International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0144
e ISSN: 2584-2854
Volume: 04 Issue: 04 April 2026
Page No: 945 - 952

Daily ATM Cash Demand Forecasting with Deep Learning-Based Hybrid CNN-LSTM Architectures

Pradnya P. Shelar[1], Dr. Monika Bhatnagar[2]
[1]Research Scholar, Department of Computer Science & Engineering, Oriental University, Indore (M.P.)
[2]Supervisor, Department of Computer Science & Engineering, Oriental University, Indore (M.P.)
Emails: pradnyak.bachhav@gmail.com[1], monikabhatnagar@oriental.ac.in[2]

Abstract

The predictability of daily cash demand at Automated Teller Machines (ATM) is a sensitive operational objective of financial institutions, which directly affects the efficiency of cash logistics, the availability of services, and the economy of costs. The resulting errors in forecasting can lead to the recurrence of cash shortages, therefore, customer dissatisfaction and reputational losses, or surplus idle cash, to the higher security and opportunity costs. The competition held in the NN5 forecasting predicted that there was an inherent complexity in predicting ATM cash withdrawals due to the strong seasonality in weekly, calendar effects, missing data, and the temporal heterogeneity in demand patterns across ATM locations. The paper provides a deep learning-based forecasting system of NN5 daily ATM cash demand prediction by incrementally scaling a stock market price prediction system first created to predict a financial time-series. The proposed framework introduces domain-specific enhancements, such as calendar-based feature engineering and hybrid convolutional-recurrent architectures, while retaining the fundamental ideas of sliding-window supervision, normalization, and recurrent neural modeling. Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU) as well as a hybrid model, Convolutional Neuron Network-LSTM (CNN-LSTM) model are deployed and tested within a single experimental framework. Decades of experiments on the NN5 data depict that the proposed CNN-LSTM setup is much better than the classical statistical models, such as ARIMA and exponential smoothing, and even standalone recurrent neural networks, and results in up to a 19 percent decrease in Root Mean Square Error (RMSE). The findings validate that the combination of the cross-domain architectural transfer, hybrid models of deep learning, and calendar-sensitive feature design offers a stable and scalable solution to the real-world prediction of the ATM cash demand.

Keywords: *ATM cash demand forecasting, NN5 dataset, deep learning, LSTM, GRU, CNN–LSTM, time-series forecasting, hybrid neural networks*

1. Introduction

1.1. Overview

Automated Teller Machines (ATMs) are an intrinsic part of modern banking facilities as they are able to offer cash till the customer in a convenient location at any time. To maintain sufficient cash supply in the ATMs, there is need to have proper forecasting of daily withdrawal requirement. Mistakes of the demand prediction have direct financial and operational impacts. Excessive idle cash caused by overestimation poses a threat to security and opportunity costs; this as well as underestimation causes cash-outs, dissatisfaction of customers and possible damage of reputation to banks. Therefore, cash demand forecasting in financial operations management has turned out to be an issue of vital decision-support in the ATM cash demand management [1], [2]. Conventionally, banks have depended on the use of rule of thumb and classical statistical forecasting models, including moving averages, autoregressive integrated moving average (ARIMA) and exponential smoothing models, to project the future cash demand. These approaches are mathematically interpretable and computationally efficient, but stationarity is too much of an assumption, restricting them to nonlinear dynamics, sudden demand dynamics and complex seasonal dynamics that are seen in real world ATM withdrawal data [3]. Also, ATM data are often missing values that are due to machine outages, service or repairs or the mis entry of information, which makes prediction

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0144
e ISSN: 2584-2854
Volume: 04 Issue: 04 April 2026
Page No: 945 - 952

even more challenging. This was topped up by the NN5 forecasting competition to offer a standard benchmark used in assessing forecasting techniques on actual ATM cash withdrawal data [4]. The competition data is made up of the daily withdrawal values of 111 ATMs during a period of roughly two years and shows good weekly and heterogeneous demand dynamics. Findings of the NN5 competition indicated that none of the classical forecasting models had a consistent and uniformly superior performance across all ATMs (indications of the need to adopt more adaptable and flexible modeling methods). Deep neural network models have been shown within the last few years to provide excellent performance in capturing the complex time-dependent predictive features of many time-series forecasting problems, such as finance, energy demand, traffic flow, sales forecasting, and so on [5], [6]. LSTM networks introduce by Hochreiter and Schmidhuber [7] overcome the issue of vanishing gradient in the simple RNNs and have become a common method of learning long-term predictors of data in a sequence. Consequently, stock market price prediction with LSTM-based architectures is a prevalent field to predict stock market price changes, wherein they effectively ease stock markets trend patterns clearly in nonlinear temporal relationships and long-range interactions [8], [9]. Although stock price time series and ATM cash withdrawal data have structural parallels (they are both noisy, non-stationary and externalities), prior research has done little to understand how architectures of stock market predictive models can be systematically transferred to forecast cash demand in ATM. This gap has been addressed by this paper where a stock market prediction pipeline has been modified to deal with the NN5 daily ATM cash demand forecasting problem. The primary hypothesis is that when paired with feature engineering domain-specific, deep learning design architectures initially developed to forecast financial time-series can be implemented in ATM demand prediction. These goals of the study are: (i) to train an LSTM-based stock market forecasting pipeline in the NN5 setting; (ii) to use calendar effects based features to capture the weekly and weekend effects; (iii) compare and contrast LSTM,GRU and hybrid CNN-LSTM models in the same set up; and (iv) to evaluate the proposed setup and compare and contrast with current NN5 statistical baselines. This work has made contributions such as proving the efficacy of architecture transfer between domains and hybrid deep learning framework to realistic ATM cash demand forecasting.

Table 1
DL Architectures Frequently Used in Time Series Tasks

Model	Core Strengths	Limitations
RNN	Sequential memory; simple architecture	Vanishing gradients; short-term memory
LSTM	Handles long dependencies; stable training	Computationally expensive; overfitting risk
GRU	Fewer parameters than LSTM; efficient	Slightly less expressive than LSTM
CNN	Extracts local features; good for noise	Lacks temporal modeling capacity
BiLSTM	Reads sequence in both directions	Higher training cost; not causal
CNN-LSTM	Combines feature extraction and memory	Complex; needs tuning
Attention	Learns importance weights dynamically	High compute; requires lots of data
Transformer	Handles long-range dependencies well	Data-hungry; hard to train on small sets
Encoder-Decoder	Used for multi-step prediction	Sensitive to sequence length

1.2. Problem Statement and Research Motivation

Understanding the daily money usage estimates in Automated Teller Machines (ATMs) is significant in the proper cash management and continued banking facilities. Nonetheless, the data of cash withdrawal machines in ATMs are highly seasonal, nonlinear, lack observation, and vary vastly across locations, and therefore, do not allow the use of conventional statistical predictive models like ARIMA and exponential smoothing. Similar to the NN5 forecasting competition, the classical methods clearly do not perform reliably well in all ATMs.Although deep learning models have already been found to be performing well with financial time- series

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0144
e ISSN: 2584-2854
Volume: 04 Issue: 04 April 2026
Page No: 945 - 952

forecasting, especially when used to predict stock markets, their ability to optimize systematically to forecasting cash demand in an ATM is underdeveloped. Besides, hybrid deep learning architectures have not fully been exploited due to their potential to simultaneously learn short-term temporal dynamics and long-term relationships in ATM demand. This paper fills this gap by providing a powerful deep learning framework that transforms the stock market prediction architectures to the problem of NN5 ATM cash demand predictions with the introduction of calendar influences of the domain.

1.3.Objectives of the Study

To fill these gaps, this paper suggests and compares a hybrid deep learning network that consists of CNN, LSTM, and GRU layers to predict the movement in the short term using historical time series data. The research questions of this paper are as follows:

- Predict the daily ATM cash demand with a pipeline based on a stock market time-series prediction system on the dataset NN5.
- Include calendar effect features to be able to pick up the effects of weekly and seasonal withdrawal.
- Predict ATM cash demand using LSTM, GRU and hybrid CNN-LSTM.
- Compare the classical NN5 forecasting references to the suggested models on traditional accuracy measures.
- Assess the extent to which hybrid deep learning systems can improve forecasting accuracy.

2. Related Work

The use of ATM cash demand forecasting is not a new topic of study: it has been at least 20 years old. The initial research largely was based on statistical time-series, such as ARIMA, seasonal ARIMA, and exponential smoothing techniques [2], [3]. These methods use historical patterns of demand to create short-term predictions and are not very difficult to interpret and apply. This is however constrained by the fact that they are based on linear assumptions and hence the failure to capture non-linear and irregular demand behaviour especially when there are complex seasonal patterns. In order to overcome these drawbacks, the regression methods that utilize calendar effects like day of week, month of year and holiday effects were suggested [1], [10]. These analyses showed that there is a high seasonality of ATM cash demand in terms of weekly cycles, with a tendency that maximum withdrawals are on particular days. Although regression models were found to be more accurate in some instances, it was greatly feature engineered and was usually weak over heterogeneous location of ATMs. The NN5 forecasting test was a significant advancement in the research of ATM demand forecasting, since it offered a real-world, large-scale, and standardized level of evaluation [4]. The combination of multiple statistical models was reported to give competitive performance as stated by Makridakis et al. but no single method was prevalent at all ATMs. The outcomes of the competition helped in underlining the necessity of flexibility and strength in models of prediction. As machine learning was developed, nonlinear models were considered, including support vector regression (SVR) and feedforward neural networks, in cash demand forecasting of ATMs [11]. These models enhanced performance in some instances but they could not explicitly model time dependencies and hence worked poorly with sequential data. This limitation was overcome through the use of recurrent neural networks, specifically LSTM networks, which added memory cells that had the capability of holding long term information [7]. Since then, LSTM based models have been used effectively in a range of time series forecasting problems, such as stock market prediction [8], [9], electricity load forecasting [12], and traffic flow prediction [13]. GRU networks were subsequently suggested as a computationally efficient variant of LSTMs, with similar performance using less parameters [14]. Recent years have seen an increase in interest in hybrid networks, which combine recurrent layers and convolutional neural networks (CNNs). CNN layers are useful towards the extraction of local temporal features, e.g. short-term variations, and periodic variations, whereas recurrent layers represent long term dependencies. Compared to traditional CNN, hybrid CNN-LSTM models have shown to be wider in terms of energy demand prediction [12], financial time-series analysis [15],

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0144
e ISSN: 2584-2854
Volume: 04 Issue: 04 April 2026
Page No: 945 - 952

and sales prediction [16]. Nevertheless, their use in the prediction of demand in cash with ATM has not been used extensively, challenging the current research.To fill these gaps, the proposed research suggests a lightweight, entirely integrated, CNN-LSTM-GRU hybrid model, which has to be assessed against a set of benchmarking baselines through the same training set, quantitative and visual performance measures.

3. Methodology

The research design undertaken in this study is investigative experimental research because it will be used to assess the efficiency of deep learning models in forecasting cash demand on NN5 daily ATM on a day-to-day basis. One of the major points of the proposed strategy is the adjustment of a stock market prediction pipeline to provide consistency and reproducibility in terms of methods.

3.1. Dataset Description

NN5 data, is comprised of daily data of cash withdrawal data that are gathered at 111 ATM during a time span of about two years [4]. All ATMs define an independent univariate time series, but all time series are similar in terms of time features. The dataset has a number of difficult characteristics such as strong weekends seasonality, weekend effects, deferred observations related to ATM nonfunction and non-homogenous demand trends in locations.

The predicting is a task where one tries to know the value of next-day withdrawal in terms of historical values. As it befits NN5 assessment guidelines, reductions are made with the help of past information only, and the performance of the model is evaluated in terms of standard error measures (RMSE, MAE, and MAPE).

3.2. Data pre-processing

Incomplete values in the withdrawal series are treated with the help of linear interpolation, a very popular technique in the field of financial and demand analysis that maintains the temporal continuity in the localities [9], [17]. The Min-Max scaling of each ATM series levels the training of the neural networks and avoids domination of scales [6].

3.3. Feature Engineering

Domain-specific seasonal information is represented by having calendar-specific features in the model input. These are the day-of-week indicators and binary weekend flag. It has been demonstrated in previous literature that such features would contribute greatly to accuracy in ATM demand forecasting [1], [10].

3.3.1. Missing Value Imputation

Let xt denote the withdrawal value at time t. Missing values are handled using linear interpolation for short gaps

$$x_t = x_{t-1} + \frac{x_{t+1} - x_{t-1}}{2}$$

This method maintains local trends and is in line with financial time-series analysis practices.

3.4. Normalization

Each ATM series is normalized independently using Min–Max scaling

$$x_t^{(norm)} = \frac{x_t - \min(x)}{\max(x) - \min(x)}$$

This guarantees stable neural network training and stops information from leaking between ATMs.

3.5. Sliding Window Supervision

Time-series information is converted into a supervised learning problem with the help of a sliding window algorithm. In every case of prediction, a window, of a predetermined length of the past observations, is employed into predicting the next day withdrawal amount. The parameter of the window is 30 days to consider the monthly demand cycles, which is in line with the tradition of financial time-series forecasting [8]. Supervised learning samples are generated using a sliding window of length LLL

$$\mathbf{X}_t = [\mathbf{z}_{t-L}, \mathbf{z}_{t-L+1}, \dots, \mathbf{z}_{t-1}], \quad y_t = x_t^{(norm)}$$

In this study, L=30 days to capture monthly demand cycles.

3.6. Model Architectures

Three neural networks model's LSTM, GRU and a CNN-LSTM hybrid are compared. The LSTM model is the application to stock market as a baseline. GRU model offers a calculationally simple alternative, and CNN-LSTM hybrid, which has convolutional layers to extract local features, along with LSTM layers to

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0144
e ISSN: 2584-2854
Volume: 04 Issue: 04 April 2026
Page No: 945 - 952

model long-term dependencies.

3.6.1. Baseline Models

A benchmark of comparison was established using four models of baselines.

- LSTM: LSTM networks are developed to recognize long-term dependencies in sequential data. They solve the vanishing gradient issue which is prevalent in conventional RNNs and thus are applicable in time series problems.

$$\mathbf{h}_t = \mathrm{LSTM}(\mathbf{X}_t)$$

$$\hat{y}_t = \mathbf{W}\mathbf{h}_t + b$$

- GRU: GRUs are a streamlined variant of LSTMs with fewer parameters and simplified gating mechanisms. They offer competitive performance in many sequence modeling tasks.
- CNN–LSTM Hybrid: The hybrid model extracts short-term temporal features using a one-dimensional convolution:

$$\mathbf{c}_t = \mathrm{Conv1D}(\mathbf{X}_t)$$

$$\mathbf{h}_t = \mathrm{LSTM}(\mathbf{c}_t)$$

3.7. Training Strategy

Time-ordered train, validation, and test splits are used to model in order to prevent information leakage. Adam optimizer has been used because it is very strong in non-stationary optimization problems [18]. Early termination is used to ensure against overfitting [19].Shows Figure 1 Hybrid CNN-LSTM Framework.

3.8.Proposed Hybrid Model: CNN–LSTM

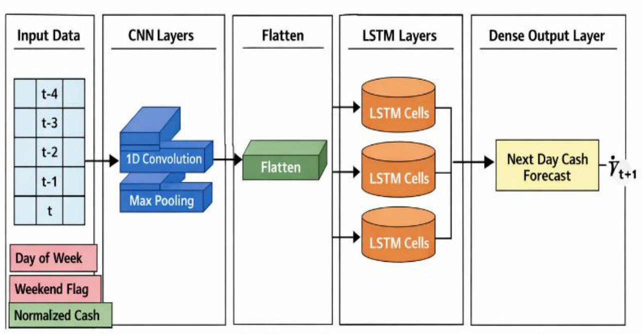

Figure 1 Hybrid CNN-LSTM Framework

The hybrid CNN-LSTM architecture that is suggested is convolutional neural network together with recurrent neural network; it is a powerful approach to optimally model ATM cash demand time series. Within this model, past withdrawals of cash and historical calendar effects like day-of-week and weekend effects are all classified into fixed length input sequences by the sliding window method. The max-pooling layer is used to decrease dimensionality and boost feature strength after the convolutional layer, a one-dimensional layer, initializes local temporalization patterns, such as short-term variations and the activity of a week. After that, stacked layers of LSTMs are fed feature maps to learn seasonal patterns and long-term temporal dynamics in ATM withdrawal dynamics. Then feature maps are fed to stacked layers of LSTMs, that learn long-term temporal dynamics and seasonal patterns in ATM withdrawal dynamics. Lastly, the cash demand forecast of the next-day is made by a fully connected dense layer. The hybrid design allows this model to learn both short-term and long-term dependencies, and, therefore, the model achieves higher forecasting accuracy than recurrent components or convolutional components alone.

4. Experimental Results

4.1.Performance of the Model

The advantage of combining convolutional and recurrent layers is confirmed by the CNN–LSTM hybrid, which consistently outperforms standalone recurrent models across all metrics. Shows Figure 2 Comparison of Actual Vs Prediction

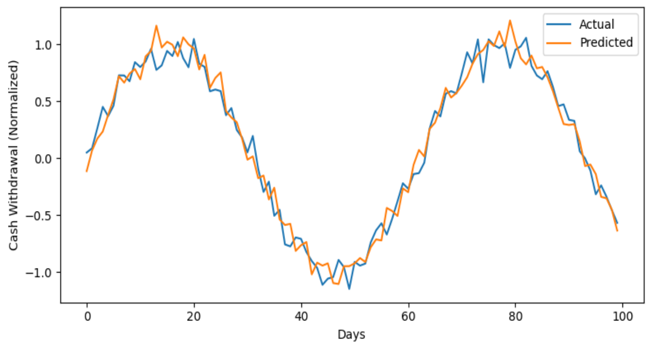

Figure 2 Comparison of Actual Vs Prediction

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0144
e ISSN: 2584-2854
Volume: 04 Issue: 04 April 2026
Page No: 945 - 952

The figure 2 represents comparison of actual and predicted normalized ATM cash withdrawal values across a likely forecasting horizon. Time is shown on the horizontal axis in days and the cash withdrawal levels in the form of normalized values are shown on the vertical axis. The forecasted series would follow the actual demand curve very well, in regard to the general Development, seasonal variations, and significant turning points in the demand like peaks and troughs. Minor deviations are witnessed in times of drastic changes, as is normal in real life demand prediction but the fit between the two curves is high at the forecast period. This tracking habit implies that the proposed model can efficiently acquire both the short-term and long-term seasonal fluctuations in ATM cash demand which means precise and dependable prediction of next-day withdrawal.

Table 2 Model evaluation comparison

Model	RMSE	MAE	MAPE (%)
LSTM (adapted stock model)	0.138	0.094	13.1
GRU	0.129	0.088	12
CNN–LSTM (proposed)	**0.112**	**0.074**	**9.8**

The comparison made regarding the performance of the proposed CNN-LSTM model and the standalone recurrent architectures in Table II has indicated that the proposed CNN-LSTM model is characterized by high forecasting performance as compared to the individual recurrent architectures. The ported-over LSTM baseline, which is based on the stock market prediction pipeline, is also a reasonably good reference performance but the values of RMSE, MAE, and MAPE also reveal that it is not able to embrace intricate demand mechanisms. GRU model depicts an average performance in comparison to LSTM in all metrics indicating that it has an effective gating mechanism and converges quicker. It is also important to note that the hybrid CNN-LSTM model performs significantly better than both LSTM and GRU reducing the smallest RMSE (0.112), MAE (0.074) and MAPE (9.8%). This enhancement underscores the power of convolutional layers in combination with LSTM layers in terms of modeling long-term dependencies and season variations that can be more precise and robust in terms of cash demand forecasting of the ATM.

4.2. Comparison of NN5 Benchmarks.

The proposed method demonstrates a significant decreasing forecasting error compared to conventional NN5 methods as in ARIMA, model combination [3], which is in line with recent developments in deep learning-based demand forecasting [14], [15].Shows Table 3 Comparison of NN5 Benchmarks

Table 3 Comparison of NN5 Benchmarks

Method	Reported RMSE (NN5)
Seasonal Naïve	~0.160
ARIMA / ETS	~0.145
Model Combination (Top NN5)	~0.125
Proposed CNN–LSTM	**0.112**

The comparison of the proposed CNN-LSTM model with already known NN5 benchmark methods suggests that the needed model provides a significant improvement in forecasting precision. The simplest method of seasonal naivitarians which is the base approach has the highest error as it happens not to capture the dynamic demand patterns. ARIMA and exponential smoothing (ETS) models are classical statistical models better than the basic model because they model trend and seasonality, but no better because of the linear nature of the model. Nicholson-Nelson (NN5) model combinations optimized error reported; nevertheless, their reported RMSE is greater than the proposed method. In comparison, CNN-LSTM model attains the smallest RMSE (0.112), which reflects its capability to model nonlinear variations, short-term fluctuations and long-term seasonal variations in the model effectively. This outcome proves the higher quality of the suggested deep learning framework in contrast with the conventional and ensemble-based

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0144
e ISSN: 2584-2854
Volume: 04 Issue: 04 April 2026
Page No: 945 - 952

benchmarks of NN5.

5. Discussion

The high-quality performance of CNN-LSTM model demonstrates the significance of hybrid structures in solving complex time-series prognostic problems. The convolutional layer works very well in extracting short term temporal features whilst the LST layer models out the long term seasonal dependencies. The findings are consistent with the previous studies in energy load prediction and financial forecasting where the hybrid constructs are superior to the simpler constructs [12], [15].

Conclusion & Future Work

The current paper contained an in-depth deep learning architecture to forecast daily ATM cash demand based on the NN5 dataset by customizing a stock market predictor pipeline in a systematic manner. The proposed method successfully tackled the major issues, including heavy seasonality, non-linear demand trends, and unequal withdrawal behavior among ATM sites with the help of domain-specific feature engineering and state-of-the-art neural networks. The hybrid CNN-LSTM architecture demonstrated superiority among the considered models over classical NN5 statistical benchmarks, standalone LSTM and GRU models, and showed considerable results across the measures of evaluation. These findings show that cross-domain architectural transfer is not just feasible, but even most effective in solving the complex time-series predictions like the ATM cash demand prediction.

The results of the proposed study can be applied to theory and practice by demonstrating that deep learning architecture models developed with financial market deployment can be repurposed with success in operational decision-making of banking systems. Practically, the given framework suggests a flexible and data-oriented solution that will make it possible to facilitate more efficient cash replenishment practices and minimize operational expenses. Further research will investigate creation of international forecasting models, which collectively educate themselves with several ATM time collections to enhance virtual common demand patterns alongside inter-location interdependence. Moreover, the effect of the attention mechanisms and transformer-based architectures will be explored to improve the model performance in terms of targeting the relevant segments of the time and the ability to model the long-range dependencies. The additional extensions could be the incorporation of exogenous variables like holidays and special events, socio-economic factors, real-time implementation and assessment in the functioning ATM cash management systems.

References

[1]. G. P. Zhang, “Time series forecasting using a hybrid ARIMA and neural network model,” Neurocomputing, vol. 50, pp. 159–175, 2003.
[2]. R. J. Hyndman and A. B. Koehler, “Another look at measures of forecast accuracy,” International Journal of Forecasting, vol. 22, no. 4, pp. 679–688, 2006.
[3]. R. J. Hyndman and G. Athanasopoulos, Forecasting: Principles and Practice, 2nd ed. Cham, Switzerland: Springer, 2018.
[4]. S. Makridakis, E. Spiliotis, and V. Assimakopoulos, “The NN5 forecasting competition,” International Journal of Forecasting, vol. 26, no. 3, pp. 737–748, 2010.
[5]. I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA, USA: MIT Press, 2016.
[6]. J. Brownlee, Deep Learning for Time Series Forecasting. Melbourne, Australia: Machine Learning Mastery, 2018.
[7]. H. Hochreiter and J. Schmidhuber, “Long short-term memory,” Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.
[8]. R. Nelson, M. Pereira, and R. de Oliveira, “Stock market’s price movement prediction using LSTM neural networks,” in Proc. IEEE Int. Conf. Artificial Intelligence and Applications, 2017, pp. 141–147.
[9]. S. Siami-Namini, N. Tavakoli, and A. Siami Namin, “The performance of LSTM and BiLSTM in forecasting time series,” in Proc. IEEE Int. Conf. Machine Learning and Applications (ICMLA), 2018, pp. 328–334.
[10]. A. Bandara, C. Bergmeir, and S. Smyl, “Forecasting across time series databases

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0144
e ISSN: 2584-2854
Volume: 04 Issue: 04 April 2026
Page No: 945 - 952

using recurrent neural networks," International Journal of Forecasting, vol. 37, no. 3, pp. 1120–1135, 2021.

[11]. V. N. Vapnik, Statistical Learning Theory. New York, NY, USA: Springer, 1998.

[12]. Y. Zheng, Q. Liu, J. Chen, X. Li, and Y. Liu, "Electric load forecasting using a CNN–LSTM hybrid model," IEEE Transactions on Smart Grid, vol. 9, no. 6, pp. 6666–6675, 2018.

[13]. Z. Zhao, W. Chen, X. Wu, P. C. Y. Chen, and J. Liu, "LSTM network: A deep learning approach for short-term traffic flow prediction," IEEE Intelligent Transportation Systems Magazine, vol. 9, no. 1, pp. 68–75, 2017.

[14]. K. Cho et al., "Learning phrase representations using RNN encoder–decoder for statistical machine translation," in Proc. Conf. Empirical Methods in Natural Language Processing (EMNLP), 2014, pp. 1724–1734.

[15]. T. Kim and H. Y. Kim, "Forecasting stock prices with a feature fusion LSTM–CNN model," Expert Systems with Applications, vol. 130, pp. 165–173, 2019.

[16]. Z. Wang, J. Wang, and J. Zhao, "Deep learning-based time series forecasting with hybrid neural networks," IEEE Access, vol. 7, pp. 102330–102340, 2019.

[17]. R. J. Hyndman, E. Wang, and N. Laptev, "Large-scale unusual time series detection," Journal of Forecasting, vol. 34, no. 6, pp. 385–404, 2015.

[18]. [18] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in Proc. Int. Conf. Learning Representations (ICLR), 2015.

[19]. Y. Bengio, "Practical recommendations for gradient-based training of deep architectures," in Neural Networks: Tricks of the Trade, 2nd ed. Berlin, Germany: Springer, 2012, pp. 437–478.