Blending Econometric and Deep Learning Approaches for Enhanced Volatility Forecasting of the KSE-100 Index

Ubaida Fatima, PhD
NED University of Engineering and Technology
Karachi, Sindh,
Pakistan

Rimsha Zafar
NED University of Engineering and Technology
Karachi, Sindh,
Pakistan

Syeda Arsala Shah
NED University of Engineering and Technology
Karachi, Sindh,
Pakistan

Abdus Samad
NED University of Engineering and Technology
Karachi, Sindh,
Pakistan

ABSTRACT
The objective of this study is to develop a hybrid forecasting model that combines statistical and machine learning techniques to predict stock market volatility in Pakistan. The dataset used spans from January 2019 to December 2023, and the model's accuracy is evaluated using Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). Traditional forecasting models often struggle to account for market uncertainty due to external economic factors, such as IMF policies and currency fluctuations. By utilising non-linear techniques, including Autoregressive Integrated Moving Average (ARIMA), Long Short-Term Memory (LSTM), Linear Regression, and Generalised Autoregressive Conditional Heteroscedasticity (GARCH), this study aims to enhance volatility predictions, thereby enabling quicker and more informed investment decisions.

Keywords
Long Short-Term Memory (LSTM), Generalised Autoregressive Conditional Heteroscedasticity (GARCH), Autoregressive Integrated Moving Average (ARIMA), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE).

1. INTRODUCTION
Stock market return projections present significant challenges for investors and stakeholders due to the inherent complexities and volatility of financial markets. The unpredictable nature of stock prices makes accurate forecasting a difficult yet essential task in financial research. Over the years, researchers have continuously explored and refined predictive techniques to enhance market analysis. Understanding stock market movements is crucial, as they play a pivotal role in economic expansion and investment decision-making. Both academic and technical studies aim to identify the most effective forecasting approaches to improve market predictability (Contreras et al., 2003)[32]. Theoretical research suggests that stock markets contribute to long-term economic growth through consumption and investment channels. Forecasting, as a dynamic and continuous process (Golden et al., 1994)[33], relies on historical data and real-time market trends to make informed assumptions about future stock movements. However, Cao et al. (2019)[24] emphasise that "the reliability and accuracy of forecasting models remain a challenge." A well-designed predictive model can help investors mitigate uncertainty, reduce risk, and make more informed decisions. In addition to benefiting investors, effective forecasting enhances the confidence of regulators and policymakers by providing insights into market trends and economic stability. Stock market forecasting methods are broadly categorised into time series analysis, machine learning models, and statistical techniques. Among these, time series analysis is widely regarded as one of the most effective techniques. While many time series forecasting strategies exist, no single approach can accurately predict stock market returns in all scenarios. Researchers have long debated the best forecasting methodology, as no universal agreement exists. Time series models analyse historical stock price data under the assumption that past patterns influence future trends. Key forecasting techniques include ARIMA (Autoregressive Integrated Moving Average), LSTM (Long Short-Term Memory), Linear Regression, and GARCH (Generalized Autoregressive Conditional Heteroscedasticity). Stock market forecasting is generally classified into technical and fundamental analysis. Fundamental analysis evaluates a stock's intrinsic value based on financial performance and industry trends, whereas technical analysis examines past market data to predict future price movements (Levy, 1966). These methods have guided investment decisions for decades. However, the Efficient Market Hypothesis (EMH)(Fama, 1970) challenges the effectiveness of past price data in predicting future movements, suggesting that stock prices follow a random walk and fully reflect available information. Empirical studies have provided mixed findings—some support the random walk theory (Konak&Seker, 2014; Tong, 2012), while others indicate that stock prices can be forecasted to some extent (Darrat&Zhong, 2000; Al-Tabbakh et al., 2018; Lo &MacKinlay, 1989; Owido et al., 2013; Radikoko, 2014). Given the rapid fluctuations in stock markets, moving averages are often used as a tool for trend analysis. Predicting market movements remains challenging due to independent variables such as macroeconomic conditions, geopolitical events, and investor sentiment. Nonetheless, accurate forecasting methods help investors mitigate risk and improve decision-making. A strong economy enhances investor confidence, often leading to increased stock prices and greater participation in equity markets. Analysts continually refine prediction techniques, seeking to maximise profits while minimising forecasting errors. Wang et al. (2018) suggest that market prediction should

be approached from both statistical and artificial intelligence perspectives. Improving forecasting models directly benefits investors by enhancing decision-making and risk management. Stock markets facilitate capital flow between savers and investors, with returns largely dependent on the predictability of price movements. The core challenge in stock market forecasting lies in identifying key influencing factors that drive stock prices and overall profitability. Research consistently highlights a positive correlation between financial markets and economic growth, reinforcing the significance of accurate market predictions. Given the volatile and uncertain nature of financial markets, forecasting stock returns remains one of the most complex challenges in investment strategy. Forecasting models have been widely applied in both developed and emerging markets. In South Asia, however, limited research exists on the application of ARIMA models, particularly in Pakistan, India, Bangladesh, and Sri Lanka. As emerging economies present high volatility and potential for strong returns, further investigation into these markets could enhance investor interest and generate improved investment strategies. By leveraging advanced forecasting techniques, financial analysts and investors can navigate market uncertainties, optimise portfolio strategies, and make more data-driven decisions.

In the dynamic realm of financial markets, volatility risk—a measure of the variability and unpredictability of asset prices—stands as a critical metric for assessing market uncertainty. Accurate forecasting of volatility is indispensable for investors, portfolio managers, and financial institutions, as it underpins risk management strategies, derivatives pricing, and the optimization of asset allocation. Traditional and modern quantitative approaches have sought to predict volatility, yet the inherent complexity of financial time series, marked by nonlinear patterns, sudden regime shifts, and clustering phenomena (e.g., periods of high volatility followed by similar periods), poses persistent challenges. This study explores the efficacy of four distinct methodologies—Long Short-Term Memory (LSTM) networks, Autoregressive Integrated Moving Average (ARIMA), Generalized volatility, aiming to illuminate their relative strengths and limitations.

GARCH models, a cornerstone of financial econometrics, explicitly address volatility clustering and time-varying variance, making them a natural choice for capturing leptokurtic asset return distributions. ARIMA, a flexible linear model adept at handling non-stationary data, provides a benchmark for traditional time series forecasting. In contrast, LSTM networks, a subclass of recurrent neural networks (RNNs), offer the capacity to model complex temporal dependencies and nonlinear interactions through their gated memory mechanisms, potentially uncovering latent patterns in vast datasets. Linear regression, while simpler, serves as a baseline to evaluate whether sophisticated models justify their complexity. By comparing these approaches—spanning classical econometrics, machine learning, and statistical regression—this study assesses their predictive accuracy, robustness to market noise, and adaptability to evolving market conditions. The findings aim to guide practitioners in selecting appropriate tools for volatility risk forecasting, balancing computational demands, interpretability, and performance in an ever-changing financial landscape.

1.1 Problem Identification

The unpredictable nature of stock markets, influenced by social, political, and economic events, makes accurate predictions increasingly challenging. Recent studies on forecasting stock returns have revealed that traditional theories are being improved upon, and new prediction techniques are emerging. Accurate predictions can lead to high returns and profit maximization. However, many investors primarily rely on fundamental approaches that focus on external factors such as company profiles, market conditions, and textual information, which often results in lower profitability.

Over the past 20 years, artificial intelligence has gained popularity among researchers for stock market prediction due to its unique features and dynamic capabilities. The field is progressively advancing towards more accurate and efficient AI-driven models. Initially, simpler linear models, such as autoregressive models and moving average models, were used. However, further research has shown that these original prediction models are limited by the significant noise present in stock data and the uncertainty of various influencing factors. Therefore, This study aims aim to provide traders with a more accurate and efficient model using Artificial Neural Networks compared to earlier models.

1.2 Background and Need

Research on stock markets, especially in emerging economies, is vital due to their high volatility and potential for significant returns, which attract large investors (Akhtar and Khan, 2016). This study aims to enhance the understanding of emerging markets and assist investors in making informed investment decisions based on market efficiency. Specifically, it investigates the efficiency of the Pakistani stock market, which has been found to exhibit low efficiency according to the Efficient Market Hypothesis (EMH), demonstrating randomness and a random walk behavior. Various econometric tools have been employed to test the weak form of market efficiency. As stock trading evolves, investors increasingly rely on Intelligent Trading Systems for price predictions, moving beyond traditional fundamental analysis. This approach enables more informed decisions, such as buying low and selling high.

This project leverages ARIMA, LSTM, Linear Regression, and GARCH models to forecast stock volatility. These models combine technical analysis, statistical methods, and machine learning to create an artificial trading system that refines stock selection and trading strategies. Evaluating these models through strategic indicators demonstrated that both the individual Linear Regression model and the Linear Regression model combined with GARCH provide accurate forecasts, with both models outperforming others in their forecasting capabilities.

1.3Aims and Objectives

The primary objective of this study is to compare the predictive accuracy of ARIMA, GARCH, and deep learning models (e.g., LSTM, RNN) and Linear Regression in forecasting stock market volatility. Specific objectives include:

1. Evaluate model performance using key metrics such as Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) to determine their effectiveness in capturing volatility effects.
2. Analyze model robustness under different market conditions, specifically comparing performance during high-volatility and low-volatility periods.
3. Investigate the potential of deep learning models in outperforming traditional methods for forecasting long-term trends and non-linear market patterns.

2. LITERATURE REVIEW

This comparative analysis compares traditional and modern deep learning approaches to forecast stock market volatility, particularly in developing markets like Pakistan. The models used in this study include ARIMA, GARCH, Linear Regression, and LSTM to assess the effectiveness of these models in predicting stock market behavior. Market volatility, a quantifiable measure of the variance of returns on a security or market index, is critical to derivative pricing, hedging strategy development, risk management structures, and optimal asset allocation models. Volatility is imperative to determine derivative prices, create hedging strategy, risk management structures, and optimal asset allocation models. Volatility clustering, the observation that episodes of high volatility are followed by additional periods of high volatility, underscores the need for effective forecasting methods in financial markets.

Riz Gunaryati, Achmad Benny Mutiara, Sulistyo Puspitodjat (2025) [2] propose a hybrid ARIMA-LSTM model with Lowes's linear regression to improve Sharia stock price forecasting. Using closing stock prices of four Indonesian Islamic banks, the model outperforms individual ARIMA and LSTM models, achieving higher accuracy (97.6%–99.5%) by effectively reducing noise and capturing complex nonlinear patterns.

ZedaXu, John Liechty, Sebastian Benthall, Nicholas Skar-Gislinge, Christopher McComb (2024) [5] propose the GARCH-Informed Neural Network (GINN), a hybrid deep learning model integrating GARCH and LSTM to enhance stock market volatility prediction. Inspired by Physics-Informed Neural Networks (PINN), GINN leverages GARCH's statistical structure with LSTM's flexibility, outperforming traditional models in forecasting accuracy. Empirical results show superior out-of-sample performance based on R^2, MSE, and MAE, highlighting GINN's effectiveness in capturing market volatility dynamics.

LenyYuliyani (2024) [3] examines stock market volatility forecasting using linear regression on (2024) analyses the top 10 issuers in the LQ45 Index. By comparing predicted and actual closing prices, the study determines optimal buy/sell decisions. Prediction accuracy is assessed using Root Mean Square Error (RMSE). Results indicate that stock price growth is not consistently upward, highlighting the model's importance for investor decision-making. The findings confirm that linear regression provides reliable forecasts for the Indonesian stock market.

Lihki Rubio, Adriana Palacio Pinedo, Adriana MejíaCastaño& Filipe Ramos (2023) [14] studied and found that hybrid models, combining ARIMA and GARCH, outperform individual models in forecasting volatility. Using the wavelet transform to decompose time series into low and high-frequency signals improves predictions. ARIMA models are suitable for forecasting low-frequency signals, while GARCH models are better for high-frequency signals. GARCH models capture strong nonlinear patterns in high-frequency signals.

Giovanni Campisi, Silvia Muzzioli, and Bernard De Baets (2023) [8] explore the predictive power of volatility indices in forecasting stock market direction using machine learning techniques. Analysing S&P 500 returns and volatility indices (2011–2022), the research evaluates models based on accuracy, ROC curve, and F-measure. Findings reveal that machine learning models outperform traditional regression models, with Random Forest and Bagging achieving the highest accuracy.

Hanna Azhar (2022) [15] explores stock return forecasting in Pakistan using ARIMA modelling to highlight the effectiveness in predicting market trends and volatility. It underscores the importance of a statistical model to improve forecasting accuracy in emerging markets.

Timothy A. Smith#1, with Alex Caligiuri& J Rhet Montana (2018) [28], Regression models are a powerful tool for estimating asset values by analyzing their characteristics and returns concerning overall market performance. This study builds upon previous research, which developed a regression model to predict the S&P 500 based on macroeconomic indicators. It enhances the original model by incorporating updated data and redefining the measurement of volatility. The revised model is used to estimate market volatility and is evaluated against the S&P 500's implied volatility through a simulation that employs the Black-Scholes framework to forecast the index's value one year ahead. While no model can capture market volatility with complete precision, the newly defined measures of volatility demonstrated superior performance compared to the traditionally used implied volatility. This offers a more refined approach to estimating volatility.

2.1 Research Gap

The aim of this study is to identify the best model for Pakistani investors to guide them in decision-making regarding risk mitigation and portfolio management. With the findings, more precise emerging market financial forecast tools, as the case for Pakistan, would be developed. Four models namely ARIMA, LSTM, Linear Regression, and GARCH, are employed here to compare their forecasting abilities. LSTM performs exceptionally well at recognizing intricate, nonlinear patterns by way of AI-facilitated deep learning and is well-apt to detect volatile price swings. ARIMA detects trend and seasonality within time series and Linear Regression is a simple yet effective trend forecast method. GARCH is focused on estimating volatility in markets that is used in risk determination. The research seeks to bridge a literature gap on emerging markets' stock prediction models by meeting the demand for improved financial forecasting models.

3. DATASET AND METHODOLOGY

The analysis began by gathering reliable data to ensure accurate results, focusing on the KSE-100 Index. In this study four forecasting models were applied: ARIMA, GARCH, LSTM, and Linear Regression, alongside two hybrid models (ARIMA-LSTM and Linear Regression-GARCH) to assess stock market volatility. The dataset included daily historical closing prices from 2019 to 2023(Figure 1).

KSE-100 Index Market Price								
Date	Price	Open	High	Low	Price Return	Open Return	Individual Volatiliy	Log Return
12/29/2023	62,451.04	62,225.33	62,644.38	61,807.54	0.643%	1.792%	0.926%	-0.641%
12/28/2023	62,052.24	61,130.13	62,750.78	61,092.27	1.953%	3.439%	0.642%	-1.934%
12/27/2023	60,863.62	59,097.62	61,009.88	58,758.48	2.861%	-4.072%	4.927%	-2.820%
12/26/2023	59,170.98	61,606.46	61,634.55	59,026.81	-4.107%	-1.960%	1.789%	4.194%
12/22/2023	61,705.09	62,838.18	62,995.22	61,569.14	-1.577%	0.153%	1.393%	1.589%
12/21/2023	62,693.57	62,742.26	62,850.10	61,750.97	0.393%	-0.191%	0.711%	-0.392%
12/20/2023	62,448.01	62,862.54	63,261.05	61,082.50	-0.613%	-3.473%	2.139%	0.615%
12/19/2023	62,833.03	65,124.30	65,132.03	62,360.78	-3.637%	-2.107%	1.582%	3.705%
12/18/2023	65,204.68	66,525.70	66,586.62	65,064.94	-1.399%	1.142%	1.724%	1.409%
12/15/2023	66,130.02	65,774.74	66,346.77	65,559.49	1.039%	0.691%	0.550%	-1.033%
12/14/2023	65,450.19	65,323.06	65,622.56	64,437.70	0.260%	-2.159%	1.405%	-0.260%
12/13/2023	65,280.16	66,764.30	67,093.96	64,427.39	-1.726%	0.973%	1.665%	1.741%
12/12/2023	66,426.78	66,121.14	66,604.04	66,121.14	0.628%	-0.062%	0.670%	-0.626%
12/11/2023	66,012.33	66,162.29	66,564.04	65,129.00	-0.319%	1.702%	1.871%	0.320%
12/8/2023	66,223.63	65,055.17	66,273.73	65,055.17	2.326%	1.397%	0.604%	-2.300%
12/7/2023	64,718.08	64,159.02	64,958.10	63,853.05	1.473%	1.509%	0.199%	-1.462%

Figure 1 Historical Data of KSE-100 Index

To test the generalizability of the models, the same analysis was done on the KMI-30 Index using the same methodologies. Although the models confirmed applicability beyond the KSE-100 Index, limitations were faced with forecasting crypto currency volatility due to data availability, as it was only accessible monthly, requiring adjustments for effective implementation.

For a thorough analysis of the KSE-100 index, data from various financial research publications were used.

Table 1 Source of Data

Data Taken From	Labels	No. Of Data	Link
KSE-100	Open, Close	5 Columns, 1241 Rows	https://www.investing.com/ [2]
KMI-30	Open, Close	5 Columns, 1241 Rows	https://www.investing.com/ [2]

3.1 Models

Employed four models for stock volatility risk prediction: Autoregressive Integrated Moving Average (ARIMA), Generalised Autoregressive Conditional Heteroscedasticity (GARCH), Long Short-Term Memory (LSTM), and Linear Regression. Following a structured, step-by-step methodology, this study evaluated each model's performance using key metrics such as Root Mean Square Error (RMSE) and Mean Absolute Error (MAE). In this study it was discovered that both individual models, such as Linear Regression, and hybrid models, like Linear Regression combined with GARCH, deliver reliable forecasts with high accuracy.

3.1.1 Auto-Regressive Integrated Moving Average (ARIMA)

The ARIMA (Auto-Regressive Integrated Moving Average) model is a popular method for forecasting time series data, particularly in financial contexts like stock prices. ARIMA is effective because it can manage a variety of data structures, including those that exhibit trends, seasonality, and noise. The key principle behind ARIMA is to identify the relationships between current observations and their historical lags to make predictions about future values.

3.1.2 Generalised Autoregressive Conditional Heteroscedasticity

Bollerslev (1986) introduced the GARCH model, which is based on the assumption that forecasts of time-varying variance rely on the lagged variance of the asset. An unexpected increase or decrease in returns at time t will lead to a rise in the expected volatility in the following period.

3.1.3 Long-Short Term Memory

LSTM networks are specifically designed to retain long-term dependencies in financial time-series data. Unlike traditional RNNS, which can suffer from the vanishing gradient problem, LSTMS utilise memory cells and gate mechanisms to selectively store or forget information. This capability allows them to analyse stock price trends over weeks, months, or even years, thereby improving the accuracy of market forecasts. With their ability to process sequential data and recognize long-term trends, LSTM models have become powerful tools in financial forecasting, enabling traders and analysts to make more informed, data-driven investment decisions.

3.1.4 Linear Regression

Linear regression is one of the most fundamental and widely used techniques in supervised learning. It helps to understand and model the relationship between a dependent variable (the target) and one or more independent variables (the predictors). The primary goal of linear regression is to establish a linear relationship between these variables, which enables predictions based on the given dataset. This technique is extensively used across various fields, including business analytics, medical research, agricultural studies, and financial forecasting.

Outcome Projections

Forecasting ARIMA on KSE-100 Index Dataset

The volatility of the KSE-100 index wavered between 2020 and 2019, with a huge spike in 2020 during the COVID-19 pandemic. This panic generated immense uncertainty and abrupt swings in stock prices. In 2021, volatility reduced because of recovery efforts, vaccine distribution, business reopening, an increase in oil prices, and growth in the tech sector. Presently, the stock market is on the rise due to optimism, political stability, economic reforms, confidence of investors, moderate oil prices, and growth in the IT and remittance sectors. However, future market performance is based on real-time happenings and world economic situations (Figure 2)

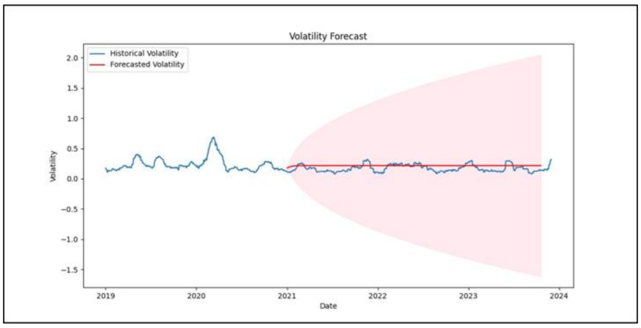

Figure 2 Volatility Trends – Historical Data and ARIMA Forecast KSE-100 Index

Forecasting GARCH on KSE-100 Index Dataset

The KSE-100 index's past and forecasted volatility is depicted in this graph. The historical volatility from 2019 to mid-2023 is represented by the blue line. It shows variations with notable spikes, especially around 2020, which may suggest times of market upheaval. The GARCH model predicts volatility from July 2023 to June 2024, as seen by the red dashed line. Before stabilising, the forecast displays a slight rising tendency after beginning at a comparatively lower level. This implies that although an initial increase in volatility is anticipated, it may eventually level off as a result of possible market corrections or less uncertainty (Figure 3).

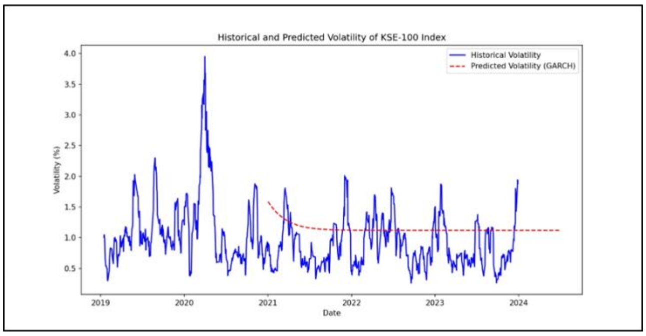

Figure 3 Volatility Trends - Historical Data and GARCH Forecast KSE-100 Index

Forecasting LSTM on KSE-100 Index Dataset

The LSTM model predicts the test dataset's volatility, and the new column has expected values. A blue line displays the true volatility, and a red dashed line indicates the LSTM model's predicted volatility (Figure 4).

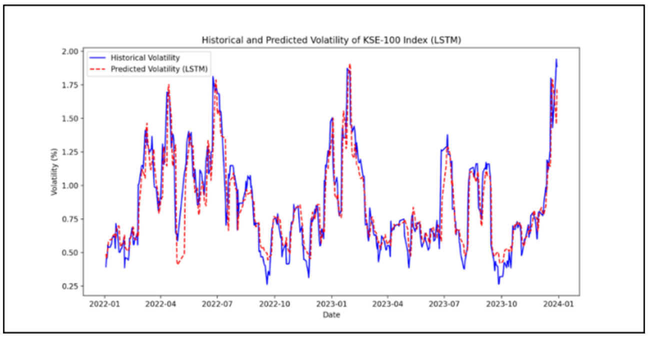

Figure 4 Volatility Trends - Historical Data and LSTM Forecast KSE-100 Index

Forecasting Linear Regression on KSE-100 Index Dataset

Out-of-sample predictions are generated for the test set using the trained model. A line plot is used to compare the anticipated volatility to the historical volatility. The real historical volatility is shown by the blue line. The expected volatility using linear regression is shown by the red dashed line (Figure 5).

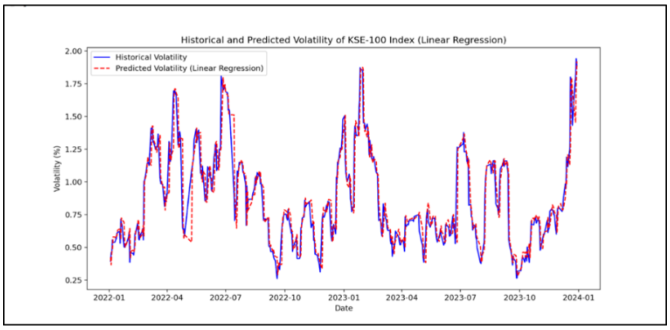

Figure 5 Volatility Trends - Historical Data and Linear Regression Forecast KSE-100 Index

Individual Error Matrix:

Table 2 Error Table

Models	RMSE	MAE
LSTM	0.001637	0.001178
ARIMA	0.052549	0.044685
GARCH	0.928456	0.853722
Linear Regression	0.001277	0.000765

Interpretation of Individual Error Matrix

The table highlights the comparative performance of different models used for volatility prediction. Traditional statistical models such as ARIMA (RMSE = 0.052, MAE = 0.044) and GARCH (RMSE = 0.92, MAE = 0.85) exhibited relatively poor predictive accuracy, as indicated by their high error metrics. These results suggest that both models were ineffective in capturing the underlying volatility patterns, reflecting a weak fit to the data. In contrast, machine learning-based approaches, particularly the LSTM model (RMSE = 0.0016, MAE = 0.0011) and the Linear Regression model (RMSE = 0.0012, MAE = 0.00076), demonstrated significantly better performance. These models achieved much lower error values, indicating a closer approximation to the actual volatility.

Among all the models, the Linear Regression model outperformed the others by consistently delivering the most accurate predictions. Its superior performance is evident from the lowest RMSE and MAE values, which underscore its ability to minimize prediction error effectively. Despite the advanced architecture of the LSTM model, Linear Regression surpassed it in both metrics, suggesting that for the given dataset and problem scope, simpler models with well-tuned parameters can sometimes outperform more complex deep learning approaches. This highlights the importance of model selection based on empirical performance rather than complexity alone. Overall, Linear Regression emerged as the most reliable and effective model for volatility prediction in this study.

4. HYBRID MODEL

Hybrid Model LSTM with ARIMA

Prediction is done using the Hybrid Model of ARIMA with LSTM. The orange line shows the prediction while the blue line shows the actual volatility. A line plot is used to compare the results (Figure 6).

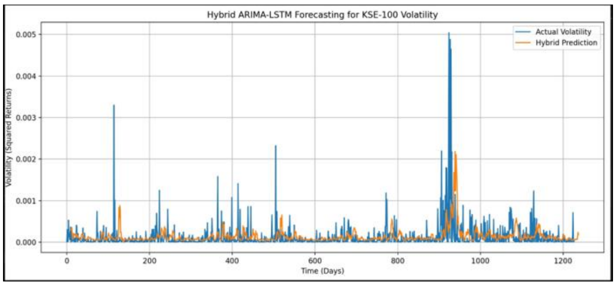

Figure 6 Volatility Trends - Actual Volatility and Hybrid Forecast KSE-100 Index

Hybrid Model Linear Regression with GARCH

Prediction is done using the Hybrid Model of Linear Regression with GARCH. The orange line shows the prediction while the blue line shows the actual volatility. A line plot is used to compare the results (Figure 7).

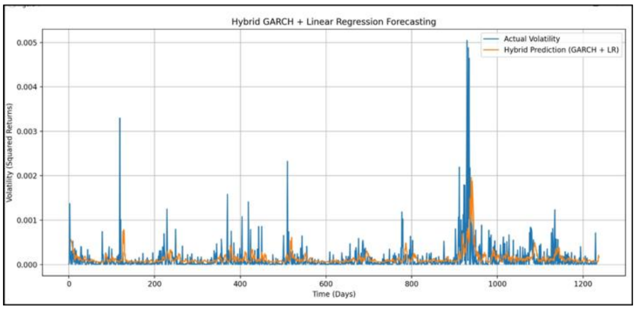

Figure 7 Volatility Trends - Actual Volatility and Hybrid Forecast KSE-100 Index

Hybrid Model Error Matrix

Table 3 Hybrid Model Error Matrix

Models	RMSE	MAE
ARIMA With LSTM	0.01249	0.000156
Linear Regression With GARCH	0.01191	0.000141

Interpretation of Hybrid Model Error Matrix

The table presents the performance metrics of two hybrid forecasting models—ARIMA integrated with LSTM, and Linear Regression combined with GARCH—both applied for predicting market volatility. These hybrid models aim to leverage the strengths of their components: ARIMA captures linear temporal structures, LSTM handles complex nonlinear patterns, while Linear Regression and GARCH contribute simplicity and robustness in modelling conditional heteroskedasticity. Both models demonstrated strong predictive capabilities, as indicated by their low Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) values, reflecting their effectiveness in tracking actual volatility trends.

Among the two, the Linear Regression with GARCH model achieved slightly better performance, with an RMSE of 0.01191 and MAE of 0.00014, compared to the ARIMA with LSTM model, which recorded an RMSE of 0.01249 and MAE of 0.000156. These metrics suggest that the Linear Regression with GARCH model produced forecasts more closely aligned with actual volatility and was marginally more successful in capturing variance in the data. Although the performance gap between the two models is relatively small, the consistently lower error values associated with the Linear Regression with GARCH model highlight its greater precision and reliability.

Therefore, based on empirical results, the Linear Regression with GARCH hybrid model can be considered the superior choice for volatility forecasting in this context. It balances simplicity, interpretability, and predictive power, making it a more practical and effective tool for real-world financial forecasting applications.

Model Generalization to KMI-30 Index

ARIMA *Model Generalization to KMI-30 Index*

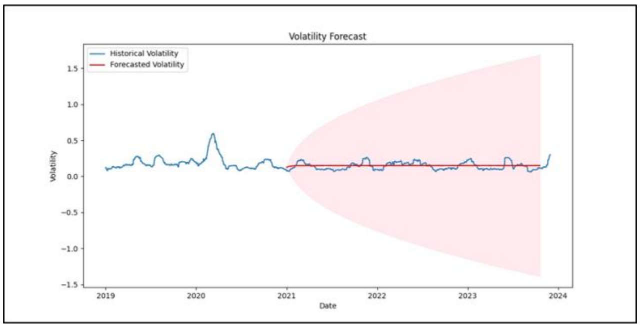

Figure 8 Volatility Trends - Historical Data and ARIMA Forecast KMI-30 Index

GARCH Model Generalization to KMI-30 Index

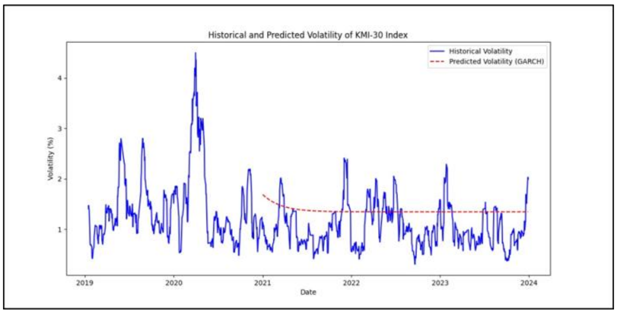

Figure 9 Volatility Trends - Historical Data and GARCH Forecast KMI-30 Index

Linear Regression Model Generalization to KMI-30 Index

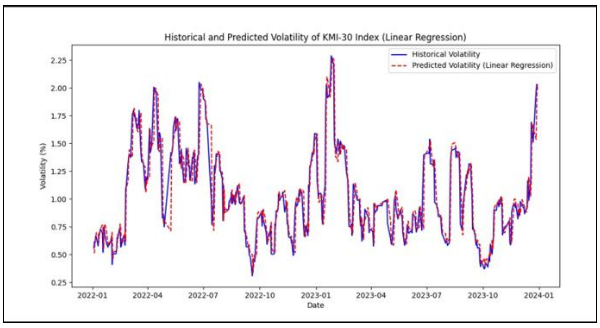

Figure 10 Volatility Trends - Historical Data and Linear Regression Forecast KMI-30 Index

LSTM Model Generalization to KMI-30 Index

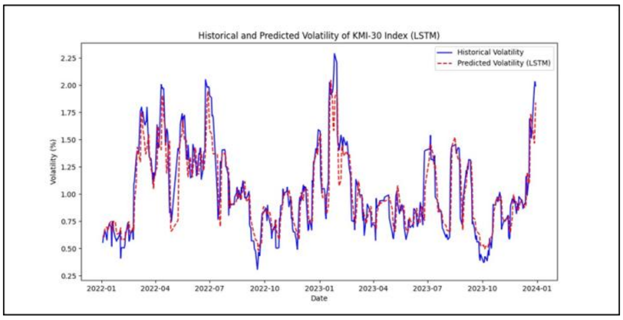

Figure 11 Volatility Trends - Historical Data and LSTM Forecast KMI-30 Index

Error Matrix for Multi-Index Validation of KMI-30 Index

Table 4 Error Matrix KMI-30 Index

Models	RMSE	MAE
LSTM	0.002427	0.002063
ARIMA	0.073007	0.064221
GARCH	1.098857	1.019484
Linear Regression	0.001509	0.000926

Interpretation of Error Matrix for Multi-Index Validation of KMI-30

The table represents the generalized error metrics of four models applied to the KMI-30 Index, highlighting their forecasting accuracy across multiple evaluation indicators. The ARIMA model (RMSE = 0.073007, MAE = 0.064221) and the GARCH model (RMSE = 1.0988, MAE = 1.019) performed poorly. These values reflect that ARIMA model provides the better result than the GARCH model. In contrast, both the LSTM model (RMSE = 0.0024, MAE = 0.0021) and the Linear Regression model (RMSE = 0.0015, MAE = 0.0009) have achieved higher predictive accuracy and reliability. Among all the models tested on the KMI-30 Index, Linear Regression outperformed the others, offering the lowest RMSE and MAE values . This suggests that Linear Regression is the most effective model for volatility forecasting, providing consistent and precise results across different market indices.

Generalization of Hybrid Models to KMI-30 Index

Hybrid Model ARIMA with LSTM

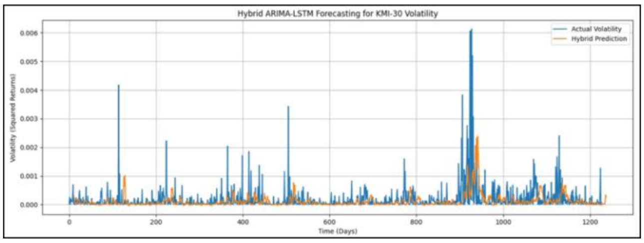

Figure 12 Volatility Trends - Actual Volatlity and Hybrid Forecast KMI-30 Index

Hybrid Model Linear Regression with GARCH

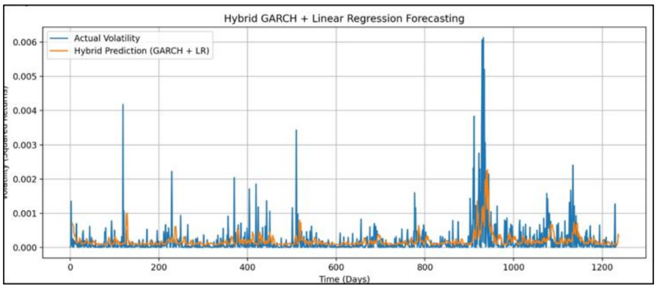

Figure 13 Volatility Trends - Actual Volatility and Hybrid Forecast KMI-30 Index

Hybrid Error Matrix for Multi-Index Validation of KMI-30

Table 5 Error Table for KMI-30 Index

Models	RMSE	MAE
ARIMA With LSTM	0.014434	0.0020835
Linear Regression With GARCH	0.014267	0.00020357

Interpretation Of Hybrid Error Matrix For Multi-Index Validation Of KMI-30

The results demonstrate that both hybrid models showed reasonable forecasting accuracy for the KMI-30 Index; however, the Linear Regression with GARCH model shows more precise and reliable results. Its lower (MAE) and (RMSE) suggest greater consistency in predictions, enhancing its effectiveness in capturing volatility dynamics in alternative indices. These findings support the applicability of the proposed models beyond KSE100, confirming their effectiveness in various market conditions.

5. LIMITATIONS OF THE WORK

This study is limited to daily prices of the stock's closing for the accuracy of the model for forecasting.

- Macroeconomic variables are not included; they may affect the forecasting of the model.
- LSTM models require significant computational resources and longer training times, which may not be feasible in all financial environments.
- Crypto currencies were excluded due to monthly data

limitations, reducing the scope for generalizing the models beyond traditional equities.

6. CONCLUSION

The study evaluated the forecasting performance of numerous models including ARIMA, GARCH, LSTM, and Linear Regression along with hybrid approaches like (ARIMA with LSTM and Linear Regression with GARCH) for predicting the volatility risk of the KSE-100 and KMI-30 indices in Pakistan. The results showed that traditional models like ARIMA and GARCH were less effective at capturing sudden market shocks and complex non-linear behavior of financial data. Alternatively, machine learning techniques and statistical approach, particularly Linear Regression and LSTM, demonstrated superior predictive accuracy, proof by their lower RMSE and MAE values.

Among all the models, Linear Regression was the most effective standalone tool, while the hybrid model like Linear Regression with GARCH outperformed all others in stability and accuracy. This highlights the perks of combining statistical methods with modern machine learning.

These findings are crucial for investors, analysts, and policymakers in emerging markets, as financial volatility is influenced by various political, economic, and global factors. The success of hybrid models offers a roadmap for improving risk management and trading strategies, enabling investors to time trades better, enhance operational efficiency in financial institutions, and help policymakers utilize volatility forecasts as early warnings for financial instability.

7. RECOMMENDATIONS

1. Adopt Hybrid Models: Investors, analysts, and financial institutions should consider using hybrid models like Linear Regression-GARCH for better prediction accuracy under volatile market conditions.
2. Integrate Real-Time Economic Indicators: Future models should adopt macroeconomic indicators namely inflation, interest rates, and political events with aim of enhance responsiveness and predictive power.
3. Use in Risk Management Systems: Financial institutions and portfolio managers can integrate these forecasting models into their risk management systems to anticipate market turbulence and manage exposure effectively.
4. Policy Implications: Policymakers can use volatility forecasts as early warning signals for economic instability, helping them take preventive measures to maintain market confidence.
5. A more thorough assessment of a variety of datasets or market scenarios should be the focus of future research in order to improve the findings' robustness and generalizability. This would make it possible to comprehend the model's performance under various scenarios in more detail.

8. REFERENCES

[1] Santosh Kumar, Surya Kant Pal, Ashok Kumar Pal, Khursheed Alam; Prediction of stock market using machine learning algorithms. AIP Conf. Proc. 30 January 2025; 3191 (1): 040009. https://doi.org/10.1063/5.0247388

[2] Gunaryati, A., Mutiara, A., & Puspitodjati, S. (2025). Enhancing Sharia Stock Price Forecasting using a Hybrid ARIMA-LSTM with Locally Weighted Scatterplot Smoothing Regression Approach. Journal of Applied Data Sciences, 6(1), 683-694. DOI: https://doi.org/10.47738/jads.v6i1.514

[3] Eka, A. P. B., Bakri, A. A., & Yuliyani, L. (2024). Utilizing linear regression to forecast the stock price fluctuations of top-rated companies. Jurnal Info Sains : Informatika Dan Sains, 14(01), 551–559.

[4] Sangeetha, J. M., & Alfia, K. J. (2024). Financial stock market forecast using evaluated linear regression-based machine learning technique. Measurement: Sensors, 31, 100950. https://doi.org/10.1016/j.measen.2023.100950

[5] Zeda Xu, John Liechty, Sebastian Benthall, Nicholas Skar-Gislinge, and Christopher McComb. 2024. GARCH-Informed Neural Networks for Volatility Prediction in Financial Markets. In Proceedings of the 5th ACM International Conference on AI in Finance (ICAIF '24). Association for Computing Machinery, New York, NY, USA, 600– 607. https://doi.org/10.1145/3677052.3698600

[6] Wu, X., Zhao, A., Wang, Y., & Han, Y. (2024). Forecasting Chinese stock market volatility with high-frequency intraday and current return information. Pacific-Basin Finance Journal, 86, 102458. https://doi.org/10.1016/j.pacfin.2024.102458

[7] García-Medina, A., Aguayo-Moreno, E. LSTM–GARCH Hybrid Model for the Prediction of Volatility in Cryptocurrency Portfolios. Comput Econ 63, 1511–1542 (2024). https://doi.org/10.1007/s10614-023-10373-8.

[8] Campisi, G., Muzzioli, S., & De Baets, B. (2024). A comparison of machine learning methods for predicting the direction of the US stock market on the basis of volatility indices. International Journal of Forecasting, 40(3), 869-880. https://doi.org/10.1016/j.ijforecast.2023.07.002.

[9] Xu, K., Wu, Y., Jiang, M., Sun, W., & Yang, Z. (2024). Hybrid LSTM-GARCH Framework for Financial Market Volatility Risk Prediction. https://doi.org/10.5281/zenodo.13643010.

[10] García Medina, Andrés & Moreno, Ester. (2023). LSTM–GARCH Hybrid Model for the Prediction of Volatility in Cryptocurrency Portfolios. Computational Economics. 53 63. 10.1007/s10614-023-10373-8. DOI: 10.1007/s10614-023-10373-8.

[11] Eliasson, E. (2023). Long Horizon Volatility Forecasting Using GARCH-LSTM Hybrid Models: A Comparison Between Volatility Forecasting Methods on the Swedish Stock Market. https://kth.divaportal.org/smash/get/diva2:1849031/FULLTEXT01.pdf

[12] Ghani, U., Zhu, B., Qin, Q., & Ghani, M. (2023). Forecasting US Stock Market Volatility: Evidence from ESG and CPU indices. Finance Research Letters, 59, 104811. https://doi.org/10.1016/j.frl.2023.104811

[13] Xie, Wangsong, Nonlinear Volatility Risk Prediction Algorithm of Financial Data Based on Improved Deep Learning, Discrete Dynamics in Nature and Society, 2022, 3037040, 8 pages, 2022. https://doi.org/10.1155/2022/3037040

[14] Rubio, L., Palacio Pinedo, A., Mejía Castaño, A. et al. Forecasting volatility by using wavelet transform, ARIMA, and GARCH models. Eurasian Econ Rev 13, 803–830 (2023). https://doi.org/10.1007/s40822-023-00243-x

[15] Azhar, H. (2022). Returns Forecasting in Pakistan: A Study by Using ARIMA Modelling (Doctoral dissertation, CAPITAL UNIVERSITY).Kakade, K., Jain, I., & Mishra, A. K. (2022). Value-at-Risk forecasting: A hybrid ensemble learning GARCHLSTM based approach. Resources Policy, 78, 102903. https://doi.org/10.1016/j.resourpol.2022.102903

[16] Jiao, X., Song, Y., Kong, Y., & Tang, X. (2022). Volatility forecasting for crude oil based on text information and a deep learning PSO-LSTM model. Journal of Forecasting, 41(5), 933-944. https://doi.org/10.1002/for.2839

[17] Liu, R., Jiang, Y., & Lin, J. (2021). Forecasting the Volatility of Specific Risk for Stocks with LSTM. Procedia Computer Science, 202, 111-114. https://doi.org/10.1016/j.procs.2022.04.015

[18] Zhang, W., Gong, X., Wang, C., & Ye, X. (2021). Predicting stock market volatility based on textual sentiment: A nonlinear analysis. Journal of Forecasting, 40(8), 14791500. https://doi.org/10.1002/for.2777

[19] Hameed, M. ., Iqbal, K. ., Ghazali, R. ., Jaskani, F. H. ., & Saman, Z. . (2021). Karachi Stock Exchange Price Prediction using Machine Learning Regression Techniques. EAI Endorsed Transactions on Creative Technologies, 8(28), e5. https://doi.org/10.4108/eai.24-8-2021.170753

[20] Muhammad Shahzeb Alia*, A. J. (24 September 2020.). Modelling and Forecasting Volatility in Pakistan Stock. International Journal of Sciences: Basic and Applied Research. 54

[21] Aamir, Muhammad & Ali, Hazrat & Khalil, U. & Bilal, Muhammad. (2020). Comparative Analysis Of Alternative Time Series Models In Forecasting Karachi Stock Exchange. 19. 7981-8002. 10.17051/ilkonline.2020.04.765193.

[22] Vidal, A., & Kristjanpoller, W. (2020). Gold volatility prediction using a CNNLSTM approach. Expert Systems With Applications, 157, 113481. https://doi.org/10.1016/j.eswa.2020.113481

[23] Hu, Y., Ni, J., & Wen, L. (2020). A hybrid deep learning approach by integrating LSTM-ANN networks with GARCH model for copper price volatility prediction. Physica A: Statistical Mechanics and its Applications, 557, 124907. https://doi.org/10.1016/j.physa.2020.124907

[24] (2019), C. e. (20 February 2019). The single-cell transcriptional landscape of mammalian organogenesis.

[25] Asghar, M.Z., Rahman, F., Kundi, F.M. et al. Development of a stock market trend prediction system using multiple regression. Comput Math Organ Theory 25, 271–301 (2019). https://doi.org/10.1007/s10588-019-09292-7

[26] Idrees, S. M., Alam, M. A., & Agarwal, P. (2019). A Prediction Approach for Stock Market Volatility Based on Time Series Data. IEEE Access, 7, 17287–17298. https://doi.org/10.1109/access.2019.2895252

[27] Kim, Ha & Won, Chang. (2018). Forecasting the Volatility of Stock Price Index: A Hybrid Model Integrating LSTM with Multiple GARCH-Type Models. Expert Systems with Applications. 103. 10.1016/j.eswa.2018.03.002. https://doi.org/10.1016/j.eswa.2018.03.002

[28] Timothy A. Smith, with Alex Caligiuri, J Rhet Montana, "Using a Multiple Linear Regression Model to Calculate Stock Market Volatility," International Journal of Mathematics Trends and Technology (IJMTT), vol. 57, no. 4, pp. 220-224, 2018. Crossref, https://doi.org/10.14445/22315373/IJMTT-V57P531

[29] Gharehchopogh, F. S., Bonab, T. H., & Khaze, S. R. (2013). A linear regression approach to prediction of stock market trading volume: a case study. International Journal of Managing Value and Supply Chains, 4(3), 25. chrome-

[30] Situm, Mario, Time Series Volatility Forecasting Using Linear Regression and GARCH (February 7, 2010). Available at SSRN: or http://dx.doi.org/10.2139/ssrn.1549385

[31] Angelidis, T., Benos, A., & Degiannakis, S. (2004). The use of GARCH models in VaR estimation. Statistical Methodology, 1(1-2), 105-128. https://doi.org/10.1016/j.stamet.2004.08.004 55

[32] Contreras, J. (3, AUGUST 2003). ARIMA Models to Predict Next-Day Electricity Prices. Retrieved from Contreras, J. (n.d.). Javier Contreras.

[33] R.L golden, c. (1994, december 1). Observations of cosmic-ray electrons and positrons using an imaging calorimeter.