Journal of Intelligent Systems and Internet of Things
Vol. 18, No. 02, PP. 60-71, 2026
DOI: https://doi.org/10.54216/JISIoT.180204

Power Consumption Prediction Using a CNN-LSTM-Attention Hybrid Deep Learning Model

Nebras Jalel Ibrahim[1,*], Samah Faris Kamil[2], Ghasaq Saad Jameel[3]
[1]Computer Center, University of Diyala, Diyala, Iraq
[2] College of Chemical Engineering, University of Technology, Baghdad, Iraq
[3]Ministry of Education, Salah Al-Din, Iraq
Emails: nebras.jalel@uodiyala.edu.iq; samah.f.kamil@uotechnology.edu.iq; ghasaq.s.jameel@aliraqia.edu.iq

Abstract

Reducing energy losses and increasing power grid efficiency need accurate prediction of power consumption accurate prediction of future energy consumption requires the use of time series data. To overcome the shortcomings of conventional techniques for forecasting energy consumption in India for the period from 2 January, 2019 to 23 May, 2020, we used an attention mechanism, which is still relatively new and not well known. In this paper, we propose a new approach for predicting energy consumption by combining local feature extraction with convolutional neural networks (CNNs), long short-term memory (LSTM) to capture long-term temporal dependencies, and attention mechanisms to deal with the issue of information loss brought on by extremely lengthy input time series data. After high-dimensional features are extracted from the input data using a one-dimensional CNN layer, temporal correlations within historical sequences are captured using an LSTM layer. In order to optimize the weighting of the LSTM outputs, strengthen the impact of important information, and enhance the prediction model as a whole, an attention mechanism is finally implemented. This integration improves the model's ability to represent complex spatio-temporal patterns. The mean absolute error (MAE) and root mean square error (RMSE) are used to assess the performance of the proposed model. The results demonstrate that the CNN-LSTM-Attention model outperforms conventional hybrid CNN-LSTM and LSTM models, demonstrating superior performance across a range of prediction scenarios. By supporting more reliable grid management, proactive intervention methods, and predictive maintenance, these developments contribute to reducing load imbalances and energy waste in India. The Future developments could see the proposed model extended to other time series prediction domains.

Received: March 31, 2025 Revised: June 24, 2025 Accepted: August 20, 2025

Keywords: Power consumption prediction; CNN; LSTM; Attention mechanism; Deep learning; Hybrid model

1. Introduction

India's rapid urban growth and industrial development have significantly increased the nation's electricity demand, leading to considerable challenges in energy forecasting and effective power grid management [1]. To enhance energy distribution, reduce operational expenses, strengthen grid reliability, and promote sustainable energy strategies, precise prediction of power consumption is essential. The complex, nonlinear, and multi-scale temporal trends found in power usage data are often challenging for traditional statistical models and even standalone deep learning methods to accurately capture.

Historically, prediction power consumption has depended on statistical models such as ARIMA, exponential smoothing, and regression analysis, along with conventional machine learning methods like Support Vector Machines (SVM) and Artificial Neural Networks (ANN) [2]. Although these approaches have attained satisfactory accuracy in different scenarios, they frequently find it challenging to completely capture the highly non-linear relationships and intricate temporal dynamics present in contemporary power consumption data, particularly when impacted by a variety of interacting factor.

In recent times, deep learning models have demonstrated exceptional performance in tasks related to sequence modeling and time series forecasting across various fields [3]. Their capability to automatically learn hierarchical features and grasp long-term dependencies from extensive datasets makes them especially effective for intricate time series prediction challenges such as forecasting power consumption. Recurrent Neural Networks (RNNs), particularly their variants like Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU), are specifically designed to effectively manage sequential data, overcoming the shortcomings of conventional techniques in capturing complex temporal patterns [4]. The main contributions of this paper are:

- Designing new hybrid model that combines CNN , LSTM and Attention mechanism. To effectively classify daily power consumption patterns in India.The architecture is designed to take advantage of the contextual significance of Attention, temporal interdependence for LSTM and spatial features for CNN present in time series data.
- Comparing the proposed CNN–LSTM–Attention model with baseline LSTM and CNN–LSTM models allows for a thorough assessment of performance gains. Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and confusion matrix are among the metrics used for Comprehensive evaluation.
- Reduces overfitting and enhances prediction accuracy in India
- using areal world dataset puplicy accessible of power consumption by Indian state between 2019 and 2020 . The challenge is more feasible for real time energy management scenarios.
- This paper contributes to more effective and intelligent energy planning by offering a framework that may be expanded for demand side management in smart cities, energy forecasting dashboards and smart grid applications.

The rest of this paper is organized as follows: Section 2 provides a comprehensive review of the existing literature on power consumption prediction and classification, emphasizing the evolution of machine learning and deep learning models, particularly hybrid models. Section 3 describes the metholodgy time series and the dataset used in this paper, deep learning models and architecture of the proposed CNN–LSTM–Attention model, outlining each component and the rationale for integrating convolutional, sequential, and attention-based mechanisms. Section 4 provides more details on the experimental setup, including evaluation metrics, and Model Performance Visualization a comparative analysis of the proposed model against baseline models. Section 5 presents the Results and Discussion, highlighting the model's predictive performance and the limitations of the current paper. Finally, Section 6 addresses Conclusion and future work and indicates possible avenues for future research.

2. Literature Review

Power consumption prediction has become really important lately because we're using more energy and need better ways to manage it. Many studies have looked at different methods to predict energy use, from basic statistical models to more complex machine learning and deep learning techniques [5]. Pillalamarri Madhavi and S. Satyanarayana [6] studied India's electrical generation and consumption for 2019-2020 using supervised ML algorithms and analyzed prediction accuracy.Manish Uppal et al. [7] analyzed Delhi’s discoms’ load during COVID-19 using ensemble methods with machine learning and weather data to handle unusual consumption patterns from the lockdown.Manu Suvarna et al. [8] used machine learning frameworks to quantify the impact of COVID-19 lockdowns on electricity consumption in seven Indian states, integrating weather and econometric data.Saikat Gochhait and D. Sharma [9] tested multiple regression models on a Maharashtra dataset, finding Gaussian Process Regression (GPR) to be the most accurate.Pradeep K et al. [10] used a Random Forest algorithm to predict energy consumption for Indian states.

Rashmi Bareth and Anamika Yadav [11] proposed an LSTM model for day-ahead forecasting using historical power data from Chhattisgarh. Rashmi Bareth et al [12]. also forecasted monthly power demand in Chhattisgarh with LSTM, showing lower error rates than traditional methods.V. Veeramsetty et al. [13] utilized Principal Component Analysis (PCA) and Recurrent Neural Networks for load prediction in an Indian substation, thereby enhancing computational efficiency.Saji et al. analyzed 10 years of Kerala's electricity demand, finding Random Forest most accurate[14].In Telangana, Jayashree S and R. A [15] compared ARIMA, LSTM, and CNN models for demand prediction. LSTM offered a better fit for observed data than ARIMA and CNN. L. D and Abhishek Srivastava [16] compared various ensemble methods (RF, XGBoost, AdaBoost, etc.) and deep learning models (LSTM, GRU, RNN, Prophet) for Delhi's 2020 load forecast, utilizing weather and calendar features. CNN was proposed by Kasun Amarasinghe et al. [17] to forecast energy loads at the building level.A CNN-LSTM neural network that can extract spatial and temporal data to accurately estimate the energy consumption of housing was proposed by Tae-Young Kim and Sung-Bae Cho [18].Yahya Hafedh Abdulameer and Abdullahi Abdu Ibrahim [19] suggested a hybrid forecasting model that makes use of CNN-based and LSTM deep learning architectures for predicting.Table 1 summarizes the key findings of the power consumption prediction methods mentioned above. There are still not many specific studies that apply and compare sophisticated hybrid deep learning architectures, like CNN-LSTM models enhanced with attention mechanisms, to large-scale regional or national power consumption data from India, especially for the recent and unusual 2019–2020 period. By providing a direct empirical comparison of these advanced models using pertinent Indian datasets, Therefore, it is worth analyzing power consumption forecasting with such a hybrid model.

Table 1: Summary of literature on power consumption in india

Reference	Region/Coverage	Dataset Description	Period	ML/DL Algorithms Used
[1]	India (national/sectors)	Power generation and consumption data	2019-2020	Supervised ML (various)
[2]	Delhi (DISCOMs)	Load and weather data	2020-2021	Ensembles, ML
[3]	7 Indian states	Electricity, weather, social distancing	2020-2021	ML models
[4]	Maharashtra (state)	Half-hourly load data from MSEDCL	Jul 2020-Aug 2022	Regression (24 models)
[5]	Delhi	Load, weather, calendar data	2020, prior data	Ensembles, DL (LSTM, GRU, etc.)
[6]	Chhattisgarh (state)	One-year historical consumption data	Not stated	LSTM
[7]	Chhattisgarh (state)	Daily/monthly load averages	2018-2022	LSTM
[8]	Indian states (multiple)	Regional energy consumption data	Not stated	Random Forest
[9]	India (substation level)	33/11 kV substation measured loads	Not stated	PCA + RNN
[13]	Kerala (state)	Electricity demand data from KSEB (10 years)	2013-2022	RF, XGBoost, others
[16]	Telangana (state)	Power consumption, industrial-HT sector data	2014-2022	LSTM, ARIMA, CNN

3. Methodology

This section present the proposed model for prediction power consumption in india including Times series,dataset description,proposed deep learning model.

3.1 Times series

All the proposed models in this paper fall into the time-series model. A time series refers to a sequence of data points indexed in time order. Figure 1 as shown time series of power consumption prediction, it typically involves historical energy consumption data collected over regular intervals (e.g., hourly, daily, monthly) [20]. Deep learning models well-suited for analyzing such data because they can capture the sequential dependencies and patterns that evolve over time. A time series can be mathematically represented as in equation (1) :

$$y(t) = y(1), y(2), \dots \dots \dots, y(n) \qquad (1)$$

Where $y(t)$ is the observation at time t and the data is typically collected at regular intervals (e.g., hourly, daily, monthly).

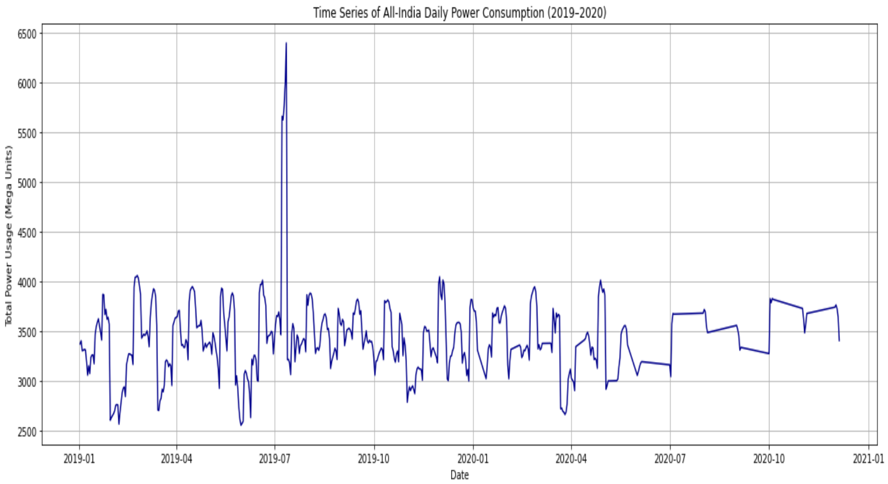

Figure 1. Time Series of All-India Daily Power Consumption (2019-2020)

3.2 Dataset Description

The dataset provides a comprehensive overview of India's energy usage by state from 2019 to 2020 . The time series data, which spans 17 months from January 2, 2019, to May 23, 2020, emphasizes the importance of the data for forecasting model training, economic assessments, and energy policy analysis. It displays a useful dataset of state-by-state power consumption in Mega Units (MU) that is accessible on Kaggle [21].This dataset's trustworthiness is increased by the fact that it comes from the Power System Operation Corporation Limited's (POSOCO) weekly energy reports. In order to promote study and analysis in the energy sector, it also makes reference to an official Ministry of Power resource that offers more comprehensive annual information, such as energy availability and demand metrics for the fiscal year. For those involved who want to enhance energy management and influence policy choices in India, the information is essential [22]. This data set include two sub-dataset : dataset-tk has 43 column or features represent India cities and row represents time series. The second dataset called Long- data incudes 6 features.

3.4 Deep Learning Models

To assess the effectiveness of deep learning for power consumption prediction, we implemented and compared the following architectures and choose the proposed model:

3.4.1 Long Short-Term Memory (LSTM) Network

An LSTM is a form of recurrent neural network (RNN) seen in Figure 2 and was created specifically to learn from sequences with long-range dependencies [23].By using memory cells and gating methods, LSTM can solve the vanishing gradient problem more successfully than conventional RNNs. These characteristics make LSTM networks ideal for time series forecasting applications like power consumption prediction since they enable them to preserve and refresh a recollection of historical data [24].

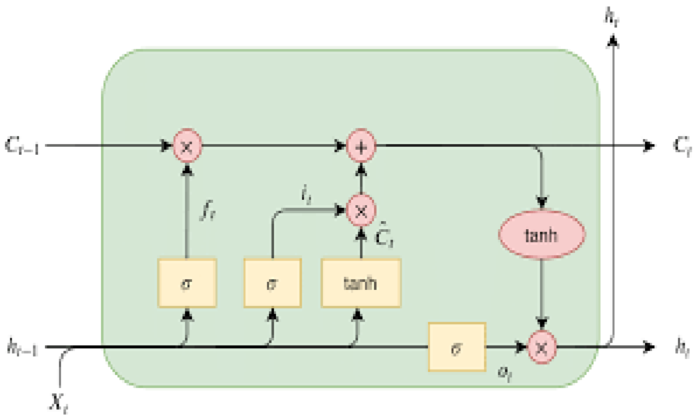

Figure 2. Structure of LSTM model

3.4.2 Hybrid CNN-LSTM Model

The Hybrid CNN-LSTM model, as depicted in Figure 3, is a sophisticated deep learning architecture that combines Long Short-Term Memory (LSTM) networks and Convolutional Neural Networks (CNNs) in order to effectively utilize both models' capabilities for processing data with temporal and spatial properties [25].The major goal of this hybrid technique is to succeed at tasks involving both local characteristics and sequential dependencies, such as time series classification, speech recognition, and video analysis, among others.This design effectively captures data patterns by using convolutional layers to extract local hierarchical features after CNNs have processed raw data [26]. The model may then monitor how particular properties evolve over time thanks to the input of these features into LSTM layers, which are experts at understanding temporal sequences and long-term relationships. This integration boosts robustness to noisy input patterns, automates feature engineering and improves prediction accuracy (often outperforming separate models).

The hybrid CNN-LSTM model is particularly well-known for applications such as predicting power consumption classes, financial forecasting, and evaluating machine remaining useful life, exhibiting its versatility in a wide range of difficult, data-intensive tasks. Overall, this model demonstrates the potential of mixing several neural architectures to address multidimensional problems in machine learning [27].

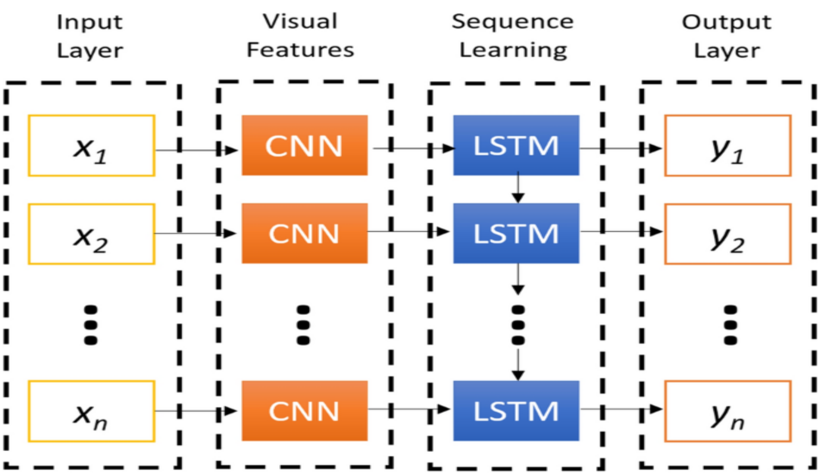

Figure 3. Hybrid CNN-LSTM model

3.4.3 Attention Mechanism

The Attention Mechanism represents a significant breakthrough in deep learning, especially in areas like Natural Language Processing (NLP), computer vision, and time series analysis. It tackles the challenges faced by traditional sequence-to-sequence models, particularly their limitations with lengthy sequences owing to fixed-size context vectors [28]. The attention Mechanism allows models to selectively concentrate on pertinent sections of data through query, key, and value components to compute attention scores and generate context vectors. Variants such as self-attention and multi-head attention provide enhanced flexibility and improved performance. The implementation of attention techniques has resulted in higher accuracy, better interpretability, and faster training times, transforming fields like machine translation, image recognition, and predictive modeling in time series. Attention is useful in power consumption prediction because it helps identify essential time steps, lowers noise and increases accuracy by focusing on relevant features [29].

3.4.4 The Proposed Hybrid Model for Power Consumption Prediction

The proposed model is made to accurately prediction power consumption in India. Figure 4. show a deep learning architecture designed for sequence processing that combines CNN,LSTM and an attention Mechanism. The model starts with an input sequence where data like time series, text, or audio enters. The first layer is a 1D CNN, which identifies local patterns by using convolutional filters along a single dimension. Next, a max pooling layer reduces the dimensionality of feature maps while keeping important features, thus enhancing robustness and reducing complexity.By processing data sequentially and capturing long-term dependencies, the LSTM layer which is coupled to the pooling layer resolves problems with conventional RNNs. By allocating various weights, the attention mechanism improves predictions for long sequences by concentrating on the most pertinent portions of the input after the LSTM.

A Dense layer transforms the attention mechanism's output into an appropriate representation for the final predictions. The model's output which can be any categorization, value, or sequence depending on the job, is generated by the last layer.This hybrid model greatly improves prediction performance while successfully reducing the drawbacks of separate deep learning models. For sophisticated power consumption predictions in actual smart grid systems, the model thus turns out to be a reliable and scalable solution.

Figure 5 shows the workflow for estimating power consumption with a hybrid deep learning model . The dataset, which includes auxiliary features and historical power usage, is presented first, followed by a description of the data pretreatment procedures, including feature engineering and handling missing values. The dataset is then divided into subgroups for testing and training. The core of the approach is a hybrid model that combines CNN, LSTM and an attention mechanism to efficiently capture temporal patterns in the data.The ultimate objective is to produce accurate power consumption projections for applications like demand forecasting and energy management by evaluating model accuracy using performance evaluation metrics like RMSE and MAE.

The proposed model reorders cities in the database based on their power onsumption, from lowest to highest, as illustrated in Figure 6. This aids in identifying high and low power consumption cities, so providing insights for future plans and strategies to efficiently manage power consumption, particularly in high consumption areas.

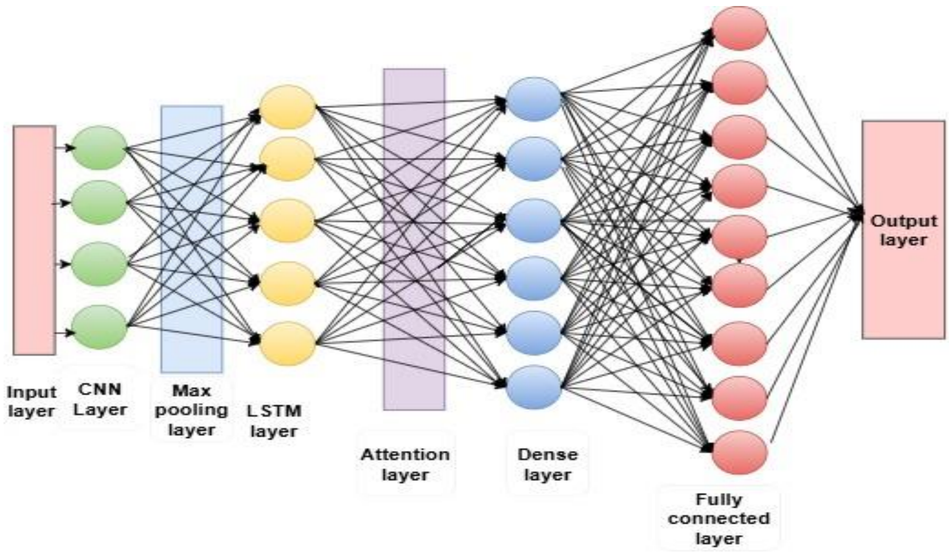

Figure 4. The proposed model's architecture for power consumption prediction

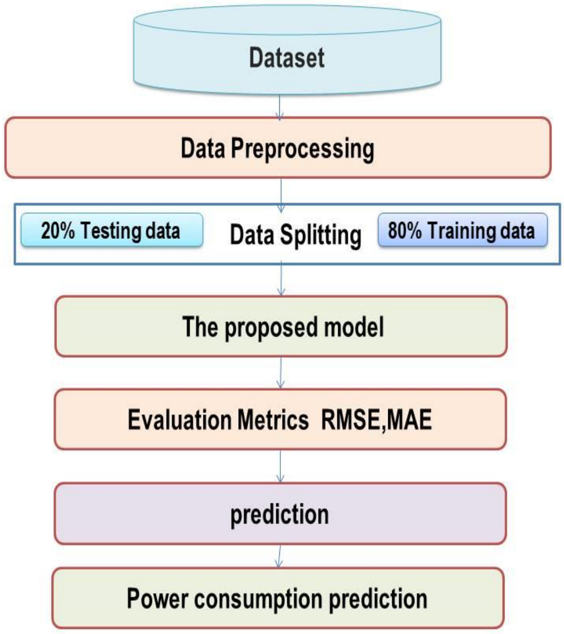

Figure 5. Methology of the proposed deep learning for power consumption prediction

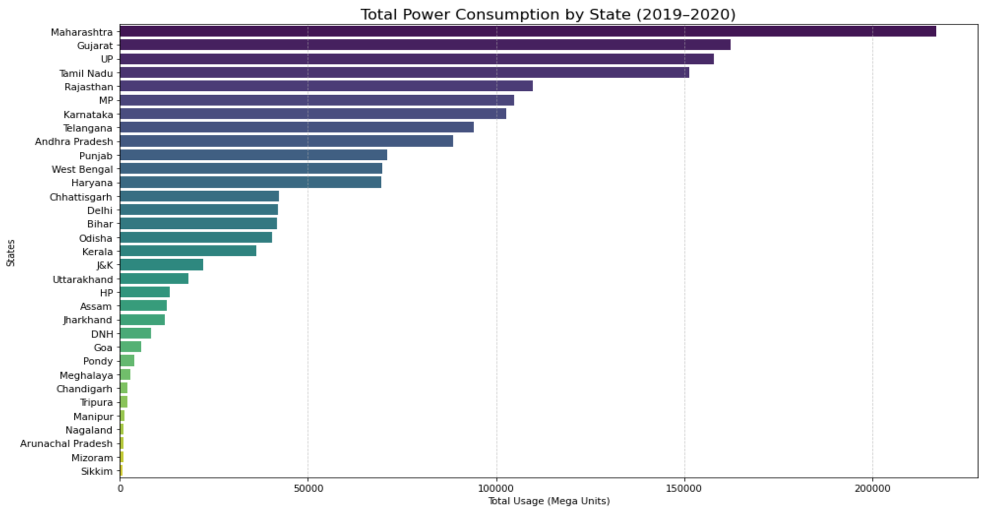

Figure 6. Power consumption in Indian states

4. Experiments

4.1 Experimental Setup

To ensure consistency and reliability in the results, all experiments were conducted under the same computational and software environment.Table 2 below summarized the experimental setup includes both hardware and software configurations. The parameters for the specified components of the proposed deep learning model are summarized in Table 3 below.

Table 2: Experimental Setup

Component	**Specification**
Hardware	Intel Core i7 processor, 16 GB RAM, NVIDIA GPU
Software Environment	Python 3.9
Libraries Used	TensorFlow 2.12, scikit-learn, pandas, matplotlib
Number of Epochs	100
Batch Size	64
Optimizer	Adam

Table 3: The parameters setting of the proposed model

Layers	Parameters
CNN Layer	64
Pooling Layer	pool size = 2
LSTM Layer	64, return sequences = True
Attention Mechanism	Self-Attention
Dropout Rate	0.3
Dense	4 units

4.2 Evaluation Metrics

Numerous factors, such as the testing portion settings, prediction methods, and the reliability of the data source, can impact a forecast's accuracy. To evaluate the quality generic measures must be used in forecasts. The most crucial statistic is the accuracy measure, which provides immediate insight into a prediction model's performance.

Root-Mean-Square Error (RMSE) is widely used on the test dataset to assess how well comparison models perform. I refer to the time step here [30]. For the performance measurement, the test dataset's average RMSE was acquired in equation (2).

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(A_t - F_t)^2} \qquad (2)$$

Mean Absolute Error (MAE) is a basic measure used in regression that checks how far off predictions are from the actual values, without worrying about whether the predictions are too high or too low [31]. You get it by averaging the absolute differences between what was predicted and what really happened in equation (3).

$$MAE = \frac{1}{n}\sum_{t=}^{n}|A_t - F_t| \qquad (3)$$

4.3 Model Performance Visualization

4.3.1 Line plots comparing actual and predicted power consumption

Three deep learning models hybrid model (CNN-LSTM), LSTM Model, and hybrid model (CNN-LSTM-attention) are compared in figure 7 using a series of line plots that display the models' predicted performance over time against actual power consumption classes. The accuracy of these algorithms in prediction power usage is demonstrated by the tight alignment of the projected and actual numbers in each plot. The best fit is notably shown by the hybrid Model (CNN-LSTM-attention), indicating that the attention mechanism improves prediction accuracy. All things considered, the investigation shows how well these models can predict outcomes despite the dynamic nature of power consumption data.

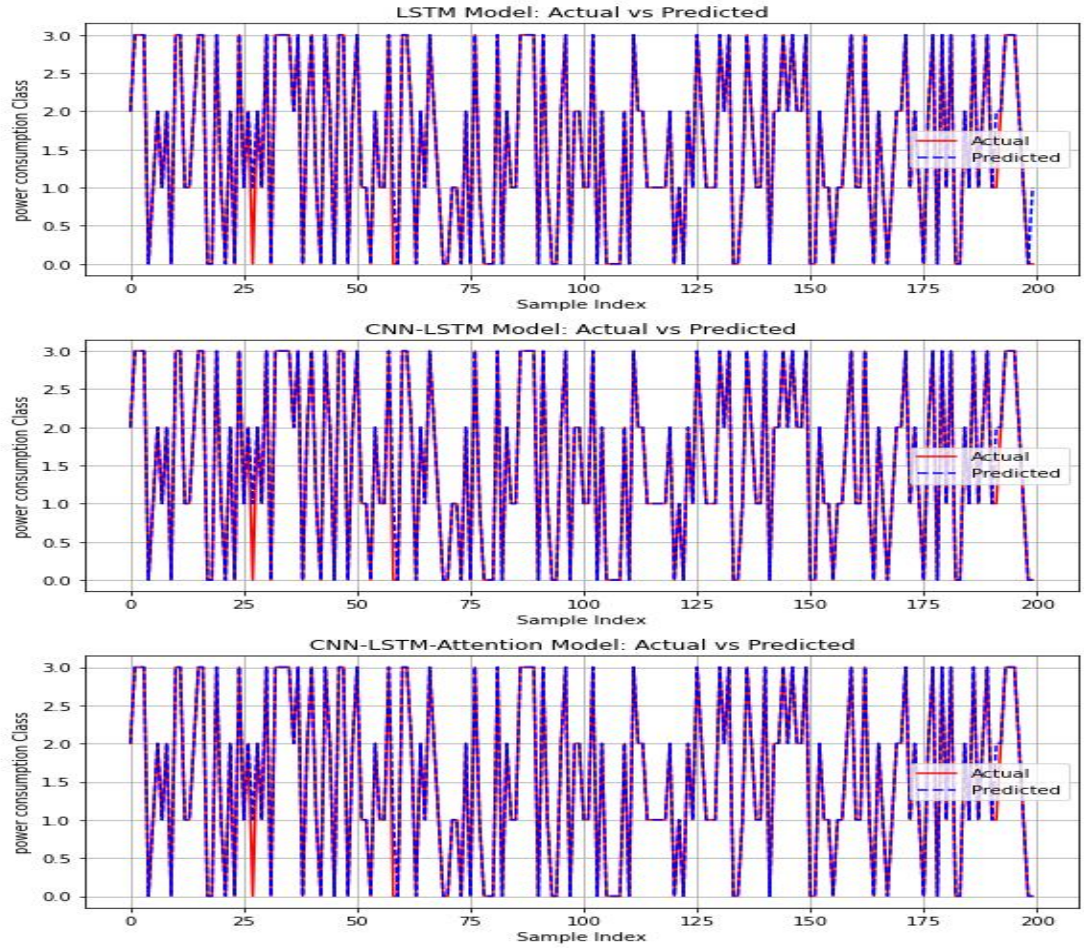

Figure 7. The model’s effectiveness in predicting power consumption

4.3.2 Loss curves for training and validation performance

Three graphs illustrating the training and validation loss curves for several deep learning models over 100 epochs are shown in figure 8. These models include a CNN - LSTM-Attention model, a hybrid model that combines CNN and LSTM, and a solo LSTM model. Plot observations show that all models learn well, as seen by declining validation and training losses. Remarkably, when compared to standalone LSTM, hybrid models typically attain lower loss values, indicating enhanced performance via the use of multiple architectures and attention mechanism.

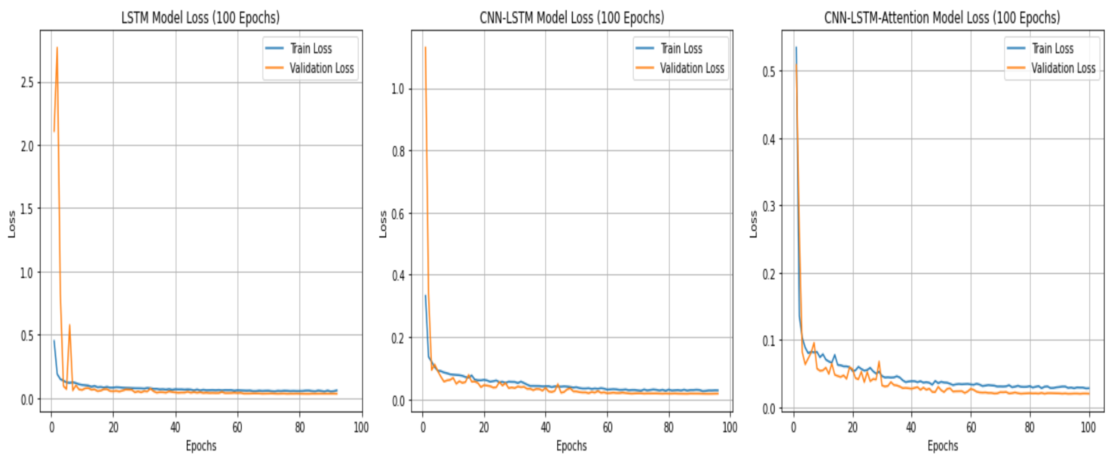

Figure 8. Training and validation loss for models

4.3.3 Confusion matrix-style heatmaps for model classification accuracy

The classification performance of three deep learning models LSTM, CNN-LSTM, and CNN-LSTM-Attention is displayed in Figure 9 by the confusion matrices. applying confusion matrices to a four-class multi-class classification job. Each model's confusion matrix is represented as a heatmap that visually conveys the number of correct and incorrect predictions. The LSTM model, while having good accuracy, demonstrated the highest number of misclassifications, particularly in classes 0, 1, and 2. In contrast, the CNN-LSTM model performed better and made fewer mistakes in all of these classes. Overall, the results show that CNN-LSTM-Attention performs better in terms of classification accuracy than both the LSTM and CNN-LSTM models.

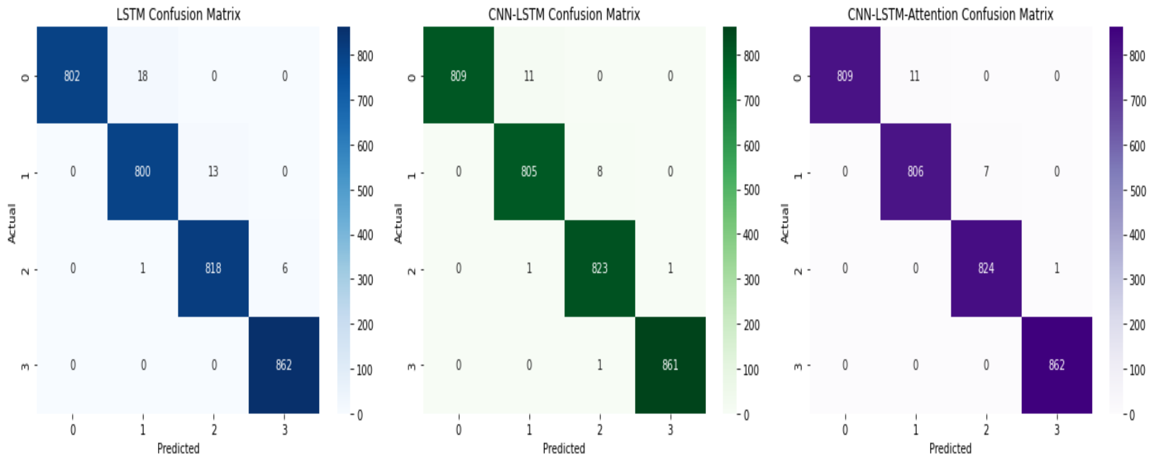

Figure 9. Confusion matrix for models

5. Results and Discussion

Table 4. below summarizes the results on the test set and shows the comparative results of deep learning models used to predicate power consumption in India from 2019 to 2020. We analyze each model's performance with standard metrics,the models were evaluated using MAE and RMSE. The proposed hybrid model CNN-LSTM-Attention is the best in terms of both MAE and RMSE and should be preferred for power consumption prediction.The Comparative result with previous studies is illustrated in Table 5. Although our study presents a new approach on predicting power consumption using deep learning,there are limitations to address. These drawbacks offer worthwhile chances for additional research in this area.The key shortcomings of our proposed solution are presented below:

1. Range of Applicability : The study's data comes from a narrow time period (2019-2020) and a particular region (India). This restricts the findings' applicability to longer term forecasts or to other regions with distinct consumption patterns.

2. Absence of External Variables: Only historical consumption values were used to train the models. The absence of significant exogenous factors including temperature, industrial activity, economic indicators, holidays, and population dynamics may have decreased forecast accuracy in situations involving abrupt external changes.

Table 4: The evaluation metrics for each model

Model	MAE	RMSE
LSTM	0.0114	0.1070
CNN-LSTM	0.0071	0.0814
The proposed model	0.0057	0.0756

Table 5: Comparative result with previous studies

Reference	Method	dataset	MAE	RMSE
[17]	CNN	Individual household electric power consumption dataset	Non	0.732
[18]	CNN-LSTM	individual household power consumption dataset	0.3493	0.6114
[19]	Hybrid GRU-CNN-BiLSTM	household electricity usage data in Baghdad	Non	0.094603
Our proposed model	LSTM-CNN-Attention	Power Consumption in india(2019-2020)	0.0057	0.0756

6. Conclusion and future work

This paper uses data from 2019 to 2020 to demonstrate the use of deep learning to prediction power consumption in India.Future power consumption forecasting is now crucial for enhancing infrastructure, usage efficiency, and collaboration between smart grids and home energy. A hybrid CNN-LSTM-Attention, A Hybrid LSTM -CNN, and a simple LSTM were the three configurations that were tested. Although they varied in accuracy, all models were able to predict electricity use. The proposed model shows better prediction accuracy compared with the LSTM,CNN-LSTM . Based on MAE and RMSE, the proposed CNN-LSTM-Attention model performed the best results that integrating attention with various deep learning methods can enhance complex time series predictions. To record temporal dynamics, long-range dependencies and local information. this paper highlights the advantages of hybrid models over conventional recurrent networks **.**For future work, we plan to explore to improve power consumption prediction models.Prediction accuracy can be increased by incorporating outside factors like humidity, temperature, and economic indicators. It will be easier to capture more consistent consumption patterns if datasets are expanded to encompass recent years, particularly those following COVID-19. Instead of using set thresholds, models should employ dynamic categorization boundaries that are adapted to particular requirements. Understanding regional consumption patterns can be enhanced by putting spatial temporal modeling tools into practice. Their application in power utilities will be supported by optimizing models for real time deployment. Continued research in this area could lead to more comprehensive solutions for reducing power consumption and cost.

References

[1] G. Dave, T. Patel, and S. Mitra, "Global scenario on renewable energy and net zero goal emphasizing Indian standpoint: a short review," *Energy Syst.*, pp. 1-22, 2025.

[2] A. Kurani, P. Doshi, A. Vakharia, and M. Shah, "A comprehensive comparative study of artificial neural network (ANN) and support vector machines (SVM) on stock forecasting," *Ann. Data Sci.*, vol. 10, no. 1, pp. 183-208, 2023.

[3] Z. Han, J. Zhao, H. Leung, K. F. Ma, and W. Wang, "A review of deep learning models for time series prediction," *IEEE Sensors J.*, vol. 21, no. 6, pp. 7833-7848, 2019.

[4] M. Waqas and U. W. Humphries, "A critical review of RNN and LSTM variants in hydrological time series predictions," *MethodsX*, p. 102946, 2024.

[5] P. W. Tien, S. Wei, J. Darkwa, C. Wood, and J. K. Calautit, "Machine learning and deep learning methods for enhancing building energy efficiency and indoor environmental quality–a review," *Energy AI*, vol. 10, p. 100198, 2022.

[6] P. Madhavi and S. V. Satyanarayana, "Data Analysis of electrical systems using machine learning algorithms," in *2022 First Int. Conf. Artif. Intell. Trends Pattern Recognit. (ICAITPR)*, 2022, pp. 1-4.

[7] M. Uppal, R. Kumari, and S. Shrivastava, "An Ensemble Approach for Short-Term Load Forecasting for DISCOMS of Delhi Across the COVID-19 Scenario," in *2021 9th IEEE Int. Conf. Power Syst. (ICPS)*, 2021, pp. 1-6.

[8] M. Suvarna, A. Katragadda, Z. Sun, Y. B. Choh, Q. Chen, and X. Wang, "Advances in Applied Energy," *Adv. Appl. Energy*, 2021.

[9] S. Gochhait and D. Sharma, "Regression model-based short-term load forecasting for load despatch centre," *J. Appl. Eng. Technol. Sci. (JAETS)*, vol. 4, no. 2, pp. 693-710, 2023.

[10] K. Pradeep, T. Likhita, V. Sai Pavan, and A. Raghu Vira Pratap, "Machine-learning based approach to predict energy consumption of India States," in *2022 6th Int. Conf. Comput. Methodol. Commun. (ICCMC)*, 2022, pp. 1153-1156.

[11] R. Bareth and A. Yadav, "Day Ahead Load Demand Forecasting based on LSTM Machine Learning Model," in *2024 Third Int. Conf. Power, Control Comput. Technol. (ICPC2T)*, 2024, pp. 404-408.

[12] R. Bareth, A. Yadav, S. Gupta, and M. Pazoki, "Daily average load demand forecasting using LSTM model based on historical load trends," *IET Gener., Transmiss. Distrib.*, vol. 18, no. 5, pp. 952-962, 2024.

[13] V. Veeramsetty, D. R. Chandra, F. Grimaccia, and M. Mussetta, "Short term electric power load forecasting using principal component analysis and recurrent neural networks," *Forecasting*, vol. 4, no. 1, pp. 149-164, 2022.

[14] A. Saji, A. Prakash, and M. S. Krishnan, "Electricity Demand Forecasting In Kerala Using Machine Learning Models," in *2023 3rd Int. Conf. Intell. Technol. (CONIT)*, 2023, pp. 1-6.

[15] S. Jayashree and A. Rajini, "Forecasting power consumption by deep learning approaches in the state of telangana," in *2023 3rd Int. Conf. Intell. Technol. (CONIT)*, 2023, pp. 1-7.

[16] D. Lakshmi, R. S. Tiwari, N. Nagpal, N. Kassarwani, G. Vishnuvarthanan, and A. Srivastava, "Comparative study on forecasting of schedule generation in delhi region for the resilient power grid using machine learning," *IEEE Trans. Ind. Appl.*, vol. 60, no. 2, pp. 2107-2116, 2023.

[17] K. Amarasinghe, D. L. Marino, and M. Manic, "Deep neural networks for energy load forecasting," in *2017 IEEE 26th Int. Symp. Ind. Electron. (ISIE)*, 2017, pp. 1483-1488.

[18] T. Y. Kim and S. B. Cho, "Predicting residential energy consumption using CNN-LSTM neural networks," *Energy*, vol. 182, pp. 72-81, 2019.

[19] Y. H. Abdulameer and A. A. Ibrahim, "Forecasting of Electrical Energy Consumption Using Hybrid Models of GRU, CNN, LSTM, And ML Regressors,".

[20] N. Bhoj and R. S. Bhadoria, "Time-series based prediction for energy consumption of smart home data using hybrid convolution-recurrent neural network," *Telemat. Inform*, vol. 75, p. 101907, 2022.

[21] *State-wise-power-consumption-in-india* [Dataset]. Kaggle. Available: https://www.kaggle.com/datasets/twinkle0705/state-wise-power-consumption-in-india

[22] B. S. Vardhan, A. Swain, M. Khedkar, I. Srivastava, and N. D. Bokde, "An overview of Indian power sector and its Energy Management," *Renew. Energy Focus*, p. 100597, 2024.

[23] Z. Liu, Z. Zhu, J. Gao, and C. Xu, "Forecast methods for time series data: A survey," *IEEE Access*, vol. 9, pp. 91896-91912, 2021.

[24] C. L. Liu, T. Y. Chang, J. S. Yang, and K. B. Huang, "A deep learning sequence model based on self-attention and convolution for wind power prediction," *Renew. Energy*, vol. 219, p. 119399, 2023.

[25] A. Tasdelen and B. Sen, "A hybrid CNN-LSTM model for pre-miRNA classification," *Sci. Rep.*, vol. 11, no. 1, p. 14125, 2021.

[26] A. J. Yousif and M. H. Al-Jammas, "Real-time Arabic Video Captioning Using CNN and Transformer Networks Based on Parallel Implementation," *Diyala J. Eng. Sci.*, pp. 84-93, 2024.

[27] T. T. Hoa, T. M. Le, and C. H. Nguyen-Dinh, "Hybrid model of 1D-CNN and LSTM for forecasting Ethereum closing prices: a case study of temporal analysis," *Int. J. Inf. Technol.*, pp. 1-13, 2025.

[28] D. Soydaner, "Attention mechanism in neural networks: where it comes and where it goes," *Neural Comput. Appl.*, vol. 34, no. 16, pp. 13371-13385, 2022.

[29] A. J. Yousif and M. H. Al-Jammas, "Semantic-based temporal attention network for Arabic Video Captioning," *Nat. Lang. Process. J.*, vol. 10, p. 100122, 2025.

[30] A. Wan, Q. Chang, A. B. Khalil, and J. He, "Short-term power load forecasting for combined heat and power using CNN-LSTM enhanced by attention mechanism," *Energy*, vol. 282, p. 128274, 2023.

[31] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, "Convolutional neural networks: an overview and application in radiology," *Insights Imaging*, vol. 9, no. 4, pp. 611-629, 2018.