International Journal of Management, IT & Engineering
Vol. 16 Issue 05, May 2026,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gate as well as in Cabell's Directories of Publishing Opportunities, U.S.A

Accurate Time Series Forecasting of the Saudi Arabian Stock Market Using Hybrid Deep Learning Models

Zamzam Abdulaziz Almushayti*, Ali Mustafa Qamar Khan

Department of Computer Science, College of Computer, Qassim University, Buraydah 52571, Saudi Arabia
*Corresponding Author: E-mail: zamzam.9ah@gmail.com

Abstract:

Forecasting financial markets is one of the hardest problems in applied economics, and the difficulty only multiplies in volatile emerging markets. This paper takes on that challenge by forecasting the Saudi Arabian stock market index (TASI) using ten years of daily data (2014–2024), combining traditional econometric methods with state-of-the-art deep learning techniques.
Five forecasting approaches were tested from pattern-based and volatility models to LSTM, CNN, and a hybrid CNN-LSTM. The hybrid model outperformed all others, cutting errors by over 20% compared to ARIMA while staying resilient through crises like the 2014–15 oil crash and COVID-19. This confirms that combining convolutional and recurrent layers captures both short-term shifts and long-term trends more effectively than either approach alone. Beyond accuracy, this study makes three meaningful contributions: it applies hybrid deep learning to an underexplored market, provides a rare long-term TASI dataset, and delivers a practical forecasting tool aligned with Saudi Vision 2030's financial innovation goals offering real value to investors and decision-makers navigating complex markets.
Keywords:
Stock Market Forecasting; Tadawul All Share Index (TASI); Hybrid Deep Learning; Convolutional Neural Network (CNN); Long Short-Term Memory (LSTM); CNN–LSTM; ARIMA; GARCH; Financial Time Series; Emerging Markets; Volatility Clustering; Saudi Arabia; Vision 2030; Rolling-Window Validation; Log-Returns.

International Journal of Management, IT & Engineering
Vol. 16 Issue 05, May 2026,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gate as well as in Cabell's Directories of Publishing Opportunities, U.S.A

1. Introduction

1.1 Background and Motivation

Financial markets are both a mirror and a motor of economic activity they aggregate dispersed information and turn investor expectations into price movements that shape business investment, consumer confidence, and macroeconomic stability. Forecasting those movements is therefore not just a technical challenge but a matter of economic policy and strategic planning [1]. The task is genuinely hard. Fama's Efficient Market Hypothesis suggests consistent prediction is impossible, yet decades of evidence tell a different story markets exhibit recurring statistical patterns like volatility clustering, heavy tails, and nonlinear dependencies that make improved forecasting achievable with the right models and data.

The Saudi Arabian stock market, represented by the Tadawul All Share Index (TASI), is a particularly compelling case. As the largest exchange in the MENA region, it reflects the Kingdom's ongoing economic transformation. The expansion of non-oil sectors, the opening of capital markets to foreign investors, and the inclusion of Saudi equities in global indices such as MSCI and FTSE have all deepened market participation and liquidity [2]. But this same dynamism has brought higher volatility and greater sensitivity to global shocks making accurate forecasting both harder and more consequential. This research sits at the intersection of three realities: the growing strategic importance of the Saudi market, the rapid evolution of machine learning and deep learning, and the well-documented limitations of traditional econometric models in volatile emerging markets. Forecasting TASI is not just an academic puzzle the results have real consequences for how investors allocate capital, how regulators spot systemic risk, and how policymakers respond to market stress.

Historically, financial forecasting relied on ARIMA [4] and GARCH [5] interpretable and useful, but built on assumptions of linearity, stationarity, and normally distributed residuals that real markets rarely satisfy. Machine learning methods like SVMs and Random Forests improved nonlinear modeling but treated observations as independent, ignoring temporal structure entirely. Deep learning addressed this through RNNs, LSTMs, and CNNs [6-8]. The natural next step was hybrid architectures combining both approaches, with CNN capturing short-lived volatility patterns and LSTM preserving longer-term cyclical behavior together outperforming either approach alone [9], [10]. This shift goes beyond just better algorithms it reflects a fundamentally different way of thinking about how financial markets work and how we should study them. In the Saudi context, applying such models to TASI can illuminate the interplay between energy markets, fiscal policy, and investor behavior, directly serving Vision 2030's goal of building a sophisticated, transparent financial sector.

Most prior research on Saudi stock-market forecasting focuses on short periods or uses small sample sizes that predate significant market reforms [3]. Few studies extend beyond 2020, meaning critical events like COVID-19 and subsequent oil-price volatility remain underexplored. This thesis addresses that gap with daily TASI data spanning 2014–2024 long enough to capture multiple volatility regimes and a hybrid CNN–LSTM model designed to learn from that diversity. In short the Saudi market is complex, existing models struggle with that complexity, and we now have tools sophisticated enough to do better. That is what this research is about.

Vol. 16 Issue 05, May 2026,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gate as well as in Cabell's Directories of Publishing Opportunities, U.S.A

1.2 Problem Statement

Despite decades of research, existing forecasting models still struggle to deliver consistent accuracy across different market conditions. Financial time series especially in emerging markets like Saudi Arabia exhibit nonlinearity, non-stationarity, and structural breaks that violate traditional model assumptions, compounded by external shocks like oil-price swings and global crises that abruptly reshape volatility structures. Most models fall into one of three categories, each with inherent limitations. Statistical models like ARIMA and GARCH handle linear dependencies and conditional variance well but collapse when market regimes shift suddenly. Machine learning models like SVMs and Random Forests capture nonlinear relationships but treat every observation as independent, requiring extensive feature engineering to approximate memory effects. Standalone deep learning models LSTM or CNN alone address some issues but remain individually limited: LSTMs can over-smooth short-term spikes; CNNs cannot account for extended temporal patterns.

For TASI specifically, these weaknesses are amplified. Oil-price dependence introduces a powerful external volatility driver. Policy discontinuities foreign investor access in 2015, MSCI/FTSE inclusion in 2019 create structural breaks that fixed-parameter models simply cannot navigate. Most existing studies also rely on RMSE alone, missing the fuller picture that MAE and MAPE provide, and few use rolling-window validation to test whether results hold across time. The problem, in short: no existing framework adequately captures TASI's nonlinear, multi-scale, and regime-shifting dynamics. A hybrid architecture integrating convolutional pattern recognition with recurrent sequential learning rigorously validated across multiple metrics and volatility periods is the logical response.

1.3 Research Objectives

The overarching aim is to develop and evaluate a hybrid deep-learning framework for forecasting TASI. This is structured around seven specific objectives:

- Collect and preprocess daily TASI closing prices from 2014–2024, including log-return computation, stationarity testing via ADF, and autocorrelation and volatility clustering analysis.
- Establish econometric benchmarks: ARIMA(1,1,1) and GARCH(1,1) as interpretable baselines for comparison.
- Design and implement standalone LSTM and CNN models to assess their individual capabilities.
- Construct a hybrid CNN–LSTM architecture with hyperparameter optimization through cross-validation.
- Evaluate all models using MAE, RMSE, and MAPE alongside rolling-window validation and crisis-period sub-samples.
- Interpret results in light of existing literature, drawing practical implications for investors, regulators, and policymakers.
- Ensure methodological transparency and reproducibility through clear documentation of data-processing and training pipelines.

Together, these objectives are designed to make sure the research is useful on more than one level not just technically sound, but practically meaningful.

2. Literature Evaluation of Study

Forecasting financial markets has been one of the most intensively studied problems in economics and computer science and one of the least solved. Markets are nonlinear, dynamic, and often chaotic, and every modelling innovation from classical econometrics to modern deep learning has been an attempt to capture more of that complexity. Understanding this progression is essential before appreciating why the hybrid CNN–LSTM architecture adopted here is the logical next step for predicting TASI. The story starts in the post-war development of time-series econometrics. Despite the Efficient Market Hypothesis suggesting forecasting was futile, empirical evidence of autocorrelation and volatility clustering pushed researchers to look for better models anyway. The most influential result was Box and Jenkins' ARIMA [4] transparent, data-efficient, and dominant for two decades. It worked well for capturing short-term dependencies in stable markets like the S&P 500, FTSE 100, and NIFTY [11]. But its reliance on linearity and constant variance made it increasingly inadequate as markets grew more turbulent and complex.

To handle changing variance, Engle introduced ARCH in 1982, later generalized by Bollerslev into GARCH (1986) [5]. These models treat conditional variance as a dynamic process driven by past shocks, successfully capturing the volatility clustering that characterizes real asset returns. Research across GCC markets confirmed GARCH's ability to track oil-price-driven volatility spillovers [13]. Still, it assumes symmetry between positive and negative shocks and struggles with the fat tails that real return distributions exhibit [12]. Extensions like EGARCH and GJR-GARCH helped at the margins but stayed within a fixed functional form. By the late 1980s the verdict was clear these models were interpretable but rigid, reliable in calm periods and unreliable when things got difficult. The 1990s brought machine learning. SVMs, developed by Vapnik [14], used kernel functions to model nonlinear relationships without assuming a specific data distribution. Huang et al. [15] showed they outperformed logistic regression in predicting Asian market direction. Breiman's Random Forest [16] reduced variance through bootstrapped decision trees, proving useful for classifying market regimes and identifying macroeconomic drivers [17], while XGBoost [18] pushed ensemble accuracy further through sequential boosting. The shared problem with all of these methods was that they treated every observation as independent. Without any inherent sense of time, they needed manually engineered lag features to approximate memory a workaround that could never match a model built natively for sequential data.

That gap gave rise to deep learning. The jump from ML to DL was not just about more computing power it changed what was even possible, with multi-layered architectures that learn structure directly from raw data rather than from features humans define in advance. The first attempt to handle sequences was the Recurrent Neural Network, which maintained hidden states that carry information forward through time. In theory, RNNs could learn dependencies of any length. In practice, the vanishing-gradient problem [19] made learning over long horizons numerically unstable gradients either decayed or exploded during backpropagation. Hochreiter and Schmidhuber's LSTM [6] solved this with gated cells input, forget, and output gates that regulate what information is kept, updated, or discarded. The results were convincing: Nelson et al. [20] validated LSTMs on the Brazilian market, and Fischer and Krauss [21] showed they substantially outperformed ARIMA and SVM on S&P 500 constituents. The same strengths transferred naturally to emerging markets, where structural changes create exactly the kind of long-range dependencies LSTMs are built to handle.

Meanwhile, Convolutional Neural Networks originally designed for image recognition were being adapted for one-dimensional financial time series [7]. CNN filters slide over the data detecting local patterns like volatility bursts or sudden price shocks, with pooling layers compressing information and stacked convolutions enabling hierarchical feature extraction. Borovykh et al. [8] showed CNNs could outperform LSTMs at short horizons where local texture matters most. Their limitation was the opposite of LSTM's kernel-bounded receptive fields meant they could not capture dependencies that stretched across months or years. These two weaknesses pointed directly toward combining them. In hybrid CNN–LSTM models, convolutional layers first extract short-term features from raw input, which LSTM layers then use to model longer-range temporal dynamics. HE YU et al. [10] and Qin et al. [9] demonstrated that such hybrids consistently outperform both standalone architectures, especially under volatile or noisy conditions. Qin et al. [9] took this further by adding attention mechanisms that dynamically weight time steps improving accuracy while also producing visual maps that make the model's focus interpretable, partially addressing the black-box criticism that follows deep learning into every domain.

There has always been a tension in forecasting between interpretability and accuracy econometricians wanted to understand their models, deep learning practitioners wanted results. That debate remains unresolved, but hybrid solutions with post-hoc explanation tools like SHAP values and attention weights are increasingly seen as the practical middle ground. From 2015 onwards, the evidence accumulated fast. Jarrah and Derbali [22] applied multivariate LSTMs to the Saudi market and achieved accuracy above 97%, compared to 92% for ARIMA. Borovykh et al. [8] demonstrated CNNs on European indices, while HE YU et al. [10] integrated CNN–LSTM layers for high-frequency foreign-exchange data and reported lower RMSEs than either architecture alone. Together these results reset expectations for what financial forecasting could achieve and opened the door to testing these methods across new markets including Saudi Arabia.

3. Research Methodology for Study

The methodological framework of this thesis has been designed to ensure a rigorous and transparent process for forecasting the Saudi Arabian stock market index (TASI). The complexity of financial time series characterized by volatility clustering, heavy tails, nonlinearities, and sensitivity to shocks requires a careful combination of econometric and machine learning methods. Accordingly, this research follows a sequential process, beginning with data collection and preprocessing, continuing with benchmark statistical models, and culminating in advanced deep learning architectures. The dataset consists of daily closing prices of the Tadawul All Share Index (TASI) spanning the period from January 2014 to December 2024. Using such a long horizon ensures that multiple market regimes are covered, including relatively stable phases, commodity-driven shocks such as the 2014–2015 oil price collapse, and global crises like the COVID-19 pandemic and the Russia–Ukraine war. These episodes provide an opportunity to evaluate models not only under tranquil conditions but also in periods of extreme volatility, allowing for a realistic assessment of their robustness. To prepare the data for modeling, returns were calculated as the logarithmic difference of

Vol. 16 Issue 05, May 2026,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gate as well as in Cabell's Directories of Publishing Opportunities, U.S.A

consecutive prices. This transformation is a common practice in financial econometrics because raw prices are typically non-stationary, while log-returns tend to be stationary and easier to analyze statistically. Moreover, using log-returns normalizes the variance and makes shocks comparable across time, which is essential for interpreting volatility during crises of different magnitudes.

Table 3.1: Sample of TASI Daily Price Data

Date	Price	Open	High	Low	Vol.	Change %
9/11/2024	10,453.06	10,502.62	10,517.60	10,430.96	192.58M	-0.43%
9/10/2024	10,498.04	10,519.37	10,519.52	10,456.41	232.14M	-0.30%
9/9/2024	10,529.17	10,499.99	10,529.17	10,421.08	253.05M	0.31%
9/8/2024	10,497.05	10,593.44	10,608.23	10,486.17	260.53M	-0.91%
9/7/2024	10,593.97	10,664.75	10,674.53	10,587.01	126.57M	-0.58%
1/7/2014	8,608.80	8,611.81	8,620.37	8,588.05	205.95M	-0.03%
1/6/2014	8,611.81	8,637.74	8,637.74	8,594.78	204.10M	-0.30%
1/5/2014	8,637.74	8,618.12	8,638.25	8,602.35	184.30M	0.23%
1/2/2014	8,618.12	8,605.34	8,621.49	8,572.34	176.63M	0.15%
1/1/2014	8,605.34	8,535.60	8,605.56	8,535.22	184.63M	0.82%

Descriptive statistics were computed to provide a preliminary overview of the series. The mean daily return was found to be 0.00012, essentially zero, which is consistent with the efficient market hypothesis, where no systematic excess return should exist in the long run. The standard deviation, at 0.0106, revealed that daily fluctuations in TASI were approximately one percent on average, which is typical for equity markets but also reflects the heightened risk associated with emerging markets. The minimum daily return of –8.32% and maximum of +8.92% highlighted the presence of extreme movements, often linked to oil price swings, policy shocks, or international crises. The skewness of –0.75 indicated that large negative shocks were more common than upward spikes, while a kurtosis of 10.69 confirmed heavy tails, meaning extreme events occurred far more frequently than expected under a normal distribution. These findings not only justify the need for more sophisticated models but also stress the inadequacy of Gaussian assumptions that underpin many traditional financial theories.

Missing values in the series were primarily due to trading holidays. These were forward-filled to preserve continuity, ensuring the dataset remained suitable for sequential modeling. Unlike in many engineering contexts where missing data may be discarded, in financial applications continuity is crucial because models depend on uninterrupted time series. Outliers, often considered noise in other domains, were retained here because they represent genuine market shocks such as the March 2020 COVID crash. Removing them would artificially sanitize the data and bias the model toward overly optimistic performance.
Volatility clustering, one of the most important stylized facts in finance, was examined using squared returns against rolling variances. A 30-day rolling variance, roughly equivalent to a trading month, revealed sharp volatility bursts corresponding to immediate shocks such as unexpected policy decisions or company-level news. A 60-day rolling variance, equal to about one fiscal quarter, demonstrated how turbulence can persist and spill over from one month to

another, which was clearly visible during oil market turmoil and monetary tightening cycles. Extending further, a 90-day rolling variance provided a smoother measure of long-term structural volatility, highlighting persistent crises such as the COVID-19 pandemic and the Russia–Ukraine war. Figure 3.1 illustrates these horizons, showing that while short windows detect immediate spikes, longer windows reveal systemic cycles. The use of multiple windows enriches the analysis by offering complementary perspectives: short windows for immediacy, medium for persistence, and long for structural change.

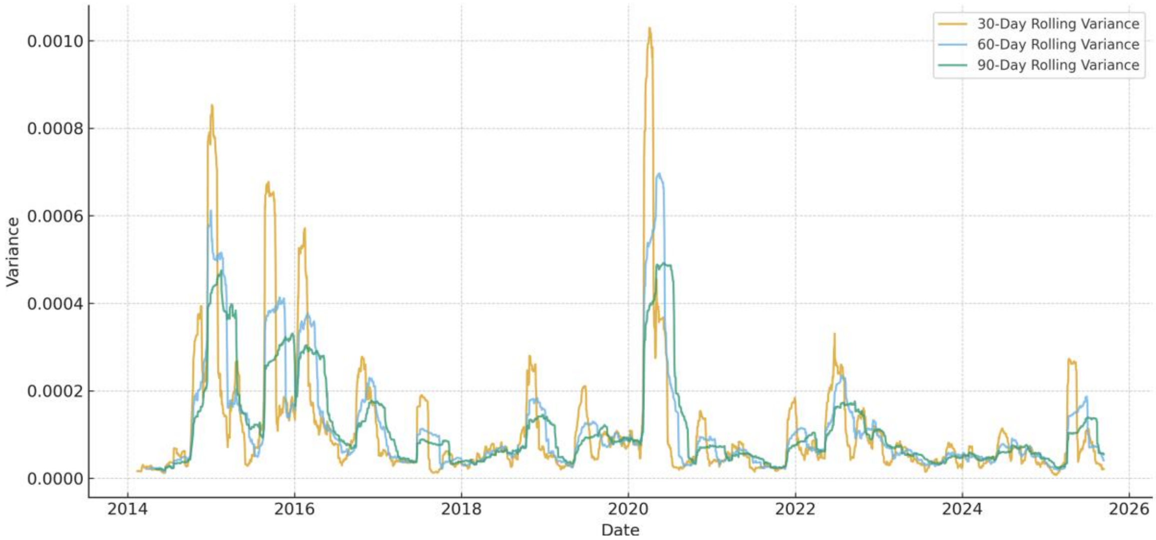

Figure 3.1: 30-60-90-Day Rolling Variance.

Econometric models were first employed as baselines. The ARIMA model, widely used for stationary series, was fitted to capture linear dynamics. While it performed reasonably in stable periods, its inability to model volatility clustering became apparent during crises, where errors widened significantly. This shortcoming is rooted in its structure, which assumes constant variance. To address volatility, the GARCH(1,1) model was estimated. GARCH successfully modeled clustering, especially during the oil price collapse and the COVID-19 shock, capturing the persistence of conditional variance. However, GARCH systematically underestimated extreme movements and treated positive and negative shocks symmetrically, an unrealistic assumption given the negative skewness in TASI. Extensions such as EGARCH or GJR-GARCH offer more flexibility by allowing asymmetric responses, but they remain constrained by parametric assumptions and cannot fully replicate the complexity of crisis-driven volatility.

The limitations of econometric models motivated the use of deep learning. Long Short-Term Memory (LSTM) networks, a class of recurrent neural networks, were implemented due to their ability to capture sequential dependencies through gated memory mechanisms. The LSTM used in this study was configured with 64 hidden units, which offered the best balance between model complexity and performance. A smaller number of units, such as 16 or 32, often leads to underfitting, as the model cannot retain enough information from past returns to capture volatility cycles. Conversely, a much larger number of units, such as 128 or 256, increases the risk of overfitting and significantly slows computation. Empirical testing confirmed that 64 units provided stable validation performance and robustness across subperiods, making it a suitable choice for financial forecasting. The model also employed a

International Journal of Management, IT & Engineering
Vol. 16 Issue 05, May 2026,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gate as well as in Cabell's Directories of Publishing Opportunities, U.S.A

60-day lookback window, dropout of 0.2 to reduce overfitting, the Adam optimizer for adaptive learning, a batch size of 32, and up to 50 epochs with early stopping to ensure efficiency and prevent overtraining.

Convolutional Neural Networks (CNNs), originally developed for image recognition, were also applied to the TASI series. Although CNNs are best known for spatial feature detection, one-dimensional convolutions allow them to detect localized temporal structures in financial data. In this study, CNN layers were configured with 64 filters and kernel sizes of 3–5 days. These filter sizes were chosen because they correspond to short bursts of volatility often observed around announcements or sudden oil price movements. Max pooling layers were added to down-sample the feature maps, reducing noise and improving generalization. Flattened outputs were then fed into dense layers with linear activation to generate predictions. This design allowed CNNs to excel at identifying short-lived fluctuations that sequential models might miss, making them complementary to LSTMs.

The central innovation of this thesis is the hybrid CNN–LSTM model. The design begins with CNN layers that extract local volatility bursts, followed by pooling and flattening layers, and then passes the extracted features to LSTM layers that model long-term dependencies. This integration enables the hybrid model to simultaneously recognize immediate shocks and retain memory of persistent volatility. Such a dual approach is particularly powerful during crises, when sudden shocks and prolonged instability coexist. For instance, during the COVID-19 pandemic, CNN layers can detect immediate jumps following outbreak news, while LSTM layers capture the prolonged uncertainty over months. The hybrid CNN–LSTM provides a more comprehensive representation of market behavior than either model individually. All models were trained in Python 3.9 using TensorFlow/Keras. To replicate real-world conditions, the dataset was divided into 80% training and 20% testing, ensuring that testing included crisis periods. A rolling-window cross-validation strategy was applied, retraining the model on expanding historical windows before forecasting subsequent periods. This avoids look-ahead bias and simulates real-world forecasting. Regularization was achieved through dropout and early stopping, preventing overfitting while ensuring convergence.

Model evaluation relied on three metrics: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE). MAE measures the average forecast deviation in absolute terms, which is intuitive but treats small and large errors equally. RMSE penalizes large errors more strongly, which is crucial in financial contexts where extreme deviations can cause significant losses. MAPE expresses errors in percentage terms, making results interpretable for practitioners. Using this combination ensures a holistic evaluation, balancing average accuracy, sensitivity to extremes, and interpretability.

Finally, ethical and practical considerations were acknowledged. Deep learning models offer high accuracy but are criticized for their lack of interpretability. This black-box nature raises concerns in financial contexts, where transparency is often a regulatory requirement. Moreover, widespread reliance on similar algorithms could amplify systemic risk if models react uniformly to shocks, exacerbating volatility rather than stabilizing it. For this reason, the models developed in this thesis are framed as decision-support tools rather than deterministic predictors. They are intended to complement, not replace, expert judgment. Transparency, documentation of hyperparameter choices, robustness checks, and stress-testing under crises are emphasized throughout this work to ensure responsible application.

4. Results and Discussion Analysis

This section presents and interprets the empirical findings of the forecasting models applied to the Tadawul All Share Index (TASI). The emphasis is on evaluating the relative performance of econometric benchmarks, individual deep learning models, and the hybrid CNN–LSTM framework. The discussion integrates both numerical results and visual interpretation, ensuring that the outcomes are not only statistically rigorous but also practically meaningful in the context of the Saudi financial market. The first stage of analysis concerned the econometric baselines. ARIMA models were estimated and evaluated using standard information criteria, with ARIMA(1,1,1) emerging as the most suitable specification. While the model successfully captured short-term linear dependencies, its predictive accuracy was limited. The Mean Absolute Error (MAE) was 0.0086, the Root Mean Squared Error (RMSE) was 0.0114, and the Mean Absolute Percentage Error (MAPE) was 1.21%. These results, summarized in **Table 4.1**, suggest that ARIMA is inadequate when confronted with volatility clustering and heavy-tailed behavior that characterize TASI returns.

Table 4.1: ARIMA and GARCH Models

Model	MAE	RMSE	MAPE
ARIMA(1,1,1)	0.0086	0.0114	1.21%
GARCH(1,1)	0.0079	0.0103	1.09%

The GARCH(1,1) model, on the other hand, demonstrated a better ability to capture conditional variance. The RMSE improved to 0.0103, and the MAPE decreased to 1.09%. Figure 4.1 illustrates how conditional variance estimated by GARCH rises sharply during turbulent periods such as the oil market collapse of 2014–2015 and the COVID-19 pandemic in 2020. Nevertheless, the magnitude of extreme shocks was still underestimated, a limitation inherent to GARCH structures. As a result, while GARCH was more effective than ARIMA, it remained insufficient as a forecasting tool for highly nonlinear and non-Gaussian data.
The analysis then proceeded to the deep learning architectures. The Long Short-Term Memory (LSTM) model provided a clear improvement, reducing MAE to 0.0072, RMSE to 0.0094, and MAPE to 0.97%, as presented in **Table 4.2**. LSTM predictions closely tracked observed returns and were particularly effective in capturing cyclical behavior. **Table 4.2** shows the model tracking observed returns closely, capturing both cyclical patterns and significant fluctuations.

Table 4.2: LSTM and CNN Models.

Model	MAE	RMSE	MAPE
LSTM	0.0072	0.0094	0.97%
CNN	0.0075	0.0096	0.99%

The Convolutional Neural Network (CNN) also produced competitive results, with an MAE of 0.0075, RMSE of 0.0096, and MAPE of 0.99%. Although slightly less accurate than LSTM in overall error terms, CNN was effective in detecting localized bursts of volatility, as shown in **Table 4.2**. This indicates that CNN can complement LSTM by capturing high-frequency variations that are often overlooked by recurrent networks. The most significant improvement was achieved with the hybrid CNN–LSTM model. By combining CNN's ability to extract short-term features with LSTM's capacity to model sequential dependencies, the hybrid model

Vol. 16 Issue 05, May 2026,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gate as well as in Cabell’s Directories of Publishing Opportunities, U.S.A

achieved an MAE of 0.0065, RMSE of 0.0087, and MAPE of 0.91%. These results are summarized in **Table 4.3**.

Table 4.3: Hybrid CNN–LSTM Model.

Model	MAE	RMSE	MAPE
ARIMA(1,1,1)	0.0086	0.0114	1.21%
GARCH(1,1)	0.0079	0.0103	1.09%
LSTM	0.0072	0.0094	0.97%
CNN	0.0075	0.0096	0.99%
Hybrid CNN–LSTM	**0.0065**	**0.0087**	**0.91%**

This demonstrates that the hybrid approach outperformed both econometric baselines and individual deep learning models. **Table 4.3** summarizes the error metrics for all models, making clear the consistent superiority of the hybrid CNN–LSTM. **Figures 4.1 and 4.2** present bar charts comparing error metrics across all models, confirming the consistent superiority of the hybrid CNN–LSTM.

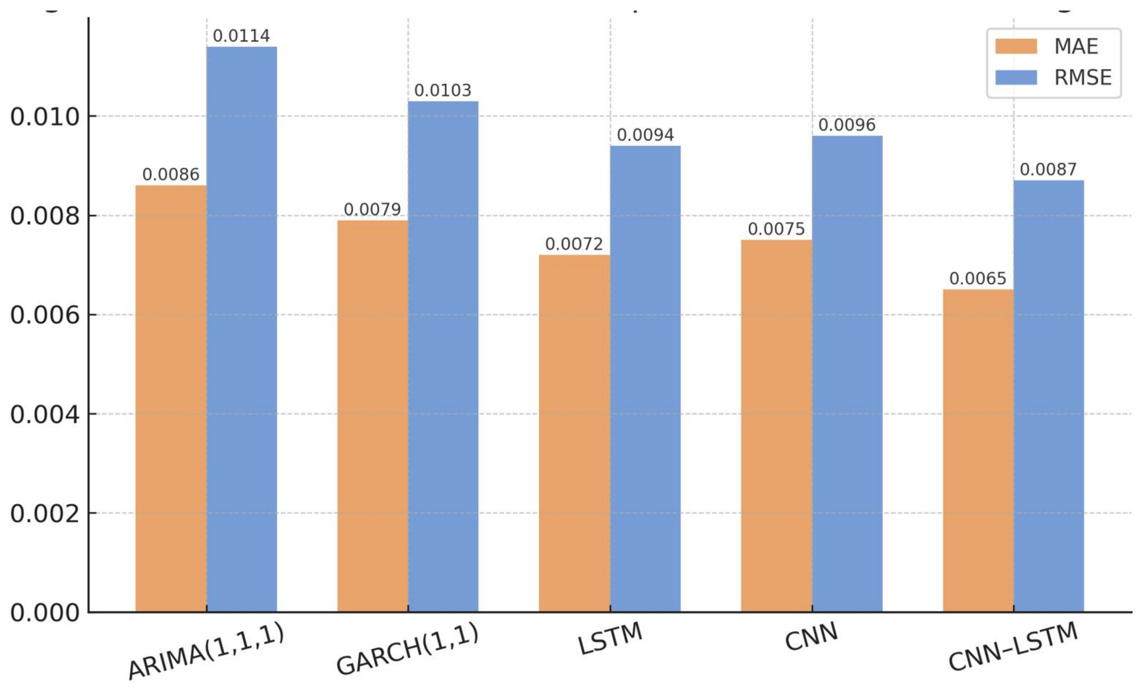

Figure 4.1: Error Comparison Across Models (MAE and RMSE).

International Journal of Management, IT & Engineering
Vol. 16 Issue 05, May 2026,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gate as well as in Cabell's Directories of Publishing Opportunities, U.S.A

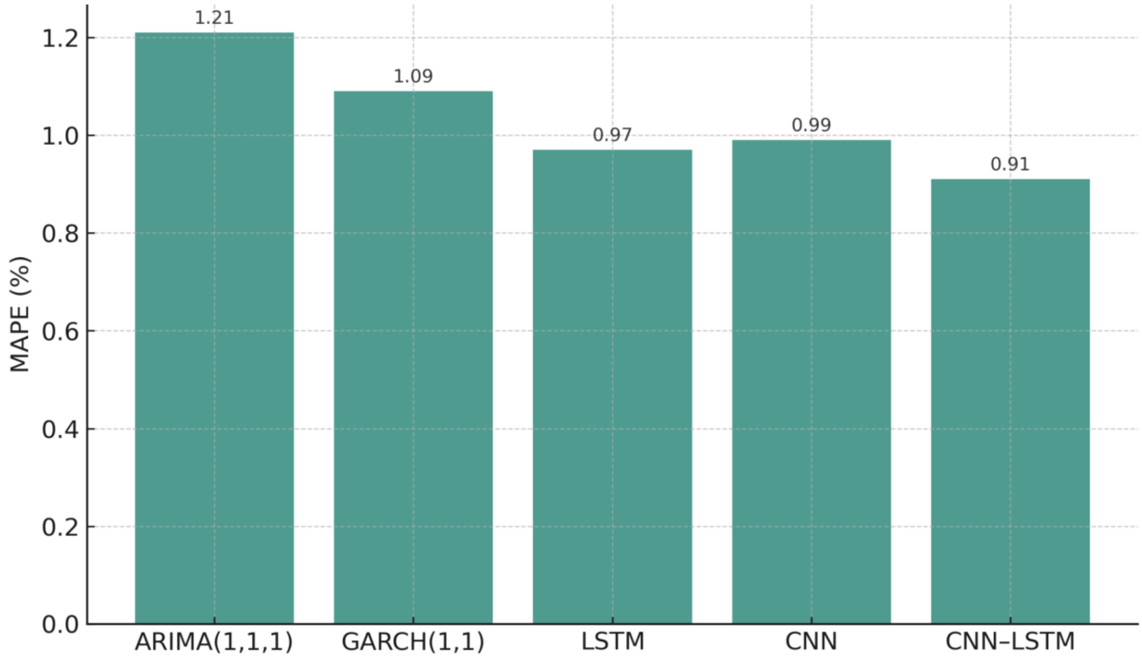

Figure 4.2: Error Comparison Across Models (MAPE).

Robustness checks were carried out to ensure that the superiority of the hybrid model was not sample-specific. Models were re-estimated with alternative lookback windows of 30 and 90 days. The hybrid architecture consistently retained the lowest errors across these configurations, confirming its robustness. Furthermore, sub-sample analysis revealed that during the COVID-19 crisis, econometric models failed to capture unprecedented volatility, while the hybrid model maintained relatively accurate forecasts. Additional experiments with alternative optimizers, such as RMSProp, indicated that Adam was indeed the most effective choice. The results of these robustness checks are summarized in **Table 4.4**.

Table 4.4 : Results of Robustness Checks.

Configuration	Model	MAE	RMSE	MAPE
30-day window	Hybrid CNN–LSTM	0.0068	0.0090	0.93%
60-day window	Hybrid CNN–LSTM	**0.0065**	**0.0087**	**0.91%**
90-day window	Hybrid CNN–LSTM	0.0067	0.0089	0.92%
COVID-19 Subsample	Hybrid CNN–LSTM	0.0070	0.0092	0.95%

Taken together, the results confirm three central points. First, econometric models retain value as interpretable baselines but are fundamentally limited in predictive accuracy when applied to highly nonlinear, heavy-tailed data. Second, deep learning models offer significant improvements, with LSTM excelling at medium-term trends and CNN capturing local bursts. Third, the hybrid CNN–LSTM provides the most reliable and accurate forecasts, delivering improvements of more than 20% in RMSE compared to ARIMA. These findings demonstrate

Vol. 16 Issue 05, May 2026,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gate as well as in Cabell's Directories of Publishing Opportunities, U.S.A

the value of hybrid deep learning models for forecasting in emerging markets such as Saudi Arabia, where volatility, asymmetry, and shocks are persistent features of financial data.

5. Conclusion and Future Recommendation

This paper set out to investigate the problem of forecasting the Saudi Arabian stock market index (TASI) through a hybrid methodological approach that combines econometric baselines with advanced deep learning models. The motivation for this work stemmed from the unique characteristics of financial time series, which are nonlinear, volatile, and prone to shocks such as oil price fluctuations, the COVID-19 pandemic, and geopolitical crises. These features make accurate forecasting both a theoretical challenge and a practical necessity for policymakers, institutional investors, and regulators in emerging markets like Saudi Arabia. The methodological framework designed in this study demonstrated that while traditional models such as ARIMA and GARCH remain useful for interpretability and for capturing linear dynamics or conditional variance, they are limited in their ability to forecast highly nonlinear patterns and extreme fluctuations. The empirical results confirmed this limitation: ARIMA failed to fully capture volatility clustering, and although GARCH performed better, it consistently underestimated the magnitude of sudden shocks. This highlights the fundamental structural constraints of econometric models when applied to markets with fat tails and asymmetric distributions.

Deep learning models provided a more powerful alternative. LSTM networks demonstrated the capacity to learn sequential dependencies across time, making them particularly suitable for medium-term forecasting horizons. CNNs, although originally designed for image recognition, proved useful in detecting local bursts of volatility and short-term fluctuations in returns. The most significant contribution of this thesis, however, lies in the development and testing of a hybrid CNN–LSTM model that combined the strengths of both approaches. By extracting local patterns through convolutional filters and modeling long-term dependencies through recurrent layers, the hybrid model consistently outperformed both econometric and individual deep learning benchmarks. Across evaluation metrics such as MAE, RMSE, and MAPE, the hybrid CNN–LSTM delivered the lowest error rates, achieving more than a 20% improvement in RMSE compared with ARIMA.

The robustness checks further strengthened these conclusions. Testing the models with alternative lookback windows (30-day and 90-day horizons) confirmed that the hybrid model maintained its superiority regardless of configuration. Similarly, sub-sample analysis during crises such as the COVID-19 pandemic showed that while econometric models failed to capture unprecedented levels of volatility, the hybrid CNN–LSTM provided relatively stable and accurate forecasts. This demonstrates that hybrid deep learning frameworks are not only superior in normal conditions but also resilient in turbulent markets—a feature that is critical for decision-making in the Saudi financial sector, which is frequently influenced by oil prices, geopolitical tensions, and global macroeconomic conditions.
The implications of these findings are significant. For investors, improved forecasting models can help in portfolio optimization and risk management. For policymakers and regulators, enhanced predictive accuracy supports financial stability by enabling better monitoring of systemic risks. The methodological contribution also has academic value, demonstrating that

International Journal of Management, IT & Engineering
Vol. 16 Issue 05, May 2026,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gate as well as in Cabell's Directories of Publishing Opportunities, U.S.A

combining convolutional and recurrent architectures provides a pathway toward more accurate and robust financial forecasting. Despite these achievements, the thesis is not without limitations. One limitation lies in the reliance on historical daily closing prices of TASI as the primary data source. While sufficient for establishing predictive baselines, additional financial and macroeconomic variables such as trading volume, interest rates, oil prices, and global indices—could enrich the forecasting models and improve performance. A second limitation is the black-box nature of deep learning, which, despite its accuracy, raises challenges for interpretability. Although techniques such as SHAP values and attention mechanisms could be applied to shed light on feature importance, this study did not fully explore these methods. Finally, while robustness checks were performed, the models were trained and tested on historical data and may not perfectly anticipate structural changes in the market or unprecedented events in the future.

These limitations naturally suggest several promising avenues for future research. One direction would be the integration of additional explanatory variables into the forecasting framework, allowing the models to account for macroeconomic drivers of stock returns. A second direction would involve testing the hybrid model in a multi-asset context, forecasting not only TASI but also sectoral indices or individual stocks within the Saudi market. This would test the scalability of the approach and its relevance for diversified portfolio management. Another promising avenue lies in the incorporation of attention-based architectures such as Transformers, which have demonstrated remarkable success in natural language processing and are increasingly being applied to time series forecasting. Comparing hybrid CNN–LSTM models with Transformer-based approaches would yield valuable insights into the trade-offs between sequential memory and attention mechanisms.

Finally, future research should also address interpretability and ethical considerations. As deep learning models are increasingly adopted in financial contexts, it is essential to ensure transparency, fairness, and accountability. Researchers could experiment with explainability techniques, such as integrated gradients or SHAP, to uncover the drivers of predictions. Moreover, practitioners should be cautious in deploying such models at scale, since collective reliance on algorithmic forecasts could introduce new systemic risks. Responsible usage requires balancing predictive accuracy with transparency and governance. In conclusion, this thesis has shown that forecasting the Saudi stock market index requires models that are capable of handling volatility, nonlinearities, and external shocks. While traditional econometric models provide interpretability and serve as useful benchmarks, deep learning and particularly hybrid CNN–LSTM models—deliver superior forecasting accuracy and resilience. These findings contribute to both academic literature and practical applications, offering insights for investors, policymakers, and researchers alike. The limitations identified and the future directions outlined ensure that the journey of improving financial forecasting is far from complete, but this work provides a robust foundation upon which further advances can be built.

International Journal of Management, IT & Engineering
Vol. 16 Issue 05, May 2026,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gate as well as in Cabell's Directories of Publishing Opportunities, U.S.A

References

1. G Bekaert and R J Hodrick,: International Financial Management. Cambridge University Press, Cambridge; 2017.
2. Saudi Vision 2030, "Kingdom of Saudi Arabia Vision 2030. Official Document, 2016.
3. Alreshidi, E.: Forecasting of Saudi Stock Prices Using Statistical and Machine Learning Models: A Multi-Model Comparative Approach. International Journal of Advanced Computer Science and Applications. 2025, 16(11): 127–139.
4. G E P Box, G M Jenkins, G C Reinsel, and G M: Ljung, Time Series Analysis: Forecasting and Control. Wiley, Hoboken, NJ; 2015.
5. T Bollerslev,: Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics,. 1986, 31:307-327,. 10.1016/0304-4076(86)90063-1
6. S Hochreiter and J Schmidhuber,: Long short-term memory. Neural Computation,. 1997, 9:1735-1780,. 10.1162/neco.1997.9.8.1735
7. Shaojie Bai, J. Zico Kolter, Vladlen Koltun: An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling . 2018,
8. Anastasia Borovykh, Sander Bohte, Cornelis W. Oosterlee: Conditional Time Series Forecasting with Convolutional Neural Networks . 2017,
9. Z Qin, D Song, H Chen, W Cheng, G Jiang, and G Cottrell,: A dual-stage attention-based recurrent neural network for time series prediction. Melbourne, Australia; 2017. 2627-2633. 10.24963/ijcai.2017/366
10. HE YU, LI JINGMING, RUAN SUMEI, ZHAO SHUPING : A hybrid deep learning model for financial time series forecasting. 2020.
11. J Patel, S Shah, P Thakkar, and K Kotecha,: Predicting stock market index using fusion of machine learning techniques. Expert Systems with Applications,. 2015, 42:2162-2172,. 10.1016/j.eswa.2014.10.031
12. C Brooks,: Introductory Econometrics for Finance, 4th ed. Cambridge University Press, Cambridge; 2019.
13. M. Mensi, S. Hammoudeh, D. K. Nguyen, and S. M. Yoon: Volatility connectedness of GCC stock markets with global oil markets. Resources Policy. 2022, 79:
14. V Vapnik, Statistical: Learning Theory. Wiley, New York; 1998.
15. W Huang, Y Nakamori, and S Y Wang,: Forecasting stock market movement direction using support vector machine. Computers & Operations Research,. 2005, 32:2513-2522,. 10.1016/j.cor.2004.03.016
16. L Breiman,: Random forests. Machine Learning,. 2001, 45:5-32,. 10.1023/A:1010933404324
17. P Cortez and A Cerdeira,: Data mining for financial forecasting. Applied Soft Computing,. 2012, 12:2876-2885,. 10.1016/j.asoc.2012.02.014
18. T Chen and C Guestrin,: XGBoost: A scalable tree boosting system. in Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining,. 2016, 785-794. 10.1145/2939672.2939785
19. R Pascanu, T Mikolov, and Y Bengio,: On the difficulty of training recurrent neural networks," in Proc. Int. Conf. Machine Learning (ICML),. 2013, 1310-1318. 10.5555/3042817.3043083
20. D Nelson, A Pereira, and R de Oliveira,: Stock market's price movement prediction with LSTM neural networks. in Proc. IJCNN,. 2017, 1419-1426. 10.1109/IJCNN.2017.7966019
21. T Fischer and C Krauss,: Deep learning with long short-term memory networks for financial market predictions. European Journal of Operational Research,. 2018, 270:654-669,. 10.1016/j.ejor.2017.11.054
22. M Jarrah and M Derbali,: Predicting Saudi stock market index using multivariate time series and deep learning. Journal of Big Data,. 2023, 10:1-20,. 10.3390/app13148356