Science in Information Technology Letters
Vol. 6., No. 1, May 2025, pp. 1-15
ISSN 2722-4139
http://pubs2.ascee.org/index.php/sitech

Machine learning-based residential load demand forecasting: Evaluating ELM, XGBoost, RF, and SVM for enhanced energy system and sustainability

Modawy Adam Ali Abdalla [a,1], Ahmed Mohamed Ishag [a,b,2,*], Hassan Ahmed Osman [c,3], Mohamed Elhindi [d,4], Nasreldin Ibrahim [e,f,5], Aissa Snani [g,6], Gomaa Haroun Ali Hamid [a,7], Abdallah Hammad[h,i,8]

[a] Department of Electrical and Electronic Engineering, College of Engineering Science, Nyala University, Nyala 63311, Sudan
[b] School of Engineering and Applied Science, Kampala International University, Kampala, P.O. Box 20000, Uganda
[c] Department of Chemistry, College of Education, Nyala University, Nyala 63311, Sudan
[d] College of Energy and Electrical Engineering, Hohai University, Nanjing 211100, China
[e] School of Electrical and Information Engineering, Tianjin University, Tianjin, Tianjin 300072, China
[f] Department of Electrical Engineering, Faculty of Engineering, University of Al Fashir, Al Fashir 61111, Sudan
[g] LabGED, Department of Computer Science, University of Badji Mokhtar, Annaba 23000, Algeria
[h] Department of Electrical Engineering, College of Engineering, University of Bisha, 61922, Saudi Arabia
[i] Department of Electrical Engineering, Faculty of Engineering at Shoubra, Benha University, Egypt
[1] brojacter88@yahoo.com; [2] Ahmed.ishag@kiu.ac.ug ; [3] hassanahmed94@yahoo.com;
[4] mohamedalhindy1394@gmail.com; [5] nasreldinibrahim2@gmail.com; [6] aissa.snani@univ-annaba.org;
[7] gomaaharoun1982@gmail.com; [8] abdallah.hammad@ub.edu.sa
* Corresponding Author

ARTICLE INFO

Article history
Received April 29, 2025
Revised May 12, 2025
Accepted May 12, 2025

Keywords
Residential energy consumption
Sustainable energy system
Load forecasting
Extreme learning machines
Machine learning algorithms

ABSTRACT

Accurate forecasting of electrical power load is essential for properly planning, operating, and integrating energy systems to accommodate renewables and achieve environmental sustainability. Therefore, this study introduces different machine learning (ML) methods, including support vector machines (SVM), random forests (RF), extreme learning machines (ELM), and extreme gradient boosting (XGBoost) to predict hourly electricity demand using electricity consumption and temperature data for train and test ML models. The data is processed by autocorrelation function (ACF) and cross-correlation function (CCF) to determine the appropriate lag time for the inputs. Furthermore, ML model accuracy is assessed using coefficient of determination (R^2), mean absolute error (MAE), and root mean square error (RMSE). Results show that the ELM model achieved the highest R^2 in both summer (0.971) and winter (0.868), outperforming the other models in accuracy R^2 and error reduction (MAE and RMSE). ELM also more effectively captured load fluctuations. The result of this research has applications for load demand forecasting in the proper planning and operation of the residential grid. The results help estimate load demand and provide useful guidance for residential grid planning and management by determining the best techniques for precisely estimating load demand and identifying domestic energy consumption patterns.

1. Introduction

Power generation is facing a major challenge from the rising demand for electricity, which will increase the utilization of petroleum and harmful emissions [1]–[5]. Due to urbanization and population growth, the residential sector is predicted to have the fastest-growing electricity consumption among all sectors, with an annual increase of 1.2% by 2040. In addition to consuming more than 40% of global power, residential buildings are accountable for one-third of all CO2 emissions [6]–[9]. This rapid increase in residential energy consumption places immense strain on the power grid [10], [11]. Therefore, accurate electricity consumption prediction is essential in properly planning, operating, and integrating energy systems to accommodate renewable energy sources and achieve environmental sustainability, thus reducing power grid strain [12]–[15]. Consequently, analyzing electricity consumption data and forecasting accurate electricity load demand in buildings is essential for optimizing energy use and achieving energy savings.

On the other hand, energy consumption changes dramatically during the year, and the nature of consumption varies from season to season, even day to day, or from time to time due to season differentiation of the year and the effect of utilization factors. So, obtaining a real load demand profile during the year is very difficult. At the same time, predicting load demand is quite important for each energy provider and consumer, which will help make plans to schedule loads based on the available generation and is a useful reference for improving energy management. Therefore, several methods have been used to forecast loads, such as traditional and modified traditional techniques, including the stochastic time series, exponential smoothing, multiple regression, adaptive load forecasting, regression method, etc. [15], [16]. However, these methods have low prediction accuracy, and the prediction error is high when the weather changes suddenly. In integrated energy systems, one of the most crucial pillars for attaining the energy balance among the grid and consumers is effective load prediction. In addition, accurate load forecasting is essential to formulate the planning and operation strategies of energy generation, transmission, and distribution systems [17]. Some academics looked into how accurate load forecasting affected electrical energy systems. For example, Ranaweera et al. [18] examined the effect of inaccurate load forecasts, enhancing daily peak load, and the influence of the various year seasons on the power systems. In the relevant study, the authors confirmed that load forecasting accuracy would increase the economic benefits of the energy systems and markets. In another study, Monforte et al. [19] examined load forecasting and developed frameworks for improving energy utilization. They identified the alternative methods of load forecasting that provide value to the grid and its consumers. The authors identified that the best load forecasting methods are based on ML. Recently, ML methods have been used widely for load forecasting.

Wang et al. [20] utilized RF to forecast building energy consumption. RF model has been compared with regression tree (RT), and support vector regression (SVR); they found that the RF outperformed SVR and RT in performance index, which were 5–5.5% for SVR and 14–25% for RT, respectively. In another work, Shao et al. [6] examined energy consumption forecasting through SVM, and the authors showed that the MSE and R2 are equal to 2.22 % and 0.94, respectively. Furthermore, Abbasi et al. [21] investigated the prediction of load by using the XGBoost algorithm. In the study, the daily load data were converted to weekly data time series, and XGBoost was utilized to extract and select the characteristics from the data with the aim of load forecasting. In the relevant study, the mean absolute percentage error value equals 10.08%, and the MAE value equals 88.90%, with 97.21% accuracy. Li et al. [22] investigated a novel ensemble approach for short-term demand prediction according to three different kinds of forecast methods: ELM, partial least squares regression, and wavelet transform. The authors demonstrate how the ELM approach can greatly enhance prediction accuracy. To obtain the proper planning and operation of energy management strategies, the load profile must be achieved over a period of time. Nevertheless, it is extremely challenging to obtain actual yearly electricity consumption curve. Consequently, designers employ a single day's average daily electricity consumption [23]. However, the daily average load for one day does not give an accurate description of the load profile due to energy consumption changing dramatically during the year, and the nature of consumption varies from season to season, even day to day, or from time to time.

Accurate electricity demand forecasting is critical for efficient power management and system design. While previous studies have extensively employed ML algorithms due to their high accuracy, certain research gaps persist [24]–[27]. Despite tremendous progress in forecasting energy usage in residential buildings, there are still issues with the processing and quality of energy consumption data. In particular, current methods frequently fall short of identifying and utilizing the insightful information in this data. Thus, enhancing data processing and predicting accuracy are the primary goals of this research. The data is first processed and examined utilizing ACF and CCF to improve feature extraction and temporal analysis. A number of ML methods, such as RF, SVM, XGBoost, and ELM, are also used to improve energy consumption prediction. Model performance is thoroughly assessed using measures like R^2, MAE, and RMSE to ensure a thorough evaluation of their efficacy. This method seeks to fully utilize energy consumption data for more precise and trustworthy load forecasting by addressing the current constraints.

2. Method

To properly design, operate, and assess the energy system's performance and efficiencies, the demand for household electrical power must be accurately predicted. Making crucial and accurate decisions, such as buying and producing electrical energy, is aided by precise forecasts. Fig. 1 depicts the research methodology used in this work, which includes multiple crucial phases.

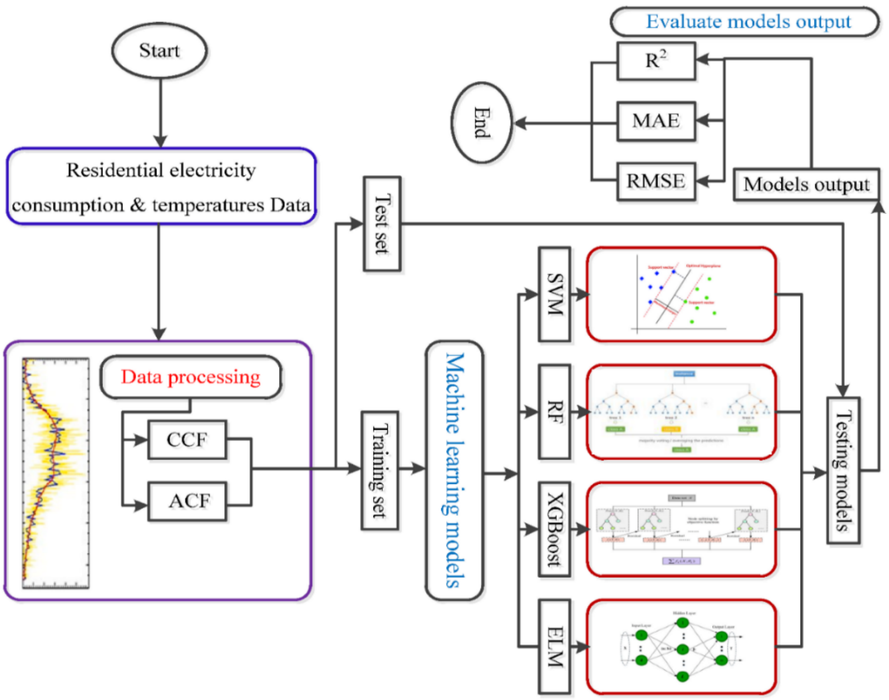

Fig. 1. The electricity consumption forecasting model schematic diagram

The forecasting process uses temperature and load demand data as inputs. Furthermore, these inputs are preprocessed using ACF and CCF to enhance the data accuracy and identify pertinent features. Thus, ML approaches, including SVM, RF, ELM, and XGBoost, are used to forecast load demand. The ML performance is assessed by different key metrics: R^2, MAE, and RMSE. This structured methodology ensures a robust approach to enhance the precision and reliability of electricity demand in the residential grid.

2.1. Machine Learning Algorithms Descriptions

2.1.1. Support Vector Machine (SVM)

Cortes [28] proposed the SVM model, which is a kernel-based supervised learning method designed to construct the optimal separating hyperplane across distinct categories. This process relies on binary classification in the field of arbitrary characteristics and is thus suited for predicting problems. SVM might learn and predict the nonlinear relationships between data in higher dimensions, minimizing the measured training error and distributing error extent to achieve generalized regression efficiency [29]. SVM is actually a productive learning technique founded on efficient optimization theory. The following equation can be used to represent SVM [30].

$$f(x) = \varpi\phi(x) + b \tag{1}$$

where $\phi(x)$ is a function that can convert x into the high-dimensional feature spaces, x is the input data, ω and b are the weights vector and a threshold, respectively, which can be estimated by minimizing the following regularized risk function [30].

$$R(C) = \frac{1}{2}||\varpi|\quad|^2 + C\frac{1}{n}\sum_{i=1}^{n} L(d_i, y_i) \tag{2}$$

where $\frac{1}{2}\|\omega\|^2$ is the regularized term, C is the error penalty parameter, n is the number of the observations, d_i is the desired value, and $C\frac{1}{n}\sum_{i=1}^{n} L(d_i y_i)$ is an error of empirical, in which the function L_ε can be expressed as [30].

$$L_k(d, y) = |d - y| - \varepsilon|d - y \geq \varepsilon \text{ } or \text{ } 0 \text{ } otherwise \tag{3}$$

where ε is the tube size. The approximated function $f(x)$ can be expressed by introducing Lagrange multipliers as follows [28].

$$f(x, \alpha_i, \alpha_i^*) = \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) K(x, x_i) + b \tag{4}$$

where α_i, α_i^* are Lagrange multipliers, and $K(x, x_i)$ is the kernel function. The kernel function uses the Gaussian kernel function $K(x_k, x)$, which is formulated as [30].

$$K(x_k, x) = \exp\left(-\gamma |x_k - x|^2\right) \tag{5}$$

where γ is the parameter of the kernel function

2.1.2. Random Forest (RF)

Breiman [31] presented the RF model, which used his "bagging" idea to set a collection of decision trees with controlled variation. A group of distinct classification and regression trees (CART) that are trained using bagging and random variable selection make up the ensemble prediction model known as RF [20]. It is also a set of decision trees that are based on the numbers of random vectors sampled separately with almost the same frequency for all tree predictors [32]. The RF method is frequently used for regression and prediction issues and is designed to give accurate estimation while avoiding overfitting the data. The CART utilized in RF depends on the Gini coefficient chosen. The criteria for each child node attaining its maximum purity, with all observations on that child node falling into the same categorization, is used to determine the Gini coefficient. The following formula can be used to determine the Gini coefficient of CART [33].

$$Gini(p) = 2p(1-p) \tag{6}$$

If a feature $A = a$ is utilized, D it is separated into two sections upon passing each segmentation point: D_1 the sample set that satisfies $A = a$ and D_2 the sample set that satisfies $A \neq a$. The Gini coefficient with the characteristic requirement is.

$$Gini(D, A) = \frac{D_1}{D} Gini(D_1) + \frac{D_2}{D} Gini(D_2) \tag{7}$$

where $Gini(D, A)$ denotes the uncertainty of D. Each CART in RF is intended to find the segmentation point of the feature with the fewest Gini coefficients by continually iterating over each possible segmentation point of the feature in the tree and dividing the data set into two subsets until the stop condition is met.

2.1.3. Extreme Learning Machine (ELM)

The ELM model was first presented by Huan [34], who claimed it could learn faster than feed-forward neural network (FFNN) algorithms with greater generalization capabilities at incredibly fast learning speeds. For single-layer feed-forward networks (SLFNs), ELM is a sophisticated data-driven technique that defines the output weights of SLFNs analytically and chooses hidden nodes at random. An input layer, a hidden layer (made up of neurons), and an output layer make up the ELM model. Classical FFNN models require tuning of all model parameters, whereas the ELM model does not require tuning of the hidden layer. With the lowest training error and the smallest weight norm, the network outperforms gradient descent and backpropagation techniques in terms of enumerated capacity. In general, the ELM model with ℓ hidden nodes and an activation function $g(x)$ can be expressed as follows [35].

$$\sum_{\mathrm{i}=1}^{\ell} \beta_{\mathrm{i}} g_{\mathrm{i}} \left(w_{\mathrm{i}} \cdot x_{\mathrm{j}} + B_{\mathrm{i}} \right) = z_{\mathrm{j}},\ j = 1, 2, \ldots.N \tag{8}$$

where $w_{\mathrm{i}} = \left[w_{\mathrm{i1}}, w_{\mathrm{i2}}, w_{\mathrm{i3}}, \ldots., w_{\mathrm{in}} \right]^{\mathrm{T}}$ is the weight vector connecting the ith hidden node and the input nodes, $\beta_{\mathrm{i}} = \left[\beta_{\mathrm{i1}}, \beta_{\mathrm{i2}}, \beta_{\mathrm{i3}}, \ldots., \beta_{\mathrm{im}} \right]^{\mathrm{T}}$ is the weight vector connecting the ith hidden node and the output nodes, B_i is the threshold of the ith hidden node and $g_i\left(w_i \cdot x_j + B_i\right)$ is hidden layer output function, and $w_i \cdot x_j$ denotes the inner product of w_i, x_j, and z_j is the ELM model output.

The above equation can be reformulated in the following matrix form:

$$H\beta = Z \tag{9}$$

$$H\left(w_1,\ldots,w_\ell,B_1,\ldots,B_\ell,x_1,\ldots,x_\ell\right) = \begin{bmatrix} g\left(w_1 \cdot x_1 + B_1\right) & \ldots & g\left(w_\ell \cdot x_1 + B_\ell\right) \\ \ldots & \ldots & \ldots \\ g\left(w_1 \cdot x_N + B_1\right) & \ldots & g\left(w_\ell \cdot x_N + B_\ell\right) \end{bmatrix}_{N\times\ell} \tag{10}$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \ldots \\ \beta_\ell^T \end{bmatrix}_{\ell\times m} \tag{11}$$

$$Z = \begin{bmatrix} Z_1^T \\ \ldots \\ Z_N^T \end{bmatrix}_{N\times m} \tag{12}$$

2.1.4. Extreme Learning Machine (ELM)

The XGBoost model presented by Chen [36] is an innovative implementation method with gradient boosting machines, specifically K classification and regression trees (CART). The technique was inspired by the concept of "boost," which aggregates all the predictions of a group of "weak" learners in order to build a "strong" learner using progressive training methods. The XGBoost model is designed to avoid overfitting while simultaneously minimizing computational expenses. This is accomplished by simplifying the goal functions, which allow for the combination of forecasting and regularization components while preserving appropriate computational speed. During the training period, simultaneous calculations are also performed for the functionalities in the XGBoost model. The general prediction function of the XGBoost model at step t can be expressed as follows [37].

$$f^{(t)} = f^{(t-1)} + f_t\left(x_i\right) \tag{13}$$

where $f^{(t)}$ is predictions at steps t, $f^{(t-1)}$ is the predictions at steps t–1, $f_t\left(x_i\right)$ is the learner at step t, and x_i is the input variable.

The analytic expression is used in the XGBoost model to prevent over-fitting issues without compromising the arithmetic speed of the model, which is presented as follows:

$$obj^{(t)} = \sum_{k=1}^{n} l\left(\overline{y_i}, y_i\right) + \sum_{k=1}^{t} \Omega\left(f_i\right) \tag{14}$$

where l is the loss function, n is the used observations number, and Ω is the regularization term, which is formulated as:

$$\Omega\left(f\right) = \mu T + \frac{1}{2}\delta\left\|\omega\right\|^2 \tag{15}$$

where μ is the minimum loss, and δ is the regularization parameter

2.2. Statistical Metrics and Performance Estimation

The efficacy of machine learning models to forecast the hourly load was evaluated in the current study using a few statistical criteria. It is recommended that this study make use of the R2, MAE, and RMSE. The best value of R2, which indicates good agreement, is 1, which quantifies the degree of agreement between the observation and prediction data. However, the systematic error was measured using the MAE and RMSE, and 0 is the ideal value. The calculation of the above-mentioned indices is as follows [38]–[41].

$$R^2 = \left[\frac{\sum_{i=1}^{n} (X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{n} (X_i - \overline{X})^2 \sum_{i=1}^{n} (Y_i - \overline{Y})^2}} \right]^2 \tag{16}$$

$$MAE = \frac{\sum_{i=1}^{n} |X_i - Y_i|}{n} \tag{17}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (X_i - Y_i)^2} \tag{18}$$

where X_i and Y_i represent the observed and predicted values at i hour, while $\overline{X}$ and $\overline{Y}$ represent the observed and predicted mean values, respectively.

2.3. Load Data Used in this Study

Residential households' electricity consumption fluctuates throughout the course of a given day in rather predictable ways [42]–[47]. Electricity is used in almost every household in the United States. When most people are sleeping at night, the least amount of electricity is used. Every day, around 5:00 am, the least amount of power is often requested, and depending on the season, the peak demand happens at some point during the day before declining in the late evening. Because homes utilize air conditioning on hot days, the US residential sector uses a lot of electricity during summer, when the afternoon load rises. Although the amount of energy used during the winter months is less variable than it is during the summer, there are still morning and evening peaks due to the need for space heating and water heating [48]. In the United States, residential users use between 29.7 and 40.1 kWh per day and between 893 and 1200 kWh per month [49]. The Northwest Energy Efficiency Alliance (NEEA) provided the load profile data used in this investigation [50]. This dataset includes hourly resolution power demand for a year, covering energy end users in 101 Pacific Northwest-US residences [51].

2.4. Data Processing

Processing of data used: to determine the appropriate lag time for the inputs, this investigation applied a mathematical approach proposed by Sudheer et al. [52]–[54]. The approach relies on the hypothesis that potential contributing variables with varying time lags could be found through data series interpretation and statistical analysis, including ACF and CCF. Using ACF and CCF, this study established the best lag time of the inputs. The summer and winter seasons' maximum three-lag times for energy consumption were established for ACF > 0.6 and ACF > 0.4, respectively. However, the maximum four-lag time of mean temperature in the summer and winter was found to be CCF > 0.6 and CCF > 0.4, respectively. During the applications, the models were trained using 80% of the original data and tested using 20% of the end data.

3. Results and Discussion

Scatter plots of predicted and actual load demand for a variety of ML models, such as RF, ELM, SVM, and XGBoost, are shown in Fig. 2 and Fig. 3, respectively, for the winter and summer seasons. The forecast accuracy of each model under both seasonal settings is shown graphically in these figures.

A number of statistical measures, including the R^2, RMSE, and MAE, were utilized to evaluate the degree of agreement and the magnitude of error between the projected and actual load demand in order to quantify the accuracy of these predictions. The R^2 values, which show the percentage of variance in the observed data that can be accounted for by the model's predictions, are especially instructive.

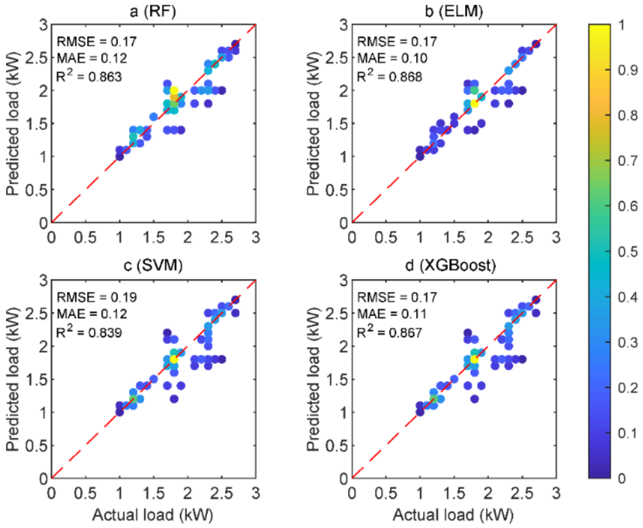

Fig. 2. Hourly load evaluations from various ML models throughout the testing phase over the winter are displayed in scatter plots

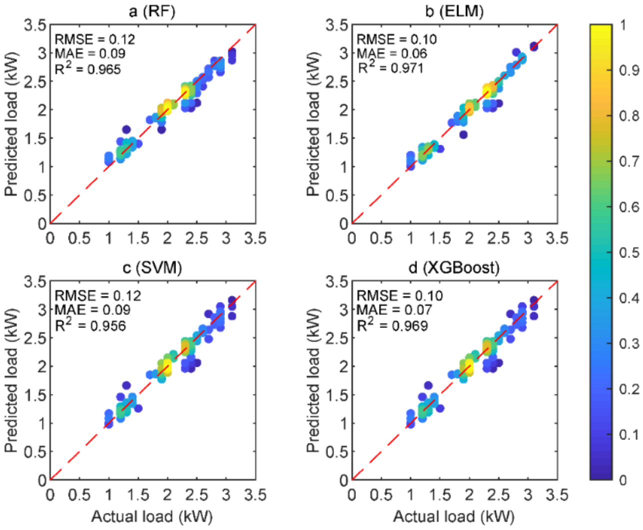

Fig. 3. Hourly load evaluations from various ML models throughout the testing phase over the summer are displayed in scatter plots

Furthermore, according to several ML algorithms, Fig. 4 and Fig. 5 compare the actual and anticipated electricity consumption for the winter and summer. The findings show that the ELM algorithms outperform others throughout winter and summer in capturing the peak load demand. All things considered, the ELM model shows great potential and capability for simulating electricity use.

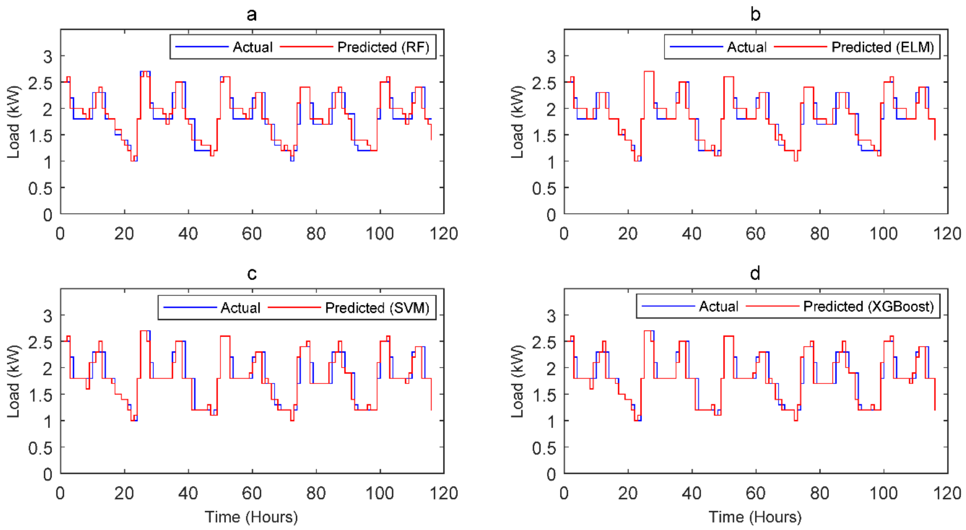

Fig. 4. Comparison of the forecasted and real electricity consumption over the winter months using several ML algorithms

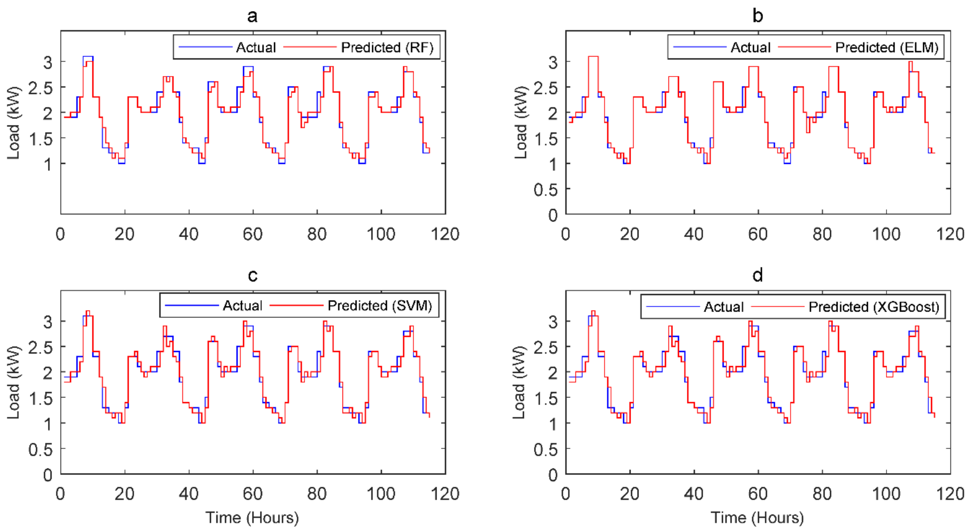

Fig. 5. Comparison of the forecasted and real electricity consumption over the summer months using several ML algorithms

3.1. Winter Season Analysis

During the winter season (Fig. 2), the R^2 values for each model show in Table 1.

Table 1. Winter Season Performance Metrics

Model	R^2	MAE	RMSE
RF	0.863	0.12	0.17
XGBoost	0.867	0.11	0.17
SVM	0.839	0.12	0.19
ELM	0.868	0.10	0.17

As shown in the Table 1, ELM yields the highest R^2 value (0.868) in the winter season, closely followed by XGBoost (0.867). This indicates that both models offer excellent predictive accuracy, but ELM slightly outperforms the others in capturing the variation in load demand. Furthermore, the ELM model demonstrates the smallest MAE (0.10) and RMSE (0.17), suggesting that it consistently delivers more accurate predictions with lower errors compared to the other models. The RF model, while still effective, shows slightly higher values for MAE (0.12) and RMSE (0.17), indicating it is less precise than ELM.

3.2. Summer Season Analysis

During the summer season (Fig. 3), the R^2 values for the models improve significantly, as show in Table 2.

Table 2. Summer Season Performance Metrics

Model	R^2	MAE	RMSE
RF	0.965	0.09	0.12
XGBoost	0.969	0.07	0.10
SVM	0.956	0.09	0.12
ELM	0.971	0.06	0.10

In the summer season, the ELM model achieves the highest R^2 value (0.971), followed closely by XGBoost (0.969). These results suggest that the predictive accuracy of all models improves during the summer, likely due to more consistent patterns in electricity consumption. With the lowest MAE (0.06) and RMSE (0.10), ELM once again performs exceptionally well, demonstrating its capacity to produce incredibly precise and dependable forecasts. Even while it is still useful, the RF model's greatest error scores for both MAE (0.09) and RMSE (0.12) demonstrate how limited it is in comparison to ELM in terms of forecasting load demand in the summer.

3.3. Discussion

The ELM model's superior R^2 values and the lowest MAE and RMSE show that it outperforms the other ML algorithms (RF, SVM, and XGBoost) in both the winter and summer months. This indicates that the ELM model is especially helpful for precise load forecasting in home energy systems since it is very good at identifying the underlying patterns in load demand and reducing prediction errors. Since temperature and other climatic conditions usually lead to higher and more consistent demand, the improvement in predictive performance during the summer months across all models is probably the result of more predictable patterns of power consumption. On the other hand, winter load demands could be more variable due to variables like heating needs, which are harder to forecast precisely.

The RF model continually performs worse than ELM and XGBoost in terms of error metrics (MAE and RMSE), despite its encouraging results. This implies that although RF is able to identify certain trends in the data, it is not as good at simulating load demand as the other models, especially in the winter and summer. Despite being a robust model, SVM routinely outperforms XGBoost and ELM, particularly in the winter, with a slightly lower R2 and greater MAE and RMSE. These results imply that the ELM model is very appropriate for predicting residential load demand in terms of real-world implementation. It is a great option for energy generation and purchase decision-making processes because of its accuracy in modeling peak load demand. Further increasing its usefulness in energy system planning and operation is the fact that it performs well in both the winter and the summer, suggesting that it can be dependably utilized all year round.

Lastly, this study shows how crucial it is to choose the right machine learning model for load forecasting, with ELM turning out to be the best option for both the winter and summer. To further enhance the accuracy of load demand projections, research should investigate the incorporation of more intricate elements, including time-series data or outside variables.

4. Conclusion

This study evaluated different ML model performances, such as RF, SVM, XGBoost, and ELM, to predict the load demand of the residential sector. To train and test the models, hourly temperature and electricity consumption data were used. The ACF and CCF were used in the preprocessing steps to determine the proper fall timings. MAE, RMSE, and R^2 were the evaluation criteria that showed the ELM model consistently beat other models. ELM had the lowest MAE (0.10) and RMSE (0.17) during the winter, with an R^2 of 0.868, suggesting better accuracy and less prediction error. Furthermore, demonstrating its resilience to seasonal fluctuations, ELM achieved the maximum R^2 value of 0.971 in the summer, with the minimum MAE (0.06) and RMSE (0.10). Among the different ML algorithms examined, the ELM algorithm proved to be the most dependable due to its exceptional ability to capture patterns of low-variation electricity usage and peak load demand. Although RF and SVM models had significantly lower accuracy and greater error metrics, XGBoost also showed outstanding prediction skills, coming in just below ELM. While winter forecasts presented additional difficulties because of increased demand fluctuation, the seasonal study showed that predictive accuracy was better during the summer months due to more stable patterns of electricity usage. This study emphasizes how precise load forecasting for residential systems has real-world applications in resource optimization, energy generation planning, and purchasing decisions. Since ELM is the best model for forecasting household electricity load demand, this study provides important information about how to choose appropriate machine learning models for energy management systems. Future research could investigate the incorporation of other elements, such as behavioral patterns, socioeconomic data, and time-series features, to improve machine learning models' prediction capabilities. Furthermore, combining real-time forecasting skills with hybrid machine learning techniques may enhance flexibility in response to changing patterns in energy usage, facilitating the creation of intelligent and more effective energy systems.

Declarations

Author contribution. All authors contributed equally to the main contributor to this paper. All authors read and approved the final paper.

Funding statement. None of the authors have received any funding or grants from any institution or funding body for the research.

Conflict of interest. The authors declare no conflict of interest.

Additional information. No additional information is available for this paper

References

[1] M. Wang and M. A. A. Abdalla, "Optimal Energy Scheduling Based on Jaya Algorithm for Integration of Vehicle-to-Home and Energy Storage System with Photovoltaic Generation in Smart Home," *Sensors*, vol. 22, no. 4, p. 1306, Feb. 2022, doi: 10.3390/s22041306.

[2] M. A. A. Abdalla, W. Min, A. H. Gomaa Haroun, and M. Elhindi, “Optimal Energy Scheduling Strategy for Smart Charging of Electric Vehicles from Grid-Connected Photovoltaic System,” in *2021 7th International Conference on Electrical, Electronics and Information Engineering (ICEEIE)*, Oct. 2021, pp. 37–42, doi: 10.1109/ICEEIE52663.2021.9616634.

[3] A. Rehman, A. Rauf, M. Ahmad, A. A. Chandio, and Z. Deyuan, “The effect of carbon dioxide emission and the consumption of electrical energy, fossil fuel energy, and renewable energy, on economic performance: evidence from Pakistan,” *Environ. Sci. Pollut. Res.*, vol. 26, no. 21, pp. 21760–21773, Jul. 2019, doi: 10.1007/s11356-019-05550-y.

[4] A. Hossain Bagdadee, Z. Li, and M. Adam Ali Abdalla, “Constant &Reliable Power Supply by the Smart Grid Technology in Modern Power System,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 561, no. 1, p. 012088, Oct. 2019, doi: 10.1088/1757-899X/561/1/012088.

[5] M. Elhindi, Z. Chen, H. Ma, X. Sun, and M. A. Ali Abdalla, “Characteristic analysis of demagnetization fault in FSCW-PMM with four-layer windings,” *J. Magn. Magn. Mater.*, vol. 623, p. 172990, Jul. 2025, doi: 10.1016/j.jmmm.2025.172990.

[6] M. Shao, X. Wang, Z. Bu, X. Chen, and Y. Wang, “Prediction of energy consumption in hotel buildings via support vector machines,” *Sustain. Cities Soc.*, vol. 57, p. 102128, Jun. 2020, doi: 10.1016/j.scs.2020.102128.

[7] O. Abbaker Ahmed Mohammed *et al.*, “Optimal α-Variable Model-Free Adaptive Barrier Function Fractional Order Nonlinear Sliding Mode Control for Four Area Interconnected Hybrid Power System With Nonlinearities,” *IEEE Access*, vol. 12, pp. 61586–61609, 2024, doi: 10.1109/ACCESS.2024.3391611.

[8] M. Elhindi, M. A. A. Abdalla, A. Omar, A. Pranolo, A. Mirghani, and A. A. Omer, “Comparative Electromagnetic Performance Analysis of Double Stator and Single Stator Superconducting Generators for Direct-Drive Wind Turbines,” *Int. J. Robot. Control Syst.*, vol. 4, no. 2, pp. 678–690, May 2024, doi: 10.31763/ijrcs.v4i2.1385.

[9] M. A. A. Abdalla *et al.*, “Optimizing Energy Usage and Smoothing Load Profile via a Home Energy Management Strategy with Vehicle-to-Home and Energy Storage System,” *Sustainability*, vol. 15, no. 20, p. 15046, Oct. 2023, doi: 10.3390/su152015046.

[10] M. Wang, D. Pan, M. A. A. Abdalla, H. Wang, F. Zuo, and X. Sheng, “A Review of Robust State Estimation of Distribution Network,” in *2023 2nd Asian Conference on Frontiers of Power and Energy (ACFPE)*, Oct. 2023, pp. 475–481, doi: 10.1109/ACFPE59335.2023.10455394.

[11] O. A. A. Mohammed, L. Peng, G. H. A. Hamid, A. M. Ishag, and M. A. A. Abdalla, “Effective Energy Management Strategy with Model-Free DC-Bus Voltage Control for Fuel Cell/Battery/Supercapacitor Hybrid Electric Vehicle System,” *Machines*, vol. 11, no. 10, p. 944, Oct. 2023, doi: 10.3390/machines11100944.

[12] M. Abdallah, B. Mohammadi, H. Nasiri, O. M. Katipoğlu, M. A. A. Abdalla, and M. M. Ebadzadeh, “Daily global solar radiation time series prediction using variational mode decomposition combined with multi-functional recurrent fuzzy neural network and quantile regression forests algorithm,” *Energy Reports*, vol. 10, pp. 4198–4217, Nov. 2023, doi: 10.1016/j.egyr.2023.10.070.

[13] F. Kurniawan, S. Sulaiman, S. Konate, and M. A. A. Abdalla, “Deep learning approaches for MIMO time-series analysis,” *Int. J. Adv. Intell. Informatics*, vol. 9, no. 2, p. 286, Jul. 2023, doi: 10.26555/ijain.v9i2.1092.

[14] M. A. A. Abdalla, W. Min, W. Bing, A. M. Ishag, and B. Saleh, “Double-layer home energy management strategy for increasing PV self-consumption and cost reduction through appliances scheduling, EV, and storage,” *Energy Reports*, vol. 10, pp. 3494–3518, Nov. 2023, doi: 10.1016/j.egyr.2023.10.019.

[15] S. Khan, A. Abdulsatar, and S. Alsamarai, “Social media and its adverse effect on academic performance of students,” *Bull. Soc. Informatics Theory Appl.*, vol. 3, no. 1, pp. 38–44, May 2019, doi: 10.31763/businta.v3i1.163.

[16] G. R. Sosa, M. Z. Falah, D. F. L, A. P. Wibawa, A. N. Handayani, and J. A. H. Hammad, "Forecasting electrical power consumption using ARIMA method based on kWh of sold energy," *Sci. Inf. Technol. Lett.*, vol. 2, no. 1, pp. 9–15, May 2021, doi: 10.31763/sitech.v2i1.637.

[17] J. Zhu, H. Dong, W. Zheng, S. Li, Y. Huang, and L. Xi, "Review and prospect of data-driven techniques for load forecasting in integrated energy systems," *Appl. Energy*, vol. 321, p. 119269, Sep. 2022, doi: 10.1016/j.apenergy.2022.119269.

[18] D. K. Ranaweera, G. G. Karady, and R. G. Farmer, "Economic impact analysis of load forecasting," *IEEE Trans. Power Syst.*, vol. 12, no. 3, pp. 1388–1392, 1997, doi: 10.1109/59.630486.

[19] G. Newsom and E. G. Brown, "Improving Solar and Load Forecasts by Reducing Operational Uncertainty," California, pp. 1-119, 2019. [Online]. Available at: https://www.energy.ca.gov/sites/default/files/2021-06/CEC-500-2019-023.pdf.

[20] Z. Wang, Y. Wang, R. Zeng, R. S. Srinivasan, and S. Ahrentzen, "Random Forest based hourly building energy prediction," *Energy Build.*, vol. 171, pp. 11–25, Jul. 2018, doi: 10.1016/j.enbuild.2018.04.008.

[21] R. A. Abbasi, N. Javaid, M. N. J. Ghuman, Z. A. Khan, S. Ur Rehman, and Amanullah, "Short Term Load Forecasting Using XGBoost," in *Advances in Intelligent Systems and Computing*, vol. 927, Springer, Cham, 2019, pp. 1120–1131, doi: 10.1007/978-3-030-15035-8_108.

[22] S. Li, L. Goel, and P. Wang, "An ensemble approach for short-term load forecasting by extreme learning machine," *Appl. Energy*, vol. 170, pp. 22–29, May 2016, doi: 10.1016/j.apenergy.2016.02.114.

[23] T. Khatib, A. Mohamed, and K. Sopian, "A review of photovoltaic systems size optimization techniques," *Renew. Sustain. Energy Rev.*, vol. 22, pp. 454–465, Jun. 2013, doi: 10.1016/j.rser.2013.02.023.

[24] K. Haruna, A. S. Uju, I. A. Lawal, and R. Abubakar, "Placement model for students into appropriate academic class using machine learning," *Sci. Inf. Technol. Lett.*, vol. 4, no. 1, pp. 1–11, May 2023, doi: 10.31763/sitech.v4i1.1024.

[25] A. Y. Saleh and F. N. Binti Mostapa, "Suicide and self-harm prediction based on social media data using machine learning algorithms," *Sci. Inf. Technol. Lett.*, vol. 4, no. 1, pp. 12–21, May 2023, doi: 10.31763/sitech.v4i1.1181.

[26] M. Klaiber, "A Fundamental Overview of SOTA-Ensemble Learning Methods for Deep Learning: A Systematic Literature Review," *Sci. Inf. Technol. Lett.*, vol. 2, no. 2, pp. 1–14, Dec. 2021, doi: 10.31763/sitech.v2i2.549.

[27] D. P. Ismi and N. Khoirunnisa, "Enhancing the performance of heart arrhythmia prediction model using Convolutional Neural Network based architectures," *Sci. Inf. Technol. Lett.*, vol. 5, no. 2, pp. 1–12, Nov. 2024, doi: 10.31763/sitech.v5i2.1794.

[28] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, Sep. 1995, doi: 10.1007/BF00994018.

[29] B. Mohammadi and S. Mehdizadeh, "Modeling daily reference evapotranspiration via a novel approach based on support vector regression coupled with whale optimization algorithm," *Agric. Water Manag.*, vol. 237, p. 106145, Jul. 2020, doi: 10.1016/j.agwat.2020.106145.

[30] V. H. Quej, J. Almorox, J. A. Arnaldo, and L. Saito, "ANFIS, SVM and ANN soft-computing techniques to estimate daily global solar radiation in a warm sub-humid environment," *J. Atmos. Solar-Terrestrial Phys.*, vol. 155, pp. 62–70, Mar. 2017, doi: 10.1016/j.jastp.2017.02.002.

[31] L. Breiman, "Random Forests," *Mach. Learn. 2001 451*, vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: 10.1023/A:1010933404324.

[32] A. N. S. Kinasih, A. N. Handayani, J. T. Ardiansah, and N. S. Damanhuri, "Comparative analysis of decision tree and random forest classifiers for structured data classification in machine learning," *Sci. Inf. Technol. Lett.*, vol. 5, no. 2, pp. 13–24, Nov. 2024, doi: 10.31763/sitech.v5i2.1746.

[33] G. Huang *et al.*, "Evaluation of CatBoost method for prediction of reference evapotranspiration in humid regions," *J. Hydrol.*, vol. 574, pp. 1029–1041, Jul. 2019, doi: 10.1016/j.jhydrol.2019.04.085.

[34] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, no. 1–3, pp. 489–501, Dec. 2006, doi: 10.1016/j.neucom.2005.12.126.

[35] R. Ahmed, V. Sreeram, Y. Mishra, and M. D. Arif, "A review and evaluation of the state-of-the-art in PV solar power forecasting: Techniques and optimization," *Renew. Sustain. Energy Rev.*, vol. 124, p. 109792, May 2020, doi: 10.1016/j.rser.2020.109792.

[36] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," *Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, vol. 13-17-August-2016, pp. 785–794, Aug. 2016, doi: 10.1145/2939672.2939785.

[37] J. Fan *et al.*, "Comparison of Support Vector Machine and Extreme Gradient Boosting for predicting daily global solar radiation using temperature and precipitation in humid subtropical climates: A case study in China," *Energy Convers. Manag.*, vol. 164, pp. 102–111, May 2018, doi: 10.1016/j.enconman.2018.02.087.

[38] S. H. Rafi, Nahid-Al-Masood, S. R. Deeba, and E. Hossain, "A Short-Term Load Forecasting Method Using Integrated CNN and LSTM Network," *IEEE Access*, vol. 9, pp. 32436–32448, 2021, doi: 10.1109/ACCESS.2021.3060654.

[39] R. Munif and A. Prahara, "Betta fish classification using transfer learning and fine-tuning of CNN models," *Sci. Inf. Technol. Lett.*, vol. 5, no. 1, pp. 16–25, Apr. 2024, doi: 10.31763/sitech.v5i1.1378.

[40] H. Jayadianti, B. A. Arianti, N. H. Cahyana, S. Saifullah, and R. Dreżewski, "Improving sentiment analysis on PeduliLindungi comments: a comparative study with CNN-Word2Vec and integrated negation handling," *Sci. Inf. Technol. Lett.*, vol. 4, no. 2, pp. 75–89, Nov. 2023, doi: 10.31763/sitech.v4i2.1184.

[41] A. H. Pratomo, N. H. Cahyana, and S. N. Indrawati, "Optimizing CNN hyperparameters with genetic algorithms for face mask usage classification," *Sci. Inf. Technol. Lett.*, vol. 4, no. 1, pp. 54–64, May 2023, doi: 10.31763/sitech.v4i1.1182.

[42] M. A. A. Abdalla, W. Min, and O. A. A. Mohammed, "Two-Stage Energy Management Strategy of EV and PV Integrated Smart Home to Minimize Electricity Cost and Flatten Power Load Profile," *Energies*, vol. 13, no. 23, p. 6387, Dec. 2020, doi: 10.3390/en13236387.

[43] M. Elhindi, M. A. A. Abdalla, and D. Liu, "Investigating the impact of magnetic air gap variations on short circuit characteristics of partially high temperature superconducting generator," *Phys. C Supercond. its Appl.*, vol. 623, p. 1354550, Aug. 2024, doi: 10.1016/j.physc.2024.1354550.

[44] N. Ibrahim, N. Dong, and M. A. A. Abdalla, "On dual-loop model-free adaptive iterative learning control and its application," *J. Process Control*, vol. 147, p. 103376, Mar. 2025, doi: 10.1016/j.jprocont.2025.103376.

[45] N. Ibrahim and N. Dong, "A Novel Dual-Loop Model-Free Adaptive Iterative Learning Control and Its Application to the Refrigeration Systems," *Int. J. Robust Nonlinear Control*, vol. 35, no. 6, pp. 2213–2234, Apr. 2025, doi: 10.1002/rnc.7790.

[46] M. Adam, A. Abdalla, G. Kalcon, and A. Mohammed, "Design of Travelling Wave Relay for Protection of Transmission Lines," *Int. J. Res. Stud. Electr. Electron. Eng.*, vol. 3, no. 4, pp. 29–41, 2017, doi: 10.20431/2454-9436.0304005.

[47] A. Fithriyyah, T. Purwaningsih, S. Konate, and M. A. A. Abdalla, "Mapping dengue vulnerability: spatial cluster analysis reveals patterns in Central Java, Indonesia," *Sci. Inf. Technol. Lett.*, vol. 4, no. 2, pp. 109–122, Nov. 2023, doi: 10.31763/sitech.v4i2.1203.

[48] "Hourly electricity consumption varies throughout the day and across seasons," *U.S. Energy Information Administration (EIA)*, 2020. [Online]. Available at: https://www.eia.gov/todayinenergy/detail.php?id=42915.

[49] E. Jahani, K. Cetin, and I. H. Cho, "City-scale single family residential building energy consumption prediction using genetic algorithm-based Numerical Moment Matching technique," *Build. Environ.*, vol. 172, p. 106667, Apr. 2020, doi: 10.1016/j.buildenv.2020.106667.

[50] B. Larson *et al.*, "Residential building stock assessment: Metering study," 2014. [Online]. Available at: http://neea.org/docs/default-source/reports/residential-building-stock-assessment--metering-study.pdf?sfvrsn=6.

[51] N. Mims Frick *et al.*, "End-Use Load Profiles for the U.S. Building Stock: Technical Advisory Group Meeting #11." p. 152, Aug. 02, 2022. [Online]. Available at: https://www.osti.gov/biblio/1880324.

[52] K. P. Sudheer, A. K. Gosain, and K. S. Ramasastri, "A data-driven algorithm for constructing artificial neural network rainfall-runoff models," *Hydrol. Process.*, vol. 16, no. 6, pp. 1325–1330, Apr. 2002, doi: 10.1002/hyp.554.

[53] A. Z. Z. Abidin, M. F. I. Othman, A. Hassan, Y. Murdianingsih, U. T. Suryadi, and T. F. Siallagan, "A comparison of machine learning methods for knowledge extraction model in A LoRa-Based waste bin monitoring system," *Int. J. Adv. Intell. Informatics*, vol. 10, no. 1, p. 79, Feb. 2024, doi: 10.26555/ijain.v10i1.1026.

[54] M. A. As'ari, N. A. J. Sufri, and G. S. Qi, "Emergency sign language recognition from variant of convolutional neural network (CNN) and long short term memory (LSTM) models," *Int. J. Adv. Intell. Informatics*, vol. 10, no. 1, p. 64, Feb. 2024, doi: 10.26555/ijain.v10i1.1170.