OPEN ACCESS
International Journal of Innovations in Science & Technology
RESEARCH & INNOVATION DIVISION
IJIST

Deep Recurrent Neural Network-Based Forecasting of Electricity Consumption and Anomalies Detection

Waqas Ahmed[1], Tayyab Ali[2], Bacha Hussain Bukhari[3], Amjad Khattak[1]

[1]Department of Electrical and Power Engineering, University of Engineering & Technology, Peshawar.

[2] Department of Electrical and Power Engineering, Iqra National University, Peshawar.

[3]Department of Mechatronics Engineering, University of Engineering & Technology, Peshawar.

*Correspondence: waqaspir23@gmail.com

Citation | Ahmed. W, Ali. T, Bukhari. B. H, Khattak. A, "Deep Recurrent Neural Network-Based Forecasting of Electricity Consumption and Anomalies Detection", IJIST, Vol. 07 Issue. 03 pp 1856-1870, August 2025

Received | June 26, 2025 **Revised |** August 07, 2025 **Accepted |** August 09, 2025 **Published |** August 10, 2025.

The construction industry is among the greatest fuel consumers of the world and a major source of carbon dioxide. Owing to this, the environmental effect can be minimized when focusing on conserving energy in buildings. Misuse of energy in the effort of equipment and errors of humans in the work is never considered budget. The use of smart buildings will address this problem since it will track the use of energy, detect abnormal behavior, and remind the managers that they are supposed to take energy-conservation actions. The current paper considers the issue of anomaly detection in the hourly electricity consumption level on a real basis and gives a two-step process with a Long Short-Term Memory (LSTM) network. In the first step, there will be forecasting of energy consumption, and, following this, the anomalies will be identified with the assistance of an LSTM Autoencoder. The article draws comparisons between highly complex time-dependent feature extraction algorithms like Rough Autoencoder (RAE), Deep Temporal Dictionary Learning (DTDL). The other algorithms could not perform better than the proposed method, the range of R-squared value was 95.11, MAE was 38.5, the MSE was 2448.94, and the RMSE was 49.49. Besides, the paper evaluates the means through which the AI-based anomaly detection solutions can provide forecasts of the electricity consumption, and the LSTM networks and autoencoders were tested to be more appropriate in forecasting the electricity consumption than the other deep learning algorithms.

Keywords: Forecasting, Energy Efficiency, Predictive Modeling, Building Energy Monitoring, Energy-saving Strategies, Anomaly Identification

Introduction:

Buildings lose considerable energy because of equipment faults and human errors, making energy-efficient strategies essential to reduce both consumption and carbon emissions. Building Energy Monitoring Systems (BEMS) with smart building technologies are used to identify abnormal energy usage and guide savings measures. The research proposes an AI-based model for electric power consumption prediction and anomaly detection.

Anomaly detection identifies behaviors that deviate from the intended or expected patterns and can be applied in domains such as fraud detection, healthcare, and cybersecurity. It can utilize both labeled and unlabeled data, labeled datasets provide information to define and recognize anomalies, while unlabeled datasets contain data whose anomaly status is unknown. Such anomalies are of point kind, contextual and collective. Point anomaly is where anomalous single data point occurs from the rest, such as an atypical monthly power cost, as described by author [1].

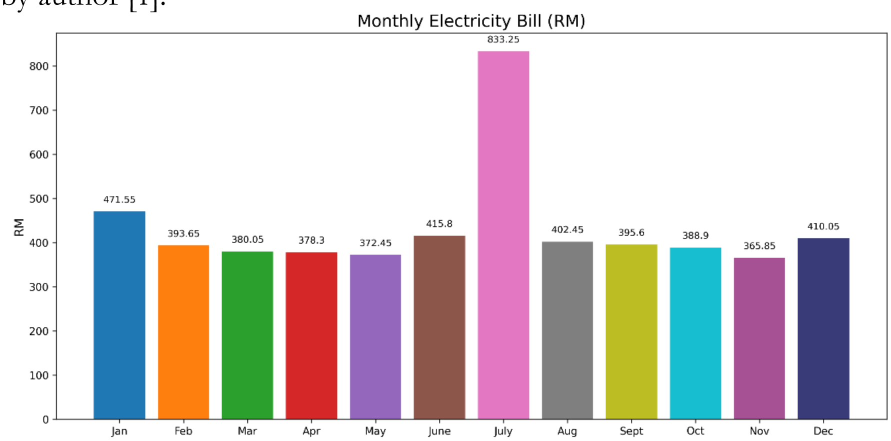


Figure 1. Point analyst in a monthly electricity bill (author's own illustration)

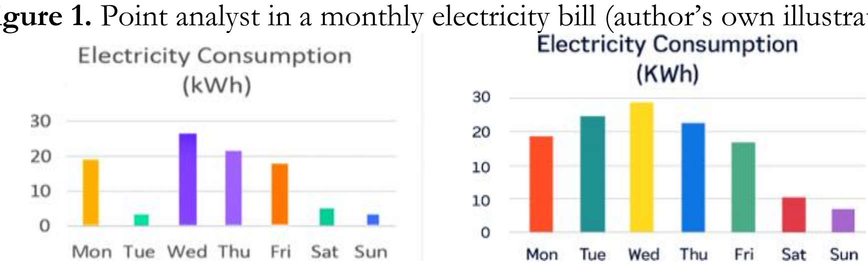


Figure 2. Daily Electricity Consumption (a) Low consumption during weekends is normal; (b) Low consumption during normal weekdays is anomalous.

Contextual anomalies consider the context in which the data occurs. For example, low power usage on a weekday may be considered abnormal, whereas the same pattern would be normal on weekends or public holidays. Collective anomalies occur when a group of data points shows abnormal behavior, even if individual points seem normal. This study focuses on detecting anomalies in time series data by incorporating contextual and collective components alongside periodic patterns.

Anomaly detection is widely studied across applications like fraud detection, healthcare, and aviation. Conventional methods often overlook the sequential nature of data, which limits their effectiveness in detecting anomalies within time series datasets. Sequences, which can be symbolic or time-based, present unique challenges because anomalies can appear as entire sequences or subsequences. Traditional anomaly detection methods struggle to handle time series data effectively, particularly in applications such as building energy consumption analysis.

This article makes two main contributions:

It introduces an LSTM-based prediction model for forecasting future energy consumption sequences.

It uses an LSTM autoencoder for anomaly detection in building energy data, addressing challenges with long time series. The article evaluates model performance across different prediction lengths.

Literature Review:

Anomaly Detection in Building Energy Consumption:

Historical building energy data are typically used to identify point anomalies. In contrast, the proposed study focuses on both anomaly detection and energy consumption prediction. A blend of neural network ARIMA model was utilised to forecast the consumption in [2], and checking the deviations was carried out through the two-sigma rule to detect unusual occurrences. In [3], the authors proposed an unsupervised prediction-based point anomaly detection algorithm on which they have conducted a learning scheme where every reading has been assigned an anomaly score. The author also introduced graphical apparatus: time series and measure of anomalies.

Power consumption variations have been used to identify irregularities in recurring building activities. In [4], simultaneous anomalies were detected by analyzing power usage both in the frequency domain and over a moving one-week time window. The frequency-domain and moving one-week time window approach proposed in [4] incorporated the assumption that data were continuously available, which often led to false positive findings. In one study, one-step-ahead forecast was suggested to indicate abnormalities [5], however, day of the week and holiday factor that could be a potential false positive reducing influence was not considered.

In the study by [5], contextual information was also considered, acknowledging that certain values may be normal in one context but truly anomalous in another. It stimulated the amount of anomaly detection to consider relative inconsistencies when it comes to collective sequential consumption. In other studies [6], multivariate methods are used to construct models to explain aspects like climate and building structure, with the expectation of helping managers make energy efficiency decisions. An article [7] addresses how the energy consumed and supplied is influenced by external variables and explains how it relies on an ARMA model to state the boundaries of energy usage and find anomalies. A three-part system [8] identifies anomalies in real-time data while continuously adapting to emerging context-specific situations.

Anomaly Detection using Artificial Neural Network:

In addition to the methods described, the paper also explains how the researchers utilize artificial neural network strategies. Another study on anomaly detection using machine learning (RML) developed a system that leverages time series data from environmental sensors to achieve real-time anomaly detection [5]. Their proposed approach relies on an autoregressive data model and its corresponding future prediction horizon. Different prediction models employed by the researcher to come up with a one step ahead of a predication comprises of the naive predictor, the nearest cluster (NC), multi-layer perceptron (MLP) and the single-layer linear network. A value is considered normal if it falls within the prediction range, which is determined using the standard deviation of the residuals from the applied model; otherwise, it is classified as anomalous. The paper, [9] proposes the potentially useful approach anomaly detection in the data on power consumption, which is a combination of the ARIMA and Recurrent Neural Network (RNN) architecture more suitable to the data area. Each technique employs a predictive model that is defined by calculating the difference between the predicted and actual consumption levels. The two models are used on the same timesteps so that their residuals fall close together to identify a probable anomaly. The author believes that the integration of both prediction models of RNN with the ARIMA leads to increased anomaly detection behavior as compared to independent anomaly detection behavior.

Anomaly Detection with LSTM:

The specific use of the stacked LSTMs networks in this research is to detect abnormalities in the ECG context through formulating an error vector, which indicates the deviation between the predicted and actual values. The error vectors are assumed to be distributed as a Gaussian to determine a cut-off point between normal and abnormal behavior. The presence of LTM models has been experienced in the detection of abnormalities present in the ECG readings. The same method was also applied to other time series datasets [10], with power usage data achieving a precision of 0.94 and a recall of 0.17. In intrusion detection, the problem described is brought down in [11], where they introduce a solution to it using an LSTM RNN approach to detect point anomalies and collective irregularities. A circular buffer maintains a record of prediction errors and may be used to label subsequences as collective anomalies, with an 86 percent true positive rate.

Objectives:

This study pursues the following objectives:

To forecast hourly electricity consumption using an LSTM-based regression network.

To design an anomaly detection model based on an LSTM autoencoder capable of identifying both contextual and collective anomalies in building energy data.

To compare the proposed framework with existing statistical, machine learning, and deep learning approaches (ARIMA, RNN, RAE, and DTDL) to validate its effectiveness.

Novelty Statement:

The novelty of this research lies in combining time series forecasting and anomaly detection in a single recurrent deep learning framework. Unlike prior studies that focused only on prediction [2] or anomaly detection [3][4], our approach addresses both tasks simultaneously using real hospital data augmented with synthetic anomalies. Furthermore, the evaluation across multiple prediction horizons provides fresh insights into how input–output sequence lengths affect anomaly detection accuracy.

Methodology:

Overview:

This study aims to develop an anomaly detection model using an LSTM-based autoencoder, followed by the creation of a power consumption forecasting model employing the LSTM algorithm from Deep Learning. This section details the comprehensive methodological procedure.

Dataset:

This research is based on one year of data collected hourly on electricity usage from a hospital located in Phoenix, USA. OpenEI provides the dataset, which consists of 8,760 data points representing hourly power consumption in kilowatt-hours (kWh). The data was either used for forecasting or augmented with synthetic anomalies to test the effectiveness of anomaly detection, making the dataset a combination of real and artificial data.

Data Preparation:

Even though there could be practical anomalies in the data, they were considered too small and unimportant to be applied to the forecasting. It was assumed that the original data contains no noise, and any noise present results solely from the insertion of synthetic anomalies. No labeled datasets were available to work on anomaly detection thereby synthetic, and application specific datasets have been utilized, and it is done to emulate the conditions in the real world. The system was tested with these artificial datasets because it came with both event-based and periodic anomalies, providing an opportunity to test the performance of the algorithms within various contextual conditions.

Event-Based Dataset (abnormal peak):

The event-based anomalies dataset was created to test the algorithms' ability to detect sudden changes in consumption, such as spikes or drops. These anomalies included a peak consumption period on a weekday (hours 7500-7515) and surge peaks (hours 7900 and 8100), which could indicate faulty appliances or extra machines running.

Periodical Dataset (off-periodic pattern):

Because the original dataset follows a regular periodic pattern in power consumption, a separate dataset of periodic anomalies was generated to identify irregularities in the usage curve. The aim of this dataset with off-periodic anomalies was to study the algorithm's abilities to model implicit periodicity and to analyze the behaviors of the modeling when periods in such periodical data were missing. Such anomalies can occur in real life due to malfunctioning or faulty appliances, or because of an unexpected additional load.

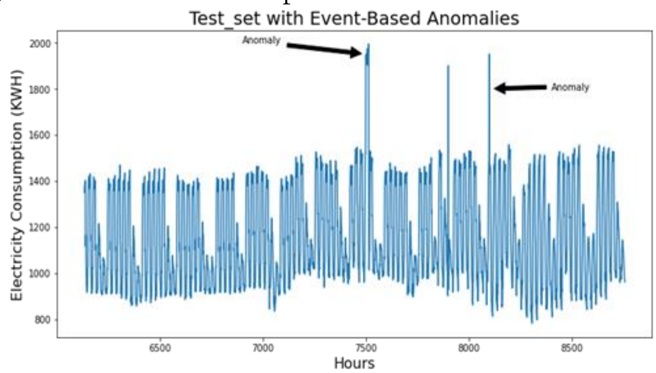


Figure 3. Synthesized Event-Based Anomalies in Test Dataset.

The test set of the original dataset with an off-periodic usage pattern was created by replacing the original power consumption hourly records with random values ranging from 900 to 2000, taken on 7424th to 7600th hours. The periodic pattern was also disturbed on 8265th to 8600th hours with a step of 5 hours by adding random values ranging from 500 to 2000 as the power consumption. Figure 4 shows the synthesized off-periodic pattern test data of the original dataset. As discussed earlier, the exploration analysis of our dataset indicates periodic seasonality of daily and weekly patterns in the original dataset. The creation of an off-periodic pattern dataset was an attempt to emulate the scenario of having consumption out of the daily, weekday, and weekend periodic patterns.

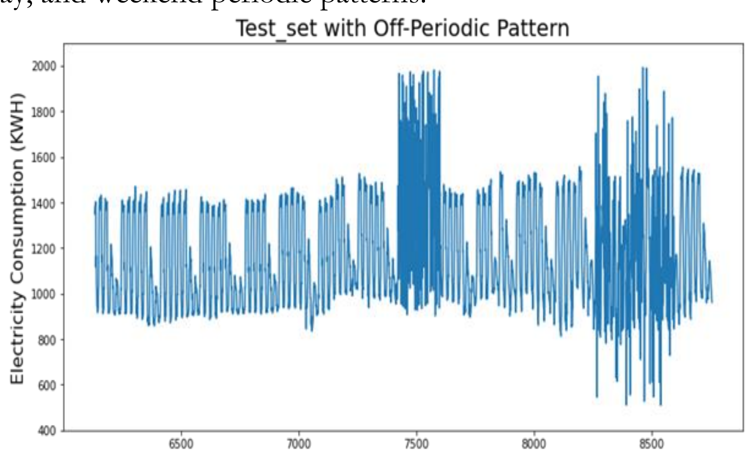


Figure 4. Synthesized Anomalies Based on Off-Periodic Pattern in Test Dataset.

LSTM Network for Predictions:

The objective of the study was to forecast electricity consumption using an LSTM regression network. The block diagram of the suggested arrangement of this component of the system is presented in Figure 5, which reveals both input and output components, the pre-processing of the data, an LSTM network, and the model optimization component. The expected power consumption sequence data are the sequence data of expected electricity consumption, giving the output data of the input factory data.

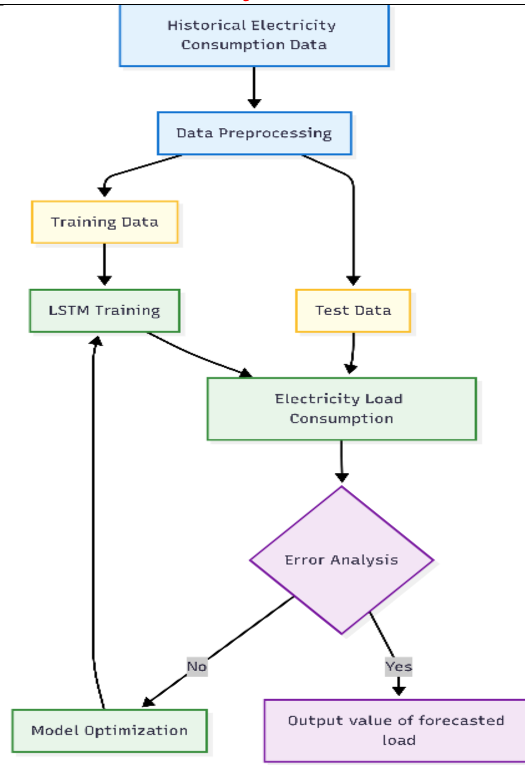


Figure 5. Proposed Framework for the Forecasting Model.

The LSTM algorithm is designed to address long-term dependency issues, allowing it to handle time series data with significant delays, unlike other machine learning algorithms that suffer from vanishing gradients. LSTM networks are built to retain information over extended periods [12]. The computational complexity of a deep neural network (DNN) is determined by the sum of its layers, with LSTM complexity defined as $W = (4IH + 4H^2 + 3H + HO)$, where I is the input count, O is the output count, and H is the number of neurons in the hidden layer. The computation grows linearly with input size, but LSTM processing is more efficient compared to other models [13][14]. The LSTM regression model was implemented using TensorFlow Keras, a high-level API in Python.

$$Y_t = f(Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}) \qquad (1)$$

The LSTM model predicts current energy consumption using historical data. The time series was converted into X and Y matrices and then reshaped into 3D tensors with 60-time steps of electricity consumption, structured as (samples, time steps, input features). To enhance the fitting of the model and evade the issue of training, the data was standardized to ensure that the standard deviation was optimum, i.e., one, and the mean was also optimum, i.e., zero, on a scale of 0 1.

Data Splits:

To train the forecasting model itself, the original dataset was used according to the information that did not contain any anomalies. The dataset possesses 8760 data points with respect to the time series annual data of power consumption per hour. The construction of training and test sets was completed through the use of the dataset. The divisiveness is achieved under different circumstances. The first trial used a nine-tenth data set of the data to train and then evaluated it using a ten-tenth data set test set, followed by a 70- 30 and 50- 50 combination.

LSTM Network Anomaly Detection:

The second objective of the study was to develop an anomaly detection model capable of cleaning the collected energy consumption data. Figure 6 represents the flowchart of the suggested system of anomaly detection.

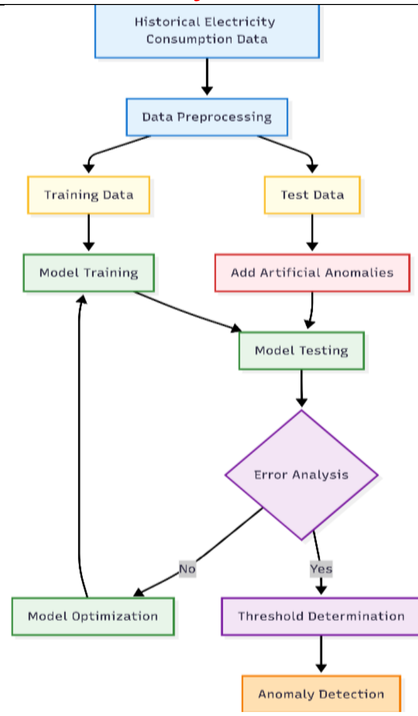


Figure 6. Flowchart of the Proposed Anomaly Detection Model

To detect event-based and off-periodic pattern anomalies, we used an LSTM-based autoencoder model. The model compresses input data, reconstructs it, and calculates the reconstruction error to identify anomalies. The encoder creates a compressed representation, while the decoder reconstructs it over time steps. We trained the model using standard data and fine-tuned hyperparameters to minimize data loss. An anomaly detection threshold was set by analyzing the reconstruction loss distribution, with records exceeding this threshold flagged as anomalies.

Experiments and Results:

Overview:

This section presents the experiments conducted and their corresponding results. The training of the LSTM forecasting algorithms is included, and the experimental findings for anomaly identification are covered. The analysis of some state-of-the-art temporal time series feature extraction models on the given dataset and the applicability of some recent AI-based anomaly detection methods are discussed, respectively.

LSTM for Forecasting:

Model Training:

The LSTM model used for forecasting power consumption was trained under three different setups: first, using 90% of the data for training and 10% for testing; second, a 70-30 split; and third, a 50-50 split. This train set was also divided into a training set and a validation set, which was relatively good, and it serves as an overfitting prevention strategy. In the architectural configuration, the LSTM model undergoes training, testing, and validation on the dataset. The networks were then corrected through parameter tuning, which includes the batch size, number of hidden layers, hidden units, activation function, and so on. The optimum architecture was then chosen after testing multiple distinct architectural versions. Training options that have been specified in the best-performing algorithm are tabulated in Table 1. The model was built with five hidden layers, and the layers were specified to have 150 hidden units. The solver was set to Adam optimizer, and MAE was used as a loss metric and trained for 5 epochs. The Initial learning rate was set by the adaptive method. The training progress is plotted as shown in Figure 7 (a), (b), and (c) with 90%-10%, 70%-30%, and 50%-50%

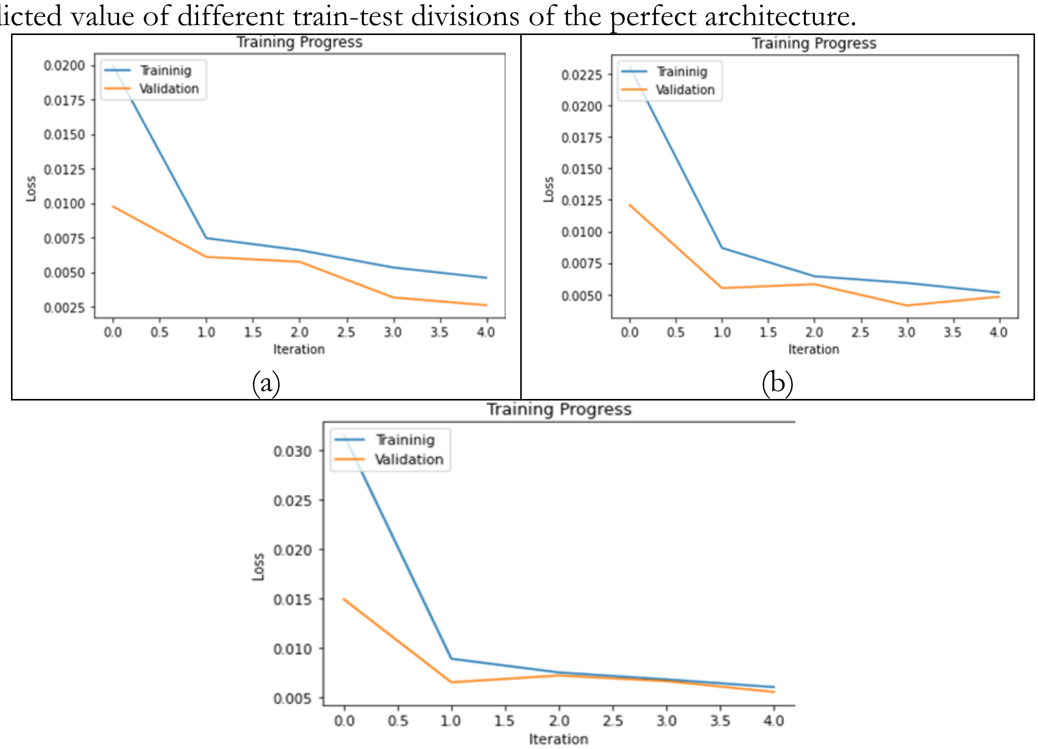
combinations, respectively. The proposed model made use of the Keras framework and ran inside Google Collaboratory, a research instrument developed by Google for deep learning analysis and network development. Google Collaboratory provided a significant advantage by offering free access to GPUs. In this setup, an NVIDIA TESLA T4 GPU with 16GB of memory was used. Google Collaboratory provides a significant advantage by offering most AI engineers free access to GPUs.

Table 1. LSTM Layers and Options Specification

Hyper-parameter	Hyper-parameters setting
Hidden layers	5
HIDDEN Units	150
ACTIVATION FUNCTION	ReLU
TRAINING OPTIMIZER	Adam
EPOCHs	5
Batch Size	5
Learning RATE	Adaptive
Validation split	0.33

Forecasting Future Power Consumption:

After coming up with our final working model, future values of power consumption were forecasted on the basis of the future values of time steps using the predict () dense class on the Keras model. The LSTM network generates predictions sequentially, producing one step at a time. The first prediction is based on the last observation from the training data, and subsequent predictions are made iteratively, one step at a time. The former values of prediction were then used as the initial values in the subsequent processes of prediction. Figure 8 (a), (b), and (c) show some of the algorithm variations of the LSTM network before achieving the very optimized parameters. Figure 9 (a), (b), and (c) depict the training data of the plots with the predicted value of different train-test divisions of the perfect architecture.



(c)

Figure 7. Training progress on yearly training data taken at: (**a**) 90%-10% training-test data division; (**b**) 70%-30% training-test data division; (**c**) 50%-50% training-test data division.

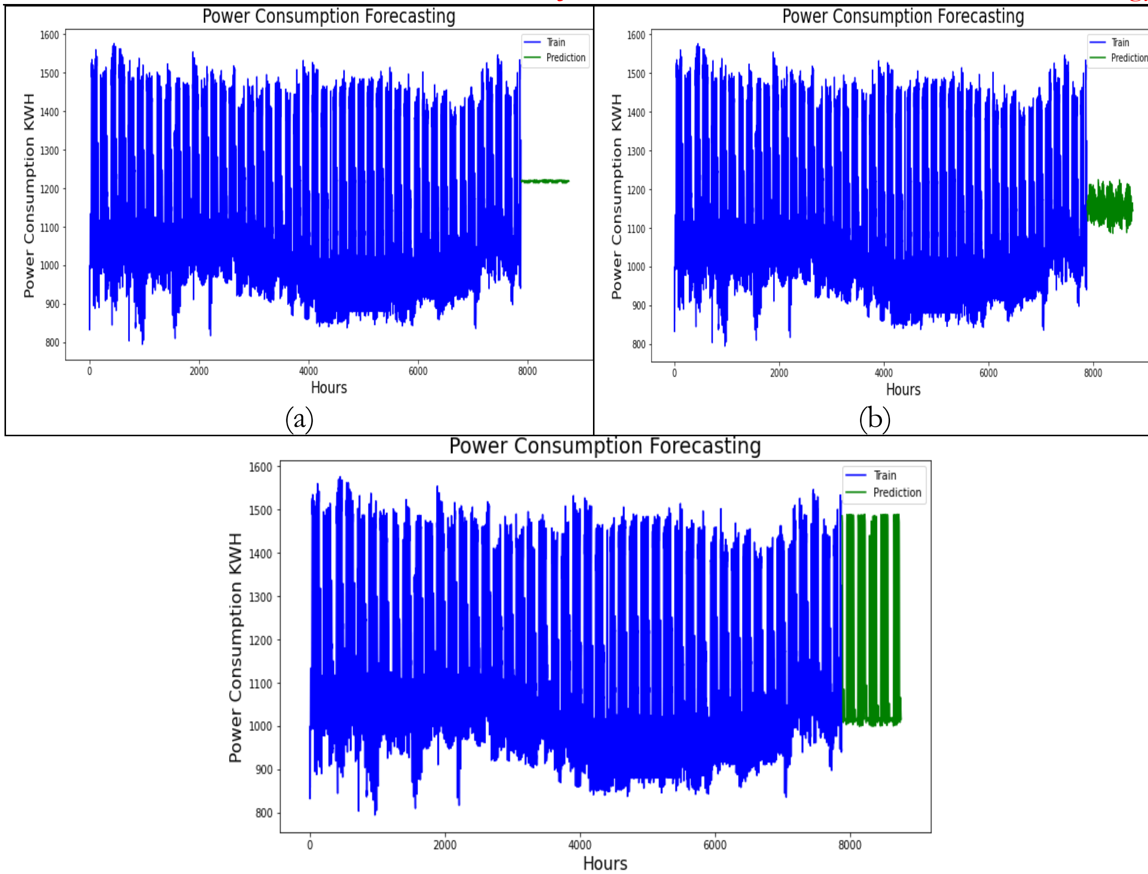


Figure 8. Training progress on yearly training data taken in: (**a**) Step-1 setting; (**b**) Step-2 setting; (**c**) Step-3 setting

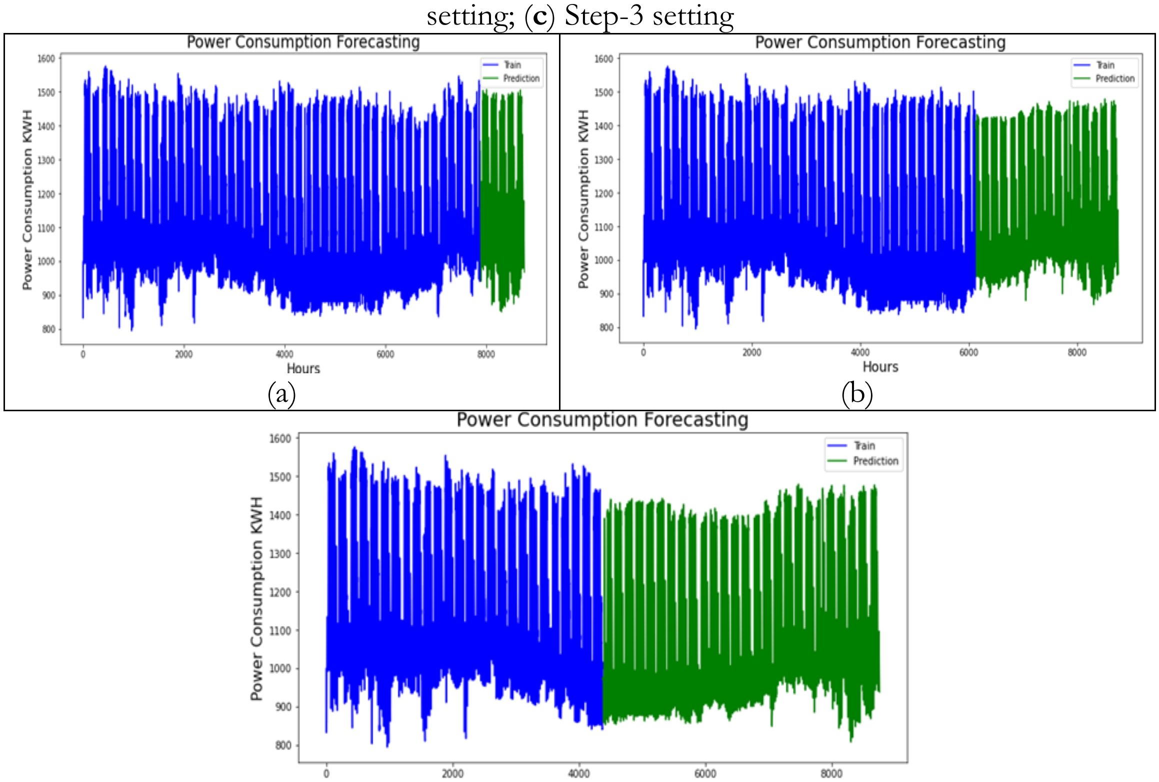


Figure 9. Training progress on yearly training data of the optimum network taken at: (**a**) 90%-10% training-test data division; (**b**) 70%-30% division; (**c**) 50%-50% division.

Comparison of Test Data with Forecasted Values:

The results of the optimal network forecast were subsequently plotted against the test data, as shown in Figures 10(a), (b), and (c).

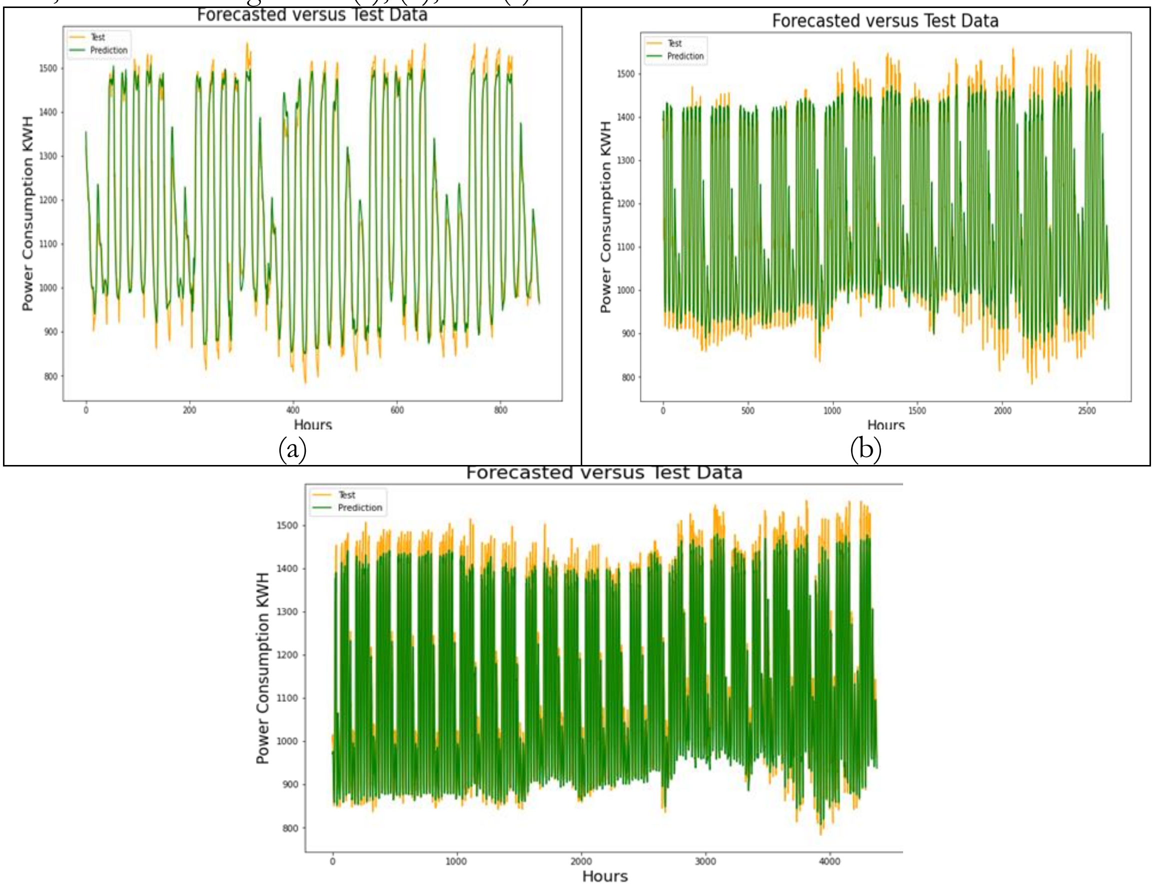


Figure 10. Test data and forecasted values of the optimum network at: (**a**) 90%-10% training-test data division; (**b**) 70%-30% division; (**c**) 50%-50% division

Performance Evaluation:

The model evaluation has shown a remarkable performance as reported in Table 2. The accuracy achieved in each setup, reflected by higher R^2 values and lower errors, is quite satisfactory, indicating that the LSTM forecasting model performs well on both the training and test datasets.

Table 2. LSTM for forecasting Evaluation

Metrics	90%-10% Division	70%-30% Division	50%-50% Division
R-2 (%)	95.11	94.31	92.49
MAE	38.45	36.65	41.57
MSE	2448.94	2514.71	3582.23
RMSE	49.49	50.15	59.85

LSTM Autoencoder for Anomaly Detection:

Model Training:

For anomaly detection, we trained an LSTM autoencoder on the electricity consumption readings in the training set, as 70% of the dataset. We assumed that there were no anomalies and that readings were normal. We specified the LSTM autoencoder model where the input sequences were of one-time step, one feature, and the output sequences were of the same time some feature. Figure 11 shows the training progress, and Table 3 enlists the specified training options of the algorithm.

Table 3. LSTM Layers and Options Specification

Hyper-parameter	Hyper-parameters setting
Hidden layers	5
HIDDEN Units	16
ACTIVATION FUNCTION	ReLU
TRAINING OPTIMIZER	Adam
EPOCHs	5
Batch Size	5
Learning RATE	Adaptive
Validation split	0.33

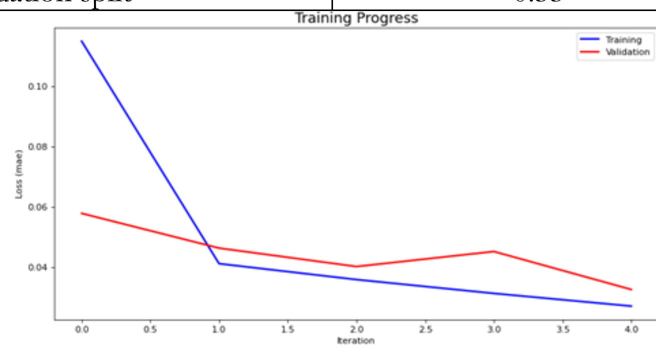


Figure 11. LSTM Autoencoder Model Training Progress.

Anomaly Detection:

We set an anomaly threshold based on the highest value in the training set's computed loss distribution. If the reconstruction error in the test set exceeded this threshold, the data point was flagged as an anomaly. Figures 12 and 13 show the test loss and reconstruction error threshold for event-based and periodical anomalies, respectively.

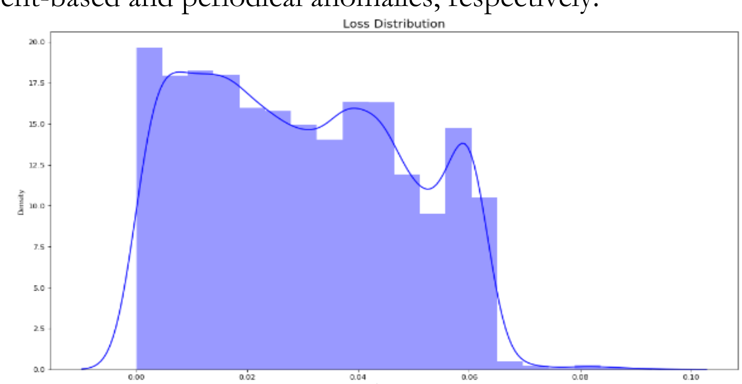


Figure 12. Training Loss Distribution.

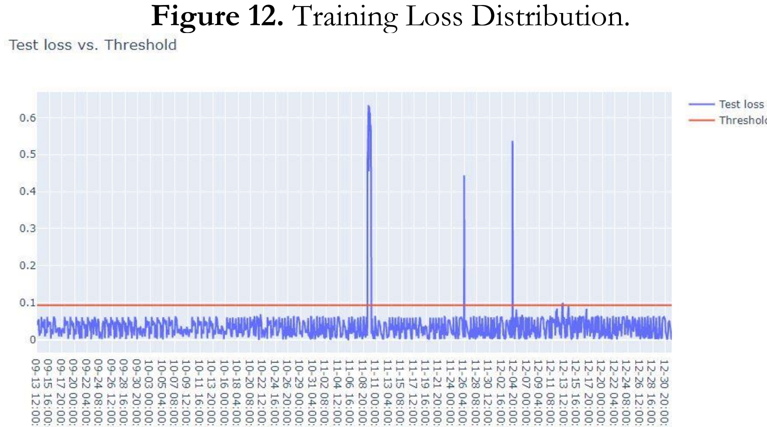


Figure 13. Test loss versus threshold in event-based anomaly

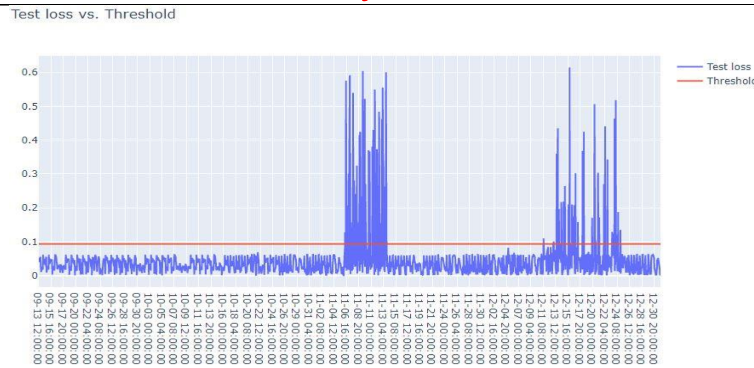


Figure 14. Test loss versus threshold in the periodic anomaly.

Anomaly Detection:

The anomalies detected in event-based and periodical data were then plotted as shown in Figures 15 and 16.

Figure 15. Anomaly Detection in Event-Based Dataset.

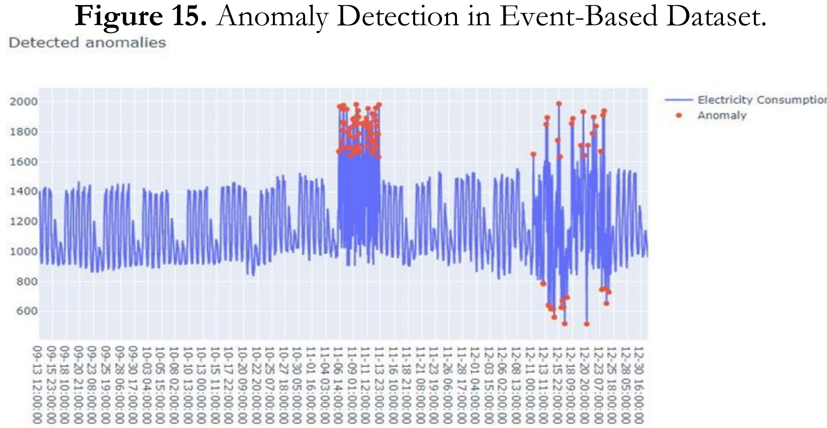


Figure 16. Anomaly Detection in Periodical Dataset.

Comparison with temporal time series feature extraction techniques:

In this study, we used advanced time-dependent feature extraction methods to improve forecasting accuracy. The RAE approach combines autoencoders with rough set theory to extract discriminative features from the electricity consumption data. The model learns a compressed representation of the data, which is then decoded to reconstruct the original input, with the LSTM achieving high accuracy on these encoded features. We also employed DTDL, which iteratively improves dictionary representation to capture temporal patterns in the data. The DTDL method enhances forecasting by merging dictionary learning with deep learning techniques.

Comparative results indicate that although both RAE and DTDL perform effectively, the proposed LSTM model surpasses them, achieving higher forecasting accuracy (R^2) and lower error rates (MAE), as presented in Table 4.

Table 4. Comparative analysis of temporal time series features extraction methods

Metrics	RAE	DTDL	Proposed Model
R-2 (%)	84.84	86.25	95.11
MAE	65.04	64.37	38.45
MSE	7586.5	6877.95	2448.94
RMSE	87.1	82.93	49.49

Applicability of Some Recent AI-based Anomaly Detection Methods:

The sort and amount of consumption anomalies that may be recognized are just as important to the usefulness and strength of AI-based anomaly detection as the accuracy of identifying abnormal consumption of energy. Based on their ability to handle and analyze hybrid data, deep learning-based detection systems for anomalies have significant outcomes in terms of the accuracy of identifying abnormal consumption [15]. The recently published AI-based anomaly detection methods are listed in Table 5 along with their advantages and limitations.

Table 5. Recently published AI-based anomaly detection methods

Ref.	Method	Strengths	Limitations
[16]	One-class Random Forest	Annotated data are not required	Low accuracy and only excessive consumption are detected
[17]	Variational recurrent autoencoder	Annotated data are not required	Assessment of performance is challenging
[18]	Gradient boosting machine	Power utilizes forecasting and anomaly detection	Only detecting suspicious consumption rates and having poor interpretability.
[19]	Deep autoencoder (DAE)	High accuracy, Power utilization forecasting, and anomaly detection	High cost of computing, only excessive usage is detected

Discussion:

The experimental results demonstrated that the proposed LSTM-based forecasting and anomaly detection framework achieved strong performance, with R^2 values exceeding 92% across all data splits (Table 2). The MAE remained consistently low, between 36.65 and 41.57, and the RMSE ranged from 49.49 to 59.85, highlighting the robustness of the model in predicting hourly electricity consumption. The anomaly detection component using an LSTM autoencoder also proved effective, identifying both event-based and off-periodic anomalies with minimal false positives.

When compared with existing studies, the superiority of the proposed approach becomes evident. For instance, [2] applied ARIMA, GRU, and hybrid ARIMA-LSTM approaches for peak energy prediction, reporting higher RMSE values (~82) than those achieved in this study. Similarly, the hybrid ARIMA-RNN approach proposed by [20] produced an MAE of 64.37, whereas the LSTM framework in this research reduced the MAE to 38.45 (Table 4). This improvement of nearly 40% underscores the advantage of sequence-to-sequence deep learning in capturing long-term temporal dependencies.

Moreover, advanced feature extraction techniques such as RAE and DTDL, while effective in extracting discriminative features from time-series data, still fell short compared to the proposed model. The RAE method achieved an R^2 of 84.84, while DTDL reached 86.25, both considerably lower than the 95.11% obtained in this work (Table 4). These comparisons suggest that direct application of the LSTM architecture can outperform complex hybrid frameworks, offering a balance between computational efficiency and accuracy.

The findings also align with recent AI-based anomaly detection studies (Table 5), where deep learning approaches such as DAE [19] achieved high accuracy but at a high computational cost. In contrast, the LSTM autoencoder employed in this research offered high precision in detecting abnormal consumption patterns while maintaining lower complexity. This demonstrates the practical potential of the model for real-world deployment in Building Energy Monitoring Systems (BEMS).

Overall, the discussion indicates that while traditional statistical and hybrid methods provide valuable insights, deep recurrent neural networks such as LSTM offer superior accuracy and generalization for both forecasting and anomaly detection tasks in energy consumption data.

Conclusion:

This research focused on predicting energy usage and detecting abnormalities in electricity data using LSTM-based models. We took prediction based on the LSTM method and anomaly detection based on the LSTM Autoencoder to real hospital power consumption data. The actual experiment scores showed near projections and low MSE, and strong anomaly detecting skills. Further optimization of models towards their performance hyperparameters was done to achieve better performance of the models.

One of the problems of anomaly identification is the fact that raw world datasets do not exist that include anomalies. The solution is tied to the manual addition of synthetic anomalies to the normal data, which may be a weakness in the real-world comparison of anomalies. The study also compares state-of-the-art feature extraction models like RAE and DTDL, and AI-based anomaly detection approaches. Our proposed model outperformed DTDL and RAE, with a 1.10x and 1.12x improvement in R^2, respectively. The proposed model also achieved a significantly lower MAE (38.45), MSE (2.80x and 3x lower), and RMSE (49.49) compared to DTDL and RAE.

References:

[1] V. K. Varun Chandola,Arindam Banerjee, "Anomaly detection: A survey," *ACM Comput. Surv.*, vol. 41, no. 3, pp. 1–58, 2009, doi: https://doi.org/10.1126/science.1235338.

[2] P. P. B Peak, "Peak Electrical Energy Consumption Prediction by ARIMA, LSTM, GRU, ARIMA-LSTM and ARIMA-GRU Approaches," *Energies*, vol. 16, no. 12, p. 4739, 2023, doi: https://doi.org/10.3390/en16124739.

[3] H. Janetzko, F. Stoffel, S. Mittelstädt, and Daniel A. Keim, "Anomaly detection for visual analytics of power consumption data," *Comput. Graph.*, vol. 38, pp. 27–37, 2014, doi: https://doi.org/10.1016/j.cag.2013.10.006.

[4] M. Wrinch, T. H. M. El-Fouly, and S. Wong, "Anomaly detection of building systems using energy demand frequency domain analysis," *IEEE Power Energy Soc. Gen. Meet.*, 2012, doi: 10.1109/PESGM.2012.6344790.

[5] B. S. M. David J. Hill, "Anomaly detection in streaming environmental sensor data: A data-driven modeling approach," *Environ. Model. Softw.*, vol. 25, no. 9, pp. 1014–1022, 2010, doi: https://doi.org/10.1016/j.envsoft.2009.08.010.

[6] O. D.-P. Angel L. Zorita, Miguel A. Fernández-Temprano, Luis-Angel García-Escudero, "A statistical modeling approach to detect anomalies in energetic efficiency of buildings," *Energy Build.*, vol. 110, pp. 377–386, 2016, doi: https://doi.org/10.1016/j.enbuild.2015.11.005.

[7] J. Ploennigs, B. Chen, A. Schumann, and N. Brady, "Exploiting generalized additive models for diagnosing abnormal energy use in buildings," *BuildSys 2013 - Proc. 5th ACM Work. Embed. Syst. Energy-Efficient Build.*, Nov. 2013, doi: 10.1145/2528282.2528291.

[8] and T. L. Y. Jiang, C. Zeng, J. Xu, "Real time contextual collective anomaly detection over multiple data streams," *Work. Outlier Detect. Descr. under Data Divers*, 2014.

[9] K. Hollingsworth *et al.*, "Energy Anomaly Detection with Forecasting and Deep

Learning," *Proc. - 2018 IEEE Int. Conf. Big Data, Big Data 2018*, pp. 4921–4925, Jul. 2018, doi: 10.1109/BIGDATA.2018.8621948.
[10] P. Malhotra, L. Vig, G. M. Shroff, and P. Agarwal, "Long Short Term Memory Networks for Anomaly Detection in Time Series," *Eur. Symp. Artif. Neural Networks*, 2015.
[11] N.-A. L.-K. Loic Bontemps, Van Loi Cao, James McDermott, "Collective Anomaly Detection based on Long Short Term Memory Recurrent Neural Network," *arXiv:1703.09752*, 2017, doi: https://doi.org/10.48550/arXiv.1703.09752.
[12] S. Siami-Namini, N. Tavakoli, and A. Siami Namin, "A Comparison of ARIMA and LSTM in Forecasting Time Series," *Proc. - 17th IEEE Int. Conf. Mach. Learn. Appl. ICMLA 2018*, pp. 1394–1401, Jul. 2018, doi: 10.1109/ICMLA.2018.00227.
[13] Y. Himeur, A. Alsalemi, F. Bensaali, and A. Amira, "Smart power consumption abnormality detection in buildings using micromoments and improved K-nearest neighbors," *Int. J. Intell. Syst.*, vol. 36, no. 6, pp. 2865–2894, Jun. 2021, doi: 10.1002/INT.22404;JOURNAL:JOURNAL:1098111X;WGROUP:STRING:PUBLICATION.
[14] C. L. Yun Bai, Jingjing Xie, Chao Liu, Ying Tao, Bo Zeng, "Regression modeling for enterprise electricity consumption: A comparison of recurrent neural network and its variants," *Int. J. Electr. Power Energy Syst.*, vol. 126, p. 106612, 2021, doi: https://doi.org/10.1016/j.ijepes.2020.106612.
[15] A. A. Yassine Himeur, "A Novel Approach for Detecting Anomalous Energy Consumption Based on Micro-Moments and Deep Neural Networks," *Cognit. Comput.*, vol. 12, pp. 1381–1401, 2020, doi: https://doi.org/10.1007/s12559-020-09764-y.
[16] M. I. Khawaja Moyeez Ullah Ghori, "Performance analysis of machine learning classifiers for non-technical loss detection," *J. Ambient Intell. Humaniz. Comput.*, vol. 14, no. 11, 2020, doi: 10.1007/s12652-019-01649-9.
[17] J. Pereira and M. Silveira, "Unsupervised Anomaly Detection in Energy Time Series Data Using Variational Recurrent Autoencoders with Attention," *Proc. - 17th IEEE Int. Conf. Mach. Learn. Appl. ICMLA 2018*, pp. 1275–1282, Jul. 2018, doi: 10.1109/ICMLA.2018.00207.
[18] S. Touzani, J. Granderson, and S. Fernandes, "Gradient boosting machine for modeling the energy consumption of commercial buildings," *Energy Build.*, vol. 158, pp. 1533–1543, 2018, doi: https://doi.org/10.1016/j.enbuild.2017.11.039.
[19] J. W. Cheng Fan, Fu Xiao, Yang Zhao, "Analytical investigation of autoencoder-based methods for unsupervised anomaly detection in building energy data," *Appl. Energy*, vol. 211, pp. 1123–1135, 2018, doi: https://doi.org/10.1016/j.apenergy.2017.12.005.
[20] M. W. Diederik P. Kingma, "An Introduction to Variational Autoencoders," *arXiv:1906.02691*, 2019, doi: https://doi.org/10.48550/arXiv.1906.02691.