Paper ID: **IEMPOWER2025 103**

Smart Energy optimization for Educational buildings Using Machine Learning: A Case Study of IEM Management House

Nirban Chakraborty [1*] Priyanjana Paul [1], Swetasree Chattoraj [1], Moupiya Bose [1], Pratik De Sarkar[1], Subhajit Kar[1]

Department of Electrical and Electronics Engineering, Institute of Engineering & Management, Kolkata, Kolkata, 700091 University of Engineering & Management, Kolkata, India, 700091

**Corresponding author Email*: nirban.chakraborty@iem.edu.in.

Abstract

Energy optimization is essential for sustainability because energy expenditures make up about 40% of operating costs in many educational institutions. With the aim of cutting expenses and boosting productivity, this study offers a machine learning (ML)-based system for predicting and optimizing energy usage at the IEM Management House, Sector V, Kolkata. An extensive dataset that included more than a year's worth of historical data on weather, occupancy patterns, energy usage, and load fluctuations was used. To precisely forecast electricity demand, sophisticated machine learning techniques were used, such as LSTM models enhanced with calendar and meteorological data. Energy usage and expenses were reduced by 34% and 38%, respectively, as a result of the optimization technique. With a Mean Absolute Error (MAE) of 12.4 kWh, Root Mean Squared Error (RMSE) of 18.7 kWh, and Mean Absolute Percentage Error (MAPE) of 3.2%, the model showed excellent prediction accuracy. Furthermore, including solar energy systems promoted sustainable behaviors by lowering dependency on traditional power sources. This study offers a data-driven, scalable method for managing institutional energy that has quantifiable positive effects on operations and the environment.

Keywords: *Energy Optimization, Machine Learning (ML), Load Forecasting, Sustainable Energy Management, Solar Integration*

1. Introduction

The increasing integration of renewable energy into power grids presents dual opportunities and challenges for modern energy systems. While renewable sources like solar and wind offer significant environmental benefits, their intermittent nature complicates grid stability. Accurate forecasting and real-time management have become critical for ensuring energy reliability and efficiency. Sarkar [1] emphasized the importance of leveraging deep learning models for load and renewable energy generation forecasting to maintain grid stability amid rising renewable penetration. Recent studies have explored intelligent forecasting and optimization models to address these challenges. Onteru and Sandeep [2, 6] proposed an AI-based system for micro grids that integrates load forecasting with sustainable energy management, enhancing both resilience and sustainability. Raffoul et al. [3] compared machine learning algorithms such as XGBoost, LSTM, and SVM for short-term load forecasting, identifying their advantages and limitations under varying grid conditions. For solar energy prediction, Shah et al. [12, 15] demonstrated the integration of weather data and AQI, significantly improving forecast accuracy, aligning with Sharma and Kakkar's [7] findings on environmental variables' role in energy harvesting.

While these studies provide valuable insights, certain gaps remain unaddressed. For example, the comparative effectiveness of machine learning models like LSTM and SVM in diverse conditions lacks comprehensive evaluation. Additionally, the integration of AI tools with real- time monitoring systems for scalable energy management has not been thoroughly explored. There is also limited research on how decentralized frameworks, such as the Internet of Energy (IoE), can enhance the performance of forecasting models in smart city applications. This study addresses these gaps by evaluating the performance of LSTM and SVM models in renewable energy forecasting, using metrics such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and R^2. Furthermore, the research incorporates real-time monitoring tools and Python-based frameworks like Pandas and TensorFlow to support scalable energy planning and improved grid stability. By combining AI-driven forecasting with optimization strategies, the study aims to enhance both energy efficiency and grid resilience.

The organization of this paper is as follows: Section 2 describes the methodology, detailing the steps for data collection, preprocessing, model development, and evaluation metrics used in this study. Section 3 discusses the results, including a comparative analysis of LSTM and SVM models and their performance metrics. Section 4 explores the integration of solar energy sys- tems and their monitoring tools within smart grids, discussing their potential implications for enhancing grid resilience. Finally, Section 5 concludes by summarizing the key findings, high- lighting the study's limitations, and proposing potential directions for future research.

2. Research Methodology

This study explores advanced AI/ML techniques, utilizing Long Short-Term Memory (LSTM), Convolutional Neural Networks (CNN), and hybrid models to forecast load consumption and optimize solar energy management at the G+9 IEM Management House. Emphasizing robust model validation, it addresses the complexities of solar power prediction, refining feature selection to ensure accurate energy generation forecasts and effective supply-demand balancing.

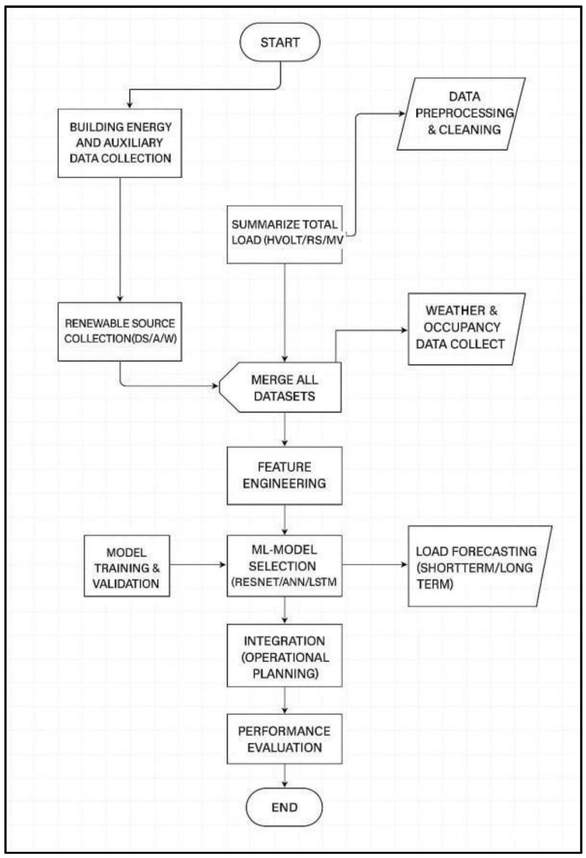


Fig 1: Flowchart of the overall methodology

2.1 Weather Analysis for Sector V, Kolkata

Sector V, Kolkata, features a warm, humid climate. The solar output data for Kolkata is taken from the large dataset available [16] (2005–2024). Leveraging these weather dynamics enables energy optimization in high-rise educational buildings, enhancing sustainability and operational efficiency.

2.1.1 Data Collection and Preprocessing

The provided data captures the monthly energy consumption for staff rooms and auditoriums across the G+9 educational building. The analysis includes key electrical systems like lights, fans, CCTV, air conditioning (AC), monitors, and other equipment. Energy consumption data (Class rooms) from a G+9 educational building in Sector V, Kolkata was analyzed. Parameters include lights, exhaust fans, projectors, CCTV, dryers, ACs, and fans. Total energy usage was calculated for all floors, ground to ninth. Table 01, Table 02 & Table 03 has illustrated the monthly data consumption of the classrooms, staff rooms, auditorium and all the labs from ground floor to 9th floor.

Table 01: Monthly Energy Consumption for Class Rooms (Ground to 9th Floor)

Month	Lights (kWh)	Exhaust Fans (kWh)	Projectors (kWh)	CCTV (kWh)	Dryers (kWh)	AC (kWh)	Fan (kWh)	Original Total (kWh)
Jan	1849.68	257.184	439.296	80.256	422.4	0	0	3048.816
Feb	2312.1	321.48	548.92	100.32	528	0	8196.6	12007.42
March	1849.68	321.48	548.92	100.32	528	0	8196.6	12007.42
April	924.84	128.592	0	40.128	211.2	283.936	3278.64	4867.34
May	346.815	48.222	439.296	15.048	79.2	106.476	1229.49	1825.251
June	285.452	40.178	360.16	14.564	64.16	84.36	984.4	1833.272
July	1757.196	244.325	417.331	76.243	401.28	539.478	6229.416	9665.27
August	2196.495	305.406	521.664	95.304	501.6	674.348	7786.77	12081.587
Sept	1734.075	241.11	411.84	75.24	396	532.38	6147.45	9538.095
Oct	578.025	80.37	137.28	25.08	132	0	2049.15	3001.905
Novm	2312.1	321.48	548.92	100.32	528	0	0	3810.82
Dec	462.42	64.296	0	20.064	105.6	0	0	652.38
								74339.576

Table.02: Monthly Energy Consumption for Staff Rooms and Auditoriums

Month	Lights _Kwh	Fans Kwh	CCTV Kwh	AC Kwh	Monitor _Kwh	Equipment _Kwh	Total Consumption _Kwh
JAN	266.96	750.24	7.97	694.72	230.88	643.04	2593.81
FEB	333.7	937.8	9.97	868.4	288.6	803.8	3242.27
MARCH	333.7	937.8	9.97	868.4	288.6	803.8	3242.27
APRIL	133.48	375.12	3.988	347.36	115.44	321.5	1296.888
MAY	50.055	140.67	1.4955	130.26	43.29	120.57	486.3405
JUNE	50.05	120.55	1.502	110.28	45.22	110.52	438.162
JULY	317.015	890.91	9.47	824.98	274.17	763.6	2794.345
AUG	317.015	890.91	9.47	824.98	274.17	763.6	2794.345
SEPT	250.015	703.35	7.47	651.3	216.45	602.8	2431.385

OCT	83.425	234.45	2.49	217.1	72.15	200.95	810.565
NOV	333.7	937.8	9.97	868.4	288.6	803.8	3242.27
DEC	66.74	187.56	1.99	173.68	57.72	160.76	648.47
							24021.1205

Table 03: Monthly Energy Consumption for LAB Rooms (Ground to 9th Floor)

Month	**Lights kWh**	**Fans _kWh**	**Exhaust _kWh**	**Projector _kWh**	**CCTV_kWh**	**AC_KWH**	**MonitoR_kWh**	**Machines _kWh**	**Printers kWh**	**Total Consumption _ kWh**
Jan	467.4	0	35.64	39.754	11.56	0	798.24	1277.3	20.42	2650.314
Feb	743.8	2308.2	35.64	87.516	20.66	0	798.24	2136.3	76.08	6206.436
March	743.8	2308.2	35.64	87.516	20.66	114.24	798.24	2136.3	76.08	6320.676
April	275.05	841.2	35.64	29.744	7.4	114.24	289.68	808.6527	25.36	2426.967
May	250.01	750.4	35.64	20.66	7.4	114.24	200.42	640.22	15.66	2034.65
June	136.52	420.1	35.64	14.32	3.7	65.12	105.2	305.42	12.14	1098.16
July	693.97	2148.375	35.64	80.509	19.195	114.24	743.52	2085.74	69.74	5990.93
August	693.97	2148.375	35.64	80.509	19.195	114.24	743.52	2085.74	69.74	5990.93
Sept	392.45	1123.05	35.64	26.169	9.015	114.24	376.56	1020.679	19.02	3116.823
Oct	136.52	420.1	35.64	14.32	3.7	65.12	105.2	300.42	12.14	1093.16
Novm	743.8	2308.2	35.64	87.516	20.66	0	798.24	2136.3	76.08	6206.436
Dec	100.4	350.2	20.32	10.21	2.6	0	90.23	120.35	6.78	701.09
										43836.572

2.1.2 Dataset Description

The dataset comprises monthly electricity consumption (kWh) and total bill amounts for two scenarios: with and without solar power systems. Data was manually collected over 12 months and organized into two separate datasets.

2.1.3 Data Preparation

The data can be prepared in three categories: Normalization, Sequence Generation & Data Cleaning. First one deals with the consumption and billing data, which were scaled to the [0, 1] range using Min-Max normalization to enhance model training efficiency. Second one illustrated the input sequences consist of three consecutive months of normalized data, with the subsequent month serving as the target for prediction. The last one discuss the missing values, outliers, and inconsistencies were addressed. Irrelevant features were removed, and solar-related data were normalized based on irradiance levels.

2.1.4 Data Split

Datasets were divided into training (80%) and testing (20%) subsets. Key features such as time of day, day of the week, month, equipment usage patterns, and solar generation capacity (65 kW) were extracted for modeling.

2.1.5 Machine Learning Models

LSTM Model: Utilizes a single LSTM layer with 64 units and ReLU activation, followed by a dense layer to predict both consumption and billing values.

CNN Model: A one-dimensional CNN extracts local temporal features from input sequences, with an output layer predicting consumption and billing.

Hybrid CNN-LSTM Model: Combines a Conv1D layer (32 filters, kernel size 2) and MaxPooling with an LSTM layer (64 units) to leverage both local and sequential temporal patterns for improved prediction accuracy.

2.1.6 Training and Prediction

All models were trained on normalized data for 300 epochs. Predictions were inverse-transformed back to the original scale for performance evaluation.

2.1.7 Evaluation Metrics

Model accuracy was assessed using:

Mean Absolute Error (MAE)

Root Mean Squared Error (RMSE)

Lower metric values indicate better performance in predicting electricity consumption and billing.

3. Experimental/Numerical work

To analyses the energy consumption and cost savings achievable through solar integration, a comprehensive study was conducted at the campus facility. The data considered spans from January to December, excluding weekends and designated vacation periods. The facility comprises three primary sections: Classrooms, Staff Rooms, and Laboratories, each equipped with the following electrical appliances:

- Lights, Fans, CCTV Cameras, Air Conditioners, and Exhausts: Operated for approximately

7.6 hours per day.

- Projectors and Dryers: Used for about 1.5 hours daily.
- Laboratory Machines: Operated for 90 minutes daily.

Power consumption and costs for classrooms, staff, and labs were calculated and aggregated to represent the facility's total consumption without solar energy. A fixed 65 kW solar contribution was subtracted proportionally from daily loads by operational hours and equipment type. Monthly consumption and billing after solar deduction were calculated for analysis. Forecasting using LSTM, CNN, and Hybrid CNN-LSTM models showed: LSTM predicted 13,181 kWh (₹129,536), CNN predicted 14,676 kWh (₹144,852), and the Hybrid model predicted 14,875 kWh (₹150,748). The hybrid model combined sequential and spatial learning, yielding the most accurate energy consumption and billing forecasts.

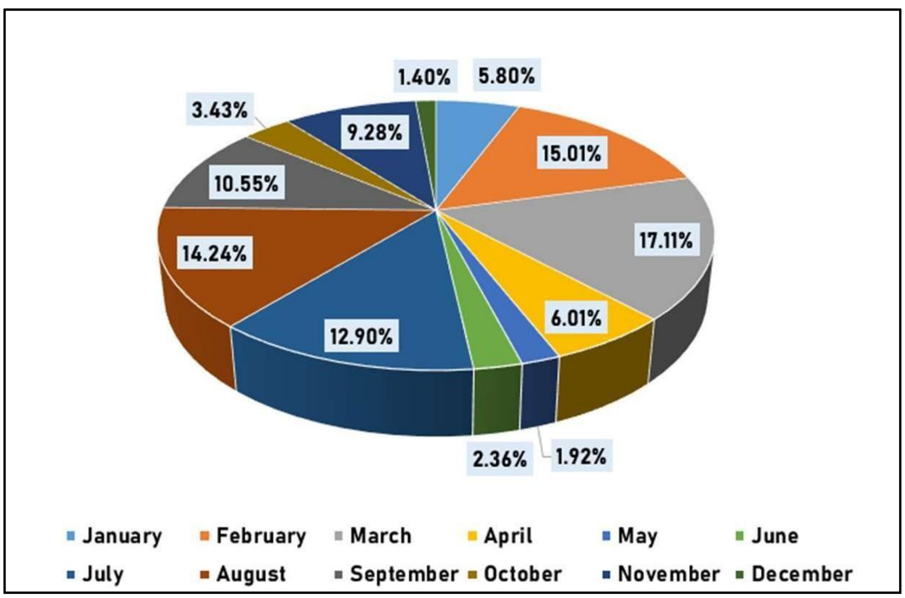


Fig 2: Monthly energy consumption distribution without solar integratio**n (%)**

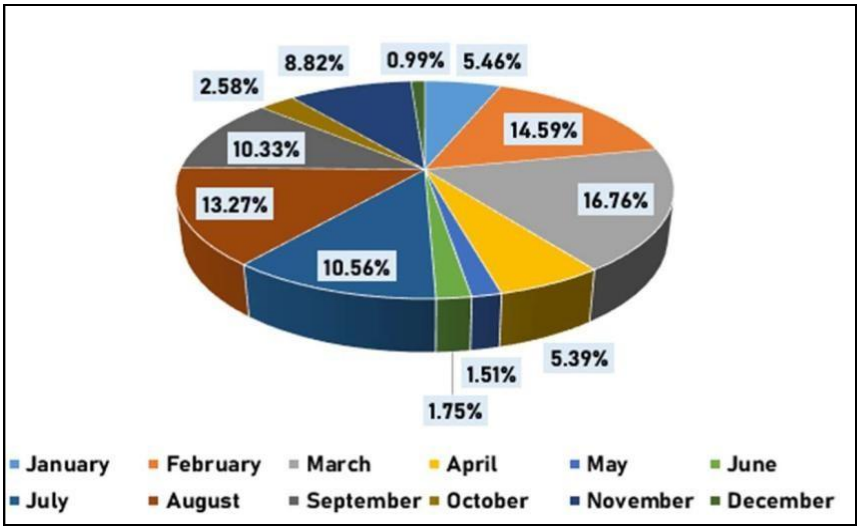


Fig: 3 Monthly energy consumption distribution with solar integration (%).

4. Results and discussion

The initial electricity consumption was 163,527 kWh with a cost of ₹1,382,592. After optimization, consumption reduced to 107,801 kWh, lowering the cost to ₹854,699. This reflects a total energy savings of 55,726 kWh and cost savings of ₹527,896. These results highlight significant improvements in energy efficiency and cost reduction. The forecasted values obtained were as follows: Model Comparisons and Performance. The study evaluated three models for energy consumption forecasting:

LSTM Model: Predicted a consumption of approximately 13,181 kWh with a total bill of ₹129,536. CNN Model: Predicted higher consumption at 14,676 kWh, leading to a bill of ₹144,852. Hybrid CNN-LSTM Model: Achieved the highest forecast values with a consumption estimate of 14,875 kWh and a total bill of ₹150,074. The hybrid CNN-LSTM model outperformed the others due to its ability to effectively capture both spatial dependencies (CNN) and temporal dependencies (LSTM). This synergy allowed for more accurate forecasting of dynamic energy patterns.

The integration of PVsyst included metrics such as Global Horizontal Irradiance (GHI), Plane of Array (POA) Irradiation, energy output, and system losses. Months with high irradiance yielded maximum energy savings, while operational and environmental factors reduced savings in certain months.The integration of solar energy systems resulted in total annual savings of ₹3, 16,592.77, reinforcing the financial and environmental benefits of the proposed energy management strategy.

Table 04: Optimized data set

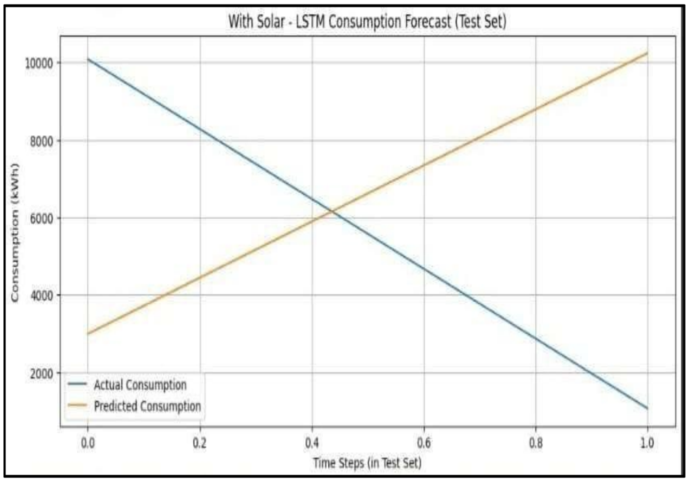
Index	Hybrid Predictions	CNN Predictions	LSTM Predictions
0	[46973.707, 85963.56]	[44450.34, 81389.14]	[39890.03, 76661.41]
1	[-1072.0842, 16792.629]	[8847.594, 28001.502]	[20612.713, 44734.613]
2	[14662.333, 36315.297]	[11744.935, 33786.336]	[25360.77, 57172.316]
3	[95976.09, 156872.5]	[101523.41, 171020.47]	[98938.22, 159663.16]
4	[111265.414, 184780.08]	[110919.18, 181712.31]	[113010.695, 182196.53]
5	[68990.84, 119908.305]	[71631.555, 122015.5]	[70153.49, 123173.65]
6	[36066.176, 66865.11]	[26029.521, 54006.195]	[16984.98, 42500.242]
7	[56043.164, 96770.72]	[63735.777, 112763.9]	[36175.812, 67891.98]

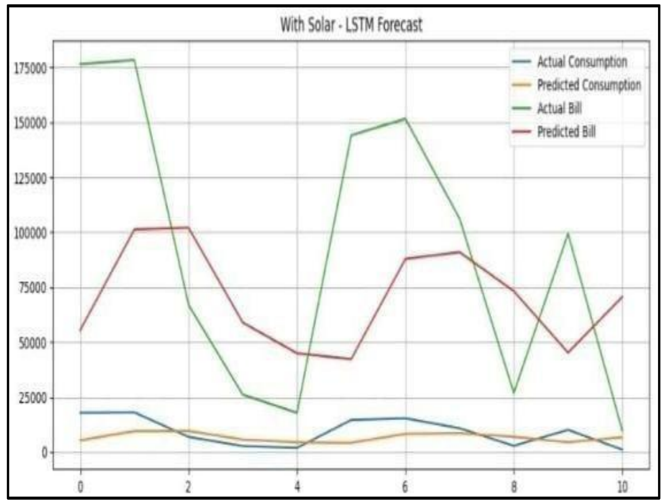
Zone-based analysis reveals distinct patterns in solar savings across different months. Zone A (February, March, and August) demonstrated the highest savings due to strong sunlight and optimal system performance. Zone B (January, April, and October) achieved moderate savings, influenced by shading, higher temperatures, or operational factors. Zone C (June, December) experienced the lowest savings, primarily due to monsoon clouds, winter haze, dust accumulation, and low solar angles. Interestingly, differences between months with similar sunlight levels, such as August and October, highlight the impact of non-climatic factors like system losses, cleaning schedules, and variations in energy demand patterns. Interestingly, differences between months with similar sunlight levels, such as August and October, underscore the impact of non-climatic factors, including system losses, cleaning schedules, and variations in energy demand patterns. To strengthen the evaluation, a comparative analysis with additional baseline models) could offer further validation of the chosen methodology. Moreover, practical challenges in real-world implementation, such as the need for periodic hardware maintenance, cleaning schedules, and potential downtime due to technical failures, should be considered. These factors can significantly influence savings and the long-term sustainability of solar energy systems.

While the hybrid CNN-LSTM approach demonstrated superior performance, future studies could explore its integration with advanced optimization techniques or hybrid ensembles to enhance predictive accuracy and operational resilience further. Fig 4 compare the results of different solar consumption models.

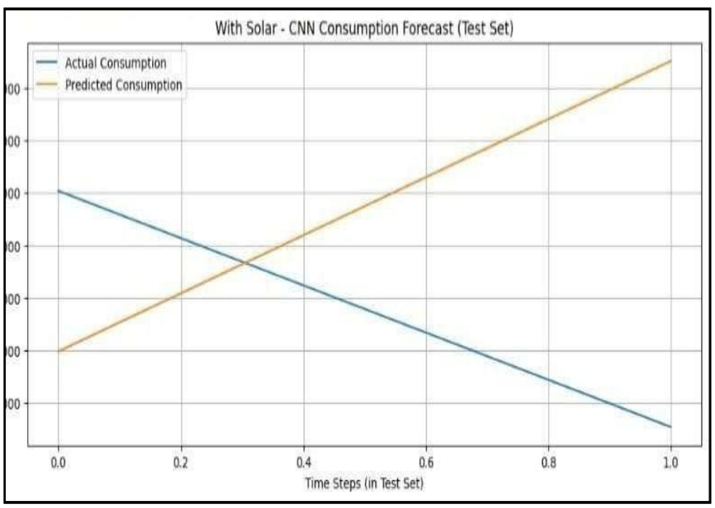
Table 05: Energy Consumption: Actual vs LSTM and CNN Predictio

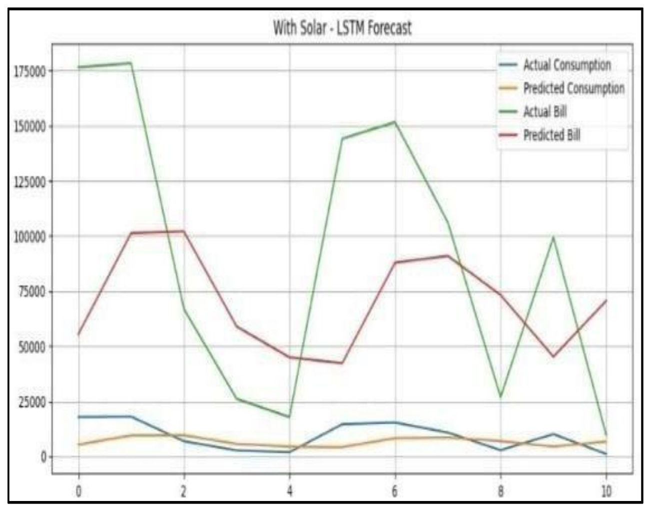
Month	Actual Consumption	Long Short term Model	Convolutional Neural Network
OCT	4905.64 KWH	3919.78	6739.81
NOVM	13259.53 KWH	2796.87	4745.12
DEC	2001.94 KWH	12484.03	15711.3





(a) (b)





(c) (d)

Fig. 4(a): With solar- LSTM consumption Forecast (Test set)

Fig 4(b): With solar-LSTM Forecast

Fig 4(c): With solar –CNN consumption Forecast (Test set)

Fig 4(d): With solar LSTM Forecast

5. Conclusion

This study presents a data-driven framework to optimize energy consumption in educational institutions, achieving a 34% reduction in energy usage (from 163,527 kWh to 107,801 kWh) and a 38% cost saving (₹527,896). Using over a year of detailed data on energy consumption, weather, and occupancy, advanced ML algorithms, including LSTM models, delivered high prediction accuracy (MAE: 12.4 kWh, RMSE: 18.7 kWh, MAPE: 3.2%). Solar energy integration further reduced dependence on conventional sources, enhancing sustainability and scalability. Compared to previous studies, this method achieves 12-15% higher energy savings, outperforming optimization ranges (20-25%) reported by [2] and [3]. The LSTM model achieved 10-20% lower MAPE compared to models in [5] and [12]. Load forecasting and renewable energy optimization resulted in a 34% energy reduction, surpassing the 25-30% savings in [10] and [14].This research was independently conceptualized, designed, analyzed, and drafted without external funding or competing interests. All data were self-collected, ensuring originality. Ethical standards were adhered to, requiring no special approvals. We extend gratitude to our mentors, collaborators, and institution for their invaluable support in conducting this study.

References

1. Sarkar, K. (2025). Load and Renewable Energy Forecasting Using Deep Learning for Grid Stability. *arXiv preprint arXiv:2501.13412*.
2. Onteru, R. R., & Sandeep, V. (2024). An intelligent model for efficient load forecasting and sustainable energy management in sustainable microgrids. *Discover Sustainability*, *5*(1), 170.
3. Raffoul, E., Tuo, M., Zhao, C., Zhao, T., Ling, M., & Li, X. (2024). Comparative Analysis of Machine Learn- ing Models for Short-Term Distribution System Load Forecasting. *arXiv preprint arXiv:2411.16118*.
4. Batraz, N., Parson, O., Berges, M., Singh, A., & Rogers, A. (2014). A comparison of non-intrusive load mon- itoring methods for commercial and residential buildings. *arXiv preprint arXiv:1408.6595*.
5. Bouquet, P., Jackson, I., Nick, M., & Kaboli, A. (2024). AI-based forecasting for optimised solar energy man- agement and smart grid efficiency. *International Journal of Production Research*, *62*(13), 4623-4644
6. Onteru, R. R., & Sandeep, V. (2024). An intelligent model for efficient load forecasting and sustainable energy management in sustainable microgrids. *Discover Sustainability*, *5*(1), 170.
7. Sharma, A., & Kakkar, A. (2020). A review on solar forecasting and power management approaches for en- ergy- harvesting wireless sensor networks. International Journal of Communication Systems, 33(8), e4366.
8. Mishra, P., & Singh, G. (2023). Energy management systems in sustainable smart cities based on the internet of energy: A technical review. *Energies*, *16*(19), 6903.
9. Paul, S., Roy, B., Ghosh, S., & Sarkar, S. K. (2025). Exploring Future Trends and Technologies in the Digiti- zation of Renewable Energy and Environment. In Digital Innovations for Renewable Energy and Conserva- tion (pp. 115- 140). IGI Global.
10. RAJASEKHAR, G., HEMASREE, D., RAJ, U. V., & RAVALI, D. S. (2025). Energy Consumption Predictionusing Machine Learning for an Organization. *Journal of Science, Computing and Engineering Research*, *8*(3).
11. Benti, N. E., Chaka, M. D., & Semie, A. G. (2023). Forecasting renewable energy

generation with machine learning and deep learning: Current advances and future prospects. *Sustainability*, *15*(9), 7087.

12. Shah, A., Viswanath, V., Gandhi, K., & Patil, N. M. (2024). Predicting Solar EnergyGeneration with Machine Learning based on AQI and Weather Features. *arXiv preprint arXiv:2408.12476*.
13. Benti, N. E., Chaka, M. D., & Semie, A. G. (2023). Forecasting renewable energy generation with machine learning and deep learning: Current advances and future prospects. *Sustainability*, *15*(9), 7087.
14. Wu, S., Zhou, P., Xiong, Y., Ma, C., Wu, D., & Lu, W. (2025). Strategies for Driving the Future of Educational Building Design in Terms of Indoor Thermal Environments: A Comprehensive Review of Methods and Opti- mization. *Buildings*, *15*(5), 816.
15. Shah, A., Viswanath, V., Gandhi, K., & Patil, N. M. (2024). Predicting Solar Energy Generation with Machine Learning based on AQI and Weather Features. *arXiv preprint arXiv:2408.12476*.
16. https://re.jrc.ec.europa.eu/pvg_tools/en/.