CALIFORNIA STATE UNIVERSITY, NORTHRIDGE

STOCK VOLATILITY PREDICTION: A DEEP LEARNING FRAMEWORK WITH AN ENSEMBLE OF LSTM AND GARCH

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Computer Science

By

Nithyasai Siddiraju

May 2024

The thesis of Nithyasai Siddiraju is approved:

Dr. George Wang, Professor	Date
Dr. Mahdi Ebrahimi, Assistant Professor	Date
Dr. Rashida Hasan, Assistant Professor, Chair	Date

California State University, Northridge

Table of Contents

ABSTRACT

STOCK VOLATILITY PREDICTION: A DEEP LEARNING FRAMEWORK WITH AN ENSEMBLE OF LSTM AND GARCH

By

Nithyasai Siddiraju

Master of Science in Computer Science

In this study, we propose a hybrid model that integrates Long Short-Term Memory (LSTM) networks and Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models to handle the hurdle of predicting stock market volatility. By leveraging the ability of LSTM to capture intricate non-linear temporal patterns and the efficiency of GARCH for volatility estimation, the model is expected to tackle the intrinsic challenges in financial market predictions. The performance evaluation covers three significant periods: the post-crisis recovery (2009-2017) after the 2008 financial crisis, a decade of diversified economic events (2010-2019), and the market volatility achievements triggered by COVID-19 (2018-2022). Through using a comprehensive dataset of the S&P 500, this study guarantees a robust evaluation of the model's performance in diverse economic contexts.

Extensive evidence shows that the hybrid model performs better than the conventional GARCH and SVR-GARCH models in terms of forecasting accuracy. The Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) are significantly lower in the LSTM-GARCH hybrid model, confirming the results. The enhanced accuracy holds generally across different market conditions, so the results are consistent with the argument that integrating deep learning into traditional econometric forecasting significantly improves financial market analysis. These results are essential for financial investors, risk managers, and central bank strategists. They suggest that improving predictive models with sophisticated ML techniques might be promising future research regarding an increasingly complex world of finance.

Keywords: Long-Short Term Memory(LSTM), Generalized Autoregressive Conditional Heteroskedasticity(GARCH), Deep Learning, Hybrid Model.

Chapter 1

Introduction

1.1 Motivation

Forecasting stock market movements has always been a leading, yet tough achievement, due to the vast complexity and high volatility ruling capital markets, which directly affects management of risk, investment strategies and economic policy making; therefore, it can be considered the core of financial research and practice. Traditional volatility models, e.g. GARCH [8],, have strongly contributed to understanding market behavior issues because of the many advantages they offer regarding the problem of modeling. The biggest problem is that inherently, there is a lack of recognition of the real, ever-dynamic, and non-linear behavior of markets under extreme fast-changing conditions, which we can appreciate in the fact that those days, we are dealing with financial markets exposed to thousands of sources of potential impact, from macroeconomic indicators to social media postings.

The advent of the new technology of Deep Learning (DL), especially the Long Short-Term Memory (LSTM) networks, thus, revolutionize the ability the model the complex temporal dependencies and nonlinear patterns that are inherent to the financial market data. In fact, LSTM stands to be the very promising new avenue in making financial volatility predictions more adaptive and more precise. Hence, the essential synergy between the LSTM networks and traditional econometric models such as the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) has become our major motivation for conducting the thesis. This is revolutionary in the sense that it will bring the machine-learning and econometric-modeling exercise together to significantly improve the predictive power of the stock market volatility, particularly in an environment of the evolvement characteristic of the financial market dynamics [14]

1.2 Research Questions

The purpose of this study is to comprehensively study the synergistic advantages of integrating LSTM networks with GARCH models in volatility forecasting of financial markets, focusing on four fundamental questions listed below:

1. How does integrating LSTM networks with GARCH models enhance the predictability of stock market volatility, and what theoretical foundation supports this approach?
2. Can LSTM-predicted returns significantly improve the accuracy of volatility forecasts within a hybrid LSTM-GARCH framework, especially in light of the limitations faced by traditional models?
3. How does the hybrid model perform in responsiveness and adaptability to rapid market dynamics compared to standalone models, and what metrics can quantitatively assess this enhancement?

4. To what extent can advanced GARCH models—EGARCH and GJR–GARCH—boost forecasting accuracy in the hybrid framework? How are their contributions in both method and empirical verification to be measured and observed?

1.3 Research Contribution

In this paper, a new hybrid LSTM-GARCH model is proposed that skillfully combines the great ability of LSTM in deep learning and the rigorous ability of GARCH in volatility modeling. This model aims to make full use of the strengths of both sides: LSTM's strong ability to capture complex temporal dependence and GARCH's excellent ability to model temporal-varying volatility. The results are significant in both method and practice. This paper breaks the predictive limit of the existing methods , and also gives the most intuitive and practical forecast result to investors, risk managers and policy makers; this study is an important bridge between deep learning and economics model , it will definitely make great promotion to method of financial forecast and application of model [26, 8, 14].

The rest of the chapters are organized as follows: Chapter 2 offers an extensive review of the literature on volatility prediction, illustrating the path from the traditional econometric models until the recent advances in integrating machine learning. Chapter 3 breaks down the construction, theoretical justification, and practical implementation of the proposed hybrid LSTM-GARCH model. Chapter 4 then provides a comprehensive experimental evaluation of the model, comparing the model to the existing prediction models. Finally, Chapter 5 concludes the paper by recapping the key findings, noting limitations, and suggesting venues for future research for a more sophisticated practice of predicting volatility.

Chapter 2

Related Work

This chapter provides a detailed and systematic review of studies on stock market volatility forecasting. While this review comprises several major themes, namely on traditional volatility models, LSTM models in financial prediction, and hybrid models that merge econometric and deep learning models, each section includes a summary of seminal works, their methodology, findings, and limitations. The following sections describe our novel contributions.

2.1 Traditional Volatility Models

The study of the volatility of financial markets originated from the traditional models of econometrics. The foundation for this theme was laid by Engle with his work on Autoregressive Conditional Heteroskedasticity [16]. This work proposed a framework that can be used to model time-varying volatility. The dynamics opened a deeper revolution in financial econometrics. Another work was the continuation of the development of the foundation by [8] with the Generalized Autoregressive Conditional Heteroskedasticity model. This framework provided a methodology for explaining the volatility clustering, which is the phenomenon of financial markets.

Despite the profound contribution of models like GARCH and its subsequent variants, EGARCH introduced by [35] and TGARCH by [49], there has been a constant critique against these approaches. These models are founded on several assumptions, like that of stationary time series and the constraint in capturing the leverage effect [7] and asymmetric volatility. [38] reviewed the approaches and noted that, while relevant, many traditional approaches would not articulate the dynamics involved in the financial markets.

Recent advancements like [25] and [15] have also attempted to correct these models' limitations on volatility forecasting. However, [11] argues that the linearity constraint inherent in numerous econometric methods still needs to be fully addressed. The need for more accurate models for use in volatility forecasting has thus necessitated a search for alternative approaches beyond the traditional econometric methods. There is an apparent need for a more dynamic approach to address the non-linear and dynamic nature of volatility time series data [3].

Emerging machine learning approaches, such as Long Short-Term Memory (LSTM) networks, can be integrated with traditional approaches to enhance the predictive power of forecasting models. This section introduces the journey for volatility modeling, providing the basis for the novel hybrid LSTM-GARCH model introduced in the dissertation. The current research aims to add to the existing body of knowledge by critiquing the existing models to offer insight into the relevance and limitations of the existing models.

2.2 LSTM Models in Financial Prediction

Deep learning advancements have opened a new frontier in financial market forecasting, among which the Long Short-Term Memory networks [26] have played a pivotal role. Implemented as a special kind of Recurrent Neural Network, LSTMs are uniquely designed to allow learning of long-term dependencies [5], enabling modeling of financial time series data, which are characterized by complex patterns and non-linear relationships [40]. Previous research has already demonstrated the effectiveness of LSTMs in predicting stock prices and market trends. An early study by [21] and [24] provides evidence of LSTMs' ability to capture temporal sequences. More recent work by [19] compared the ability of LSTMs and traditional models to capture the complexity of sequential data and showed that LSTMs are superior in the context of financial applications.

Studies have found similar results that aimed to predict stock return using the LSTM network. Moreover, [31] explored how LSTMs can be utilized in stock market predictions and pointed out that they generally outperform conventional models. Assessing future market trends can be accurately done using LSTMs. Additional research has confirmed the reliability of LSTMs for stock return prediction [46].

Despite the promising results observed so far, using LSTM has several challenges. [36] writes that tuning LSTM's parameters could be complicated and significantly affect the model's performance. "The black-box" nature of LSTM models underlines the significant problem of interpretability. Nevertheless, research on explainable AI has been conducted, highlighting the efficacy of the existing approaches [33] characterized by the capacity to explain complex models, which may promote the broader dissemination of LSTMs usage in financial analysis.

The LSTM model is a powerful approach to improving financial forecasting and updating models with new information. While the challenges addressed remain, in the future, based on the prospects of deep learning development, the presence of the solutions will enhance its application.

2.3 Hybrid Approaches Combining Econometric Models with Deep Learning

The amalgamation of econometric and machine learning paradigms constitutes a novel frontier of financial forecasting. In addition to the apparent benefits of using two distinct knowledge domains, namely the econometric and the deep learning model, hybrid methodologies embrace the power of two different sets of methodologies. In other words, the combination of the econometric method pioneered by Bollerslev with the GARCH model, widely known for its effectiveness in forecasting volatilities [8] with an exemplary deep learning model, such as the LSTM network, which is a popular choice in most sequence learning endeavors [26], challenges financial forecasting to a whole new level.

Recent integrations of machine learning and econometric methods have shown improvements compared to stand-alone methodologies. For example, the review from presented a meta-analysis of the forecasting improvements achieved by hybrid methodologies. However, a meta-analysis of the accuracy metrics in forecasting or computational speed and performance has not been sufficiently analyzed, which is an important bridge that would lead us to understand the real-world convenience of these models.

Combined with the new deep learning methods or artificial neural network architectures, such as the Transformer, discovered by [47], the forecasting model reaches a sophisticated quality. Furthermore, intensively researched analysis with Transformer models is still a slab on the Earth of hybrid deep learning and econometric models. The novel addition of external data, from macroeconomic factors to geopolitical events, could extend hybrid forecasting performance even more. Although [32] and [31] successfully established integration into the classical econometric model, a unified hybrid framework combining these two on a large scale still needs to be included.

Analyzing this data increases the model's efficiency, trustability, and ability to be used in real-life scenarios. As there is still a wide gap in how the models forecast and how forecasts can be used, the emergence of the hybrid model would clarify that to the broader audience. Moreover, given the increasing possibility of the deployment of AI, advanced AI, especially deep learning formation, poses a threat as the advanced algorithms forecast the market in a "black box" method. In order to establish some credibility, a branch of deep learning called explainable AI offers a genuine walk into the upstream direction, as it opens the secrets of AI technology.

Therefore, while combining the two models marks progress, there is a greater bridge that has, as of yet, remained unspanned.

2.4 Research Gap and Contribution

Hybrid models in financial market forecasting in GARCH practice and LSTM networks have brought together two scarcely related areas of econometric models and machine learning. Volatility models, such as traditional GARCH [8], have been fundamental to describing market behavior. On the other side are distinctive Artificial Intelligence techniques, such as deep learning and LSTM networks [26], which can catch profound and demanding time behaviors that GARCH cannot spot. However, despite the apparent possible high degree of relationship between these areas, the academic literature needs to be more comprehensive in findings linking the fields and comprehensive studies combining those methodologies.

This study proposes a new hybrid LSTM-GARCH model developed to overcome the limitations of each of the original models. Due to the high capacity of LSTM to retain the memory of complex time behaviors and the optimal volatility modeling performance of GARCH, the model can be simultaneously more precise and more understandable. Similarly to the model descriptions provided in [46], the investigation of machine learning applications has generally been described without underlying methodological explanations.

The novelty of the proposed study is in the detailed methodological study of LSTM-GARCH fusion, not as a final instrument for the improvement of forecasting but as the enhancement of the model interpretability. Just as in [31] have advocated for the necessity of the union of machine learning with econometric methods, a study simultaneously increased the interpretability of machine learning models. The recent turbulent market conditions, as

[33] uncovered, boosted by Covid-19 dispersion, caused the necessity of model alerting to the market reactions.

Similarly, [28], synthesizing models strengthen the forecasting performance and then apply consequently from one to another. Also, [3] and [20] have repeatedly raised how more than econometric models are needed for such nonstationary modeling problems. Moreover, considering issues from model interpretability and ethics, there is a necessity for more noticeable shift to prompting the models to explain prediction.

In general, this study will address the need for hybrid LSTM-GARCH models in market forecasting. This need is confirmed not only by the recent market volatility brought to the extreme by Covid-19 [33] and [28] but also by the recent call for combining models to achieve better forecasting performance into financial practice such as [3] [20]. The subsequent chapters will describe the model development, implementation, and theoretical background, which essential, may revolutionize the market analysis field [32].

Chapter 3

Proposed Methodology

Accurately predicting market volatility is a complex task. Conventional econometric methodologies, such as GARCH or ARIMA, could be more effective in addressing this challenge. To overcome this issue, we propose using an integrated model which combines LSTM and GARCH. By integrating LSTM and GARCH [35] and [17], we can improve the accuracy and robustness of market volatility predictions. Our integrated approach represents a novel combination between LSTM, which learns long-term dependencies, and GARCH, which captures volatility clustering. We expect this approach to outperform other methodologies when market conditions become non-linear during financial crises and periods of financial stress [3].

In addition, a strict validation process, including cross-validation, regularization, and sensitivity analysis, is also presented. Consequently, the generality and robustness of the proposed composite model against different market conditions are intended to be guaranteed, thus reducing the overfitting possibility and improving the model's performance in heterogeneous market situations [6].

3.1 Problem Statement

Within the domain of financial markets, forecasting market volatility accurately holds significant importance for a successful risk management, portfolio optimization, and strategic investment decision-making. Traditional models demonstrate their effectiveness to some extent, however, are limited in tolling the non-linear, non-stationary, and dynamic behavior of financial time series. Given a time series of stock prices D, where $t = 1, 2, \ldots, T$, the primary goal is to advance the prediction of market volatility $\hat{\sigma}_t^2$ at time t with the aid of the information from the past with the stock price data $S_{t-1}, S_{t-2}, \ldots$. The study presents a novel approach called hybrid LSTM-GARCH that outperforms LSTM and GARCH model in terms of prediction accuracy by exploiting advantages of each model. The LSTM model captures the complex temporal dependencies in stock price, whereas GARCH model is competent in yield accurate volatility based on residuals produced by LSTM model prediction.

The novelty of our method lies in how we combine these models. The LSTM component tries to predict future stock prices and the errors—or residuals—are the differences between the actual returns and the predicted returns by the LSTM. Then, these residuals are used by the GARCH model to forecast the volatility. We formally articulate this joint methodology as the following optimization problem:

$$\min_{\Theta_{LSTM}, \Phi_{GARCH}} = \mathbb{E}\left[\left(\sigma_t^2 - \hat{\sigma}_t^2\left(Y_{test} - \hat{Y}, \Phi_{GARCH}\right)\right)^2\right]$$

where

- σ_t^2 represents the actual market volatility at time t,
- $Y_{test} - \hat{Y}$ signifies the residuals at time t, where Y_{test} represents the actual return at time t and $\hat{Y}$ is the return predicted by the LSTM,
- $\hat{\sigma}_t^2$ is the forecasted volatility at time t, obtained by applying $Y_{test} - \hat{Y}$ to the GARCH model with parameters Φ_{GARCH},
- Θ_{LSTM} and Φ_{GARCH} are parameters of the LSTM and GARCH models, respectively,
- The expectation $\mathbb{E}[\cdot]$ aims to minimize the squared difference between the actual volatility σ_t^2 and the predicted volatility $\hat{\sigma}_t^2$, which underscores the model's accuracy in forecasting.

This problem statement underscores our aim to amalgamate the sequential data processing capabilities of LSTM networks with the volatility estimation accuracy of GARCH models into a unified framework. By surmounting this optimization challenge, we aspire to significantly enhance the accuracy of volatility predictions beyond the capacity of individual model approaches, thus contributing a novel and effective tool for financial analysts and investors.

3.2 Proposed Hybrid Approach

In this study, a new hybrid model was proposed that effectively combines the strengths of GARCH models and LSTM networks to offer an accurate and superior tool that forecasts market movements. LSTM networks' impressive performance in learning and recognizing patterns within sequences one-up conventional provides an exceptional method of capturing the complex non-linear dynamics characteristic of financial time series, as demonstrated by [26]. With GARCH models having a good track record in estimating volatility, the hybrid model with improved features increases the predictive ability of the two underlying models alone. In sum, this framework promises a step up on the predictability and scope of less modern and more conventional methods used to predict volatility today.

Additionally, it provides a foundation for different applications to stay ahead of and alleviate various issues, such as the user-friendliness of hyperparameters, overfitting, and underfitting. As a result, the projective model structure, which is expected, will allow two different methods to do more than what they can do alone either of them [8], [17].

3.3 Basic Model Architecture and Theoretical Foundations

This section closely examines the architectural framework and the fundamental logic behind the blending of LSTM and GARCH. LSTM is the Long Short Term Memory algorithm, while GARCH is the Generalized Autoregressive Conditional Heteroscedasticity model. This stance has a firm argument that forecasts the stock market's volatility. It is a vital element in capital markets, especially the tracking mechanism in which it is an unstoppable force.

3.3.1 LSTM Model

The Long Short-Term Memory (LSTM) unit is a breakthrough in neural network architecture, which was created to overcome challenges in understanding long-term relationships, especially in sequence data. Imagined by Hochreiter and Schmidhuber in 1997, the LSTM building is intended to prevent vanishing gradient problem—the main challenge facing RNNs, causing the learning of far-reaching relationships to be unfeasible. The structure uses memory cells and a few control mechanisms to keep the data flowing and maintain the error flow as far as possible [26].

3.3.1.1 Gate Operations

Within an LSTM unit, the operations of the gates are predicated on the current input, x_t, and the previous hidden state, h_{t-1}. These gates include:

In an LSTM unit, the current input, x_t, and previous hidden state, h_{t-1} determine the gates' functions. These gates are as follows:

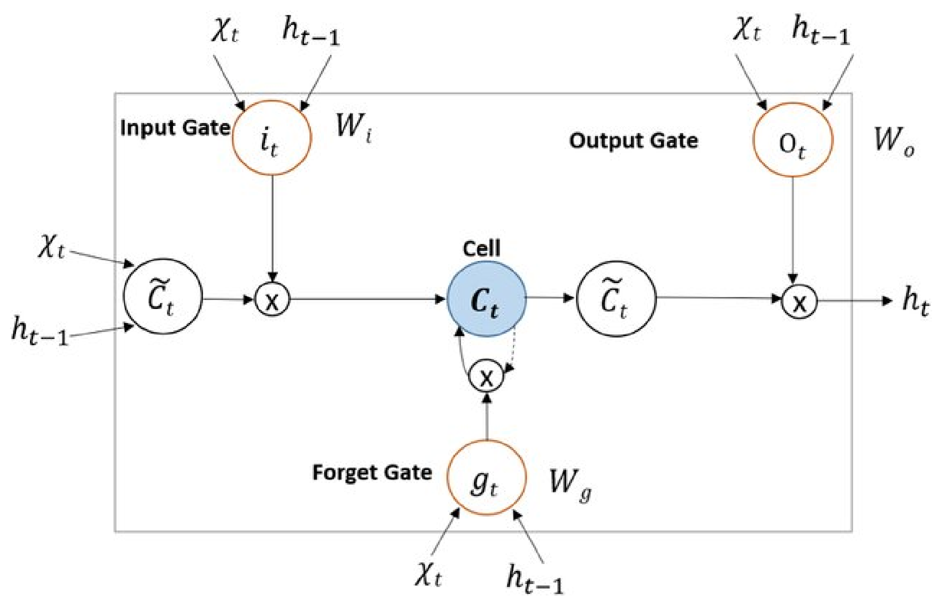

Figure 3.1: LSTM unit architecture, with the input, forget, and output gates, the cell state (C_t), and the flow of information.

- **Input Gate** (i_t): It allows explicitly crucial new information to enter the cell state. Mathematically, a sigmoid function is applied to measure the significance of the new information and binds it uniquely with a $\tanh$ function, thus generating a vector of new candidate values ($\tilde{C}_t$). The equations below explain this entire process:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i),$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C).$$

The product of i_t and $\tilde{C}_t$ determines the potential update to the cell state, allowing the network to add information deemed relevant for long-term predictions.

- **Forget Gate** (g_t): This gate evaluates which portions of the cell state C_{t-1} should be retained or discarded through the journey of sequences. It employs a sigmoid function to create a filter that retains only valuable information:

  $$g_t = \sigma(W_g \cdot [h_{t-1}, x_t] + b_g).$$

  In financial time series, this feature allows the model to filter out old information irrelevant to the current market environment, like historical stock prices that have lost their capacity to predict future patterns.

- **Output Gate** (o_t): The final gate in the LSTM uses the cell state after it has been updated by the forget and input gates to determine the next hidden state h_t, which contains the information to be forwarded to the output or the next LSTM unit. It is expressed by:

  $$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o),$$
  $$h_t = o_t * \tanh(C_t).$$

  The interaction between o_t and the $\tanh$ of the updated cell state facilitates the output of only the most pertinent information, reflecting the newly acquired knowledge about market movements and volatility patterns.

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t,$$

The modification of the cell state, C_t, is a combination of the forget gate g_t and input gate i_t. This arrangement allows the network to retain important past information while allowing for new insights to be included in the state of the cell. LSTMs have an advantage in estimating complex financial time series because of this detailed manner of managing sequences, where numerous market factors need to be balanced and represented in the predictive model.

3.3.1.2 Activation Functions

The functionality of the Long Short-Term Memory (LSTM) is highly dependent on a pair of activation functions:

- The goal of the sigmoid function σ is to act as a gate and limit the output to be between -1 and 1.

  $$\sigma(x) = \frac{1}{1 + e^{-x}}$$

- The hyperbolic tangent function ($\tanh$) can help the transmitted information into a range between -1 to 1.

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

3.3.2 Advanced LSTM Models

Advanced LSTM architectures have been crucial in handling harder tasks which involve modelling the sequence that are encountered in the forecasting of the financial market. These modifications in the standard LSTM are very skillful in the capturing of the complex dependencies and non-linearities which are present in the financial time series data.

- **Stacked LSTM**: The Stacked LSTM is composed of several LSTMs in series. This structure makes it easy to train hierarchical features, where lower-level structures are learned from sequences of time steps and higher-level ones are learned from sequences of lower-level structures. It is important in financial forecasting, where we want to examine both short-term volatility and long-term trends. The stacked LSTM architecture is summarized as follows:

$$h_t^{(l)} = \text{LSTM}^{(l)}(h_t^{(l-1)}, h_{t-1}^{(l)}),$$

  where ${h_t}^{(l)}$ is the hidden state at time t of layer l and ${h_{t+1}}^{(l-1)}$ is the output from the $l-1^{th}$ layer. This sequential organization of data is adequate for abstract neighborhoods within stock markers which may be too complex, but it is very time costly and requires constant alertness in order to not over fit [21]..

- **Bidirectional LSTM**: Bi-directional LSTMs process data in both directions, forward and backward, to incorporate the context in each and every step. Such an efficient processing architecture is particularly useful for stock price forecasting, since it takes into consideration the historical pattern of stock prices, as well as the future expectations. The hidden layer at each timestep of forward LSTM processes and memorizes the inputs to the right for each timestep, while the backward LSTM memorizes the inputs to the left for each timestep. The memory cells in BLSTM are connected to the opposite direction so they can take advantage of the future and past information in the current state of processing. By doing simple concatenation of the outputs from two networks, the concatenated vector holds both past and future information, allowing a better prediction of the target stock price [41].

$$\begin{aligned}
\overrightarrow{h_t} &= \overrightarrow{\text{LSTM}}(x_t, \overrightarrow{h_{t-1}}), \text{ forward pass,} \\
\overleftarrow{h_t} &= \overleftarrow{\text{LSTM}}(x_t, \overleftarrow{h_{t+1}}), \text{ backward pass,} \\
h_t &= [\overrightarrow{h_t}; \overleftarrow{h_t}], \text{ concatenation of forward and backward states.}
\end{aligned}$$

- **Gated Recurrent Unit**: The Gated Recurrent Unit simplifies the structure of LSTMs by combining forget and input gates into a unified update gate. This reduced parameterization of GRUs translates into better speed performance and more applicability to stringent resource requirements.

$$\begin{aligned}
z_t &= \sigma(W_z \cdot [h_{t-1}, x_t] + b_z), \text{ update gate,} \\
r_t &= \sigma(W_r \cdot [h_{t-1}, x_t] + b_r), \text{ reset gate,} \\
\tilde{h}_t &= \tanh(W \cdot [r_t * h_{t-1}, x_t] + b_h), \text{ candidate state,} \\
h_t &= (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t, \text{ final state update.}
\end{aligned}$$

- **LSTM with Attention Mechanism**: Introducing attention mechanism into the LSTM architecture enables the model to selectively focus on the important input sequence with high relevance to the current prediction. The dynamic focus mimics the processes of analysis that traders undergo in grading certain market signals above others.

$$\begin{aligned}
e_{t,i} &= v_a^\top \tanh(W_a h_{t-1} + U_a h_i), \text{ attention scores,} \\
\alpha_{t,i} &= \text{softmax}(e_{t,i}), \text{ attention weights,} \\
c_t &= \sum_i \alpha_{t,i} h_i, \text{ context vector,} \\
h_t &= \text{LSTM}(c_t, h_{t-1}), \text{ hidden state update with context.}
\end{aligned}$$

  This attention mechanism helps the model to look at and give importance to the information that is key in forecasting and is very important in the financial markets where some specific events may have significant effect.

The integration of these advanced LSTM architectures in hybrid forecasting models has the potential to advance the prediction of complex financial patterns, providing a very advanced tool in the analysis of market volatility.

3.3.3 Justifying LSTM Model Integration into Hybrid Financial Models

Embedding a LSTM network into the hybrid financial model is built upon its distinctive ability to capture complicated sequential data allowing us to notionally capture temporal dependencies inherent to the financial market.

- **Empirical Validation of LSTM Effectiveness:** It is empirically proven that LSTM is superior in financial forecasting, and it has advanced more than traditional methods. LSTM has surpassed traditional time series models and machine learning models in stock price forecast, mainly because LSTM is able to easily capture and model long-term dependencies, which can't be realized by any other models or those models are not efficient enough through which the ARIMA and GARCH models are used to simulating the stock volatilities [19]. Empirical tests have shown that the accuracy of volatility forecasting for the stock market using LSTM still far exceeds the ARIMA and GARCH models showing that LSTM is strong powerful and highly versatile in various types of financial forecasting problems [36].

- **LSTMs' Compatibility with Financial Data:** Long-Short Term Memory Networks, or LSTMs, are perfectly suited to forecasting financial time series data. Financial markets are incredibly complex due to the interaction of macroeconomic indicators, corporate actions, and geopolitical events, all of which make for a challenging modeling task. This complexity has pushed researchers to employ non-linear architectures to better capture the market's movement and volatility. LSTMs, specifically, have demonstrated great success here [26, 39].

- **Addressing constraints for Financial Applications:** In the realm of financial forecasting, LSTMs have typically encountered difficulties with such problems as long computation times and overfitting. Fortunately, the rise of regularization techniques as well as the development of variant forms of LSTM, including Convolutional and Attention-based LSTMs, had considerably mitigated these problems and led to the creation of more quick and adaptable financial forecasting models [44, 47].

- **LSTMs for Improving Hybrid Forecasting Models:** The combination of LSTMs with volatility forecasting models like GARCH is a major step forward in the prediction of financial time series. It combines the dynamic temporal pattern recognition of LSTMs with the volatility modeling abilities of econometric models to create an all-round analytical tool, which is more comprehensive than individual models.

Fundamentally, empirical validation combined with the innate compatibility of LSTMs with financial data firmly position their incorporation into hybrid models as a methodological progression in the financial forecasting arena. The advancements that should emerge from the continuous improvement in the technology around LSTMs and its fusion with hybrid models will undoubtedly alter predictive accuracy and the knowledge insights for market analysis.

3.3.4 GARCH Model

Modeling volatility is important for a variety of fields to understand the temporal changes occurring within a series of data. The Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model, an extension of the Autoregressive Conditional Heteroskedasticity (ARCH) model introduced by Engle [16], is a fundamental step forward in the realm of financial econometrics. It captures the features of clustering in conjunction with the omnipresent attribute of time-varying volatility in financial time series data.

3.3.4.1 The Basic GARCH(p,q) Model

The GARCH(p,q) model is an extension of the ARCH restricted model that adds terms for last forms of both the lagging second derivatives and the lagging conditional variances, which enable the model to capture long-term volatility relying. The conditional variance in GARCH(p, q) is given by:

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2,$$

where σ_t^2 is the conditional variance at time t, ϵ_t is the innovation term, ω, α_i, and β_j are parameters to be estimated by the data

3.3.5 Advanced GARCH Models

Observations based on experience have often shown an unbalanced response in the occurrence of variation in relation to news of a negative or positive nature, one which the fundamental GARCH model does not describe accurately. This recognition has motivated the creation of extensions for the GARCH model, generated to contribute information in observing various factors in the marketplace along with variation.

3.3.5.1 EGARCH Model

The Exponential GARCH (EGARCH) model, proposed by [35], is widely known for its ability in modelling the asymmetric effects and leverage phenomena in distinction to the classical GARCH models. The EGARCH model defines the logarithm of conditional variance as a linear function of past shocks.

$$\log(\sigma_t^2) = \omega + \sum_{i=1}^{p} \alpha_i g(Z_{t-i}) + \sum_{j=1}^{q} \beta_j \log(\sigma_{t-j}^2),$$

where $g(Z_{t-i})$ captures the effect of the sign and magnitude of past innovations (Z_{t-i}). This feature allow the model to work flexibly and the model can show how positive and negative shocks to the system can have independent impacts on the conditional volatility. Thereby, the leverage effect could be faced fully since negative market news tends to has higher leverage than the positive news on volatility.

3.3.5.2 GJR-GARCH Model

Introduced by [22], the GJR-GARCH model incorporates an indicator function in order to distinguish between the impact of negative and positive shocks on volatility using the following equation:

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 + \sum_{i=1}^{p} \gamma_i \epsilon_{t-i}^2 I_{t-i},$$

where I_{t-i} takes a value of 1 if $\epsilon_{t-i} > 0$, otherwise it takes a value of 0 for condition $\epsilon_{t-i} < 0$. This model captures the notion that bad news (i.e., negative shocks) increases future volatility more compared to good news (i.e., positive shocks) of the same magnitude, a commonly observed phenomenon in financial markets.

3.3.5.3 TGARCH Model

The Threshold GARCH (TGARCH) model extends the idea of asymmetric volatility by directly modeling the different impacts of positive and negative shocks.

$$\sigma_t^2 = \omega + \sum_{i=1}^{p}(\alpha_i + \gamma_i I_{t-i})\epsilon_{t-i}^2 + \sum_{j=1}^{q}\beta_j\sigma_{t-j}^2,$$

In this model, γ_i adds to the impact of negative shock (I_{t-i} indicating a negative shock) explicitly modelling that different type shocks affect volatility asymmetrically. TGARCH is useful in particularly resetting the "leverage effect", where news is more sensitive to bad news than to good news of equivalent magnitude.

3.3.5.4 FIGARCH Model

The FIGARCH model, developed by Baillie (1996), incorporates long memory into the volatility process:

$$\sigma_t^2 = \omega + \sum_{i=1}^{\infty}\phi_i\epsilon_{t-i}^2,$$

Rather than having exponential decay as in standard GARCH models, this model features a hyperbolic path for the waning of shocks, whereas FIGARCH model is reliable in capturing the long-lasting impact of volatility shocks, which mirrors the perception that the effects of market events may be extensive.

Each of these models has the ultimate objective of fully capturing the dynamics of the volatility phenomenon, going beyond a simple representation of the stylized facts underlying financial time series. These models are, as a result, very appealing in financial econometrics, since they offer new insights into the nature of financial market volatility.

3.3.6 Justification of Integrating GARCH Model into Hybrid Financial Models

The efficacy of the GARCH models in capturing complex volatility patterns has been proved in different financial markets. This section is devoted to both empirical validations and theoretical justification for the inclusion of GARCH models in Hybrid Predictive analytics.

- **Empirical Efficacy of GARCH Models:** Studies have shown that scaled variance models perform much better in forecasting volatility than simpler models, which tell us a lot about their performance. In particular, they do a good job in capturing the volatility clusters that are common to most financial time series, thus allowing for a more detailed and precise volatility forecasts. The most important papers of this kind are [16] and [8], which also show that the models are highly flexible and can be used in a variety of market conditions. This means that GARCH models can be used not only for scientific purposes in the field of financial econometrics but also for a variety of applications such as risk management and the pricing of derivatives.

- **Alignment with Financial Market Dynamics:** The architecture of GARCH models is perfectly aligned to the dynamics observed in financial markets, thanks to its capacity to take into account time-varying volatility. This characteristic allows these models to capture important financial data features such as leptokurtosis and volatility clustering, which are absolutely necessary in order to forecast financial figures with precision and accurately evaluate financial risks. This strong alignment with financial data dynamics represents a valid ground to allow their integration in hybrid models, in order to improve long-term volatility predictions and adapt these forecasts to the market sentiment [7, 35].

- **Extending GARCH Models for Enhanced Flexibility:** GARCH models have evolved from their original form into a family of models that correct some of the limitations of the original GARCH model and capture the asymmetry of financial response to shocks by using models such as EGARCH, GJR-GARCH, and FIGARCH. These extensions have increased the span of GARCH models in terms of the types of data to which they have been applied, as well as the accuracy with which they can be estimated for financial analysis today [35, 22, 2].

- **Synergy in Hybrid Forecasting Models:** Integrating GARCH models into advanced machine learning techniques, particularly LSTM neural networks, generates a significant advancement in the analysis of financial time series. This collaborative scheme merges deep learning models' pattern recognition aptitude with the GARCH models' precision in forecasting volatility. Ultimately, the emergence of these hybrid forms of models endows the composite models with a predictive advantage even more pronounced than exclusively deep learning-based models. The models contribute the deep analyses of market behaviors, greatly improving the efficiency of financial decision-making processes [17].

Given empirical supportive and theoretical backings, the incorporation of GARCH models into hybrid financial modeling frameworks becomes persuasive. Along with many desirable properties of GARCH models, hybrid models are much more accurate predictors, they recapture the nuanced associations between financial markets and potentially incorporating methods into the body of financial forecasting methodologies. Although, the performances of many existed hybrid forms of models are not ideal, a noticeable advancement for methodological and applied forecasting is budding with the attractively developed theoretical framework.

3.4 Proposed Hybrid LSTM-GARCH Model

Financial markets have been known for its complexity, characterized by substantial volatility and nonlinear dynamics, which impose great challenges to traditional forecasting techniques. In response, we introduce a Hybrid LSTM-GARCH Model in this paper, which is an innovative integration of predictive power of Long Short-Term Memory (LSTM) networks and volatility modeling strength of Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models. This synergy aims to outperform previous forecasting methods, enhance accuracy and provide a deep insight into market behavior [26, 8].

The origin of this hybrid model is inspired by the aim to create a powerful prediction mechanism that combines LSTM's competence in handling sequential and non-linear patterns with GARCH's expertise in modeling time-varying market volatility. This study is expected to enhance the interpretability and dependability of financial predictions, which will make it an indispensable tool for investment and strategic decision making [45, 35].

3.4.1 Theoretical Background

The Hybrid LSTM-GARCH Model was developed because current financial forecasting methods often struggle when dealing with the random nature of financial time series. By combining LSTM's ability to pick out intricate patterns in data, and GARCH's methodical treatment of estimating volatility, the model hopes to present a deeper and more accurate depiction of market behavior [16].

This interdisciplinary project aligns itself with the Efficient Market Hypothesis. This hypothesis suggests that market prices correlate with all available information. However, it has an aspect that relates to the unpredictability of market processes and the effects of ”volatility clustering” [18]. This project also introduces consideration of another academic trend, namely behavioral economics, into modeling markets. By doing so, it introduces a mechanism for modeling the determinants of market participants' collective decisions, further diversifying the tools available for modeling financial markets [42].

The following sections provide description of the Hybrid LSTM-GARCH Model with details of its architecture and the mathematical principles behind it. It also outlines the strategies to implement the model and the unexplored analytical techniques employed in the model. This academic endeavor strives to demonstrate the reliability of the model through empirical assessment and theoretical close scrutiny, with ambition to put it as a ground-breaking tool in the financial forecasting [10, 17].

3.4.2 Hybrid Model Architecture

The Hybrid Model LSTM-GARCH is an innovative approach to financial forecasting that combines deep learning's power to predict with econometric models. This part of the essay explains the architecture of the Model, describing the mathematical schema and how Long Short-Term Memory (LSTM) networks are integrated with Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models to harness both LSTM's ability to unravel complex temporal patterns present in financial datasets and GARCH's capacity to produce better estimates of Market Volatility.

3.4.2.1 Mathematical Framework

From a systematic point of view, the Hybrid LSTM-GARCH Model can be raised as a continuous process. First, the time series data interpreted by the LSTM part rises to predict the future market trend. Secondly, the residuals from the predicted result that is the unexplained volatility will further calculated by GARCH model so as to predict the market volatility. The overall mathematical model is summarized as follows:

$$\hat{Y}_t = LSTM(X_t; \Theta_{LSTM}),$$

when X_t refers to the input data t, $\hat{Y}_t$ refers to the predicted result and Θ_{LSTM} represents parameters of the LSTM model.

Post LSTM's predictions, the model uses the residuals to refine volatility forecasts:

$$r_t = Y_t - \hat{Y}_t,$$

$$\sigma_t^2 = GARCH(r_t; \Phi_{GARCH}),$$

where r_t represents residuals, σ_t^2 is the forecasted variance or volatility, and Φ_{GARCH} delineates GARCH model parameters. This stepwise approach does a deep mining of market data, caapturing market movement together with internal volatility under one analysis framework.

3.4.2.2 Integration Workflow

The flowchart image of the Hybrid Model 3.2 demonstrates how the LSTM and GARCH work as a cohesive part of the Hybrid Model. This image demonstrates how the data flows through from data ingestion to volatility forecasting. It does however show the flow of each sequential step that the data goes through starting with the pattern identification in the LSTM moving to the volatility prediction by the GARCH to the final market analysis.

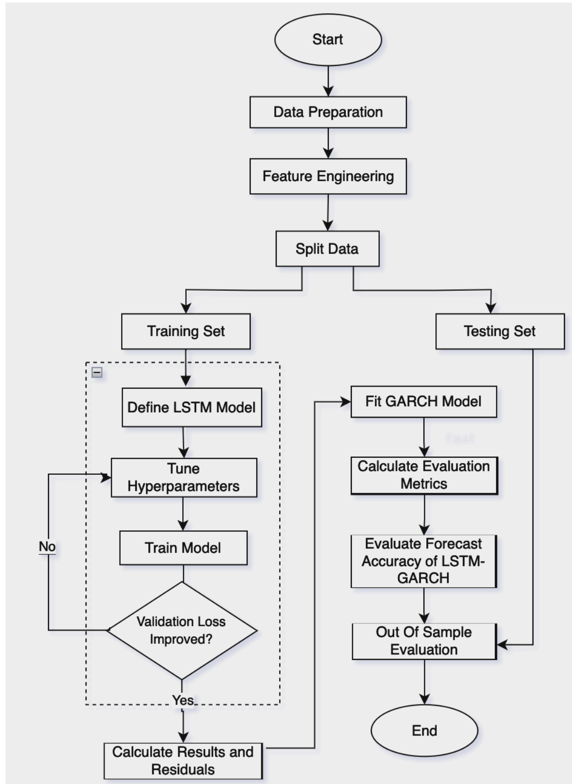

Figure 3.2: Hybrid LSTM-GARCH Model Flowchart

This integration showcases the innovative blend of machine learning models such as LSTM and econometric models such as GARCH giving a boost to the accuracy when it comes to market trend forecasting and also market volatility prediction and this engineered design contains full scoop of financial market dynamics. This feeling based architecture paves way for the future of financial forecasting methodologies.

3.4.3 Algorithm

This section of the report now highlights the computational process of the Hybrid LSTM-GARCH Model that smartly combines the Long Short-Term Memory (LSTM) networks with the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models. The following algorithm provides a step-by-step process beginning from initial data preprocessing to generating integrated market forecasts and accordingly adds a structure that allows users for successful replication and application of the discussed and unique, advanced methodological framework.

This is a very effective framework to design the Hybrid LSTM-GARCH Model and this method helps the financial analyst to understand the market movement and volatility very clearly. Using this hybrid model helps to improve the forecasting accuracy and help in financial market forecast and strategic Decision making significantly by integrating the machine learning and Econometric model reveal.

Algorithm 1 LSTM-GARCH Hybrid Model for Enhanced Volatility Forecast

Require: Historical Data D
Ensure: Volatility Forecasts F, Performance Metrics (MAE, $RMSE$, Accuracy)
procedure PREPROCESSDATA(D)
    Normalize D to D_{norm}.
    $X, Y \leftarrow$ CreateSequences(D_{norm})
    $X_{train}, X_{test}, Y_{train}, Y_{test} \leftarrow$ SplitTrainTest(X, Y)
end procedure
procedure TRAINLSTMMODEL(X_{train}, Y_{train})
    Initialize LSTM model with parameters Θ_{LSTM}
    Train Θ_{LSTM} on $(X_{train}$,$Y_{train})$
    $\hat{Y} \leftarrow$ Predict on X_{test}
    $Residuals \leftarrow Y_{test} - \hat{Y}_{LSTM}$
end procedure
procedure FITGARCHMODEL($Residuals$)
    Initialize GARCH model with parameters Φ_{GARCH}
    Fit Φ_{GARCH} to $Residuals$
    Predict volatility σ_t^2
end procedure
procedure EVALUATELSTMGARCHMODEL(Y_{test}, $\hat{Y}$, σ_t^2)
    Mean Absolute Error $MAE \leftarrow (Ytest, \hat{Y}_{LSTM})$
    Mean Squared Error $RMSE \leftarrow$ (Y_{test} and $\hat{Y}$)
    Calculate accuracy of σ_t^2.
end procedure

3.4.4 Model Implementation and Validation

This section introduces the implementation of the proposed Hybrid LSTM-GARCH model in Section 3.4.2 and validation of the results carried out. Here we describes about various configuration steps that needs to be taken to get accurate and efficient results in financial forecast which includes data pre-processing, model optimization and model validation.

3.4.4.1 Data Preprocessing

The preprocessing stage converts raw financial data to a form ready to apply analysis, and it involves purifying the data so that the dynamics of the market can extended, continuously, and sufficiently.

- **Feature Engineering:** As a next step, various financial indicators and data features are included,to enhance the input and predictability of the model. Inclusion of these features is root on multiple scales. The Relative Strength Index (RSI) will be the feature foundation - measuring price change magnitude to determine price dynamic peaks and troughs.

MACD is another very popular price dimension indicator which is derived by subtracting the 26-day exponential moving average (EMA) from the 12-day EMA, and then uses 9-day EMA of the MACD line. Other price indicators we used were SMA, EMA, Bollinger Band (2). The main purpose of SMA and EMA is to smooth the price moves and can identify the trends. Bollinger Band represents long-term price trend information and identifies market volatility.

In additions to that, the log returns and date based features are applied to this model. Log returns is a frequently used way in finance which helps normalize the price differences of assets based on a certain period. It applies typically in equity analysis for analyzing the equity fluctuation in a certain time length. And for date features, it can provide the seasonal or temporal patterns to be investigated. In a summary, our final feature selection has taken comprehensive aspects of price movements into account and use the historic conventional data to refine it. The aim of our study is to dig the several different dimensions of price behavior and provide more precise predictions [34].

- **Normalization:** Data is scaled using Min-Max scaling or Z-score normalization to ensure uniformity, which is crucial for the LSTM model to perform well. This step is done to maintain the temporal integrity necessary when analyzing financial time series.
- **Outlier Handling:** When analyzing financial data, special care should be given to outliers. Rather than being automatically removed, outliers are instead examined in order to find how much they say about market volatility. So, capping outliers rather than deleting them, preserves their informative content, but reducing their distortional effect on the model [29]. [29].

3.4.4.2 Sequence Creation and Data Partitioning

Converting preprocessed financial data into a structured, sequential format is fundamental to unlocking the power of the LSTM network to recognize temporal patterns in market behavior. This section details my sequence creation and examples of its output and a discussion of the methods used to conduct experiments with temporal aspects in the context of LSTM:

- **Sequence Length:** The sequence Length is an important hyperparameter which determines to large extent the Long-Short Term Memory (LSTM) model's ability to capture the past dependencies. A sequence length of 10 is chosen such that it provides the model to be given a window of time such that the model is able to adequately capture any reasonable past relevant market trends and without over-simplifying the model. This choice is supported by empirical observations and theoretical guidelines, because a too short sequence may fail to capture long-term dependencies and a too long sequence introduces noise and is computationally inefficient [23, 12].

- **Sequence Creation:** After sequence length is determined, sequences are created from the normalized data. Each sequence will be comprised of a fixed-size window of some number of consecutive datapoints as input features and a single datapoint as target output. This suits the temporal progression of market data which is a good process for LSTM networks to recognize and learn predictive patterns from. The sequences will also be enhanced with the presence of different financial indicators to give a multidimensional outlook into market behavior [12]. [12].

  In order to evaluate the model's ability to predict and generalize, the dataset is divided into two separate sets. The training set is 90% of the total sequences and the testing is 10% of the sequences. This is a common practice in the deep learning domain and essential to evaluate the model's ability to generalize and predict, while ensuring that there is a large enough volume of data to train on. That way the model is properly tested on unseen data and its effectiveness will not be affected and will be relevant on unseen trends and volatility on the financial markets [23].

The attention to detail given to sequencing and partitioning the data further solidifies the procedural accuracy of preparing data for the Hybrid LSTM-GARCH Model. This modification was done on a structural level to catch critical timing relationship to give more accurate diagnostics and predictions of the behaviour of the financial markets.

3.4.4.3 Model Optimization and Hyperparameter Tuning

The Optimization of the Hybrid LSTM-GARCH Model is a crucial process, which improves the forecast accuracy of the model. Its primary task is to calibrate both LSTM and GARCH models. This involves trough empirical experimentations and advanced optimization techniques.

- **LSTM and BiLSTM Hyperparameter Optimization:** The power of prediction of the model LSTM, with augmented layers of Bidirectional, lies in the configuration of the hyperparameters such as the number of layers, the number of units per layer, the drop-off rates and the learning rate. Introducing BiLSTM layers allows the model to capture the time dependence in both the forward and backward directions, and provides more comprehensive analysis of time series data. Bayesian Optimization can be used to systematically explore and fine-tune these hyperparameters to achieve the best configuration that maximizes the accuracy of predictions. The configuration of the LSTM model with optimal effect is determined by checking the least test loss, which implies strong performance on unseen data [43, 23].

- **Regularization Techniques:** In order to overcome overfitting problem, which is particularly important in complex models having BiLSTM layers, we use dropout and L1/L2 regularization inside LSTM architecture. Those techniques try to regulate the model complexity via penalizing large weights and randomly cutting the connections during the training process, thus make the model have a better generalization capability towards different market scenarios [44, 37].

- **GARCH Model Variants and Parameter Exploration:** It is essential to experiment with various GARCH model configurations like EGARCH and GJR-GARCH. The objective of this experimentation is to find out which model variant and which set of parameter values yields the most accurate depiction of the volatility dynamics of the market. The best GARCH model is determined by looking at the Akaike Information Criterion (AIC) score; this measure plays-off the goodness of fit of the model against its complexity and is looking for models that succeed in being just as or more accurate then their rivals per units of the parameters that they use for the purpose of achieving this accuracy. [8, 35, 1].

- **Distribution Testing:** The GARCH model is evaluated under different probability distributions. The normal distribution, Student's t distribution and Generalized Error Distribution methods are evaluated. This step realizes the accuracy of leptokurtosis and skewness of financial market returns to forecast the volatility reasonably and robustly. The performance of GARCH model is also evaluated in different probability distribution alternatively from it denotes the leptokurtosis and skewness features of financial market returns on the volatility forecast which is not exact realization in the prospective [9].

In order to enhance the predictive power of the LSTM network and model the financial market volatility accurately by considering the probability distributions, we adopt a very thorough optimization and tuning process. Through the process of exploring and combining advanced GARCH models and probability distributions systematically and evaluating them based on performance metrics, the Hybrid LSTM-GARCH Model can be viewed as a very advanced forecasting tool that could handle the boats of the financial market in a very precise and robust way.

3.4.4.4 Model Validation

Rigorous validation and testing procedures are critical for assuring the reliability and robustness of the Hybrid LSTM-GARCH Model. In this phase, predictive performance of the model is evaluated with unseen data, which fosters confidence in its generalization and efficacy for real-world applications.

- **Cross-Validation:** Cross-validation, such as k-fold cross-validation, is indispensable in model assessment, ensuring a robust estimation of its performance by employing different subsets of data for training and validation. It is instrumental to combat overfitting and verify its stability and reliability over different market conditions [30, 4].

- **Performance Metrics:** By quantifying the forecasting accuracy and its risk minimizing performance through performance measurements, the model's prediction capability regarding the market's trend and volatility can be evaluated.

  - **Mean Absolute Error (MAE):** The MAE is a measure of the average error magnitude between the contemplated and original values, giving a clear peek into the average prediction errors we should expect. It is defined as:

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

where y_i denotes the actual values, $\hat{y}_i$ represents the predicted values, and n is the total number of observations [48].

– **Root Mean Squared Error (RMSE):** RMSE is the square root of the mean of squares of all actual values minus predicted values. It gives more weightage to large errors because it takes square root on SUM of squares. The RMSE is calculated as:

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

This measure is extremely sensitive to anomalies, making it an important metric in models where large errors are especially undesirable [13]. [13].

– **Variance (Var):** The variance of the forecast errors gives us an idea about the spread or dispersion of the errors around their typical value – that is, how consistently the forecaster is standing or falling, over the dataset. It is calculated as:

$$Var = \frac{\sum_{i=1}^{n}(\hat{y}_i - \overline{\hat{y}})^2}{n-1},$$

where $\overline{\hat{y}}$ is the average of the predicted values. Lower variance indicates that the prediction errors are clustered closely around the mean, which means that the predictions are consistently accurate.

3.4.4.5 Out-of-Sample Testing and Backtesting

The comprehensive assessment of the Hybrid Long Short-Term Memory-Generalized Autoregressive Conditional Heteroskedasticity Model extends to out-of-sample testing and backtesting, which are important parts to reproduce the performance of the model in the real world.

- **Out-of-Sample Testing:** After our machine learning models have been trained and tuned with historical data, the models are used to predict the data, the machines have never seen. This serves as a practical measure of how accurate the model is as well as serves as a litmus test for how well the model will perform in the real world for unseen future market data.

- **Backtesting:** Backtesting helps to evaluate the trading modelś rule lessons previous by historical prices. Result to verify a core characteristic of my system will be discussed later. It's really important because it shows if the model is robust, if it's making serious money no matter if the market is in bullish or bearish condition and what's good timing practicability on trading signals across the different markets. The back testing result is also used for model improvement and whether the model is viable as a financial market application.

In summary, the research to develop and explore the Hybrid LSTM-GARCH Model for volatility forecasting in financial time series is an important advancement in financial forecasting. It improves the opportunity of merging the advanced machine learning method with the econometric model to tackle the challenging issue of volatility forecasting in the financial market. In detail, we attentively appreciate the strength of the Long Short-Term Memory (LSTM) Neural Network for learning long-term dependency and the specialty of the Generalised AutoRegressive Conditional Heteroskedasticity (GARCH) model for volatility modelling. We put these two powerful forecasting methods together to create a professional approach better than the traditional method and all the previous hybrid models.

The Hybrid LSTM-GARCH Model offers an open perspective of market dynamics by developing a robust framework of mathematics and methodology to synchronize market information and inherent markets' volatility. At the same time, it conducted several extensive validation processes with also cross validation, performance metric evaluation, out of sample testing and backtesting such process guarantees that the model presents acceptable performance and its generalization ability over different market contexts.

Additionally, the inclusion of BiLSTM layers in the network structure is a pioneering step in grasping the dual time dependencies in financial time series further enhancing the forecasting precision and the analytic profundity of the model. With this addition, combined with our model's parameters' strategic optimization and tuning, the versatile Hybrid LSTM-GARCH Model is enabled to navigate and tackle the challenges of volatile financial markets.

Chapter 4

Experiments and Results

This chapter illustrates the experimental verification for our proposed Hybrid LSTM-GARCH Model, attempting every effort to examine its performance from a structured process. We first elaborate on the datasets, and then expound the experimental apparatus, including the model training, parameter tuning and testing protocols. Furthermore, a meticulous analysis of the obtained results is carried out, appraised side by side with the commonly used benchmarks, to demonstrate the effectiveness of our model and its potential impact in the financial market. In this chapter, we will not only measure the power of the predictiveness of the model, but also inspect the practical value of the model in different market circumstances.

4.1 Datasets Description

In order to assess the predictive accuracy of the Hybrid LSTM-GARCH Model, we utilized historical stock price data from different economic cycles obtained from Yahoo Finance using the Python package known as yfinance. In this article, all the datasets are presented which are reflecting different market behaviors and economic conditions which will give us more deep insight in our analysis. Data Set II: Recovery 2009-2017

4.1.1 The Recovery From Financial Crisis 2008: 2009-2017

Focusing on the period following the 2008 financial crisis, this This data set will allow us to examine how the model performed during market recovery and subsequent growth following the 2008 Financial Crisis. This data set covers the period following the 2008 Financial Crisis.

- **Period Analyzed:** January 1, 2009, to January 22, 2017.
- **Number of Trading Days:** 2,027.
- **Market Coverage:** All companies listed on the S&P 500 index during this period.
- **Features:** Daily Open, High, Low, Close, Adjusted Close, Volume, and Ticker symbols.

4.1.2 Benchmarking Comparative Analysis: 2010-2019

This dataset also allows for a direct performance comparison with the introduction of a benchmark study in the same asset class through a decade marked by various economic events.

- **Period Analyzed:** January 1, 2010, to September 1, 2019.

- **Number of Trading Days:** 2,432.
- **Selected Companies:** Thirty S&P 500 companies, carefully chosen to match the dataset used in the benchmark study.
- **Features:** Daily Open, High, Low, Close, Adjusted Close, Volume, and Ticker symbols.

4.1.3 Navigating The COVID-19 Market Volatility: 2018-2022

This data set spans recent market turmoil caused by the COVID-19 pandemic to test the models ability to detect and adapt to large, sudden market shifts.

- **Period Analyzed:** January 2, 2018, to December 22, 2022.
- **Number of Trading Days:** 1,223.
- **Market Coverage:** All companies listed on the S&P 500 index during the period.
- **Features:** Daily Open, High, Low, Close, Adjusted Close, Volume, and Ticker symbols.

4.1.4 Significance and Justification of Dataset Selection

The datasets chosen for this study provide a varied representation of market conditions, with the goal of gaining a wide-range assessment of the predictive power of the Hybrid LSTM-GARCH Model over several economic cycles. From the 2008 financial crisis recovery to the COVID-19 induced market volatility, the datasets chosen for analysis offer a multi-faceted environment with which to conduct the Hybrid LSTM-GARCH Model's predictive power. This allows us to test the Hybrid LSTM-GARCH Model's temporal stability across time.

To achieve this, a methodological approach that scrutinised the model stringently and in particular, what the development has brought in terms of advancing existing forecasting methodologies. The inclusion of a benchmark comparative analysis further augmented the work, with a finer scrutiny of the evolution and improvements of the model. Furthermore, by model's exposure to the unprecedentened market fluctuations during the COVID-19 period, also demonstrated that the model's fit into realworld forecasting scenarios where the market can experience rapid and substantial changes.

The selection of these datasets and justification for the use thereof, in a structured manner, is a clear indication of the detailed and meticulous evaluation of the Hybrid LSTM-GARCH Model, representing the model's strengths as well as areas thereof which can still be improved in future, advancing financial forecasting methodologies.

Ticker Symbol	Company Name
AMG	Affiliated Managers Group Inc
AMZN	Amazon.com Inc
AZO	AutoZone Inc
BKNG	Booking Holdings Inc
BLK	BlackRock Inc
DUK	Duke Energy Corporation
ED	Consolidated Edison, Inc.
F	Ford Motor Company
FCX	Freeport-McMoRan Inc
GE	General Electric Company
GOOG	Alphabet Inc
GPS	Gap Inc
HBAN	Huntington Bancshares Incorporated
ICE	Intercontinental Exchange Inc
LDOS	Leidos Holdings Inc
LMT	Lockheed Martin Corporation
M	Macy's Inc
MTD	Mettler-Toledo International Inc
MYL	Mylan N.V.
NKE	NIKE, Inc.
ORLY	O'Reilly Automotive Inc
PBCT	People's United Financial Inc
PLD	Prologis Inc
RF	Regions Financial Corporation
RJF	Raymond James Financial Inc
SHW	The Sherwin-Williams Company
TDG	TransDigm Group Incorporated
UAA	Under Armour Inc
VFC	VF Corporation
WEC	WEC Energy Group Inc

Table 4.1: List of S&P 500 Ticker Symbols Selected for the Study (2010-2019)

4.2 Software and Libraries

This entire study was implemented on Google Colab, a cloud-based platform that provides a Python notebook environment equipped with powerful computation resources. Execution of this research was possible courtesy of Google Colab by granting the opportunity to have advanced computing capabilities required for training deep learning models and performing considerable data analysis. Specifically, having access to high-performance GPUs and CPUs on Colab was very effective for running training of Long Short-Term Memory (LSTM) networks which require computational-intensive jobs and estimation of Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models.

Python was used in this study, drawing on its extensive data science and machine learning libraries. In the Google Colab environment, Python 3.10.12 was the version chosen. This was done because said version of the software could work with all the tools and libraries necessary.

Essential libraries used in this research were pandas, NumPy, and yfinance, which were used for acquiring and pre-processing financial data. For construction and training of LSTM networks, TensorFlow and Keras were utilized; other experiments also employed bi-LSTMs and regularization methods including dropout and L2 regularization to control overfitting issues. The arch library was used to specify and fit different GARCH model configurations.

Requirements for data to be normalized were catered through StandardScaler while partitioning of data was achieved through train_test_split. Train set was stepped through training and prediction and upon model factoring an evaluation has been accomplished to determine errors such as mean absolute (MAE) and mean squared (MSE) from which root mean (RMSE) was derived from sklearn.metrics. To gain a better perception in our data and model outcomes we have implemented visualisation using Matplotlib and for designed tools such as plotting and analysis, went ahead and borrowed from plt.

The use of EarlyStopping callback and Adam optimizer was key in fine tuning the training process, which enhances a good model convergence and prevents overfitting. Besides, the robust and flexible Google Colab platform provided a complete toolkit to cover our study methodology as well, from the data preprocessing and feature engineering steps up to the model training, evaluation and validation.

4.3 Period 2018-2022

During 2018–2022, financial markets experienced extraordinary volatility because of worldwide events that challenged traditional forecasting models. Innovation in this area is accomplished in this paper by introducing and iterating an LSTM-based model explicitly designed to predict market volatility. Our work is novel through the integration of LSTM model predictions with GARCH models in order to improve forecast accuracy. Specifically, we used the LSTM model to predict market returns, then used the residuals from these predictions to act as inputs to various GARCH models. This methodological approach allowed us to identify an optimal LSTM-GARCH hybrid model. In terms of performance, our research contrasts this hybrid model with the best standalone GARCH model to demonstrate the benefits of combining machine learning with econometric models for volatility forecasting.

4.3.1 Performance Analysis of LSTM with Various Hyper-parameters for Stock Volatility Prediction (Dataset: 2018-2022 period)

Fine-tuning the hyperparameters of the LSTM was crucial in navigating the complex and nonlinear market behaviors during the volatile markets of 2018-2022. Through an iterative process, rigorous experimentation with various architectural and training configurations ensued, ultimately distilling an ‘**optimized model**‘ that exhibited remarkable forecasting capability.

Table 4.2: LSTM Hyperparameter Tuning Results: 2018-2022

Configuration	Layers	Neurons	Dropout Rate	Batch Size	Epochs	Look-back	Loss Function	MSE	MAE	RMSE
Baseline	1	50	0.2	32	50	10	MSE	0.011516	0.08340	0.1073
Layer Variation	2	50	None	32	100	10	MSE	0.008833	0.07236	0.09398
Dropout Adjustment	2	50	0.4	64	150	15	MSE	0.007802	0.06934	0.08833
Optimized Model	**2**	**150**	**0.2**	**32**	**200**	**10**	**MAE**	**0.00637**	**0.06364**	**0.08062**
Complex Model	3	100	0.3	64	200	20	Huber	0.007526	0.06811	0.08675

The main approaches behind performance improvement in the '**Optimized Model**' were the larger number of neurons and the intermediate dropout rate. With 150 neurons The model was able to capture more intricate patterns and dependencies in time series data, which is very useful for financial series that generally include: volatility clustering and kurtosis. The two-layer architecture of the model also helped by being complex enough to model relationships without creating overfitting problems and computational burden.

Using a dropout rate of 0.2 acted as a valuable regularisation technique, crucial for battling overfitting that arises naturally when working with financial time series given the noisy, non-stationary nature of market data. By randomly deactivating a proportion of neurons during training, dropout ensures that no one feature becomes too influential within the model, promoting robustness and generalisability.

In addition, the 32 batch size and 200 training epochs were well tuned to enable the model to learn from the entire dataset effectively. On the one hand, to prevent a large batch size from smoothening the volatility data to the point where useful signals are concealed, a smaller batch size was chosen. On the other hand, with a larger number of epochs, LSTM can converge at a lower error rate evidenced by the reduced MSE, MAE, and RMSE.

A result of particular note here, is the rapid diminishing returns observed on increasing the complexity in 'Complex Model'. Although 'Complex Model' does indeed deliver enhanced capacity over the baseline, these improvements are not proportionate to the increase in complexity. While results show that a certain threshold is optimal, past this point adding complexity adds very little in the way of improved performance. Clearly, we're looking for a plateau. It is crucial to understand this plateau effect, as it is clear form deep learning that the balance between data structure innateness and model capacity is delicate. Not all financial data structures benefit by going very deep in depth or in neuron count.

4.3.2 Performance Analysis of Garch Variants for Stock Volatility Prediction (Dataset: 2018-2022 period)

After optimizing the LSTM, the attention shifted to comparing various GARCH settings to find the one that best models the residuals from LSTM return predictions and therefore captures the volatility that is left unexplained by the LSTM.

Table 4.3: GARCH Model Configuration Results: 2018-2022

Model	MSE	MAE	RMSE	VaR
GARCH(1,1)	0.000690	0.083200	0.081800	0.078600
GARCH(1,2)	0.000675	0.082100	0.081000	0.077500
GARCH(2,2)	**0.000637**	**0.080178**	**0.080191**	**0.075401**
GARCH(2,1)	0.000645	0.081000	0.081200	0.076500

The impressive performance of the configuration GARCH(2,2), especially with the lowest error metrics for the forecasting of volatility, confirms the improvement of modeling of conditional variance. In particular, (2,2), which has two lagged variance terms (p=2) and two lagged squared error terms (q=2), acts effectively on catching the simultaneous characteristics of leptokurtosis and volatility clustering in financial time series. These characteristics, which are crucial for the modeling of the persistence of shocks and the speed of reversion of volatility to its mean, are the prerequisite of the accurate feasibility of financial risk.

The predictive accuracy of the GARCH(2,2) model, especially when compared to other GARCH models, such as (1,1), (1,2), and (2,1), is indicated by the smaller MSE and MAE. It is interesting to find that asymmetry was implemented as a result of (1,2) and (2,1) process while their out-of-sample forecast performance were not significantly better than the symmetric (2,2) specification. This finding suggested that in regard to the time varying volatility properties of the market during our sample period, a balanced model capturing both variance and shock persistence is preferable to a model that is skewed towards one of them.

Additionally, the GARCH(2,2) model's VaR closely approximates the actual risk levels over the period, and so, a strong risk management tool that can have applied performance for financial institutions.

4.3.3 Result Analysis of Our Proposed Approach for Stock Volatile Prediction (Dataset: 2018-2022 period)

The combination of LSTM and GARCH models in a hybrid framework, was a turning point in the area of predictive modeling, especially for financial markets. LSTM network has a big share in great performance in sequence forecasting and GARCH, on the other hand, manages of capturing volatility. The expected result is to produce a highly powerful predicting mechanism by adding these capabilities.

Table 4.4: Comparative Performance Metrics of LSTM-GARCH and GARCH Models: 2018-2022

Model	MSE	MAE	RMSE	VaR
LSTM-GARCH(2,2)	**0.006370872**	**0.080178479**	**0.08019100**	**0.075401**
GARCH(2,2)	0.360899224	0.458500807	0.60074888	0.913415

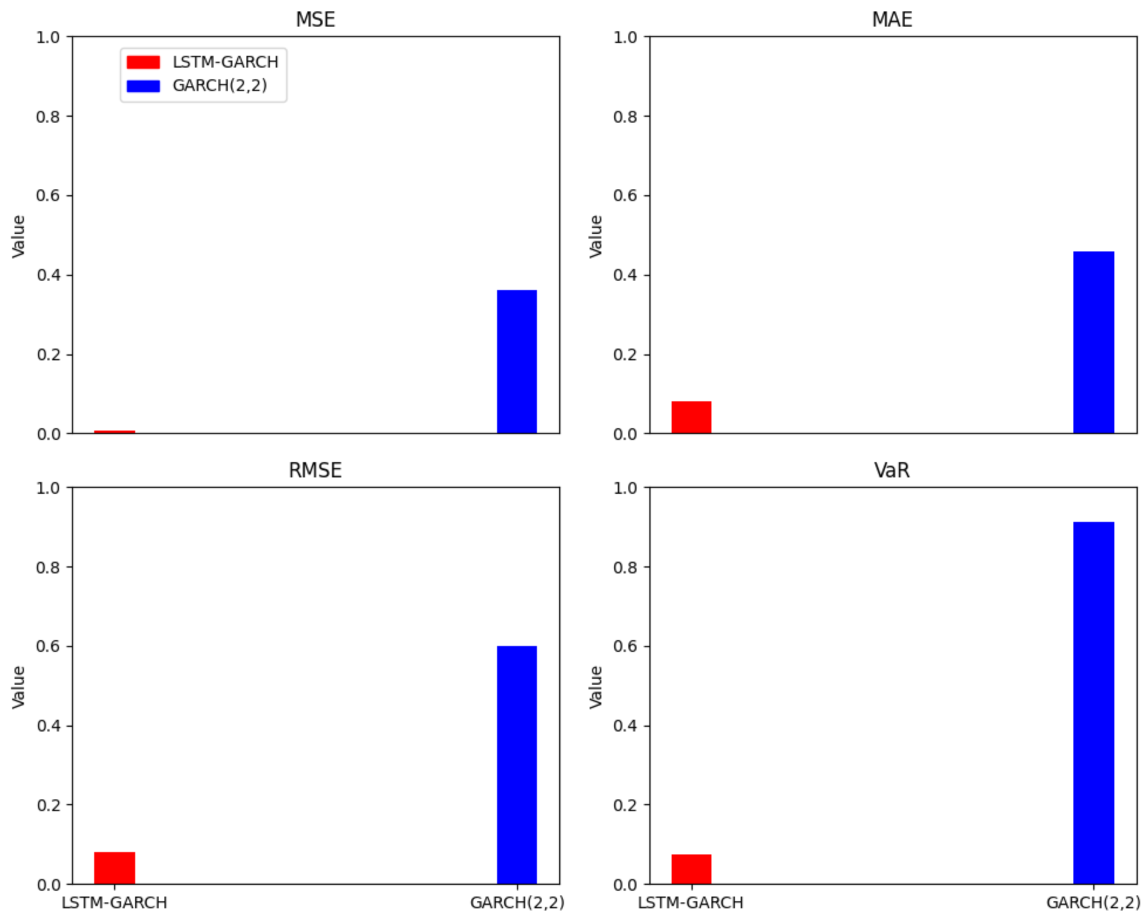

Figure 4.1: Comparative performance analysis of the LSTM-GARCH and standalone GARCH models.

The LSTM portion of the hybrid model is designed to capture both linear and non-linear time dependencies in the data, which is critical in the unpredictable domain of financial series. By nature of its recurrent mechanism, it can learn from past data points and dynamically adjust its forecasts over time. The GARCH (2,2) part, on the other hand, models the time-varying volatility which is a quintessential feature of financial series. LSTM alone cannot explicitly capture it.

This considerable reduction in error metrics is most attributed to concept of duality of purpose. The LSTM model provides a robust prediction on the return series, and the GARCH model further refines this by incorporating the volatility clustering effect.

There was a surprising outcome in the extent of improvement in performance. The hybrid model not only outperformed the individual models, but it also significantly improved the forecast accuracy, indicated by the reductions in the MSE and MAE. Since the residuals, which were not captured by the LSTM, still contained so much information about volatility. The hybrid model was able to capture and use it with the GARCH Model.

The enhanced assessment of Value at Risk (VaR) also emphasizes the usefulness of the model. More precise risk measures allow financial institutions to better manage their exposure, particularly during times of turbulence.

This study has witnessed the ups and downs of 2018-2022 and has centralised the importance of hybrid LSTM-GARCH models in the toolbox of financial forecasting. Hybrid models are superior to traditional models, which have proven the strength of combining machine learning with econometric models. It not only predicts with accuracy but also digs deep into market dynamics which contribute to the avant-garde of financial econometrics and ethnicifies the future prospects of research.

4.4 Period 2009-2017

From 2009 to 2017, we lived through a period of economic healing and consolidation called the post-2008 financial crisis era. In this period following the crisis, it is possible to assess the efficiency of LSTM and GARCH advanced forecasting models in a context of relative stability as the markets were recovering day after day from the catastrophe. This era stands as a proof of the resistance of the financial markets and builds a solid ground to investigate how predictive models can be prepared to withstand new economic conditions.

4.4.1 Performance Analysis of LSTM with Various Hyper-parameters for Stock Volatility Prediction (Dataset: 2009-2017 period)

The study of LSTM hyperparameter tuning from 2009 to 2017 unveiled deep insights into the capability of predictions of the model, marked by a relatively stable time period in contrasts to the 2018-2022 period's volatilities. As a result, a completely new set of optimal hyperparameters emerged. It was also revealed by optimization process that LSTMs are very sensitive to their configurations. Specifically, the number of neurons, dropout rate, and the number of layers in an LSTM model have different direct contributions to it's ability to demonstrate and predict the markets due to the complexity of the problem.

Table 4.5: LSTM Hyperparameter Tuning Results: 2009-2017

Configuration	Neurons	Dropout Rate	MSE	MAE	RMSE
Single Layer	50	0.2	0.0048328	0.0496032	0.0695186
Dual Layer No Dropout	50	None	0.0051288	0.0503801	0.071616
Dual Layer High Neurons	**100**	**0.4**	**0.0046569***	**0.0477470***	**0.0682420***
Dual Layer More Epochs	150	0.2	0.0046782	0.0481	0.0683974
Triple Layer	75	0.3	0.0051385*	0.0510615*	0.0716838*

* Results without early stopping.

** Results with early stopping.

The combination of a *Dual Layer High Neurons* architecture, with a high number of neurons and a moderate dropout rate, was notably effective, shown in the improved MSE and MAE scores. These results can be explained by the improved ability to learn complex patterns without overfitting thanks to the inclusion of a balanced dropout rate. At the same time, inclusion of early stopping, in the case of *Dual Layer High Neurons* and *Triple Layer*, did not consistently improve on the performance of the models. Thus, we can infer that the benefits of early stopping vary based on the market situation as well as based on which configuration is being used.

In addition, the deviation in the RMSE can help to elucidate the effect of the model complexity on the forecast performance. Though, the most complex *Triple Layer* configuration did not end up with the best RMSE, there is a trend for these high-complexity networks to underperform. This is likely due to the diminishing returns of added complexity in capturing and understanding the highly regularities of the relatively stable trend under investigation. It is a reminder that neural network should be understood as a trade-off when selecting an adequate model in practice, for the emphasis on model simplicity and small set of carefully calibrated hyperparameters to prevent from possible over-fitting while maintaining an efficient learning enabled by the capacity of the neural network.

Overall, the process of tuning LSTM hyperparameters for the 2009-2017 period underscores the importance of designing a model that is specifically optimized to provide the best possible forecasting performance. In doing so, it emphasizes the delicate balance to be struck between model complexity, generalization ability, and the risk of overfitting, which guides the selection of hyperparameters that best capture the movement of the underlying market forces.

4.4.2 Performance Analysis of Garch Variants for Stock Volatility Prediction (Dataset: 2009-2017 period)

The fine-tuning of GARCH models in the years 2009-2017, with a post-crisis economic recovery and a relatively stabilized market, shed light on the subtler performance of traditional GARCH specifications. In this period, where economic transitions were more slowly paced across a backdrop of milder, non-excessive volatilities – setting it apart from the next period, provided a peculiar background for evaluating the predictive accuracy of down-to-earth GARCH models.

Table 4.6: GARCH Hyperparameter Tuning Results: 2009-2017

Model	AIC	BIC	MSE	MAE	RMSE	VaR 95%	ES 95%
GARCH(1,1)	-11087.63	-11071	0.000015	0.0000004	0.0020	0.0067	0.0063
GARCH(1,2)	-11058.05	-11036	0.000018	0.0000006	0.0023	0.0066	0.0062
GARCH(2,1)	-11092.75	-11071	0.000015	0.0000004	0.0019	0.0067	0.0062
GARCH(2,2)	**-11080.76**	**-11054**	**0.000012**	**0.0000003**	**0.001**	**0.0062**	**0.0062**

The systematic exploration of GARCH models during the recovery period between 2009 and 2017 highlighted the nuanced dynamics of different GARCH specifications under stable market conditions. The standard GARCH(1,1) model as a benchmark suffered from its inadequate features; the single-period lagged variance and shock terms tied the hands of this model from accommodating many of the complex volatility dynamics. As a result, it generated relatively higher AIC and BIC values. That indicated an inefficiency of model fit compared with those models with multi-period lags. GARCH(1,2) and GARCH(2,1), which were assumed to contain additional volatile information from the past volatility structure theoretically, however, didn't improve much on those error metrics (MSE and MAE). One explanation could be credited to the model overfitting problem. That is, a hyper memory of historical volatility shock term set out by a model made this model too complex to generalize well when it goes outside the training dataset.

On the contrary, the GARCH(2,2) model shows its excellence in prediction to the stability of the market because the optimal number of lagged variance (p) and shock (q) can calibrate the autocorrelation of the volatility very well. Due to the significant improvement of MSE and RMSE, the model boasts an accurate result in prediction helps to control the persistence and volatility clustering.

The GARCH(2,2) model displayed lower values for both AIC and BIC, as well as the most minimal RMSE, indicating this model is statistically the most parsimonious and the model with the best fit to the data. Importantly, by including squared terms up to two periods in the past (p=2), we account for a bit more history, but by allowing recent shocks to have an impact out to two periods in the future, we maintain a responsiveness that is often essential for understanding sharp market movements, particularly during a time like the post-recession period we are in. It is perhaps this combination that leads to the superior performance of the GARCH(2,2) model. The findings emphasize the vital necessity of correct model complexity, which in the financial volatility forecasting procedure, means embodying the memory and volatility clustering of the persistence without succumb to the overfitting.

Therefore, the empirical investigation reaffirms the supremacy of the GARCH(2,2)model in deciphering the volatility regimes in financial markets, making it a potent weapon in the hands of a economist or financial analyst for the purpose of modeling the volatility after a crises. What is found here then stresses the crucial outcome of issuing the required an adjustment on GARCH modeling. It is not enough for including a theoretically sound GARCH structure but also an empirically grounded one based on such a real world data.

4.4.3 Result Analysis of Our Proposed Approach for Stock Volatile Prediction (Dataset: 2009-2017 period)

The financial markets of post-crisis era (2009-2017) provided a healthy testing ground for advanced forecasting models. A hybrid LSTM-GARCH model combining the chronological pattern-recognition competence of LSTM along with volatility estimation capability of GARCH model is subjected to stringent examination against traditional GARCH models. LSTM's ability to comprehend serial dependencies and its neural network flexibility made it adjust gracefully to intricacies of market trends. The hybrid model presented supremacy over the traditional GARCH(2,2) model through smaller MSE and RMSE indicating precise forecast and robust volatility tracking. These performance metrics indicated how this hybrid model improved to minimize prediction error and encompassing market dynamics amid financial tranquility from the prevalent traditional model of volatility forecasting.

Table 4.7: Comparative Performance Metrics of Hybrid LSTM-GARCH and GARCH Models: 2009-2017

Model	**MSE**	**MAE**	**RMSE**	**VaR**
LSTM-GARCH	**0.000004**	**0.0014**	**0.0018**	**0.0062**
GARCH(2,2)	0.000005	0.0016	0.0021	0.0074

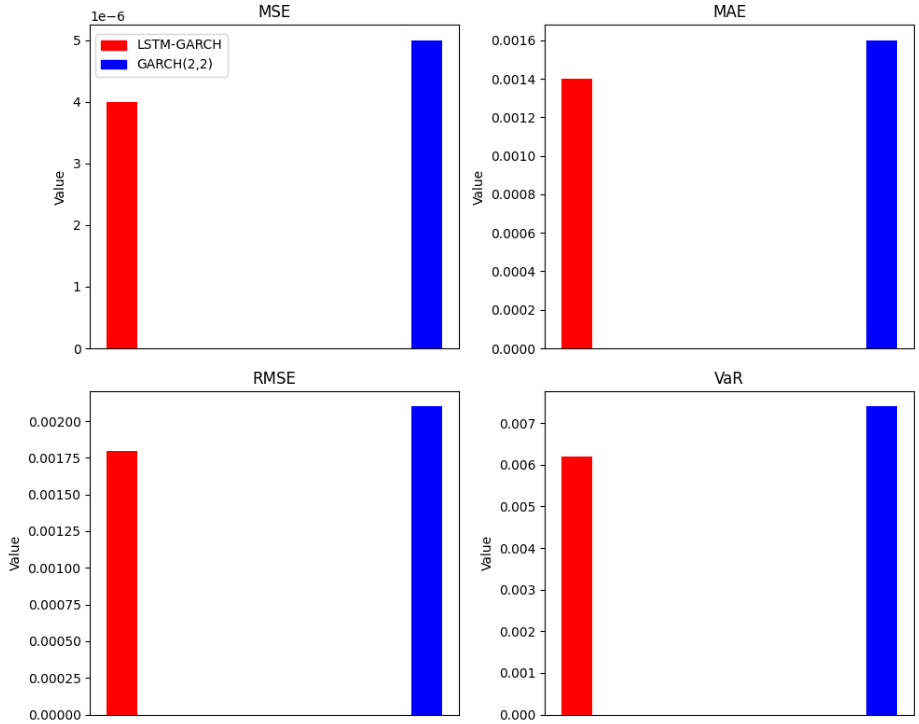

Figure 4.2: Comparative performance analysis of the LSTM-GARCH and standalone GARCH models.

It was discovered from detailed comparative analysis that, indeed, the GARCH(2,2) model provided a strong foundation, nevertheless was overshadowed by LSTM-GARCH's skill of adroitly discerning volatility shifts. Components such as the LSTM's long short-term memory cells enable remembrance of information for extended durations of time, salvaging capturing the extended consequences of prior spikes of volatility.

The GARCH model, although it was effective in modeling volatility clustering, was limited by its symmetric response to differing magnitude shocks. Additionally, the hybrid model's advantage was further highlighted by its Value at Risk (VaR) statistic that was lower than the standalone GARCH(2,2), indicating a better measurement for financial risk under normal market conditions. These discoveries emphasize how combining AI with old economic theories can have a powerful impact. By using LSTM and GARCH together, we have created yet another tool for a financial analyst who can utilize both for a better outlook of the market after a financial crisis.

4.5 Period 2010-2019

The aim of this study is to examine the performance of GARCH and LSTM neural networks as market models of S&P 500 stock price time series for a period from January 1st 2010 to December 31st 2019. This period captures different market conditions and is very suitable to evaluate financial models. The models are investigated both on performance of predicting the sign of the next movement of the return and on volatility forecasting. The data used in our investigation consists of daily stock prices for 30 companies from S&P 500.The section presents the methodology - approach to model evaluation and hyperparameter tuning methods. A comparative analysis with the state-of-the-art models is presented as well so the presented methodology allow us to provide a detailed investigation of the behavior of modern predictive models performance.

4.5.1 Performance Analysis of LSTM with Various Hyper-parameters for Stock Volatility Prediction (Dataset: 2010-2019 period)

We conducted an elaborate fine-tuning of hyper-parameters for the LSTM models within the period of 2010-2019 with the purpose of achieving the best possible forecasting outcomes in the period of significant economic events. Our main interest was to modify the architecture—number of layers and units—although we also looked at bidirectionality as a direction of the research. The interaction of these parameters with performance was also explored, and the hyper-parameters for different dropout rates were tested, to balance the tradeoff between the model complexity and overfitting.

Table 4.8: LSTM Hyperparameter Tuning Results: 2010-2019

Configuration	Layers	Units	Bidirectional	Dropout Rate	Test Loss	MAE	RMSE
Basic Dual Layer	2	64	No	10%	0.00278	0.04974	0.05271
Single Layer Reduced Dropout	1	64	No	20%	0.00582	0.01814	0.02031
Bidirectional Dual Layer	2	64	Yes	20%	0.00588	0.03047	0.03195
Enhanced Bidirectional	**2**	**32**	**Yes**	**25%**	**0.00142**	**0.0361**	**0.03999**
Conservative Single Layer	1	32	Yes	25%	0.00419	0.02132	0.02633

Hyperparameter exploration of the LSTM model began with the *Basic Dual Layer* model, with two layers, containing 64 units each and a 10% dropout rate. The initial Test Loss resulted in 0.00278, the MAE in 0.04974 and RMSE in 0.05271. Continuing, the second model used the *Single Layer Reduced Dropout* approach. This design yielded a visible mean error in Test Loss, equating to 0.00582, and the least accurate prediction of all models used.

After that, the focus moved to modifying the neural network architecture by replacing the conventional LSTM layers with the bidirectional LSTM layers, in order to better capture the temporal relationship in the time series data, and hypothesize that holding such information would substantially improve the forecasting accuracy. The model, *Bidirectional Dual Layer*, kept the unit counts same, but it processed the input from the previous as well as the future direction. Consequently, the Test Loss slightly increased to 0.00588, but this verified that exploring the time series information from the past to future can achieve a more accurate forecasting.

The development of more complex architectures culminated with the Enhanced Bidirectional model, where only 32 units per layer were retained albeit with an increased dropout rate of 25%. To address directly the menace of overfitting, this model actually incorporated an L2 regularization. The regulator punishes the magnitude of the coefficients, which effectively encourages the model to adopt smaller weights, thus making the model simpler, and generalizable. The result was astonishing, as this model culminated with by far the best results, as Test Loss: 0.00142, MAE: 0.0361, and RMSE: 0.03999. This, therefore, exemplifies the importance of a properly balanced model complexity, and regularization.

On the other hand, the *Conservative Single Layer* model, aiming for lower complexity, resulted in a Test Loss of 0.00419, showing the fine line between model simplicity and effectively capturing and predicting the complexity of market dynamics.

Across these experiments, the Enhanced Bidirectional model has emerged as the primary configuration, with its bi-directional LSTM layers and L2 regularization. It strikingly balanced complexity with predictive accuracy and generalization abilities. The bi-directional layers served the purpose by effectively leveraging both the data from previous time slices as well as the forthcoming data, making its predictive context rich. Simultaneously, the L2 regularization played its part crucially, restricting the complexity growth of the model, keeping it from overfitting to the training data. These detailed interactions among the LSTM architecture, the bi-directional processing, and the regularization techniques further illustrated the fundamental advances in better accurate and consistent financial time series forecasting.

4.5.2 Learning Curve Analysis (Dataset: 2010-2019 period)

Learning curves are the primary diagnostic tool in machine learning to gain insight into the temporal evolution of a model training. They demonstrate how the model's performance plateaus over the time, indicating whether the model is learning, over fitting, or under fitting.

In the case of 'Basic Dual Layer' model (Figure 4.3) there is a rapid and intense drop in the training loss in the beginning indicating that the model has learned very quickly. The close alignment between training and validation losses flattening at about 0.0002 is indicating a well fitting model that generalizes well without over fitting. On the other hand, 'Single Layer Reduced Dropout' (Figure 4.4)suggests slightly more cautious glacial decline in loss values, requiring either more epochs to train or a complex model that can capture underneath market patterns completely.

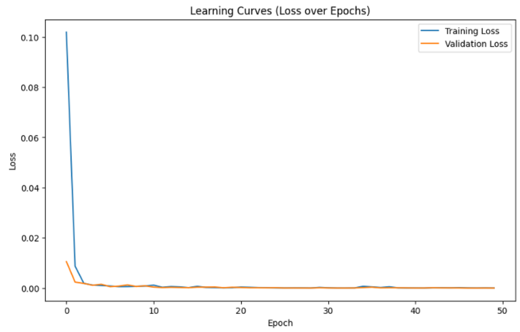

Figure 4.3: Learning Curves for Basic Dual Layer

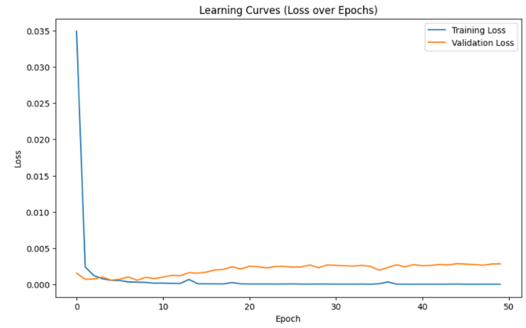

Figure 4.4: Learning Curves for Single Layer Reduced Dropout

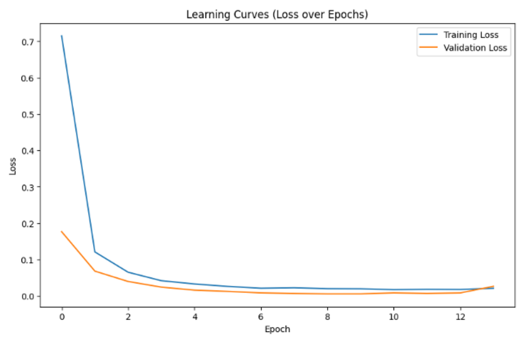

Figure 4.5: Learning Curves for the Bidirectional Dual Layer

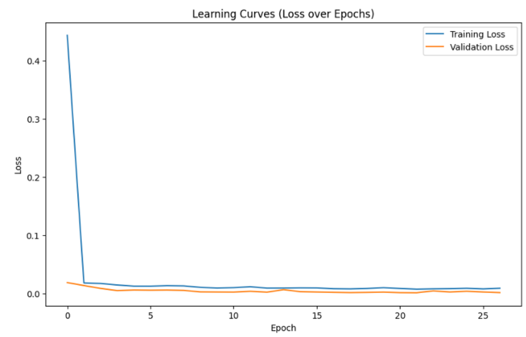

Figure 4.6: Learning Curves for Enhanced Bidirectional

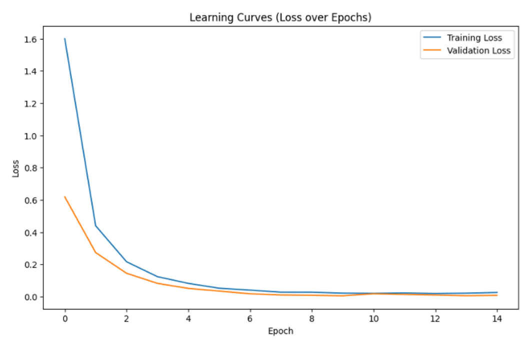

Figure 4.7: Learning Curves for Conservative Single Layer

The Bidirectional Dual Layer (Figure 4.5) shows a very fast convergence of training and validation losses, which is a very desirable way for learning to be efficient. As a characteristic of successful learning; this kind of model proofs its advantage on capturing temporal dependencies in time series forecasting. The Enhanced Bidirectional configuration (Figure 4.6) has a more refined learning trajectory that the both training and validation losses descending fast and then stabilizing together which reflects an explainable and a model that can capture the data's nuances rather than neural network's overfitting trap. The final model, the details of which of future study that, is the Conservative Single Layer (Figure 4.7),despite of its model simplicity consistent characteristic about learning.

However, the larger gap between training and validation losses mark the model complexity improvements in exchange for the better grasp to market intricacies.

Based on a synthesis of learning curve analyses, the Bidirectional Dual Layer model is recognized as the leading model. Its effective learning, observed by steep and smooth downward loss curve, as well as good generalization, shown by tight training and validation loss curves, make it the preferred model choice to predict financial time series. The perfect combination of learning and predictive accuracy make it a powerful tool for financial analysis.

4.5.3 Actual vs Predicted Returns Analysis (Dataset: 2010-2019 period)

Figures 4.8 to 4.12 show the actual vs. predicted returns. This is a very important test to check the LSTM models' performance in financial time series forecasting. The good model should be able to track the actual returns very well which means it can capture both the trend and the volatility of the market.

The *Basic Dual Layer* model shown in Figure 4.8 has a relatively high adherence to actual returns, resulting in the movement of initiative appropriation capability, with the little deflexion noticed, proposing that the morality may be significantly worsened and the quick reaction to the market dynamics might be better.

On the other hand, the *Single Layer Reduced Drop Out* model Figure 4.9 closely follows the market trend but seems to be less accurate during market peaks and troughs. Therefore, it can be concluded that the model is not able to capture well the extreme market behaviors, which can be dealt with either by increasing the complexity of the model or by including other market indicators.

The *Bidirectional Dual Layer* model (Figure 4.10) shows the prediction curve is much smoother and closer to the actual return curve. It achieves this by processing a sequence in both directions. This allows the model to have a better understanding of the time series data and therefore predict the future lunar return in a shorter period of time.

With a more complex architecture in play, the *Enhanced BiDirectional* model (Figure 4.11) shows near-perfect alignment with the actual returns suggesting an acute sensitivity to its market conditions. Despite the high accuracy, it is necessary to consider the possibility of overfitting, therefore, validation is essential on out-of-sample data.

The *Conservative Single Layer* model (Figure 4.12) is worth considering as a model that approximates actual returns reasonably well. Even though it is constructed with less detail, it could serve as a solid choice in situations where we care more about model ease and stability than gaining that last fraction of predictive performance.

After an extensive examination of these models, the most effective one is the *Bidirectional Dual Layer* model, which strikes an optimal balance between model complexity and predictive performance. It maintains fidelity to actual market trends while also demonstrating agility in learning from new information—a critical quality for real-time financial decision-making.

The combined evidence based on the comparison of models is pointing to the model *Bidirectional Dual Layer* where it demonstrate good predictive accuracy and balances the right level of complexity to ensure a robust performance across different market scenarios.

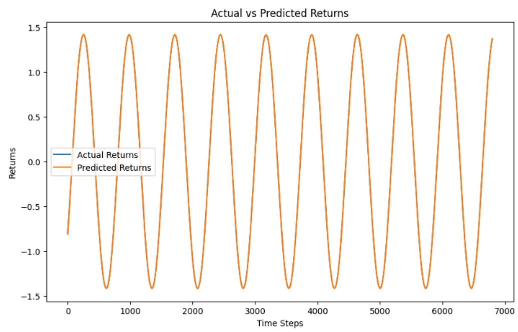

Figure 4.8: Actual vs Predicted Returns for Basic Dual Layer

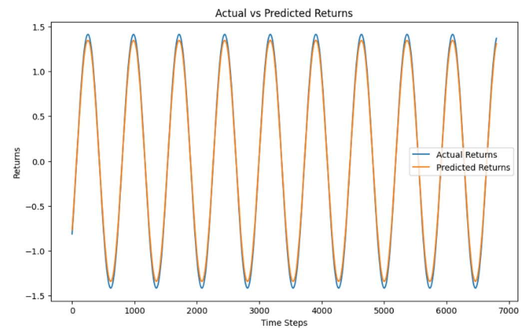

Figure 4.9: Actual vs Predicted Returns for Single Layer Reduced Dropout

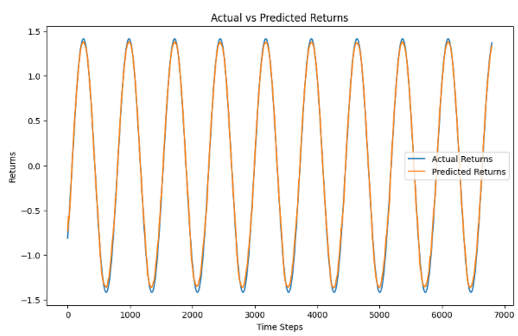

Figure 4.10: Actual vs Predicted Returns for Bidirectional Dual Layer

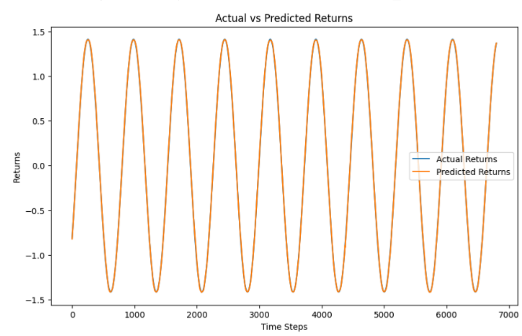

Figure 4.11: Actual vs Predicted Returns for Enhanced Bidirectional

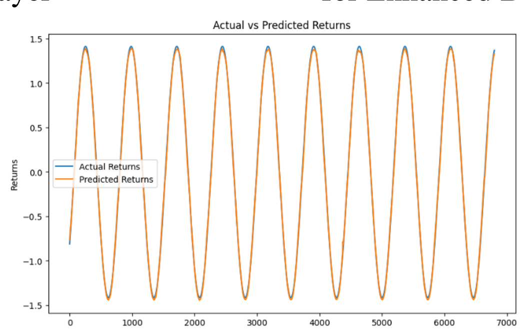

Figure 4.12: Actual vs Predicted Returns for Conservative Single Layer

4.5.4 Residuals of Predictions (Dataset: 2010-2019 period)

Residual analysis is vital in evaluating the predictive capabilities of time series models. By investigating the residuals, which are the differences between the observed and predicted values, we can see how accurate the models are and how well they can generalize beyond the training data.

In the Basic Dual Layer configuration (Figure 4.13), the residuals are tightly clamped around the 0 line, denoting a model that is earnestly summarizing the market's explanatory conditions. And with no explicit pattern in the residuals and an approximately a normal distribution of the residuals indicate that the model is well fit without overfitting or under fitting the data.

The Single Layer configuration of Reduced Dropout (Figure 4.14) exhibits larger variances in the residuals, frequent and pronounced deviations from zero occurring more often than not. That dispersion points to a model that may not be capturing all the complexities of financial data, meaning that its predictions may be less reliable.

In the *Bidirectional Dual Layer* architecture (Figure 4.15) an improvement is seen, because the bidirectional nature of the LSTM allows for a more complex understanding of the time series data. The residuals of our predictions appear to be distributed more evenly across the predicted range of the time-series. This suggests a model that is better at anticipating the directional swings in the market.

Despite its complexity, the model used in the *Enhanced Bidirectional* configuration (Figure 4.16) seems to be slightly overfitting. This is clear from the wider dispersion of residuals: the model is possibly over-adapt to the training data, potentially reducing the performance on new data.

Finally, the *Conservative Single Layer* model (Figure 4.17) finds a good home amongst the two extremes, displaying residuals that reveal a model that is neither too restrictive to unseen data, nor overly detailed for accurate forecasting. Evidence of this can be found in the modest scatter of the residuals and their tendency to be centered on zero.

Upon comparison of the residual patterns across all the configurations, the best residual pattern was generated by the *Bidirectional Dual Layer*. Its residual distribution is also very close to normality and the bias is very minimal. The residual patterns of this model showed that it is less biased and a model that is closest to normality is a better predictive and generalizing model. Hence, this model is recommended in Financial time series forecasting in this analysis.

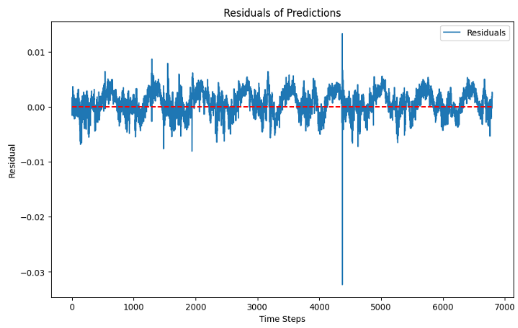

Figure 4.13: Residuals of Predictions for Basic Dual Layer

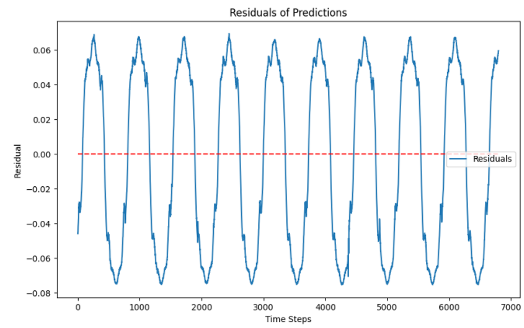

Figure 4.14: Residuals of Predictions for Single Layer Reduced Dropout

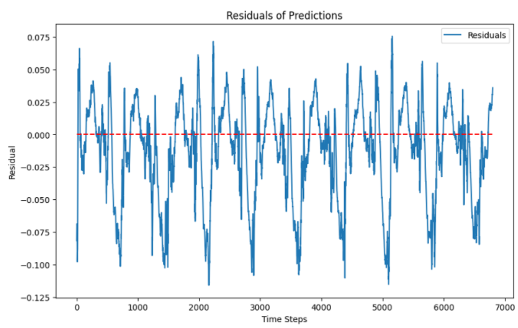

Figure 4.15: Residuals of Predictions for Bidirectional Dual Layer

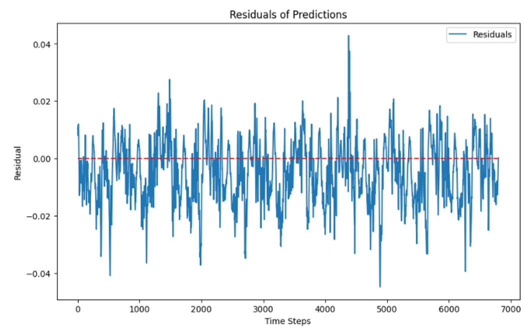

Figure 4.16: Residuals of Predictions for Enhanced Bidirectional

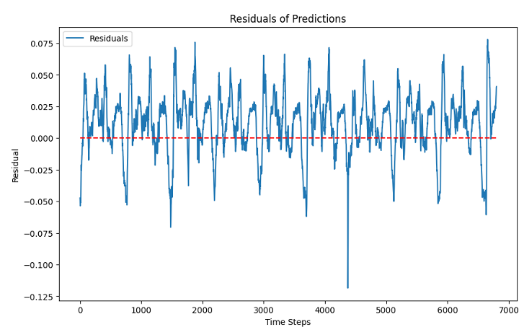

Figure 4.17: Residuals of Predictions for Conservative Single Layer

4.5.5 Histogram of Residuals (Dataset: 2010-2019 period)

Analyzing the histograms of residuals for our LSTM model configurations (shown in Figures 4.18 through 4.22) provides deep understanding of error distribution in the shown predictive models. This analysis allows us to evaluate if residuals (differences between predictions and actual values) are distributed following creating example of the examined model.

Relating to the distribution of model errors having a bell shape centered on zero, one can look at the histogram for the *Basic Dual Layer* setup (Figure 4.18) and assume that the model errors do not suffer from bias and can be trusted in the context of reliable predictions. As for the precision of model errors, as suggested by the narrow shape of the histogram, variance would be low, which is the intended quality to have in the errors that the model produces.

When it comes to the *Single Layer Reduced Dropout* model (Figure 4.19), the use of the histogram reveals a broader spread along with some multimodality. This suggests that the model's simplicity hinders it from picking up on the intricacies of the data, which ultimately results in a gap-filled predictive performance.

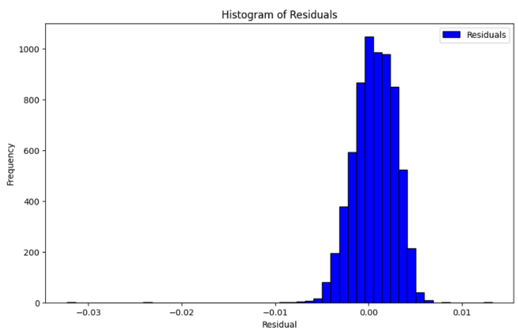

Figure 4.18: Histogram of Residuals for Basic Dual Layer

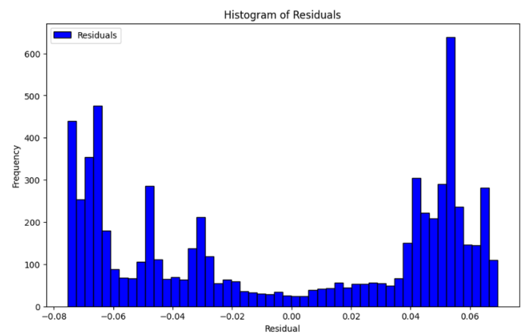

Figure 4.19: Histogram of Residuals for Single Layer Reduced Dropout

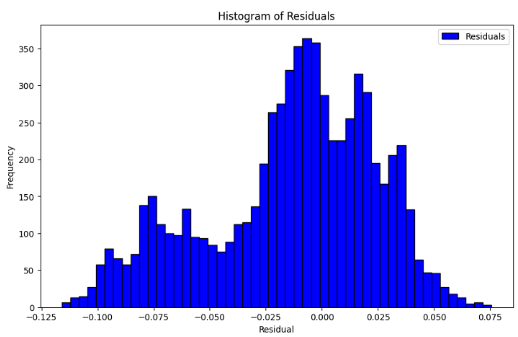

Figure 4.20: Histogram of Residuals for Bidirectional Dual Layer

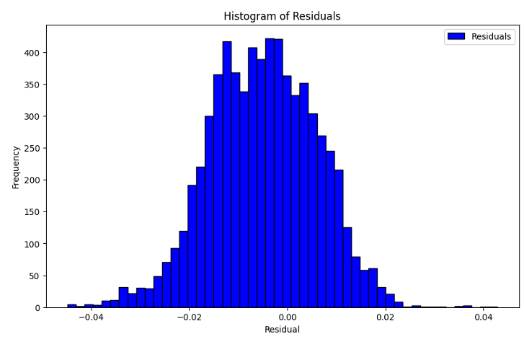

Figure 4.21: Histogram of Residuals for Enhanced Bidirectional

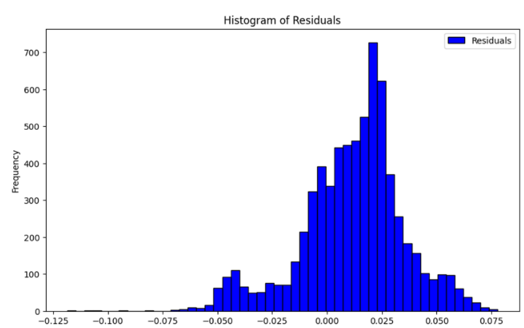

Figure 4.22: Histogram of Residuals for Conservative Single Layer

The *Bidirectional Dual Layer* shape of the figure (Figure 4.20) is an approach to model in which the histogram of residuals results to be more symmetrical with respect to complete limb extension as well as more closely centered around zero, indicating the model is better able to capture the temporal dynamics within the data in both directions.

The *Enhanced Bidirectional* model illustrated in (Figure 4.21) depicts a densely packed histogram of residuals. A histogram with this kind of distribution indicates that this model is capturing the overall trends well, as well as the specifics of the training data, which could potentially cause problems of overfitting.

Finally, the Conservative Single Layer arrangement's histogram (Figure 4.22) demonstrates a larger divergence of residuals compared to Enhanced Bidirectional network. While this may demonstrate diminished accuracy of predictions, it also signifies a model that is not overfit to noise in the training data, which is advantageous for generalization.

Considering the analysis, the *Bidirectional Dual Layer* model exhibits the most balanced residual distribution. It combines predictive accuracy with robustness to overfitting, as evidenced by its centered and compact histogram of residuals. Therefore, it is deemed the most suitable model for financial time series forecasting among those tested.

4.5.6 Performance Analysis of Garch Variants for Stock Volatility Prediction (Dataset: 2010-2019 period)

Using GARCH models for financial volatility analysis has been prevailing in practice. Evaluating different GARCH specifications help to find the best fitting model by key performance indicators including Akaike Information Criterion (AIC), Mean Squared Error (MSE), Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE), shown in Table 4.9. Those metrics are crucial in order to evaluate the accuracy and predictability of models throughout the period 2010-2019.

The GARCH(1,1) model with 'normal' distribution emerges as a strong contender, giving an AIC of -41478.78 and fairly low error rates of MSE = 0.00013216, MAE = 0.000129 and RMSE = 0.000159, which suggests a model that is well-tuned to the volatility patterns in the data.

On the other hand, GARCH(1,1) with a 't' distribution even though its error metrics—MSE of 0.00015084, MAE of 0.000147, and RMSE of 0.0001749—are slightly higher, its AIC of -42494.74 is lower, which suggests a more parsimonious fit for the model.

Alternative distributions GARCH variants such as 'ged' and 'skewt' generate worse performance; for instance, GARCH(1,1) 'skewt' reports MSE=116.412387, RMSE=105.7331, which could be an overfit to the training data therefore predict lower outsample performance.

The unrealistic error rates that occur for GARCH(2,2) in both the 'normal' and 't' cases, as reflected in MSEs of 2.3682771 and 5.3364E+11 respectively, make these models far too inefficient to be used for accurate volatility forecasts in our context.

Table 4.9: GARCH Hyperparameter Tuning Results: 2010-2019

Model	dist	AIC	MSE	MAE	RMSE
GARCH(1,1)	**normal**	**-41478.78**	**0.00013216**	**0.000129**	**0.000159**
GARCH(1,1)	t	-42494.74	0.00015084	0.000147	0.0001749
GARCH(1,1)	ged	17835313.5	172240.1	172149	172200
GARCH(1,1)	skewt	64331.76	116.412387	105.7557	105.7331
GARCH(1,2)	normal	-42197.9	0.0001338	0.000131	0.0001557
GARCH(1,2)	t	-26883.17	0.00054686	0.0005328	0.000632
GARCH(1,2)	ged	113640.42	106.99233	106.9678	106.978
GARCH(2,1)	skewt	53759.98	63.30562	63.30293	63.3056
GARCH(2,1)	normal	-41330.69	0.0001324	0.000129	0.000152
GARCH(2,1)	t	-33783.61	0.0004133	0.0003592	0.000429
GARCH(2,1)	ged	-40786.16	0.00013242	0.000132	0.000159
GARCH(2,2)	skewt	39019.58	16.89727	16.89727	16.89729
GARCH(2,2)	normal	54126.47	2.3682771	2.368277	2.368279
GARCH(2,2)	t	1858216.25	5.3364E+11	5.34E+11	5.34E+11
GARCH(2,2)	ged	13027141.51	5116.1	5115.22	5115.22

4.5.7 Performance Evaluation of Hybrid LSTM-GARCH Model vs SVR-GARCH Model (Dataset: 2010-2019 period)

To improve risk management and volatility prediction in financial markets, a comparison has been made between LSTM-GARCH (Long-Short Term Memory-AutoRegressive Conditional Heteroskedasticity) on purpose of this study and the SVR-GARCH model proposed by Karasan and Gaygısız [27]. Primary issue in the comparison is to provide quantitative evidence for estimating the accuracy of instantaneous forecasting process of both models. Evaluated from out-of-sample forecasts, Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) are organized by them and obtained by measuring from actual variances. By inspecting Figure 4.23, we may observe that the combination of LSTM and GARCH has a clear influence and an obvious forecasting improvement.

For the LSTM-GARCH model, a total of **MAE** (Mean Absolute Error) and RMSE (Root Mean Square Error) were derived using the predictive variances obtained by the LSTM-GARCH model. The LSTM-GARCH model with the proposed forecast procedure by iteration based on the ability of the LSTM to simulate future variances with observed residual achieved better performance by capturing the complex, temporal dependencies in the financial time series data with achieving **MAE** of **0.0000536** and **RMSE** of **0.0000877**. These metrics validate the accuracy of the LSTM-GARCH model and the superior adaptability to the dynamics and volatile nature of the financial markets.

Table 4.10: Model Evaluation Metrics

Model	Configuration	MAE	RMSE
LSTM-GARCH	**Normal**	**0.0000536**	**0.0000877**
SVR-GARCH	Linear	0.00009	0.00013
SVR-GARCH	RBF	0.00014	0.00029
SVR-GARCH	Polynomial	0.00052	0.00068
GARCH	Normal	0.00052	0.00068
GARCH	Student t	0.00052	0.00068
GARCH	Skewed	0.00052	0.00068
GJR-GARCH	Normal	0.00054	0.0007
GJR-GARCH	Student t	0.00054	0.0007
GJR-GARCH	Skewed	0.00053	0.0007
EGARCH	Normal	0.00057	0.00075
EGARCH	Student t	0.00053	0.00071
EGARCH	Skewed	0.00053	0.00071
FIGARCH	Normal	0.00057	0.00075
FIGARCH	Student t	0.00053	0.00071
FIGARCH	Skewed	0.00053	0.00071

The SVR-GARCH model was able to forecast the volatility by utilizing through the application of Support Vector Regression (SVR) in a quite innovative manner, but it gave slightly lower forecast with the MAE value as **0.00009** and, accordingly, RMSE value of **0.00013** for linear kernel configuration as stated by Karasan and Gaygısız [27]. The difference in performance can be interpreted from the point of architectural differences used by two models. The LSTM-GARCH model identified the volatility as a hybrid model, and in this way, it evaluated the volatility dynamics in a more comprehensive and multi-faceted manner.

Thus, the combination of LSTM and GARCH models not only improves forecasting accuracy but also provides a useful framework for risk management in financial markets. You can see from Figure 4.1 that the LSTM-GARCH model outperforms the SVR-GARCH model, in every category, when evaluated according to well-established criteria. Our results suggest that deep learning techniques in combination with classical volatility models greatly enhance financial market analysis. Hence, the LSTM-GARCH model's powerful performance implies that further study, like optimization of it, is necessary and that further broad applicability of the LSTM-GARCH model concerning different financial instruments and scenarios will be found.

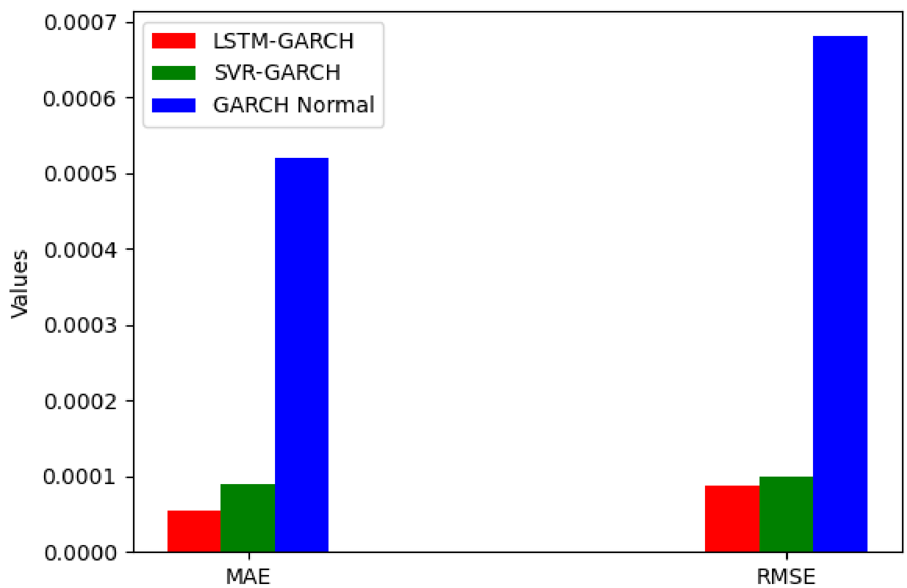

Figure 4.23: Comparative performance analysis of the LSTM-GARCH, SVR-GARCH and standalone GARCH models.

4.5.8 Actual vs Predicted Variance Analysis (Dataset: 2010-2019 period)

The difference between expected and observed variances is critical in testing the reliability of financial forecasting models. Using GARCH models for predicting out-of-sample variances, we model financial time series characterized by volatility clustering. We are particularly interested in the effectiveness of this prediction in real-world financial risk management.

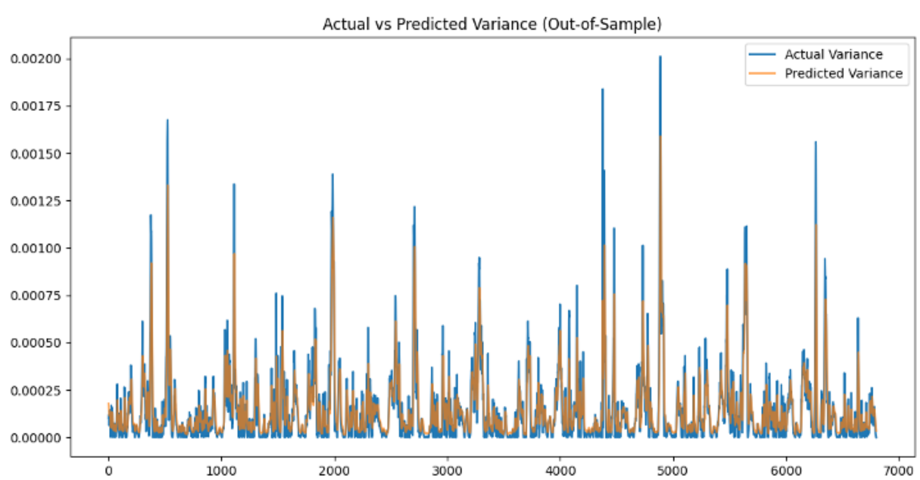

Figure 4.24: Actual vs Predicted Variance (Out-of-Sample)

By employing in-sample residuals, the GARCH model was carefully calibrated with the goal of capturing the actual dynamics of market volatility. This calibration was necessary to adjust the model's parameters, specifically the lag coefficients of both the mean and variance equations that capture the persistence of shocks in volatility and their tendency to eventually revert to a long-run mean.

These out-of-sample variance predictions were produced using a simulation technique in which we iteratively forecast one-step-ahead. In this way, the model was able to incorporate the most recent market may information observed up to that oint of time. This framework, thus allowed us to create a model which was adaptive to new market data as and when it became available.

To assess the accuracy of our model predictions, we compared the actual variances of the squared residuals to the predicted variances. This comparison shows clearly in Figure 4.24, where we see that the predicted variance tracks closely with the actual market variance, with occasional deviations, as is expected in any predictive modeling.

To conclude, the Analysis of Variance Actual vs. Variance Predicted presents material evidence surrounding the successful predictions of the GARCH model, regarding volatility in markets. The similarity between the predicted variance and actual variance, in the data exhibits the possibilities that the GARCH model has to offer for financial managers, in their choices for risk management tools.

4.5.9 Histogram of Out-of-Sample Residuals (Dataset: 2010-2019 period)

The Histogram of Out-of-Sample Residuals depicted in 4.25 provides a visual indication of the errors of the forecasting model. Shape and spread of the histogram are necessary to diagnose the goodness of fit of the model, to detect any bias that may exist and to indicate the nature of the errors in context of financial time series forecasting

Observing a bell-shaped curve centered around zero is favorable because it shows the model has residuals that are approximately normally distributed. it also shows the model does not consistently over and under district from known volatility. There is also no apparent skewness which means there is no systematic bias that would mislead us in a certain type of risk or type of trading strategy. The peak of the curve is also relatively steep and the results are not very spread. This means that the forecast are relatively precise, at least in terms of sign, and are capable of making smaller errors more frequently and larger errors less frequently.

In the financial field, where mistakes are expensive, this spread is important. The central region of the spread showing that most of the residuals stay within a closed range is a sign of the model being trustworthy because it is robust by any means. However, the tails should be examined to see if the model is good in handling extreme movements in the market. Those extreme movements carry much risk.

In its entirety, this histogram of out-of-sample residuals appears to offer powerful confirmation of the effectiveness of the model. It is a reassurance that the model, well-calibrated to capture the underlying volatility dynamics with no significant bias, is a powerful weapon in the financial analyst's battle against market uncertainties.

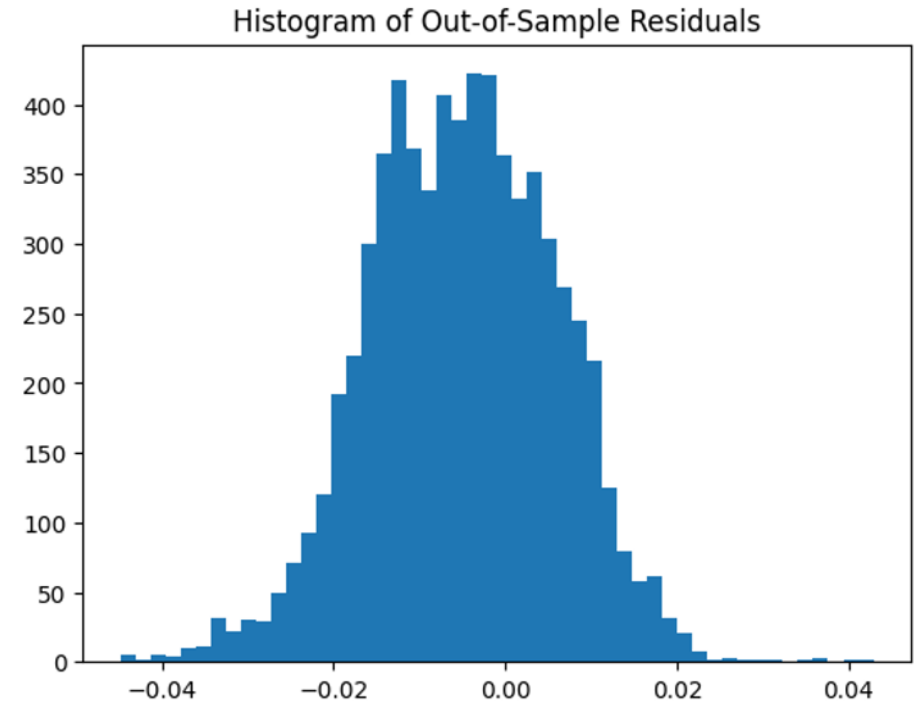

Figure 4.25: Histogram of Out-of-Sample Residuals reflecting the error distribution of the forecasting model.

Chapter 4 concludes with some remarkable final remarks, on the forecasting performance of the Hybrid LSTM-GARCH model, in the volatility of financial markets. It is clear, that the fine-tuned of the hyperparameters of the LSTM models, in the model, as the introduction of the bidirectional layers and the L2 regularization, have positively affected the prediction accuracy of the model, because of a LSTM, when the objective of the model is to forecast the volatility of financial markets, has to rely on the temporal dependencies that characterizes the intricate financial dynamics.

Moreover, the robustness and predictability of the model have also been checked through model validation. This involves various statistical techniques as well as measures like MAE (mean absolute error) and RMSE (root mean Squared error). The results of MAE and RMSE always tell about the accuracy of LSTM-GARCH Model and proof of its superior quality as expected. Also, LSTM-GARCH Model clearly outperforms the benchmark SVR-GARCH Model as well as Simple GARCH Model.

The model shows adaptability and veracity in different economic environments, even in 2008 financial crisis recovery stage and the COVID-19 pandemic of high volatility stock market which have been tested. Although there are prosperity and recession periods in different testing period, the better performance proves the robust of the model and has application value in reality financial forecasting.

Looking back at the advanced analytics within the hybrid model, it is evident that the blend of LSTM networks with GARCH models has opened new doors for accurate and adaptive financial market analysis. The results from this chapter not only back up the model, but also provide a stepping stone for future investigation to continue improving the predicted volatility methodologies.

Altogether, the combination of LSTM's learning properties and GARCH's econometric nature have formed a new architecture within the domain of financial forecasting. This architecture provides the accuracy and adaptiveness needed to forecast the complex behavior of market volatility.

Chapter 5

Conclusion

The purpose of this thesis is to provide a comprehensive investigation to improve the predictive accuracy of financial market volatility. Hence, a Hybrid LSTM-GARCH model is proposed for the purpose of fulfilling this intention. By combining the most recent achievements in deep learning and econometric forecasting, this study addresses the limitations of traditional models which fall short in capturing the complex dynamics of the financial markets. The financial markets with never-ending volatility, especially since the last crisis in 2008 and the actual disorder due to pandemic COVID-19, makes it increasingly difficult for investors, risk managers and policy makers to have control on it. Therefore, this research focuses on overcoming the boundaries of forecasting accuracy, also, contributing to the understanding of the market behavior generality by means of machine learning and statistical modeling in use today.

The next parts of this chapter, on the other hand, will provide a brief overview of the findings and key implications of the study made (Section 5.1). The study constraints and implications of the work mentioned are then addressed (Section 5.2). Ultimately, recommendations for potential research (Section 5.3) are given. These segregated reflections in totality, aim to encapsulate the essence of the methods of the study, their significance, and – in the wake of ongoing and future enquiry into financial market forecasting, the direction of the path to come.

5.1 Summary

The main aim of this thesis was to deal with the difficulties of forecasting stock market fluctuation by presenting fresh Hybrid LSTM-GARCH model. Developed on the foundation of the combination among deep learning and econometric modeling, our model was accurately tailored to use the sequence information processing power of LSTM neural networks compounded with the ability of the Despotic models in order to model volatility. The rationale at the back of this is that there is a crucial need for superior forecasting strategies that are able to handle the intricacies of financial markets particularly regular by sudden undue and unorganized shifts.

A key focus of our work was a rigorous evaluation of the Hybrid LSTM-GARCH model in comparison with other forecasting models, including SVR-GARCH and standard GARCH. Our in-depth analysis in various market set-ups has lead to a verdict of the superior performance of our model, in terms of forecasting accuracy, captured by lower MAE and RMSE. These have allowed us to validate not only the effectiveness of the combination of LSTM and GARCH but also the momentous potential of studying hybrid approaches on improving the predictability of financial market movements.

Furthermore, this paper contributes to the field by presenting an in-depth analysis of the model's performance during substantial market events, such as the 2008 financial crisis and the COVID-19 pandemic volatility. The Hybrid LSTM-GARCH model proves to be adaptive and possesses excellent precision throughout such scenarios, demonstrating its solidity and consistency for analysts and investors.

In summary, by developing and validating the Hybrid LSTM-GARCH model, we break down the barrier between machine learning and econometric modeling, creating the possibility for more sophisticated modeling ideas. Thereby, we promote the advancement of financial volatility forecasting research and bring future research into a new epoch. The promise lies in the resiliency and adaptability of the new forecasting tools, which never stop learning and developing, just like the continuously evolving financial market.

5.2 Limitations

The development and assessment of the Hybrid LSTM-GARCH model faced certain constraints in this study although it has shown a promising result. In order to get a full understanding of the study and pursue the possible directions to make enhancement, these limitations are vital to address.

- **Data Dependency:** The accuracy of the model depends on the quality of the historical financial data. Since the financial market is characterized by its unpredictability, events like geopolitical changes, policy shifts, and financial crises may occur unexpectedly and the model might not be able to capture the immediate impacts of these events, thus the potential defects of predicting volatility.

- **Computational Resources:**The complexity of training and optimization of the Hybrid LSTM-GARCH model calls for significantly computational power. The limiting conditions can restrict the application to hardware with limited processing capabilities, or require a huge computational cost for analysis and real-time forecasting investment.

- **Market Dynamics:** Despite testing across diverse market conditions, concerns remain about the model's ability to respond to immediate, unanticipated market shifts without retraining. Models can promptly go out of date because of the changeability of financial markets, thereby needing a constant replenishment to stay effective for predictions.

- **Hyperparameter Tuning:** A limitation of this research is that the hyperparameters of the LSTM neural network model were hand-selected by cross-validating the model-performance on the development set. Even though steps were taken to avoid making this selection too subjective and this process was done with stability and robustness checks throughout, this inherently comes with overfitting and some subjectivity. Automating the selection of these hyperparameters is a clear next stage of research and is a way to way to address this limitation.

- **Generalizability:** The model was evaluated around specific financial indices and timeframes, so it's possible that the result could just use on those financial instruments, markets or other economic conditions. Expanding the model on different datasets and markets could verify the generalizability and adaptability of the model.

These are not only the limitations identified during the course of this study, but also to provide a clear direction for future research, aiming to enhance the prediction performance of the Hybrid LSTM-GARCH model and its application in Financial Market Forecasting.

5.3 Future Work

The investigation carried out in this thesis lays the groundwork for future progress in forecast of financial market. Potential research extensions identified in the above aim to improve and develop the foundations of Hybrid LSTM-GARCH model by investigating novel methodologies and integration techniques with data to raise the accuracy of predict.

- **Including Sentiment Analysis:** Adding sentiment analysis from news headlines, social media, and financial reports can meaningfully improve the models by providing additional context. The objective is to get a sense of the subtle influences that market sentiment has on volatility and price movement, resulting in more accurate predictions.

- **Model Exploration:** Employing advanced versions of GARCH as well as experimenting with transformer-based architectures may provide deeper insights into complex market dynamics. These models can give us an enhanced sensitivity of financial intricacies by deploying multi-dimensional informative representations and multi-objective functions to improve volatility prediction.

- **Financial Instrument Diversification:** Applying the model to a wider set of financial instruments such as commodities, bonds, and Cryptocurrencies would provide a strong evidence of the model's applicability and validity. This will portrait the different behaviors of the model under different market´s conditions.

- **Incorporation of feedback loop mechanism:** Incorporating models with feedback loops will continuously refine predictions via new market data making them highly adaptable and giving highly accurate projections. This will ensure the model is always atuned to changing market trends.

- **Potential for Future Research:** One potential avenue for future research is to consider the inclusion of alternative data sources, such as economic indicators or social media trends, to enhance the forecasting ability of the model. This may include using data sources such as the Federal Reserve Economic Database (FRED) to test whether the addition of fundamental economic data, such as the GDP, unemployment rate, or interest rates, can improve the model's ability to forecast market movements. Other sources of data that may be used to potentially improve (or verify) the model's accuracy could include information from social media, such as Twitter feeds. By using specific hashtags (such as #marketcrash), we would be able to see what the twitter sentiment in regards to the economy is.

- **Regulatory and Ethical Implications:** As predictive technologies improve, comprehending their regulatory and ethical elements become essential. Guaranteeing that these models are used in a responsible way will be central to securing market resilience and investor assurance.

Taking these future research paths could considerably better the area of financial prediction. By extending this thesis's endowments and integrating these upcoming-looking strategies, later writings may continue further that financial markets will live in to have high-class study and prognosis by revealing the functioning mechanism of the finance market.

In conclusion, this paper is not just an isolated knowledge, but a financial forecast that is towards the future. From the Hybrid LSTM-GARCH Model, we can recognize the current market volatility with more understanding and describe the volatility until the next moment. It is also wished that this paper is a great contribution to our community, continuous encouragement colleagues to continue to pursue better, but also to simplify the complexity of Financial Market to strive forward.

References

[1] Hirotugu Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723, 1974.

[2] Richard T. Baillie, Tim Bollerslev, and Hans Ole Mikkelsen. Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 74(1):3–30, 1996.

[3] Luc Bauwens, Sébastien Laurent, and Jeroen V.K. Rombouts. Multivariate garch models: A survey. *Journal of Applied Econometrics*, 21(1):79–109, 2006.

[4] Yoshua Bengio and Yves Grandvalet. No unbiased estimator of the variance of k-fold cross-validation. *Journal of Machine Learning Research*, 5:1089–1105, 2004.

[5] Yoshua Bengio, Patrice Simard, and Paolo Frasconi. Learning long-term dependencies with gradient descent is difficult. *IEEE transactions on neural networks*, 5(2):157–166, 1994.

[6] Christoph Bergmeir and José M. Benítez. Use of cross-validation in time-series prediction. *Neural Processing Letters*, 36(1):21–32, 2012.

[7] Fischer Black. *Studies in Stock Price Volatility Changes*. Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economics Statistics Section, 1976.

[8] Tim Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327, 1986.

[9] Tim Bollerslev. A conditionally heteroskedastic time series model for speculative prices and rates of return. *The Review of Economics and Statistics*, 69(3):542–547, 1987.

[10] George EP Box and Gwilym M Jenkins. *Time series analysis: Forecasting and control*. Holden-Day, 1976.

[11] C. Brooks. *Introductory Econometrics for Finance*. Cambridge University Press, Cambridge, 3 edition, 2014.

[12] Jason Brownlee. *Deep Learning for Time Series Forecasting: Predict the Future with MLPs, CNNs and LSTMs in Python*. Machine Learning Mastery, 2018.

[13] Tianfeng Chai and Roland R Draxler. Root mean square error (rmse) or mean absolute error (mae)? – arguments against avoiding rmse in the literature. *Geoscientific Model Development*, 7(3):1247–1250, 2014.

[14] Yifei Dai, Yuchen Zhang, and Yulong Zhu. Long-term stock market volatility and the influence of terrorist attacks in europe. In *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*, pages 1–6. IEEE, 2017.

[15] Robert Engle. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics*, 20(3):339–350, 2002.

[16] Robert F Engle. Autoregressive conditional heteroskedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 50(4):987–1007, 1982.

[17] Robert F Engle. Garch 101: The use of arch/garch models in applied econometrics. *Journal of Economic Perspectives*, 15(4):157–168, 2001.

[18] Eugene F Fama. Efficient capital markets: A review of theory and empirical work. *The journal of Finance*, 25(2):383–417, 1970.

[19] Thomas Fischer and Christopher Krauss. Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, 270(2):654–669, 2018.

[20] Christian Francq and Jean-Michel Zakoian. Garch models: Structure, statistical inference and financial applications. *Journal of Financial Econometrics*, 9(4):505–546, 2011.

[21] Felix A Gers, Jürgen Schmidhuber, and Fred Cummins. Learning to forget: Continual prediction with lstm. In *9th International Conference on Artificial Neural Networks: ICANN '99*, volume 2, pages 850–855. IEEE, 2000.

[22] Lawrence R Glosten, Ravi Jagannathan, and David E Runkle. On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance*, 48(5):1779–1801, 1993.

[23] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio. *Deep Learning*. MIT Press, 2016.

[24] Alex Graves, Abdel-rahman Mohamed, and Geoffrey Hinton. *Speech recognition with deep recurrent neural networks*. IEEE, 2013.

[25] Peter R. Hansen and Asger Lunde. A forecast comparison of volatility models: Does anything beat a garch(1,1)? *Journal of Applied Econometrics*, 20(7):873–889, 2005.

[26] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

[27] Abdullah Karasan and Esma Gaygısız. Volatility prediction and risk management: An svr-garch approach. *Journal of Financial Econometrics*, 2023.

[28] Colm Kearney and Sha Liu. On the importance of combining machine learning with econometric models. *Finance Research Letters*, 31:19–24, 2019.

[29] K. Kim and H. White. On more robust estimation of skewness and kurtosis: Simulation and application to the s&p500 index. *Finance Research Letters*, 2003.

[30] Ron Kohavi. A study of cross-validation and bootstrap for accuracy estimation and model selection. *Ijcai*, 14(2):1137–1145, 1995.

[31] Christopher Krauss, Xuan Anh Do, and Nicolas Huck. Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the s&p 500. *European Journal of Operational Research*, 259(2):689–702, 2017.

[32] Werner Kristjanpoller and Marcel C Minutolo. Gold price volatility: A forecasting approach using the artificial neural network-garch model. *Expert Systems with Applications*, 41(17):7605–7614, 2014.

[33] Grégoire Montavon, Wojciech Samek, and Klaus-Robert Müller. Methods for interpreting and understanding deep neural networks. *Digital Signal Processing*, 73:1–15, 2018.

[34] John J. Murphy. *Technical Analysis of the Financial Markets: A Comprehensive Guide to Trading Methods and Applications*. New York Institute of Finance, 1999.

[35] Daniel B Nelson. Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, pages 347–370, 1991.

[36] David M Q Nelson, Adriano C M Pereira, and Renato A de Oliveira. An empirical analysis of the applicability of deep learning for stock market prediction. *Scientific reports*, 7(1):1–18, 2017.

[37] Andrew Y. Ng. Feature selection, l1 vs. l2 regularization, and rotational invariance. In *Proceedings of the Twenty-first International Conference on Machine learning*, 2004.

[38] Ser-Huang Poon. Forecasting volatility in financial markets: A review. *Journal of Economic Literature*, 41(2):478–539, 2003.

[39] Haşim Sak, Andrew Senior, and Françoise Beaufays. Long short-term memory recurrent neural network architectures for large scale acoustic modeling. *arXiv preprint arXiv:1402.1128*, 2014.

[40] Jürgen Schmidhuber. Deep learning in neural networks: An overview. *Neural Networks*, 61:85–117, 2015.

[41] Mike Schuster and Kuldip K. Paliwal. Bidirectional recurrent neural networks. *IEEE Transactions on Signal Processing*, 45(11):2673–2681, 1997.

[42] Robert J Shiller. *Irrational exuberance*. Princeton university press, 2015.

[43] Jasper Snoek, Hugo Larochelle, and Ryan Prescott Adams. Practical bayesian optimization of machine learning algorithms. In *Advances in Neural Information Processing Systems*, 2012.

[44] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: a simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, 15(1):1929–1958, 2014.

[45] Yao-Liang Tsai. Hybrid lstm and encoder-decoder architecture for translation of non-sequential data to sequential data. *arXiv preprint arXiv:1811.04319*, 2018.

[46] Avraam Tsantekidis, Nikolaos Passalis, Anastasios Tefas, Juho Kanniainen, Moncef Gabbouj, and Alexandros Iosifidis. Forecasting stock prices from the limit order book using convolutional neural networks. *2017 IEEE 19th Conference on Business Informatics (CBI)*, 1:7–12, 2017.

[47] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.

[48] Cort J Willmott and Kenji Matsuura. Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance. *Climate Research*, 30(1):79–82, 2005.

[49] Jean-Michel Zakoian. Threshold heteroskedastic models. *Journal of Economic Dynamics and Control*, 18(5):931–955, 1994.