International Journal of Islamic and Middle Eastern Finance and Management

719

Stock market predictive modeling using recurring neural network with long short-term architecture

Fahim Afzal
Business School, Hohai University, Nanjing, China

Farman Afzal
Institute of Business and Management, University of Engineering and Technology, Lahore, Pakistan

Muhammad Kamran
Operations Department, Powerledger, Perth, Australia, and

Haiying Pan
Business School, Hohai University, Nanjing, China

Received 30 December 2024
Revised 19 June 2025
23 July 2025
Accepted 2 August 2025

Abstract

Purpose – The fast development of artificial intelligence (AI) and machine learning techniques can help develop precise and trustworthy forecasting models with precision to predict the volatile and irregular stock behavior in the market. The purpose of this paper is to develop a robust predictive model that has the potential to use AI in enhancing the accuracy of stock market profit assessment. This study evaluates the effectiveness of a long short-term memory (LSTM) model for predicting the stock market behavior by incorporating key economic factors.

Design/methodology/approach – This study constructs an LSTM model that incorporates economic factors influencing stock market direction, including interest rates, inflation, GDP and unemployment rates. These indicators are selected based on their theoretical relevance and empirical significance. A multivariate LSTM model is developed and validated against traditional autoregressive integrated moving average (ARIMA) and machine learning-based random forest (RF) models. Model performance is evaluated using root mean squared error, mean absolute error and directional accuracy.

Findings – The experimental results show that the performance of LSTM model varies across four different cross-validation folds, with the fourth fold appearing to have better generalization performance than the first fold, as both the mean training (0.0450687) and validation (0.01293) values are lower in the fourth fold than in the previous folds. This study highlights that the recurring neural network-based LSTM model, when combined with a holistic approach provides accurate and interpretable predictions for stock market indices of highly volatile markets. The results demonstrate that the LSTM model significantly outperforms ARIMA and RF models in terms of predictive accuracy. The LSTM effectively captures nonlinear patterns in financial time series and shows strong generalization capabilities.

Practical implications – This study suggests that AI-driven techniques can be adapted to predict stock behavior with precision. In addition, such models may be effective for other market indices displaying similar behavioral patterns, aiding stakeholders in making informed investment decisions.

Originality/value – This study contributes to the body of knowledge by proposing a practical AI-based stock market prediction model that incorporates economic factors directly impacting stock market movements. This integral approach helps develop more accurate prediction models, providing insight

This paper forms part of a special section "Artificial intelligence (AI) and machine learning (ML)", guest edited by Mamunur Rashid, Tonmoy Taufic Choudhury and Rashedul Hasan.

International Journal of Islamic and Middle Eastern Finance and Management
Vol. 19 No. 3, 2026
pp. 719-739

1753-8394
DOI 10.1108/IMEFM-12-2024-0632

into various elements of the stock market. The inclusion of baseline model comparisons and detailed residual analysis enhances the methodological rigor and applicability of the approach.

Keywords Stock market prediction, RNN, LSTM, Volatility, Financial modeling, Machine learning

Paper type Research paper

1. Introduction

The stock exchange represents a dynamic environment where market values fluctuate continuously, reflecting the ever-changing nature of economic and financial conditions. Predicting stock prices is no easy task, especially in highly volatile stock markets, such as those in Asia. Some authors support and believe in historical trends, whereas others argue that historical data sometime provides misleading predictions. Over the past decade, information-sharing sources such as social media and financial news have facilitated a clearer understanding of stock performance, which can sometimes trigger unpredictability challenges. However, information technology now provides more effective media sentiment analysis to enhance the understanding of stock performance. Such IT-based techniques help in matching news sentiments with stock market trends, providing a better step for accurate predictions (Seng and Yang, 2017; Shahi *et al.*, 2020; Sheth and Shah, 2023).

To predict stock prices and returns, the majority of the earlier studies used statistical time-series approaches based on historical data (Efendi *et al.*, 2018). The most widely used methods include moving average, Kalman filtering, exponential smoothing, autoregressive conditional heteroscedasticity model, autoregressive moving average and autoregressive integrated moving average (ARIMA) models (Chen and Chen, 2015; Wei *et al.*, 2011; Yeh *et al.*, 2011). Following this, with the rise of artificial intelligence (AI) and soft computing, studies on stock market prediction have given these methodologies greater consideration. A helpful technique to depict these processes is using machine learning. Machine learning techniques are proving to be noticeably faster and more accurate than current prediction strategies. It projects a market value that is very close to the tangible value, thereby increasing accuracy (Jiang, 2021). This promotes the development of accurate stock market prediction models that use AI and machine learning approaches. Several published studies have tried to create sophisticated forecasting models or systems that can effectively predict stock market movements (Sedighi *et al.*, 2019). In broad terms, stock market forecasting is considered one of the most important but challenging fields in financial research (Chen and Hao, 2017). In emerging markets like the Pakistan Stock Exchange (PSX), prior studies have primarily used traditional models such as ARIMA or GARCH. However, studies like Lin and Lobo Marques (2024) suggest that deep learning approaches, especially long short-term memory (LSTM), offer superior performance in capturing complex temporal patterns, which motivates our approach.

The stock price changes are unpredictable, and there are numerous interrelated causes for this behavior. Interconnected elements, including statistics on the global economy, the unemployment rate, monetary policies of major countries, disastrous events and several other factors like inflation, GDP, unemployment rate, consumer confidence and market sentiments, could all be contributing factors (Moradi *et al.*, 2021). In the stock market, everyone wants to boost their profits while lowering their risks. The key difficulty is gathering the various pieces of information, organizing them in one location and creating a reliable model for reliable predictions. Huge profits for the seller and the broker might result from an accurate stock prediction. Prediction is often considered to be chaotic rather than random, meaning it can be predicted by carefully studying the history of the relevant stock market. To mitigate risks and maximize returns, market participants strive to assess these

elements thoroughly. However, the discipline of stock market forecasting faces significant challenges in collecting a diverse range of data and developing reliable prediction models.

This unpredictability of stock price changes is due to numerous interrelated factors, including global economic impact, the unemployment rate, monetary policies of major countries, disastrous events and several other factors, like inflation, GDP, unemployment rate, consumer confidence and market sentiments, which could all be contributing factors (Moradi *et al.*, 2021). In the stock market, everyone wants to boost their profits while lowering their risks.

Due to its effective and precise measures, machine learning has been used in the field of stock prediction, which has attracted the attention of many researchers (Qiu and Song, 2016). In machine learning, the data set used is crucial. The data set should be as precise as possible, given the practical constraints, because even minor alterations to the data can have a significant impact on the final product. Several researchers have developed further nonlinear models that use artificial neural networks, fuzzy-neural systems, genetic algorithms, evolutionary and/or particle swarm methodologies for predicting stock markets over the past three decades, as AI has grown (Nti *et al.*, 2020; Rekha and Sabu, 2022; Shi and Zhuang, 2019).

As AI technology analyses data more accurately than humans do, it is possible to predict approaching financial crises with greater accuracy. This can help financial institutions make more informed decisions and mitigate the risks associated with market volatility. The use of AI in the financial sector holds the potential to improve the accuracy of stock market predictions significantly (Kurani *et al.*, 2023; Nazareth and Reddy, 2023). Better stock predictions enable investors to take precautions on their investment against the risk of loss. AI algorithms possess the intense capacity to precisely analyze big data sets, detecting complexities and correlations that often escape human intelligence, leading to tactful perception in making investment decisions (Tölö, 2020). However, it is crucial to appreciate the astounding results of AI tools; even with the accomplishment of comprehensive precision, it is still a grueling task. Thus, investors need to remain vigilant and have a diverse portfolio as a preventive measure. Lin and Lobo Marques (2024) demonstrated the utility of LSTM models in capturing the non-linear relationships between macroeconomic variables and market dynamics in the European stock markets. While their focus was on developed economies, our study uniquely extends this approach to the PSX, providing new insights from an emerging market perspective.

The primary goals of this comprehensive study are to investigate the efficiency of AI in forecasting stock market returns and to assess the LSTM machine learning methodology in explaining these predictions. The purpose of this study is to evaluate the degree of effectiveness with which the LSTM model can predict the trends of the stock market. Moreover, this study conducts a thorough assessment of fundamental economic factors that can impact the prediction of this model, intending to improve precision and further clarify the factors influencing stock market dynamics. The purpose of this research is to develop a robust predictive model that has the potential to use AI in enhancing the accuracy of stock market profit assessments for investors. This research aims to carry out a comprehensive analysis of historical stock market data from the PSX. Contrary to most previous research, which has relied solely on data from the stock market, our method underscores the importance of considering multiple features that influence stock market movements, particularly in emerging markets, given their pronounced volatility.

Hence, this study includes crucial macroeconomic components, including interest rates, GDP, inflation and unemployment rates, acknowledging their significant impact on the volatile stock market. This integral approach helps develop more accurate prediction models,

providing insight into various elements that influence the stock market, specifically in circumstances related to the PSX. Machine learning models, such as LSTMs, often excel in accuracy, but their scaled outputs can be challenging to interpret. To address this, we use Min-Max scaling during model training for stability. Following prediction generation, we use the Min-Max Scaler's inverse transform method to restore predictions to their original scale. This ensures that our predictions are not only accurate but also immediately comprehensible, empowering decision-makers to act on AI-driven insights with confidence in the financial domain. To validate model performance, the study also benchmarks LSTM predictions against those of ARIMA and random forest (RF) models, offering a comparative evaluation across statistical and machine learning paradigms.

2. Data preparation

This study uses data from the Islamic stock indices of the PSX, which fully comply with the criteria for Islamic stocks defined by the PSX Shariah Advisory Board. It is pertinent to note that, following the directive from the State Bank of Pakistan, all commercial banks are expected to operate under an Islamic banking framework by June 2027. Therefore, in the future, only Islamic financial services would be offered to the larger market. Considering the transition of the financial ecosystem, it is vital to consider Islamic stock indices for predicting future stock trends. Subsequently, these indices would be the accurate estimators of stock behavior. Four prominent stock indices, namely, the KSE100, SE30, KSEAll and KMI30, were used for the period spanning from 2013 to 2023, providing a comprehensive representation of the market conditions. A total of 2,472 daily closing prices were gathered from Yahoo Finance for data analysis. These daily closing prices were transformed into log returns, which were subsequently used for LSTM modeling.

While every market is subject to the influence of numerous external factors, the stock market, in particular, is significantly impacted by economic indicators such as the Interest Rate (referred to as KIBOR in Pakistan), GDP, Inflation, and Unemployment Rate, etc. Ongoing research efforts have resulted in the continuous development of novel models, and AI-based models have emerged as notably practical tools for stock market predictions. It is worth noting that many prior studies exclusively relied upon stock prices to forecast future trends; however, the inclusion of influential economic factors holds supreme importance. Thus, this study explicitly incorporates essential economic indicators spanning from 2013 to 2023 (namely, KIBOR, GDP, Inflation and Unemployment Rate) known to influence the functioning of the stock market significantly. Integrating macroeconomic indicators, such as interest rate, inflation and industrial production, enables the incorporation of broader economic dynamics that traditional technical analysis alone may overlook.

2.1 Data normalization

A Shapiro–Wilk test was conducted to assess the normality of the data. The low *p*-values indicate that the data does not follow a normal distribution. The QQ-plots Figure 1, illustrates significant distributional disparities within the data. The pronounced deviations in the tails of the distribution provide clear evidence of nonnormality. Subsequent to the confirmation of nonnormality through the test results, the data underwent further processing to normalize it.

A complete analysis of summary statistics for three financial and economic indicators is given in Table 1, which also highlights the central tendencies and variability within each data set. The four stock indices, KSE100, KSE30, KSEAll and KMI30, all have different means and standard deviations, reflecting the different degrees of volatility they show.

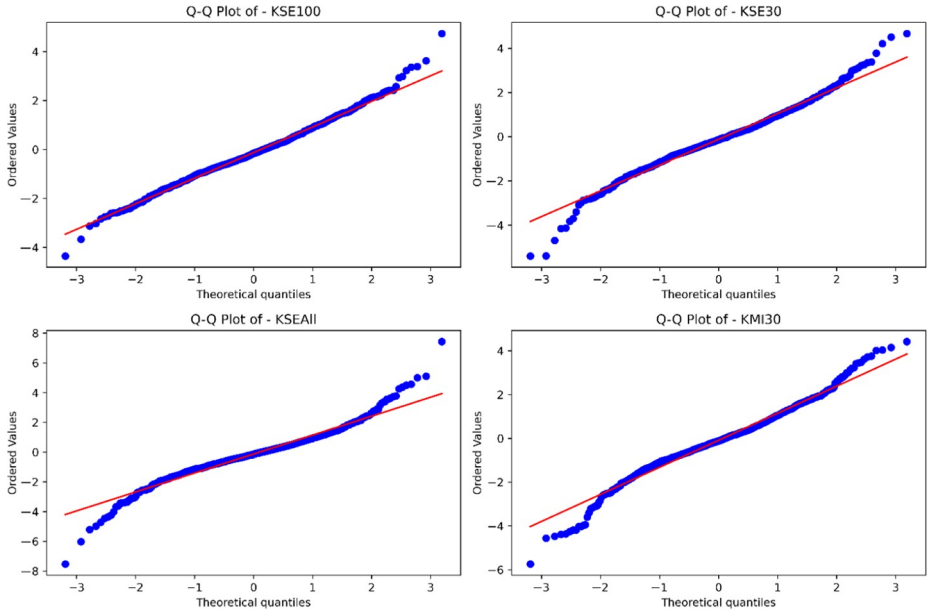


Figure 1. Q-Q plots shows a graphical representation of notable distributional differences in the data. The severe deviations in the distribution's tails provide clear evidence of nonnormality
Source: Authors' own work

2.2 Feature selection

Economic indicators have a substantial impact on stock market events. This study examines the significance of four prominent economic indicators – Kibor, GDP, Inflation and Unemployment across various stock indices. The data spanned economic indicators from 2013 to 2023. Figure 2 provides a clear illustration of the relevance of the features in influencing stock market price fluctuations. Notably, across all indices, Kibor emerged as the most prominent feature, exhibiting a higher feature score compared to the other indicators. Nevertheless, it is essential to acknowledge the significance of the other indicators, even if their scores are relatively modest. This is particularly crucial in the case of Pakistan, where these indicators hold noteworthy relevance in interpreting the complexities of stock market behavior. The comprehensive inclusion of all four indicators stems from their individual importance to the target market. Each indicator imparts distinct market insights and underlying information, contributing to a deeper understanding of market dynamics.

Random Forest Regressor (RFR) is used to calculate the feature importance scores, as it provides scores that are relatively easy to interpret and understand. In addition, it directly assigns scores to each feature, unlike other AI models. Moreover, it has the strength to capture the complex relationship between the variables.

The RFR equation to calculate the scores is as follows:

For each feature "j" in the data set, the importance score "Imp(j)" is calculated as the sum of the normalized total decrease in Gini impurity over all decision trees "t" in the forest, where "f(j, t)" is the number of splits on feature "j" in tree "t":

Table 1. Summary statistics

	KSE100	KSE30	KSEALL	KMI30
Count	2,472	2,472	2,472	2,472
Mean	0.000267	−0.000050	0.000249	0.000258
Std	0.010513	0.011700	0.009115	0.012384
Min	−0.071024	−0.077801	−0.057981	−0.078312
25%	−0.004841	−0.005965	−0.004369	−0.006033
50%	0.000421	−0.000087	0.000390	0.000080
75%	0.006044	0.006119	0.005431	0.006929
Max	0.046840	0.047279	0.037986	0.061936
Shapiro–Wilk	0.99	0.98	0.95	0.96
Test statistic *p*-values	8.46E-05	1.01E-09	7.86E-17	3.32E-11
	KIBOR	GDP	inflation	Unemp
Mean	9.98	3.65	10.07	4.46
Std	3.76	2.24	8.13	1.49
Min	6.23	−0.94	2.53	1.8
25%	6.89	3.12	3.93	3.6
50%	8.81	4.07	8.35	4.08
75%	11.45	5.77	15.23	6.3
Max	23.27	6.1	28.3	6.8

Note(s): This table includes counts, means, and standard deviations, offering insights into the central tendency and variability of the data. Count means number of observations in this case, there are 2,472 observations for each variable, Mean is a measure of central tendency, Standard deviation (std) measures the amount of variation or dispersion in a set of values. 25, 50, 75% are percentiles that provide information about the distribution of the data, "Min" is the minimum value and "Max" is the maximum value observed. The Shapiro–Wilk test is used to assess normality of data. The "*p*-values" represent the likelihood of detecting the specified test statistic if the data were drawn from a normally distributed population. Lower *p*-values indicate that there is more evidence against the null hypothesis of normal distribution
Source(s): Authors' own work

$$Imp(j) = \frac{1}{T}\sum_{T=1}^{t} \sum splits\ on\ feature\ j \frac{f(j,\ t)}{n_t}\ .\Delta I\ (Gini,\ t) \tag{1}$$

3. Long short-term memory model

LSTM model consists of following layers:

- Input Layers (Input sequences);
- LSTM Layers (Model stack, sequential learning pattern); and
- Dense Layers (Output).

$$q_t = \sigma\left(v_f \cdot \left[k(t-1),\ u_t\right] + d_f\right) \tag{2}$$

$$j_t = \sigma\left(v_f \cdot \left[k(t-1),\ u_t\right] + d_j\right) \tag{3}$$

$$\widetilde{C} = tanh\left(v_c \cdot \left[k(t-1),\ u_t\right] + d_c\right) \tag{4}$$

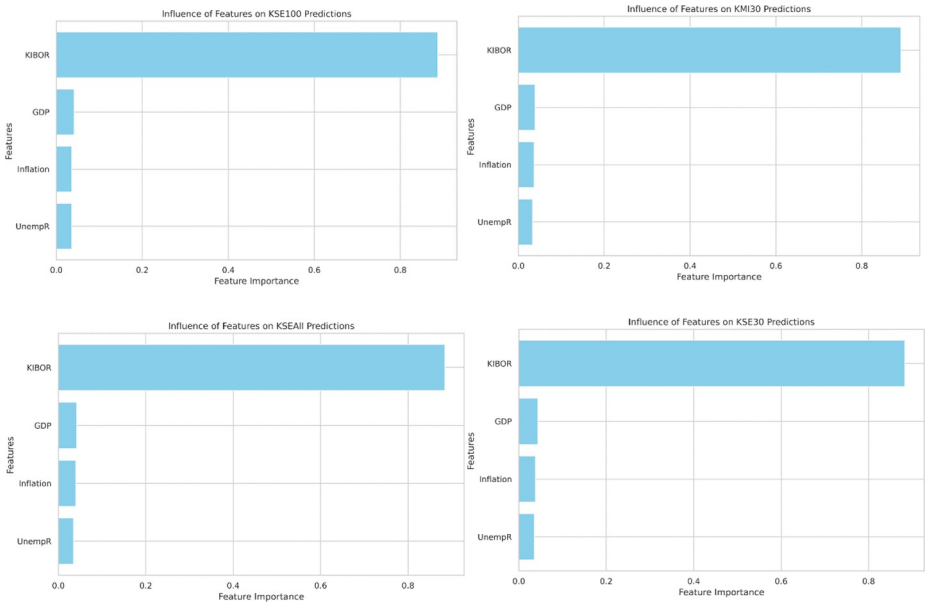


Figure 2. Graphical illustration of the features' importance in affecting changes in KSE100, KSE30, KSEAll and KMI30
Source: Authors' own work

$$C_t = g_t \cdot C(t-1) + j_t \cdot \widetilde{C}t \tag{5}$$

$$m_t = \sigma\left(v_m \cdot \left[k(t-1),\ u_t\right] + d_o\right) \tag{6}$$

$$k_t = m_t \cdot tanh(C_t) \tag{7}$$

where

$\mathbf{q_t}$ = Forget gate vector at time *t*;
$\mathbf{j_t}$ = Input gate vector at time *t*;
$\tilde{\mathbf{C}}_\mathbf{t}$ **(tilde C)** = Candidate cell state vector at time *t*;
$\mathbf{C_t}$ = Updated cell state vector at time *t*;
$\mathbf{m_t}$ = Output gate vector at time *t*;
$\mathbf{k_t}$ = Hidden state (output) vector at time *t*;
$\mathbf{u_t}$ = Input vector at time *t* (includes both stock prices and macroeconomic indicators);
$\mathbf{k_{t-1}}$ = Hidden state vector from the previous time step (*t–1*);
$\mathbf{v_f}$, $\mathbf{v_j}$, $\mathbf{v_c}$, $\mathbf{v_m}$ = Weight matrices for forget gate, input gate, candidate cell state and output gate, respectively;
$\mathbf{d_f}$, $\mathbf{d_j}$, $\mathbf{d_c}$, $\mathbf{d_o}$ = Bias vectors for each gate (forget, input, candidate and output);
σ = Sigmoid activation function;
tanh = Hyperbolic tangent activation function; and
$\mathbf{[k_{t-1}, u_t]}$ = Concatenation of the previous hidden state and the current input.

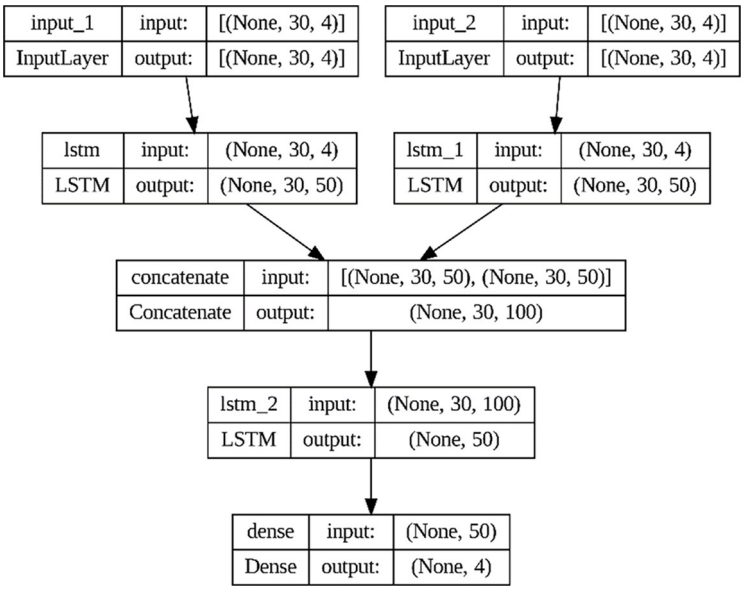


Figure 3. LSTM architecture processes two distinct forms of input data – historical stock prices and economic indicators
Source: Authors' own work

3.1 Architecture of the long short-term memory model

LSTM architecture in Figure 3 demonstrates its versatility by handling two different types of input data, historical stock prices and economic indicators. The model starts with two input layers, each designed to process sequences of data with specific lengths and feature dimensions. Afterward, two LSTM layers are introduced to process each type of input separately. These LSTM layers, each with 50 units, use the Rectified Linear Unit activation function and are configured to generate sequences, preserving temporal context. As the model progresses, a concatenation layer smoothly combines the processed sequences, effectively integrating the information obtained from both historical stock prices and economic indicators. This crucial step helps capture complex relationships and connections between the two types of data. Further down the architecture, a final LSTM layer with 50 units is added to capture higher-level temporal dependencies within the combined data. To make predictions about future stock prices, the model concludes with an output layer that uses the linear activation function. It is important to note that the number of output units matches the dimensionality of the stock price data, ensuring that predictions maintain the same format.

After configuring the LSTM model's architecture, the data set is partitioned into separate training and testing subsets. The training data set constitutes 80% of the data, whereas the remaining 20% forms the testing data set. The model compilation is executed by employing specific hyperparameters, including the Adam optimizer and the mean squared error (MSE) loss function. These parameters remain constant throughout various epochs and batch sizes, ensuring consistent conditions even when these latter variables are modified. The LSTM

Table 2. Value of loss function and MAE for different number of epochs

Sr. no.	Training	Testing	Epoch	Batch	Optimizer	Loss function	Mean training loss	Mean validation loss	SD training loss	SD validation loss	Cross validation: Mean training MAE	Cross validation: Mean validation MAE
1	80	20	50	64	Adam	MSE	0.0034	0.0012	0.0167	0.0027	0.0531	0.019
2	80	20	50	32	Adam	MSE	0.0014	0.0006	0.006	0.0005	0.0436	0.0165
3	80	20	100	64	Adam	MSE	0.0013	0.0005	0.0081	0.0007	0.0534	0.016
4	*80*	*20*	*100*	*32*	*Adam*	*MSE*	*0.0008*	*0.0004*	*0.004*	*0.0003*	*0.0451*	*0.0129*

Note(s): In above table training and testing means percentage of the dataset used for training and testing the model, Epoch means number of passes through the entire training data set during the training process, Batch refers to the number of samples used, Optimizer is the optimization algorithm used during training. In this case, it is Adam. Loss Function used to evaluate the performance of the model. Here, MSE is used, Mean Training and Mean Validation Loss are the average loss on the training and validation dataset over all epochs, Std Dev Training and Std Dev Validation Loss indicating the variability of the training and validation loss values, mean-training-MAE refers Mean Absolute Error for the mean of the training set in cross-validation folds, mean-validation-MAE refers to the Mean Absolute Error for the mean of the validation set in cross-validation folds

Source(s): Authors' own work

model's performance is systematically assessed across varying configurations of LSTM layers, epochs and batch sizes, as outlined in the comprehensive Table 2.

The model selection process hinges upon the analysis of training loss function values, specifically the MSE, derived from model fitting. The chosen model is subsequently leveraged to generate predictions for the forthcoming 365 days. Subsequently, cross-validation is used to evaluate the selected model's predictive capacity using the mean absolute error (MAE). Notably, the predictions are obtained through each of the LSTM models under consideration. To validate the superiority of the selected model, an extensive cross-validation precision testing procedure is performed. By comparing the MAE values and comprehensively assessing the chosen model's cross-validated performance, its efficacy is confirmed.

The future prediction has been made using the following equation:

$$Z^{t+1} = g\left(u_t,\ k_{t-1},\ c_{t-1}\right) \tag{8}$$

where z^{t+1} represents the predicted value for the next time step (t+1), g denotes the LSTM model's prediction function, u_t signifies the input features at the current time step (t), $k_{t-1}$is the hidden state of the LSTM at the previous time step (t−1) and c_{t-1} denotes the cell state of the LSTM at the previous time step (t−1).

In the context of this study, the LSTM model predictions are based on the economic indicators. Equation (8) captures the input features and hidden states to generate future predictions.

3.2 Model interpretability

To enhance the interpretability of our financial prediction model, we used a crucial innovation during the postprocessing stage. While our LSTM-based model uses Min-Max scaling for optimal training, making the predictions amenable to machine learning algorithms, we ensure their real-world relevance by applying an inverse transformation. The inverse transformation method, represented mathematically as:

$$S = S1 * (max - min) + min \tag{9}$$

where S represents the original, interpretable value, S_1 denotes the scaled value obtained from the model's prediction and max and min signify the maximum and minimum values from the original data used for scaling.

This innovative step ensures that our model's predictions are not confined to a numerical range but are returned to their original scale, immediately accessible and actionable for financial decision-makers. By making the outputs interpretable, our innovation empowers stakeholders to navigate financial complexities with confidence and precision.

3.3 Cross-validation

The cross-validation technique has been used to evaluate the performance of LSTM model. The cross-validation process has been carried out using MAE and RMSE, with a statistical process as follows:

$$MAE = \frac{1}{\mathrm{k}} \sum_{i=1}^{k} MAE_i \tag{10}$$

where k is the number of cross-validation folds (in this case, $k = 5$). MAE_i is the MAE for the ith fold, calculated by comparing the actual stock index values $Z_{test,i}$ and the predicted stock index values $\widehat{Z}_{pred,i}$:

$$MAE_i = \frac{1}{N}\sum_{J=1}^{N} \mid Z_{test,i,j} - \widehat{Z}_{\text{pred, i, j}} \tag{11}$$

where N is the number of samples in the testing set of the ith fold, $Z_{test,i,j}$ is the jth actual stock index value in the testing set of the ith fold and $\widehat{Z}_{pred,i,j}$ is the jth predicted stock index value in the testing set of the ith fold:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y_i})^2} \tag{12}$$

where:

y_i = Actual (true) value at time step i;

$\hat{y}_i$ = Predicted value at time step i;

n = Total number of observations; and

Directional accuracy (%) = Measures the ability of the model to predict the correct direction of movement.

3.4 Baseline models

To assess the relative performance of the LSTM model, we implemented two widely used baseline models:

3.4.1 Autoregressive integrated moving average. The ARIMA model is a traditional statistical model for time series forecasting. Its general form is:

$$ARIMA(p, d, q) = y_t = c + \phi_1 y_{t-1} + \ldots. + \phi_p y_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \ldots. + \theta_q \epsilon t - q \tag{13}$$

where:

p = lag order;

d = degree of differencing;

q = order of moving average; and

ε_t = white noise.

3.4.2 Random forest regression. RF is a robust ensemble learning algorithm using multiple decision trees. It predicts as:

$$\widehat{y_t} = \frac{1}{N}\sum_{i=1}^{N} f_i(X_t) \tag{14}$$

f_i = prediction from the ith decision tree;

N = total number of trees;

X_t= input features at time t;

The RF model used the same economic indicators as input features to ensure a fair comparison.

4. Results and discussions

The proposed technique is trained and evaluated using a data set of PSX stock indices obtained from Yahoo Finance. It is divided into training and testing sets and produces the following results after passing through the various models. Table 2 shows the precision of our training and testing for all the epochs. First, we discovered nonnormality in the data using the Shapiro–Wilk test, and then the data was normalized. We used the Random Forecast Regressor to calculate the feature importance scores, and we discovered that across all indices, KIBOR emerges as the most prominent feature, with a higher feature score than the other indicators. Specific hyperparameters, such as the Adam optimizer and the MSE loss function, are used to execute the model compilation. A lower MSE implies a better fit of our model to the data; these parameters are constant throughout epochs and batch sizes, maintaining consistent conditions even when these latter variables are changed. In addition to making the predictions suitable to machine learning algorithms, our LSTM-based model uses Min-Max scaling for optimal training. By applying an inverse transformation, we ensure that the predictions of our model are not constrained to a numerical range but are instead returned to their original scale, which ensures that the predictions are relevant to real-world situations.

Table 2 contains information about different configurations of an LSTM model, including training and testing splits, training parameters and evaluation metrics. The first configuration seems to have relatively high training and validation losses, as well as MAE values. The standard deviation of losses is relatively high, indicating that the model's performance varies significantly during training. However, second configuration performs better than the first configuration. It has lower training and validation losses, lower standard deviations and lower MAE values. The smaller batch size might contribute to improved generalization. The third configuration features a LSTM layer size, and more epochs compared to the first configuration.

Although the training loss is lower, indicating that the model may be learning more effectively, the validation loss and MAE remain high. The standard deviations of losses are also elevated, suggesting potential instability in training. The fourth configuration performs well compared to the other configurations. It has the lowest training and validation losses, as well as the lowest standard deviations of losses. As both training and validation losses decrease, it suggests that our model is learning the data's patterns effectively and is likely to generalize well to unseen data. Figure 4 shows graphs of loss functions for all four folds.

We have used the cross-validation technique to evaluate the performance of the LSTM model. Table 2 shows the mean training MAE and mean validation MAE for different cross-validation folds in the context of LSTM model evaluation. These values are typically used to assess the model's performance and generalization ability. Here's an analysis of four different folds: In first fold, the mean training MAE is 0.05313, which is relatively higher than the mean validation MAE. This suggests that the model might be overfitting to the training data as it performs better on the validation data.

However, the validation MAE of 0.019 indicates good predictive performance on unseen data. In second fold, both the mean training (0.04355) and the validation (0.01645) MAE values are lower compared to first fold. This suggests an improvement in model's performance and generalization. The lower training MAE indicates that the model is learning the underlying patterns in the data effectively. Similar to first fold, in third fold, the mean training MAE is higher than the mean validation MAE. However, both values are relatively lower than those of first fold, indicating improved performance. In fourth fold, both the mean training (0.0450687) and validation (0.01293) MAE values are lower than in the previous folds. A low validation MAE suggests that the model generalizes well to unseen data in this

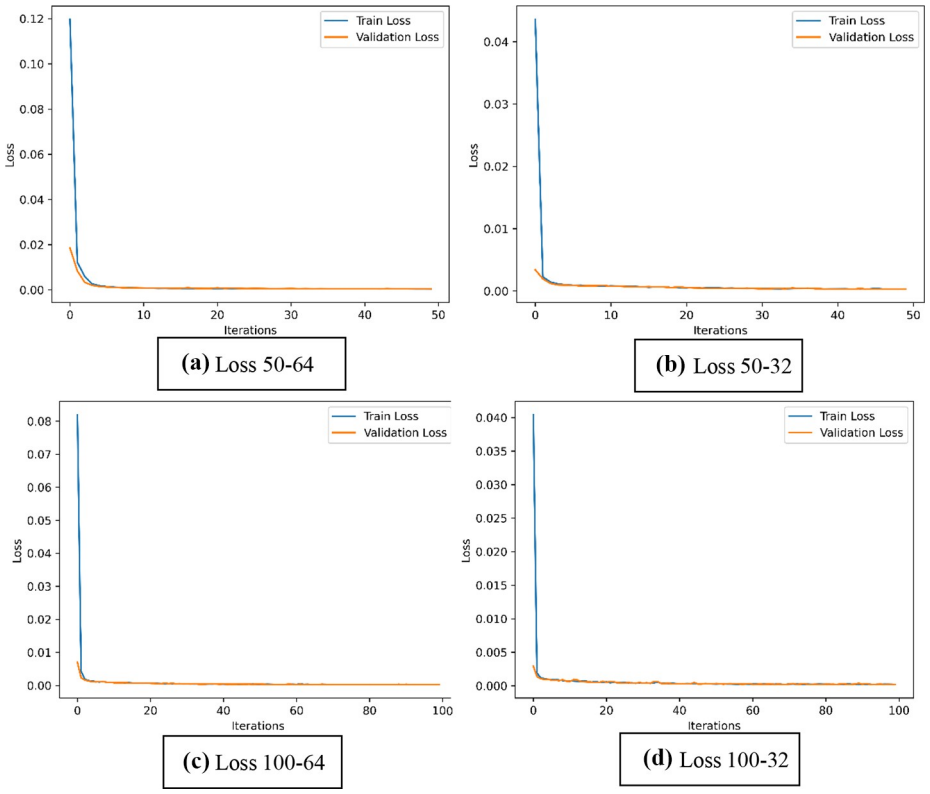


Figure 4. Graphical description of training and validation losses of four configurations
Source: Authors' own work

particular fold. The LSTM model's performance varies across different cross-validation folds. Folds second, third and fourth appear to have better generalization performance compared to first fold. The decreasing trend in both training and validation MAE values across the folds indicates that the model is learning and improving over time, which is a positive sign. For different data sets, we can observe that training with more epochs can improve our testing result, while also allowing us to achieve better forecasting and prediction values. MAE measures the average absolute difference between the model's predictions and the actual target values. A lower MAE indicates that, on average, the model's predictions are closer to the actual values. Table 2 and Figure 5 also demonstrates that, increasing the number of training epochs for our model improves the forecasting's accuracy. The length of the data and the number of epochs both significantly affect the testing outcome.

Thus, last configuration, i.e. fourth fold is selected because it seems to be the most promising among the options provided, as it has lower losses, lower standard deviations, lower training MAE values, lower cross-validation MAE values and good cross-validation performance, indicating that the LSTM model is making accurate predictions on average also indicating better model performance and generalization also suggests that the model's performance is stable and reliable. Our findings align partially with the broader patterns

observed by Chan *et al.* (2025), who noted the varying predictive power of economic indicators across market conditions. This reinforces the importance of market-specific dynamics in LSTM-based forecasting.

Figure 5 and Table 3 show the future predictions and precision of our training and testing for all the epochs for KSE100, KSE30, KSE-All and KMI30 indices of PSX.

4.1 Comparison of models

To benchmark the performance of the LSTM model, two baseline models were implemented:

ARIMA: A traditional statistical forecasting model suitable for univariate time series analysis. The model parameters (p, d, q) were selected based on AIC minimization.

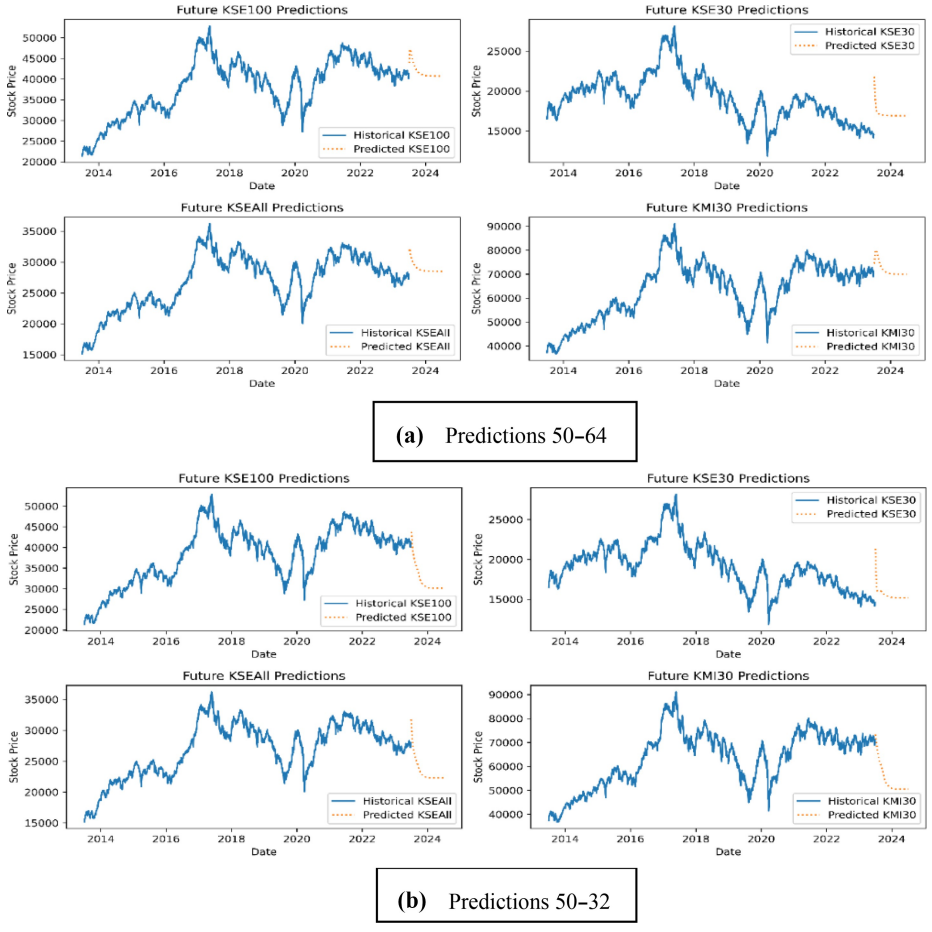
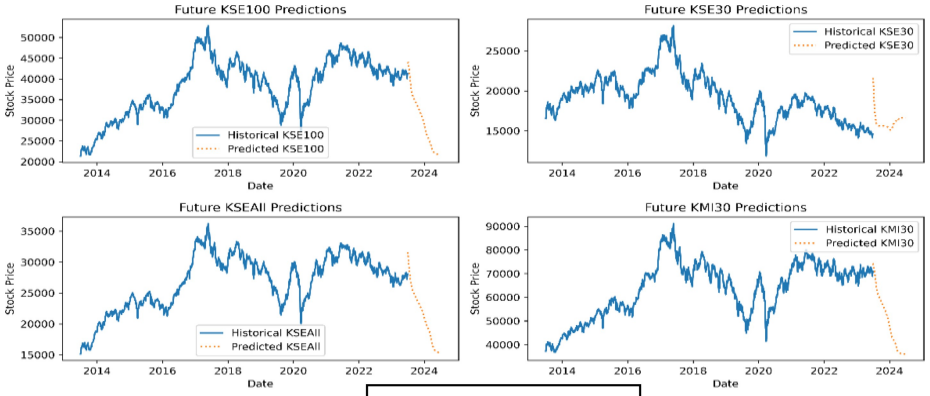
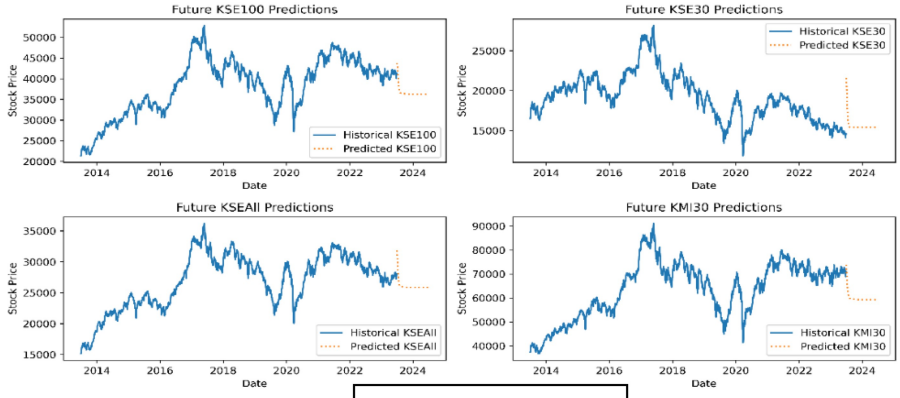


Figure 5. Future predictions results of training with two different number of epochs – (50 and 100) and with two different batches – (32 and 64)
Source: Authors' own work



(c) Predictions 100–64



(d) Predictions 100–32

Figure 5. continued

RFR: A nonparametric ensemble learning method known for its robustness and ability to handle non-linear relationships. The model was trained with the same feature set as the LSTM.

All models were trained and tested on identical training/test splits. For evaluation, we used MAE, RMSE and directional accuracy.

To enhance interpretability, MAE and RMSE were computed on the original scale of the stock prices. This allows direct comparison of forecast error magnitudes. For instance, the LSTM model yielded an MAE of 165.24 points on the PSX index, reflecting high precision (see Table 4).

Although the LSTM model outperforms traditional models, the margin of improvement is moderate. Future work could explore ensemble learning or hybrid modeling approaches to further boost accuracy, especially in uncertain economic environments.

Table 3. Predicted values of stocks

	Inverted-Predictions 50–32				Inverted-Predictions 50–64			
Date	KSE100	KSE30	KSEAll	KMI30	KSE100	KSE30	KSEAll	KMI30
July 1, 2023	43891	21679	31809	73782	43921	21802	31818	74440
July 2, 2023	43980	21583	31859	73605	44298	21728	31965	74553
July 3, 2023	44062	21396	31867	73381	44813	21541	32089	74768
July 4, 2023	44113	21193	31817	73059	45282	21324	32181	74904
July 5, 2023	44126	20938	31772	72692	45715	21100	32214	75136
:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:
June 27, 2024	33249	15901	24049	53840	40747	16909	28487	69971
June 28, 2024	33249	15901	24049	53840	40747	16909	28487	69971
June 29, 2024	33250	15901	24049	53841	40747	16909	28487	69971
June 30, 2024	33250	15901	24049	53841	40747	16909	28487	69971
	Inverted-Predictions 100–32				Inverted-Predictions 100–64			
Date	KSE100	KSE30	KSEAll	KMI30	KSE100	KSE30	KSEAll	KMI30
July 1, 2023	43834	21574	31829	73929	43759	21623	31614	74288
July 2, 2023	43632	21334	31591	73130	43821	21496	31441	74150
July 3, 2023	43467	21017	31269	72220	43908	21304	31238	73939

(*continued*)

Table 3. Continued

Date	Inverted-Predictions 100–32				Inverted-Predictions 100–64			
	KSE100	KSE30	KSEAll	KMI30	KSE100	KSE30	KSEAll	KMI30
July 4, 2023	43285	20654	30910	71278	43964	21034	31020	73635
July 5, 2023	43120	20270	30522	70305	43976	20726	30768	73320
:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:
June 27, 2024	36280	15405	25886	59333	21771	16638	15379	36121
June 28, 2024	36280	15405	25886	59333	21771	16638	15379	36121
June 29, 2024	36280	15405	25886	59333	21771	16638	15379	36121
June 30, 2024	36280	15405	25886	59333	21771	16638	15378	36121

Note(s): This table presents some of predicted values of four stock indices in four different configurations such as 50–32, 50–64, 100–32, 100–64
Source(s): Authors' own work

Table 4. Comparison of model performance

Model	MAE (original scale)	RMSE	Directional accuracy (%)
ARIMA	215.31	290.11	51.20
Random Forest	198.45	275.09	54.80
LSTM	165.24	240.33	60.50

Note(s): Table presents accuracy of all three models used in the study
Source(s): Authors' own work

Table 5. Comparison of model performance

Horizon	MAE	RMSE	Directional accuracy (%)
1-day ahead	165.24	240.33	60.50
5-day ahead	189.11	258.7	57.90
10-day ahead	215.45	276.5	55.10

Note(s): Table presents forecast accuracy of using LSTM model
Source(s): Authors' own work

4.2 Multi-horizon forecasting performance

To assess robustness over time, the LSTM model was tested across different prediction horizons: one-day ahead, five-day ahead and 10-day ahead.

These results show that while error increases with horizon length, the model maintains relatively high directional accuracy even at 10-day intervals.

Table 5 presents a comparison of forecast accuracy of the LSTM model against ARIMA and RF, we applied the Diebold–Mariano test. The null hypothesis of equal predictive accuracy was rejected at the 5% significance level for both comparisons, indicating that the LSTM model provides significantly better forecasts.

Our findings have several practical implications. For investors and portfolio managers, the integration of economic indicators with LSTM models can enhance short- and medium-term market timing strategies, particularly in emerging markets like Pakistan. Policymakers may also use such predictive frameworks to anticipate market stress periods and monitor the financial system's sensitivity to macroeconomic developments. Moreover, this model can serve as a decision support tool when designing counter-cyclical regulatory policies. While forward-filling and interpolation help synchronize data frequency, they may introduce bias or dampen short-term variations, potentially affecting model sensitivity to real-time economic shifts. Future models could test alternative imputation strategies or mixed-frequency modeling techniques.

5. Conclusions

In conclusion, our study has demonstrated the performances of an recurring neural network-based LSTM model for stock market prediction, our LSTM-based model, with careful configuration and optimization, demonstrated robust performance in predicting stock indices. The chosen configuration, informed by comprehensive evaluations and cross-validation analyses, stands out as the most promising for accurate and reliable forecasting in the context of the PSX.

In this study, we investigated significance and impact of four economic indicators using RFR and found KIBOR most significant with high score Figure 2. Furthermore, our results clearly show that a higher number of training epochs enhances the model's prediction skills, highlighting the effect of training epochs on forecasting accuracy. LSTM architectures have been implemented and their performances are analyzed by using various evaluation metrics to identify the best model. For stability, we apply Min-Max scaling during model training. Then, after prediction generation, we use the "inverse-transform" method of the MinMaxScaler to return forecasts to their original scale. This ensures that our forecasts are accurate and understandable right away, enabling decision-makers to act with confidence on AI-driven insights in the financial sector. We have monitored the concepts of training loss and validation loss. These metrics helped to assess the performance and generalization ability of our model. Lower training losses and validation losses indicate that there is low difference between model's prediction and actual values.

We have tested LSTM model's performance across four different cross-validation folds and key finding shows that the model achieved its lowest training and validation losses in fourth fold, with 100 epochs and a batch size of 32 (Table 2). Also, in the context of LSTM model performance and evaluation, the mean training MAE and mean validation MAE are used. We have tested LSTM model's performance across four different cross-validation folds and the fold (100–32) is the best suited model and experimental results showed both the mean training (0.0450687) and validation (0.01293) MAE values are lower than in the previous folds. A low validation MAE suggests that the model generalizes well to unseen data in this particular fold.

LSTM models, which offer improved accuracy and responsiveness to dynamic market situations, are going to transform stock market forecasting in the future. Future revisions of LSTM models are projected to deliver more accurate projections, enabling investors to make knowledgeable decisions in real time by leveraging their inherent capacity to capture long-term dependencies and nonlinear interactions in time series data. Enhancements in computational efficiency and feature learning and representation may enable the extraction of more accurate and pertinent market indicators, as well as real-time predictions. LSTM models can help improve risk management techniques by offering more accurate predictions of possible market downturns or swings. This can be critical for investors and financial institutions in making educated risk-mitigation decisions.

The suggested methodology can be easily adapted to use with other wide market indices when the data displays a similar pattern of behavior. The proposed model can help interested parties better understand the market environment before making investment decisions. The results of our model have been encouraging. According to the testing results, our model can track the progress of opening prices for four indices. Our findings show that AI-driven methodology can be used to various market indices displaying.

We have included major macroeconomic factors including interest rates, inflation, GDP and unemployment rates in this analysis as we recognize their direct impact on stock market movements. Other factors that influence financial markets include political events, global trends and investor mood. LSTM models may not fully account for all of these complexities, resulting in inaccurate predictions. So in future these factors can also be considered for more accurate predictions using AI methods.

LSTM models perform well for large data sets, these models are susceptible to overfitting when trained on small data sets and generalize poorly to new, unseen data. Although LSTM models are intended to capture long-term dependencies, producing accurate long-term forecasts in financial markets remains a difficult issue due to the numerous unknown aspects. Despite the model's strong performance, it is limited by the availability of high-frequency

economic data; future studies could address this by incorporating alternative data sources. While our study provides encouraging results, several limitations remain. First, the model is calibrated on historical data, and its performance in highly volatile or unprecedented market conditions (such as geopolitical crises or pandemics) remains to be tested. Second, the chosen economic indicators, while grounded in theory, may not capture all relevant market-driving factors. Future research could explore the integration of sentiment analysis or global economic indicators to improve predictive accuracy further. Third, deep learning models like LSTM require substantial data, which may limit applicability in markets with shorter histories.

A key challenge in using economic indicators for stock market prediction lies in the timeliness of such data. Many economic variables are reported monthly or quarterly, whereas stock markets operate on a daily basis. In our model, we addressed this mismatch via feature engineering (e.g. forward filling, interpolation, lag structures). However, such approaches may not fully compensate for delayed information flow, which could reduce model effectiveness in real-time deployment. Future research could explore incorporating high-frequency proxies for economic activity, such as sentiment indices or alternative data sources.

Our results demonstrate that incorporating macroeconomic indicators into an LSTM framework can significantly enhance the accuracy of stock price forecasting in emerging markets, such as the PSX. The model consistently outperformed traditional benchmarks across multiple prediction horizons. These findings contribute to both academic understanding and practical forecasting approaches for emerging financial markets.

References

Chan, L.K., Lin, C.Y. and Yeh, J.-H. (2025), "Market efficiency and stability under short sales constraints: evidence from a natural experiment with high-frequency resolution".

Chen, M.-Y. and Chen, B.-T. (2015), "A hybrid fuzzy time series model based on granular computing for stock price forecasting", *Information Sciences*, Vol. 294, pp. 227-241.

Chen, Y. and Hao, Y. (2017), "A feature weighted support vector machine and K-nearest neighbor algorithm for stock market indices prediction", *Expert Systems with Applications*, Vol. 80, pp. 340-355.

Efendi, R., Arbaiy, N. and Deris, M.M. (2018), "A new procedure in stock market forecasting based on fuzzy random auto-regression time series model", *Information Sciences*, Vol. 441, pp. 113-132.

Jiang, W. (2021), "Applications of deep learning in stock market prediction: recent progress", *Expert Systems with Applications*, Vol. 184, p. 115537.

Kurani, A., Doshi, P., Vakharia, A. and Shah, M. (2023), "A comprehensive comparative study of artificial neural network (ANN) and support vector machines (SVM) on stock forecasting", *Annals of Data Science*, Vol. 10 No. 1, pp. 183-208.

Lin, C.Y., and Lobo Marques, J.A. (2024), "An exploratory comparison of stock price prediction: using multiple machine learning approaches based on global stock indices", *Entrepreneurship, Innovation, and Technology: A Holistic Analysis of Growth Factors*, Springer Nature Switzerland AG, via Palgrave Macmillan, Cham, pp. 317-336.

Moradi, M., Appolloni, A., Zimon, G., Tarighi, H. and Kamali, M. (2021), "Macroeconomic factors and stock price crash risk: do managers withhold bad news in the crisis-ridden Iran market?" *Sustainability*, Vol. 13 No. 7, p. 3688.

Nazareth, N. and Reddy, Y.Y.R. (2023), "Financial applications of machine learning: a literature review", *Expert Systems with Applications*, Vol. 219, p. 119640.

Nti, I.K., Adekoya, A.F. and Weyori, B.A. (2020), "A systematic review of fundamental and technical analysis of stock market predictions", *Artificial Intelligence Review*, Vol. 53 No. 4, pp. 3007-3057.

Qiu, M. and Song, Y. (2016), "Predicting the direction of stock market index movement using an optimized artificial neural network model", *Plos One*, Vol. 11 No. 5, p. e0155133.

Rekha, K.S. and Sabu, M.K. (2022), "A cooperative deep learning model for stock market prediction using deep autoencoder and sentiment analysis", *PeerJ Computer Science*, Vol. 8, p. e1158.

Sedighi, M., Jahangirnia, H., Gharakhani, M. and Farahani Fard, S. (2019), "A novel hybrid model for stock price forecasting based on metaheuristics and support vector machine", *Data*, Vol. 4 No. 2, p. 75.

Seng, J.-L. and Yang, H.-F. (2017), "The association between stock price volatility and financial news–a sentiment analysis approach", *Kybernetes*, Vol. 46 No. 8, pp. 1341-1365.

Shahi, T.B., Shrestha, A., Neupane, A. and Guo, W. (2020), "Stock price forecasting with deep learning: a comparative study", *Mathematics*, Vol. 8 No. 9, p. 1441.

Sheth, D. and Shah, M. (2023), "Predicting stock market using machine learning: best and accurate way to know future stock prices", *International Journal of System Assurance Engineering and Management*, Vol. 14 No. 1, pp. 1-18.

Shi, C. and Zhuang, X. (2019), "A study concerning soft computing approaches for stock price forecasting", *Axioms*, Vol. 8 No. 4, p. 116.

Tölö, E. (2020), "Predicting systemic financial crises with recurrent neural networks", *Journal of Financial Stability*, Vol. 49, p. 100746.

Wei, L.-Y., Chen, T.-L. and Ho, T.-H. (2011), "A hybrid model based on adaptive-network-based fuzzy inference system to forecast Taiwan stock market", *Expert Systems with Applications*, Vol. 38 No. 11, pp. 13625-13631.

Yeh, C.-Y., Huang, C.-W. and Lee, S.-J. (2011), "A multiple-kernel support vector regression approach for stock market price forecasting", *Expert Systems with Applications*, Vol. 38 No. 3, pp. 2177-2186.

Corresponding author
Farman Afzal can be contacted at: farmanafzal@gmail.com