Prediction and Analysis of Household Energy Consumption Integrated with Renewable Energy Sources using Machine Learning Algorithms in Energy Management

Nirbhi Jain, Shreya Sharma, Vijyant Thakur *, Mounica Nutakki **‡, Srihari Mandava ***

*B.tech, School of Electrical Engineering, Vellore Institute of Technology, Vellore, TamilNadu, India

** Research Scholar, School of Electrical Engineering, Vellore Institute of Technology, Vellore, TamilNadu, India

***Associate Professor, School of Electrical Engineering, Vellore Institute of Technology, Vellore, TamilNādu, India

(mounica.nutakki2020@vitstudent.ac.in, mandavasrihari@vit.ac.in)

‡ Srihari Mandava; Mounica Nutakki, Vellore, TamilNadu, India, Tel: +9751296300,

mandavasrihari@vit.ac.in

Received: 06.06.2023 Accepted:04.08.2023

Abstract- The project aims to develop an Energy management system (EMS) that can adapt to changes in the building's environment and optimize energy consumption without the need for manual intervention. In an IoT-based smart home context, this work intends to provide predictive models that are driven by data gathered from various sensors to simulate the appliances energy usage. This paper specifically consider two approaches to the prediction problem. First is to determine the important features in forecasting the appliance energy consumption and a general energy consumption model is created utilising machine learning (ML) techniques. A publicly accessible dataset made consisting of historical readings from various humidity and temperature sensors as well as information on the overall amount of energy consumed by household appliances in an smart home that is used to evaluate the performance of the suggested models. By exhibiting significant variation with both the training as well as test data for certain characteristics, among the proposed ML models interms of accuracy score logistic regression beats other models in the prediction outcomes comparison . In order to accurately predict future energy use based on historical energy usage data, we secondly create a time-series models using MA, ARIMA, SARIMAX and LSTM (univariate and multivariate) approach. The suggested predictive models will, in general, allow customers to reduce the power consumption of household appliances and also for utility to more accurately plan and estimate future energy demand to support green urban development.

Keywords EMS, Time-Series forecasting, Machine Learning, LSTM, EDA, RMSE.

1. Introduction

A smart house automates all of its embedded technology and is the domestic extension of building automation. A home that has big home appliances like refrigerators, washers, dryers, and freezers, as well as entertainment systems, TVs, computers, security, and camera systems that can communicate with one another and be controlled remotely by a schedule, mobile device, phone, or internet. The residents of the home use a wall-mounted terminal or a mobile device connected to internet cloud services to control these systems, which are made up of switches & sensors connected to a central hub.

Smart houses provide convenience, low operating costs, security, and energy efficiency. Installing smart devices provides convenience and can save time, money, and energy. These technologies are flexible and responsive to the always changing needs of the homes' residents. Its infrastructure is typically flexible enough to function with a range of equipment from different vendors and standards. By using microcontroller enabled actuators and sensors, the fundamental architecture allows tracking embedded devices in the household, and evaluating the state of the house.

This project aims to utilize machine learning models to estimate the energy usage of household appliances and predict high and low energy consumption based on appliances' energy use, along with weather, humidity, and temperature attributes. Our study uses data from smart metres over a time period of one minute over 350 days of home appliances in kwh and local weather conditions to try and estimate this change in consumption of energy based on weather. Tracking the trends of energy consumption can help with more effective energy control by analysing the data collected from IoT devices.

The main objective is to investigate the relationship between appliance energy consumption and various predictors and compare the predictability rate of different machine learning algorithms. The data collected from a single household is pre-processed and used to train machine learning models such as linear regression, decision trees, and neural networks to predict energy consumption based on the given predictors.

The results and analysis chapter provides a detailed comparison of the predictability rate of each algorithm and highlights the impact of different predictors on energy consumption. This work has significant implications for energy management and sustainability, providing insights into how households and energy providers can optimize energy usage by understanding the impact of various predictors on energy consumption.

Fig 1: Smart home

Rest of the paper describes as follows. Section II is about the literature work. In section III, a detailed explanation regarding the problem statement of proposed work and the methods used are discussed in section IV. Data modelling of the proposed work in shown in section V where as in section VI modelling of proposed methods have been discussed. In Section VII, results as well as the performance evaluation of proposed models is shown and finally section VIII is conclusion.

2. Literature Survey

The need for efficient and sustainable energy management systems has become increasingly important in recent years, driven by concerns about climate change and the depletion of non-renewable energy sources.

- Pereira et al. (2020) discuss the use of AI and ML techniques for energy management systems. They highlight the ability of these techniques to predict energy demand, identify usage patterns, and optimize consumption. The study reviews existing literature on AI and ML for energy management and presents a new approach using deep learning. Their proposed method outperforms previous techniques in terms of prediction accuracy, but limitations include the need for large amounts of data and high computational costs [1].

- Yu et al. (2021) provide a comprehensive literature review of the use of deep reinforcement learning (DRL) in smart building energy management. Their goal is to highlight the benefits of DRL in optimizing building energy consumption and reducing energy waste. The techniques used in the paper are DRL architectures, including DQNs, policy gradient methods, and actor-critic methods. The paper does not propose a new technique but rather compares and discusses previous techniques used in the literature. Limitations of the paper include the lack of experimental validation of the proposed directions for future research and the challenges in implementing explainable DRL models in practical applications [2].

- The goal of Ahrarinouri et al. (2020) is to review the use of multi-agent reinforcement learning (MARL) in energy management systems for residential buildings. The techniques used in the paper are MARL algorithms such as Q-learning, actor-critic, and deep reinforcement learning. The paper does not propose a new technique but rather reviews the previous techniques used in the literature. The comparison of proposed and previous techniques is not explicitly mentioned in the paper. Limitations of the paper include the lack of experimental validation of the suggested directions for future research and the challenges in implementing MARL algorithms in practical applications [3].

- Wijesingha et al.'s (2021) article aims to create a smart residential energy management system (REMS) utilizing machine learning techniques. Regression analysis, neural networks, decision trees, and support vector machines are among the approaches employed in the article. The research does not offer a novel approach but rather examines the application of machine learning techniques in energy management and presents a real example of a smart REMS that employs machine learning. The paper's drawbacks include a lack of experimental validation of the suggested system and the difficulties in implementing such a system in real-world contexts [4].

- Nie et al.'s (2021) purpose is to investigate the usage of gradient-boosting regression trees (GBRT) for forecasting household energy consumption. The strategy utilized in the research is a GBRT-based approach that involves predicting future energy use using previous energy consumption data and meteorological data. The suggested GBRT-based strategy is compared to other machine learning approaches used in energy consumption prediction, such as regression analysis, artificial neural networks, and support vector machines, in the study. The GBRT model predicts household energy usage with excellent accuracy and surpasses other models in terms of prediction error and computing efficiency. The paper's limitations include a lack of explanation of the GBRT model's interpretability and the potential biases in the historical data utilized for training [5].

- Krishna Prakash and D. Prasanna Vadana's (2017) research aims to assess current literature on residential energy management systems (REMS) and evaluate their strategy utilising machine learning techniques such as neural networks and fuzzy logic. The authors compare their technique to existing machine learning models and highlight its potential for usage in smart grid systems and lowering consumer energy prices. The research does not suggest a new approach, but rather assesses the performance of existing strategies that have been employed in the literature. The paper's limitations include a lack of experimental validation of their technique on a large dataset, as well as difficulties in implementing their approach in practical applications [6].

- The paper by Nguyen et al. (2019) proposes a KNN-based approach for appliance classification in HEMS to accurately identify energy consumption for efficient energy management. The authors compare the proposed method to existing approaches and demonstrate its superior accuracy and efficiency. They suggest that incorporating additional features such as voltage and current could further enhance the method. However, the paper's limitations include the need for further testing on larger datasets and the potential for inaccuracies in appliance classification if the features used are not representative of the appliance's true energy consumption [7].

- Yu et al. (2019) propose a load forecasting model based on smart metre data and a gradient boosting decision tree method to improve the efficiency of energy management systems. The paper provides a thorough examination of existing load forecasting systems, highlighting their shortcomings. In terms of accuracy, the suggested model based on smart metre data and GBDT outperforms existing techniques. The study's shortcomings include the need for additional testing on different datasets and the possible difficulties of using the proposed model in real-world energy management systems. Despite these drawbacks, the work offers crucial insights into how ML algorithms might be used to address issues with sustainability and energy management [8].

- In the context of Medan, Indonesia, the work by Aisyah and Simaremare (2021) makes a significant advance to our understanding of the relationship between weather conditions and energy consumption. The influence of temperature, humidity, and rainfall on the region's electricity demand can be better understood through the use of statistical modelling and correlation analysis. However, the study is constrained by its concentration on a particular region and use of information from a short period of time. To investigate the effects of other factors on power demand and to validate the results in different contexts, more study is required. The study does not take into consideration future research improvements in consumer behaviour or energy policy that might have an impact on power demand in the future. Despite these drawbacks, the results of this study can aid energy suppliers and legislators in boosting sustainability and energy efficiency in the area [9].

An in-depth description of the ant lion optimizer algorithm's inspiration and pseudocodes is given in [10]. Then, an exact model of Turkey's monthly electrical energy usage for the spring and summer seasons is presented. the quadratic model used in the fitness function impedes efforts to increase modelling accuracy, whereas the exponential model applied in the fitness function helps to reduce modelling errors. For the estimation of monthly electrical energy consumption, it is also discovered that the inputs of daily average temperature, daily average wind speed, daily average humidity, daily average rainfall, and daily total global solar radiation are appropriate and effective. Additionally, it is anticipated for both seasons that the developed electrical energy consumption models will ease the operation of the power system, aid in the adoption of demand-side management strategies for electricity suppliers, and boost the potential for energy savings. Similar modelling analysis for the autumn and winter seasons needs to be done in more research in order to see how they affect electrical energy consumption modelling.

Table 1: Literature Work

Ref	Aim	Model	Limitations
[4]	short-term electricity load	hybrid mode decomposition algorithms	didn't compare with other methods
[19]	short-term household load forecasting	PSO	facing issue with energy storage
[20]	long term load forecasting	different ML models	long term LSTM based model
[21]	short-term load forecasting	ANN, ATR and RNN	accuracy is neglected
[22]	long- and short-term load forecasting	CNN, LSTM and GRU	methods used only for short term load forecasting

3. Problem Statement

Several studies have shown the potential of machine learning models to predict household energy consumption based on various predictors such as weather, humidity, and temperature. These studies demonstrate the potential of machine learning models to predict household energy consumption and the importance of considering various predictors such as weather, humidity, and temperature. The motivation behind this project is to build on these findings and investigate the potential of machine learning models to estimate energy usage of household appliances, which can provide more granular insights into energy consumption patterns and enable more targeted energy optimization strategies. In order provide insight into the advantages and disadvantages of each machine learning method for forecasting energy use, the research also compares the prediction rates of various algorithms. This analysis can assist determine the best technique for estimating energy usage in various scenarios and serve as a foundation for further study.

4. Methodology

The methodology for this research project involved several key steps. First, the data was collected from various sources, including weather stations, energy suppliers, and building management systems. This data was then pre-processed to ensure that it was in a usable format for our analysis[11-12]. This involved steps such as removing missing or erroneous values, scaling the data, and converting it to a suitable format.

Next, time series analysis is used to identify patterns and trends in the data. This involved techniques such as decomposing the time series into its seasonal, trend, and residual components, as well as performing autocorrelation and partial autocorrelation analysis. various visualization techniques are also used, such as line plots and heatmaps, to gain insights into the data and identify any potential outliers or anomalies. After identifying patterns and trends in the data, then it is proceeded to develop several machine learning models. These models were designed to predict future energy consumption based on historical data and external factors such as weather patterns. Several popular machine learning algorithms are implemented, including linear regression, decision tree, random forest, and gradient boosting, and evaluated their performance using various metrics such as root mean squared error (RMSE) and coefficient of determination (R2).

In addition to the machine learning models, a deep learning model using a Long Short-Term Memory (LSTM) neural network is also developed. This model was specifically designed to handle time series data and was trained on historical energy consumption data to predict future energy consumption. To validate the performance of our models, the dataset was splitted into training and testing sets and evaluated their performance on the testing set. We also conducted a comparative analysis of all the models to identify which one performed best in terms of RMSE and R2. Finally, the results of our analysis are discussed and provided insights into the potential applications of our models. We also identified several design constraints and trade-offs that need to be considered when developing energy prediction models, such as the need for accurate and reliable data, the trade-off between model complexity and performance, and the need to balance short-term and long-term predictions. Overall, the methodology employed in this project was designed to provide a comprehensive analysis of energy prediction models using a range of techniques, including data pre-processing, time series analysis, and machine learning. By combining these techniques, we were able to develop accurate and reliable energy prediction models that can be used to improve energy efficiency and reduce costs for building owners and energy suppliers.

5. Data Analysis

The dataset used in this research project consists of energy consumption data collected from a variety of sources containing a one-minute readings in the year 2016 for total 350 days of domestic appliances in kW from a smart meter, as well as the weather data in the city of San Francisco, California region. This dataset includes energy statistics from household appliances as well as weather information. ML models are used to estimate energy usage and predict energy consumption using data on temperature, humidity, visibility, pressure, windspeed, dew point, and probability of precipitation. The main purpose is to forecast future energy use using current and historical data.

A few insights that discovered were that energy usage soared from June to October and that furnaces dropped during the months of June and July, maybe owing to hot weather. During the months of August and September, the wine cellar also required a significant quantity of energy. This overall shows a huge energy consumption trend during the major months of June, July, and August. To view a better consumption pattern, separate and define additional columns from date-time, and look at a trend across months, weeks, and hours in a day. Fig 2 provides a more complete and in-depth explanation. Over the course of a week, we can clearly observe that the home office consumes less on weekends and more in the wine cellar. During the weekend, the kitchen and microwave use less electricity. Time-series patterns may be seen in the home office, refrigerator, wine cellar, living room, and furniture. This is because these appliances must maintain a steady interior temperature or change to a pleasant temperature based on the season. The use of energy is low during the day and high at night. The creation of energy is highest during the day and lowest at night. This is assumed to be because energy creation is encouraged during the day because there are no people at home, and consumption increases at night when residents come home.

In Fig 3, As there are so many data points, we used the daily average. From July through September, energy usage soars. Energy generation has no big peaks, but instead steadily climbs from January to July and then gradually drops.

Fig 2: Average consumption per hour

Fig 3: House overall corresponds to the sum of other consumption

6. Modelling

After pre-processing and visualizing the dataset, the next step is to model using different ML models along with time series analysis-based models.

6.1 Machine Learning Models

6.1.1 Logistic Regression:

Logistic Regression [13] is a commonly used statistical model for binary classification applications. Given the input characteristics, the logistic function transfers every real-valued input to a probability output between 0 and 1, which reflects the likelihood of the positive class (y=1). This is how the logistic function is defined (1):

$$\mathrm{P}(\mathrm{y}=1|\mathrm{X}) = \frac{1}{1+e^{-z}} \quad (1)$$

Where z in (2) represents the linear combination of the predictor variables and their weights (w):

$$Z = \omega_0 + \omega_1 X_1 + \omega_2 X_2 + \cdots \dots \dots \dots \dots + \omega_n X_n \quad (2)$$

Here the predictor variables are replaced with different weather parameters in order to understand how they affect the energy consumption patterns and also improves the accuracy of the predictions.

6.1.2 Decision tree classifier

This algorithm recursively partitions the data into smaller subsets based on the values of the input features and creates decision nodes shown in Fig 4 that determine the best feature and value to split the data at each step [14]. Here our main aim was to create a tree that accurately predicts the target variable by minimizing the impurity of each node. We strive to restrict the depth of the tree since deep trees are prone to overfitting; this frequently helps to enhance the model's generalization performance.

Fig 4: Decision tree classifier

6.1.3 K-Nearest Neighbour:

K-Nearest Neighbor being a non-parametric classification algorithm [15] uses the distance between data points in order to make predictions. KNN helps predict the class of a test instance by locating the K training examples which are closest in the feature space and also by assigning the test instance to the majority class among its K nearest neighbors. After comparing we could see that the Euclidean Distance formula with K=5 gives better accuracy for the energy dataset. To formulate the given distance metric we use the formula (3):

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (3)$$

6.1.4 AdaBoost Classifier:

The Adaboost algorithm is a form of ensemble learning that combines numerous weak classifiers to create a strong classifier as shown in Fig 5. As AdaBoost is sensitive to noisy data as well as outliers, even with a large number of weak classifiers [16], it makes it a powerful algorithm that achieves high accuracy for classification tasks.

Fig 5: AdaBoost Classifier

6.1.5 Random Forest Classifier:

Random Forest classifiers are based on the idea of decision trees, which are a simple yet powerful model for making predictions. In a random forest [17], multiple decision trees are trained on different subsets of the training data and with different subsets of the input features. This randomness in the selection of features and training data helps to reduce overfitting and improve the generalization performance of the model.

6.2 Time Series Analysis

6.2.1 ARIMA:

ARIMA is one of the simple and most effective algorithms mainly used for performing time series forecasting [18]. It is a statistical analysis traditional model [19] that uses time series data to either for better understanding the dataset or for predicting the future trends. ARIMA is used in this project to predict future overall energy consumption per day in a smart home.

- The first step in developing an ARIMA model is to create a stationary time series. When a linear regression model in ARIMA is referred to as "auto-regressive," it signifies that it employs its own lags as predictors.
- In this project ARIMA is used to forecast the future values of the total energy used in the smart home (column= 'use') by training the model on a dataset of the size of 351 days.
- The p and q parameters of the model have been chosen to be based on the PACF and ACF plots respectively and since the data is stationary, we could choose the d value as zero, but we have also experimented with first-order differencing.

6.2.2 SARIMAX:

SARIMAX (Seasonal Autoregressive Integrated Moving Average with Exogenous Variables) is a time series forecasting model that takes into account the influence of external variables on the time series as well as seasonal and non-seasonal components [20]. The following elements are used in the model:

- Seasonal ARIMA model: To account for seasonal fluctuations in the time series data, SARIMAX uses a seasonal ARIMA model.
- Exogenous variables: SARIMAX enables the addition of outside factors that may have an impact on the time series being predicted. SARIMAX can capture the impacts of variables that are external to the time series being projected by including exogenous variables.

6.2.3 Univariate LSTM:

The univariate LSTM model is a type of LSTM model that is designed to process and predict time series data with only one variable [21-22]. Here, exploring the application of LSTM for univariate time series analysis, where it is mainly focused on predicting energy consumption based on historical data.

6.2.4 Multivariate LSTM:

Multivariate LSTM models are powerful tools for predicting energy consumption, as they can incorporate information from multiple sources to generate more accurate predictions [23-24] The Multivariate LSTM model is designed to take into account multiple variables that are expected to impact energy consumption, including weather parameters such as precipitation, temperature, humidity, and wind speed [25-26]. The inclusion of weather parameters can have a significant impact on the accuracy of the model, as weather conditions can strongly affect energy consumption patterns[27-28].

7 Results and Analysis

The project aimed to develop reliable and efficient energy forecasting models using ML and time series analysis. To achieve this goal, several algorithms such as Random Forest, Logistic Regression, KNN, Decision Tree, AdaBoost Classifier, ARIMA, SARIMAX, Univariate LSTM and Multivariate LSTM were applied to the dataset. The findings of the study can have significant implications for energy management practices in various sectors such as residential, commercial, and industrial. The ARIMA and SARIMAX model graphs also show a decrease in the discrepancy between projected and actual values over time, as depicted in Figures 6 and 7. Fig 7, which is a SARIMAX multivariate model, demonstrates how the test and model data fit together. Comparing the univariate and multivariate models, it is evident that including additional weather characteristics improves the accuracy of the model significantly.

Fig 6: Test, train, model graph comparison for ARIMA model

Fig 7: Test, train, model graph comparison for SARIMAX model

Figures 8 and 9 demonstrate how the discrepancy between projected and actual values is decreasing. Fig 9, which is an LSTM Multivariate, demonstrates how the test and model data fit together. When we compared the univariate and multivariate models, we could see that including all of the weather characteristics improved the fit significantly.

Fig 8: Test, train, model graph comparison for LSTM Univariate model

Fig 9: Test, train, model graph comparison for LSTM Multivariate model

Table 2 presents the RMSE (Root Mean Square Error) and R2 (Coefficient of Determination) values for various ML models used to forecast energy consumption patterns. RMSE measures the difference between actual & predicted values, with lower values indicating a better fit between the model and the data. R2 quantifies the percentage of the target variable's variation that can be predicted from the independent variables. Higher values denote a better fit.

Table 2: RMSE and R2 Values for Different ML Models

Model	R^2 Score	RMS Score
Logistic Model	0.670	0.280
KNN	0.520	0.114
Random Forest	0.695	0.072
Decision Tree Classifier	0.331	0.331
AdaBoost Classifier	0.730	0.06

Table 3: Accuracy Compression of Proposed Models

Model	Accuracy
Logistic Model	94
KNN	92.4
Random Forest	90.6
Decision Tree Classifier	86.9
AdaBoost Classifier	84

The AdaBoost Classifier and Random Forest models fared the best, with RMSEs of 0.064 and 0.072, respectively, suggesting a smaller discrepancy between predicted and actual values. Furthermore, these models obtained high R2 values of 0.730 and 0.695, indicating a solid match between the model and data. Decision Tree Classifier on the other hand, fared the poorest, with an RMSE of 0.331 and a R2 score of 0.331.

The ROC curve highlights the sensitivity of the classifier model by displaying the ratio of true positives to false positives. The model is better the higher the AUC.AUC-ROC curves are used to visually depict the trade-off between specificity and sensitivity for every potential cut-off for a test being run. The area under the ROC curve in Figure 15 shows the benefit of using the test for the underlying question. AUC-ROC curves are an alternative indication for classification problems at varying threshold levels. A test's effectiveness will rise as the AUC-ROC curve gets closer to the upper left corner.

Fig 10: ROC curve and Accuracy curve of Logestic Regression model

Fig 11: ROC curve and Accuracy curve of AdaBoost Classifier model

Fig 12: ROC curve and Accuracy curve of Decision Tree Classifier model

Fig 13: ROC curve and Accuracy curve of KNN model

Fig 14: ROC curve and Accuracy curve of Random Forest Classifier model

Fig 15: ROC curve and Acuuracy curve of AdaBoost Classifier model

Therefore, to merge the True Positive Rate (TPR) & False Positive Rate (FPR) into one metric, the logistic regression & KNN were used to produce the two initial metrics with a variety of various thresholds as illustrated in Figures (10–14) and then they were plotted on a single graph resulting in figure 15. The resulting curve metric we consider is the area under this curve, which is called as AUC-ROC. When TPR and FPR are equal, as seen by the dashed diagonal line in figure 15, an AUC of 0.5 is indicated. Each model's ROC-AUC curve is displayed in figures [7-11] and is better as long as the AUC is greater than 0.5.

Table 4: RMSE and R2 Values for Time-Series ML Models

Model	R^2 Score	RMSE Score
Logistic Model	0.77	0.266
KNN	-2.379	0.509
Random Forest	-0.315	0.243
Decision Tree Classifier	0.106	0.173
AdaBoost Classifier	0.700	0.110

Table 4 is representing RMSE as well as R2 values for all the time series analysis-based models. This includes a base model which MA or Moving Average, ARIMA, SARIMAX, LSTM Univariate, and LSTM Multivariate. On the results bases, it can be observed that the LSTM Multivariate performs the best with an RMSE score of 0.110, which indicates a lower difference between predicted and actual values. Also, this model had a higher R2 score of 0.700, indicating a strong fit in between the model as well as the dataset. On, the other hand, the ARIMA model performed the worst with an RMSE of 0.509 and an R2 score of -2.379.

8 Conclusion

Numerous alternative prediction techniques can be used to predict energy usage, according to a literature review on prediction analysis in home energy management systems. These techniques include hybrid, machine learning, and statistical techniques. The application of statistical techniques is often straightforward and is based on historical data. They aren't usually particularly reliable, especially when it comes to estimating energy use in homes with fluctuating energy needs. Although machine learning techniques can be more accurate than statistical techniques, they are more complicated. The use of machine learning techniques allows for the identification of trends in previous data that may be used to forecast future energy usage. Methods that integrate machine learning and statistics are called hybrid methods. By combining the advantages of the two methodologies, this can increase the accuracy of predictions.

Based on the results obtained from the various machine learning models and time series analysis, it can be concluded that the LSTM multivariate model performed the best in predicting energy consumption for the given dataset. The values of RMSE obtained for the LSTM model were lower than the other models, indicating that it was better at predicting energy consumption values. The R2 score for the LSTM model was also high, indicating that it had a good fit with the data. In conclusion, this project demonstrated the performance of ML models and time series analysis in forecasting the future energy consumption for a building. The results obtained can be used by building owners and energy managers to optimize energy usage and reduce costs. However, further study is required to improve the reliability and generalizability of the models and to incorporate additional parameters that affect energy consumption.

Acknowledgements

Authors may acknowledge to any person, institution or department that supported to any part of study.

References

[1] L. S. B. Pereira, R. N. Rodrigues, E. A. C. A. Neto, "Modeling of Energy Management Systems using Artificial Intelligence," 2020 IEEE International Systems Conference (SysCon), 2020, pp. 1–6.

[2] L. Yu, S. Qin, M. Zhang, C. Shen, T. Jiang, and X. Guan, "A review of deep reinforcement learning for smart building energy management," IEEE Internet of Things Journal, vol. 8, no. 15, pp. 12046–12063, 2021.

[3] M. Ahrarinouri, M. Rastegar, and A. R. Seifi, "Multiagent reinforcement learning for energy management in residential buildings," IEEE Transactions on Industrial Informatics, vol. 17, no. 1, pp. 659–666, 2020

[4] J. Wijesingha, B. R. Hasanthi, I. Wijegunasinghe, M. Perera, and K. Hemapala, "Smart residential energy management system (rems) using machine learning," pp. 90–95, 2021

[5] Nie, M. Roccotelli, M. P. Fanti, Z. Ming, and Z. Li, "Prediction of home energy consumption based on gradient boosting regression tree," Energy Reports, vol. 7, pp. 1246–1255, 2021.

[6] Krishna Prakash and D. Prasanna Vadana, "Machine learning based residential energy management system," in Proceedings of the 2017 IEEE International

Conference on Computational Intelligence and Computing Research (ICCIC), Tamil Nadu, India, pp. 14–16,2017.

[7] D. Nguyen, T. D. Do, M. H. Le, N. T. Le, and W. Benjapolakul, “Appliance classification method based on k-nearest neighbors for home energy management system,” in 2019 First International Symposium on Instrumentation, Control, Artificial Intelligence, and Robotics (ICASYMP), pp. 53–56, IEEE, 2019.

[8] X. Yu, Z. Xu, X. Zhou, J. Zheng, Y. Xia, L. Lin, and S.-H. Fang, “Load forecasting based on smart meter data and gradient boosting decision tree,” in 2019 Chinese Automation Congress (CAC), pp. 4438–4442, IEEE, 2019.

[9] Aisyah and A. A. Simaremare, “Correlation between weather variables and electricity demand,” in IOP Conference Series: Earth and Environmental Science,vol. 927, p. 012015, IOP Publishing, 2021.

[10] Yesilbudak, Mehmet, Ozge Sagliyan, and Ayse Colak. "Monthly electrical energy consumption modeling using ant lion optimizer." 2019 8th International Conference on Renewable Energy Research and Applications (ICRERA). IEEE, 2019.

[11] Harrouz, A., Temmam, A., & Abbes, M. (2018). Renewable energy in algeria and energy management systems. International Journal of Smart Grids, ijSmartGrid, 2(1), 34-39.

[12] Atasoy, T., Akınç, H. E., & Erçin, Ö. (2015, November). An analysis on smart grid applications and grid integration of renewable energy systems in smart cities. In 2015 international conference on renewable energy research and applications (ICRERA) (pp. 547-550). IEEE.

[13] Tso, Geoffrey KF, and Kelvin KW Yau. ”Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks.” Energy 32.9 (2007): 1761-1768.

[14] Wahid, Fazli, and D. Kim. ”A prediction approach for demand analysis of energy consumption using k-nearest neighbor in residential buildings.” International Journal of Smart Home 10.2 (2016): 97-108.

[15] An, G., Jiang, Z., Cao, X., Liang, Y., Zhao, Y., Li, Z., ... & Sun, H. (2021). Short-term wind power prediction based on particle swarm optimization-extreme learning machine model combined with AdaBoost algorithm. IEEE access, 9, 94040-94052.

[16] Wang, Z., Wang, Y., Zeng, R., Srinivasan, R. S., & Ahrentzen, S. (2018). Random Forest based hourly building energy prediction. Energy and Buildings, 171, 11-25.

[17] Sagiroglu, S., Terzi, R., Canbay, Y., & Colak, I. (2016, November). Big data issues in smart grid systems. In 2016 IEEE international conference on renewable energy research and applications (ICRERA) (pp. 1007-1012). IEEE.

[18] Wang, Xiping, and Ming Meng. ”A Hybrid Neural Network and ARIMA Model for Energy Consumption Forcasting.” J. Comput. 7.5 (2012): 1184-1190.

[19] Tarsitano, Agostino, and Ilaria L. Amerise. ”Short-term load forecasting using a two-stage sarimax model.” Energy 133 (2017): 108-114.

[20] Oymak, A., & Tur, M. R. (2022). A Short Review on the Optimization Methods Using for Distributed Generation Planning. International Journal of smart grid-IJSMARTGRID, 6(3), 54-64..

[21] Zhang, Q., Sun, Y., & Cui, Z. (2010, December). Application and analysis of ZigBee technology for Smart Grid. In 2010 International Conference on Computer and Information Application (pp. 171-174). IEEE.

[22] Kim, Tae-Young, and Sung-Bae Cho. ”Predicting residential energy consumption using CNN-LSTM neural networks.” Energy 182 (2019): 72-81

[23] Succetti, F., Di Luzio, F., Ceschini, A., Rosato, A., Araneo, R., & Panella, M. (2021, September). Multivariate prediction of energy time series by autoencoded LSTM networks. In 2021 IEEE International Conference on Environment and Electrical Engineering and 2021 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe) (pp. 1-5). IEEE.

[24] Khan, Z. A., Ullah, A., Haq, I. U., Hamdy, M., Mauro, G. M., Muhammad, K., ... & Baik, S. W. (2022). Efficient short-term electricity load forecasting for effective energy management. Sustainable Energy Technologies and Assessments, 53, 102337.

[25] Zhao, X., Gao, W., Qian, F., & Ge, J. (2021). Electricity cost comparison of dynamic pricing model based on load forecasting in home energy management system. Energy, 229, 120538.

[26] Bui, V., Le, N. T., Nguyen, V. H., Kim, J., & Jang, Y. M. (2021). Multi-behavior with bottleneck features LSTM for load forecasting in building energy management system. Electronics, 10(9), 1026.

[27] Panapongpakorn, Tanwalai, and David Banjerdpongchai. ”Short-term load forecast for energy management systems using time series analysis and neural network method with average true range.” 2019 First International Symposium on Instrumentation, Control, Artificial Intelligence, and Robotics (ICA-SYMP). IEEE, 2019.

[28] Atasoy, T., Akınç, H. E., & Erçin, Ö. (2015, November). An analysis on smart grid applications and grid integration of renewable energy systems in smart cities. In 2015 international conference on renewable energy research and applications (ICRERA) (pp. 547-550). IEEE.