Real-time forecasting of FTSE/JSE-top40 using deep neural models: GPT-SNN-PPO vs. LSTM

Claris Shoko[1*], Ntebogang Moroke[2], Caston Sigauke[3], and Katleho Makatjane[4]

To cite this article:
Shoko, C., Moroke, N., Sigauke, C., & Makatjane, K. (2026). Real-time forecasting of FTSE/JSE-top40 using deep neural models: GPT-SNN-PPO vs. LSTM.. Romanian Journal of Economics, 62 (1), 28-44.

Abstract. Background: *Stock market data is characterised by high volatility, non-linearity, and sensitivity to various factors, such as trading volumes and investor sentiment. Time series models, such as GARCH and ARIMA, fail to capture the non-linear and intricate patterns in stock market data. Deep learning algorithms, particularly LSTM models, have demonstrated the ability to predict complex stock market data in recent years, including the South African TSE/JSE Top 40 index, which comprises the top 40 companies on the Johannesburg Stock Exchange.* ***Objective:*** *This study evaluates the efficacy of the GPT-SNN-PPO and the LSTM models in stock price forecasting using the FTSE/JSE Top 40 South African stock market index. The core of the research lies in evaluating which model best adapts to the dynamic, volatile nature of real-time stock market data. The GPT-SNN-PPO model suggested in this study combines elements of a Generative Pre-trained Transformer (GPT) for feature extraction, a Spiking Neural Network (SNN) for processing, and a Proximal Policy Optimisation (PPO) algorithm for reinforcement learning. This combination aims to leverage the strengths of different neural network paradigms to improve real-time forecasting accuracy. Performance evaluation of the two architectures was conducted using the MSE, MAE, and MFE.* ***Results:*** *The results show the superiority of the LSTM model over the GPT-SNN-PPO architecture across short-, medium-, and long-term forecasting horizons, both in-sample and out-of-sample. Based on the LSTM model, a win was predicted for the 7th of July 2025, a loss on the 15th of August 2025, and a win on the 1st of September 2025. This study's results support the literature, recommending the LSTM as the best architecture for predicting and forecasting stock markets.* ***Originality:*** *The novelty of the study is in the in-sample and out-of-sample validation of the LSTM model at different horizons to enhance the generalisability of the findings. The study is also the first to propose the GPT-SNN-PPO architecture, which, although inferior, performs almost as well as the LSTM.*

Keywords: *ES, FTSE/JSE Top 40, GPT-SNN-PPO, LSTM, VaR.*
JEL classification: *C58*

1. Introduction

The domestic economy's structure makes it difficult to predict changes in South Africa's FTSE/JSE Top 40 Index. The index is quite sensitive to changes in global commodity prices and the South African rand's value (Makatjane and Moroke, 2022), as more than 65% of the firms in the index operate in the mining and financial services industries. Local factors make these global pressures much

[1] University of Botswana, Gaborone, Botswana; *Corresponding author: shokoc@ub.ac.bw.
[2] North West University, Mmabatho, South Africa; ntebo.moroke@nwu.ac.za.
[3] University of Venda, Thohoyandou, South Africa; caston.sigauke@univen.ac.za.
[4] University of the Western Cape, South Africa; kmakatjane@uwc.ac.za.

worse. For example, Eskom's load shedding causes frequent power disruptions, and recurrent sovereign rating downgrades prompt sudden shifts in investor sentiment, triggering waves of anxiety and reactive trading. The "long memory" characteristics of such market shocks, as shown by Chikobvu and Sigauke (2020), are characterised by persistent volatility. Unlike what short-memory models would predict, volatility in these cases typically exhibits a hyperbolic decline, as first noted by Bollerslev (1988). This problem has persisted for a long time, a sign of systemic financial stress that threatens the whole basis of conventional econometric models.

According to Makatjane and Moroke (2022), the economy of South Africa is a prime example of a compound vulnerable state characterised by interconnected vulnerabilities to commodity-cycle amplification, infrastructure-induced economic discontinuities, and liquidity flow volatility. Methods for making predictions, particularly those based on linear time series such as the Autoregressive Integrated Moving Average (ARIMA) and Generalised Autoregressive Conditional Heteroscedasticity (GARCH) models, have serious shortcomings when applied to this data. The lack of computing capability is not the source of these limitations; rather, they stem from a methodological mismatch with economies functioning in permanent disequilibrium. These models fail in these settings due to nonlinear feedback loops driven by both internal and external shocks. For example, the rand may have a short-term boost in value due to higher export profits if mining output were to experience a boom. The economic advantages may be offset by inflationary supply-chain disruptions and a drop in productivity when this surge occurs simultaneously with severe infrastructure breakdowns, such as Eskom's Stage 6 load-shedding. Conventional models consistently fail to recognise these multidimensional dynamics for what they are: statistical outliers, not signs of profound structural interactions (Dufrenot and Matsuki, 2021).

This erroneous assumption reveals two serious problems with the modelling process. In disequilibrium economies, the first one is dynamic misspecification. The mediating role of energy restrictions in the link between commodity prices and currency strength has not been well parameterised. For instance, electricity shortages might cause industrial bottlenecks, which can turn profits from the mining industry into inflationary pressure, a transmission mechanism that variance-based techniques do not consider. Regime blindness is the second. Standard elasticity equations may no longer apply when economic systems undergo phase transitions due to simultaneous shocks, such as a commodities boom and a power grid breakdown. The problem is that old models like ARIMA still rely on the idea of historical mean reversion, leading them to misinterpret structural breaks as temporary volatility (Chikobvu and Sigauke 2020). The GARCH model (Engle, 1982) assumes that parameters remain constant over time and that shocks do not persist, which helps it capture uneven volatility patterns. When applied to repeated structural breakdowns, these assumptions make the model useless. For instance, GARCH-type models failed to account for the enduring volatility during the 2019 sovereign rating downgrade. Chikobvu and Sigauke (2020) found that these models treated the persistent market repricing as temporary, mean-reverting swings driven by chronically elevated sovereign risk premia. The fact that models trained on stable market circumstances missed the fundamental shift in risk environment demonstrates a serious issue with regime identification.

Nonetheless, the COVID-19 pandemic showed that standard forecasting systems cannot handle simultaneous systemic shocks. Forecast errors ranged from 22% to 37% across key indices when commodity prices fell with currency devaluation and energy supply interruptions (Djankov and Panizza, 2022). These errors surged because typical modelling structures evaluated each failure as a single event rather than its cascading and interrelated nature. Energy shortages exacerbated supply constraints on commodities, weakening the rand. Dynamic misspecification is a methodological blind spot that leads to prediction failures because models cannot adjust when a structural breakdown occurs. Specifically, traditional models lack the tools to adjust how variables are related to one another when regimes change, such as when a commodity transitions from a downturn to a supercycle. They also do not react strongly enough when the market structure changes after a political or infrastructure crisis. During times of tectonic economic realignment, this rigidity makes econometric models useless. Because of this, people often treat structural breaks as statistical outliers, making it harder to identify volatility regimes and worsening market efficiency and policy effectiveness.

This study evaluates the predictive efficacy of two deep learning architectures, GPT-SNN-PPO and LSTM, for real-time forecasting of the FTSE/JSE Top 40 Index across short-, medium-, and long-term intervals. GPT-SNN-PPO is a hybrid model that uses transformer-based feature extraction, spiking neural networks, and reinforcement learning. It can abstract context, filter events, and adjust to changes in the regime in real time. LSTM models, on the other hand, are effective at finding long-term dependencies and temporal correlations. This makes them useful for simulating steady patterns in financial time series. The research assesses which model most precisely depicts the intricate and fluctuating dynamics of South Africa’s Top 40 market, using both historical and out-of-sample data over several periods in real time.

The GPT-SNN-PPO architecture in this study is based on the idea of combining three distinct inductive biases to address the challenges of predicting financial time series when the market is volatile and changes regimes. The GPT part works as a high-capacity context compressor, extracting long-range connections and complex feature interactions from past market data. The model's self-attention mechanism enables it to identify structural breaks and long-term memory effects that recurrent models often miss. The Spiking Neural Network layer brings in an event-driven representation, where membrane potentials and threshold-based spiking filter temporal input. This mechanism allows for temporal sparsity, is resistant to noise bursts, and is sensitive to unexpected changes. These are all very useful in financial settings where shocks occur frequently; moreover, the addition of Proximal Policy Optimisation enables real-time adaptation and improves both forecasting accuracy and spike dynamics. Instead of limiting reinforcement learning to trading actions that occur after the fact, our PPO’s structure stabilises updates when they are not stationary by automatically adjusting both the forecasting policy and SNN parameters, such as thresholds and leak rates. The resulting system thus produces a coherent pipeline in which GPT reduces representation error, SNN selects salient events, and PPO enables robust adaptation in the face of distributional drift.

This architecture differs from standard hybrid models, which typically have only two components. Transformer-based reinforcement learning (TRF) models combine sequence modelling and policy optimisation, although reinforcement learning is often limited to trade choices; forecasting remains supervised and based on continuous activations. Transformer-SNN models enhance temporal sparsity but lack automatic adaptation to changing data distributions, which reduces stability during regime shifts. In contrast, SNN-RL models emphasise event-driven adaptations but struggle to capture long-term dependencies, limiting their effectiveness for financial series with consistent trends. Our GPT-SNN-PPO paradigm addresses these problems by combining context abstraction, event-driven filtering, and online adaptation into a unified architecture. This three-way integration tackles the three fundamental problems of traditional approaches: inaccurate characterisation of complicated relationships, failure to detect structural fractures, and delayed response to market changes. Unlike simple method stacks, GPT-SNN-PPO is a theoretically grounded architecture that can address multiple difficulties simultaneously, whereas two-component hybrids cannot. In a nutshell, the following hypothesis is tested H_a: The GPT-SNN-PPO model provides significantly more accurate forecasts than the LSTM model across one or more forecasting horizons.

2. Literature review

Empirical Deficiencies in Volatility Regime Recognition

The foundational GARCH outline (Engle, 1982) is significant for modelling asymmetric volatility clustering, but it relies on the assumptions of parameter stationarity and shock transience. These assumptions make it insufficient to address persistent structural breaks. An illustrative example is the 2019 sovereign credit downgrade, in which GARCH-type models fundamentally misjudged volatility's persistence. These models saw the lasting changes in the market, caused by a permanent rise in sovereign risk premia, as just temporary ups and downs that would go back to normal. Such an error indicates a significant shortcoming in regime identification: frameworks intended for stable financial systems failed to acknowledge the structural changes in South Africa's risk environment (Al-Najjar 2016). This deficiency was further intensified during the COVID-19 crisis. The pandemic resulted in concurrent declines in commodity markets, significant currency devaluation, and substantial energy disruptions caused by Eskom. Standard models exacerbated forecasting errors during these concurrent shocks, with errors ranging from 22% to 37% across major indices (Djankov and Panizza, 2022). This divergence occurred because conventional frameworks treated these interconnected events as independent occurrences, rather than recognising their cascading nature. Electricity shortages exacerbated commodity supply bottlenecks, subsequently increasing currency volatility—feedback loops that conventional models consistently neglected.

Operational Distortions from Model Misspecification

Inaccuracies in forecasting directly affect policy and market operations. Resource misallocation represents a significant distortion. Neglecting infrastructure-commodity interdependencies in ARIMA models, such as the impact of load shedding on supply disruptions in the platinum sector, can lead to misdirected fiscal responses and poorly prioritised stabilisation measures. A further consequence is the mispricing of risk. During the 2021 civil unrest, GARCH-type models notably underestimated the persistence of regime change, leading to a systematic underestimation of equity tail risks (Makatjane, 2024). The results show a major problem: traditional equilibrium-based models do not adequately capture the ongoing instability in emerging economies like South Africa, where capital flight increases due to simultaneous currency and energy crises.

Forecasting models must change to address these constraints. Effective models should explicitly parameterise shock interdependencies, such as commodity prices, currency fluctuations, and infrastructure stability, and include regime-switching processes to identify structural shifts. Models must also enable continual recalibration, perhaps using reinforcement learning methods such as Proximal Policy Optimisation (PPO), which dynamically adjusts to crisis-derived reward signals (Schulman et al., 2017). The evidence supports growing research indicating that context-aware computing is better for triple-exposure forecasting than incremental enhancements to classic econometric models (Davies et al., 2021). This knowledge problem stems from classic models that assume market structure is stable, whereas developing countries are constantly changing due to sudden shifts in commodity prices, currency fluctuations, and supply chain disruptions caused by infrastructure issues.

Gaps in Machine Learning and AI-Based Forecasting

Volatility models like Fractionally Integrated GARCH (FIGARCH) (Baillie et al., 1996) better handle persistent volatility. However, they remain limited because there are no real-world protocols for their use in high-frequency environments such as the Johannesburg Stock Exchange, where millisecond latency affects trading viability. Initially, we praised Long Short-Term Memory (LSTM) networks for their ability to identify sequential relationships in financial data. However, their reliance on historical continuity presents a challenge when the market changes. Elhag (2025) reports that load-shedding announcements led to a 37% increase in prediction error, underscoring how easily sudden changes in the system can throw off these models.

Transformer-based architectures, known for their self-attention and ability to understand context, have pushed the limits of financial forecasting. However, their high computational demands often lead to latency issues, particularly during sudden spikes in U.S. interest rates or fluctuations in commodity prices. In many cases, sub-second responsiveness is needed, but transformer models are often not sufficiently robust for real-time use (Vaswani et al., 2017). This conflict between analytical expertise and operational agility is still a major problem in developing markets. While Proximal Policy Optimisation (PPO), originally designed for robotics and reinforcement learning in game theory, shows promise for adapting policies to changing conditions, its potential for financial forecasting remains underutilised (Schulman et al., 2017). Its ability to adapt to new environments suggests it could address some of the structural rigidity seen in traditional models, though this application is still in its infancy. There is a growing consensus that algorithms alone cannot capture the rhythm and nuance of volatility in emerging markets such as South Africa. What is needed is an integrative approach that combines technical agility with situational awareness. Effective forecasting systems must be able to interpret the underlying drivers of volatility spikes, policy uncertainty, grid instability, investor sentiment, and institutional fragility and translate these insights into timely, actionable intelligence.

3. Methodology and data

This section details the architecture, theoretical foundations, and implementation procedures of the GPT-SNN-PPO and LSTM models used to forecast the FTSE/JSE Top 40 Index in real time. This study uses the daily closing prices of the FTSE/JSE Top 40 Index for the period from 2 January 2012 to 30 June 2025, obtained from the Johannesburg Stock Exchange. The data were pre-processed to handle missing values, and the series was transformed into logarithms to stabilise variance. For model training, the data were further normalised using Min-Max scaling, ensuring that both GPT-SNN-PPO and LSTM models received standardised input suitable for accurate forecasting and were further divided into 80% training and 20% testing.

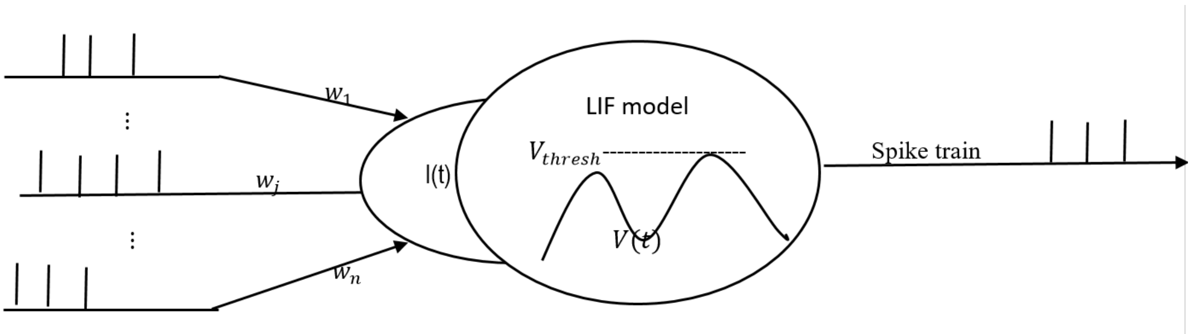
GPT-SNN-PPO architecture

The GPT-SNN-PPO is a more recent approach that combines the strengths of Generative Pre-trained Transformers (GPT), Spiking Neural Networks (SNN), and Proximal Policy Optimisation (PPO). The GPT models can be used for various tasks, including generating deep learning models for financial time series. On the other hand, SNN offers advantages in terms of energy efficiency and computational speed. While combining GPT could be used to generate and optimise the SNN, PPO helps the SNN learn the behaviour of complex environments, such as financial markets. Thus, combining these three offers a more robust and adaptive approach to real-time forecasting of the FTSE/JSE Top 40.

Spike Neural Network model: Leaky Integrate and Fire (LIF) model

Various spiking-neurone models clarify changes in cell membrane potential and the mechanisms underlying pulse generation, with varying levels of specificity (Abusnaina and Abdullah, 2014). The Integrated firing (I and F) model was proposed by Lapique et al. (1907), as cited by Gerstner and Kistler (2002), where the theoretical idea of I and F neurones suggests that the dynamic behaviour of biological neurone membrane potential should show three traits: leakage, accumulation, and threshold excitation. Therefore, the I and F neurones do not fully align with the biological principles. The I and F model is a simplified version of the Leaky Integrate-and-Fire (LIF) model; thus, the LIF model maintains the three properties while simplifying the action potential generation process, reducing computational complexity, and improving biological interpretation (Qin et al., 2025). However, Figure 1 displays the architecture of the LIF model, and Wu et al. (2019) describe its membrane potential dynamics using the differential equation in Equation (1).

Figure 1. Schematic diagram of the LIF model



Source: authors.

$$\begin{cases} \tau \dfrac{dV(t)}{dt} = -V(t) + rI(t) \\ s(t) = 1, \text{ if } V(t) > V_{threshold} \\ s(t) = 0, \text{ otherwise} \end{cases}, \quad \textbf{(1)}$$

where τ is the time constant or leaky integrator that determines the rate at which the membrane potential changes, $V(t)$ is the membrane potential of the neuron at time t, s represents the spike train, and r is the membrane resistance. The firing of a spike is triggered when the membrane potential $V(t)$ exceeds the threshold $V_{threshold}$. I(t) represents the sum of the current supplied from the j^{th} neuron, which, according to Karimah et al. (2025), is given as

$$I(t) = \sum_j w_{ij} \sum_f \delta(t - t_j^f), \quad \textbf{(2)}$$

where w_{ij} is the weight between the j^{th} and the i^{th} neuron, δ is the Dirac function, and t_j^f is the firing time of the neuron j. The dynamics of the neurons in the i^{th} hidden layer is updated by the discrete version of the LIF model as follows

$$\begin{cases} V^i(t+1) = \beta V^i(t) + \omega^i s^{i-1}(t) - V_{threshold} s^i(t) \\ s_j^i(t) = 1, \text{ if } V_j^i(t) > V_{threshold} \\ s_j^i(t) = 0, \text{ otherwise} \end{cases}, \quad \textbf{(3)}$$

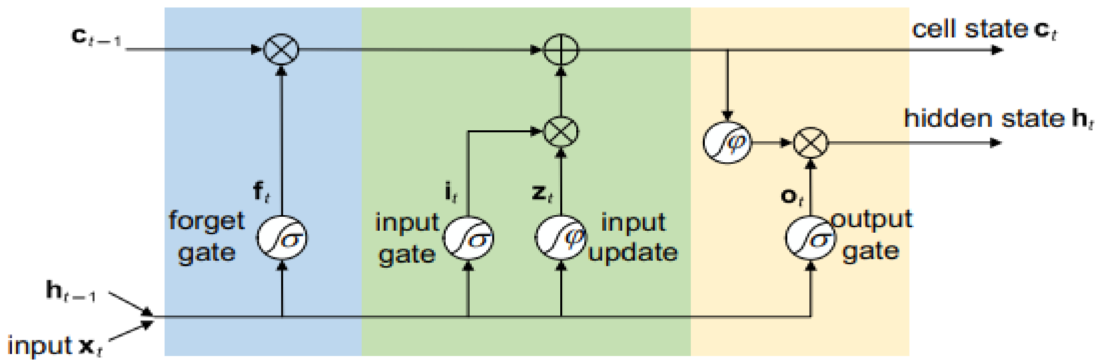
where $V^i = \left[v_1^j, v_2^j, \ldots, v_n^j\right]^T$ is the vector of the neurons of the i^{th} layer $S^i = \left[s_1^j, \ldots, s_n^j\right]^T$ is the train vector of the i^{th} layer, and $\omega^i = \left\{\omega_{jk}^j\right\}$ is the weight matrix from $(i-1)^{th}$ layer and $1-\beta$ represents the leaky rate. Because time series forecasting involves many time steps, the LIF cannot keep information from the past for very long. Therefore, to mitigate this problem, the Recurrent LIF (RLIF) is used, which is based on a recurrent neural network (RNN) with a LIF-based feedback loop. With respect to Qin et al. (2025), the RLFI enables the network to leverage associations across different time steps to predict what will happen at the current time step.

Long-short-term memory forecasting

Long short-term memory networks (LSTMs) are a form of RNN that are successful at learning long-term dependencies (Malashin et al., 2024), and Figure 2 shows an example of an LSTM. Memory cell gates are functional modules found in the memory block, which is a self-connected subnet. The gates of the multiplicative units control the flow of information, while the memory cell monitors the status of the neural network over time. Notably, there are three different kinds of gates: forget gates, output gates, and input gates. Ehteram et al. (2024) state that the input gate regulates the amount of new information that enters the memory cell, the forget gate regulates the amount of information that remains in the memory cell through recurrent connections, and the output gate regulates the amount of information that is used to determine the output activation of the memory block before flowing into the neural network. The commonly used activation functions for LSTMs are the sigmoid function $\sigma(x) = \frac{1}{1+\exp-x}$, such that $0 \le \sigma(x) \le 1$, and the tanh function $\varphi(x) = 2\sigma(2x) - 1$, such that, $-1 \le \varphi(x) \le 1$.

The operation of the LSTM is divided into two steps: Step 1: a sigmoid layer (input gate), which determines what should be updated, and a $\tanh$ layer that creates new candidate values of Z_t that are added to the memory cell state for the current moment at time t. In Step 2, these two parts are combined to trigger an update of the memory cell state to the new cell state C_t. This is done by multiplying the corresponding elements of c_{t-1} and the output of the forget gate layer (f_t), and then adding $i_t * Z_t$, where i_t denotes the output of the input gate. The last step is deciding what to output by multiplying the value obtained from a $\tanh$ function and the output of a sigmoid layer O_t. The cooperation of the memory cell and the gates has endowed the LSTM with a powerful ability to predict time series with long-term dependencies. The schematic diagram for the LSTM architecture is presented in Figure 2.

Figure 2. Long-short-term forecasting



Source: Hua et al., 2019.

Based on the work by Silva et al. (2023), we present a summary of the comparison of the SNN versus the LSTM in Table 1 below.

Table 1. A comparison of the SNNs versus LSTM

Feature	Spiking Neural Networks (SNNs)	Long Short-Term Memory (LSTM)
Architecture	Biologically inspired, event-driven, uses spikes	Recurrent neural network with memory cells
Computational efficiency	Potentially more energy-efficient due to sparse spiking activity	Can be computationally intensive, especially with large datasets and architectures
Learning mechanism	Spike-based learning, synaptic plasticity	Gradient descent, backpropagation
Time series dependency handling	Can capture temporal dependencies, but may require specialised architectures.	Excellent at capturing long-term dependencies in sequential data
Training complexity	It can be challenging to optimise hyperparameters and train effectively	Well-established training procedures, but still sensitive to hyperparameters
Real-world applications	Emerging field, promising in signal processing and neuromorphic computing	Widely used in forecasting tasks (e.g., stock markets, weather)
Ease of implementation	A relatively new field with fewer tools and libraries	Widely supported in TensorFlow, PyTorch, and other frameworks
Typical data representation	Uses spikes (discrete events)	Uses continuous values (floating-point numbers)

Source: Compilation from Silva et al. (2023).

Error metrics

The efficacy of the SNN and LSTM architectures across the 5, 24, and 60-day timeframes is assessed using the mean square error (MSE), mean prediction error (MFE), and mean absolute error (MAE). The Mean Absolute Error (MAE) is the preferred error measure due to its reduced sensitivity to outliers in comparison to the Mean Squared Error (MSE) and Mean Forecast Error (MFE); and it is suitable in this study because the time series data for the closing price of FTSE/JSE-Top40 exhibits significant tails, primarily attributable to the COVID-19 epidemic. The pandemic had substantial repercussions for finance; therefore, these outliers must be acknowledged, making MAE the most suitable statistic for performance evaluation as documented by Khan et al (2022). The Mean Squared Error (MSE) represents the average of the squared deviations of the predicted values from the actual values and magnifies them by squaring, making it advantageous when substantial errors are very unwelcome. Finally, the Mean Forecast Error is a statistical measure employed to assess the overall bias of a forecasting technique. Xian et al (2025) document that the MFEt is determined by averaging the forecast errors across a series of time intervals or data points. The MAE, MSE, and MFE are defined by the following formulas, respectively.

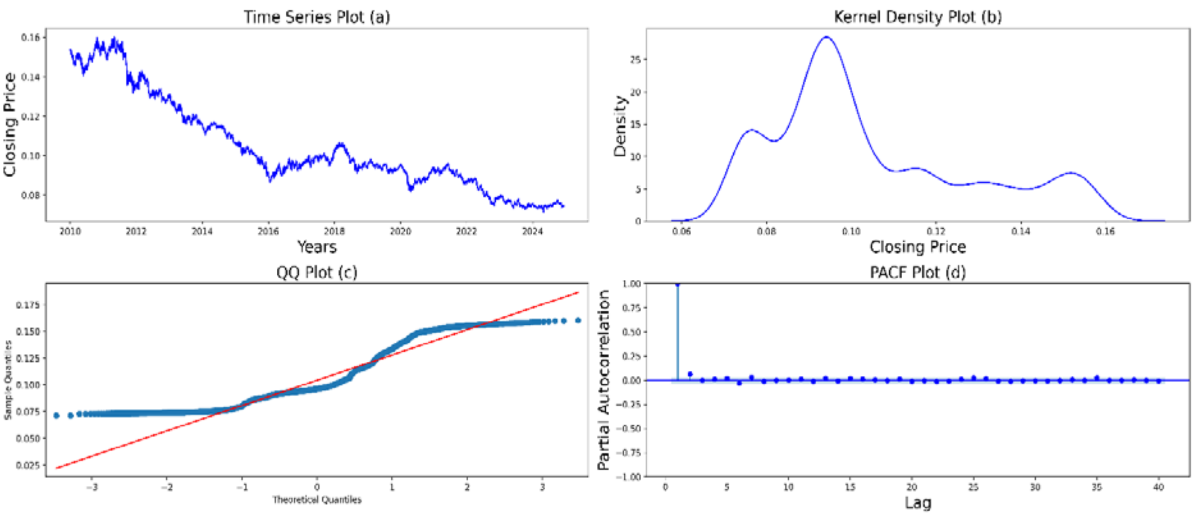
$$\text{MAE} = \frac{1}{n}\sum_{t=1}^{n}| y_t - \hat{y}_t |, \quad \text{MSE}=\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2, \quad \text{MFE}=\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)$$

where n is the number of data points, y_t is the actual value, and $\hat{y}_t$ is the predeicted value.

4. Research results

Figure 3 presents an extensive exploratory study of the closing price of the FTSE/JSE Top 40. Especially around 2020, the time series plot (a) clearly shows periods of high and low volatility, as well as sudden changes, indicating that the variance is not constant and the data are not stable over time. The multi-modal distribution reveals structural shifts and departures from normality, while the kernel density plot (b) suggests potential fat tails. The non-normality of the data is further confirmed by the Q-Q plot (c), which shows substantial departures from the predicted normal distribution at both tails, typical of financial markets. At lag 1 and 2, the PACF Figure 3 (d) displays notable partial autocorrelations, implying the existence of short-term memory models. The series shows temporal dependence, nonlinearity, and heteroscedasticity, generally suggesting that proper modelling and forecasting require sophisticated models.

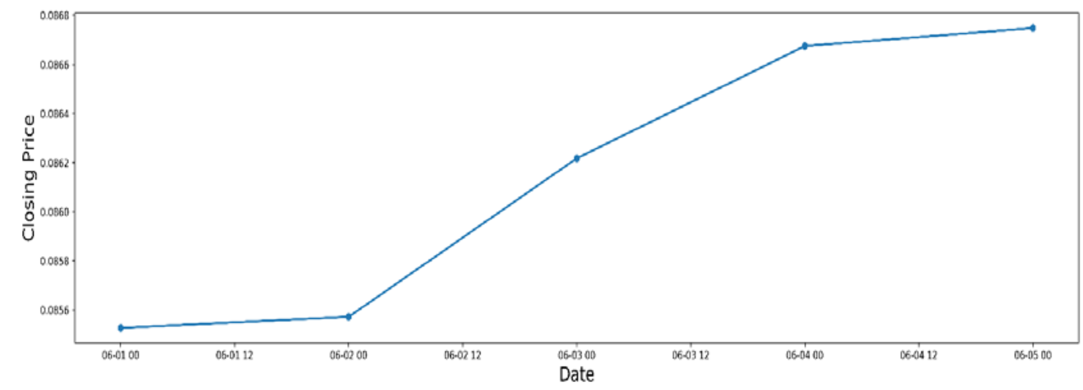
Figure 3. Univariate data analysis for the Botswana Company Index



Trend analysis results

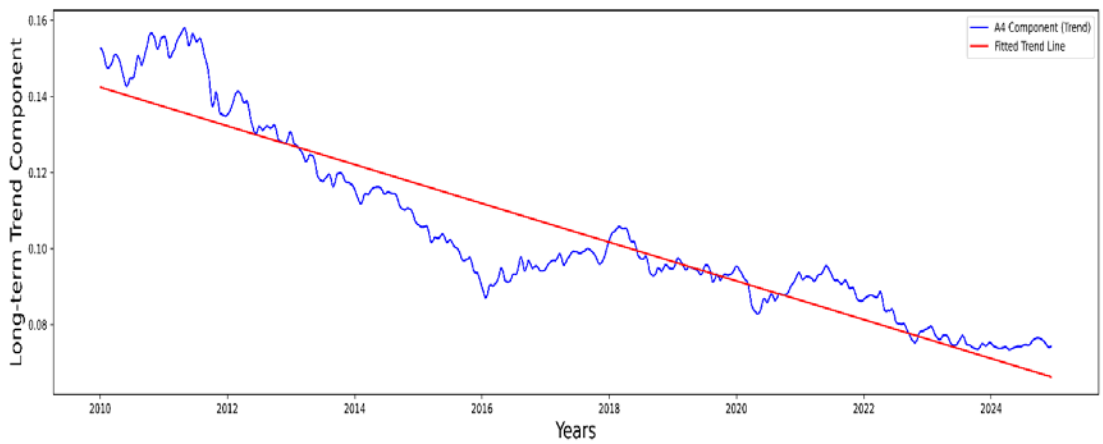
Financial time series data clearly shows trends in weeks, months, quarters, etc. All exchanges are closed on weekends and holidays, so trading is not allowed. On weekdays, activity is high, and Figure 4 shows the day-of-the-week effect on the FTSE/JSE Top 40. Monday usually marks the lowest level of market activity; the final two working days of the week typically show the greatest level. Trade activity gradually rises Monday through Friday, and the lack of trading on weekends and holidays results in low opening stock prices on Monday. The weekend effect—also known as the Monday effect, the day-of-the-week influence, or the Monday seasonal effect, describes how stocks often do better on Friday than on Monday (Strohsal et al., 2019), hence the highest value of the closing price of FTSE/JSE-Top40 on Friday as seen in Figure 4. These results contradict those of Makatjane and Moroke (2022), who found that the Friday values were smallest in the FTSE/JSE All Share index they used.

Figure 4. Day of the week pattern



To significantly capture the trend in the series, we fit a discrete wavelet transform, and the decomposition results for the FTSE/JSE-Top40 closing prices reveal that the long-term trend component (A4) dominates with a value of 71,983.5800, indicating that the overall price movement is primarily driven by low-frequency, structural trends influenced by macroeconomic factors. The detailed components D4 (1.7726), D3 (0.8068), D2 (0.4548), and D1 (0.3281) progressively capture medium- to very high-frequency fluctuations, representing short-term volatility, cyclical variations, and noise. Their relatively low magnitudes compared to A4 suggest that short-term shocks, speculative jumps, and high-frequency noise contribute minimally to the overall price dynamics. This implies that FTSE/JSE-Top40 prices during the observed period exhibit trend-dominated, relatively stable behaviour, with limited influence from short-term volatility clustering or irregular disturbances. Hence, the process exhibits no serial correlation, as seen in Figure 3(d). To accurately estimate the overall trend, a simple linear regression model is fitted, and we find the trend coefficient to be -0.1761, indicating a long-run decrease in closing prices, as seen in Figure 5.

Figure 5. Wavelet trend decomposition from 2010 to 2025



Short-, medium- and long-term forecasting with GPT-SNN-PPO VS long-short term memory

To begin the analysis, each architecture is trained with a mini-batch size of 32, stochastic gradient descent (SGD), a learning rate of 0.01, and a momentum of 0.9. Early stopping is implemented with a patience range of 25 to 75 epochs and a minimum absolute improvement between and 10^{-6}, with a maximum of 1000 training epochs. We therefore begin with the GPTSNNForecaster, a proprietary neural network model that combines Transformer architectures with spiking neural networks for time-series forecasting. An embedding linear layer translates raw features into a higher-dimensional space (d-model) to embed the input sequence, and a multi-layer Transformer encoder captures temporal relationships using self-attention on this embedded sequence. The model processes the latest time step representation output of the Transformer using a Leaky Integrate-and-Fire (LIF) neuron from the Norse library, as described by Rast et al. (2011). We produce the final forecast by passing the spiking output through a linear layer, where the Transformers' sequential learning and spiking neurons' energy-efficient, dynamic processing make this combined model a revolutionary forecasting methodology.

When combining GPT-SNN with Proximal Policy Optimisation (PPO), the training results indicate that our PPO agent is progressing through a stable but slow learning phase. The average episode length (ep_len_mean) of approximately 1,910 steps indicates that the agent consistently completes long episodes without premature termination, suggesting it navigates the environment without critical failures. However, the negative average episode reward of -666 indicates that the agent is still accumulating penalties or performing poorly under the task's reward structure. The training loss is relatively low at 2.87, and the value loss of 7.77 suggests that the value network is learning expected returns but is not yet very accurate, as confirmed by the near-zero explained variance, indicating that the critic's value estimates do not match actual outcomes. The policy gradient loss is nearly zero (–8.17e–05), meaning that the policy is changing very little with each update, and this is reinforced by the very low KL divergence (approx_kl), indicating almost no divergence from the old policy. Additionally, a zero-clip fraction implies that policy updates are well within the clipping range—a sign of conservative updates. The entropy loss has reduced to -1.14, indicating that the policy is becoming more deterministic and less exploratory, as expected, as learning stabilises. However, the low frames per second (fps) of 11 indicates slow training speed, likely due to computational constraints such as using a central processing unit (CPU) or complex environment computations. The performance results are reported in Table 2.

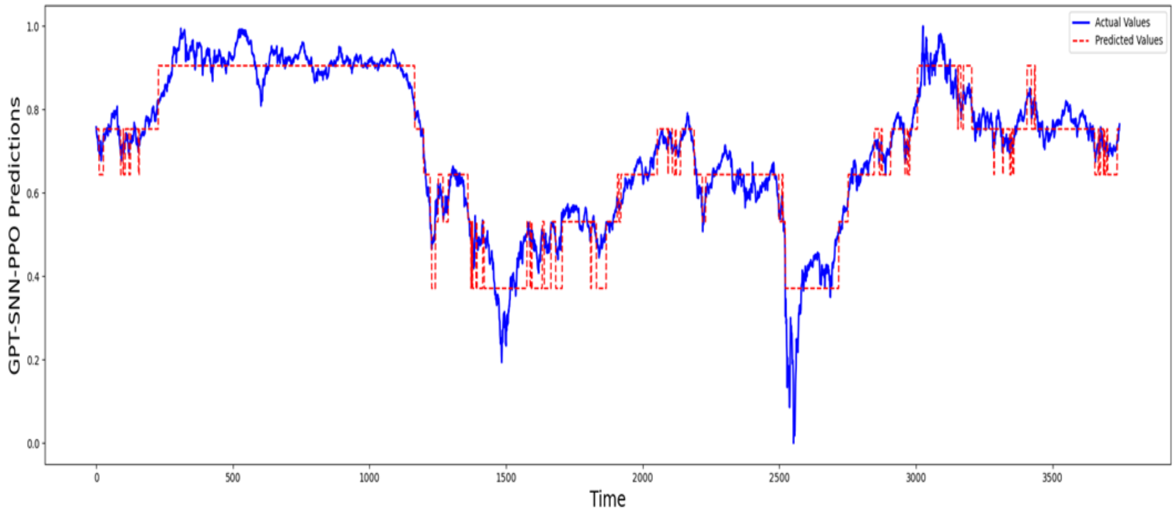
Table 2. Forecasting performance of GPT-SNN-PPO

Architecture	Horizon	MSE	MFE	MAE
GPT-SNN-PPO	5	0.0470	0.0041	0.0008
	24	0.0467	0.0043	0.0007
	60	0.0447	0.0038	0.0009

Low and steady error measurements show that the GPT-SNN-PPO architecture has strong predictive performance across short-, medium-, and long-term horizons. The model produces reliable, essentially objective short-term forecasts, achieving a mean squared error of 0.0470, a mean forecast error of 0.0041, and a mean absolute error of 0.0008 at a 5-step horizon. The MSE improves slightly to 0.0467 when the forecasting horizon spans 24 steps; the MFE rises slightly to 0.0043; and the MAE drops to 0.0007, indicating a slight tendency to underestimation. Hence, the model demonstrates its ability to generalise effectively beyond the short term, maintaining its predictive accuracy across a longer horizon.

With the lowest recorded MSE of 0.0447 and the MFE and MAE stabilising at 0.0038 and 0.0009, respectively, the model's performance continues to improve at a 60-step horizon. Given that forecasting algorithms seldom maintain or improve accuracy over longer periods, these outcomes are very noteworthy because the slight decrease in bias and the continuous low absolute errors at long horizons imply that the GPT-SNN-PPO architecture is competent at capturing underlying temporal dynamics and long-range interdependence within the data. These results generally show the model's strong generalisation capacity and dependability for both short- and long-term risk forecasting. To assess the model's overall performance, Figure 6 compares the predicted values with the original Top40 index. Although the predicted series closely follows the general trend of the original data, the model does not fully capture the series' underlying dynamics. This is the consequence of relatively low policy learning, persistently negative reward signals, and a stable yet modest learning rate observed during training. Collectively, these indicators suggest that while the model approximates the trend, it falls short of accurately reproducing the finer patterns of the original series.

Figure 6. Fitted conditional mean for GPT-SNN-PPO architecture



Part of the observed discrepancy between the model's forecasts and the real Top40 index is ascribed to external structural factors not included in the model, which often introduce non-linearities and regime changes not captured by a simple price-based model. Moreover, macroeconomic factors such as interest rates, inflation, gross domestic product (GDP) growth, and global demand shocks greatly affect stock prices. Finally, influencing rapid market responses and volatility spikes intrinsically difficult to predict with historical price data alone are geopolitical events such as wars in areas of oil production, sanctions, and trade disputes. Finally, the model is also influenced by news cycles, risk aversion, and financial market expectations; investor mood and speculative behaviour also cause price swings that depart from basic value. Such features of unpredictability compromise the model's generalising capacity but help us explain the complexity and noise in the oil market. Thus, by combining sentiment analysis, macroeconomic data, and event-driven characteristics, future models will be more accurate.

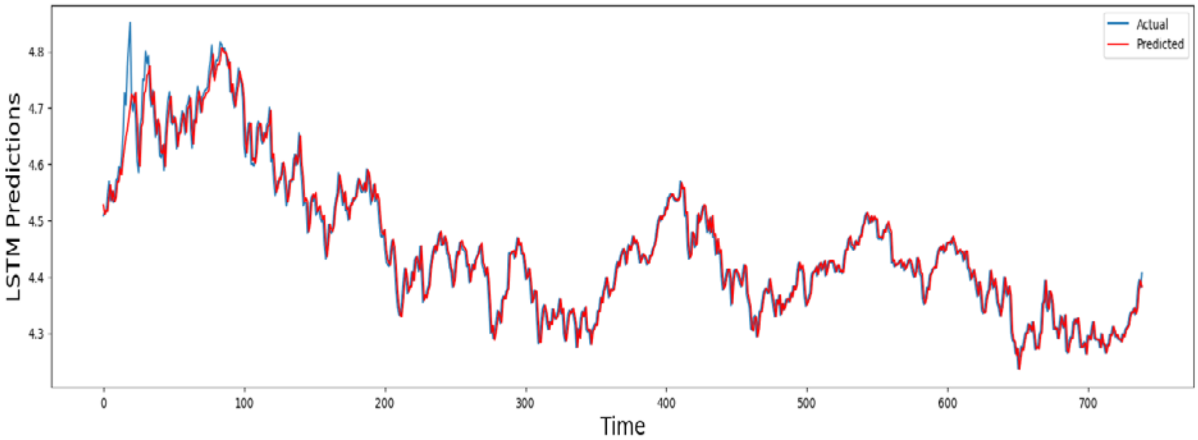
The performance of the LSTM model across three forecasting horizons, 5-, 24-, and 60-time steps, demonstrates relatively strong and consistent predictive capability. At a short horizon of 5 steps, the model achieves an MSE of 0.0231, an MFE of 0.0009, and an MAE of 0.0183. These metrics suggest that the model closely tracks short-term fluctuations in the data, with a slight positive bias. As the forecast horizon increases to 24 and 60 steps, the MSE improves slightly to 0.0186 and 0.0197, respectively, while the MFE remains stable at 0.0006, indicating consistent directional accuracy over longer horizons. These results are reported in Table 3. Interestingly, the MAE improves at the 24-step (0.0055) and 60-step (0.00082) horizons, suggesting a modest average error in estimating the true values. However, these values remain close to zero, reinforcing the model's stability and precision in longer-term forecasting. Overall, the LSTM exhibits reliable predictive performance, with minimal directional bias and robust generalisation across varying forecast intervals.

Table 3. The forecasting performance of the LSTM

Architecture	Horizon	MSE	MFE	MAE
LSTM	5	0.0231	0.0009	0.0183
	24	0.0186	0.0006	0.0055
	60	0.0197	0.0006	0.00082

The fitted series closely aligns with the actual FTSE/JSE Top 40 series when using the LSTM architecture, as illustrated in Figure 7. This suggests that the LSTM model can capture the underlying temporal patterns in the data more effectively than alternative approaches.

Figure 7. Fitted conditional mean for LSTM architecture



Comparative analysis of GPT-SNN-PPO and LSTM architectures

Predictive performance, temporal alignment, and resilience across many prediction horizons show considerable variations between the GPT-SNN-PPO and LSTM designs. Demonstrating a modest capacity to follow the overall trend of the Top40 index, the GPT-SNN-PPO model merges a transformer-based GPT backbone with spiking neural dynamics and is optimised using the Proximal Policy Optimisation method. Low policy gradient values, flat entropy loss, and negative mean rewards during training, however, indicate unsatisfactory policy convergence in learning terms. These metrics suggest underperformance in capturing the complex volatility and nonlinearity inherent in oil market data, as well as inadequate investigation. Although directionally consistent, the fitted series differs from the real series, especially during periods of strong market swings.

In contrast, the LSTM model exhibits superior performance in mimicking the actual series, as illustrated in Figure 7. This architecture excels at learning long-term dependencies and temporal correlations, which is particularly advantageous when modelling sequential financial time series. The LSTM model shows better results in predicting the actual data, as shown in Figure 7 and also has lower MSE and MFE for all periods tested, short-term (5), medium-term (24), and long-term (60) as compared to the GPT-SNN-PPO. The close match between the predicted and actual values shows that the LSTM model is better at understanding both trends and fluctuations compared to the GPT-SNN-PPO model; hence, this achievement is attributed to the LSTM's ability to retain memory of past observations over time, while the GPT-SNN-PPO model, despite its theoretical advantages in attention-based sequence modelling, may require more extensive tuning, data augmentation, or hybrid input features (e.g., macroeconomic or sentiment data) to match the LSTM's performance in this context.

Tables 2 and 3 illustrate the forecasting performance of the GPT-SNN-PPO and LSTM models, respectively, at various horizons (5, 24, and 60 steps ahead). The GPT-SNN-PPO model exhibits a Mean Absolute Error (MAE) between 0.0007 and 0.0009. In contrast, the LSTM model shows a higher MAE at shorter horizons, specifically 0.0183 at 5 steps and 0.0055 at 24 steps, which significantly decreases to 0.00082 at the 60-step horizon. The results demonstrate that both models attain a significant level of predictive accuracy, especially at extended time horizons. The direct conversion of these forecast errors into predictive risk for market participants reflects their economic significance. The Mean absolute error of 0.0008 indicates that, on average, the model's forecast diverges by merely a small fraction from the actual closing value of the FTSE/JSE Top 40 index. This data indicates several implications for investors. Initially, it demonstrates high accuracy, thereby reducing the likelihood of substantial mispricing or incorrect trading signals. The low MAE diminishes predictive risk, which is particularly significant in algorithmic or high-frequency trading contexts, where minor forecast errors can aggregate into considerable financial repercussions. The reliability indicated by low error values enhances the credibility of the models as decision-support tools in portfolio management and risk assessment. Both architectures provide highly accurate forecasts; however, the Mean Forecast Error (MFE) values indicate small but persistent directional biases: GPT-SNN-PPO shows a positive bias, and LSTM shows a bias close to zero. From an economic perspective, even minor biases can impact systematic trading decisions by causing slight over- or under-estimations of upward movements. Consequently, while the very low MAE values indicate high forecasting reliability, investors and analysts should consider them alongside bias measures to avoid overconfidence in the models' outputs.

Out-of-sample forecasting performance of the LSTM in real time

The LSTM forecasting model demonstrates strong real-time predictive capabilities for the JSE Top 40 Index by utilising recent historical closing prices, which are log-transformed and normalised using a MinMaxScaler to ensure numerical stability and facilitate model convergence. Leveraging a recurrent neural network architecture, the model effectively captures temporal dependencies within a 120-day historical window, enabling accurate short-, medium-, and long-term forecasts. Evaluation metrics across various prediction horizons indicate a gradual decline in forecasting accuracy as the horizon increases, with the mean squared error rising from 0.00083 over a 5-day horizon to 0.01898 at 60 days, as shown in Table 4. Correspondingly, the mean absolute error increases from 0.0278 to 0.1188. Notably, the mean forecast error remains consistently negative across all horizons, suggesting a slight tendency of the model to underpredict actual closing prices. The same results of an increase in the errors for the LSTM in an out-of-sample forecasting exercise are also found by Shoko et al (2024).

To operationalise these forecasts, the LSTM architecture is embedded within a Streamlit web application, enabling real-time deployment and interactive forecasting. This interface allows end-users to input the most recent market data and receive immediate predictions, accompanied by 99% prediction intervals and actionable trading signals. Using Streamlit offers several practical advantages, including ease of deployment, an intuitive user interface, and seamless visualisation of results. Its lightweight structure supports rapid prototyping and browser-based accessibility, making forecasting tools available to a broader audience without requiring programming expertise. Despite a slight decline in predictive accuracy over longer horizons, the LSTM model remains highly effective for short- to medium-term forecasting, and its performance could be further enhanced by incorporating external macroeconomic indicators, investor sentiment, and political risk factors, which play a significant role in driving financial market volatility.

Table 4. Out-of-sample forecasting for LSTM in real time

Architecture	Horizon	MSE	MAE	MFE
LSTM	5	0.000830	0.027775	-0.027775
	24	0.008374	0.078646	-0.075148
	60	0.018980	0.118798	-0.095304

To quantify uncertainty, 99% prediction intervals are constructed using the empirical standard deviation of the residuals and the corresponding normal distribution critical value. This approach ensures the forecasts are not only accurate but also probabilistically interpretable, enabling robust, real-time risk management. Makatjane et al. (2025) emphasised that the use of the prediction interval width further helps evaluate the model's stability, especially in predicting extreme market fluctuations and preventing underestimation of potential losses. Moreover, the inclusion of forecast intervals and trading signal visualisations further enhances the model's practical utility, providing a comprehensive and academically sound basis for financial time series forecasting, as shown in Figure 8. In this way, we can alert investors to the likelihood of significant gains when they sell or buy, and vice versa.

The return values reflect the percentage change in the JSE Top 40 index price over the 60 business days following each trading signal. For instance, a return of -3.32% indicates that the index is expected to decline by this amount after a buy signal is issued, meaning the anticipated upward movement is unlikely to materialise. This is also supported by the downward trend seen in Figure 7, specifically in the 99% prediction interval width. Similarly, a return of -1.09% after another buy signal suggests a modest decrease during the holding period. These negative returns following buy signals imply that our LSTM can forecast a downward trend from 20 June 2025 to 11 September 2025, resulting in losses for investors who act on these signals.

Figure 8. An out-of-sample prediction interval width with trading signals

These results are reported in Table 5. In contrast, the negative returns recorded after sell signals (such as -2.17% and -0.88%) correspond to successful predictions of price declines, which will bring profits to those executing short positions or exiting holdings. This demonstrates that the model is more effective at forecasting downward movements in the J403 index over the 60-day horizon. The magnitude of these returns is consistent with moderate market fluctuations typical of a broad equity index. Overall, the results suggest that while the model's sell signals hold practical value for risk management and tactical decision-making, its buy signals require further refinement to improve predictive accuracy in capturing upward market momentum.

Table 5. Expected returns levels on each trading signal

Date	Signal	Returns	Outcome
2025/06/25	1	-0.03324	Loss
2025/07/10	-1	-0.02167	Win
2025/08/15	1	-0.01091	Loss
2025/09/01	-1	-0.00875	Win

Future trading and risk management techniques for the JSE Top 40 index depend heavily on observed trends in predictive accuracy and price swings. The ability of our LSTM to forecast declining trends suggests its potential use as a tool for spotting market declines or periods of higher volatility, thereby allowing investors to take defensive action, such as timely position exits or hedging. The inadequacy in predicting increasing price movements, however, emphasises the necessity of prudence when depending on purchase signals because false-positive or early entry might cause unnecessary losses and capturing growth possibilities and boosting general trading success depend on the sensitivity of the model to signals, indicating a bullish market is expected to improve in the future. To better navigate the complex dynamics of the J403 index, especially during periods of uncertainty or structural change, investors and portfolio managers can also consider adding supplementary metrics or qualitative insights, such as market sentiment analysis and public news from social media or newspapers, to augment the model. Achieving balanced, practical insights that properly guide investment choices in the South African stock market will ultimately depend on improving forecasting techniques.

5. Discussion, limitations, recommendations, and conclusion

This research evaluates the efficacy of two machine learning architectures, GPT–SNN–PPO and LSTM, for predicting the closing price of the FTSE/JSE Top 40, a financial series characterised by temporal dependencies, nonlinearity, and heteroscedasticity. The GPT–SNN architecture, functioning as a spike-based learner, excels at capturing temporal relationships; however, it requires specialised techniques to navigate complex dynamics. To address this, the model is improved using Proximal Policy Optimisation (PPO), yielding the new GPT–SNN–PPO hybrid. The LSTM design, on the other hand, uses memory cells to capture long-term dependencies, making it a suitable choice for time series forecasting. We use MSE, MFE, and MAE to evaluate performance. Prioritising MSE and MFE for the final evaluation is based on their sensitivity to outliers, which is important because COVID-19 shocks create severe spikes in the data. According to these measures, LSTM always beats GPT-SNN-PPO in-sample and out-of-sample throughout all time intervals (5-day, 24-day, and 60-day). LSTM's ability to generalise to unseen data and its promise for real-time prediction are highlighted by its out-of-sample performance, which shows lower error than its in-sample results. The model predicts that traders will succeed on July 10, 2025, experience a loss on August 15, and then achieve another victory on September 1, 2025, aligning with the market's natural fluctuations.

This study does not explicitly test for stationarity, and the assumption of stationarity is addressed only indirectly through MinMax rescaling. This may limit the robustness of the results, especially in the presence of structural shifts and regime changes. A more comprehensive error analysis, including residual distributions and forecast error variability, is also absent, which constrains the interpretability of model performance. Furthermore, while the models achieve low error values, the economic meaning of forecast errors, such as what an MAE of 0.0008 implies for investors or market participants, is not fully explored. Small numerical errors may still translate into meaningful financial gains or losses, depending on market volatility. Another limitation concerns real-time forecasting constraints, as the study assumes that data are processed without delays or transaction costs, which is rarely the case in actual markets.

Subsequent research should address these limitations by integrating macroeconomic indicators, sentiment data, and event-driven features into hybrid architectures. Including these omitted variables allows the model to capture exogenous shocks, behavioural effects, and structural market changes that price-based inputs alone cannot account for, thereby improving predictive accuracy and robustness. Moreover, the trained model could be applied in transfer learning tasks to forecast related financial instruments or indices. Transfer learning leverages knowledge from the FTSE/JSE Top 40 to improve learning efficiency and performance on new datasets, reducing training time while maintaining high predictive performance. These extensions provide financial market practitioners and academics with more reliable and adaptable forecasting tools for complex, dynamic markets.

This research assesses the predictive efficacy of two deep learning models—GPT–SNN–PPO and LSTM —on the closing price of the FTSE/JSE Top 40 Index, a financial series characterised by temporal dependencies, nonlinearity, and heteroscedasticity. Although GPT–SNN–PPO effectively captures temporal dynamics and adapts to changing patterns through its spike-based learning and reinforcement learning components, the LSTM consistently outperforms it across all forecasting horizons. The improved performance of LSTM is evident in both in-sample and out-of-sample evaluations, emphasising its ability to generalise to novel data and making it especially appropriate for real-time prediction. Notwithstanding these encouraging findings, the research recognises limitations, such as its reliance solely on historical price data, the implicit assumption of stationarity via MinMax rescaling, and the lack of macroeconomic, investor sentiment, and event-driven factors. Addressing these deficiencies in forthcoming research, particularly via transfer learning, might improve forecast accuracy, resilience to structural changes, and application across analogous financial instruments. The results indicate that while hybrid architectures such as GPT–SNN–PPO have theoretical potential, LSTM offers a more dependable and pragmatic approach for long-term forecasting of the FTSE/JSE Top 40. The work offers important advice to academics and practitioners seeking to develop precise, flexible, and economically relevant forecasting models in complex financial markets.

Authors' contribution*: Methodology, Abstract, and proofreading the final document-C.Sh.; Literature review and proceeded with the final document. NM; Conclusion, improved the methodology and the analysis –Csi.Data curation, data analysis, and reporting - KM.*
Acknowledgement. *The authors thank numerous people for their helpful comments on this paper.*

References

Abusnaina, A. A., and Abdullah, R. (2014). Spiking neuron models: A review. International Journal of Digital Content Technology and its Applications (JDCTA), 8(3), 1–12

Al-Najjar, D. (2016). Modelling and estimation of volatility using ARCH/GARCH models in Jordan's stock market. Asian Journal of Finance and Accounting, 8(1), 152–169. https://doi.org/10.5296/ajfa.v8i1.9129

Baillie, R. T., Bollerslev, T., and Mikkelsen, H. O. (1996). Fractionally integrated generalised autoregressive conditional heteroscedasticity. Journal of Econometrics, 74(1), 3–30. https://doi.org/10.1016/S0304-4076(95)01749-6

Bollerslev, T., Engle. R.F. and Wooldridge J.M., (1988). A Capital Asset Pricing Model with Time-Varying Covariances. The Journal of Political Economy, 96(1), 116–131. doi:10.1086/261527

Chikobvu, D., and Sigauke, C. (2020). Statistical distribution fitting to the number of COVID-19 deaths in South Africa. Unpublished work. Available at https://www.researchgate.net/publication/342020590_Statistical_distribution_fitting_to_the_number_of_COVID-19_deaths_in_South_Africa. (Accessed on 2025 July 13).

Davies, A., Veličković, P., Buesing, L., Blackwell, S., Zheng, D., Tomašev, N., Tanburn, R., Battaglia, P., Blundell, C., Juhász, A., Lackenby, M., Williamson, G., Hassabis, D., & Kohli, P. (2021). Advancing mathematics by guiding human intuition with AI. *Nature*, *600*(7887), 70-74. https://doi.org/10.1038/s41586-021-04086-x

Djankov, S., and Panizza, U. (Eds.). (2022). COVID-19 in developing economies. CEPR Press. https://cepr.org/publications/books-and-reports/covid-19-developing-economies

Dufrenot, G., and Matsuki, T. (2021). Recent econometric techniques for macroeconomic and financial data. Springer. https://doi.org/10.1007/978-3-030-54252-8

Ehteram, M., Nia, M. A., Panahi, F., & Farrokhi, A. (2024). Read-First LSTM model: A new variant of the long short-term memory neural network for predicting solar radiation data. Energy Conversion and Management, 305, 118267.https://doi.org/10.1016/j.enconman.2024.118267

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica, 50(4), 987–1007. https://doi.org/10.2307/1912773

Elhag, M. (2025). LSTM prediction capability on the South African JSE Top 40 of historical and live data (Master's thesis). University of Cape Town. http://hdl.handle.net/11427/41537

Gerstner, W. and Kistler, W. M. (2002). Spiking neuron models: Single neurons, populations, plasticity. Cambridge University Press. https://doi.org/10.1017/CBO9780511815706

Hua, Y., Mou, L., and Zhu, X. X. (2019). Recurrently exploring class-wise attention in a hybrid convolutional and bidirectional LSTM network for multi-label aerial image classification. ISPRS Journal of Photogrammetry and Remote Sensing, 149, 188–199. https://doi.org/10.1016/j.isprsjprs.2019.01.015

Karimah HN, Lee C, Seo Y. (2025) Batchnorm-Free Binarised Deep Spiking Neural Network for a Lightweight Machine Learning Model. Electronics; 14(8):1602. https://doi.org/10.3390/electronics14081602

Khan, M. A., Khan, R., Algarni, F., Kumar, I., Choudhary, A., and Srivastava, A. (2022). Performance evaluation of regression models for COVID-19: A statistical and predictive perspective. Ain Shams Engineering Journal, 13(2), 101574. https://doi.org/10.1016/j.asej.2021.08.016

Lapique, L. (1907). Recherches quantitatives sur l'excitation électrique des nerfs traitée comme une polarisation. Journal of Physiology and Pathology, 9, 620–635.

Makatjane K, Shoko C, and Moroke N. (2025). Downside Risk of the South African Mining Index: Leveraging the Power of Long Short-Term Memory and Explainable AI. Journal of

International Crisis and Risk Communication Research, 130–146. Retrieved from https://jicrcr.com/index.php/jicrcr/article/view/3066

Makatjane, K. (2024). Forecasting Time-varying Value--at--Risk and Expected Shortfall Dependence: A Markov-switching Generalised Autoregressive Score Copula Approach. Austrian Journal of Statistics, 53(2), 81–98. https://doi.org/10.17713/ajs.v53i2.1710

Makatjane, K., and Moroke, N. (2022). Examining stylised facts and trends of FTSE/JSE TOP40: A parametric and non-parametric approach. Data Science in Finance and Economics, 2, 294–320.

Malashin, I., Tynchenko, V., Gantimurov, A., Nelyub, V., and Borodulin, A. (2024). Applications of Long Short-Term Memory (LSTM) networks in polymeric sciences: A review. Polymers, 16(18), 2607. https://doi.org/10.3390/polym16182607

Qin, P., Ding, Y., Li, Y., Ye, B., Gao, Z., Liu, Y., Cai, Z., and Qi, H. (2025). Spiking neural networks optimised by improved cuckoo search algorithm: A model for financial time series forecasting. Algorithms, 18, 262. https://doi.org/10.3390/a18050262

Rast, A., Galluppi, F., Davies, S., Plana, L., Patterson, C., Sharp, T., Lester, D., and Furber, S. (2011). Concurrent heterogeneous neural model simulation on real-time neuromimetic hardware. Neural Networks, 24(9), 961–978. https://doi.org/10.1016/j.neunet.2011.06.014

Schulman, J., Wolski, F., Dhariwal, P., Radford, A., and Klimov, O. (2017). Proximal policy optimisation algorithms. arXiv preprint. https://arxiv.org/abs/1707.06347

Silva, I., Nunes, Y. T., Moioli, R. C., & Guedes, L. A. (2023). Advancing Fault Prediction: A Comparative Study between LSTM and Spiking Neural Networks. *Processes*, *11*(9). https://doi.org/10.3390/pr11092772

Shoko, C., Moroke, N. D., and Makatjane, K. (2024). A Deep Learning Framework for Modelling Temporal Dependencies and Hierarchies in Hourly Electricity Demand Load. In P. Acharjya, S. Koley, & S. Barman (Eds.), Machine Learning and Computer Vision for Renewable Energy (pp. 42-65). IGI Global Scientific Publishing. https://doi.org/10.4018/979-8-3693-2355-7.ch003

Strohsal, T., Proaño, C. R., and Wolters, J. (2019). Characterising the financial cycle: Evidence from a frequency domain analysis. Journal of Banking and Finance, 106, 568–591. https://doi.org/10.1016/j.jbankfin.2019.06.010

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I. (2017). Attention is all you need. In Advances in Neural Information Processing Systems, 5998–6008.

Xiang, X., Li, Y., Wu, X., Liu, Z., Wu, L., Wu, B., ... and Zeng, Z. (2025). Future variation and uncertainty source decomposition in deep learning bias-corrected CMIP6 global extreme precipitation historical simulation. Frontiers in Earth Science, 13, 1601615. https://doi.org/10.3389/feart.2025.1601615

Wu, Y., Deng, L., Li, G., Zhu, J., Xie, Y., and Shi, L. (2019). Direct training for spiking neural networks: Faster, larger, better. In Proceedings of the AAAI Conference on Artificial Intelligence, 33(1), 1311–1318.