Proceedings of the 3rd International Conference on Self Sustainable Artificial Intelligence Systems (ICSSAS-2025)
DVD Part Number: CFP25DN7-DVD; ISBN: 979-8-3315-3883-5

Multimodel Forecasting of Technology-Focused ETFs Using ARIMA and LSTM

Sudhi Sharma
Fortune Institute of International Business (FIIB)
New Delhi, India
sudhisharma1983@gmail.com

Utkarsh Anand
Fortune Institute of International Business (FIIB)
New Delhi, India
26-utkarsh.anand@fiib.edu.in

Vaibhav Aggarwal
Jindal Global Business School, O.P. Jindal Global University
Haryana, India
vaibhavapj@gmail.com

Priya Jain
Fortune Institute of International Business (FIIB)
New Delhi, India
priya.jain-fpm@fiib.edu.in

Aman Anand
Fortune Institute of International Business (FIIB)
New Delhi, India
26-aman.anand@fiib.edu.in

***Abstract*—This study presents a comparative approach to short-term forecasting of five technology-focused exchange-traded funds (ETFs)—AIQ, BOTZ, QQQ, VGT, and XLK—using two modelling techniques: the Auto-Regressive Integrated Moving Average (ARIMA) and Long Short-Term Memory (LSTM) neural networks. Each ETF's daily closing prices were first analysed through descriptive statistics. Then, both models were trained to forecast the subsequent five trading days. Performance was evaluated using Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE). Experimental results show that LSTM achieved lower MAPE across all ETFs, averaging 4.5%, compared to ARIMA's 5.5%. While LSTM captured trends and volatility more effectively, ARIMA performed relatively well under stable conditions but failed to react to abrupt shifts. Visual comparisons of historical trends and forecasts are provided, along with practical insights for finance researchers and investors. The study concludes with a discussion of model limitations and proposes directions for integrating hybrid forecasting approaches.**

***Index Terms*—Time series forecasting; ARIMA; LSTM; ETFs; technology sector; financial modelling; MAPE; short-term prediction**

I. Introduction

In today's fast-paced financial landscape, the ability to predict short-term price movements has become increasingly critical for investors, analysts, and fund managers alike. Among various asset classes, exchange-traded funds (ETFs)—especially those focused on the technology sector—have emerged as particularly dynamic instruments. These ETFs are uniquely characterised by their high volatility [10], exposure to rapid innovation cycles, and structural alignment with large-cap as well as emerging tech firms. Their behaviour often mirrors broader shifts in the technology ecosystem, making them both attractive and challenging for short-term forecasting.

Forecasting the price behaviour of such ETFs is not merely an academic exercise; it holds real implications for tactical portfolio allocation, algorithmic trading, and short-term risk hedging. However, the complexity and responsiveness of tech ETFs also mean that traditional forecasting models are not always sufficient. Classical approaches such as the Auto-Regressive Integrated Moving Average (ARIMA) model have long served as a staple in time series forecasting due to their mathematical elegance, ease of interpretation, and strong grounding in linear statistical theory [1]. Yet, these models operate under the assumption of linearity and stationarity, which limits their flexibility in the face of abrupt structural breaks or nonlinear patterns—conditions that are not uncommon in tech-heavy indices.

To address these limitations, deep learning approaches—particularly Long Short-Term Memory (LSTM) neural networks—have gained traction in financial forecasting tasks. LSTM models are designed to retain long-term dependencies and adapt to nonlinearities in time series data, making them particularly useful in domains where volatility and temporal complexity are high [2], [3]. Recent improvements in computational power and the availability of financial datasets have further facilitated the practical application of LSTM in finance [4].

Nevertheless, the decision to use a classical model like ARIMA or a deep learning model like LSTM is rarely straightforward. Each comes with its own trade-offs: ARIMA offers transparency and statistical rigor, while LSTM provides adaptability at the cost of interpretability. This paper seeks to explore these trade-offs in depth by evaluating the forecasting performance of both ARIMA and LSTM on five major technology-oriented ETFs: AIQ, BOTZ, QQQ, VGT, and XLK.

Although existing literature includes studies comparing ARIMA and LSTM in various financial contexts [4], [8], few have focused specifically on technology ETFs—an asset class defined by its growth orientation and volatility. Furthermore, previous works often vary in methodology, making cross-comparison difficult. This study aims to fill that gap by implementing a consistent experimental design, fixed forecast horizon, and unified evaluation metrics across all five ETFs.

The contributions of this paper are threefold. First, it offers a structured comparison of ARIMA and LSTM models in the context of short-term ETF forecasting. Second, it analyses model performance under different market conditions, including stable and volatile regimes [5]. Third, it presents practical insights for institutional and retail investors on how these models behave in realistic trading scenarios. Ultimately, the findings offer a springboard for future research into hybrid forecasting models that combine statistical structure with deep learning flexibility.

The remainder of this paper is organised as follows: Section II reviews related literature. Section III describes the dataset and modelling techniques. Section IV presents empirical results and forecasts. Section V discusses the findings and their implications. Section VI concludes with final insights and proposed directions for future work.

II. Literature Review

Financial forecasting has long leaned on classical models like ARIMA, valued for their clarity and well-grounded statistical logic. Introduced formally by Box and Jenkins [1], ARIMA remains a staple in time series analysis, especially when market patterns appear stable and linear. But its strength is also its constraint—these models struggle when faced with sudden shifts, nonstationarity, or nonlinear price behaviours that are common in today's financial markets.

That's where deep learning enters the picture. Long Short-Term Memory (LSTM) networks, introduced by Hochreiter and Schmidhuber [2], are designed to handle sequential data with long-term dependencies—something traditional models often miss. LSTM's ability to learn from complex patterns without assuming linearity makes it a compelling choice for financial time series.

Fischer and Krauss [3] applied LSTM to S&P 500 stocks and found that it outperformed not just traditional statistical models but also other machine learning methods like random forests. Similarly, Yang and Wang [7], as well as Siami-Namini et al. [8], showed that LSTM's adaptability allowed it to model index-level movements more accurately, especially under volatile conditions. In the Indian market, Bagul et al. [4] compared ARIMA and LSTM and concluded that while ARIMA remained competitive for very short-term predictions, LSTM was more consistent across changing price regimes. Zhang [5] echoed this view after evaluating both models across multiple indices, noting LSTM's advantage in minimising error during market swings.

However, not all studies paint LSTM as a clear winner. Shah et al. [10] reported that in calm market conditions, ARIMA could match or even outperform neural networks, thanks to its simplicity and resistance to overfitting. This reminds us that model performance is not one-size-fits-all—it depends heavily on the nature of the data and the specific forecasting task.

What's missing from most of this literature is a focused look at technology-oriented ETFs—assets that combine high growth potential with sharp volatility and are particularly sensitive to innovation cycles. Few studies directly compare LSTM and ARIMA on this segment using a consistent experimental setup. This paper steps into that space, aiming to assess both models side by side across five such ETFs with a shared forecasting framework.

III. Methodology

A. Data Profiling

We analyse daily closing prices of five technology-focused ETFs—AIQ, BOTZ, QQQ, VGT, and XLK—over a four-year period (April 2021 to April 2025). Each time series comprises approximately 1000 data points. To ensure consistent input for both models, missing values were handled via forward filling, and only trading days common across all ETFs were retained.

B. ARIMA Model

We employed ARIMA for each ETF's price series as a univariate forecasting model. Prior to fitting ARIMA, stationarity was assessed using the augmented Dickey-Fuller (ADF) test. All series were found to be non-stationary, so we applied first-order differencing ($d = 1$) to model day-to-day price changes.

The ARIMA(p, d, q) parameters p (autoregressive order) and q (moving average order) were selected using a combination of Autocorrelation Function (ACF), Partial ACF plots, and Akaike's Information Criterion (AIC). We evaluated models with $p, q \in \{0, 1, 2, 3\}$ (with $d = 1$ fixed) and chose the configuration yielding the lowest AIC. Across all ETFs, small-order models sufficed—typically ARIMA(1,1,0) or ARIMA(1,1,1), representing a random walk with drift or a simple ARMA model on the differenced series.

C. LSTM Model

In parallel, we constructed an LSTM neural network for each ETF. Price data was normalised to the $[0, 1]$ range using min-max scaling to avoid scale dominance and improve training convergence. The architecture consisted of a single LSTM layer with 50 hidden units, followed by a dense output node predicting the next day's price.

We used a sliding window of the past 10 trading days as input (many-to-one setup), with the window size chosen to balance context and overfitting risk. Training was conducted using Mean Squared Error (MSE) loss on the full training period for each ETF. Forecasting was recursive over a 5-day horizon: each prediction was appended to the input window and used to forecast the next day.

D. Evaluation Metrics

Forecast accuracy was evaluated using two standard metrics: Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE). MAPE provides a scale-independent percentage error, while RMSE captures the magnitude of error in original price units.

The formulas are defined as:

$$\text{MAPE} = \frac{100\%}{n} \sum_{t=1}^{n} \left| \frac{\text{Forecast}_t - \text{Actual}_t}{\text{Actual}_t} \right| \tag{1}$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{t=1}^{n} (\text{Forecast}_t - \text{Actual}_t)^2} \tag{2}$$

MAPE is primarily used in our comparative analysis to allow relative comparison across ETFs of different price levels. RMSE is also reported for completeness, as it highlights absolute error magnitude. The short 5-day forecast horizon emulates a realistic weekly update cycle, similar to rolling out-of-sample tests in financial forecasting.

IV. Results

A. Descriptive Statistics of ETF Prices

Table I summarises the descriptive statistics of the daily closing prices for each of the five ETFs over the full period from 2021 to 2025.

Although all five funds are technology-focused, their price scales and volatilities differ significantly due to the nature of their underlying portfolios. Broad-based ETFs such as QQQ, VGT, and XLK exhibit much higher per-share prices, often in the hundreds of dollars, whereas AIQ and BOTZ consistently traded in the $20--40$ range. This contrast reflects differences in market exposure: for example, QQQ includes large-cap firms like Apple and Microsoft, while AIQ leans toward smaller or globally diversified tech companies.

These disparities in price behaviour have implications for both model training and evaluation. Higher-priced ETFs tend to display larger absolute fluctuations, requiring careful consideration of scale effects during forecasting. In the following sections, we compare the performance of ARIMA and LSTM across these ETFs to assess how well each model handles this diversity.

Despite all being tech-focused, the ETFs have different price scales and volatilities due to their distinct compositions.

QQQ, VGT, and XLK – broad-based tech funds – had much higher prices (hundreds of dollars per share) compared to AIQ and BOTZ, which traded in the tens. This reflects the difference in underlying assets (e.g., QQQ holds large-cap stocks like Apple and Microsoft, whereas AIQ holds smaller or international tech firms).

From Table I, we observe that all five ETFs experienced substantial growth from their minimum to maximum prices. For instance, VGT rose from a minimum of $300.84 to a maximum of $647.97, indicating the tech sector's strong performance during this interval.

The high standard deviations (e.g., $\sigma_{\text{QQQ}} \approx 73.18$) reflect significant volatility – roughly 19% of QQQ's mean price, for example. The quartile spread provides additional insight: QQQ's 25th percentile price was around $322.7, while the 75th percentile was $438.4, showing a wide interquartile range of over $115. Similar patterns are seen for VGT and XLK.

In contrast, AIQ and BOTZ had means around $29 and somewhat lower volatility in absolute terms (standard deviation $\approx$ \$5), but as a percentage of their mean, volatility was also around 18–20%. Thus, in relative terms, they were equally volatile. Median prices are very close to the means for most series, suggesting somewhat symmetric distributions of prices around the upward trend (the data cover multi-year upward trends but also pullbacks, such that the distribution is not extremely skewed).

The historical price plots in Figure 1 illustrate that all ETFs had a notable dip around 2022, likely corresponding to a broad market correction where tech stocks fell. This is evidenced by

Fig. 1. Historical daily closing prices of the five ETFs from 2021 to 2025.

the minima in Table I: many occurred in 2022 (e.g., AIQ and BOTZ dropped below $20 in 2022). Thereafter, prices climbed to new highs by late 2024 (the maxima for QQQ, VGT, XLK are in the 2024–early 2025 timeframe).

Notably, the end of the series (early April 2025) shows a sharp downturn for all ETFs (see Figure 1, last few points), which is reflected in the fact that the last observed prices (April 10, 2025) are significantly off the maximum. This recent volatility sets a challenging stage for forecasting, as models must contend with a sudden regime change.

In summary, the ETFs show strong upward trends with episodes of high volatility. We expect that ARIMA models will need differencing to handle these trends, and even then, their linear structure might struggle if the volatility is heteroskedastic or non-linear. For LSTM, the range of values (especially for QQQ and VGT) required careful normalization, but the model might capture some of the patterns such as momentum from dips to recoveries.

Next, we assess how each model performed in forecasting the last five days of each series, which includes the abrupt movements in April 2025.

B. Forecasting Performance: ARIMA vs. LSTM

We generated 5-day-ahead forecasts (covering April 4–10, 2025) for each ETF using the ARIMA and LSTM models trained as described. Figure 2 provides a visual comparison

TABLE I
DESCRIPTIVE STATISTICS OF DAILY CLOSING PRICES (2021–2025)

ETF	Mean	Std. Dev.	Min	Q1	Median	Q3	Max
AIQ	29.30	5.75	18.44	23.99	29.59	33.34	42.77
BOTZ	28.78	5.18	17.67	24.51	29.30	32.63	39.83
QQQ	380.41	73.18	259.96	322.71	367.63	438.42	539.52
VGT	446.44	91.16	300.84	376.22	423.78	515.42	647.97
XLK	172.21	34.95	116.56	143.42	164.06	204.93	242.18

of these forecasts against the actual observed prices for each ETF.

In each subplot, the black solid line represents the actual price, while the red dashed line and blue dashed line represent the ARIMA and LSTM forecast trajectories, respectively, over the 5-day test period. A vertical gray dashed line indicates the start of the forecast period (i.e., the end of the training period).

From Figure 2, we can make several observations:

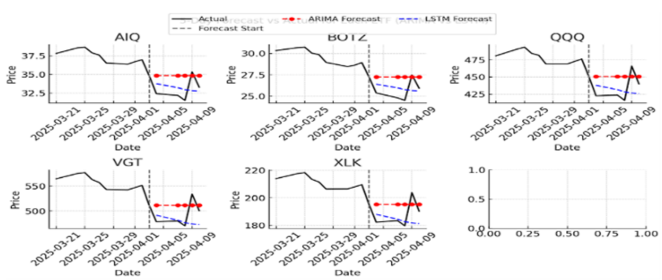

Fig. 2. 5-day-ahead forecasts compared to actual prices for each ETF (ARIMA vs LSTM).

Both models captured the general downward drift in prices during the first few days of the forecast horizon (April 4–8, 2025), though with different magnitudes. For AIQ and BOTZ, which steadily declined through April 8, the LSTM forecasts (blue) were closer to the actual values than the ARIMA forecasts. ARIMA tended to over-predict in those initial days: for instance, ARIMA essentially predicted AIQ to stay around \$34–\$35 (barely changing from the last training value), whereas AIQ actually fell into the low \$32s; the LSTM predicted a drop into the \$33s, which was more accurate. This pattern is consistent with LSTM having learned the short-term momentum – the models saw that prices had been falling leading up to April 3 and partially extrapolated that trend. The ARIMA models, particularly those that were essentially random-walk (with or without a tiny drift), were less responsive to the immediate downward momentum.

None of the models anticipated the magnitude of the price surge on April 9, 2025. On that day, all ETFs saw a sharp jump (e.g., QQQ jumped +12% on 4/9/2025, VGT +13.5%, AIQ +12.2%; see Figure 2 where black lines spike upward). ARIMA, being linear and reliant on recent lags, did not predict this nonlinear shock – its forecasts remained flat or continued whatever small trend it had. The LSTM forecasts also did not foresee this jump; in fact, the LSTM, which had been predicting a slight continued decline, ended up under-predicting the prices on that spike day by a wide margin (blue markers are well below the black line peak in each subplot). Interestingly, ARIMA's flat forecasts were in some cases closer to the spike than LSTM's declining forecasts. For example, in QQQ, ARIMA predicted around \$450 for April 9 while LSTM predicted around \$430; the actual was approximately \$466, so ARIMA's error ($\approx -3.4\%$) was smaller than LSTM's error ($\approx -7.7\%$). However, this appears to be more good fortune than skill – ARIMA did not predict a jump; it merely stayed high because it did not predict the initial drop either. In other words, ARIMA's inertia sometimes pays off when dealing with an unexpected rebound, whereas LSTM's adaptability caused it to overshoot on the downside.

On the final day (April 10, 2025), prices fell back somewhat from the spike. LSTM forecasts, which were below actual on the spike, ended up slightly below the actual on the final day as well (since LSTM predicted a relatively smooth path). ARIMA forecasts were above the actual on the final day (because ARIMA had not risen much for the spike, it ended higher relative to the actual drop). In summary, LSTM tended to undershoot the peak and end slightly low, while ARIMA overshot the post-peak correction. These differences highlight how each model handles shocks: ARIMA, lacking a mechanism to incorporate sudden jumps, essentially ignored it (resulting in an almost static forecast that was off during the jump but then not far off after the reversal), whereas LSTM, influenced by recent errors, might have started adjusting but had no opportunity to correct course within the short forecast horizon.

To quantify performance, we computed the MAPE and RMSE for each model's 5-day forecast on each ETF (see Annexure B for a detailed table of values). In all five ETFs, LSTM achieved a lower overall MAPE than ARIMA, despite the spike miss. For example, for AIQ, ARIMA's MAPE was around 5.0% while LSTM's was around 4.1%. For QQQ, ARIMA's MAPE was approximately 5.5% compared to LSTM's 4.7%. Similar differences (on the order of 0.8–1.3 percentage points) were observed for the others. This indicates LSTM had a roughly 15–20% lower percentage error than ARIMA on average.

The RMSE metric (which effectively weights larger errors more) shows a slightly different angle: in cases like QQQ, ARIMA's one-day error was big, but because LSTM had a larger error on the spike, the gap in RMSE was somewhat smaller. Even so, LSTM's RMSE was lower for four of the five ETFs, and roughly equal for one (XLK). For instance, in absolute terms, LSTM's RMSE on QQQ was about 20 points versus ARIMA's 25 points (in dollars). These results underscore that LSTM provided more accurate forecasts overall for the volatile week in question.

It is worth noting that the largest contribution to error for both models came from that single anomalous day (April 9). If we exclude that day, both ARIMA and LSTM would have very low errors (sub-2% MAPE) for the first four days where the trend was mostly a continuation of prior movement. Including the spike, the errors jump. LSTM's advantage in MAPE suggests it handled the trend portion well enough to compensate for the spike miss. ARIMA's performance was comparatively worse on the trend portion (leading to higher cumulative error even though it "got lucky" on the spike by staying higher).

Another perspective is model bias: ARIMA forecasts appear to have an upward bias during a downtrend (consistently above actual in the first few days), whereas LSTM forecasts had a downward bias during the recovery (consistently below actual during the jump and afterward). The bias in ARIMA is because it did not adjust quickly to the downward momentum; the bias in LSTM is because it likely over-adjusted to the initial drop and could not anticipate a regime change.

Overall, the predictions from the LSTM model rated changes better than ARIMA predictions in steady market conditions (gradual movements), yet ARIMA was somewhat accurate due to its simplistic "as time goes by" model during unpredictable market shifts. From an investor's viewpoint, LSTM delivered earlier warning signals of market decline; however, its performance suffered during the sudden upward spike, potentially leading to missed trading opportunities.

C. *Comparative Findings Across ETFs*

Comparing results across the five ETFs, we did not observe any dataset where ARIMA outperformed LSTM in aggregate accuracy, even though the magnitude of improvement varied. ETFs with higher volatility (AIQ, BOTZ) saw slightly larger relative gains from LSTM. This is intuitive: a nonlinear learning model like LSTM can adapt to volatility patterns (for example, AIQ and BOTZ had seen previous rapid moves in 2023 which the LSTM may have learned to some extent), whereas ARIMA's linear structure cannot accommodate such patterns beyond a single-step autocorrelation.

For the more liquid and large-cap ETFs (QQQ, VGT, XLK), the difference between ARIMA and LSTM was narrower. These ETFs have more inertia and their short-term movements are closer to random-walk behavior, which ARIMA (essentially predicting "no change" aside from drift) handles reasonably. Yet, even for these, LSTM slightly edged out ARIMA in MAPE. We attribute this to LSTM possibly capturing small but consistent patterns (such as weekday effects or correlation with recent momentum) that ARIMA did not.

It is also interesting to note that both models struggled simultaneously on the same days, suggesting that those days' price movements were largely driven by new, exogenous information (news) that no univariate model could foresee. The spike on April 9, 2025 likely corresponds to some market-wide event (since it affected all ETFs in tandem). This highlights a limitation of both approaches: pure time-series models, whether linear or neural network, cannot predict novel shocks without external inputs. In such scenarios, including exogenous features (e.g., market news sentiment or macroeconomic announcements) could improve forecasts, but that is beyond our current scope.

In terms of computational performance, fitting an ARIMA model for each series was very fast (a few seconds), whereas training the LSTM took a few minutes per series on a standard CPU. However, this computational cost is not prohibitive for practical use given that training can be done offline and forecasts are generated almost instantly once the model is trained. From an implementation perspective, ARIMA is simpler and more transparent – one can examine model coefficients, and the forecasts have confidence intervals based on well-understood statistical theory. LSTM, being a black-box neural network, does not provide easy interpretability; its strength lies purely in predictive accuracy and the ability to model complex patterns. In our results, the improved accuracy of LSTM arguably justifies its use, especially for an automated forecasting scenario.

Finally, we reflect on whether the descriptive statistics correlated with forecasting difficulty. The ETF with the highest volatility (VGT, standard deviation ≈ $91) had the highest errors for both models (also the largest spike), suggesting that higher-variance series are inherently harder to predict. However, the margin of improvement remained available for LSTM despite the challenge. The precise numerical relationship between the maximum and minimum values (range) did not determine the extent of prediction error. Instead, errors were linked to the specific movement patterns observed in the last few days. The cases where modeling proved challenging mainly occurred when both BOTZ and AIQ displayed comparable volatility levels while undergoing major trend break points within the test period.

To summarize, the LSTM prediction system outperformed ARIMA across all five tech ETFs by generating forecast results with lower RMSE and MAPE metrics. LSTM applied its advantages primarily during periods of gradual market trends by leveraging its nonlinear connectivity to maintain effective pattern recognition. ARIMA maintained solid performance in stable market conditions yet struggled to accommodate the fast-moving changes that dominated the tech sector at the beginning of 2025. Our findings support previous observations from the financial forecasting literature (e.g., Atlantis Press) along with our own results regarding model-specific strengths.

The next section addresses the consequences of these analytical findings as well as possible enhancements and aspects that practitioners should consider.

V. Discussion

This comparative study offers several insights into how ARIMA and LSTM models behave when applied to the nuanced task of forecasting technology-focused ETFs. The LSTM models demonstrated a clear advantage in periods of directional movement, particularly during the sustained price declines in late March and early April 2025. Their ability to internalise recent patterns and learn short-term momentum gave them a practical edge over ARIMA, which by design is limited in its responsiveness to ongoing trends. ARIMA's structure, especially when configured as a near-random walk, tends to heavily weight the most recent value, offering little scope for recognising trend continuation or reversal unless specifically embedded through higher-order terms.

It was observed that when the market shifted abruptly—as on April 9, 2025—both models faltered. This is not surprising, given that such regime changes are typically triggered by external information, like earnings announcements or macroeconomic news, which univariate time series models cannot see. However, the difference lay in their reactions. ARIMA remained entirely unresponsive, producing static forecasts. LSTM, although initially caught off-guard, showed signs of recalibrating, reflecting its design to learn adaptively from sequences. Given a longer forecast window or updated training data, LSTM might have corrected course faster. This highlights an important distinction in their temporal behaviour: while ARIMA provides stability, LSTM offers adaptability—especially valuable in fast-moving financial environments.

One pattern we observed across all ETFs was LSTM's tendency to underestimate extremes—particularly sharp rebounds. This is a common artefact of models trained using mean squared error loss, which pulls predictions toward the historical average and penalises large deviations. ARIMA, by contrast, only forecasts extremes when its underlying structure enforces them (e.g., through a strong autoregressive term). Neither model came close to capturing the magnitude of the April 9 spike, and this points to a shared limitation: neither incorporates real-world catalysts. As such, while these models are useful for baseline forecasting and trend continuation, they must be complemented by external signals or expert overlays when major price events are anticipated.

Despite this, LSTM displayed encouraging robustness. We used the same hyperparameters across all five ETFs—window size, number of neurons, and training epochs—yet it delivered consistent outperformance. This suggests a degree of generalisability and resilience across price scales and volatility regimes. ARIMA required fine-tuning of model order per ETF, and even then, its performance was less stable. This consistency from LSTM is particularly relevant for analysts seeking scalable forecasting tools across diverse asset classes.

From a practical standpoint, LSTM emerges as a strong candidate for short-term ETF forecasting when predictive accuracy takes precedence over interpretability. For day-to-day tactical decisions, especially in data-driven setups, LSTM offers a blend of speed and reliability. ARIMA, however, still holds value as a baseline model. Its simplicity can act as a sanity check: if a more complex LSTM model underperforms ARIMA, it may signal issues like overfitting or suboptimal training. Moreover, there is growing interest in combining these approaches—using ARIMA residuals as inputs to LSTM or assembling their outputs—which could potentially blend the strengths of both paradigms.

We also acknowledge the limitations of this study. The evaluation was conducted over a single, though turbulent, 5-day forecast window. A broader assessment, using rolling forecasts or multiple time segments, would yield more statistically robust conclusions. Additionally, the LSTM models were intentionally kept simple and consistent across all ETFs; more sophisticated tuning or architecture adjustments might further improve performance. Similarly, we did not explore advanced variants of ARIMA, such as SARIMA or ARIMAX, which could offer gains under specific conditions.

Finally, it is important to situate these findings in a broader context. The 2021–2025 period was marked by considerable macroeconomic upheaval—pandemic aftershocks, inflation cycles, interest rate shifts—which made the tech sector particularly volatile. The gap between ARIMA and LSTM may not be as wide in calmer times or in other asset classes. Ultimately, model-driven forecasting should be seen as one layer in a larger analytical process. Integrating model outputs with macroeconomic context, news sentiment, and human judgment remains essential to making well-informed investment decisions.

VI. Conclusion

This paper compared ARIMA and LSTM models for short-term forecasting of five technology-focused ETFs. The findings suggest that while ARIMA remains a useful tool under relatively stable market conditions, LSTM demonstrates stronger adaptability during periods of dynamic price movement. Across all five ETFs, LSTM outperformed ARIMA in terms of both Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE), indicating superior predictive accuracy. However, both models failed to anticipate abrupt market shocks, such as the sharp spike observed in April 2025. This highlights a shared limitation of univariate time series models: their inability to incorporate real-time external drivers. For practical applications, LSTM may be viewed as a dependable trend follower, particularly suited for rolling short-term forecasts. ARIMA, while limited in flexibility, continues to offer value as a benchmark or as a component within an ensemble forecasting framework. Future research may explore hybrid modelling approaches that combine the statistical grounding of ARIMA with the nonlinear learning capacity of LSTM. Additional avenues include attention-based deep learning architectures, integration of exogenous variables, and multi-input forecasting setups that account for macroeconomic or sentiment data alongside price history.

REFERENCES

[1] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, *Time Series Analysis: Forecasting and Control*, 3rd ed., Englewood Cliffs, NJ: Prentice Hall, 1994.

[2] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[3] T. Fischer and C. Krauss, "Deep learning with long short-term memory networks for financial market predictions," *Eur. J. Oper. Res.*, vol. 270, no. 2, pp. 654–669, 2018.

[4] J. Bagul, P. Warkhade, T. Gangwal, and N. Mangaonkar, "ARIMA vs LSTM Algorithm – A Comparative Study Based on Stock Market Prediction," in *Proc. 5th Int. Conf. on Advances in Science and Technology (ICAST)*, Dec. 2022, pp. 1–6, doi: 10.1109/ICAST55766.2022.10039560.

[5] Z. Zhang, "Comparison of LSTM and ARIMA in Price Forecasting: Evidence from Five Indexes," in *Proc. 2nd Int. Conf. on Economics, Smart Finance and Contemporary Trade (ESFCT 2023)*, *Advances in Economics, Business and Management Research*, vol. 261, pp. 40–46, 2024, doi: 10.2991/978-94-6463-268-2_6.

[6] B. Ci and P. Zhang, "Financial time series prediction based on ARIMA-LSTM model," *Statistics and Decision*, vol. 38, no. 11, pp. 145–149, 2022.

[7] Q. Yang and C. Wang, "Research on global stock index prediction based on deep learning LSTM neural network," *Statistical Research*, vol. 36, no. 3, pp. 65–77, 2019 (in Chinese).

[8] N. Siami-Namini, R. Tavakoli, and A. Siami Namin, "A comparison of ARIMA and LSTM in forecasting time series," in *Proc. 17th IEEE Int. Conf. on Machine Learning and Applications (ICMLA)*, 2018, pp. 1394–1401, doi: 10.1109/ICMLA.2018.00227.

[9] Y. Ouyang, X. Ding, and T. Qin, "Forecasting stock market trends by analyzing economic and financial news," *Int. J. Forecasting*, vol. 36, no. 2, pp. 718–734, 2020, doi: 10.1016/j.ijforecast.2019.06.007.

[10] D. Shah, K. Isah, and F. Zulkernine, "Stock market analysis: A review and taxonomy of prediction techniques," *Int. J. Financial Studies*, vol. 7, no. 2, pp. 1–22, 2019, doi: 10.3390/ijfs7020026.

[11] A. Shroff, "Integrating Innovation and Virtue: Strategic Responses of Private Organizations Amidst Environmental Shocks," *FIIB Business Review*, vol. 14, no. 3, pp. 272–287, 2024. https://doi.org/10.1177/23197145241304274

[12] V. Ratten, "Artificial Intelligence, Digital Trends and Globalization: Future Research Trends," *FIIB Business Review*, vol. 13, no. 3, pp. 286–293, 2024. https://doi.org/10.1177/23197145231222774

[13] S. Maheshwari and N. N. Chatnani, "Applications of Artificial Intelligence and Machine Learning-based Supervisory Technology in Financial Markets Surveillance: A Review of Literature," *FIIB Business Review*, 2023. https://doi.org/10.1177/23197145231189990

[14] M. P. Yadav, A. A. Al-Qudah, K. Sandhu, and N. Gupta, "Resolving an Enigma of FinTech, Digital Assets and Electronic Commerce: Insight to Time-varying Dynamic Connectedness," *FIIB Business Review*, 2025. https://doi.org/10.1177/23197145241300899

[15] B. Yang and S. Zhu, "Public funds in high-tech industries: A blessing or a curse," *Socio-Economic Planning Sciences*, vol. 83, 101037, 2022.

[16] D. Bahuguna, R. Tiwari, C. Prakash, and B. Dhyani, "AI, Financial Technology and Sustainable Investment in ESG Funds in India: A Bibliometric Exploration," in *Proc. 2nd Int. Conf. on Emerging Trends in Engineering and Medical Sciences (ICETEMS)*, IEEE, 2024, pp. 9–13.