International Journal of Power Electronics and Drive Systems (IJPEDS)
Vol. 15, No. 4, December2024, pp. 2138~2144
ISSN: 2088-8694, DOI: 10.11591/ijpeds.v15.i4.pp2138-2144 ❒ 2138

A comprehensive evaluation of machine learning algorithms for precise energy consumption forecasting in smart homes

Lakshmana Phaneendra Maguluri[1], M. Shankar[2], R. Aruna[3], D. Chitra Devi[4], M. J. Suganya[5]
[1]Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, India
[2]Department of Computer Science & Technology, Madanapalle Institute of Technology & Science, Madanapalle, India
[3]Department of Electronics and Communication Engineering, AMC Engineering College, Bengaluru, India
[4]Department of Computer Science and Engineering, S.A. Engineering College, Chennai, India
[5]Department of Electrical and Electronics Engineering, Panimalar Engineering College, Chennai, India

Article Info

Article history:

Received Feb 7, 2024
Revised Aug 31, 2024
Accepted Sep 6, 2024

Keywords:

Load forecasting
Machine learning algorithms
Performance metrics
Smart grid
Smart home

ABSTRACT

Energy is one of the most critical and costly resources, playing a vital role in our daily lives. As technology advances, the demand for energy also increases. This work proposes a model for predicting energy consumption in smart homes, consisting of data preprocessing, performance evaluation, and application. Once the data is processed, it is fed into the prediction module, where various machine-learning algorithms are applied to forecast energy consumption. As smart home environments grow in complexity, selecting the most effective machine learning algorithm becomes increasingly crucial. The persistent challenge lies in manually discerning the best-performing algorithm, given their potential variance in efficacy across diverse use cases or datasets. In the dynamic landscape of energy conservation and cost-effective power generation, precise forecasting of energy consumption is essential, playing a pivotal role in advancing energy sustainability and bolstering economic stability. This introduction explores the intricate terrain of predicting energy utilization within smart homes, a domain that has seen increased interest due to the integration of machine learning algorithms. The primary focus of this exploration is the rigorous evaluation of these algorithms, using key performance metrics such as mean absolute error (MAE), mean squared error (MSE), and R-squared.

Corresponding Author:

M. J. Suganya
Department of Electrical and Electronics Engineering, Panimalar Engineering College
Varadharajapuram, Poonamallee, Chennai 600123, India
Email: sugi.mj@gmail.com

1. INTRODUCTION

Smartgrids (SG) have emerged as a viable solution to meet the growing global energy demand. The term "grid" refers to the traditional electrical infrastructure comprising transmission lines, substations, and other elements that facilitate the delivery of energy from power plants to homes and businesses [1], [2]. What sets Smart Grids apart is their capability for two-way communication between utility providers and consumers, coupled with sensing capabilities along the grid lines. Key components of a smart grid include controls, computers, automation systems, and other advanced technologies working in tandem to address the rapid surge in energy requirements shown in Figure 1. The intelligence embedded in smart grids brings forth various benefits. Noteworthy advantages include more efficient energy transmission, enhanced security measures, and the ability to mitigate peak demand, consequently leading to a reduction in electricity rates. Smart grids are also recognized for their integration of renewable energy sources, aligning with the global push towards sustainable and eco-friendly power generation. Overall, smart grids represent a significant advancement in the energy sector, ensuring a more responsive and adaptable infrastructure to meet the evolving needs of our energy-intensive world [3]-[5]. Efficient energy management systems (EMS) relies heavily on two core pillars: prediction and scheduling. These systems play a pivotal role in ensuring the optimal functioning of SG, tasked with managing power flow across SG components to minimize costs and enhance overall quality [6].

Prediction, particularly of energy consumption by various appliances, is fundamental to the SG concept. Energy consumption can be characterized as a nonlinear time series influenced by numerous complex factors [7]. As Smart Grids increasingly incorporate renewable energy sources, the accuracy of energy prediction methods has improved significantly. Consequently, precise prediction becomes a vital element in the strategic planning of the entire smart grid [8]-[10]. Various approaches are employed for energy consumption prediction, with machine learning (ML) emerging as the most popular. ML techniques contribute to the refinement of energy forecasts, aligning with the evolving landscape of Smart Grids and reinforcing their efficiency in managing diverse energy sources [10]-[15].

Previous studies examining the integration of a smart grid to realize the concept of a smart city. The authors elaborate on energy-related policies essential for implementing a smart city through the smart grid. Additionally, they explore prevailing conceptualizations, or "imaginaries," associated with the smart city facilitated by the smart grid. These imaginaries encompass economic imperatives, environmental solutions, and the experimental challenges inherent in smart grid technology [16]-[20]. In a related discussion, other studies emphasize the pivotal roles of the internet of things (IoT) and smart grid in actualizing smart city initiatives. The authors underscore the significance of energy across various sectors while addressing diverse challenges in smart city development, such as heterogeneity, unplanned urban growth, and the adaptability of residents [21], [22]. Other researchers also state that the smart grid as the foundational element and backbone of the smart city. The author characterizes the smart grid as an amalgamation of the conventional power grid with information and communication technology, emphasizing its pivotal role as the anchor of the smart city [23]-[25].

The main contributions of this paper are: i) Conducting a literature review of previous research on energy consumption forecasting in smart homes, exploring their contributions and inferences; ii) Providing a detailed framework for energy consumption forecasting; iii) Analyzing various methodologies used in home energy consumption forecasting from multiple perspectives, discussing their findings and limitations; iv) The primary focus of this exploration is the rigorous evaluation of these algorithms, using key performance metrics such as mean absolute error (MAE), mean squared error (MSE), and R-squared; and v) This comprehensive approach aims to enhance understanding and advance the field of energy consumption forecasting in smart homes.

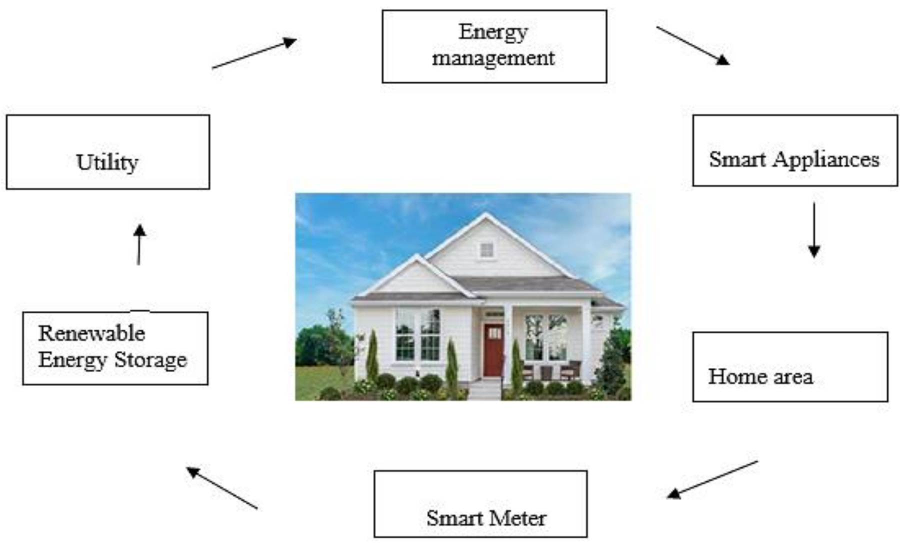

Figure 1. Home energy management system

2. ENERGY CONSUMPTION FORECASTING IN SMART HOMES

Several critical questions persist in the adoption of machine learning approaches for electricity usage forecasting. These include determining the optimal number of variables to be measured and recorded, specifying the total duration of continuous recording, and establishing the time resolution of the data. The outcomes presented in prior literature are contingent on the specific dataset employed for analysis. Table 1 compares four public datasets, emphasizing distinctions in their characteristics. Ideally, comprehensive model evaluation would be conducted across multiple datasets. However, it is common for results to be reported on a single dataset, with model performance assessed through various metrics such as MAE, MSE, and R-squared are presented as relative comparisons between different studied approaches shown in Figure 2.

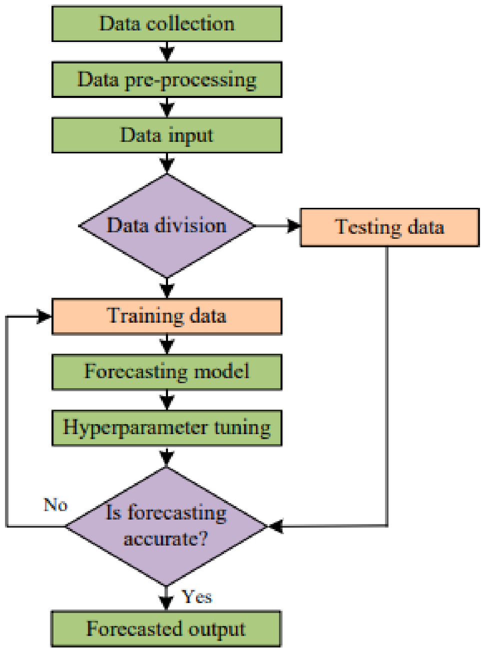

Figure 2. Process flow diagram for load forecasting

3. ERROR METRICS AND COMPUTATIONAL PERFORMANCE

Root mean squared error or RMSE and MAE serve as prevalent metrics for assessing model accuracy. However, these measures exhibit a scale-dependency, rendering the results incomparable across time series of varying magnitudes. To address this limitation, an accuracy metric known as the mean absolute scaled error (MSE) was introduced by [25]. MSE scales the error relative to a naive forecast, providing a more standardized evaluation that allows for meaningful comparisons between different time series datasets shown in (1) and (2).

$$\text{MSE}=\frac{1}{N}\sum_{i=0}^{N}(y_i - \widehat{y_i})^2 \tag{1}$$

Where MAE is the mean absolute error, a commonly used accuracy metric.

$$\text{MAE}=\frac{1}{N}\sum_{i=0}^{N}|y_i - \widehat{y}\text{i}| \tag{2}$$

- R-squared error

R-squared (R^2), commonly referred to as the coefficient of determination, measures the fraction of variability in the dependent variable that can be accounted for by the independent variables in a regression model. It serves as an assessment of how well the model conforms to the dataset. The R^2 is determined by dividing the explained variance by the total variance, providing insight into the model's effectiveness in capturing the data's patterns shown in (3).

$$\text{R2} = 1\text{-}\frac{\frac{1}{N}\sum_{i=0}^{N}(y_i - \widehat{y_i})^2}{\frac{1}{N}\sum_{i=0}^{N}(y_i - \overline{y_i})^2} \tag{3}$$

- N Number of observations

$yi - \ Pridicted \ values$

4. RESULTS AND DISCUSSION

The performance metrics bar graph provides a comprehensive visual representation of the evaluation results for three different regression models: linear regression, decision tree regressor, and random forest regressor. The graph encompasses MAE, MSE, and R-squared values, crucial indicators for assessing the models' predictive accuracy. In the subplot for each regression model, scatter plots depict the relationship between actual and predicted values, offering insight into the models' overall performance. Notably, the random forest regressor stands out for its effectiveness in capturing the underlying patterns, as reflected in the tight clustering of points around the diagonal. The fourth subplot consolidates the performance metrics in a bar graph format. Each model is represented by a distinctive color—blue for linear regression, orange for decision tree, and green for random forest. The bar graph reveals specific values for MAE, MSE, and R-squared, allowing for a direct comparison of the models' performance across these key metrics.

In conclusion, the graphical presentation effectively communicates the comparative performance of the regression models, aiding in the selection of the most suitable model based on the outlined metrics. Figure 3 shows the performance of the linear regression model, serving as a baseline but with limitations in capturing complex patterns. Figure 4 illustrates the decision tree regressor, which improves on Linear Regression by handling non-linear relationships but still exhibits higher MAE and MSE values. Figure 5 provides a comparison of all three models, while Figure 6 highlights the random forest regressor's superior performance with significantly lower MAE and MSE and a higher R-squared value. This demonstrates the effectiveness of Random Forest in improving predictive accuracy, as detailed in Table 1.

Table 1. Performance evaluation metrics

Sl.no	Model	MAE	MSE	R-squared
1	Linear regression	52.544733	8312.759514	0.169313
2	Decision tree regressor	39.447682	8486.242716	0.151977
3	Random forest regressor	32.851355	4704.308789	0.529902

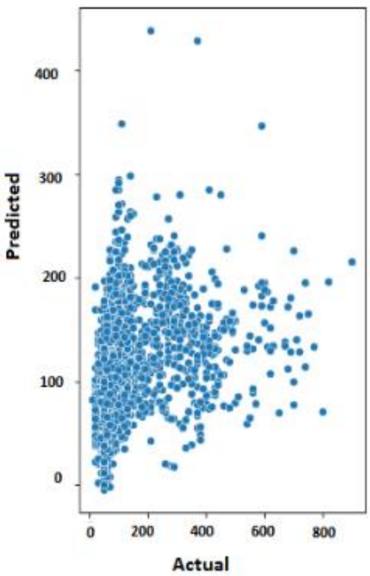

Figure 3. Linear regression

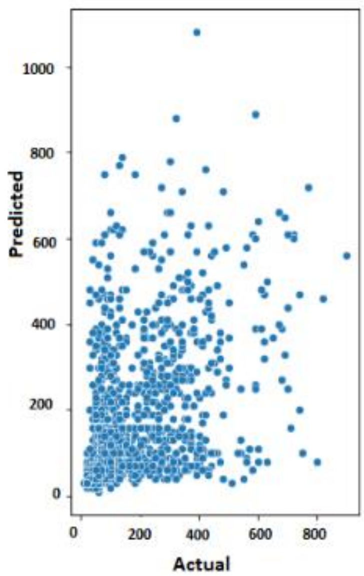

Figure 4. Decision tree regressor

Figure 7 showcases violin plots that depict the relative performance of the different forecasting methods. Each method is represented in a different color, with the thickness of the plot indicating the frequency of errors at a given value. The violin plots further emphasize the random forest regressor's superior performance, showing a more concentrated distribution of lower errors, thereby reinforcing its effectiveness in delivering more accurate and reliable predictions compared to both the linear regression and decision tree regressor models.

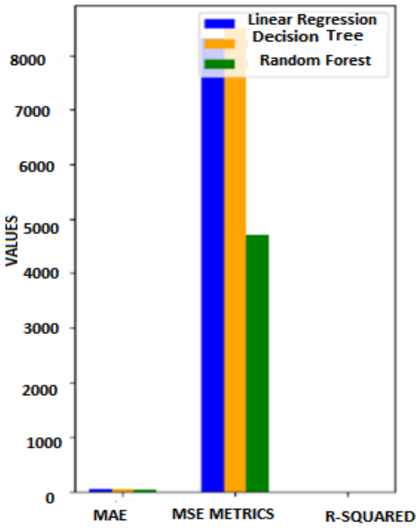

Figure 5. Performance metrics

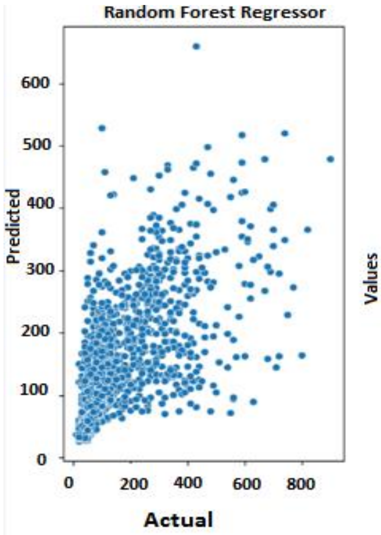

Figure 6. Random forest regressor

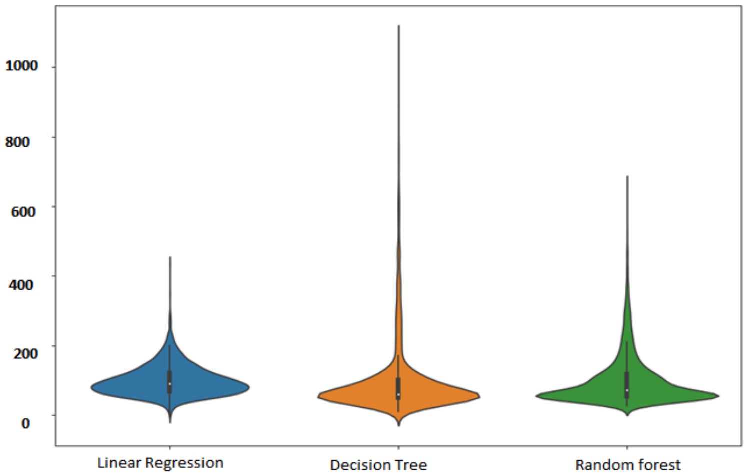

Figure 7. Violin plot of predicated values by different algorithms

5. CONCLUSION

The evaluation of three regression models linear regression, decision tree regressor, and random forest regressor reveals varying levels of predictive accuracy. The linear regression model presents a MAE of 52.54, a MSE of 8312.76, and an R-squared value of 0.17. This indicates a relatively modest performance with a limited ability to explain the variance in the data. The decision tree regressor shows some improvement, with an MAE of 39.45 and an R-squared of 0.15, although its MSE of 8486.24 suggests higher variability in its predictions compared to linear regression.

In contrast, the random forest regressor demonstrates superior performance across all metrics. It achieves an MAE of 32.85, significantly lower than the other models, and an MSE of 4704.31, indicating more precise predictions. Furthermore, its R-squared value of 0.53 highlights a substantially better fit to the data, explaining more than half of the variance. These results underscore the random forest model's ability to deliver more accurate and reliable predictions, making it the preferred choice for this regression task. Its ensemble approach leverages multiple decision trees, enhancing its robustness and predictive power over the other evaluated models.

REFERENCES

[1] S. N. N. Ibrahim, H. Zainuddin, Y. I. A. A. Rahim, N. A. Mansor, N. Muhammad, and F. L. M. Khir, "Simulation and prediction of residential grid-connected photovoltaic system performance," *International Journal of Power Electronics and Drive Systems*, vol. 14, no. 1, pp. 506–515, 2023, doi: 10.11591/ijpeds.v14.i1.pp506-515.

[2] A. Rafati, M. Joorabian, and E. Mashhour, "An efficient hour-ahead electrical load forecasting method based on innovative features," *Energy*, vol. 201, p. 117511, Jun. 2020, doi: 10.1016/j.energy.2020.117511.

[3] S. Debdas, S. Mohanty, S. Kundu, S. Mohanty, B. Biswas, and A. Pal, "Short-Term Load Forecasting Using Time Series Algorithm," in *Proceedings - International Conference on Artificial Intelligence and Smart Systems, ICAIS 2021*, 2021, pp. 1518–1524, doi: 10.1109/ICAIS50930.2021.9395948.

[4] S. Pasari and A. Shah, "Time Series Auto-Regressive Integrated Moving Average Model for Renewable Energy Forecasting," in *Sustainable Production, Life Cycle Engineering and Management*, 2020, pp. 71–77.

[5] B. Nørregaard, Z. Ma, and B. N. Jørgensen, "Syddansk Universitet Global Smart Grid Transmission : Comparison of Europe , the U . S ., and China Global Smart Grid Transmission : Comparison of Europe , the U . S ., and China," in *Proceedings of The 10th International Green Energy Conference*, 2015, no. May.

[6] S. Deepa, S. Praba, V. Deepalakshmi, L. Jayaprakash, and M. Manimurugan, "A Fuzzy Ga Based STATCOM for Power Quality Improvement," *International Journal of Power Electronics and Drive Systems (IJPEDS)*, vol. 8, no. 1, p. 483, Mar. 2017, doi: 10.11591/ijpeds.v8.i1.pp483-491.

[7] C. Zedak, A. Belfqih, J. Boukherouaa, and F. El Mariami, "An intelligent energy management system for optimum design and real-time operation," *International Journal of Power Electronics and Drive Systems*, vol. 14, no. 1, pp. 480–495, 2023, doi: 10.11591/ijpeds.v14.i1.pp480-495.

[8] Y. W. Lee, K. G. Tay, and Y. Y. Choy, "Forecasting Electricity Consumption Using Time Series Model," *International Journal of Engineering & Technology*, vol. 7, no. 4.30, p. 218, 2018, doi: 10.14419/ijet.v7i4.30.22124.

[9] S. Hosein and P. Hosein, "Load forecasting using deep neural networks," in 2017 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT), 2017, doi: 10.1109/ISGT.2017.8085971.

[10] F. Mahia, A. R. Dey, M. A. Masud, and M. S. Mahmud, "Forecasting electricity consumption using ARIMA model," 2019, doi: 10.1109/STI47673.2019.9068076.

[11] Z. Deng, C. Liu, and Z. Zhu, "Inter-hours rolling scheduling of behind-the-meter storage operating systems using electricity price forecasting based on deep convolutional neural network," *International Journal of Electrical Power and Energy Systems*, vol. 125, 2021, doi: 10.1016/j.ijepes.2020.106499.

[12] G. Sideratos, A. Ikonomopoulos, and N. D. Hatziargyriou, "A novel fuzzy-based ensemble model for load forecasting using hybrid deep neural networks," *Electric Power Systems Research*, vol. 178, p. 106025, Jan. 2020, doi: 10.1016/j.epsr.2019.106025.

[13] L. Li, C. J. Meinrenken, V. Modi, and P. J. Culligan, "Short-term apartment-level load forecasting using a modified neural network with selected auto-regressive features," *Applied Energy*, vol. 287, p. 116509, Apr. 2021, doi: 10.1016/j.apenergy.2021.116509.

[14] Y. Dong, X. Ma, and T. Fu, "Electrical load forecasting: A deep learning approach based on K-nearest neighbors," *Applied Soft Computing*, vol. 99, 2021, doi: 10.1016/j.asoc.2020.106900.

[15] M. Dehghani, Z. Montazeri, O. P. Malik, G. Dhiman, and V. Kumar, "BOSA: Binary orientation search algorithm," *International Journal of Innovative Technology and Exploring Engineering*, vol. 9, no. 1, pp. 5306–5310, 2019, doi: 10.35940/ijitee.A4215.119119.

[16] G. Dhiman, M. Garg, A. Nagar, V. Kumar, and M. Dehghani, "A novel algorithm for global optimization: Rat Swarm Optimizer," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 8, pp. 8457–8482, 2021, doi: 10.1007/s12652-020-02580-0.

[17] G. Dudek, P. Pelka, and S. Smyl, "A Hybrid Residual Dilated LSTM and Exponential Smoothing Model for Midterm Electric Load Forecasting," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 7, pp. 2879–2891, 2022, doi: 10.1109/TNNLS.2020.3046629.

[18] N. C. Batista, R. Melício, and V. M. F. Mendes, "Services enabler architecture for smart grid and smart living services providers under industry 4.0," *Energy and Buildings*, vol. 141, pp. 16–27, Apr. 2017, doi: 10.1016/j.enbuild.2017.02.039.

[19] A. R. Khan, A. Mahmood, A. Safdar, Z. A. Khan, and N. A. Khan, "Load forecasting, dynamic pricing and DSM in smart grid: A review," *Renewable and Sustainable Energy Reviews*, vol. 54, pp. 1311–1322, 2016, doi: 10.1016/j.rser.2015.10.117.

[20] I. Ozturk, "Energy Dependency and Energy Security: The Role of Energy Effıcıency and Renewable Energy Sources (The Mahbub Ul Haq Memorial Lecture)," *The Pakistan Development Review*, pp. 309–330, Dec. 2022, doi: 10.30541/v52i4Ipp.309-330.

[21] M. S. Hossain, "Cloud-supported cyber-physical localization framework for patients monitoring," *IEEE Systems Journal*, vol. 11, no. 1, pp. 118–127, 2017, doi: 10.1109/JSYST.2015.2470644.

[22] E. Jarmouni, A. Mouhsen, M. Lamhammedi, and H. Ouldzira, "Energy management system and supervision in a micro-grid using artificial neural network technique," *International Journal of Power Electronics and Drive Systems*, vol. 12, no. 4, pp. 2570–2579, 2021, doi: 10.11591/ijpeds.v12.i4.pp2570-2579.

[23] M. S. Bin Tamrin and M. R. Ahmad, "Simulation of adaptive power management circuit for hybrid energy harvester and real-time sensing application," *International Journal of Power Electronics and Drive Systems*, vol. 11, no. 2, pp. 658–666, 2020, doi: 10.11591/ijpeds.v11.i2.pp658-666.

[24] R. Mathumitha, P. Rathika, and K. Manimala, "Intelligent deep learning techniques for energy consumption forecasting in smart buildings: a review," *Artificial Intelligence Review*, vol. 57, no. 2, 2024, doi: 10.1007/s10462-023-10660-8.

[25] S. Zhao and S. Zhao, "Wind Power Interval Prediction via an Integrated Variational Empirical Decomposition Deep Learning Model," *Sustainability (Switzerland)*, vol. 15, no. 7, 2023, doi: 10.3390/su15076114.

BIOGRAPHIES OF AUTHORS

Lakshmana Phaneendra Maguluri is a highly accomplished professional with a diverse educational background and a profound expertise in Computer Science and Engineering. He gets Ph.D. in Computer Science and Engineering from Annamalai University, Chidambaram, established as solid academic foundation in Tamil Nadu, post-graduation in Information Technology from Gandhi Institute of Technology and Management and under graduation in Computer Science and Engineering from Seshadri Rao Gudlavalleru Engineering College JNTUK ratified faculty and served as assistant professor at Gudlavalleru Engineering College. Currently, he works as an associate professor at Koneru Lakshmaiah Educational Foundation. He can be contacted at email: phanendra51@gmail.com.

M. Shankar received his B.Tech. degree in Information Technology, M.E. degree in Computer Science and Engineering from Anna University, Chennai, India. He is currently working as assistant professor at Madanapalle Institute of Technology and Science, Madanapalle. He has 14 years of teaching experience and he is currently pursuing Ph.D. in Anna University. His research topics include machine learning and artificial intelligence. He can be contacted at email: shankarm@mits.ac.in.

R. Aruna is currently working as an associate professor in the ECE Department at AMC Engineering College, Banerghatta Road, Bengaluru. She graduated in Engineering at Annamalai University, Chidambaram, Tamil Nadu, India. She secured a master of engineering in ECE Department at College of Engineering, Guindy, Anna University, Chennai, Tamil Nadu, India. She secured a Ph.D. in the Electronics Department at Jain University, Bengaluru, Karnataka, India. She is in the field of signal processing domain at AMCEC, Bengaluru, Karnataka, India. She has been in the teaching profession for more than 26 years. She has presented 31 papers in national and international journals, conferences, and symposiums. Her main areas of interest include image, signal, video processing, and communication. She can be contacted at email: aruna.ramalingam@gmail.com.

D. Chitra Devi is an associate professor in the Department of Computer Science and Engineering, S.A. Engineering College, Chennai, India. She received her M.Sc., in Mathematics from Madras University, Chennai in 2000 and M.E. in Systems Engineering and Operations Research from College of Engineering, Chennai, India in 2009. She was awarded a Ph.D. in Computer Science and Engineering from the College of Engineering, Chennai, India in 2020. She has 19 years of teaching experience. Her area of research interests includes cloud computing. Green computing, data mining, machine learning, and big data analytics. She has published many research articles in reputed journals. She can be contacted at email: chitradanya@gmail.com.

M. J. Suganya is an assistant professor in the Electrical and Electronics Engineering Department at the Panimalar Engineering College, India. She received his B.E degree in Electrical and Electronics Engineering from Annamalai University and M.E degree in Power Electronics and Drives from the Government College of Engineering, Tirunelveli in 2008 and 2010, respectively. She has been an assistant professor in Panimalar Engineering College, India since 2017. Her research interests include the field of power electronics, motor drives, renewable energy, artificial intelligence, intelligent control, and digital library. She can be contacted at email: sugi.mj@gmail.com.