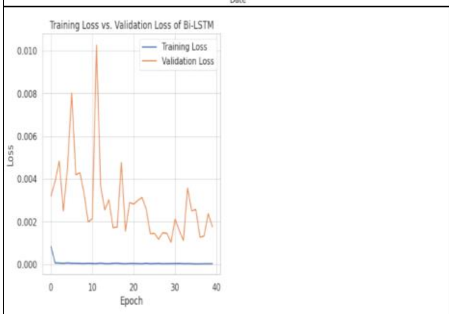
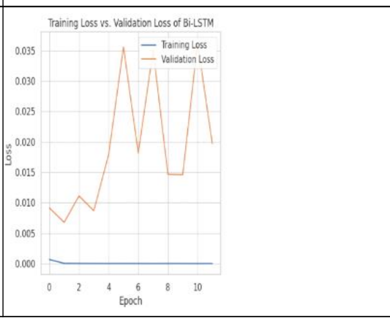
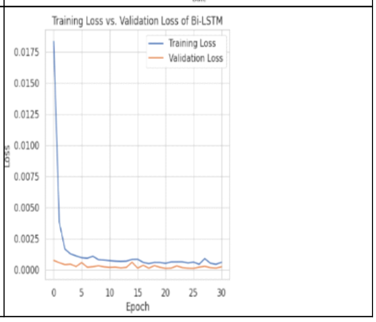
3.13.3 August 28

By the third day, the model's performance on the NVDA stock price showed even more improvement as the prediction closely aligned with the actual peaks and troughs of the price. Even the validation loss was clearly reduced by the latter epochs, indicating clear improvement compared to the previous day.

The model’s performance in predicting SMCI prices indicated clear challenges, as the predictions showed a visible lag in capturing actual price movement from the beginning of 2024 to the latter part of the year, which indicates that the model struggled with the high volatility and complex patterns inherent to the SMCI stock. The loss curves also reflect this mismatch between the model's tendency to overestimate troughs and underestimating peaks. The training loss remained stable at a lower value from the beginning of the epochs, indicating that the model learned well from the training data, while validation loss indicated considerable fluctuations throughout the curve and failure to stabilize, which clearly confirms the model's incapability to generalize to unseen data.

Although the model seemed to struggle with sharp fluctuations in the price changes of AI stocks, the predictions clearly showed an ability to replicate the actual price trend closely. The loss curves also provided valuable insight into the model's performance, as both learning and validation loss curves remain at a lower value. The occasional minor spikes in the validation curve indicate the model's discrepancy in predicting larger spikes.

Table 3.9 BI-LSTM Results August 28 stock price prediction curves (top row), training loss vs validation loss (bottom row)

NVDA	SMCI	AI
		
		

3.14 GRU

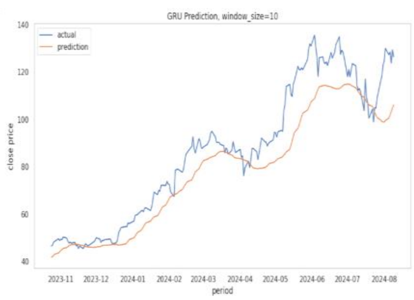
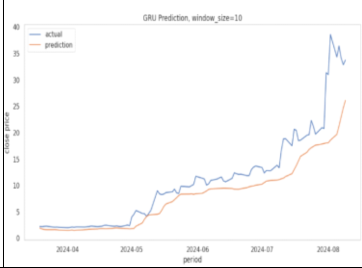
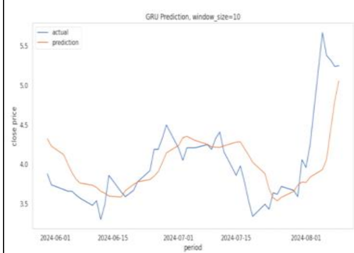
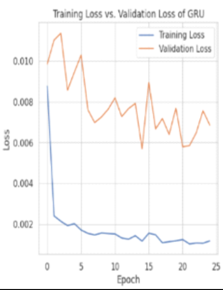
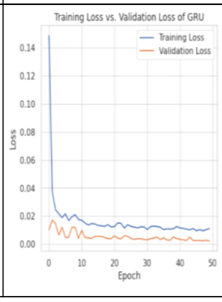
3.14.1 August 26

Compared to the LSTM models, the GRU model showed a moderate fit between the actual and predicted values. The prediction curve only followed the general trend of the actual price without any of the dynamics of the price change. Hence, this smoother representation of the price change made it difficult to capture the daily volatility of the price. This was also visible through the learning curves, as the validation loss showed more variability, with occasional high spikes indicating the inconsistency of the model's performance.

As for ASTS, the model performed well in identifying the overall trend of the stock price movement. However, similar to the NVDA price prediction, the model underestimated peaks during significant uprisings. From the loss curves, this can also be visible, as the training loss rapidly decreased and stabilised while the validation loss shows notable fluctuations at a higher loss rate, which clearly indicates the model's inability to capture the noise of the ASTS stock price.

The model showed some capabilities in predicting the price direction of LUNR compared to the other two stocks, although certain lags can be identified as the prediction curve underestimated the recent peaks while overestimating troughs in the mid-July 2024 period. The loss curves showed clear improvement compared to the loss curves of NVDA and ASTS, as both training and validation loss remained at a lower level. However, the validation loss showed some variability and inconsistency.

Table 3.10 GRU Results August 26 stock price prediction curves (top row), training loss vs validation loss (bottom row)

NVDA	ASTS	LUNR
		
		

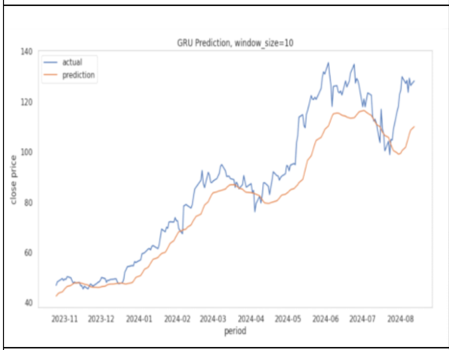
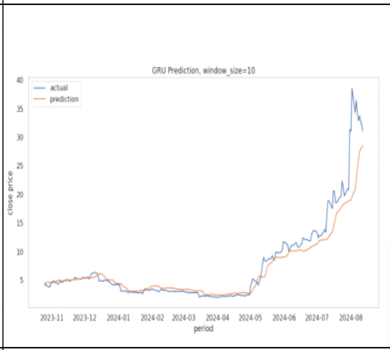
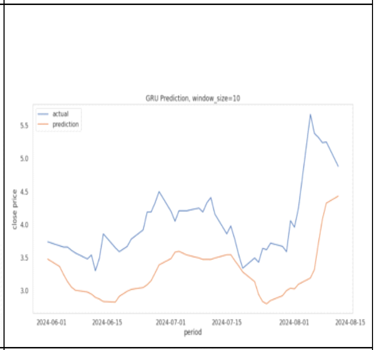
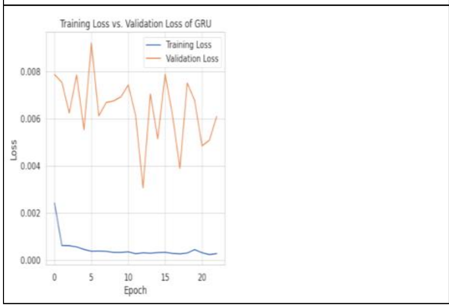
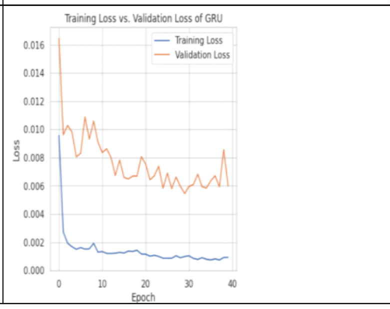
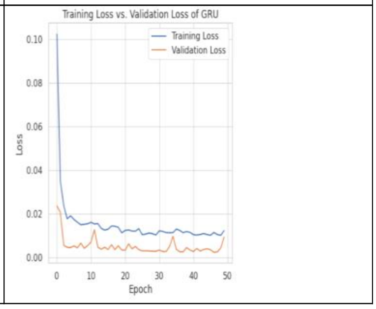
3.14.2 August 27

Unlike LSTM models, the GRU model's prediction performance of NVDA stock price resulted in an even more pronounced lag in predicted values, especially during the sharp upward movement of the price, which indicated that the GRU model struggled to identify the momentum and volatility characterised in NVDA stock price. The loss curves further illustrated this incompetence of the model. On the previous day, the loss curves demonstrated a more consistent decrease, which indicated better performance. However, the current loss curves, especially the validation loss, showed more fluctuations and higher variability through the epochs. This indicates that the model's ability to generalise training data to unseen data has further diminished.

Nevertheless, the model's performance in predicting ASTS stock prices showed considerable improvement compared to the previous day. As depicted by the graph, the predictions closely aligned with the actual values at the beginning of the time period. However, towards the end, the model's predictions lagged slightly behind. This was highly visible during the sharp increase in price in the recent past. Moreover, the training loss curves showcased a sharp downward trend at the first epoch, although validation loss exhibited periodic spikes towards the end of the training process, indicating occasional struggles with generalising the model under more complex and unpresentational price patterns.

The model still displayed noticeable delays in identifying rapid price movements of the LUNR stock. Although the prediction line showed adequate alignment with the actual values, which were highly deviated on the previous day. As for the learning curves, validation loss showed certain fluctuations at a lower rate, indicating occasional overfitting and variance in the model generalisation. The validation loss slightly increased by the last epoch, suggesting the diminished predictive power when encountering more varied and complex situations.

Table 3.11 GRU Results August 27 stock price prediction curves (top row), training loss vs validation loss (bottom row)

NVDA	ASTS	LUNR
		
		

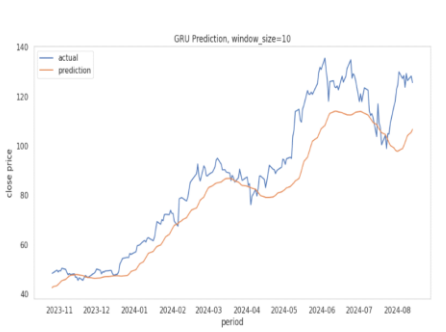
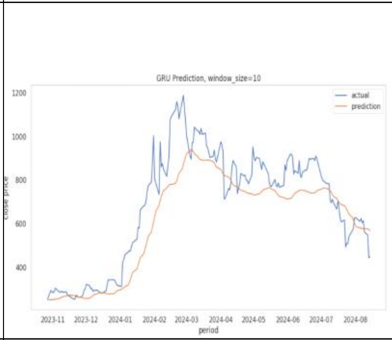
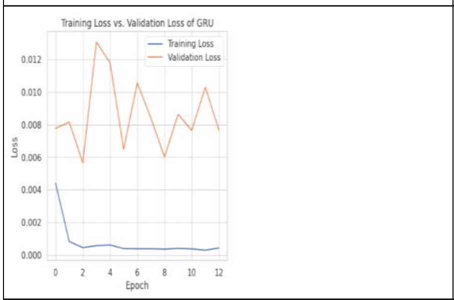
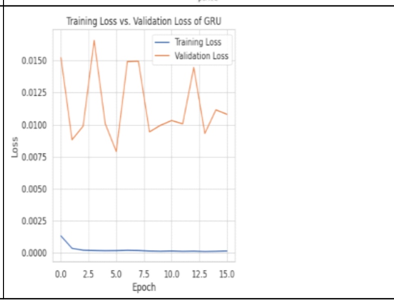
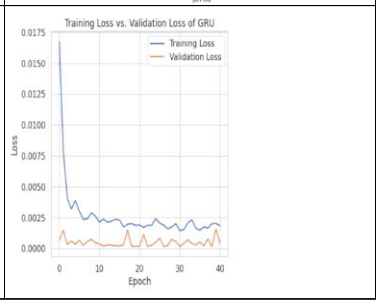
3.14.3 August 28

Similar to the prediction from the previous day, the model showcased the ability to represent the direction of the stock price movement in a general form. However, the prediction line exhibited clear lags and a flat representation of the short-term volatility, leading to less precise predictions, especially during periods of rapid price change. This was also demonstrated by the loss curves, as the validation loss stood at a noticeably higher value, indicating clear issues with the model in identifying peaks and troughs of the NVDA stock price.

As for the SMCI, the model was able to capture the long-term trend and the general direction, but the model failed to replicate the amplitude of price change properly, which was particularly pronounced during periods of rapid price changes. The loss curves also highlighted this lag in the prediction as the validation loss remained at a higher rate with significant fluctuations across the epochs. This suggests that the model is clearly incapable of consistently capturing the market pattern.

Unlike NVDA and SMCI, the model revealed some strength in capturing the overall trend of the AI stock price movement. However, it also highlighted limitations in accurately predicting finer fluctuations while also struggling with the amplitude of the fluctuations, where clear overestimations of troughs and underestimations of peaks are visible due to the lagging of the prediction line. However, the loss curves remained relatively low, indicating that the model's performance on new data is consistent but not improving significantly beyond a certain point. This indicates that even though the model has identified the general direction of the price movement, it struggles to adapt to more complex and less frequent price movements.

Table 3.12 GRU Results August 28 stock price prediction curves (top row), training loss vs validation loss (bottom row)

NVDA	SMCI	AI
		
		

3.15 Performance Evaluation

3.15.1 August 26

According to the error metrics of each model on predicting the next day price of each of the top three meme stocks, each model revealed distinctive differences in their predictive accuracy. In terms of NVDA, the single-layered LSTM model showed the best performance with the lowest RMSE (4.872), MAE (3.524) and MAPE (0.037), indicating a high level of accuracy in its predictions. This model's ability to closely track NVDA's price movement suggests that it can easily interpret the stock's long-term patterns. On the other hand, the GRU model struggled the most with NVDA price prediction as the error metrics reflected a higher error rate as RMSE (11.331), MAE (8.757) and MAPE (0.093), which reflects a significant prediction error and notable deviation from the actual price. As for ASTS, the single-layered LSTM model remained as the top-performing model with lower error rates as RMSE (1.9.3), MAE (0.827) and MAPE (0.093). This suggests that the model is more adaptable to stocks with shorter trading histories and more volatility. Similar to NVDA, the GRU model demonstrated the worst performance for ASTS. The error metrics showed higher RMSE (4.548) and MAPE (0.238), highlighting relative inefficiency in predicting the stock price movement accurately.

When analysing LUNR, all the models exhibited uniform performance with relatively minor differences in error metrics. The single-layered LSTM model continued to outperform other models with low error rates. Although the differences are minor, the single-layered LSTM model demonstrated its superior capability in predicting the price of novel highly volatile meme stocks. In contrast, the GRU model consistently showed poor performance. However, the errors seemed to be only marginally higher, indicating a reasonably balanced prediction accuracy across all models.

Overall, the single-layered LSTM model consistently proved to be the best-performing model, especially for NVDA and ASTS. The GRU model generally underperformed across all three stocks, particularly for NVDA and ASTS. As GRU demonstrated higher deviations from the actual value. Meanwhile, the Regular Stacked and Bi-LSTM models showed intermediate performance which managed to capture some of the essential patterns.

Table 3.13 August 26 Models' Performance Comparison

	Single Layerd			Regular Stacked			Bi-LSTM			GRU		
	NVDA	ASTS	LUNR	NVDA	ASTS	LUNR	NVDA	ASTS	LUNR	NVDA	ASTS	LUNR
RMSE	4.872	1.903	0.423	6.86	2.141	0.419	7.088	3.349	0.435	11.331	4.548	0.433
MAE	3.524	0.827	0.286	5.042	0.954	0.306	5.441	1.706	0.291	8.757	2.868	0.301
MAPE	0.037	0.093	0.068	0.051	0.104	0.074	0.057	0.118	0.07	0.093	0.238	0.072

3.15.2 August 27

Similar to the previous day's performance, the single-layered LSTM model continued to demonstrate superior performance in predicting NVDA's closing price with the lowest RMSE (4.533), MAE (3.244), and MAPE (0.034). Meanwhile, GRU exhibited the poorest performance with the highest RMSE (10.566), MAE (8.064), and a relatively high MAPE of (0.085). this indicated a significant difficulty of the GRU model in identifying long-term dependencies and complexities in NVDA's price data.

For ASTS, both the single-layered LSTM model and the Bi-LSTM model achieved the lowest RMSE (2.244), with the single-layered LSTM model showing slightly better performance in MAE (0.95) compared to Bi-LSTM (1.026). This indicates that the single-layered LSTM model was marginally more precise than the Bi-LSTM model in prediction. Meanwhile, the regular stacked LSTM model showed slightly higher error rates, indicating a fair but less optimal fit for the price data. Moreover, GRU continued to underperform with considerably higher error rates.

For LUNAR, a single-layer LSTM model continued to demonstrate higher performance across all models with the lowest RMSE (0.404), MAE (0.27) and MAPE (0.064). This consistently low error rate indicates the model's robustness in predicting relatively new stocks. The regular stacked and Bi-LSTM models displayed a slightly higher error rate but were relatively close to the performance of the single-layer LSTM, indicating a decent predictive capability. However, GRU continued to underperform with a high error rate, highlighting its struggle in effectively capturing intricate price movements.

Table 3.14 August 27 Models' Performance Comparison

	Single Layerd			Regular Stacked			Bi-LSTM			GRU		
	NVDA	ASTS	LUNR	NVDA	ASTS	LUNR	NVDA	ASTS	LUNR	NVDA	ASTS	LUNR
RMSE	4.533	2.244	0.404	8.131	2.664	0.414	9.035	2.244	0.44	10.566	2.827	0.872
MAE	3.244	0.95	0.27	6.463	1.173	0.283	6.481	1.026	0.299	8.064	1.328	0.76
MAPE	0.034	0.092	0.064	0.067	0.125	0.067	0.066	0.119	0.071	0.085	0.146	0.182

3.15.3 August 28

By the third day, the regular stacked LSTM model demonstrated better performance in achieving lower RMSE (4.868), MAE (3.411) and MAPE (0.036). which suggested that the regular stacked LSTM model performed better in predicting the closing price of NVDA. Single-layer LSTM model also performed well with slightly higher error metrics such as RMSE (5.664), MAE (4.179) and MAPE (0.043), making it a decent alternative for price prediction. However, Bi-LSTM and GRU models underperformed for NVDA, with GRU showing the highest RMSE (11.816) and MAE (9.07), indicating a significant difficulty in fitting the data.

In the case of SMCI, all the models struggled to accurately predict stock prices with relatively high error rates. Out of all the models, the Bi-LSTM model had the highest RMSE (165.433), MAE (131.25), and MAPE (0.162). The regular stacked LSTM model followed closely with similar high error rates. Even though the performance is not commendable, the single-layered LSTM displayed slightly better performance than the other two LSTM models, followed by the GRU model.

For the AI stock price prediction, all the models showed similar, acceptable performance. Out of them, single single-layered LSTM model stood as the best-performing model with lower RMSE (1.6), MAE (1.197) and MAPE (0.044), followed by regular stacked LSTM and BI-LSTM with competitive error metrics, making them the next most viable options for price prediction of AI stock. Although GRU stayed relatively effective, it showed particularly higher error metrics compared to the LSTM model.

Table 3.15 August 28 - Models' Performance Comparison

	Single Layerd			Regular Stacked			Bi-LSTM			GRU		
	NVDA	SMCI	AI	NVDA	SMCI	AI	NVDA	SMCI	AI	NVDA	SMCI	AI
RMSE	5.664	112.726	1.6	4.868	158.366	2.061	7.534	165.433	1.91	11.816	22.695	2.008
MAE	4.179	86.946	1.197	3.411	124.121	1.471	5.196	131.25	1.606	9.07	92.582	1.619
MAPE	0.043	0.111	0.044	0.036	0.156	0.051	0.053	0.162	0.063	0.095	0.128	0.063

Over the course of three days of testing, the single-layered LSTM model consistently demonstrated superior performance in predicting meme stock prices. Additionally, the regular stacked and Bi-LSTM models exhibited variable effectiveness, with one model occasionally outperforming the other. In contrast, the GRU model consistently delivered the poorest results.

3.16 Real Price Prediction

For three consecutive days, from August 26 to August 28, 2024, the model was trained and tested on historical data and technical data for the selected three meme stocks. The results of this evaluation are as follows:

3.16.1 August 26

For NVDA, the single-layered LSTM model predicted the price for August 27th as 122.50, which was the closest to the actual closing price of 128.3, which indicates the model's strong predictive ability compared to other models. The regular stacked LSTM model predicted 118.55 while the Bi-LSTM model predicted 115.59, both of which are vaguely farther from the actual price yet relatively accurate. The GRU model demonstrated the worst performance by predicting the closing price as 109.28, indicating the model's inability to capture NVDA's price movement accurately.

For ASTS, the single-layered LSTM model predicted the closest price to the actual price as 30.19. The Bi-LSTM model also performed reasonably well, with a prediction of 28.64. The regular stacked LSTM and GRU models predicted 27.26 and 26.13, respectively. Both models slightly underestimated the actual price.

For LUNR, all models showed relatively close price predictions to the actual value of 4.88. Single-layered LSTM and GRU were able to predict the closest price to the actual as 5.13 and

5.05, respectively. Regular stacked and Bi-LSTM models slightly overpredicted the price as 5.53 and 5.19, although still remained close to the actual value.

Table 3.16 Prediction Comparison with Actual - August 26

	Single	**Regular Stacked**	**BI-LSTM**	**GRU**	**Real**
NVDA	122.499	118.552	115.590	109.281	128.300
ASTS	30.190	27.262	28.635	26.130	31.110
LUNR	5.139	5.536	5.195	5.057	4.880

3.16.2 August 27

For NVDA, the single-layer LSTM model performed exceptionally well in predicting the closing price of August 28 by aligning its prediction to the actual price of 125.67, which was almost identical to the actual price of 125.61. Despite the single-layer LSTM model's exceptional performance, the rest of the models underperformed in price prediction. The regular stacked price was 115.69, the Bi-LSTM was 111.988, and the GRU was 110.07, all of which significantly undervalued the price predictions of the stock. Suggesting that complex models may struggle with complex price predictions like NVDA.

Table 3.17 Prediction Comparison with Actual - August 27

	Single	**Regular Stacked**	**BI-LSTM**	**GRU**	**Real**
NVDA	125.670	115.698	111.988	110.076	125.610
ASTS	27.833	28.555	29.655	29.413	27.870
LUNR	5.299	5.357	5.236	4.431	4.670

For ASTS, the predictions were slightly more varied. The single-layered LSTM model predicted the price consistently close to the actual price of 27.83, exhibiting its strong predictive capabilities. The regular stacked LSTM and GRU models also performed reasonably well with predictions of 28.55 and 29.41, respectively. In comparison the Bi-LSTM model, which overestimated the price the most as 29.66, suggesting that although it may capture some of the stock’s volatility but struggled with precision compared to simpler models.

In terms of LUNR, all models displayed similar predictions with slight differences; out of the predictions, Bi-LSTM and single-layer LSTM predictions were the closest to the actual value, 5.23 and 5.29, respectively. Meanwhile, the regular stacked LSTM model slightly overvalued the price as 5.35, and GRU slightly undervalued the price as 4.43, both of which are generally still closer to the actual value.

3.16.3 August 28

Similar to the previous days for NVDA, the single-layer LSTM model predicted a price of 116.50, which was the closest to the actual value of 117.589. The BI-LSTM model also presented a closer value to the actual price of 118.48. The rest of the models performed very poorly in price prediction. The regular stacked LSTM overestimated the price by several points to 126.46, and GRU predicted the lowest price to be 106.68, which was the farthest from the actual value.

For SMCI, all models' predictions were considerably far apart from the actual value. Out of the predictions, the single-layered LSTM model predicted a price of 468.32, which was the closest to the actual value of 448.82. The Bi-LSTM prediction was even higher, at 488.44. The regular stacked LSTM and GRU models significantly overestimated the price, at 539.50 and 568.65, respectively.

For AI stock, all the LSTM models performed substantially well in predicting the closing price for August 29. The single-layered LSTM model prediction was 23.25, which was the closest to the actual value of 23.05. The regular stacked and Bi-LSTM models had similar predictions of 23.83 and 23.72, both of which are only slightly higher than the actual prices. Unlike the GRU model, its prediction of 25.20 was the farthest from the actual value.

Table 3.18 Prediction Comparison with Actual - August 28

	Single	**Regular Stacked**	**BI-LSTM**	**GRU**	**Real**
NVDA	116.505	126.461	118.482	106.680	117.589
SMCI	468.328	539.504	488.448	568.655	448.820
AI	23.257	23.832	23.723	25.204	23.059

Overall, across all three days, the single-layer LSTM model emerged as the most reliable and effective model, offering a balance between simplicity and predictive accuracy. Meaning, the model achieved highly accurate predictions with a lower number of training parameters that resulted in reduced training time, lower memory usage and great model stability during training. This model delivered strong and consistent results for stocks like NVDA and ASTS, a stock which has extensive historical data, while also demonstrating strong performance with LUNR, which has a shorter public trading history. This adaptability reflects the single-layered LSTM model's ability to adapt to stocks with well-defined historical patterns as well as more recent stocks with limited data, effectively capturing core temporal patterns without overfitting to noise, even when data availability is limited.

Despite being more complex, the regular stacked LSTM and Bi-LSTM models did not succeed in utilising their additional layers to improve predictions significantly. In fact, they often introduced more variance and instability to the predictions. Especially with more volatile stocks like SMCI. The added complexity may have led to overfitting since the models captured noise and short-term fluctuations rather than meaningful trends. Furthermore, the Bi-LSTM model's bidirectional learning has failed to add significant value, as the meme stock price prediction data relies heavily on sequential, forward-looking patterns.

Compared to the LSTM models, the GRU model generally underperformed in predictive accuracy, making it the least efficient model in comparison. This could be due to the model's aggressive gating mechanism, which may have discarded useful temporal information, making it less suitable for datasets where subtle shifts in sentiments and technical trends act as key factors. This proves the view that not all recurrent architectures are equally suited for sentiment-sensitive financial predictions.

Therefore, the findings suggest that in sentiment-driven short-term stock price predictions, especially for meme stocks, a simpler model like the single-layered LSTM model is more preferable since it strikes a balance between model generalisation, training efficiency, and predictive reliability. Especially when working with highly volatile stock data that is inherently noisy and dynamic. Further, in real-world applications such as lightweight forecasting systems in retail trading, the single-layered LSTM model offers a more interpretable and resource-efficient solution without diminishing performance.

4 CHAPTER FOUR

Discussion and Analysis

4.1 Introduction

This chapter discusses and analyses the results from the sentiment-driven meme stock price prediction results. It evaluates the findings from the WSB sentiment analysis and the predictive performance of the selected RNNs, namely, single-layered LSTM, regular stacked LSTM, Bi-LSTM and GRU. This discussion considers the predictive accuracy of each model across multiple days and stocks and reflects how sentiment extracted from Reddit's WSB subreddit reflected the direction of the actual stock price direction. Particular attention is given to the relationship between model complexity and prediction stability in the context of volatile, retail-driven assets. The chapter also highlights key observations regarding model strengths, weaknesses and the unique challenges in forecasting sentiment-influenced stock movements, ultimately identifying the most suitable method relying on the results in the niche but growing area of financial forecasting.

4.2 WSB Sentiment Analysis

Several notable trends emerged on the sentiment analysis of the WallStreetsBets discussion (26/08/2024 –28/08/2024). The analysis reflected retail investors' collective mindset and preferences within the community. One of the most dominant trends portrayed on WSB was the retail investor's strong enthusiasm for artificial intelligence and upcoming technology-related companies. Especially for those who have been perceived as leaders or beneficiaries of the current AI boom. For example, Nvidia (NVDA), the world's largest graphics processing unit (GPU) producer, consistently emerged as the most mentioned stock in all three days. This growing hype around NVDA could be due to its GPUs focused on supercomputers that are primarily used for generative AI. (Merlo, 2023). This trend suggests that retail investors on WSB are driven by optimism and speculative excitement surrounding AI and technology. Therefore, retail investors often prioritise these stocks over others due to their perceived alignment with the ongoing global technological boom. This further emphasises the herding behaviour of retail investors who are driven by perceived growth potential and market narrative, where they flock towards trending stocks relating to AI and technology.

However, alongside the consistency for stocks like NVDA, there was a notable increase in sentiment volatility for stocks like SMCI, ASTS, LUNR and AI over the three days. While mentions and sentiment intensity for NVDA grew daily, indicating escalating interest and enthusiasm. Conversely, for other stocks, the mentions and sentiments fluctuated with varying levels of bullish, bearish and neutral sentiment. These varying sentiments reflect the speculative nature of retail investors. Since ongoing discussions on WSB are constantly influenced by the latest narratives, news, and market trends rather than any rational reasoning. As an example, other stocks, such as SMCI, ASTS, and AI stocks, displayed inconsistent sentiment patterns. Despite being a trending topic, this highlighted the scepticism or mixed opinions on them. This emphasises the reactive and opportunistic behaviour of retail investors who rally around meme stocks that prioritise short-term trading signals and sentiment-driven price movements over long-term fundamentals.

Further, the sentiment analysis highlighted a clear difference between the tech sector and other sectors. Non-tech stocks such as Boeing (BA) generally received more bearish and neutral sentiment, indicating a lack of enthusiasm toward such stocks compared to tech-related stocks. This strong difference highlights the community's preference for high-risk, high-reward stocks that resonate with current market trends. Especially the way the Redditors get on board to invest in novel stocks such as ASTS SpaceMobile (ASTS), a company that aims to build the first space-based cellular broadband network that is directly accessible by standard mobile phones and Intuitive Machines (LUNR), a novel company focused on space exploration and space technology rather than established companies in traditional sectors indicate that the investors' behavior is driven by speculative potential rather than a balanced view of the broader market.

4.3 Meme Stock Price Prediction Models

The performance of the considered models - single-layered LSTM, regular stacked LSTM, Bi-LSTM and GRU- showed notable differences across the top three stocks detected from the WSB sentiment analysis from August 26 to August 28. The model showed noteworthy differences in all five stocks (NVDA, ASTS, LUNR, SMCI). The single-layer LSTM model consistently outperformed other models in predictions, delivering lower RMSE and MAE scores. On the other hand, regardless of their complexities, the regular stacked LSTM and Bi-LSTM showed mixed results. Although Bi-LSTM performed slightly better in predicting LUNR prices, it struggled with more volatile stocks like SMCI. While GRU generally underperformed in comparison, showing higher error rates and highly varied predictions across most stocks, especially for NVDA and SMCI, across all three days. Hence, a pattern could be identified where the single-layer LSTM model showed more consistency and less prediction variance. In contrast, the deeper models introduced more volatility, especially when handling highly volatile stocks like SMCI.

In terms of computational complexity, the deep learning models based on RNNs employed in this study are inherently computationally intensive due to their sequential processing nature. For each training sample, the time complexity of a standard LSTM is approximately $O(n * h^2 + n * h * d)$, where the n represents the number of time steps, h represents the number of hidden units, and d indicates the input dimension. Stacked LSTMs increase the computational cost linearly with the number of layers, whereas Bi-LSTM effectively double it by processing input sequences in both forward and backwards directions. Moreover, these models require the maintenance of hidden and cell states across sequences and epochs, which increases memory consumption. As a result, training times tend to get longer, and the models demand higher computational resources compared to simple conventional machine learning methods. Regardless of the higher computational needs, their uses were justified given the sequential and sentiment-driven nature of the meme stock data, where capturing temporal dependencies and complex non-linear patterns was highly essential.

Out of all the models considered, the single-layered LSTM model consistently performed well for almost all the stocks. This can be attributed to the simplicity and effectiveness of capturing essential time series patterns. As an example, for NVDA, single-layered LSTM was the only model that was able to predict the next day closing price approximately closer to the actual value, while the rest of the models were drastically apart from the reality, as the single-layered architecture allowed the model to focus on critical and recent trends without introducing unnecessary complexities that result in overfitting. Similarly, the model performed well with LUNR, a stock that started public trading fairly recently. The model was able to capture the patterns effectively despite its shorter historical data, which demonstrates that the model can generalise well to both mature stocks with rich data as well as newer stocks with more recent trends.

Despite the added complexities of additional layers, both regular stacked and Bi-LSTM models did not significantly outperform the simpler single-layered LSTM model. The additional layers introduced overfitting, especially when the data set, such as meme stock data, does not require complex representations as this can lead to more variance and reduced generalisation, particularly in the case of NVDA where there is a long history of data that can contain conflicting signals, with older data patterns becoming less relevant in the current AI-driven market environment making the deeper models overfitting and leading to a more significant variance in predictions.

Although the GRU model is known for its efficiency and ability to handle sequential data, its simple architecture is not good enough to capture the subtle patterns of meme stocks. Especially when the stocks exhibited high volatility or erratic behaviour, especially for meme stocks that are highly driven by the retail investors' herding behavior, influenced by the current AI hype. Hence, the inherent reduced capacity of the model to handle complex sequences or handle intricacies of volatile stock behaviour greatly impacted its underperformance.

When considering both the WSB sentiment analysis model and next-day price prediction RNN models, both models provided separate inputs to the final study. The sentiment analysis of daily discussions on WSB primarily served to identify the most popular meme stocks and obtain the potential sentiment of retail investors towards them. The sentiment analysis was not directly incorporated into the next-day price-prediction models. Instead, the sentiment analysis functioned as a filtering mechanism to direct the predictive models to the most relevant meme stocks based on active community discussions. This approach allowed the models to prioritise more reliable aspects of the stock analysis, namely historical and technical indicators, avoiding the noise and unpredictability of the inconsistent sentiment.

For retail investors, when online discussions around a particular stock become overwhelmingly enthusiastic, especially when the stock is at its peak, it often signals that it has garnered the attention of mainstream investors, similar to what occurred with GameStop (GME), where the retail investors began to sell their stocks when the stock went viral. Prices gradually started to decline (Yahya, et al., 2021). In the case of NVDA, as shown by the sentiment analysis, it has become the hottest topic on WSB, with highly bullish sentiment. However, the stock prices were declining, and the RNN models effectively captured this trend as they based the predictions on historical price trends and technical indicators, providing a more objective view.

Hence, had the overall bullish sentiment been incorporated directly into the model, it could have skewed the prediction in favour of an upward trend. By excluding sentiment from the model, the predictions offered a more balanced outlook, giving investors room to make rational decisions based on both sentiment analysis and price predictions. These findings suggest that for meme stock forecasting where data can be noisy, sparse, and sentiment-driven driven simpler recurrent models like the single-layered LSTM often outperform deeper and more complex alternatives. Their lower computational overhead, combined with more stable predictions, makes them especially well-suited for highly volatile and retail-driven financial assets.

5 CHAPTER FIVE

Impact of Energy Prices and Macroeconomic Variables on GDP Prediction UK: Machine Learning Approach with SHAP Interpretability

5.1 Introduction

This chapter presents the analysis of the second part of the study, which investigates the impact of energy prices and key macroeconomic variables on the prediction of the UK's GDP using machine learning techniques. the study applies four predictive models, such as LR, SVR, GBM and RF, to assess the effectiveness of AI-driven methods in forecasting GDP from structured economic data spanning from 1990- 2018. In addition to evaluating the models' performance using standard error metrics, this chapter incorporates SHAP to interpret the influence of each feature on the predictions. This approach provides both predictive accuracy and model transparency, offering insights into which macroeconomic variable has the most significant effect on the GDP of the UK.

5.2 Methodology Overview

The study utilises LR, SVR, RF, and GBM models to predict the UK's GDP based on macroeconomic data from various sources (Figure 5.1). The linear regression model was selected due to its uncomplicated implementation and ability to provide clear insight into relationships between GDP and the predictors. It is a baseline model for comparing with more complex models (Kitov, 2008). SVR was employed in the study as it can easily handle non-linear relationships between variables with kernel functions (Xu, et al., 2014), making it ideal for identifying non-linear relationships between complex economic data. The RF model was employed to capture complex interactions among several variables without explicit

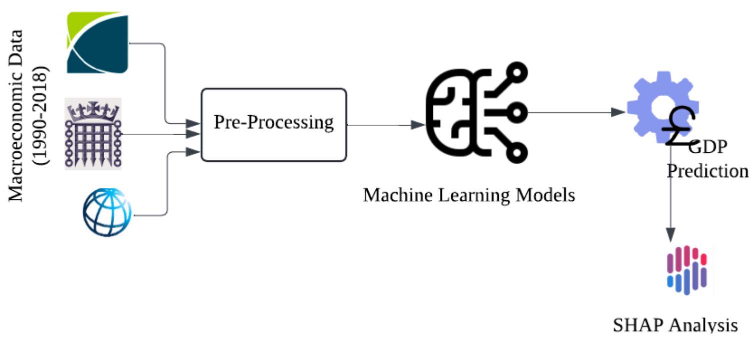


Figure 5.1 UK GDP Prediction Methodology Overview

specifications, making it suitable for multi-factor GDP prediction. Similarly, GBM was utilised to capture complex patterns with greater accuracy as it can sequentially correct errors from previous iterations using the boosting technique (Laygo-Matsumoto & Samonte, 2021). Further, SHAP was utilised to interpret the impact of each macroeconomic variable on GDP predictions to gain more precise and transparent results.

While the previous chapter employed deep learning techniques such as single-layered LSTM, stacked LSTM, Bi-LSTM and GRU to predict the next-day closing price of meme stocks, this chapter mainly utilises conventional machine learning models to predict the GDP of the UK. This shift in methodology solely reflects on the nature of the dataset and the prediction objective of each case. Since meme stock prediction includes highly volatile and sequential time series data that are influenced by sentimental and technical indicators RNNs are able to accurately capture the trend in meme stock prices. In contrast, GDP prediction is based on structured, low-frequency macroeconomic indicators. conventional machine learning models outperform deep learning methods due to the smaller dataset size and clear feature-targeted relationships without overfitting. Therefore, these methodological differences explore the effectiveness of data-driven approaches in predicting complex economic and financial variables using suitable models tailored to the characteristics of the data.

5.3 Research Paradigm

The research paradigm of this study is grounded in positivism. Positivism presumes that reality is objective and can be measured through empirical data and observations (McCloskey, 2022). This aligns with the quantitative, data-driven approach of the study since the goal is to explore relationships between macroeconomic variables and GDP using mathematical models and statistical techniques.

The ontological stance is that economic outcome (GDP) is influenced by observable and measurable macroeconomic variables such as energy prices, inflation, unemployment rates, and REER. Hence, this study assumed that these variables have objective relationships independently of the researcher's perception and that their impacts can be objectively observed by applying machine learning models to historical data.

In terms of epistemological position, this study followed an empirical stance since the relationship between macroeconomic variables and GDP is constructed through a collective analysis of real-world data from the UK (1990-2018). The study was conducted using machine learning techniques to discover complex patterns between macroeconomic variables and GDP, following a quantitative research method. Further, the study followed a deductive approach since the research was built on the hypothesis that the interconnection between macroeconomic variables and GDP can be utilised in GDP prediction.

This study followed an objective stance by employing machine learning models to interpret the interconnection between input and target variables, where personal biases and subjectivity are minimised. Performance metrics such as RMSE, MAE, MAPE and R^2 score were applied to ensure the objective stance. At the same time, SHAP contributed to the transparency and ethical integrity of the research by providing clear, interpretable insight into the way the macroeconomic variable influenced the GDP prediction process, which makes the models more understandable and reliable.

5.4 Theoretical Framework

This study's theoretical foundation was built on several critical economic and statistical theories, which provided a framework for understanding the complex non-linear relationships between GDP and macroeconomic variables.

The interrelationship between GDP and macroeconomic variables is mainly based on 'Keynesian Economic Theory', which emphasises the role of aggregated demand in determining the overall trend of economic activities, particularly in the short run (Jahan, et al., 2014). This theory argues that fluctuations in demand are driven by factors such as government spending, consumer behaviour, investments, and trade dynamics, which are key determinants of GDP or economic growth (Jahan, et al., 2014).

Considering this study, macroeconomic variables such as energy prices, inflation and unemployment rate are critical factors that shape aggregated demand. As an example, unemployment has the potential to reduce consumer spending, which results in lowered aggregate demand. This can impact business investments, which eventually constrains GDP growth. Additionally, this theory highlights the role of government intervention in an economic direction through fiscal and monetary policies in stabilising an economy during a highly volatile period. This is particularly relevant for the UK since the government significantly controls volatilities through economic regulation policies. (Koop & Lodge, 2014).

Further, ‘Neoclassical Theory’ plays a vital role in this study. The neoclassical theory focuses on the role of supply and demand in determining prices, production, and distribution in the economy. This theory assumes that economic actors are rational and make decisions based on market signals to maximise utility and profit (Pasinetti, 2000). As an example, in the short term, when energy prices rise, the cost of production will most likely increase, shifting the supply curve to the left, which reduces output and eventually slows down GDP growth.

Machine learning is another important theory applied in this study. This theory involves developing models that can learn from past data patterns and make predictions based on the relationship between variables without explicit programming. In this study, LR provided a baseline understanding of the relationship between variables, while other models (SVR, RF, GBM) provided a more sophisticated technique for capturing complex patterns.

Hence, the theoretical background of this study was built on a combination of neoclassical theory and Keynesian Economic Theory and machine learning theory. Through the combination of traditional economic theories with modern computational methods, this study seeks to offer a precise and interpretable model for GDP prediction.

5.5 Data Acquisition and Extraction

Data for this study was acquired from the Office for National Statistics (ONS) UK, House of Common Library (UK Parliament) and the World Bank for 1990-2018, ensuring accurate and consistent coverage of the key economic indicators relevant to GDP prediction. The GDP data was sourced as gross domestic product in chained volume measures[1], seasonally adjusted in million pounds (£m). Energy price data were obtained through the consumer price index (CPI) for fuel components[2], with real prices based on the year 2010, and the energy price data consisted of solid fuels, gas, electricity, liquid fuels, domestic fuels, motor and oil. Also, the unemployment rate covering individuals aged 16 and over was seasonally adjusted and provided as a percentage[3]. Inflation metrics were derived from the Consumer Price Index, including owner occupiers’ housing costs (CPIH) annual rates for all items with 2015 as the base year[4]. Net migration values[5] were obtained from (Sturge, 2024) report published in

[1] ONS-GDP UK
[2] Consumer price index: Fuel components
[3] Unemployment rate seasonally adjusted
[4] CPIH UK
[5] House of Commons Library briefing paper on Migration Statistics

the UK Parliament – House of Commons Library. Net migration statistics were included in this study to account for the demographic influence on the growth or decline of GDP. Additionally, REER data was sourced from the World Bank database,[6] which offers insight into the UK's trade competitiveness on an international scale. After extraction from different sources, the data was compiled into a single dataset for easy handling. And the dataset was examined for missing values, errors, skewness and consistency.

The following (Figure 5.2) is a visual representation of UK GDP from 1990 to 2018. This time period was specifically chosen to ensure the consistency of data throughout all variables. Further, during this period, several global incidents took place, which caused significant volatility in global economies (Trotsky, 1991), (Wilson, 1994), (Hodson & Mabbett, 2009), (McKibbin & Stoeckel, 2010). According to the graph, the GDP trend for the UK shows a general upward trend, indicating economic growth. However, there is a noticeable dip around 2008-2009 due to the global financial crisis during that period, followed by a recovery period and continuous growth (Hodson & Mabbett, 2009). This indicates the overall resilience and long-term expansion of the UK economy over a span of nearly three decades.

Figure 5.2 UK GDP Trend

[6] World Bank Group - Data

5.6 Data preprocessing

5.6.1 Correlation analysis

As a part of the preprocessing, correlation analysis was conducted on the dataset to help identify the relationship between variables and their impact on GDP. According to the correlation matrix (Figure 5.3), GDP was positively correlated with net migration (0.91), indicating that higher net migration is associated with increased GDP, likely due to the inflow of labour and growth in consumers contributing to the economy. Similarly, all energy prices presented a strong correlation with GDP (Solid fuels – 0.84, Gas – 0.78, Electricity – 0.64, Liquid fuel – 0.72, Domestic fuel – 0.76, Motor & oil – 0.84). this is a clear indication of the significant role the energy sector plays in the UK's economy in various forms, such as high demand for energy, increased industrial and economic activity, the influence of global energy markets on the UK as well as the government's 'Energy Price Guarantee Scheme' that is aimed at managing the impact of rising energy costs on households while maintaining economic stability (Waddams, 2023).

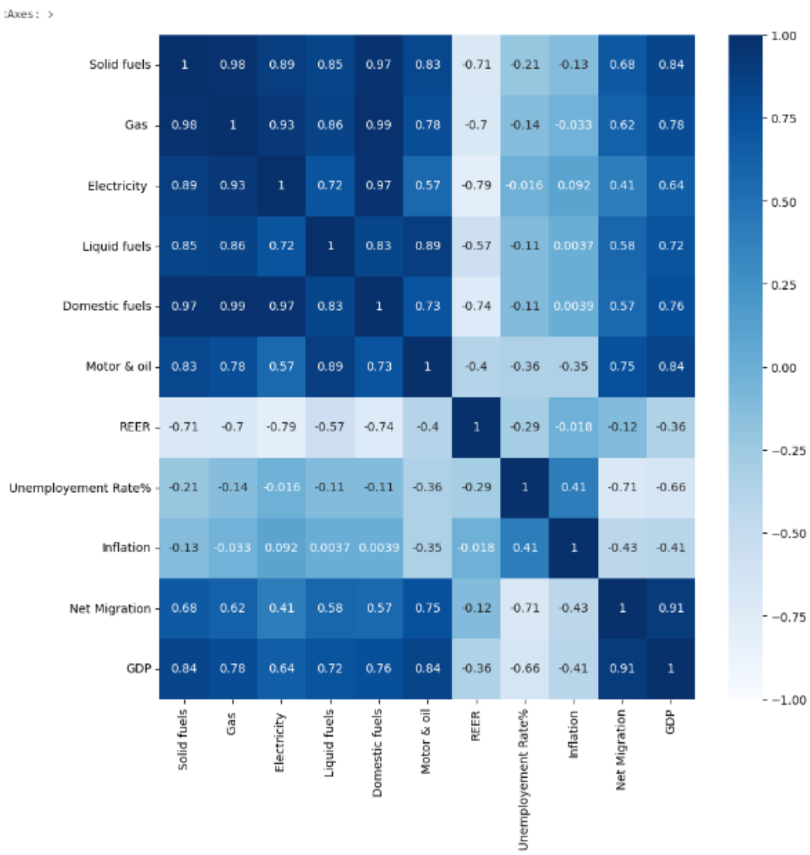


Figure 5.3 Macroeconomic Variables Correlation Matrix

Conversely, GDP showed a strong negative correlation with the unemployment rate (-0.66), highlighting the economic growth impact on job creation, resulting in lower unemployment. The moderate negative correlation between REER (-0.36) and inflation (-0.41) indicated that GDP growth is linked with enhancing export competitiveness, resulting in weaker REER and lower inflation that supports economic stability, and vice versa.

5.6.2 Pair Plot Analysis

To further understand and visually inspect the nature and pattern of the relationships between macroeconomic variables and GDP, pair plots were generated. The interpretation of each variable is as follows:

According to the pair plots of energy variables, solid fuel showed a positive relationship with GDP. The trend indicated the UK economy's reliance on solid fuel. Similarly, GAS and electricity also displayed a positive relationship. For both variables beyond 90 units, GDP indicated a strong upward trend, suggesting that gas and electricity play a critical role in the UK's economy by powering industries, businesses and households, thereby supporting the economy. Domestic fuel also demonstrated a positive relationship, suggesting that domestic fuel use is directly linked with economic productivity. Further, motor oil displayed a strong positive relationship with GDP. This is aligned with the fact that motor & oil consumption critically influences transportation, logistics and industrial activities, which are vital components of economic growth (Enders & Jones, 2016). Liquid fuel also reflected a non-linear relationship; however, between 80-100 units, GDP growth appeared to slowly decline. This may suggest the negative influence that liquid fuel has on GDP after a certain point or the influence of other external factors.

In all six observations, the plots displaced a non-linear (Figure 5.4), and a mostly positive relationship between energy prices and GDP. This confirmed the correlation matrix analysis of the positive relationship between energy price and GDP.

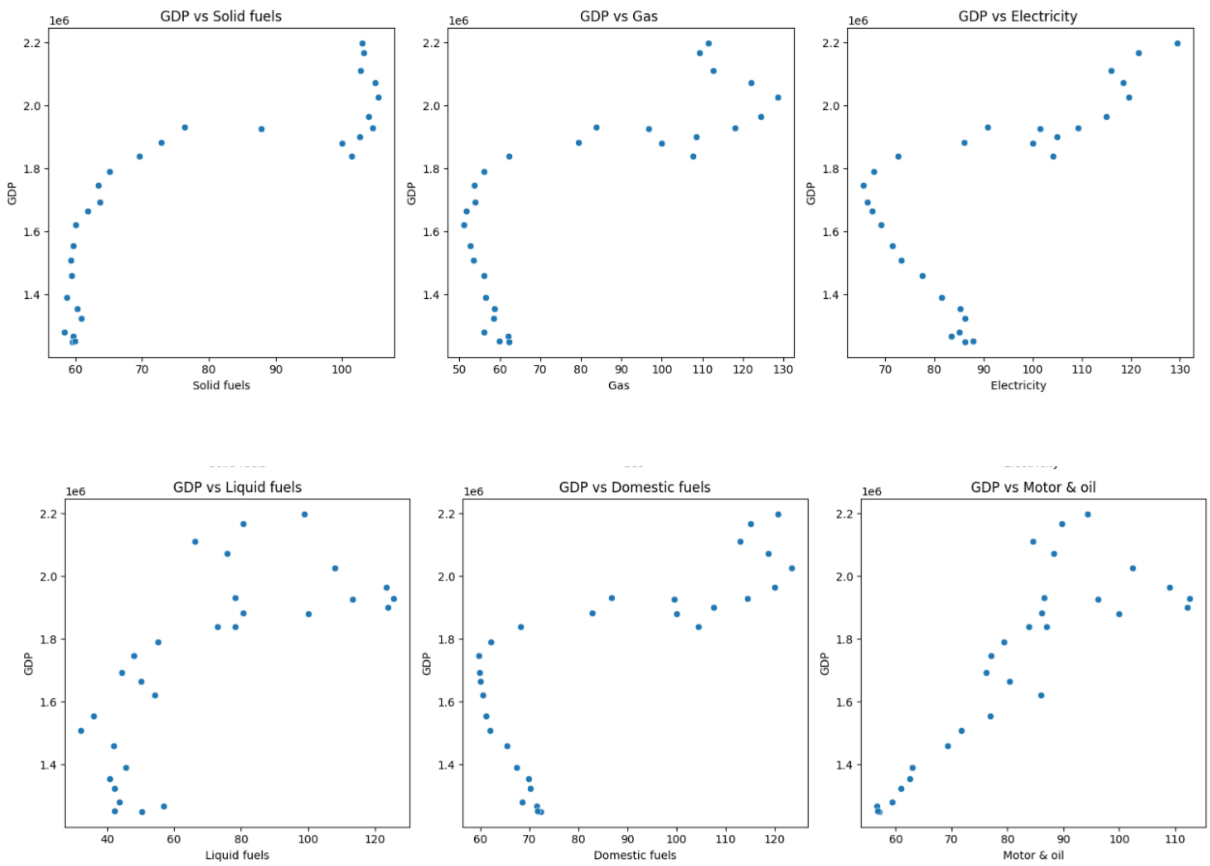


Figure 5.4 GDP and Energy Prices Pair Plots

The REER plot indicated a somewhat non-linear, antagonistic relationship with GDP, indicating that the stronger the REER relative to the trading partners, the more GDP faces a negative impact. This aligns with the economic theory that states that exchange rate movements can significantly impact economic growth (Di Nino, et al., 2011). This means that when the domestic currency strengthens with a higher REER, GDP tends to decline, potentially due to reduced export performance. However, the relationship is not perfectly linear, which implies the complexity of the REER interactions with other macroeconomic variables affecting GDP.

Regarding the unemployment rate, the plot displayed a strong negative relationship with GDP. This demonstrated the traditional economic theory—Okun's Law, which suggests that higher employment rates typically contribute to lower economic outputs. Apparently, when the unemployment rate is around 4-6%, GDP stays higher (above 2 million). Conversely, when unemployment rises to a higher level (7-10%), GDP significantly decreases, falling below 1.4 million.

The plot supports the idea that inflation is a critical factor that influences GDP, as it displays a strong negative relationship. As depicted by the plot, at a lower inflation level (between 1-3%), GDP appears to be relatively high, ranging between 1.8 and 2.2 million. However, as inflation increases beyond 3%, there appears to be a sharp drop in GDP, and by the time inflation reaches 7-8%, GDP drastically drops to the lowest observed level. Therefore, maintaining inflation at a lower level appears to be essential for maintaining strong GDP growth.

Finally, the plot displays a strong positive relationship between GDP and net migration. As net migration increases, GDP appears to constantly rise; when net migration is at a lower level (below 100,000), GDP stays around 1.4-1.6 million. Meanwhile, when net migration rose to a higher level (up to 300,000), GDP appeared to increase steadily, reaching values below 2.2 million. This indicates that higher net migration leads to a large workforce, and consumer demand contributes to the country's production of goods and services and the overall productivity of the economy.

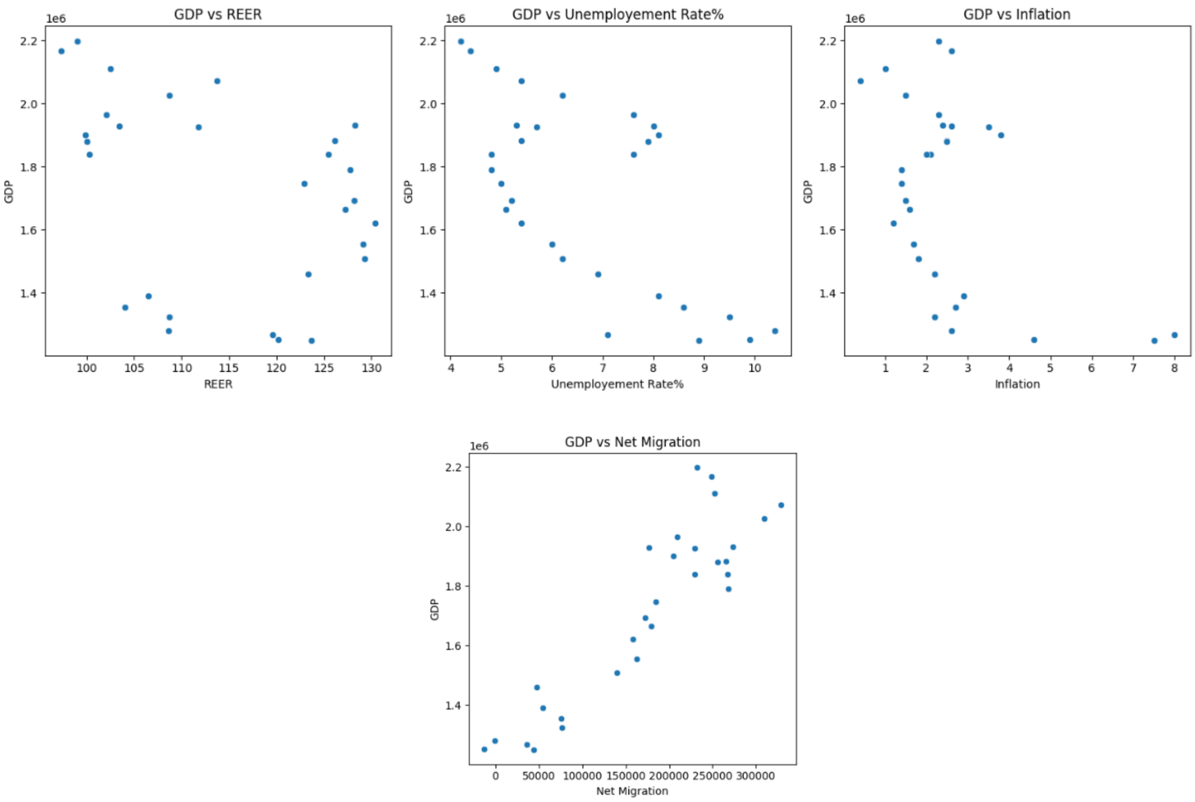


Figure 5.5 GDP REER, Unemployment Rate, Inflation and Net Migration Pair plots

5.6.3 Statistical Analysis and Splitting

Descriptive statistical analysis was conducted on the dataset since it consisted of a wide range of economic indicators and over 29 observations. According to the analysis, the features disclosed significant variabilities. As an example, energy prices showed substantial fluctuations in price levels, especially in gas prices, where the mean value was 79.59, with a standard deviation of 27.86, reflecting notable changes in energy prices over time. And the mean unemployment rate stands at 6.64% with a standard deviation of 1.77%, exhibiting varying employment conditions over the years. Similarly, net migration exhibits highly fluctuating migration patterns. However, inflation shows moderate variability, with a mean of 2.58% and a standard deviation of 1.67%.

These variations underlined the importance of ensuring that training and testing data represent overall data distribution. Hence, random splitting was employed to avoid biases and ensure the model can generalize well to new, unseen data. The dataset was split using the 'train_test_split' function of 'sklearn.model_selection', and the dataset was split, allocating 80% of the data for training and 20% for testing.

	Solid fuels	Gas	Electricity	Liquid fuels	Domestic fuels	Motor & oil	REER	Unemployement Rate%	Inflation	Net Migration	GDP
count	29.000000	29.000000	29.000000	29.000000	29.000000	29.000000	29.000000	29.000000	29.000000	29.000000	2.900000e+01
mean	77.531034	79.586207	91.134483	69.217241	84.765517	81.600000	114.767870	0.066414	0.027034	174758.620690	1.717661e+06
std	19.962663	27.857876	19.301781	29.307045	23.406276	16.714942	11.671087	0.017699	0.016324	96383.406683	2.989216e+05
min	58.300000	51.100000	65.500000	32.000000	59.700000	56.600000	97.302362	0.042000	0.010000	-13000.000000	1.248461e+06
25%	59.900000	56.100000	73.200000	44.200000	65.500000	69.300000	103.398585	0.052000	0.017000	77000.000000	1.458467e+06
50%	65.100000	62.200000	86.200000	56.700000	71.700000	83.800000	113.762294	0.062000	0.023000	185000.000000	1.788931e+06
75%	102.600000	108.500000	105.000000	80.600000	107.600000	89.800000	126.182395	0.080000	0.027000	252000.000000	1.929229e+06
max	105.500000	128.700000	129.500000	125.700000	123.600000	112.700000	130.435556	0.104000	0.080000	329000.000000	2.197841e+06

Figure 5.6 GDP Vs Macroeconomic Variables Statistical Analysis

5.6.4 Data Scaling

As the features in the dataset are individual observations in different scales the dataset was normalized using the 'MinMaxScaler' *(*5.1*)* to ensure the features are comparable, consistent and equally contribute to the machine learning models. This scaler transformed each feature to a scale between 0 and 1. The formula for MinMaxScaler is as follows:

$$X_{Scaled} = \frac{X - X_{min}}{X_{max} - X_{min}} \qquad 5.1$$

X- original feature value X_{min}, X_{max} - minimum and maximum value

5.7 Modelling

For this study, the application of four different machine learning models were explored and compared on diverse economic indicators. The models selected for this study are Linear Regression, Support Vector Regression, Random Forest, and Gradient Boosting Machines.

Each model was trained with both scaled and unscaled data to assess the impact of data normalisation on model performance. Further, hyperparameter tuning was performed for all models besides linear regression in an attempt to optimise the prediction accuracy of each model.

5.7.1 Linear regression

Linear regression is one of the most widely used statistical models in macroeconomic prediction analysis (Elliott, et al., 2008). Here, the model was trained on both scaled and unscaled independent variables (macroeconomic variables) to find the relationship with the dependent variable (GDP) by fitting a linear equation *(*5.2*)* to the observations. Mathematical expression of the model:

$$GDP = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n \tag{5.2}$$

$$GDP = \beta_0 + \beta_1(Solid\ Fuels) + \beta_2(Unemployement\ Rate) + \cdots + \beta_n(Net\ Migration) + \in$$

β_0 - intercept, $\beta_1, \beta_2 \ldots \beta_n$ – coefficient of each independent variable $\in$ - error term.

5.7.2 SVR

SVR is a machine learning model that is mainly used for classification-related tasks. It is an extension of the Support Vector Machines (SVM) model (5.3). The SVR model applies principles of SVM to regression-related problems (Peters, 2001), to enable the model to capture non-linear relationships among variables. By mapping energy prices and other economic variables into a high-dimensional space using the Kernel function, SVR is able to model complex interactions between GDP and economic indicators. Due to the ability to handle non-linear relationships, SVR can be identified as a robust choice for economic forecasting where high-level complexities are common (Xiang-rong, et al., 2010).

The SVR model was trained on both scaled and unscaled data. Both models were trained with a Radial Basis Function (RBF) kernel. Additionally, hyperparameter tuning was applied using GridSearchCV to find the best hyperparameters for the SVR model. The regularization parameter grid included – C :(1, 10, 100), the kernel Coefficient γ(‘Scale’, ‘auto’), the epsilon-tube width $\in$ (0.01, 0.1, 0.2), the kernel type (‘rbf’, ‘linear’), cross-validation was set to 10 (c=10), and scoring was set to negative mean squared error. The best model was selected based on cross-validation. The selected best set of hyperparameter features by the GridSearchCV was {'C': 1, 'epsilon': 0.01, 'gamma': 'auto', 'kernel': 'rbf'}.

$$f(x) = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)K(x_i, x) + b$$

5.3

$K(x_i, x)$ – kernel function, α_I, α_i^* - Lagrange multipliers x_i - support vectors, b – bias term

5.7.3 Random Forest

Random forest *(*5.4*)* is an ensemble learning method that combines several decision trees to enhance predictive accuracy and control over-fitting (Yoon, 2021). Every tree in the random forest model is built on a random subset of the dataset, and the final prediction is achieved by averaging the outputs of all the trees. This approach enhances the model's robustness and accuracy by reducing variance and capturing a broader range of data patterns, which makes random forests well-suited for GDP prediction based on diverse economic indicators. (Breiman, 2001). Mathematical representation of the model:

$$\hat{f}(x) = \frac{1}{T}\sum_{t=1}^{T} f_t(x)$$

5.4

$\hat{f}(x)$ – Final prediction, T – total number of trees $f_t(x)$ – prediction from the t^{th} tree.

The model was trained on both scaled and unscaled data using 100 trees. Similar to the previous model, GridSearchCV was used to find the best hyperparameters for the Random Forest model. The parameter grid includes different values for the number of estimators: [100, 200], maximum depth[None, 10, 15], minimum samples required to split an internal node: [2, 5] and minimum samples required to be at a leaf node: [1, 2]. The best-suited model was selected based on cross-validation. The best parameters were selected as {'max_depth': None, 'min_samples_leaf': 1, 'min_samples_split': 2, 'n_estimators': 100}.

5.7.4 Gradient Boosting Machines

Gradient Boosting *(*5.5*)* is considered a powerful ensemble technique that creates models sequentially, with every new model correcting the errors of the previous model. This model increases accuracy by focusing on the residuals and efficiently minimising the prediction error.

This method is primarily effective in capturing intricate patterns and interactions among the data. Making it highly suitable for economic forecasting (Yoon, 2021). The interactive nature of GBM ensures the model consistently improves while providing precise and reliable GDP prediction (Friedman, 2001).

Similar training methods were followed for GBM. Additionally, in hyperparameter tuning, different values were included in the parameter grid as learning rate: (0.1, 0.01), number of estimators, maximum depth, subsample, minimum samples required to split an internal node, minimum samples required to be at a leaf node, and maximum features. The best model was selected based on the cross-validation outcome. According to GridSearchCV, the best parameters were selected as - {'max_depth': None, 'min_samples_leaf': 1, 'min_samples_split': 2, 'n_estimators': 100}.

$$\hat{f}(x) = \sum_{m=1}^{M} \gamma_m h_m(x) \qquad 5.5$$

$\hat{f}(x)$ – Final prediction, M – Total boosting iterations, γ_m- weight applied to the m^{th} weak learner, $h_m(x)$ – m^{th} weak learner (E.g.- a decision tree)

5.8 Performance Evaluation

5.8.1 Error Metrics

The Models' performance in predicting the UK GDP was evaluated with error metrics, such as Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). These metrics provided a comprehensive overview of the model's prediction accuracy and error characteristics, allowing detailed evaluation of the models to determine which model produces reliable forecasts based on the given economic indicators.

RMSE is a commonly used metric that measures the square root of the average squared difference between predictions and actual values. It is vital to understand the absolute fit of the model to the data. The squaring puts more weight on large errors, making the metric more sensitive to outliers.

In this study, GDP is measured in large numbers, and RMSE *(5.6)* helps quantify the prediction deviation from the actual values. Hence, a lower RMSE indicates the model's robustness against large errors. Since RMSE is more sensitive to more significant errors, it helped identify the models' performance when predicting extreme GDP values and whether the model exceeded or underestimated the values by substantial margins. For example, a high RMSE indicated the model's struggle to accurately predict GDP during economic shocks or other extreme conditions.

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

5.6

y_i - Actual value of the target GDP for observation I, $\hat{y}$ - Predicted GDP for observation I, n – total number of observations

MAE *(* 5.7*)* provides the average absolute difference between the predicted values and the actual values in the same units as the output variables, allowing straightforward interpretation. In the current context, MAE helped determine how close the predicted values were to the actual GDP values without the influence of outliers. Since this metric provides the average error in the same unit as the target variable (e.g., Millions of pounds for GDP), it makes it easy for anyone to depict the typical average difference between predictions and values without the influence of outliers.

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

5.7

y_i – Actual GDP, $\hat{y}_i$ - Predicted value of observation i, n – Total number of observations

MAPE *(* 5.8*)* calculates prediction accuracy as the absolute percentage difference between predicted and actual values. The percentage differences allow easier performance comparisons across different scales and different models. Since MAPE presents the error in proportion to the actual value, it adds value to economic forecasting, where the GDP can change on a large scale over time. Thus, MAPE helps normalise the errors and provide a clear understanding of the models' performance in predicting the relative changes in GDP.

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

5.8

y_i - Actual GDP for observation i, $\hat{y}_i$ – Predicted GDP for observation I, n – number of observations, 100% - Multiplied to convert to percentage

5.8.2 R^2 Score

The R^2 score *(* 5.9*)*, also known as the coefficient of determination, measures the proportion of variance in the dependent variable that is predictable from the independent variables (Ruiz, et al., 2011). Its value ranges from 0 to 1, where a higher R^2 score indicates that the model can explain a large proportion of the variation of the outcome variable. This performance metric was applied to the prediction models as a measure of goodness of fit to measure the proportion of variance the models can predict in GDP based on the input features. Therefore, the R^2 score plays a vital role in evaluating the quality of the prediction and the relevance of selected features in explaining GDP variability.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \qquad 5.9$$

Y_i - Actual value of the dependent variable (GDP), $\hat{y}_I$ -Predicted GDP, $\bar{y}$ - Mean of the actual GDP, n – total number of observations

5.9 SHAP – Prediction Model Explanation

SHAP (Shapley Additive exPlanation) is a fairly novel framework introduced in 2017 (Lundberg, 2017) to explain the output of machine learning models. This concept has been inspired by the Shapley values from game theory, where the goal is to fairly distribute the total value of a game among the players based on their individual contribution to the total (Lundberg, 2017). Similarly, in machine learning models, features are identified as 'players' that contribute to the model's prediction. The SHAP value of each feature represents its contribution to the outcome of the model, where the impact is presented as positive or negative while quantifying the magnitude of the impact. This makes SHAP a powerful and fair method in interpreting the prediction of complex models, as it clearly depicts how each feature individually contributed to the final output, providing a detailed understanding of feature importance and model behaviour.

In this study, SHAP *(*5.10*)* plays a significant role in enhancing the interpretability of the GDP prediction models as it aids in understanding the feature contribution for the final prediction. Intern helps policymakers and users determine whether the changes in the variables have a negative or positive impact on GDP and identifies the most influential features for the GDP prediction of the UK. Moreover, SHAP aids in providing both global and local interpretability. This means that global interpretability aids in identifying features that have the most overall impact, while local interpretability helps understand the impact of features on individual predictions. Most importantly, it aids in performing realistic, comparative analysis between models while identifying whether specific models over- or under-emphasise particular features.

$$\emptyset_i = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|!\ (|F| - |S| - 1)!}{|F|!} (f(\mathrm{S} \cup \{\mathrm{i}\} - \mathrm{f(S)}) \tag{5.10}$$

$Ø_i$ - SHAP value of feature i, F – the set of all features, S – A subset of features that does not include feature i, f(S) – The model output when the features in the subset S are present. , |S|! – The factorial of the size of Subset S, |F|! = The total number of features, (|F| - |S| - 1)! – The factorial of the difference between the total number of features and the subset size.

According to cooperative game theory, the SHAP values are calculated using the above function. Moreover, this calculation can be applied in the Python environment using either 'shap.kernelExplainer' or 'shap.DeepExplainer' (Lundberg, 2017). Out of these functions, 'shap.kernelExplainer' was selected as it has been designed to be used for more straightforward functions and to generate more accurate results than sampling-based approaches (Lundberg, 2017). SHAP was applied to each version of every model (scaled, unscaled and hyperparameter-tuned models). The 'KernelExplainer' function was initialised using each model's prediction function and training data, which allowed us to calculate the marginal contribution of each feature to the model's output. The SHAP values were computed for the test dataset of each version of the models. Finally, SHAP summary and mean absolute SHAP values plots were generated to present a visual representation of the feature importance of each model.

5.10 GDP Prediction Models' Performance

5.10.1 Linear Regression (LR)

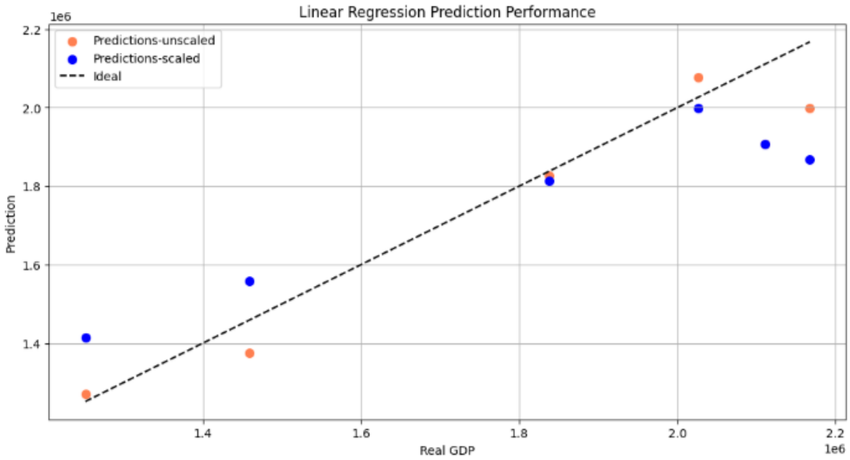


Figure 5.7 LR Scaled and Unscaled Performance

The LR model's predictions with unscaled data displayed more alignment with the ideal line. Whereas, scaling appeared to have worsened the prediction performance. In this case, predictions seemed to have deviated highly from the ideal line with higher predictions. This highlights the LR model's sensitivity to feature scaling. However, both versions of the model display some notable deviations, indicating that LR might not be able to fully capture the complex relationships between macroeconomic features and GDP.

5.10.2 Support Vector Regression (SVR)

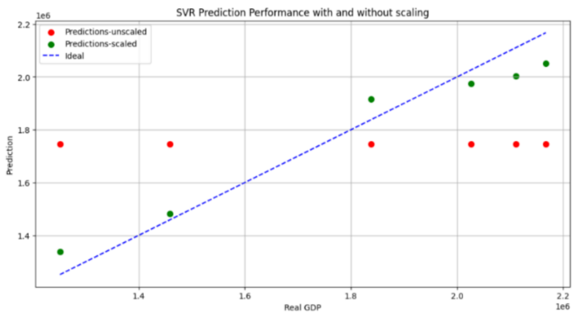


Figure 5.9 SVR Scaled Unscaled Performance

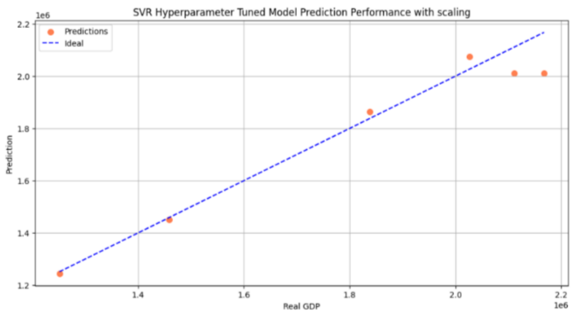


Figure 5.8 SVR Hyperparameter Tunned Performance

The SVR model in predicting GDP with unscaled features showed systematic biases and poor prediction accuracy, especially at higher GDP values. In contrast, the model displayed significantly better alignment with the ideal line with scaled features, indicating improved performance. According to the hyperparameter-tuned SVR model with scaling, the predictions appeared to be closely aligned with the ideal line, suggesting substantial improvement in the model’s performance compared to both scaled and unscaled versions without tuning. The tuning process allowed the SVR model to effectively capture the complex relationships, resulting in more precise GDP predictions. Therefore, the scaling as well as hyperparameter tuning had notably enhanced the performance of the SVR model, producing a better fit to the GDP values.

5.10.3 Random Forest (RF)

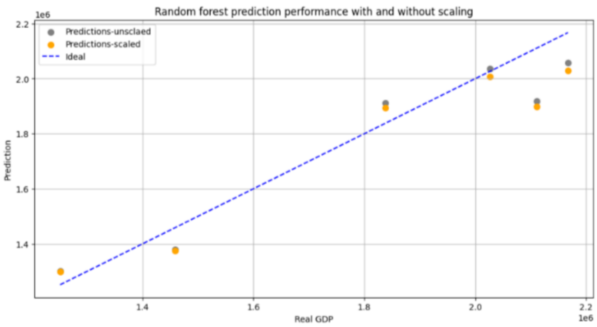


Figure 5.11 RF Scaled Unscaled Performance

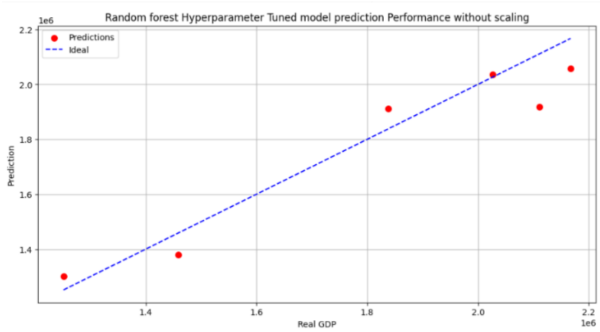


Figure 5.10 RF Hyperparameter Tunned Performance

As displayed by the plots, the RF model’s predictions with both scaled and unscaled features appear similar to each other, with prediction points clustering around the ideal line. This suggests that feature scaling did not significantly impact the model’s performance, most likely due to the tree-based structure of the RF model, which is inherently less sensitive to feature magnitudes (Khairani, et al., 2022). As demonstrated by the second plot, hyperparameter tuning had little to no impact on the RF model's performance. Although the hyperparameter-tuned version of the model showed a slight improvement over the scaled version with the original parameters, its predictions appeared largely similar to the unscaled version without tuning. This indicates that the hyperparameter tuning had no significant impact on the model's performance. In fact, the original version of the RF model with default parameters appeared to be well-suited to the dataset. Therefore, with both unscaled and hyperparameter-tuned versions exhibiting comparable performance by closely following the ideal line, it demonstrates that the model can reasonably capture the relationship between the features and GDP with limited impact from scaling or tuning.

5.10.4 Gradient Boosting Machines (GBM)

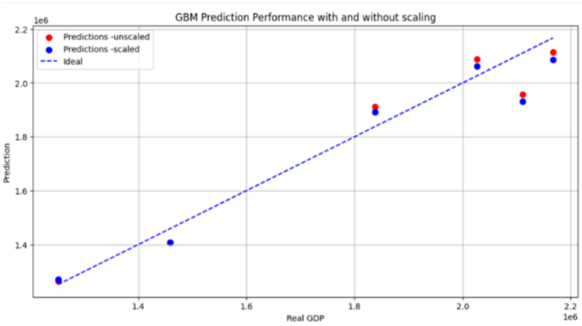


Figure 5.14 GBM Scaled, Unscaled Performance

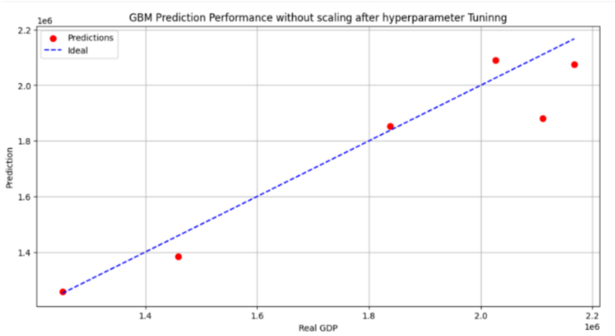


Figure 5.13 GBM Hyperparameter Tunned Performance

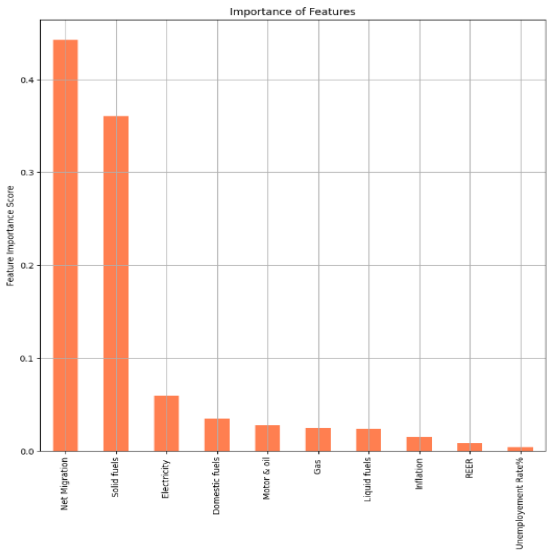


Figure 5.12 GBM Feature Importance

In terms of GBM performance, the predictions of the unscaled version of the model displayed close alignment to the ideal line in all predictions, whereas the scaled predictions seemed to underperform, especially with higher GDP values. This suggested that the original version of the features allowed the model to capture more meaningful relationships in the data. This indicates that, as a tree-based model, the GBM model is robust to scaling; in fact, scaling might have had an adverse effect by making certain features less prominent. According to the second plot, hyperparameter tuning slightly worsened the performance of the model since the prediction points show slightly higher deviation from the ideal line, especially at higher GDP values. This indicates that hyperparameter tuning did not enhance the model's performance; rather, it might have reduced the model’s ability to accurately capture the relationship between macroeconomic variables and GDP. Overall, the unscaled GBM model outperformed both scaled and hyperparameter-tuned versions, indicating that maintaining the original features and avoiding excessive tuning could lead to better model performance.

Additionally, the feature importance plot of the GBM model provided insight into the critical features that influenced its predictions. According to the plot, it is evident that net migration and solid fuels had the most substantial influence on the predictions, suggesting that they have the strongest relationship with GDP growth compared to other macroeconomic variables. This was followed by electricity, domestic fuels, and motor oil, which had moderate importance, indicating a significant but less dominant role in influencing GDP. Moreover, the remaining features, namely, gas, liquid fuels, inflation, REER and unemployment rate, displayed relatively low importance, suggesting a lesser contribution to the model's performance.

5.11 Performance Metrics Evaluation

5.11.1 Linear Regression

The performance metrics of the LR model displayed apparent differences between scaled and unscaled versions. The RMSE of the unscaled LR model was 115341.53 compared to the higher RMSE of 167933.04 of the scaled versions. Similarly, the MAE of the scaled model appeared to be significantly higher than the unscaled model. This indicated that the unscaled model provided a more accurate prediction. The difference was further confirmed by the MAPE value as the unscaled model only had a MAPE of 46.4%, while the scaled model showed a larger percentage error of 76.7%.

The R^2 score also displayed better performance unscaled version of the LR model. The unscaled model had an R^2 score of 0.8862, meaning it is able to explain approximately 88.62% of the GDP's variance. The scaled version could only explain 75.88% of the GDP's variance. Therefore, scaling appeared to have negatively impacted the model, possibly due to the scaling disrupting the inherent relationship between features and GDP.

Table 5.1 LR Performance Metrics

	RMSE	**MAE**	**MAPE**	**R^2**
Un-Scaled	115341.528	89607.197	0.046	0.886
Scaled	167933.044	136408.114	0.076	0.758

5.11.2 Support Vector Regression

Unlike the LR model, the SVR model showed clear improvement as the data went through scaling and hyperparameter tuning. According to the error metrics, the unscaled model performed very poorly with high RMSE, MAE, and MAPE scores. In contrast, scaling appeared to have a clear improvement on the performance metrics, as the RMSE dropped to 83492.05, MAE to 77210.27, and MAPE to 4.29%, which is a significant improvement from the unscaled version. Moreover, the hyperparameter-tuned version displayed further improvement by reducing the error metrics.

The R^2 scores also reflected a similar trend, with the R^2 score for the unscaled model being -0.033, indicating that the model performed worse than a simple mean prediction, as it was unable to even slightly capture the variations of the GDP. However, scaling significantly improved the performance, as the model was able to predict 94% variations in GDP. The hyperparameter tuning moderately increased the R2 score to 94.6. Hence, all the performance metrics indicated that both scaling and hyperparameter tuning significantly improved the SVR model's performance.

Table 5.2 SVR Performance Metrics

	RMSE	**MAE**	**MAPE**	**R^2**
Un-Scaled	347604.129	32338.875	0.191	-0.033
Scaled	83492.048	77219.274	0.042	0.042
Tuned	78850.795	57008.954	0.027	0.946

5.11.3 Random Forest

The performance metrics of the RF model displayed interesting observations. The unscaled version's average magnitude of residual error was approximately (RMSE) 103161, and the average absolute difference was nearly (MAE) 85672. The deviation of the predictions from the actual was nearly 4.67%. Contrastingly, the scaled version displayed higher error metric scores. This suggested that scaling did not enhance the model's performance but slightly deteriorated the prediction performance. This diminish in performance was consistent with the RF model's inherent robustness in feature scaling. Interestingly, the hyperparameter-tuned RF model displayed error metrics identical to those of the unscaled version of the model. This demonstrated that hyperparameter tuning necessarily did not improve the performance beyond the default scaled version.

The R^2 score of all three versions of the model confirmed the same. The R2 score of both unscaled and hyperparameter-tuned models demonstrated that the model was able to accurately identify 90% of GDP variations. In comparison, the scaled version showed slightly inferior performance with an R2 score of 89%. This suggests that the default confidence ratio of the RF model with unscaled data is sufficient for GDP prediction.

Table 5.3 RF Performance Metrics

	RMSE	MAE	MAPE	R^2
Un-Scaled	103160.735	85672.391	0.046	0.909
Scaled	112802.952	90445.529	0.049	0.891
Tuned	103160.735	85672.391	0.047	0.909

5.11.4 Gradient Boosting Machines

According to error metrics, the un-scanned version of the GBM model displayed better performance, with an RMSE of 80744, an MAE of 68289, and a MAPE of 3.58%. and the scaled version had slightly higher error metrics, indicating a minor decline in model performance. This is in line with the inherent sturdiness towards scaled data due to its tree-based structure. Contrastingly, the hyperparameter-tuned GBM model displayed the worst performance of all three versions of the model with a high RMSE of 109091, MAPE of 80586.23 and MAPE of 4.13%

R^2 scores also agreed with the error metrics, as the unscaled and scaled versions had close scores with slight variations of 94.4% and 89.8%, respectively. However, the hyperparameter tuned version could only recognise 89.8% of the GDP variance. Therefore, the unscaled GBM model stands out as the most effective version of the GBM model in the UK's GDP prediction.

This highlights that, in some cases, neither additional processing nor optimisation necessarily yields better performance but rather leads to potentially degraded results.

Table 5.4 GBM Performance Metrics

	RMSE	MAE	MAPE	R^2
Un-Scaled	80744.238	68289.8	0.035	0.944
Scaled	87873.275	70277.466	0.036	0.933
Tuned	109091.11	80586.23	0.041	0.898

5.12 SHAP Analysis

5.12.1 Linear Regression

The SHAP analysis provided detailed insight into each feature's influence on GDP prediction. According to the plots below, a clear distinction between the feature importance of scaled and unscaled versions can be observed. In the unscaled model, gas and domestic fuels emerged as the most influential features. Followed by net migration and solid fuels. This indicates that energy prices, especially gas and domestic fuel prices, have the highest average impact on GDP prediction.

Contrastingly, the SHAP plots of the scaled LR model exhibited different characteristics. The influence of features has been highly skewed with scaling since net migration appeared to dominate the impact on predictions, with gas, domestic, and motor & oil exhibiting nearly negligible influence on the GDP prediction. This demonstrates how scaling has significantly altered the model's interpretation of feature importance, which has led to a loss of predictive power for most features.

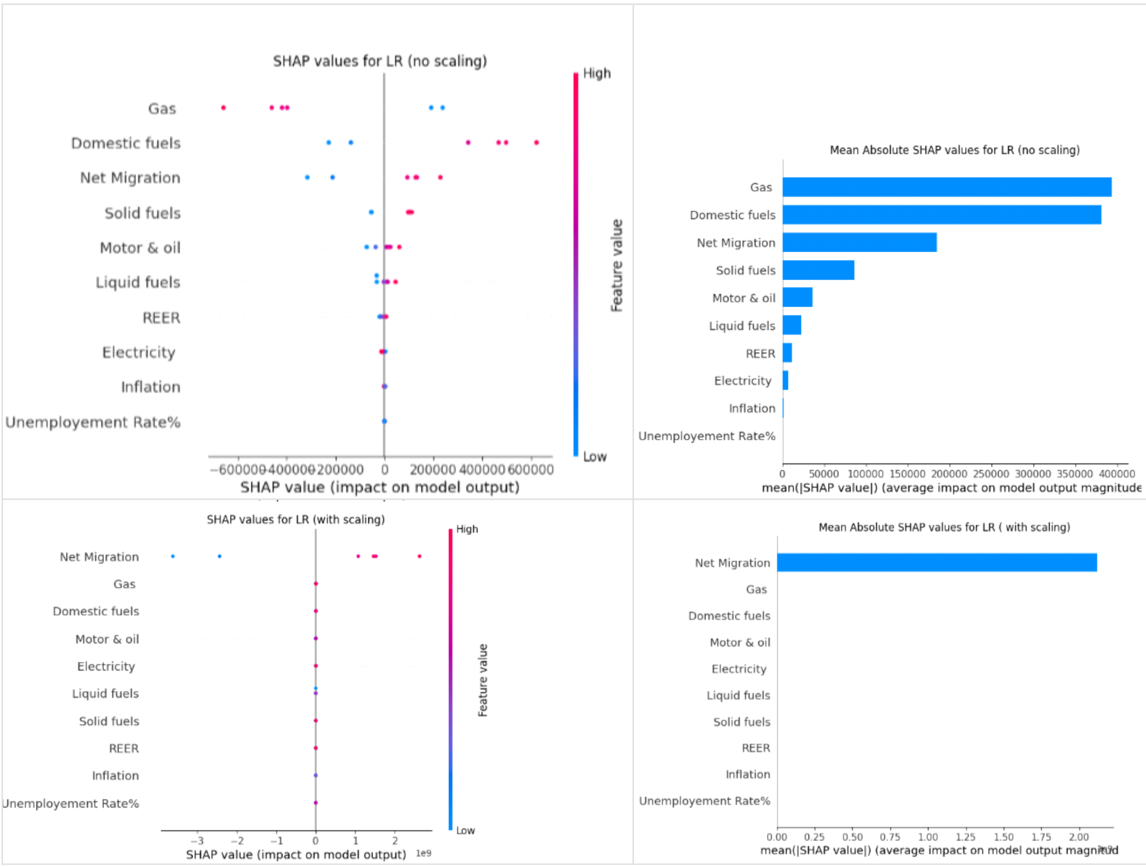


Figure 5.15 LR SHAP Plot(top right), Mean Absolute SHAP Values (top left), LR with scaling SHAP (Bottom right) Mean absolute SHAP Values (bottom left)

5.12.2 Support Vector Regression

According to both the SHAP summary plot and the mean absolute value plot for SVR without scaling, net migration appeared to have the only influence on the predictions, while other features displayed almost negligible importance.

For SVR with scaling, solid fuel, domestic fuel, gas, and net migration displayed the highest impact on the predictions, and they appeared to have a strong positive influence. The following features, such as unemployment rate, electricity, and motor oil, displayed a mix of both positive and negative influences. Further, REER and inflation displayed the least influence on predictions.

In the SVR model with hyperparameter tuning, net migration again appeared to be the most impactful feature, closely followed by solid fuel and the unemployment rate. In this version of the model, REER and net migration displayed the least influence. Interestingly, hyperparameter tuning appeared to have enhanced the relative importance of net migration while making solid fuel and the unemployment rate more influential.

Across all three versions, net migration consistently appeared to have a significant influence on predictions. Features like solid fuels, domestic fuels, and gas became more influential with scaling. With hyperparameter tuning, the unemployment rate gained significant influence, surpassing other variables.

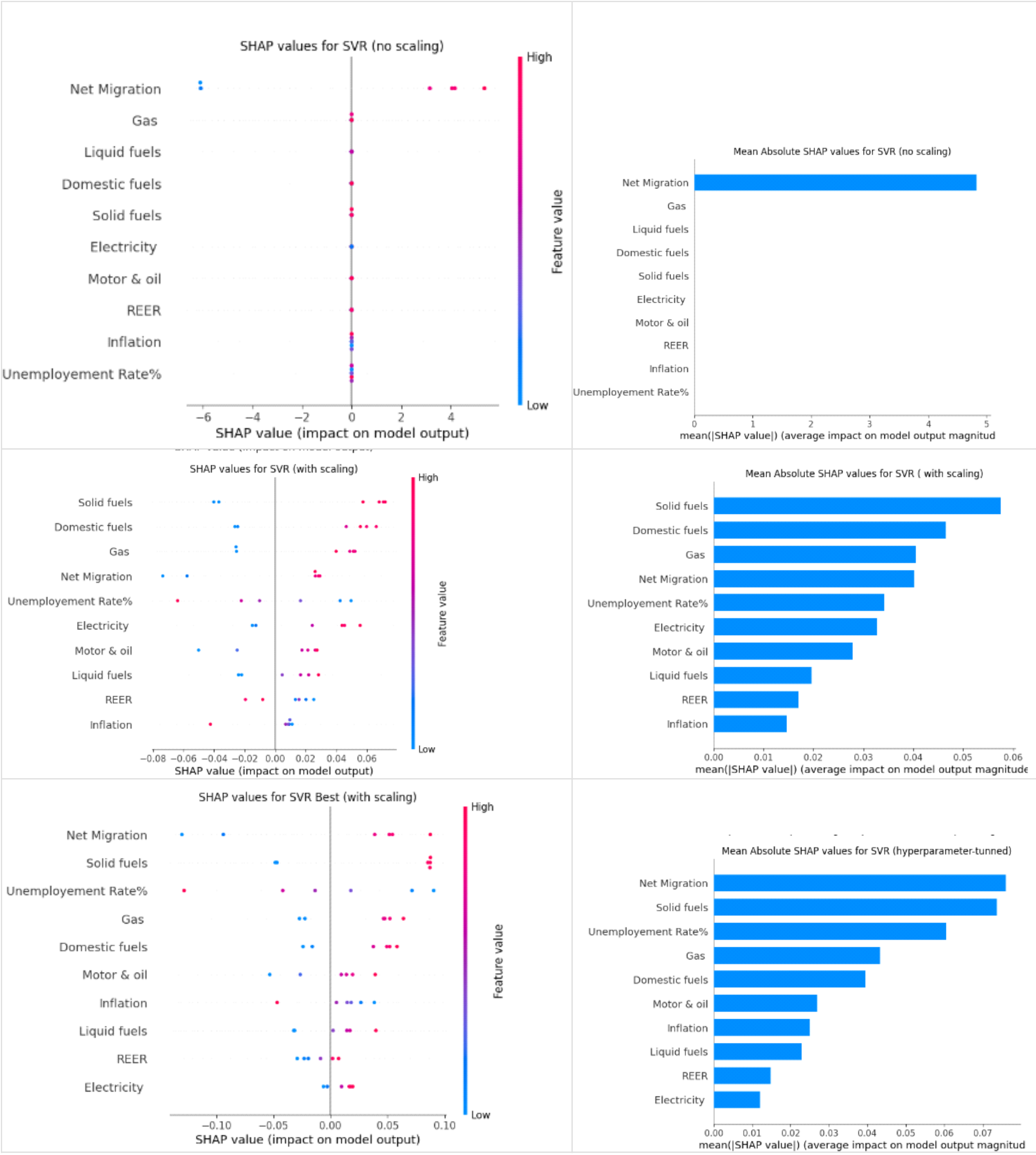


Figure 5.16 SVR SHAP Plot(top right), Mean Absolute SHAP Values (top left), SVR with scaling SHAP (middle right) Mean absolute SHAP Values (bottom right), Hyperparameter-tuned SVR SHAP plot (bottom left), Mean absolute SHAP values (bottom right)

5.12.3 Random Forest

The SHAP evaluation of the RF model revealed that solid fuel, net migration and motor & oil were the most prominent features of the predictions in all unscaled, scaled and hyperparameter-tuned versions of the model. In all settings, solid fuel exhibited the most significant positive influence on model output. The rest of the features, such as electricity, liquid fuel and gas, displayed moderate influence on the model. Meanwhile, REER, unemployment rate, and inflation displayed the least influence on the predictions, which are closer to zero. This highlights that energy prices, particularly solid fuels, play a crucial role in GDP predictions. Moreover, hyperparameter tuning or scaling had no impact on feature importance, and each feature's level of impact underwent only gone through minor changes with model alterations.

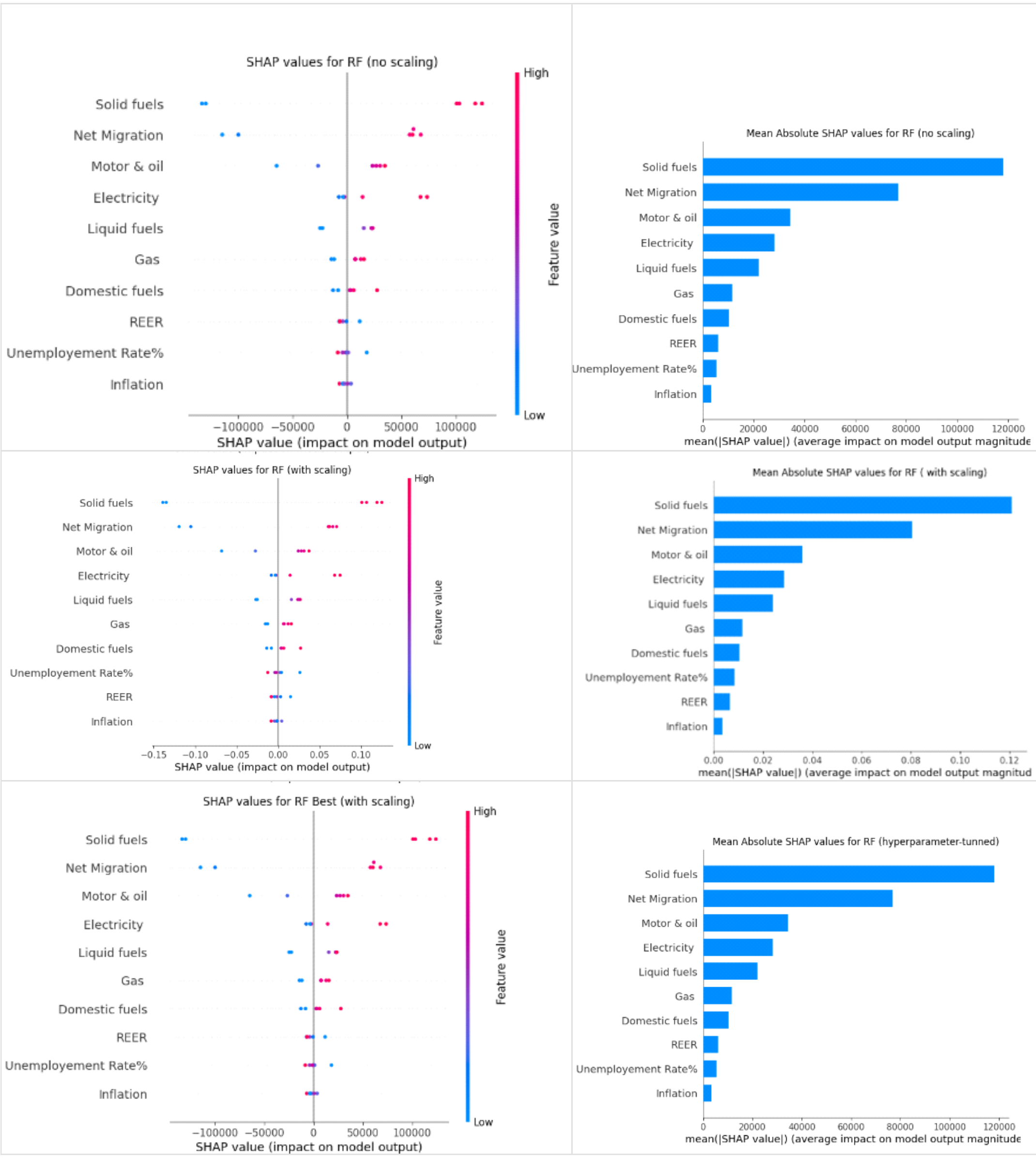


Figure 5.17 RF SHAP Plot(top right), Mean Absolute SHAP Values (top left), RF with scaling SHAP (middle right) Mean absolute SHAP Values (bottom right), Hyperparameter-tuned RF SHAP plot (bottom left), Mean absolute SHAP values (bottom right)

5.12.4 Gradient Boosting Machines

For GBM, the importance of the feature varied based on scaling and hyperparameter tuning. Across all three versions of the model, solid fuels and net migration emerged as the most influential features in predicting GDP. In the unscaled model, solid fuels had the highest impact, followed by net migrations. When scaling was applied to the dataset, solid fuel and net migration remained as the most influential features, although net migration's influence further increased with scaling. Similarly, electricity and motor & oil average impact heightened with scaling.

Interestingly, with hyperparameter tuning, motor & oil become the most important features in GDP prediction, surpassing solid fuels and net migration. However, while hyperparameter tuning improved the clarity of feature impacts, it did not necessarily enhance the prediction accuracy. Further, inflation, unemployment rate, REER, and gas consistently showed lower impacts across all versions of the model, suggesting that these features had minimal effects on GBM predictions. These SHAP analysis plots indicate the significance of energy-related factors and net migration on predictions while also reflecting on the sensitivity of the model to feature scaling and model tuning.

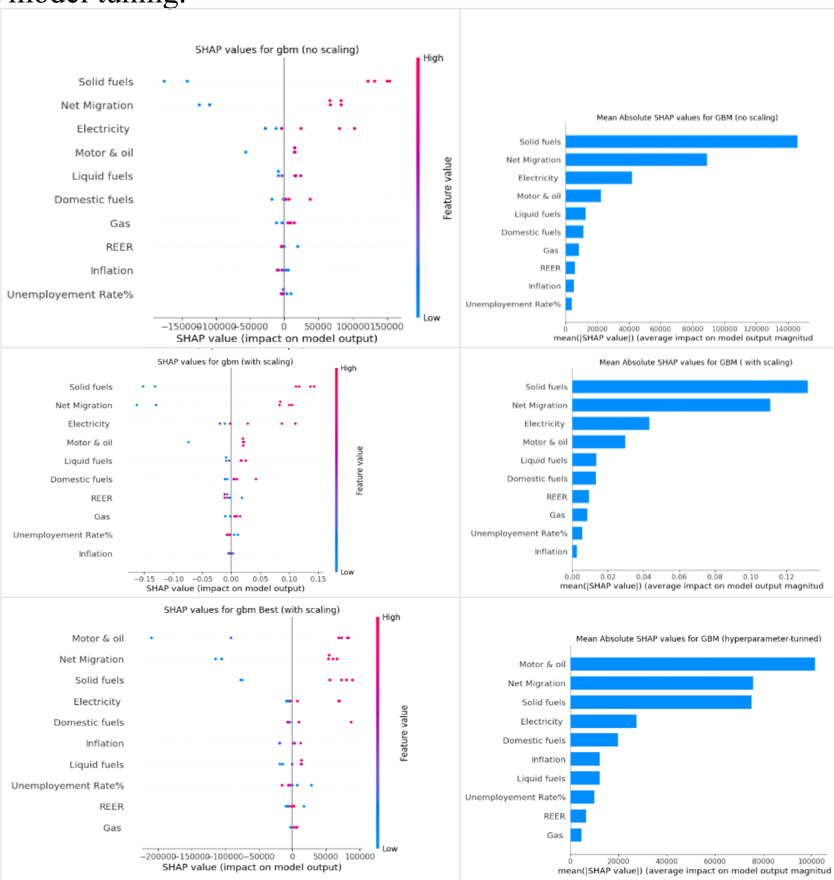


Figure 5.18 GBM SHAP Plot(top right), Mean Absolute SHAP Values (top left), GBM with scaling SHAP (middle right) Mean absolute SHAP Values (bottom right), Hyperparameter-tuned GBM SHAP plot (bottom left), Mean absolute SHAP values (bottom right)

6 CHAPTER SIX

Discussion and Analysis

6.1 Introduction

This chapter presents and discusses the results from the GDP prediction study, which applies machine learning models to analyse the influence of energy prices and key macroeconomic variables on the UK's GDP. The performance of the selected four models is evaluated using the standard error to identify the most effective approach for forecasting GDP. beyond the predictive accuracy, this chapter explores the model interpretability through SHAP, and the relative importance of each input is discussed through an economic perspective. This discussion critically assesses how different models respond to structured economic data, highlights the most influential features and reflects on the broader implications of using XAI for economic forecasting and policy analysis.

6.2 Research Overview – UK GDP Prediction

This study aimed to predict the Gross Domestic Product (GDP) of the United Kingdom using energy prices and several other macroeconomic variables from 1990 to 2018 by employing machine learning approaches. The motivation behind the research was to accurately forecast the GDP of the UK, since it is an essential factor for policymakers as well as other actors in economic planning. Traditional economic models often struggle to capture the complexity and non-linear relationship between the macroeconomic variables and the GDP.

The macroeconomic predictors included in the study are, namely, energy prices (solid fuels, gas, electricity, liquid fuels, domestic fuels, motor & oil), REER, unemployment rate, inflation and net migration. Several models were utilised for the GDP predictions; LR was employed as a baseline model, followed by SVR, RF and GBM. In order to identify the optimum version of the models, the study was executed using a different form of preprocessing, such as scaling and hyperparameter tuning.

The performance of each model was evaluated using standard performance metrics such as RMSE, MAE, MAPE and R^2 score. These metrics greatly aided in determining each model's accuracy, efficiency, and suitability for GDP predictions. Further, SHAP analysis was incorporated into each version of all the models to identify the contribution of each macroeconomic feature to the final GDP prediction in order to gain an in-depth understanding and interpretability of the predictions.

Considering the computational complexity, the conventional machine learning models used in this study were trained on annual macroeconomic indicators, resulting in a relatively small and structured dataset. The least complex model, LR, is computationally lightweight with a training complexity of $0(n^2 * d)$ to $o(n^3)$, which can be prohibitive in large data sets. The use of small datasets spanning a few decades of annual data also aided in keeping the SVR computational within manageable bounds. Both RF and GBM scale with the number of estimators and features, whereas RF is generally faster due to the parallel training of trees. On the other hand, GBM trains trees sequentially, which increases the runtime. Nonetheless, the structured nature and limited volume of the GDP data ensured that all conventional models remained computationally effective and efficient without the use of advanced hardware or long training times.

6.3 Findings

As presented by the below summary of performance metrics, the hyperparameter-tuned SVR model with scaled features demonstrated the best performance in UK GDP prediction since it denoted the lowest average difference between predictions and the actual values, average absolute difference (RMSE – 78850.79, MAE – 57008.95) and significantly low absolute percentage difference from the actual of 2.72%. Furthermore, the hyperparameter-tuned SVR model was able to accurately identify nearly 95% of the variance in the GDP of the UK.

This was closely followed by the default GBM model with unscaled features, as evident through competitive performance metrics. Although scaling did not improve the prediction performance, it rather slightly deteriorated the performance. Despite that, it was able to closely follow the trend and produce the next best predictions compared to the other models. However, the hyperparameter-tuned version further aggravated the model's performance, indicating that the default features were adequate to identify the nonlinear relationships between the macroeconomic features and the GDP.

The default RF model with unscaled features demonstrated acceptable performance. However, feature scaling diminished the model's performance, while the hyperparameter tuning demonstrated minor improvement in predictions.

Given its simple structure, the baseline LR model with unscaled data also demonstrated reasonable performance. However, scaling drastically aggravated the model's performance by increasing its RMSE from 115341.52 to 167933.04, MAE from 89607.19 to 136408.11, and MAPE from 4.64% to 7.67% while reducing the R^2 score by 13%.

Primarily, the default SVR model with unscaled data demonstrated the worst performance, with significantly contrasting performance metrics throughout all the measures. Primarily, it was evident through the R2 score since almost all the models were able to determine above 85% variance of GDP; the default SVR model gained -0.3% value, indicating that the model was not even able to detect the relationships between macroeconomic variables. This underperformance is likely due to factors such as the SVR model's high sensitivity to feature scaling, as it relies mainly on distance-based calculations with the aid of kernel functions such as radial basis function (RBF). Without normalisation, variables with large numerical ranges, such as energy prices or net migration, dominate the model's learning process, which results in preventing the model from identifying balanced relationships. Further, in the default model, the regularisation parameters (C, gamma, epsilon) were not optimised for the dataset, which likely led to the notable underfitting since the model was too constrained to learn complex patterns in the dataset. Additionally, the multicollinearity within macroeconomic variables may have further intensified the model's inability to generalise.

This poor performance led to exploring improved processing steps, such as data scaling using MinMaxScaler and hyperparameter tuning through GridSearchCV, which significantly improved the SVR model's predictive power. These challenges also underscore the importance of tailoring machine learning models to the structure and scale of economic datasets.

The scatter plot below visualises the predictions of the best version of each model and the SVR default model without scaling to demonstrate how drastically apart its predictions are from the ideal line compared to the other models.

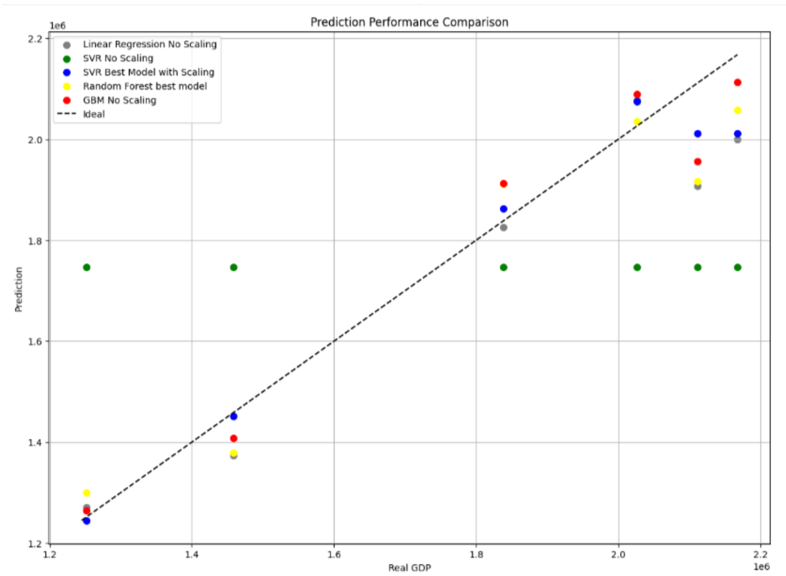


Figure 6.1 Best Version of Each Model Predictions Comparison

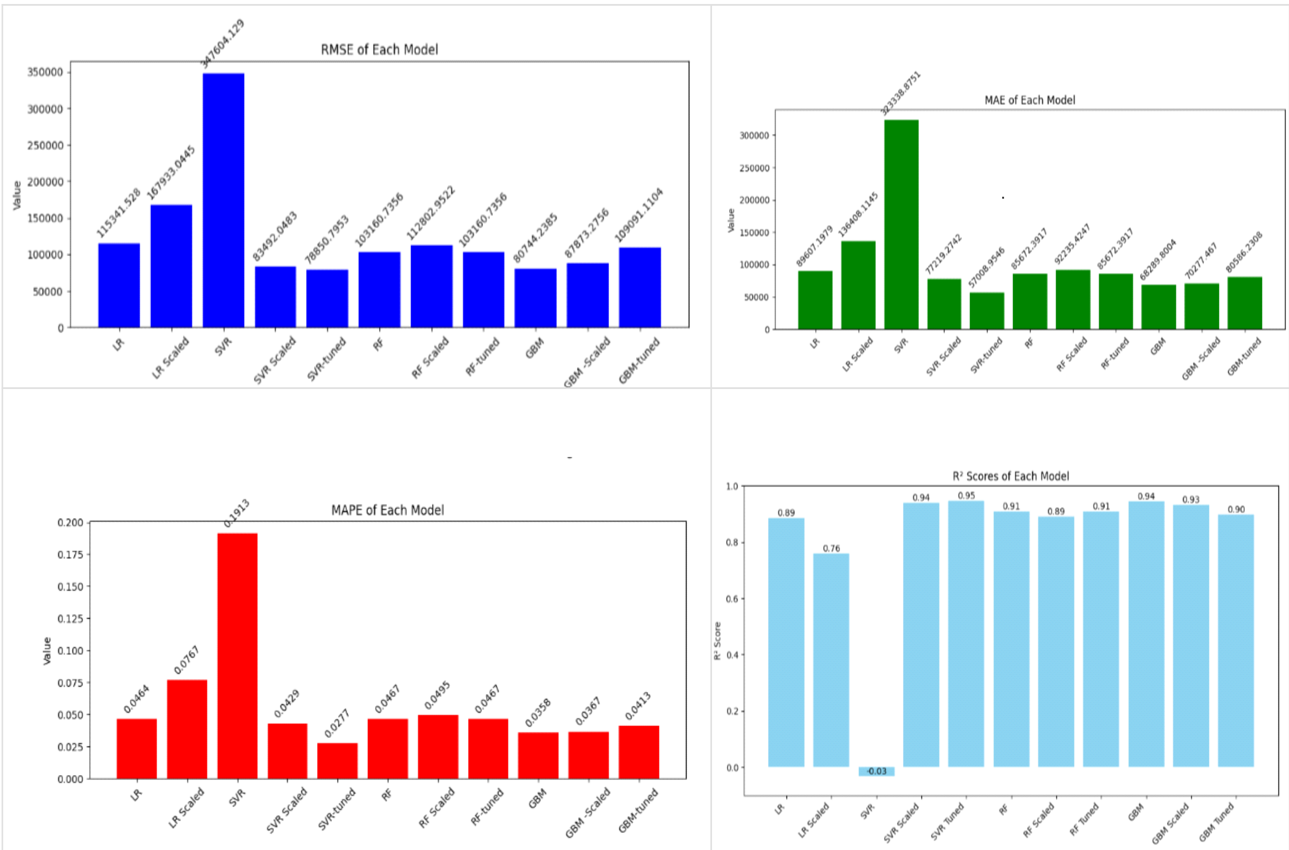


Figure 6.2 Summary of Performance Metrics of all models RMSE (top left), MAE (top right), MAPE (bottom left), R^2 score (bottom right)

6.4 Validity of the Results

Prior studies that have employed machine learning models in economic predictions have often highlighted that non-linear models such as Support Vector Regression (SVR) and ensemble methods such as Gradient Boosting Machines (GBM) demonstrate superior performance (Xu, et al., 2018), especially when handling complex non-linear relationships. As an example, studies by (Hu & Wang, 2022), (Maccarrone, et al., 2021) demonstrated that models incorporating non-linear kernels or boosting algorithms consistently outperformed linear models such as Linear regression (LR) in economic forecasting, which validates the findings of this research since the SVR model and GBM model outperformed the rest of the models with better performance.

The significant improvement in the SVR model with scaled data and hyperparameter tuning has been well documented as a valid method of model performance improvement, especially in relation to identifying intricate relationships in economic data. As the results depicted, the SVR model with hyperparameter-tuning yields the best performance with the lowest error metrics and the highest R^2 score, indicating the model's ability to explain almost 95% variance in GDP. This aligns with the findings from the study by (Chowdhury, et al., 2019), which clearly demonstrates that properly tuned SVR models could achieve high accuracy in predicting GDP by optimising the kernel and other hyperparameters.

Furthermore, prior studies suggest that feature scaling tends to have varying effects depending on the type of model being used. As depicted by the studies of (Zhang, et al., 2020), scaling could sometimes deteriorate the performance of tree-based models like random forest (RF) and GBM since these models are inherently invariant to feature scaling. This was clearly demonstrated by the current study, as scaling showed obvious deteriorations in the predictions of RF and GBM models. In contrast, the SVR model's performance drastically improved with scaling. Because, as a kernel-based model, the SVR model is highly sensitive to the scale of input features. Therefore, the model struggles to find optimal decision boundaries with differently bounded features, resulting in subpar performance (Gujarati, 2009). The significant worsening in LR performance with scaling can also be aligned with the same. Since scaling could lead to prioritising specific predictors, thereby disrupting the linear relationship assumed by the model (Trivedi, et al., 2021).

Moreover, the SVR model with default parameters showed a remarkably low R^2 score and demonstrated almost zero capability to capture any form of relationship. This resonates with the findings of prior literature that emphasise the necessity of tuning SVR hyperparameters for optimal results. It has been proven that default SVR settings could lead to suboptimal performance, particularly in environments with high-dimensional and noisy data (Xu, et al., 2014).

6.5 SHAP Results Interpretation

The SHAP analysis across all models consistently highlighted net migration as one of the most influential features in GDP prediction. This strong influence suggested that changes in net migration significantly impacted the economic direction captured by the GDP of the UK. This finding aligns with the general economic understanding that migration considerably impacts labour markets, consumption and the overall economic direction (Kudaeva & Redozubov, 2021). The high SHAP scores on net migration across all models, including the worst-performing default SVR model with unscaled features, indicated its consistent relevance in driving GDP predictions, regardless of the model or the model configurations.

In terms of energy price data, gas, domestic fuels, and solid fuels displayed a prominent impact on the models' predictions, emphasising the role of these energy sources on the GDP. In contrast, motor & oil, liquid fuels and electricity consistently exhibited moderate to low influence on GDP predictions. This could be due to the sector-specific impact of these particular energy prices, rather than having a broad impact on the entire economy as domestic fuel and solid fuels. Regardless, all the energy price data demonstrated notable influence on almost all the prediction models and their variations.

Interestingly, the unemployment rate displayed a strong influence on the SVR models with scaled and tuned parameters, which indicates that these models were more sensitive to labour market conditions. Meaning that the properly tuned and scaled SVR model is more capable of capturing the relationship between employment and GDP. However, the rest of the models experienced a comparatively low influence on the unemployment rate, implying that the linear and tree-based models may not have the capacity to effectively capture the nuanced relationship between unemployment and GDP. This noteworthy difference highlights the advantage of utilising non-linear models such as SVR when working with complex, non-linear relationships of economic features.

The SHAP analysis also revealed that inflation consistently had a low to moderate influence on GDP predictions across all models. This is an interesting discovery, especially given the theoretical importance of inflation in influencing macroeconomic performance since inflation affects purchasing power, interest rates, and overall economic stability. Yet, its minor influence on GDP predictions indicates that other factors significantly drove the GDP during the study period. This could be accounted for in this study considering the real GDP, which is adjusted to account for inflation, effectively representing the actual volume of goods and services produced in the economy without the changing price levels (Mankiw, 2019). Therefore, inflation becomes a less significant predictor compared with other variables as the effect of inflation is neutralised in real GDP (Afrouzi & Bhattarai, 2023).

Finally, REER consistently demonstrated the lowest influence on GDP predictions across all the models. This could most likely be due to its substantially low influence on the UK GDP during this period. Therefore, the rest of the macroeconomic features could have overshadowed its influence. In general, the relationship between REER and GDP is more indirect, and its influence reflects on the economy mainly in the long term (Bussière, et al., 2015).

6.6 Economic Interpretation

The SHAP analysis findings provided several economic insights into the drivers of the UK's GDP. The consistently high influence of net migration across all models reflected the significant role migration plays in the UK's economic landscape. According to (Campo, et al., 2024), migration has been a critical factor in sustaining the UK labour market and the ageing population. Especially after the expansion of the European Union, where an influx of skilled and unskilled workers contributed to the economic productivity and service sector growth. Brexit negatively impacted this since many European workers had to leave the country due to strict migration policies. The consistent influence of net migration on the prediction further supports the argument that adequate migration is a vital feature for maintaining economic growth. This is especially relevant to the migration policies following Brexit.

The significant influence of gas, domestic fuels, and solid fuels on GDP predictions underscores the importance of energy prices in shaping the UK's economic output. This was clearly discussed through the empirical research done by (Maeda, 2007), where they highlighted that there is a direct link between energy consumption and GDP growth; they also argued that energy prices have substantial effects on production costs, consumers' purchasing power and the overall economy. Similar to other economies in the UK, energy costs affect both industrial production and household consumption. The prediction's reliance on domestic fuels and gas suggests that energy price fluctuations directly influence industries and households, which impacts demand and production capabilities. The UK economy's transition away from coal-based solid fuels and increased resilience on gas due to decarbonisation policies also likely have changed the energy market intern, affecting macroeconomic outcomes (Grubb & Newbery, 2018).

In contrast, liquid fuels, motor & oil and electricity displayed a relatively moderate to low influence on GDP predictions, which may be due to the reduced consumption of oil-related energy in the UK production (Department of Business, Energy and Industrial Strategy, 2016)compared to gas and solid fuels. This can be attributed to the UK's gradual move towards cleaner energy sources and reduced dependence on oil products (Department of Business, Energy and Industrial Strategy, 2016). The low SHAP scores for electricity could also indicate a less direct impact on GDP compared to other energy types, since the electricity costs most likely be distributed across sectors, resulting in an indirect impact on the economy.

The unemployment rate's strong influence on the scaled and tuned SVR models indicated that unemployment has a non-linear relationship with the GDP, which was further confirmed as the rest of the models could not properly capture the connection. A higher unemployment rate can indicate reduced consumption that is reflected through reduced demand and results in reduced production, leading to contractions in GDP, and vice versa. As per the Beveridge Curve and Okun's Law (Okun, 1962), the unemployment rate has a significant inverse relationship with economic output, which aligns with the SHAP analysis findings on the SVR model. The stronger influence observed in the non-linear models suggests that unemployment has a varying impact across different sectors, possibly due to shifts in the labour market dynamics such as automation, skill mismatches and changing sector demands.

In terms of the low influence of inflation on the predictions of the GDP of the UK could be due to the relatively stable inflation rates experienced in the UK during most of the study period from 1190-2018. Especially after the early 1990s, when the UK adopted inflation targeting, the action taken by the authorities to maintain inflation at a target value through policies and procedures led to greater price stability. This could provide an explanation as to why inflation had a considerably low influence on the UK's GDP prediction. According (Blanchard, et al., 2015) inflation can have a less pronounced impact on real GDP when the central bank effectively manages inflation through fiscal and monetary policies.

Furthermore, the consistently low influence of REER on GDP predictions across all models indicates the minimal impact of real effective exchange rates on the overall economy during the study period. According to (Comunale, 2017), the impact of exchange rates on GDP is influenced by the economic structure. For instance, a diverse and robust economy like that of the UK is less affected by exchange rate fluctuations due to its wide-ranging exports and imports, as well as the significance of its financial services sector, which is not heavily reliant on the trade balance.

These findings provide in-depth insights into the UK's economic structure and the role of various macroeconomic factors in shaping GDP. As per the findings, net migration and energy prices remain the most significant drivers of the economic direction, while factors like REER demonstrated less impact, which aligns with the UK’s economic resilience on the currency fluctuations due to the close monitoring and maintenance by the Bank of England (O'Brien & Palma, 2023) As well it's service-based economy (Crafts & Mills, 2020).

7 CHAPTER SEVEN

Conclusion and Recommendations

7.1 Introduction

This chapter presents the overall conclusion of the two studies, drawing together the key findings from both meme stock price prediction and GDP forecasting components. The research aimed to explore the potential of machine learning techniques in two distinct yet highly data-driven domains of financial forecasting, where one is influenced heavily by public sentiment and market volatility, while the other is shaped through macroeconomic indicators and structural trends. By applying machine learning models to sentiment-driven meme stocks and machine learning algorithms to macroeconomic data, the study demonstrated the effectiveness of AI in capturing complex financial patterns. In addition to summarising model performance and insights gained through interpretability tools like SHAP, this chapter outlines the practical financial implications of the findings, acknowledges limitations and offers recommendations for future research.

7.2 Research Overview – Meme Stock Price Prediction

With the rise of online platforms, social media forums like WSB, and commission-free trading apps like Robinhood, retail investors have gained significant influence in stock trading over the past few years. Especially with new technologies and access to information becoming even more democratised, retail investors have shown that they can impact specific stocks, market trends and even stock market narratives in ways that were previously unimaginable. Hence, this study aimed to analyze daily discussions by potential investors on Reddit WSB, one of the largest online forums for retail investors, to detect the most popular meme stocks among the community and analyze the investors' sentiment towards them, as well as to predict the next-day closing price of the relevant meme stocks using RNN models based on historical price data and technical indicators. The study was based on four main models as single single-layered LSTM, regular stacked LSTM, Bi-LSTM and GRU across the top three meme stocks (NVDA, ASTS, LUNR, SMCI, AI) that had been detected on the span of three consecutive days from August 26 to August 28.

According to the results, the single-layer LSTM model consistently performed best in predicting the next-day price of almost all the stocks. Its simplicity and ability to capture core patterns without overcomplicating the learning process proved to be effective in meme stock price prediction. Conversely, more complex models like the regular stacked LSTM and Bi-LSTM introduced added variance, particularly with stocks like SMCI, where the volatility was more pronounced. While these models had the potential to capture deeper patterns, they often struggled with the increased complexity of the meme stock price data. And as for GRU, despite its reputation for simplicity and efficiency, it constantly underperformed compared to the LSTM variants, particularly with stocks that had longer historical data like NVDA. Particularly because meme stocks are known for their high volatility and nuanced price movements, which are driven by online sentiments and technical indicators, and the GRU model's fewer gates and parameters limited its ability to model the intricate patterns presented in meme stocks.

7.3 Accomplishment of Research Objectives

Develop a method to identify popular meme stocks daily:

The first objective of this study was accomplished by analysing discussions on the WallStreetBets subreddit using PRAW – Python Reddit API Wrapper. The sentiment analysis on the posts' titles and comments was performed using the VADER sentiment analysis tool, which allowed for the identification of the most frequently mentioned stocks. By setting up filters based on the number of times the stocks were mentioned, the top 10 daily popular meme stocks were identified, and out of them top three meme stocks were selected for the next day price prediction.

Implement RNN models:

To predict the next day's closing price of the selected meme stocks, several RNN models were developed and implemented. The models consisted of Single Laired LSTM, Regular Stacked LSTM, Bi-LSTM and GRU. These models were trained on historical financial data from Yahoo Finance along with several other technical indicators. The dataset for most stocks were selected based on the day they started issuing stocks to the public since most of the stocks have entered the market fairly recently, and for other stocks that have been in the market for a longer period of time, data from last ten years were selected to generate predictions for the popular stocks.

Evaluate model performance:

The prediction accuracy of the implemented RNN models was evaluated using error metrics such as RMSE, MAE and MAPE. Further, the predicted values of each stock were compared against the actual stock prices for three consecutive days, providing valuable insights into the performance of each model for each meme stock. According to both forms of evaluations, the single-layered LSTM model emerged as the best-suited model for next-day price prediction of meme stocks, while GRU demonstrated the weakest performance.

Provide recommendations to investors and future studies in the field:

Based on the model evaluations and sentiment analysis findings, practical recommendations were provided through the latter part of the study to assist the retail investors and financial analysts in making informed decisions regarding meme stocks. Additionally, the study highlighted key areas for future studies, including the incorporation of sentiment analysis data into prediction models, the improvement of model accuracy and the application of the methodology to other kinds of novel investments. Recommendations were provided for both investors and researchers to enhance the understanding and application of machine learning in the meme stock domain.

7.4 Limitations and Recommendations

Although this study provides valuable insights into the analysis and prediction of novel meme stocks, several recommendations can be made for future studies and practical applications. As mentioned in the methodology, this study excluded sentiment data from the prediction models mainly due to the lack of accessibility to Reddit's past discussions. Hence, future studies can benefit from incorporating sentiment analysis on discussion platforms like r/wallstreetbets and r/investing by collaborating with data providers or using paid data services to obtain comprehensive datasets that would enable a more thorough exploration of how retail investor sentiment directly impacts stock prices. Further, in addition to data from subreddits, future studies could explore other data sources such as news articles, earnings reports and other social media sentiments such as Twitter, LinkedIn and TikTok; incorporating a broader range of factors may help the models improve the prediction accuracy.

As expressed in the discussion, incorporating sentiment data directly into meme stock price prediction can sometimes lead to inaccuracies, especially when dealing with highly speculative or volatile retail sentiment during the peak of the stock prices. To mitigate this, future studies could follow a more refined approach, such as applying lower weights to extreme sentiment spikes to avoid over-reliance on short-term hype or smoothing sentiment over time and using time-lagged sentiment data that can help capture its long-term effects on price without overreacting to fleeting market excitement.

According to the results, simpler models like single-layer LSTM are better suited for highly volatile meme stock price prediction. Hence, future studies could experiment with more straightforward yet sophisticated models such as exponential moving average (EMA), K-nearest neighbours (KNN) or simple feedforward neural networks, as these models rely more on the recent past when generating future predictions. Therefore, these models could provide valuable insights and predictions for meme stocks, which often display unpredictable, shorter and sentiment-driven behaviour.

Additionally, the current study focuses on next-day price predictions. Expanding the prediction horizon to include weekly or monthly forecasts could provide additional insights into the effectiveness of the model, which could help investors evaluate long-term versus short-term trends to make prompt decisions in trading.

moreover, cryptocurrencies are another growing market that is highly volatile and highly influenced by retail investors and social media sentiments, which makes them ideal for sentiment-driven analysis. Platforms like r/Cryptocurrency, with millions of members and highly active discussions, serve as a rich data source for capturing investors' moods and speculations towards different cryptocurrencies. therefore, future studies could apply the model of this study on crypto market-related discussions to identify its influence on the new generation of investments.

In conclusion, this thesis has demonstrated how RNN models, particularly single-layered LSTM, can effectively predict the next-day price of highly volatile meme stocks when trained on historical and technical price data. The study highlighted the importance of balancing model complexity with the characteristics of the stock being analysed. Although the sentiment analysis was not directly integrated into the prediction model, it played a vital role in identifying the most popular meme stocks among the WSB community and their collective sentiment towards the stocks. As meme stocks continue to play a significant role in modern technology and sentiment-driven financial markets, this study contributes valuable insight into how simple, yet robust models can be tailored to the unique behaviour of meme stocks, offering a practical tool for traders to make more informed decisions in the rapidly changing financial landscape.

7.5 Research Overview – UK GDP Prediction

This study was aimed at predicting the gross domestic product of the United Kingdom using machine learning approaches by incorporating various macroeconomic variables such as energy prices, unemployment rate, real effective exchange rate, inflation and net migration. This research applied several machine learning techniques, including SVR, RF, GBM and LR as a baseline model. Data scaling and hyperparameter tuning were applied to these models to find the best fit for each model in the dataset. One of the key focuses of this research was to employ interpretability to the models by incorporating SHAP to assess the influence of macroeconomic variables on GDP prediction.

The study compared different models based on several performance metrics consistent with RMSE, MAE, MAPE and R^2 scores. The findings revealed that the hyperparameter-tuned SVR models with scaled data attain the best performance in predicting the GDP of the UK. Although the default SVR model with unscaled data attained the worst performance, it remarkably failed to detect the relationships between macroeconomic predictors and GDP. Notably, this version of the SVR model was significantly below the baseline LR model with a more straightforward structure. Additionally, SHAP analysis provided insight into the relative importance of each variable, with net migration and energy prices consistently emerging as key drivers of the UK's GDP. At the same time, inflation and REER demonstrated limited influence.

The conclusion drawn from these findings highlighted the importance of feature scaling and tuning, considering the nature and inherent structure of the selected machine learning models, and selecting the models based on the nature of the relationship between the input and target values. It also discussed the economic implications of macroeconomic variables' influence on the UK economy.

7.6 Accomplishment of Research Objectives

- Analyse the impact of key macroeconomic variables, including energy prices, on GDP prediction.

This study explored the influence of key macroeconomic variables, such as energy prices, unemployment rate, net migration, inflation and REER, on the UK's GDP predictions. The relationship between these predictors and the GDP was analysed through correlation analysis, and pair plots in the preprocessing as well as through SHAP value analysis after the predictions, allowing the identification of the most influential features. Out of these variables, net migration and energy prices, particularly gas, solid fuels and domestic fuels, were found to play an essential role in GDP prediction, highlighting their significance to economic activities in the UK.

- Application of machine learning models – Support Vector Regression (SVR), Random Forest (RF), Gradient Boosting Machines (GBM) and a statistical model - Linear Regression (LR) as a baseline model to predict GDP.

Various machine learning models, such as SVR, RF, GBM, and LR, were applied to predict the UK's GDP. The models were trained using historical macroeconomic data from 1990 – 2018, allowing for comparison across different techniques. These models were optimised using hyperparameter tuning and tested with both scaled and unscaled features to determine the impact of data preprocessing on the predictive performance of different models.

- Application of SHAP to interpret the contribution of each macroeconomic variable to the predictions of each model.

SHAP analysis was applied to each version of every model. The SHAP values were used to interpret and explain the contribution of individual macroeconomic variables to the model predictions. According to the analysis, net migration and energy prices emerged as the key contributors to the predictions. However, other variables, such as inflation and REER, displayed minimal influence on predictions. Visualisation of SHAP summary plots and mean absolute SHAP plots was employed to effectively communicate the findings.

- Compare the performance of these models using evaluation metrics such as Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and R^2 score.

The performance of the models was evaluated using several performance metrics, including RMSE, MAE and MAPE to determine the accuracy and reliability of each model. Moreover, the R^2 score was utilised to determine the model's ability to identify variance in the target value (GDP). The hyperparameter-tuned SVR model with scaled features emerged as the best performing, demonstrating low error metrics and the highest R^2 score. On the other hand, the unscaled SVR model with unscaled data demonstrated below-par performance with significantly high error metrics and a noteworthy low R^2 score. The comparison of performance metrics allowed for a clear evaluation of the model's strengths and weaknesses.

7.7 Limitations

Although this study produced a valuable contribution to the current body of literature on economic forecasting, it had to face certain limitations. For instance, the data used for the predictions span from 1990 to 2018, mainly due to the inconsistency and availability of economic data. also, the time period during and post-COVID-19 was excluded from this study since this unexpected global pandemic caused extreme disruptions to the economies, which could lead the models to be biased by this extreme fluctuation, which could compromise the validity of the predictions in more stable times.

This study relied only on annual data from 1990-2018, which limited the models' ability to capture short-term fluctuations and sudden economic shocks. The real-time or high-frequency data, such as monthly or quarterly data, were not incorporated into the study due to the difficulty of accessing real-time or short-term data on a number of macroeconomic variables and the inconsistency of the high-frequency data. therefore, annual data were selected due to their availability and completeness, which allowed for a more consistent approach. However, the use of high frequency could have enhanced the models' ability to reflect more granular changes in the economy, especially for more complex models such as GBM and RF.

Given the relatively small dataset, modern features that impact the economic direction, such as technological changes, global trade policies, and social and news sentiments, were not incorporated into the study. This was partially due to the challenges in acquiring such data, and also because of the possibility of a large number of non-standard macroeconomic features creating issues such as overfitting the models.

Finally, this study also faced limitations in a detailed analysis of the model's response to economic outliers such as recessions, booms or external shocks. The consideration of unexpected events is essential for assessing the strength and dependability in real-life scenarios. However, due to limited data availability and the exclusion of the post-2018 period to minimise the impact of COVID-19. Further, analysing outliers would require more complex methods to address heteroscedasticity and non-stationarity, which was beyond the scope of the research.

7.8 Recommendations

Based on the limitations and findings of this study, future research could consider expanding the dataset to include more extreme economic events, such as the COVID-19 pandemic, to improve the robustness of the prediction models. Additionally, incorporating high-frequency data such as quarterly or monthly observations would enable a more detailed analysis of short-term economic fluctuations and provide insight into how the short-term fluctuations and rapid changes in macroeconomic variables impact the GDP.

This study focuses on traditional macroeconomic indicators. Future studies could enhance the features by including more variables such as consumer confidence indices, international trade figures, government policy indicators and technological advancements. These additional features could improve the ability to capture complexities in the modern economy. Further, future studies could incorporate feature selection techniques to help identify the most relevant predictors and avoid any unnecessary complexities arising from too many features.

Future studies could focus on analysing outliers of the economies, such as recession and financial crises, and their impact on the GDP. This could help policymakers foresee the impact of such events when they arise using such robust models.

Furthermore, future studies could apply the methodology utilised in this study to other economies to see its applicability to diverse economic conditions. Applying this method to more developing countries would reveal interesting findings that could enhance the current understanding of the application of machine learning to economic forecasting.

While RMSE, MAPE, and R^2 score were used as core evaluation metrics for both studies due to their widespread adoption and effectiveness in regression tasks. However, future work could benefit from exploring more novel and context-specific evaluation techniques to enhance accuracy and real-world applications. These methods can include directional accuracy, which assesses whether the model accurately predicts the direction of change in stock prices, confidence intervals to capture uncertainty in forecasts, and quantile loss for a more nuanced view of prediction errors across different value ranges. Additionally, economic impact evaluation methods such as estimation of the financial cost of prediction errors as well as Sharpe ratio-style metrics, can be incorporated to generate more insights in a trading or investment context. Therefore, these advanced evaluation approaches could strengthen future research by aligning model assessment more closely with practical decision-making in finance and economic domains.

Furthermore, this study applied SHAP as an interpretation of the contribution of each macroeconomic variable to GDP prediction. Future research could explore more accessible ways to present these insights, particularly to policymakers. Therefore, developing customised visualisation techniques and domain-specific explanations could help bridge the gap between model interpretability and practical economic decision-making. Additionally, consideration of models such as Generalised Additive Models (GAMs) would produce a more straightforward understanding of the relationship between variables without compromising the accuracy.

8 References

Abonazel, M. & Abd-Elftah, A., 2019. Forecasting Egyptian GDP using ARIMA models. *Reports on Economics and Finance,* 5(1), pp. 35-47.

Abraham, R. et al., 2022. Forecasting a stock trend using genetic algorithm and random forest. *Journal of Risk and Financial Management,* 15(5), p. 188.

Abreu, I. & Mesias, F., 2020. The assessment of rural development: Identification of an applicable set of indicators through a Delphi approach. *Journal of Rural Studies,* Volume 80, pp. 578-585.

Adam, R., 2019. *Brexit: causes and consequences..* s.l.:Springer.

Afrouzi, H. & Bhattarai, S., 2023. Inflation and gdp dynamics in production networks: A sufficient statistics approach. *National Bureau of Economic Research.*

Agrawal, A. & Bhardwaj, V., 2021. *Methods of Sentiment Analysis for Hindi and English Languages.* s.l., In 2021 International Conference on Computing Sciences (ICCS), pp. 295-298.

Ahmmed, M., Babu, M., Hoque, M. & Rahman, M., 2021. A PLS-SEM approach to connect fertility, GDP, and childhood mortality with female life expectancy (FLE) in Bangladesh. *AIUB Journal of Science and Engineering (AJSE),* 20(4), pp. 151-157.

Alam, T., Sultana, N. & Rayhan, M., 2019. Structural equation modeling: an application of broadband penetration and GDP growth in Asia. *Economic Structures,* 8(17).

Ali Shah, A., Syed, G. & Shaikh, F., 2013. Effects of Macroeconomic Variables on Gross Domestic Product (GDP) in Pakistan. *Procedia Economics and Finance,* Volume 5, pp. 703-711.

Allan, G. et al., 2007. The impact of increased efficiency in the industrial use of energy: a computable general equilibrium analysis for the United Kingdom. *Energy Economics,* 29(4), pp. 779-798.

Aloosh, A., Choi, H. & Ouzan, S., 2021. Meme stocks and herd behavior. *SSRN 3909945.*

Alostad, H. & Hasan, D., 2015. *Directional prediction of stock prices using breaking news on twitter.* s.l., s.n., pp. 523-530.

Altig, et al., 2020. Economic uncertainty before and during the COVID-19 pandemic. *Journal of public economics,* Volume 191, p. 104274.

Andreas, B., Roger, E., Farmer, A. J. H. & Massimiliano, M., 2008. Factor analysis in a model with rational expectations. *Econometrics Journal, Royal Economic Society,* 11(2), pp. 271-286.

Anesti, N., Kalamara, E. & Kapetanios, G., 2021. Forecasting UK GDP growth with large survey panels. *Bank of England Working Paper No. 923.*

Angel, D., Marco, Fohlin, C. & Weidenmier, M. D., 2021. Do global pandemics matter for stock prices? Lessons from the 1918 Spanish Flu. *National Bureau of Economic Research,* Volume w28356.

Arai, M., Kinnwall, M. & Thoursie, P., 2004. Cyclical and causal patterns of inflation and GDP growth. *Applied Economics,,* 36(15), pp. 1705-1715.

Araújo, A, R. d., Vasconcelos, G. C. & Ferreira, T. A., 2007. An evolutionary morphological-rank-linear approach for time series prediction. *IEEE Congress on Evolutionary Computation,* pp. 4321-4328.

Armstrong, M., 2001. *A handbook of human resources management practices.* London: London: Book Power/ELST.

Assunção, J. & Fernandes, P., 2024. The Surprising Robustness of Partial Least Squares. *arXiv preprint.*

Atif, D., 2024. Enhancing Long-Term GDP Forecasting with Advanced Hybrid Models: A Comparative Study of ARIMA-LSTM and ARIMA-TCN with Dense Regression. *Computational Economics,* pp. 1-27.

Ayhan, F. & Elal, O., 2023. The IMPACTS of technological change on employment: Evidence from OECD countries with panel data analysis. *Technological Forecasting and Social Change,* Volume 190.

Bahelka, A. & Weerd, H., 2024. Comparative analysis of Mixed-Data Sampling (MIDAS) model compared to Lag-Llama model for inflation nowcasting. *Research Gate.*

Baker, et al., 2020. How does household spending respond to an epidemic? Consumption during the 2020 COVID-19 pandemic. *The Review of Asset Pricing Studies,* 10(4), pp. 834-862.

Baker, et al., 2016. Measuring economic policy uncertainty. *The quarterly journal of economics,* 131(4), pp. 1593-1636.

Bandi, H., Joshi, S., Bhagat, S. & Ambawade, D., 2021. *Integrated Technical and Sentiment Analysis Tool for Market Index Movement Prediction, comprehensible using XAI.* s.l., s.n.

Barnett, A., Mumtaz, H. & Theodoridis, K., 2014. Forecasting UK GDP growth and inflation under structural change. A comparison of models with time-varying parameters. *International Journal of Forecasting,* 30(1), pp. 129-143.

Barrell, R., Choy, A., Holland, D. & Riley, R., 2005. The sterling effective exchange rate and other measures of UK competitiveness. *National Institute Economic Review,* Volume 191, pp. 54-63.

Baumeister, C. & Peersman, G., 2013. The role of time-varying price elasticities in accounting for volatility changes in the crude oil market. *Journal of applied econometrics,* 28(7), pp. 1087-1109.

Bean, C., 2004. Inflation targeting: the UK experience. *Perspektiven der Wirtschaftspolitik,* 5(4), pp. 405-421.

Bellofatto, Anthony, D'Hondt, C. & Winne, R. D., 2018. Subjective financial literacy and retail investors' behavior. *Journal of Banking & finance,* Issue 92, pp. 168-181.

Benhamou, E. et al., 2021. Adaptive learning for financial markets mixing model-based and model-free rl for volatility targeting. *arXiv.*

Bhandari, P. & Frankel, J., 2017. Nominal GDP Targeting for Developing Countries. *Research in Economics,* 71(3), pp. 491-506.

Bharathi, S. & Navaprakash, N., 2024. *Prediction of Indian GDP using XGBoost Compared with Adaboost.* s.l., s.n., pp. 1-4.

Bilych, A. V., 2013. Theoretical and practical aspects of business valuation based on DCF-method (discounted cash flow). *Економіка АПК ,* Volume 8, pp. 78-84.

Blanchard, O., Cerutti, E. & Summers, L., 2015. *Inflation and activity–two explorations and their monetary policy implications,* s.l.: National Bureau of Economic Research.

Boehmer, et al., 2021. Tracking retail investor activity. *The Journal of Finance,* 76(5), pp. 2249-2305.

Bogle, S. A. & Potter, W. D., 2015. SentAMaL-a sentiment analysis machine learning stock predictive model. *Computer Engeeniering and Applied Computing,* p. 610.

Bolland, J, P. & Connor, J. T., 1997. A constrained neural network Kalman filter for price estimation in high frequency financial data. *International Journal of Neural Systems,* 8(4), pp. 399-415.

Bollen, K., 1989. A New Incremental Fit Index for General Structural Equation Model. *Sociological Methods & Research,* Volume 17, pp. 303-316.

Borjas, G., 2014. *Immigration economics.* s.l.:Harvard University Press.

Box, G. E. & Jenkins, G. M., 1976. Time series analysis: Forecasting and control. *Holden-Day.*

Breiman, L., 2001. Random Forest. *Machine Learning,* Volume 45, pp. 5-32.

Buckmann, M., Joseph, A. & Robertson, H., 2021. Opening the black box: Machine learning interpretability and inference tools with an application to economic forecasting. In Data Science for Economics and Finance: Methodologies and Applications. *Cham: Springer International Publishing,* pp. 43-63.

Burbidge, J. & Harrison, A., 1984. Testing for the effects of oil-price rises using vector autoregressions. *International economic review,* pp. 459-484.

Bussière, M., Lopez, C. & Tille, C., 2015. Do real exchange rate appreciations matter for growth?. *Economic Policy,* 30(8), pp. 5-45.

Campo, F., Forte, G. & Portes, J., 2024. The Impact of Migration on Productivity: Evidence from the United Kingdom. *The BE Journal of Economic Analysis & Policy,* 24(2), pp. 537-564.

Cao, Chunfang, Changyuan, X. & Chan, K. C., 2016. Social trust and stock price crash risk: Evidence from China. *International Review of Economics & Finance,* Volume 46, pp. 148-165.

Cao, L. & Tay, F., 2021. Financial Forecasting Using Support Vector Machines. *Neural Computing and Application,* Volume 10, pp. 184-192.

Çelik & Erkan, İ., 2020. The effects of exchange rate and interest rate exposure on the stock returns and volatility of Turkish Insurance Companie. *İstanbul İktisat Dergisi,* 70(1), pp. 141-161.

Chaudhry, Sayan & Kulkarni, C., 2021. *Design patterns of investing apps and their effects on investing behaviors.* s.l., In Proceedings of the 2021 ACM Designing Interactive Systems Conference, pp. 777-788.

Chien, T. & Hu, J., 2008. Renewable energy: An efficient mechanism to improve GDP. *Energy policy,,* 36(8), pp. 3045-3052.

Chien, T. & Hu, J., 2008. Renewable energy: An efficient mechanism to improve GDP. *Energy policy,* 36(8), pp. 3045-3052.

Chinn, M., 2006. A primer on real effective exchange rates: determinants, overvaluation, trade flows and competitive devaluation. *Open economies review,* Volume 17, pp. 115-143.

Choi, B., 2012. *ARMA model identification.* s.l.:Springer Science & Business Media.

Chowdhury, A., Hamid, M. & khi, R., 2019. Impact of macroeconomic variables on economic growth: Bangladesh perspective. Information. *Management and Computer Science,* 2(2), pp. 19-22.

Chu, B. & Qureshi, S., 2023. Comparing out-of-sample performance of machine learning methods to forecast US GDP growth. *Computational Economics,* Volume 4, pp. 1567-1609.

Claveria, O., 2021. Uncertainty indicators based on expectations of business and consumer surveys. *Empirica,* 48(2), p. 483–505.

CLEMEN, 2024. *listsof30.* [Online]
Available at: https://www.listsof30.com/p/30-wallstreetbets-wsb-terms
[Accessed 06 04 2024].

Coelho, et al., 2019. *Social media and forecasting stock price change.* s.l., In 2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC), pp. 195-200.

Comunale, M., 2017. Dutch disease, real effective exchange rate misalignments and their effect on GDP growth in EU. *Journal of International Money and Finance,* Volume 73, pp. 350-370.

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

Cooley, F, T. & Hansen, G. D., 1991. The welfare costs of moderate inflations. *Journal of Money, Credit and Banking,* 23(3), pp. 483-503.

Crafts, N. & Mills, T., 2020. Is the UK productivity slowdown unprecedented?. *National Institute Economic Review,* Volume 251, pp. R47-R53.

Crook, C. & Garratt, D., 2005. The positivist paradigm in contemporary social science research. *Research methods in the social sciences,* Volume 207, p. 214.

Das, R. & Thoudam, D. S., 2023. Multimodal sentiment analysis: a survey of methods, trends, and challenges. *ACM Computing Surveys 55,* Volume 13, pp. 1-38.

Dawkins, R., 2016. *The selfish gene.* s.l.:Oxford university press.

DeJong, D., Liesenfeld, R. & Richard, J., 2005. A nonlinear forecasting model of GDP growth. *Review of Economics and Statistics,* 87(4), pp. 97-708.

Department of Business, Energy and Industrial Strategy, 2016. *Energy and Industrial Strategy,* s.l.: Deapartment of Energy.

Di Nino, V., Eichengreen, B. & Sbracia, M., 2011. Real exchange rates, trade, and growth: Italy 1861-2011. *Bank of Italy Economic History Working Paper,* Volume 10.

Dickinson, Brian & Hu, W., 2015. Sentiment analysis of investor opinions on twitter. *Social Networking,* 4(3), pp. 62-71.

Ding, X., Yue, Z., Ting, L. & Duan, J., 2015. *Deep learning for event-driven stock prediction.* s.l., In Twenty-fourth international joint conference on artificial intelligence.

Ditta, A., Ibraheem, R. & Ayub, M., 2021. Analyzing the Impact of Real Effective Exchange Rate and Investment on Trade Deficit in the United Kingdom. Pakistan. *Journal of Humanities and Social Sciences,* 9(2), pp. 241-249.

Divya, K. & Rama Devi, V., 2014. A Study on Predictors of GDP: Early Signals. *Procedia Economics and Finance,* Volume 11, pp. 375-382.

Dollar, D., 1992. Outward-oriented developing economies really do grow more rapidly: evidence from 95 LDCs, 1976-1985. *Economic development and cultural change,* 40(3), pp. 523-544.

Edelstein, P. & Kilian, L., 2007. Retail energy prices and consumer expenditures. *ELSEVIER.*

Edey, M., 2009. The global financial crisis and its effects. Economic Papers. *A journal of applied economics and policy,* 28(3), pp. 186-195.

Elbagir, S. & Yang, J., 2019. *Twitter sentiment analysis using natural language toolkit and VADER sentiment.* s.l., In Proceedings of the international multiconference of engineers and computer scientists.

Elliott, Graham & Timmermann, A., 2008. Economic Forecasting. *Journal of Economic Literature,* 46(1), pp. 3-56.

Enders, W. & Jones, P., 2016. Grain prices, oil prices, and multiple smooth breaks in a VAR. *Studies in Nonlinear Dynamics & Econometrics,* 20(4), pp. 399-419.

European Central Bank, 2023. *European Central Bank Annual Report,* Frankfurt, Germany: European Central Bank.

Fattouh, B. & Mahadeva, L., 2013. OPEC: what difference has it made?. *Annu. Rev. Resour. Econ,* 5(1), pp. 427-443.

Flemming, J., 2010. UK macro-policy response to oil price shocks of 1974–1975 and 1979–1980. *European Economic Review,* 18(1), pp. 223-234.

Forni, M., Gambetti, L. & Sala, L., 2023. Macroeconomic uncertainty and vector autoregressions. *Econometrics and Statistics,* 15(4), pp. 101-115.

Friedman, J., 2001. Greedy function approximation: A gradient boosting machine. *The Annals of Statistics,* 29(5), pp. 1189-1232.

Fung & Glenn, M., 2003. Machine learning and data mining via mathematical programming-based support vector machines. *The University of Wisconsin-Madison.*

Gao, Y., Wang, R. & Zhou, E., 2021. prediction based on optimized LSTM and GRU models. *Scientific Programming,,* 1(1), p. 4055281.

Gao, Y., Wang, R. & Zhou, E., 2021. Stock prediction based on optimized LSTM and GRU models. *Scientific Programming,* Volume 1, p. p.4055281.

Gisser, M. & Goodwin, T., 1986. Crude oil and the macroeconomy: Tests of some popular notions: Note. Journal of Money. *Credit and Banking,* 18(1), pp. 95-103.

Gómez, V., Agustín, M. & Daniel, P., 1999. Missing observations in ARIMA models: Skipping approach versus additive outlier approach. *Journal of Econometrics,* 88(2), pp. 341-363.

Grainca, B., 2022. Methodological Basis of Unemployment. *Journal of Social, Political and Economic Research,* 9(1), pp. 163-183.

Granger, C. & Engle, R., 1987. Econometric forecasting: A brief survey of current and future techniques. *Climatic Change,* Issue 11, p. 117–139.

Grubb, M. & Newbery, D., 2018. UK electricity market reform and the energy transition: Emerging lessons. *The Energy Journal,* 36(9), pp. 1-26.

Gujarati, D., 2009. *Basic econometrics.* s.l.:s.n.

Hamilton, J., 1983. Oil and the macroeconomy since World War II.. *Journal of political economy,* 91(2), pp. 228-248.

Haque, M. S. et al., 2023. Boosting stock price prediction with anticipated macro policy changes. *Journal of Mathematics and Statistics Studies,* 4(3), pp. 29-34.

Hausmann, R., Pritchett, L. & Rodrik, D., 2005. Growth accelerations. Journal of economic growth. *Journal of economic growth,* 10(3), pp. 303-329.

Hodson, D. & Mabbett, D., 2009. UK economic policy and the global financial crisis: paradigm lost?. *JCMS: Journal of common market studies,* 47(5), pp. 1041-1061.

Ho, S. & M. Xie, V. 3. I. 1. 1., 2021. The use of ARIMA models for reliability forecasting and analysis. *Computers & Industrial Engineering,* 32(1-2), pp. 213-216.

Hossain, A., Hossen, M., Hasan, M. & Sattar, A., 2021. *GDP growth prediction of bangladesh using machine learning algorithm.* s.l., Technologies and Virtual Mobile Networks (ICICV).

Hossain, et al., 2018. *Hybrid deep learning model for stock price prediction.* s.l., In 2018 ieee symposium series on computational intelligence (ssci) IEEE, pp. 1837-1844.

Hsu, D., 2007. Experienced Entrepreneurial Founders, Organizational Capital, and Venture Capital Funding. *Research Policy,* Volume 36, pp. 722-741.

Hutto, C. & Gilbert, E., 2014. *Vader: A parsimonious rule-based model for sentiment analysis of social media text.* s.l., s.n., pp. 216-225.

Hu, Y. & Wang, Z., 2022. Support Vector Machine-Based Nonlinear Model and Its Application in Regional Economic Forecasting. Wireless. *Communications and Mobile Computing,* Volume 1, p. 3387543.

Hu, Z., 2021. Crude oil price prediction using CEEMDAN and LSTM-attention with news sentiment index. *Oil & Gas Science and Technology–Revue d'IFP Energies nouvelles,* Volume 76, p. 28.

Ibrahim, et al., 2019. Improved salp swarm algorithm based on particle swarm optimization for feature selection. *Journal of Ambient Intelligence and Humanized Computing,* Volume 10, pp. 3155-3169.

IMF, 2024. *IMF World Economic Outlook,* Washington, DC: IMF Media Center.

J. Roush, K. S. & Hu, G., 2017. *Predicting gross domestic product using autoregressive models.* London, UK, s.n., pp. 317-322.

Jafari-Samimi, A., Shiraz, B. & Fazlollahtabar, H., 2007. A comparison between time series, exponential smoothing, and neural network methods to forecast GDP of Iran. *Iranian Economic Review,* 12(19), pp. 19-35.

Jahan, S., Mahmud, A. & Papageorgiou, C., 2014. What is Keynesian economics. In: s.l.:International Monetary Fund, pp. 53-54.

Jallow, H., Mwangi, R., Gibba, A. & Imboga, H., 2025. Transfer learning for predictiGross domestic product growth based on remittance inflows using RNN-LSTM hybrid model: a case study of The Gambia. *Frontiers in Artificial Intelligence,* Volume 8, p. 151.

Jena, P. et al., 2021. Impact of COVID-19 on GDP of major economies: Application of the artificial neural network forecaster. *Economic Analysis and Policy,* Volume 69, pp. 324-339.

Jena, et al., 2021. Impact of COVID-19 on GDP of major economies: Application of the artificial neural network forecaster. *Economic Analysis and Policy,* Volume 69, pp. 324-339.

Jeong, C. (., 2024. Fine-tuning and Utilization Methods of Domain-specific LLMs. *ResearchGate.*

Jing, Nan, Z. W. & Wang, H., 2021. A hybrid model integrating deep learning with investor sentiment analysis for stock price prediction. *Expert Systems with Applications,* Volume 178, p. 115019.

Joseph, A., Potjagailo, G., Chakraborty, C. & Kapetanios, G., 2024. Forecasting UK inflation bottom up. *International Journal of Forecasting,* 40(4), pp. 1521-1538.

Kappler, M., Reisen, H., Schularick, M. & Turkisch, E., 2011. The macroeconomic effects of large exchange rate appreciations. *ZEW-Centre for European Economic Research Discussion Paper,* pp. 11-16.

Karagöz, K. & Keskin, R., 2016. Impact of fiscal policy on the macroeconomic aggregates in Turkey: Evidence from BVAR model. *Procedia economics and finance,* Volume 38, pp. 408-420.

Keogh-Brown, et al., 2020. The impact of Covid-19, associated behaviours and policies on the UK economy: A computable general equilibrium model. *SSM-population health,* Volume 12, p. 100651.

Khairani, F., Kurnia, A., Aidi, M. & Pramana, S., 2022. Predictions of Indonesia economic phenomena based on online news using Random Forest. *Sinkron: jurnal dan penelitian teknik informatika,* 6(2), pp. 532-540.

Khan, S. & Chhapra, I., 2016. Analysis of GDP and macroeconomic variables on economic growth of Pakistan. *Market Forces,* 11(1).

Khan, W. et al., 2022. Stock market prediction using machine learning classifiers and social media, news. *Journal of Ambient Intelligence and Humanized Computing,* pp. 1-24.

Khiabani, N. & Naderian, M., 2017. Speculative Oil Demand and Crude Oil Price Dynamics: A TVP-VAR Approach. *Planning and Budgeting,* 22(3), pp. 3-44.

Kilian, L., 2008. The Economic Effects of Energy Price Shocks. *Journal of Economic Literature,* 46(4), p. 871–909.

Kim, H., Kim, H. & Choi, S., 2024. Investigating the Impact of Agricultural, Financial, Economic, and Political Factors on Oil Forward Prices and Volatility: A SHAP Analysis. *Energies,* 17(5), p. 1001.

Kim, Young, H. & Won, C. H., 2018. Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models. *Expert Systems with Applications,* Volume 103, pp. 25-37.

Kirshner, J., 2001. The political economy of low inflation. *Journal of Economic Surveys,* 15(1), pp. 41-70.

Kirtac, K. & Germano, G., 2024. Sentiment trading with large language models. *Finance Research Letters,* 04.62(B).

Kitov, I., 2008. GDP growth rate and population. *arXiv.*

Koop, C. & Lodge, M., 2014. Exploring the co-ordination of economic regulation. *Journal of European Public Policy,* 21(9), pp. 1311-1329.

Korkmaz, Ö., 2019. The relationship between housing prices and inflation rate in Turkey: Evidence from panel Konya causality test. *International Journal of Housing Markets and Analysis,* 13(3), pp. 427-452.

Kudaeva, M. & Redozubov, I., 2021. *Impact of migration flows on the economic activity and labor market of Russia as a whole and regionally,* s.l.: Bank of Russia.

Kumar, Gourav, Singh, U. P. & Jain, S., 2022. Swarm intelligence based hybrid neural network approach for stock price forecasting. *Computational Economics,* 60(3), pp. 991-1039.

Lachanski, M., 2016. Did Twitter Mood Really Predict the DJIA?: Misadventures in Big Data for Finance. *Penn Journal of Economics,* pp. 8-48.

Landefeld, J., Seskin, E. & Fraumeni, B., 2008. Taking the pulse of the economy: Measuring GDP. *Journal of Economic Perspectives,* 22(2), pp. 193-216.

Laygo-Matsumoto, S. & Samonte, M., 2021. *Philippine economic growth: GDP prediction using machine learning algorithms.* s.l., s.n., pp. 15-20.

Lee, C., 2006. The causality relationship between energy consumption and GDP in G-11 countries revisited. *Energy policy,* 34(9), pp. 1086-1093.

Lee, C. & Chang, C., 2008. Trend stationary of inflation rates: evidence from LM unit root testing with a long span of historical data. *Applied Economics,* 40(19), pp. 2523-2536.

Lee, C. & Wong, S., 2005. Inflationary threshold effects in the relationship between financial development and economic growth: evidence from Taiwan and Japan. *Journal of economic development,* 30(1), p. 49.

Lee, F, C., Chen, G. & Rui, O. M., 2001. Stock returns and volatility on China's stock markets. *Journal of Financial Research,* 24(4), pp. 523-543.

Lee, K. N. S. & Ratti, R., 1995. Oil shocks and the macroeconomy: the role of price variability. *The Energy Journal,* 16(4), pp. 39-56.

Li, C. & Qian, G., 2022. Stock price prediction using a frequency decomposition based GRU transformer neural network. *Applied Sciences,* 13(1), p. p.222.

Li & Jasper, Z., 2023. Robinhood's Retail Investing App: Investor Psychology and Financial Markets. *Scholarly Review Journal Summer 2023,* Volume 3.

Li, J. & Chen, W., 2014. Forecasting macroeconomic time series: LASSO-based approaches and their forecast combinations with dynamic factor models. *International Journal of Forecasting,* 30(4), pp. 996-1015.

Lin, C. Y. & João Alexandre, L. M. ,. 9. 2., 2024. Stock market prediction using artificial intelligence: A systematic review of systematic reviews. *Social Sciences & Humanities Open,* Volume 9.

Li, R., Guy, C.K. & Leung, 2021. The relationship between energy prices, economic growth and renewable energy consumption: Evidence from Europe. *Energy Reports,* 7(1), pp. 1712-1719.

Litterman, R., 1986. Forecasting with Bayesian vector autoregressions—five years of experience. *Journal of Business & Economic Statistics,* 4(1), pp. 25-38.

Liu, S. et al., 2017. Addressing the class imbalance problem in twitter spam detection using ensemble learning. *Computers & Security,* Volume 69, pp. 35-49.

Liu, Z. et al., 2024. Hidformer: Hierarchical dual-tower transformer using multi-scale mergence for long-term time series forecasting. *Expert Systems with Applications,* 04.Volume 239.

Lundberg, S., 2017. A unified approach to interpreting model predictions. *arXiv.*

Ma, C., 2020. Financial stability, growth and macroprudential policy. *Journal of International Economics,* Volume 122, p. 103259.

Maccarrone, G., Morelli, G. & Spadaccini, S., 2021. GDP forecasting: Machine learning, linear or autoregression?. *Frontiers in Artificial Intelligence,* 4(1), p. 757864.

Maeda, A., 2007. *On the World Energy Price-GDP Relationship..* s.l., International Association of Energy Economics.

Malkiel, G, B., Rinaudo, A. & Saha, A., 2018. Option writing: Using VIX to improve returns. *Journal of Derivatives,* 26(2), pp. 38-49.

Mankiw, N., 2019. Yes, there is a trade-off between inflation and unemployment. *The New York Times.*

Manning, L. & Andrianacos, D., 1993. Dollar movements and inflation: A cointegration analysis. *Applied Economics,* 25(12), pp. 1483-1488.

Martin, L., 2019. Machine learning vs traditional forecasting methods: An application to South African GDP. *Stellenbosch Economic Working Papers.*

Mathis, T., 2023. *Infinity Investing.* [Online]
Available at: https://infinityinvesting.com/wallstreetbets-slang-meaning/
[Accessed 05 04 2024].

McCamy & Christal, 2021. Retail Investors: Why Online Investing Platforms Need More Regulation and Oversight. *Brook. J. Corp. Fin. & Com. L.,* Volume 16, p. 255.

McCloskey, D., 2022. *Beyond positivism, behaviorism, and neoinstitutionalism in economics.* s.l.:University of Chicago Press.

McKibbin, W. & Stoeckel, A., 2010. The global financial crisis: Causes and consequences. *Asian Economic Papers,* 9(1), pp. 54-86.

McLeod, D. & Mileva, E., 2011. Real exchange rates and productivity growth. *Fordham University Department of Economics Discussion Paper,* Volume 4, pp. 1-27.

McNamara, R, K., Abraham & Newman, L., 2020. The big reveal: COVID-19 and globalization's great transformations. *International Organization,* 7(1), pp. 59-77.

Merlo, L., 2023. Risk Analysis of Nvidia Corporation Stock and Equally-Weighted US Portfolio. *SSRN,* p. 4641295.

Minh, et al., 2018. Deep learning approach for short-term stock trends prediction based on two-stream gated recurrent unit network. *Ieee Access,* Volume 6, pp. 55392-55404.

Mork, K., 1989. Oil and the macroeconomy when prices go up and down: an extension of Hamilton's results. *Journal of political Economy,* 97(3), pp. 740-744.

Mork, K., Olsen, Y. & Mysen, H., 1994. Macroeconomic responses to oil price increases and decreases in seven OECD countries. *The Energy Journal,* 15(4), pp. 19-35.

Mozaffari, L. & Zhang, J., 2024. *Predictive Modeling of Stock Prices Using Transformer Model.* New York, Association for Computing Machinery.

Mubarik, Y. & Riazuddin, R., 2005. An estimate of the threshold level of inflation in Pakistan. *Karachi: State Bank of Pakistan.*

Muchisha, N., Tamara, N., Andriansyah, A. & Soleh, A., 2021. Nowcasting Indonesia's GDP growth using machine learning algorithms. Indonesian. *Journal of Statistics and Its Applications,,* 5(2), pp. 355-368.

Nava, Noemi, Matteo, T. D. & Aste, T., 2018. Financial time series forecasting using empirical mode decomposition and support vector regression. *Risks,* 6(1), p. 7.

Nelson, E. & Nikolov, K., 2003. UK inflation in the 1970s and 1980s: the role of output gap mismeasurement. *Journal of Economics and Business,* 55(4), pp. 353-370..

Nesbitt, V, K. & Barrass, S., 2004. Finding trading patterns in stock market data. *IEEE Computer graphics and applications,* 24(5), pp. 45-55.

Nguyen-Trang, T., Lethi-Thu, T. & Vo-Van, T., 2024. Interval-Valued Time Series Prediction for Vietnam Stock Indicators Based on Ensemble Long Short-Term Memory Networks. *Comput Econ,* 03.

Nikfarjam, A., Ehsan, E. & Saravanan, M., 2010. *Text mining approaches for stock market prediction.* s.l., In 2010 The 2nd international conference on computer and automation engineering (ICCAE), pp. 256-260.

Nobanee, Haitham & Ellili, N. O. D., 2023. What do we know about meme stocks? A bibliometric and systematic review, current streams, developments, and directions for future research. *International Review of Economics & Finance,* Volume 85, pp. 589-602.

Nouira, R. & Sekkat, K., 2012. Desperately seeking the positive impact of undervaluation on growth. *Journal of Macroeconomics,* 34(2), pp. 537-552.

Oancea, B., 2024. *Advancing GDP Forecasting: The Potential of Machine Learning Techniques in Economic Predictions.* Olomouc, The Czech Republic, Knowledge on Economics and Management Conference Proceedings.

O'Brien, P. & Palma, N., 2023. Not an ordinary bank but a great engine of state: The Bank of England and the British economy. *The Economic History Review,* 76(1), pp. 305-329.

OECD, 2023. *OECD Economic Outlook,* Paris: OECD Publishing.

Ogaluzor, O. I. & Ebom, B. J., 2023. THE IMPACT OF FINANCE COST ON PRICE EARNINGS RATIO AND EARNINGS PER SHARE OF LISTED OIL AND GAS

COMPANIES IN NIGERIA. *British International Journal of Applied Economics, Finance and Accounting ,* 6(7), pp. 49-60.

Okun, A., 1962. The predictive value of surveys of business intentions. *The American Economic Review,* 52(2), pp. 218-225.

Oriani, F. & Coelho, G., 2016. *Evaluating the impact of technical indicators on stock forecasting.* s.l., IEEE., pp. 1-8.

Ou, Phichhang & Wang, H., 2010. Financial volatility forecasting by least square support vector machine based on GARCH, EGARCH and GJR models: evidence from ASEAN stock markets. *International Journal of Economics and Finance,* 2(1), pp. 51-64.

Outhwaite, W. & Menjívar, C., 2019. Migration crisis and “Brexit". In: *The Oxford handbook of migration crises.* s.l.:s.n., pp. 93-110.

Oxford Analytica, 2021. *Rising US retail stock trading will draw modest reform,* oxan-db: Emerald Expert Briefings .

Pagano, S, M., Sedunov, J. & Velthuis, R., 2021. How did retail investors respond to the COVID-19 pandemic? The effect of Robinhood brokerage customers on market quality. *Finance Research Letters,* Volume 43, p. 101946.

Pagolu, V. et al., 2016. *Sentiment analysis of Twitter data for predicting stock market movements.* s.l., SCOPES, pp. 1345-1350.

Parracho, Paulo, Neves, R. & Horta, N., 2010. *Trading in financial markets using pattern recognition optimized by genetic algorithms.* s.l., In Proceedings of the 12th annual conference companion on Genetic and evolutionary computation, pp. 2105-2106.

Pasinetti, L., 2000. Critique of the neoclassical theory of growth and distribution. *PSL Quarterly Review,* 53(215).

Patel, R., Choudhary, V., Saxena, D. & Singh, A., 2021. *Lstm and nlp based forecasting model for stock market analysis.* s.l., IEEE, pp. 52-57.

Pawar, Kriti, Jalem, R. S. & Tiwari, V., 2019. *Stock market price prediction using LSTM RNN..* Singapore, Springer, pp. 493-503.

Perera, P. & Fernando, A., 2024. Impact of Energy Prices and Macroeconomic Variables on GDP Prediction UK: Machine Learning Approach. *Journal of Business and Management Studies,* 6(5), p. 113–124.

Pérez-Pons, et al., 2021. Machine learning and traditional econometric models: a systematic mapping study. *Journal of Artificial Intelligence and Soft Computing Research,* 12(2), pp. 79-100.

Persson, E. & Ståhlberg, C., 2007. PE and EV/EBITDA Investment Strategies vs. the Market: A Study of Market Efficiency.

Peters, G., 2001. A linear forecasting model and its application to economic data. *Journal of Forecasting,* 20(5), pp. 315-328.

Pindyck, R. & Rubinfeld, D., 1991. *Econometric Models and Economic Forecasts.* 2 ed. New York: McGraw-Hill Book Co.

Pirani, Muskaan & al., e., 2022. *A comparative analysis of ARIMA, GRU, LSTM and BiLSTM on financial time series forecasting.* s.l., IEEE International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE).

Provost, F. & Fawcett, T., 2013. Data Science and Its Relationship to Big Data and Data-Driven Decision Making. *Big Data,* Volume 1, pp. 51-59.

r/WallStreetBets, 2022. *r/WallStreetBets glossary.* [Online]
Available at: https://www.reddit.com/r/wallstreetbets/wiki/glossary/
[Accessed 06 04 2024].

Rahman, M., 2013. Relationship among GDP, per capita GDP, literacy rate and unemployment rate. *British Journal of Arts and Social Sciences,* 14(2), pp. 169-177.

Rashid, A. & Kocaaslan, O. K., 2013. Does Energy Consumption Volatility Affect Real GDP Volatility? An Empirical Analysis for the UK. *International Journal of Energy Economics and Policy,* 3(4), pp. 384-394.

Rayevnyeva, O., Stryzhychenko, K. & Matúšová, S., 2023. Impact of Migration Processes on GDP. *Engineering Proceedings,* 39(1), p. 86.

Reddit, 2024. *Reddit Press.* [Online]
Available at: https://redditinc.com/press
[Accessed 05 06 2024].

Reddit, n.d. *Reddit API Documentation.* [Online]
Available at: https://www.reddit.com/dev/api
[Accessed 06 04 2024].

Rehman, R. & Raoof, A., 2010. Weighted average cost of capital (WACC) traditional vs new approach for calculating the value of firm. *International Research Journal of Finance and Economics,* Volume 45, pp. 7-9.

Richardson, A., van Florenstein Mulder, T. & Vehbi, T., 2021. Nowcasting GDP using machine-learning algorithms: A real-time assessment. *International journal of forecasting,* 37(2), pp. 941-948.

Rodrik, D., 2008. The real exchange rate and economic growth. *Brookings papers on economic activity,* Volume 2, pp. 365-412.

Roush, J., Siopes, K. & Hu, G., 2017. *Predicting gross domestic product using autoregressive models.* London, UK, s.n., pp. 317-322.

Rout, Kumar, A. & al, e., 2017. Forecasting financial time series using a low complexity recurrent neural network and evolutionary learning approach. *Journal of King Saud University-Computer and Information Sciences,* 29(4), pp. 536-552.

Ruiz, V., Navarro, J. & Peña, D., 2011. Relationship between Gross Domestic Product (GDP) and Hidden Wealth over the period 2000-2008: An International Study. *Electronic Journal of Knowledge Management,* 9(3), pp. 259-270.

Sa'adah, S. & Wibowo, M., 2020. *Prediction of Gross Domestic Product (GDP) in Indonesia Using Deep Learning Algorithm.* Yogyakarta, Indonesia, s.n., pp. 32-36.

Sandhya, Pasala, Bandi, R. & Himabindu, D. D., 2022. *Stock price prediction using recurrent neural network and lstm.* s.l., In 2022 6th international conference on computing methodologies and communication (ICCMC), pp. 1723-1728.

Seven, Ünal & Yılmaz, F., 2021. World equity markets and COVID-19: Immediate response and recovery prospects. *Research in International Business and Finance,* Volume 56, p. 101349.

Shams, M., Tarek, Z., El-kenawy, E. & al., e., 2024. Predicting Gross Domestic Product (GDP) using a PC-LSTM-RNN model in urban profiling areas. *Comput.Urban Sci.,* 4(3).

Shifman, L., 2013. *Memes in digital culture.* s.l.:MIT press.

Shiller, S. & Robert, J., 2020. Narrative economics: How stories go viral and drive major economic event. *Princeton University Press.*

Shobayo, O., S, A.-L., O, P. & B, O., 2024. Innovative Sentiment Analysis and Prediction of Stock Price Using FinBERT, GPT-4 and Logistic Regression: A Data-Driven Approach. *Big Data and Cognitive Computing,* 8(11).

Sílvia, D., Francesca, M. & S, A., 2019. Forecasting the UK economy with a medium-scale Bayesian VAR. *International Journal of Forecasting,* 35(4), pp. 1669-1678.

Simionescu, M., Streimikiene, D. & Strielkowski, W., 2020. What does Google Trends tell us about the impact of Brexit on the unemployment rate in the UK?. *Sustainability,* 12(3), p. 1011.

Sims & A, C., 1980. Macroeconomics and Reality. *Econometrica,* 48(1), pp. 1-48.

Soni & Narendrakumar, A., 2019. Text Classification Feature extraction using SVM. *International Journal of Innovative Research in Computer and Communication Engineering,* 7(7), pp. 3563-3569.

Sousa, et al., 2019. Peer-to-peer and community-based markets: A comprehensive review. *Renewable and Sustainable Energy Reviews,* Volume 104, pp. 367-378.

Speranza & Grazia, M., 1996. A heuristic algorithm for a portfolio optimization model applied to the Milan stock market. *Computers & operations research,* 23(5), pp. 433-441.

Spulbar, C. et al., 2021. Advanced empirical research based on structural equation modeling (SEM) regarding the impact of tax revenue on GDP dynamics at EU-28 level. *Scientific Annals of Economics and Busines.*

Srijiranon, K., Lertratanakham, Y. & Tanantong, T., 2022. A hybrid framework using PCA, EMD and LSTM methods for stock market price prediction with sentiment analysis. *Applied Sciences,* 12(21), p. 10823.

Stallone, Angela, Antonio, C. & Massimo, M., 2020. New insights and best practices for the successful use of Empirical Mode Decomposition, Iterative Filtering and derived algorithms. *Scientific reports,* 10(1), p. 15161.

Stevenson, S., 2007. A comparison of the forecasting ability of ARIMA models. *Journal of Property Investment & Finance,* 25(3), pp. 223-240.

Stock, James, H. & Watson, M., 2002. Forecasting Using Principal Components from a Large Number of Predictors. *Journal of the American Statistical Association,* Volume 97, p. 1167–79.

Stock, J. & Watson, M., 1999. Forecasting inflation. *Journal of monetary economics,* 44(2), pp. 293-335.

Sturge, G., 2024. *Migration Statistiscs,* s.l.: House of Commons Library.

Subian, A., Mulkan, D., Ahmady, H. & Kartiasih, F., 2024. Comparison Methods of Machine Learning and Deep Learning to Forecast The GDP of Indonesia. *Sistemasi: Jurnal Sistem Informasi,* 13(1), pp. 149-166.

Sun, Jingyun, Huang, S. & Wei, C., 2023. A BERT-based deontic logic learner. *Information Processing & Management 60,* Volume 4, p. 103374.

Suomi, Aino, Schofield, T. P. & Butterworth, P., 2020. Unemployment, employability and COVID19: how the global socioeconomic shock challenged negative perceptions toward the less fortunate in the Australian context. *Frontiers in psychology,* Volume 11.

Swanson, N. & White, H., 1997. A model selection approach to real-time macroeconomic forecasting using linear models and artificial neural networks. *Review of Economics and Statistics,* 79(4), pp. 540-550.

Sweidan, O., 2004. Does inflation harm economic growth in Jordan? An econometric analysis for the period 1970-2000. *International Journal of Applied Econometrics and Quantitative Studies,* 1(2), pp. 41-66.

Tayal, D. & Satya, K., 2009. Comparative analysis of the impact of blogging and micro-blogging on market performance. *International Journal,* 1(3), pp. 176-182.

Teräsvirta, T., 2005. Univariate nonlinear time series models. *SSE/EFI Working Paper Series in Economics and Finance 593, Stockholm School of Economics.*

Teräsvirta, T., Van Dijk, D. & Medeiros, M., 2005. Linear models, smooth transition autoregressions, and neural networks for forecasting macroeconomic time series: A re-examination. *International Journal of Forecasting,* 21(4), pp. 755-774.

Tian, Y., Shi, Y. & Liu, X., 2012. Recent advances on support vector machines research. *Technological and economic development of Economy,* 18(1), pp. 5-33.

Trivedi, U., Bhatt, M. & Srivastava, P., 2021. *Prevent Overfitting Problem in Machine Learning: A Case Focus on Linear Regression and Logistics Regression.* s.l., Springer.

Trotsky, L., 1991. *The revolution betrayed: what is the Soviet Union and where is it going?.* s.l.:Mehring Books.

Ülker, E. & Ülker, S., 2019. *Unemployment rate and GDP prediction using support vector regression.* s.l., s.n., pp. 1-5.

Ulrich, K., 2004. Real GDP, Real Domestic Income, and Terms-of-Trade Changes. *Journal of International Economics,* 62(1), pp. 83-106.

Umar, Zaghum & al, e., 2021. A tale of company fundamentals vs sentiment driven pricing: The case of GameStop. *Journal of Behavioral and Experimental Finance,* Volume 30.

United Nations Department of Economic and Social Affairs (UN DESA), 2019. *World Population Prospects 2019: Highlights,* New York: United Nation.

United Nations Development Program (UNDP), 2020. *Human Development Report 2020: The Next Frontier - Human Development and the Anthropocene,* New York: UNDP.

van de Ven, D. & Fouquet, R., 2017. Historical energy price shocks and their changing effects on the economy. *Energy Economics,* Volume 62, pp. 204-216.

Vargas, M., Anjos, D., Bichara, G. & Evsukoff, A., 2018. *Deep leaming for stock market prediction using technical indicators and financial news articles.* s.l., IEEE, pp. 1-8.

Velay, Marc & Daniel, F., 2018. Stock chart pattern recognition with deep learning. *arXiv.*

Victor, V., Karakunnel, J., Loganathan, S. & Meyer, D., 2021. From a recession to the COVID-19 pandemic: Inflation–Unemployment Comparison between the UK and India. *Economies,* 9(2), p. 73.

Vuletić, M., Prenzel, F. & Cucuringu, M., 2024. Fin-GAN: forecasting and classifying financial time series via generative adversarial networks. *Quantitative Finance,* 24(2), pp. 175-199.

Waddams, C., 2023. The UK Energy Price Guarantee. *Intereconomics: Review of European Economic Policy, Sciendo,* 58(1), pp. 22-26.

Warr, B. et al., 2010. Energy use and economic development: a comparative analysis of useful work supply in Austria, Japan, the United Kingdom and the US during 100 years of economic growth. *Ecological Economics,* Volume 69, p. 1904–17.

Wei, S. & Xie, Y., 2022. On the wedge between the PPI and CPI inflation indicators. *Bank of Canada Staff Working Paper,* pp. 2022-5.

Weng, X., Lin, X. & Zhao, S., 2022. *Stock Price Prediction Based On Lstm And Bert.* Japan, s.n., pp. 12-17.

Wheeler, C., Ibeh, K. & Dimitratos, P., 2008. UK export performance research: Review and implications. *International small business journal,* 26(2), pp. 207-239.

Wilson, R., 1994. The Middle East after the Gulf War: The Regional Economic Impact. In The Middle East in the New World Order. *London: Palgrave Macmillan UK,* pp. 134-142.

Wold, H., 1985. Partial Least Squares. In: Kotz, S. and Johnson, N.L., Eds.. *Encyclopedia of Statistical Sciences,* Volume 6, pp. 581-591.

World Bank Group, 2024. *Global Economic Prospects,* Washington, DC: World Bank Group.

Wu, C. & Chen, P., 2006. *Application of support vector machines in debt to GDP ratio forecasting.* s.l., s.n., pp. 3412-3415.

Wu, J. M.-T. et al., 2023. A graph-based CNN-LSTM stock price prediction algorithm with leading indicators. *Multimedia Systems,* Volume 23, p. 1751–1770.

Wuthrich, B. et al., 1998. *Daily stock market forecast from textual web data.* s.l., In SMC'98 Conference Proceedings. 1998 IEEE International Conference on Systems, Man, and Cybernetics.

Wu, Z. & Norden, E. H., 2009. Ensemble empirical mode decomposition: a noise-assisted data analysis method. *Advances in adaptive data analysis,* 1(1), pp. 1-41.

Xiang-rong, Z., Long-ying, H. & W, Z.-s., 2010. *Multiple kernel support vector regression for economic forecasting.* Melbourne, VIC, Australia, s.n., pp. 129-134.

Xing, Frank, Z., Erik, C. & Roy, E. W., 2018. Natural language based financial forecasting: a survey. *Artificial Intelligence Review,* 50(1), pp. 49-73.

Xu, J. et al., 2024. Enhancing Ensemble Stock Price Prediction with Machine Learning Techniques.. *Reserach Gate,* 10.

Xu, et al., 2018. Recurrent broad learning systems for time series prediction. *IEEE transactions on cybernetics,* 50(4), pp. 1405-1417.

Xu, Z., Gao, Y. & Jin, Y., 2014. Application of an optimized SVR model of machine learning. *International Journal of Multimedia and Ubiquitous Engineering,* 9(6), pp. 67-80.

Yahya, A, M. & Chiu, T. N., 2021. The Meme Stock Paradox. *Corporate and Business Law Journal,* 3(1).

Yang, B. L. C. L. M., Pan, K. & Wang, D., 2016. *Application of ARIMA model in the prediction of the gross domestic product.* s.l., computer and education informationization (MCEI 2016), pp. 1258-1262.

Yang, Fuwei, Chen, J. & Liu, Y., 2023. Improved and optimized recurrent neural network based on PSO and its application in stock price prediction. *Soft computing,* 27(6), pp. 3461-3476.

Yoo, B. H., 2023. *Conditional Forecasting With a Bayesian Vector Autoregression: Working Paper,* s.l.: Congressional Budget Office.

Yoon, J., 2021. Forecasting of real GDP growth using machine learning models: Gradient boosting and random forest approach. *Computational Economics,* 57(1), pp. 247-265.

Yoon, J., 2021. Forecasting of Real GDP Growth Using Machine Learning Models: Gradient Boosting and Random Forest Approach. *Comput Econ,* Volume 57, p. 247–265.

Yoo, S. & Wagner, J., 2016. A review of the hedonic literatures in environmental amenities from open space: a traditional econometric vs. spatial econometric model. *International Journal of Urban Sciences,* 20(1), pp. 141-166.

Zdravković, M., Dudko, P. & Kucherov, M., 2023. *Can Memes Beat the Market? Forecasting Financial Asset Returns by Using Social Media Data and Machine Learning.* Switzerland, Springer, pp. 423-434.

Zhang, et al., 2020. Random forest prediction intervals. *The American Statistician.*

Zhang, J. & Yu, K. F., 1998. What's the relative risk? A method of correcting the odds ratio in cohort studies of common outcomes. *JAMA,* 280(19), p. 1690–1691.

Appendix

WSB daily popular meme stock -sentiment analysis model

Meme Stock Price prediction Models

UK GDP Prediction Models