Article History | Received: Jan. 8, 2026 | Revised: May 25, 2026 | Accepted: May 30, 2026

AI-Driven NEPSE Forecasting:
Comparing LSTM, GRU, and TCN with Sentiment

Paras Mishra[1]

Abstract

Stock price prediction in emerging markets like Nepal remains a complex and challenging task due to high volatility, limited liquidity, and the influence of investor sentiment. This study investigates the performance of three deep learning models-Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), and Temporal Convolutional Networks (TCN)-for predicting stock prices using both historical Open, High, Low, Close, and Volume (OHLCV) data and financial news headlines. Sentiment scores are extracted from news headlines using FinBERT to evaluate their impact on prediction accuracy. The models are compared based on Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), R-squared (R^2), and Mean Absolute Percentage Error (MAPE). Experimental results indicate that GRU outperforms LSTM and TCN in both scenarios, and integrating sentiment analysis further improves predictive performance. These findings highlight the potential of AI-driven models, combined with market sentiment, for informed investment decisions in the Nepalese capital market.

Keywords: Deep learning, Nepal Stock Exchange (NEPSE), sentiment analysis, stock price prediction

Cite as: Mishra, P. (2026). AI-Driven NEPSE forecasting: Comparing LSTM, GRU, and TCN with sentiment. *Journal of Business and Social Sciences Research, 11*(1), 157-163. https://doi.org/10.3126/jbssr.v11i1.99012

INTRODUCTION AND STUDY OBJECTIVES

Stock price prediction refers to predicting future values of stocks or financial instruments using their past time-series data, technical indicators, sentiments or through machine learning/deep learning algorithms (Zhao et al., 2025). Prediction has always remained a tough challenge because of the volatile, dynamic, and non-linear nature of financial markets. Effective prediction may be useful in the process of decision making for investors, businesses and policymakers. The challenge has forced researchers to look into computational techniques that may improve prediction accuracy. Neural networks which mimic the human brain can deal with non-linear and sequential data effectively. Neural network models which perform exceptionally well at modeling the long-term dependencies of financial time series include, but not limited to, the Long Short Term Memory (LSTM), Gated Recurrent Unit (GRU), and Bidirectional LSTM (BiLSTM) (Cho et al., 2014; Hochreiter & Schmidhuber, 1997).

[1]Mr. Mishra is a Lecturer, Department of Computer Science and Information Technology (CSIT) Patan Multiple Campus, Patan Dhoka, Lalitpur, Nepal. Email: parasmishra428@gmail.com

In addition to recurrent neural networks, a variety of other neural network models have been used for predicting stock prices. Very recently, the Temporal Convolutional Networks (TCNs) have been suggested as an alternative architecture for modeling the sequential time-series data using causal and dilated convolutions which effectively capture temporal dependencies (Bai et al., 2018).

News and social media sentiments have recently become important external variables influencing price changes in stocks especially in emerging capital markets like Nepal. Price changes caused by investors' reaction towards news on economics, policies, and corporate disclosure are not always represented in price history. Combining news sentiments in addition to time series prices presents a chance to improve prediction accuracy through including market psychology in predictions (Adhakari & Maheshwari, 2025). The success of integrating sentiments in prediction depends on deep learning architectures but this has never been investigated for Nepalese capital markets yet. Even though Nepal Stock Exchange (NEPSE) is the main platform for capital market activities in Nepal, it has been difficult for the investors to predict price changes in stocks. The NEPSE market is characterized as highly volatile due to low liquidity, changes in policies, information asymmetry, and high investor sentiments (Adhakari & Maheshwari, 2025). The use of conventional prediction methods and technical analysis is insufficient in modeling the non-linear nature of stock prices. The use of artificial intelligence and deep learning techniques is very effective in forecasting stock prices around the world; however, the use of AI and deep learning in the context of Nepal in relation to the stock market has not been well exploited yet. Mishra (2026) made an excellent empirical case to prove the fact that the predictive performance of GRU models, especially with sentiment scores derived from FinBERT, is far better than others in the volatile market of Nepal.

The difference creates a need for powerful methods of forecasting through AI which can effectively model the market trend and the sentiment of investors in improving forecast precision and making better decisions in NEPSE. This research work, therefore, intends to comparatively assess the performance of LSTM, GRU, and TCN algorithms considering both with and without news sentiment data for stock prices in making their impact on prediction accuracy evident. The main goal of this study is to assess and compare the predictive capacity of these deep learning models in forecasting NEPSE stock prices so as to determine the best AI based method for Nepal's capital market.

LITERATURE REVIEW

In the operation of Back Propagation Neural Network (BPNN), stock market behavior analysis works through supervised learning, which allows the learning of complex and non-linear relations between input variables and stock prices (Wu et al., 2022). The structure of the network includes input layer, containing different inputs such as historical prices and technical indicators; processing hidden layers and output layer, from which predictions can be made. In the forward propagation step, inputs flow through the network and output, which is the price prediction, is produced. Error, which is the difference between predicted and actual stock price, is computed and used for back propagation through the network in order to update the connection weights of the network using an appropriate optimization method, such as gradient descent. This forward calculation and backward error adjustment process is repeated for a certain number of iterations, allowing BPNN to "learn" the complex patterns, which makes BPNN an appropriate approach in stock price forecasting and volatility prediction since there exist non-linear and complex relations between variables (Ke et al., 2024; Liagkouras & Metaxiotis, 2025). BPNN has been proved practically valuable since it has been shown to have highly accurate predictions with very small mean absolute percentage errors, for example (Pungkasanti et al., 2026).

Traditional machine learning approaches such as Back Propagation Neural Networks (BPNN) have been compared with recurrent architectures like Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU), with findings suggesting that deep models generally outperform conventional methods in capturing temporal patterns in stock behavior on NEPSE data (Ghimire, 2025). Studies conducted on NEPSE and specific Nepalese stocks have shown that LSTM networks provide improved forecasting accuracy, particularly when used to assist investment decision-making by modeling historical trends (Giri, 2025). Comparative investigations extending to models such as Bidirectional LSTM (BiLSTM) and Support Vector Regression (SVR) report that sequence-based networks like BiLSTM can capture intricate temporal dependencies more effectively than traditional regression models (Gurung, et al., 2024). Moreover, broader deep learning comparisons in Nepalese markets have begun exploring advanced architectures and external features, including research that integrates macroeconomic variables and sentiment analysis to enhance predictive performance (Saud & Shakya, 2019). Despite these developments, the majority of previous research in the Nepalese context has mostly focused on historical pricing data, with comparatively little empirical attention paid to the effects of news emotion integration in deep forecasting models. The current study was motivated by a significant vacuum in the literature where sentiment-augmented models such as LSTM, GRU, and Temporal Convolutional Networks (TCN) applied to NEPSE with and without sentiment remain underexplored.

Long Short-Term Memory (LSTM)

Long Short-Term Memory (LSTM) networks works best for modeling long-term dependencies by solving the problem of gradient problem (Kushwaha et al., 2021). So, they are widely used in time series forecasting because of their ability to deal with long range time sequence. In traditional LSTM networks, the sequential learning mechanism captures long-term dependencies through its forget, input, and output gates, enabling efficient modeling of nonlinear temporal dynamics, as illustrated in the figure. However, raw stock price data often contain noise, abrupt fluctuations, and mixed-frequency components that can limit the learning efficiency of standalone LSTM models. Integrating the Wavelet Transform (WT) enhances this process by decomposing the signal into multi-resolution components, allowing the LSTM to learn from cleaner, frequency-refined patterns.

Gated Recurrent Unit (GRU)

GRU combines the forget gate and input gate into a single update gate to reduce the number of parameters while maintaining the comparable performance. GRUs are particularly advantageous when computational efficiency and faster convergence are required. Research has shown that GRUs can achieve similar accuracy to LSTMs in financial forecasting tasks with reduced training complexity (Saud & Shakya, 2019). The GRU architecture simplifies recurrent learning by using only two gating mechanisms: the reset gate and the update gate. The reset gate controls how much past information to forget, while the update gate regulates how much new information enters the hidden state. This streamlined gating structure enables GRUs to capture long-term dependencies with lower computational complexity compared to LSTMs.

Temporal Convolutional Network (TCN)

Temporal Convolutional Networks (TCN) as introduced as a strong alternative to recurrent models for sequence modeling, demonstrating their effectiveness in capturing long-term dependencies with parallelizable convolutional operations (Shahi et al., 2020). TCNs use causal convolutions, ensuring that the output at any time step depends only on current and past inputs. This preserves temporal order and prevents information leakage from future data. Additionally, dilated convolutions allow TCNs to exponentially increase their receptive field, enabling the capture of long-term dependencies without requiring deep network stacks. Residual connections further enhance training stability and allow deeper architectures without gradient vanishing problems. TCNs also benefit from their ability to process sequences in parallel, making them computationally efficient compared to recurrent models. The use of dilation patterns, as illustrated in the figure, allows the network to progressively expand its temporal coverage across layers. This hierarchical structure helps the model learn both short-range patterns and broader temporal trends effectively. Overall, TCNs provide a stable and scalable architecture for sequence modeling tasks.

METHODS

The research uses a comparative deep learning architecture to predict NEPSE stock prices. The approach consists of integration of historical price data (Open, High, Low, Close, and Volume) of stocks from Merolagani with sentiment scores computed from ShareSansar financial news by means of FinBERT, which surpasses VADER in financial sentiment analysis. The research covers three Nepali financial datasets (PCBL, ADBL, and NABIL) to analyze performance across different volatility levels. The scope includes hyperparameter tuning, model training, error metric evaluation, and comparative analysis between standalone and hybrid architectures. In preprocessing phase, Min-Max normalization, imputation, and combination of sentiment attributes is done. For feature engineering, MACD is used to capture momentum and trend signals. For feature decomposition, Discrete Wavelet Transform (DWT) is used to separate the stock price series into noise (high frequency) and trend (low frequency) parts to improve model learning. The three architectures, namely standalone LSTM, GRU, and TCN, are trained with and without wavelet-based features. The LSTM model makes use of gate functions (input, forget, and output gates) for memory in sequence and dropout for regularization.

Table 1
Model Performance without Sentiment

Model	MSE	RMSE	MAE	R^2	MAPE
LSTM	0.000039	0.006251	0.004451	0.998482	1.088112e+11
GRU	0.000037	0.006085	0.004882	0.998562	6.290818e+10
TCN	0.001144	0.033819	0.027726	0.955583	6.469374e+11

Note. Author's Calculation

Table 2
Model Performance with Sentiment

Model	MSE	RMSE	MAE	R^2	MAPE
LSTM	0.000071	0.008455	0.006004	0.997262	2.221430e+11
GRU	0.000030	0.005484	0.003968	0.998848	1.778609e+11
TCN	0.002978	0.054575	0.043303	0.885917	1.166810e+12

Note. Author's Calculation

DATA ANALYSIS

This part of the paper covers the forecasting results using three AI-driven model for the sampled company.

Model Performance without Sentiment

Table 1 illustrates the result for forecasting results without sentiment.

Of the three models analyzed without sentiment information, GRU and LSTM performed extraordinarily accurate in prediction, where the former performed marginally better than the latter in terms of metrics. GRU obtained the minimum MSE (0.000037), RMSE (0.006085), and maximum R^2 (0.998562), reflecting the fact that GRU is more proficient in analyzing sequential data in the NEPSE prices. The performance of LSTM is remarkable, having similar metrics (MSE: 0.000039, RMSE: 0.006251, R^2: 0.998482). On the other hand, the performance of TCN is comparatively poor, showing higher errors (MSE: 0.001144, RMSE: 0.033819, MAE: 0.027726) and lower R^2 of 0.955583. It is notable that the extremely high MAPE value in all the models may be indicative of the sensitivity to the extremely small price value.

Model Performance with Sentiment

Table 2 illustrates the result for forecasting results with sentiment.

When using sentiment features, the best performing model was GRU, which had the lowest values for MSE (0.000030), RMSE (0.005484), and MAE (0.003968) while attaining the highest value for R^2 (0.998848). This shows that GRU successfully used the news sentiments to boost prediction efficiency. It is surprising to see that LSTM performed poorly when it came to including the sentiment features because the performance worsened; there was an increase in MSE (0.000071) from (0.000039) and R^2 dropped to (0.997262). There were no signs of improvement in TCN performance as the error rates increased drastically (RMSE: 0.054575, R^2: 0.885917). The inclusion of sentiment has proved beneficial only for GRU.

CONCLUSION

In conclusion, the paper demonstrates that deep learning models, particularly GRU, are highly effective for predicting stock prices in the Nepalese capital market. While LSTM and TCN models offer reasonable predictive capabilities, GRU consistently achieved the best performance, especially when combined with sentiment analysis derived from news headlines using FinBERT. The integration of sentiment data enhanced the model's ability to capture market dynamics influenced by investor sentiment, reducing prediction errors and improving accuracy. These findings highlight the potential of AI-driven approaches for informed decision-making in Nepalese stock trading and suggest that incorporating alternative data sources, such as news sentiment, can further strengthen forecasting models in emerging markets.

Funding
The authors declared having received no funding for this work.

Conflict of interest
The authors declared having no conflict of interest associated with this publication.

REFERENCES

Adhikari, C. P., & Maheshwari, D. K. (2025). Sentimental effects on stock prices in the Nepalese stock market. *Sahayaatra, 8*(1), 79-94. https://doi.org/10.3126/sahayaatra.v8i1.81047

Bai, S., Kolter, J.Z., & Koltun, V. (2018). *An empirical evaluation of generic convolutional and recurrent networks for sequence modeling.* https://ui.adsabs.harvard.edu/abs/2018arXiv180301271B

Cho, K., van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder–decoder for statistical machine translation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 1724–1734. https//doi.org/10.3115/v1/D14-1179

Ghimire, M. (2025). Application of deep learning algorithms and genetic programming for forecasting stock prices in Nepal: A comparative study. *SADGAMAYA, 2*(1), 29-37, https//doi.org/10.3126/sadgamaya.v2i1.80325

Giri, S., Dam, S.R., Kamar, R., & Tulachan, S.B. (2025). Stock price prediction for investment decision using long short-term memory. *Technical Journal, 5*(1), 69–74. https//doi.org/10.3126/tj.v5i1.86892

Gurung, M., Neupane, P., & Shrestha, S. (2024). Stock market prediction in Nepali stock market using machine learning model. *KEC Journal of Science and Engineering, 8*(1), 158–168. https//doi.org/10.3126/kjse.v8i1.69290

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation, 9*(8), 1735–1780. https//doi.org/10.1162/neco.1997.9.8.1735

Ke, Z., Xu, J., Zhang, Z., Cheng, Y., & Wu, W. (2024). *A consolidated volatility prediction with back propagation neural network and genetic algorithm.* International Conference on Image Processing, Computer Vision and Machine Learning (ICICML), Shenzhen, pp. 1671-1675, https://doi.org/10.1109/ICICML63543.2024.10958142

Kushwaha, A.K., Kar, A.K., & Vigneswara Ilavarasan, P. (2021). *Predicting retweet class using deep learning.* Ed(s) Vincenzo Piuri, Sandeep Raj, Angelo Genovese, Rajshree Srivastava, In Hybrid Computational Intelligence for Pattern Analysis, Trends in Deep Learning Methodologies, 89-112. https://doi.org/10.1016/B978-0-12-822226-3.00004-0

Liagkouras, K., & Metaxiotis, K. (2025). *Artificial neural network with back propagation, long short-term memory and transformer models for stock market forecasting: An experimental study.*16th International Conference on Information, Intelligence, Systems & Applications (IISA), Greece, pp. 1-7, https//doi.org/10.1109/IISA66859.2025.11311244

Mishra, P. (2026). AI-driven stock price prediction: A case study on Agricultural Development Bank Limited (ADBL) using LSTM, GRU and TCN with and without news-headling sentiment. *Aadim Journal of Multidisciplinary Research Information & Technology, 2*(1.0), 114-120. https://doi.org/10.3126/ajmr.v2i1.97551

Pungkasanti, P.T., Christanto, F.W., Sabat, F.T.H., Dewi, C., & Firdaus, E.A. (2026). Python-based stock price prediction using backpropagation neural networks: A case study on ANTM. *Bulletin of Electrical Engineering and Informatics, 15*(1), 723-730. https://doi.org/10.11591/eei.v15i1.9760

Saud, A.S., & Shakya, S. (2019). Analysis of gradient descent optimization techniques with gated recurrent unit for stock price prediction: A case study on banking sector of Nepal Stock Exchange. *Journal of Institute of Science and Technology, 24*(2), 17-21. https://doi.org/10.3126/jist.v24i2.27247

Shahi, T.B., Shrestha, A., & Neupane, A. (2020). Stock price forecasting with deep learning: *A comparative study. Mathematics, 8*(9), 1441. https://doi.org/10.3390/math8091441

Wu, Y., Zhu, D., Liu, Z., & Li, X. (2022). An improved BPNN algorithm based on deep learning technology to analyze the market risks of A+H shares. *Journal of Global Information Management, 30*(7), https://doi.org/10.4018/JGIM.293277

Zhao, C., Chen, S., Zhang, L., Wang, Y., & Li, J. (2025). A hybrid stock prediction method based on periodic/non-periodic decomposition, Time2Vec, and Transformer. *EPJ Data Science, 13*(17), 1–22. https//doi.org/10.1140/epjds/s13688-024-00517-7