



OPEN ACCESS

This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Department of Artificial Intelligence, University of Management and Technology, Lahore, Pakistan

Correspondence to: Muhammad Faraz Manzoor, Faraz.manzoor@umt.edu.pk

Additional material is published online only. To view please visit the journal online.

Cite this as: Manzoor MF. Energy Load Forecasting with Machine Learning: Models, Metrics, and Future Directions. Premier Journal of Artificial Intelligence 2025;4:100018

DOI: <https://doi.org/10.70389/PJAI.100018>

Received: 15 June 2025

Revised: 5 July 2025

Accepted: 19 July 2025

Published: 5 August 2025

Ethical approval: N/a

Consent: N/a

Funding: No industry funding

Conflicts of interest: None

Author contribution: Muhammad Faraz Manzoor – Conceptualization, Writing – original draft, review and editing

Guarantor: Muhammad Faraz Manzoor

Provenance and peer-review: Unsolicited and externally peer-reviewed

Data availability statement: N/a

Energy Load Forecasting with Machine Learning: Models, Metrics, and Future Directions

Muhammad Faraz Manzoor

ABSTRACT

Energy load forecasting plays a crucial role in the efficient management and operation of smart grids, enabling utilities to optimize energy distribution and improve grid reliability. Recent advancements in machine learning (ML) techniques have significantly enhanced the accuracy and adaptability of energy load prediction models. This review explores various ML models used for energy load forecasting, including traditional models, deep learning approaches such as long short-term memory (LSTM) networks and gated recurrent units (GRUs), and ensemble methods. The review discusses the strengths and limitations of each approach, highlighting its applicability to different forecasting timeframes and data characteristics. Additionally, it examines the performance evaluation metrics commonly used to assess model accuracy and reliability. Although multiple studies have examined forecasting techniques, there is a lack of comprehensive evaluation that connects model choice with practical deployment constraints such as data quality, real-time scalability, and interpretability. This review addresses this gap by systematically analyzing the challenges and emerging solutions in the context of smart grid applications. Despite the progress made, challenges related to data quality, computational complexity, and model interpretability remain significant barriers. The review concludes with an exploration of emerging trends and future directions in energy load forecasting, including hybrid models, federated learning, and reinforcement learning, which offer promising solutions to overcome existing limitations and improve forecasting performance in smart grid systems. Overall, the findings suggest that while no single model is universally optimal, integrating external factors, improving data quality, and adopting hybrid or explainable artificial intelligence (AI) approaches are critical for building more accurate, scalable, and interpretable forecasting systems.

Keywords: Energy load forecasting, Smart grids, Deep learning models, Ensemble methods, Predictive accuracy

Introduction

Energy load forecasting is important in smart grid management for the purpose of optimal demand side planning, resource and energy allocation, and cost optimization.¹ The accuracy of prediction has been shown to be a necessity due to the growing integration of renewables and demand for more electricity. However, existing traditional forecasting methods, from statistical and econometric models, have limited success in generating such nonlinear and dynamic energy consumption patterns.² In this context, machine learning (ML) is known as a game changer in the field, providing the ability to learn from historical consumption data and external factors to improve forecasting accuracy.

To better understand forecasting strategies, it is important to distinguish between types of energy load demands. Energy load forecasting can be further categorized based on specific end-use types such as cooling, heating, and general electricity demand, each having distinct characteristics and influencing factors. Cooling load forecasting is highly sensitive to outdoor air temperature and solar radiation, with significant variations during summer months and in commercial buildings. Heating load forecasting, on the other hand, is more influenced by low temperatures and seasonal shifts in winter, especially in residential zones. General electricity demand forecasting, while encompassing both heating and cooling needs, also includes lighting, appliances, and industrial usage patterns. These differences necessitate tailored forecasting models that account for the unique drivers and temporal patterns associated with each load type.

Importance of Energy Load Forecasting in Smart Grids

In order to balance supply and demand and to minimize operational costs and improve energy efficiency, smart grids are based on accurate energy load forecasting. This enables utility companies to generate power appropriately and not face energy crises or get dependent on backup power sources. Forecasting also provides for grid stability by knowing when peak demand will occur, and the supply is changed accordingly.³ As smart grids integrate more distributed energy resources, such as solar and wind power, forecasting becomes even more critical to manage intermittent generation and ensure grid reliability. If there is no right forecasting, energy providers might face supply versus transport mismatch, high operational cost, and poor energy distribution.⁴

Role of ML in Improving Forecasting Accuracy

ML has revolutionized energy load forecasting by addressing the limitations of traditional methods.^{5,6} Reinforcement Learning (RL) models are different from statistical models, which presume that data has a particular form to be modeled, but ML techniques can learn complicated patterns in data from history without pre-programmed declaration.⁷ Past consumption trends can be analyzed in a robust way that gives good predictions using supervised learning models like decision tree (DT) or random forest.⁸ Deep learning (DL) models, including the long short-term memory (LSTM) network and gated recurrent unit (GRU), excel in capturing long-term dependencies and seasonal variations. Moreover, hybrid models consisting of multiple ML techniques have demonstrated great potential in further increasing forecasting performance. By incorporating real-time data

along with socioeconomic and weather-related factors, ML-based approaches offer a more responsive and precise method for forecasting energy load.⁹

Scope and Objective of the Review

This review examines smart grid energy load forecasting using a succinct but comprehensive discussion of ML techniques to predict the energy load in this smart grid. This is to evaluate the effectiveness of different ML models and how well they can predict performance and its limitations. In this context, a summary of certain advantages of using ML in forecasting and some of the implementation challenges and possible future research directions is presented. This study synthesizes the recent advancements in the field and proposes insights into how ML can enable smart grids to operate with higher efficiency and reliability, thus helping to adopt more sustainable energy management practices.

To guide the reader, the structure of the review is as follows: Section Research Methodology outlines the methodology adopted for this study, including the systematic literature review (SLR) process, search strategy, selection criteria, and inclusion/exclusion parameters. Section ML Techniques for Energy Load Forecasting presents a detailed literature review of ML techniques applied to energy load forecasting. Section Comparison of Predictive Techniques and Common Evaluation Metrics provides a comparative analysis of various forecasting methods, including traditional models, DL architectures, and ensemble techniques, evaluating their performance across different forecasting scenarios. Section Challenges and Future Directions discusses the key challenges, such as data quality, computational demands, and interpretability, and outlines future directions, including hybrid approaches, federated learning, and RL. Finally, Section Conclusion concludes the review by summarizing the

main findings and emphasizing the need for continued research to enhance the accuracy and robustness of energy load forecasting in smart grids.

Research Methodology

This study follows a SLR methodology to comprehensively analyze existing research on energy load forecasting techniques, as shown in Figure 1. The methodology is structured to ensure transparency, reproducibility, and rigor in identifying, evaluating, and synthesizing relevant studies on ML models applied to energy load forecasting.^{10,11}

Research Objective (RO) and Research Question (RQ) Formulation

This study uses a structured framework of ROs, RQs, and motivations to define the focus of the study in energy load forecasting, as shown in Table 1.

Literature Search Strategy

A search string was formulated to ensure that the search process captured only systematic and comprehensive studies relevant to energy load forecasting. The search string was designed to target literature focusing on the application of ML and DL techniques in forecasting energy consumption across various time horizons and contexts. The keywords and Boolean operators were defined as follows:

(“energy load forecasting” OR “electricity demand forecasting” OR “power load prediction”) AND (“machine learning” OR “deep learning” OR “time series prediction”) AND (“short-term” OR “long-term” OR “real-time”) AND (“smart grid” OR “energy systems” OR “utilities”)

This search string was applied across multiple academic databases, including IEEE Xplore, Springer-Link, ScienceDirect, and the Multidisciplinary Digital Publishing Institute (MDPI), to identify and collect relevant peer-reviewed articles.

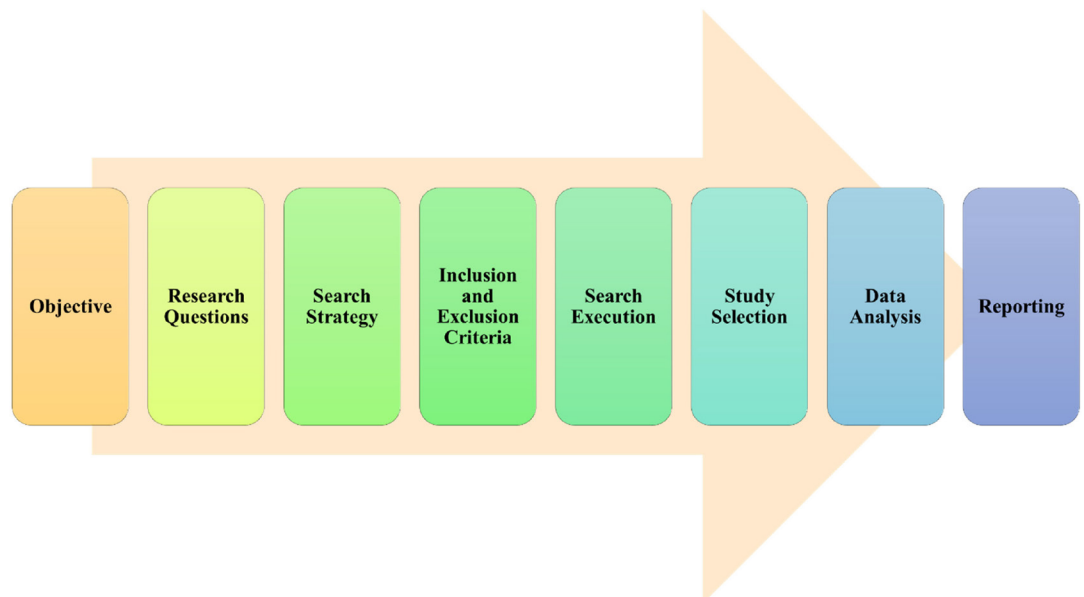


Fig 1 | Systematic review process

Study Selection Criteria

In order to limit the studies considered in this SLR to the most relevant and high-quality research, we applied a structured and multistage filtering process.^{10,12} In this study, the selection was conducted in a sequential manner using rigorous screening at multiple levels, such as title-based screening, abstract-based filtering, and full-text review, as shown in Figure 2. This process refined the selection to studies that contain direct contributions of data, models, or insights relevant to ML and DL techniques for energy load forecasting.^{13,14}

Title-Based Screening

In the first step of the selection process, a title-based screening was applied to the studies, where they were assessed based on their titles. Shortlisted research articles explicitly referred to techniques related to energy load forecasting, electricity demand prediction, time series modeling, ML, and DL. Also, studies with vague, overly general, or unrelated titles were immediately excluded to retain only relevant literature focused on forecasting within energy systems.

Abstract-Based Screening

After the title screening, the abstracts of the shortlisted papers were carefully reviewed to determine their relevance. If the study focused on ML or DL techniques specifically applied to energy load forecasting, then the study was included. It also prioritized papers discussing applications in smart grids, utility demand prediction, renewable energy integration, and real-time load management. Suitable for further review were abstracts that provided technical insights into model architectures, datasets used, and performance evaluation metrics. At this phase, studies that lacked technical depth, focused on unrelated domains, or

did not employ data-driven forecasting methods were excluded.

Full-Text Review

Full-text review of the remaining studies completed the last stage of the selection process. This was an important step to ensure that the selected papers clearly described their methodology, experimental setup, and analytical results. Articles were included if they provided quantitative performance evaluation in the form of metrics such as root mean squared error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), R-squared (R^2), or computational efficiency. Additionally, prioritized were studies that discussed challenges, limitations, and future research directions in the context of energy load forecasting. To improve the credibility and reliability of the reviewed literature, only articles published in high-impact journals or peer-reviewed conferences were considered. To ensure that only high-quality, methodologically sound, and impactful studies were included, papers with complete methodologies but lacking experimental validation, or those using outdated forecasting techniques, were excluded.

Inclusion and Exclusion Criteria

During the process of selecting high-quality and relevant studies, a set of predefined inclusion and exclusion criteria was applied to the study selection, as shown in Table 2. Refining the literature involved using these criteria to filter out studies that did not align with the ROs. In this SLR, the following specific inclusion and exclusion parameters were established, as detailed in the Table 2.

Study Selection Results

A total of 45 papers were selected and analyzed to address the RQs outlined above. Table 2 presents the

SR#	RO	RQ	Motivation
1	Identify and analyze the most commonly used ML and DL techniques for energy load forecasting.	What are the most frequently applied ML and DL techniques for energy load forecasting?	Understanding the dominant models helps in recognizing methodological trends, strengths, and limitations, which can guide future research and practical deployment.
2	Examine the primary application areas, forecasting timeframes, and data types used in energy load prediction.	In what contexts, time horizons, and data environments is energy load forecasting commonly applied?	Analyzing use cases across different settings highlights domain-specific requirements, temporal challenges, and data dependency of forecasting models.
3	Investigate the key challenges in energy load forecasting and explore future directions to improve accuracy and scalability.	What are the major challenges in energy load forecasting, and what future directions can enhance model performance?	Identifying core challenges and emerging solutions can improve accuracy, real-time feasibility, and interpretability of forecasting systems in smart grids.

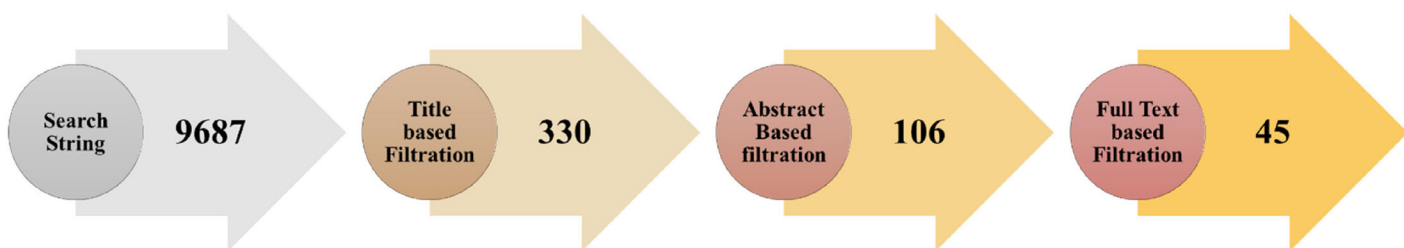


Fig 2 | Study selection process

Table 2 | Inclusion exclusion criteria

Criteria	Inclusion	Exclusion
Study type	Peer-reviewed journal articles, conference papers, and high-impact workshop papers	Non-peer-reviewed articles, book chapters, theses, white papers, or blogs
Research focus	Studies focusing on ML and DL techniques for energy load forecasting	Studies focusing on general data analytics or unrelated ML tasks
Dataset availability	Papers that use publicly available datasets or provide details of proprietary datasets used for model training and evaluation	Studies that do not specify datasets, making comparison and validation difficult
Technical depth	Papers providing detailed methodology, including model architectures, training procedures, and performance evaluation metrics	Studies lacking methodological clarity or experimental validation
Language	Papers published in English	Papers published in languages other than English

Table 3 | Study selection results

Phase	Process	Selection Stage	Elsevier	Springer	IEEE	MDPI	Others	Total
1	Search	Search string	319	1900	4719	2400	349	9687
2	Screening	Title	32	79	70	104	45	330
3	Screening	Abstract	21	29	18	30	8	106
4	Screening	Full text	11	10	16	5	3	45

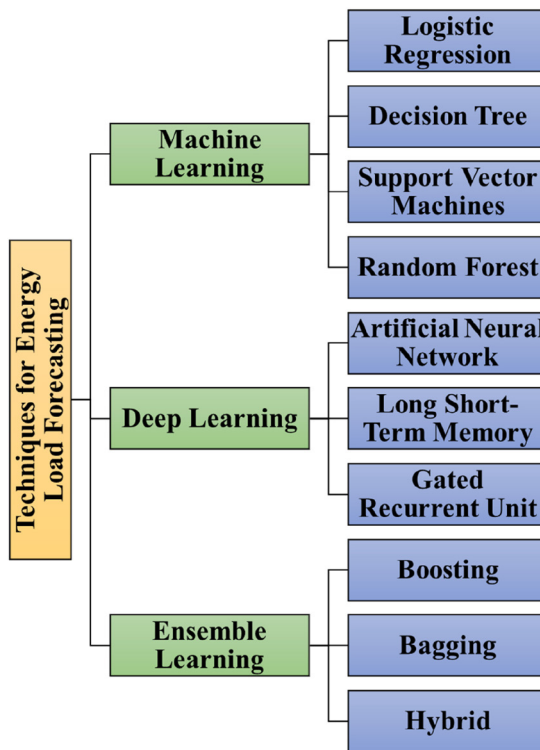


Fig 3 | Taxonomy of techniques for energy load forecasting

distribution of studies based on their sources, while Table 3 provides a detailed list of the selected papers.

ML Techniques for Energy Load Forecasting

ML techniques are used to model complex and non-linear energy consumption patterns required by energy load forecasting. These techniques used historical data, real-time grid data, and external variables (weather, economic indicators) to leverage the prediction accuracy (Figure 3). In general, these ML models are used: traditional ML models, DL models, and

ensemble models. Each of these categories has advantages and challenges different enough to make it appropriate for a different forecasting scenario. Detailed discussion of these approaches is given in this section, including how they are applied, how effective they are, and how well they perform.

Traditional ML Models

Energy load forecasting has traditionally relied on interpretable and computationally efficient ML models that perform well with structured data. Common examples of such models include the linear regression (LR), DT, support vector machine (SVM), and random forest (RF). These models are favored for their transparency and ease of implementation, particularly in short-term forecasting scenarios. For instance, Söderberg and Meurling¹⁵ conducted a comparative study on various ML models, including regression analysis and DTs, for predicting household energy consumption, highlighting their effectiveness in capturing consumption patterns. Similarly, Lahouar et al.¹⁶ developed a one-day-ahead load forecasting model using regression RF, demonstrating its accuracy and robustness in short-term load prediction. These studies underline the practical value of traditional ML approaches, particularly when interpretability and quick deployment are critical.

Linear Regression (LR)

One of the simplest models for energy load forecasting is linear regression with the assumption that a relationship with input variables (e.g., past energy consumption, temperature) and output variable (future energy demand) is linear (Figure 4). For example, Sarker et al.¹⁷ utilized linear regression combined with inverse matrix analysis to forecast energy demand in remote areas of Bangladesh, demonstrating its efficacy in regions with limited data availability. On the other hand, Mahmud¹⁸ employed linear regression analysis to predict electricity loads in isolated areas lacking

historical data, underscoring the model’s practicality in diverse settings. The model is represented as:

$$Y = \beta_0 + \sum_{i=1}^n \beta_i X_i + \varepsilon \tag{1}$$

where Y is the predicted energy load, X_i are input features, β_i are coefficients, and ε represents error terms. While LR is computationally efficient, its main limitation is its inability to model complex, nonlinear relationships in energy demand.

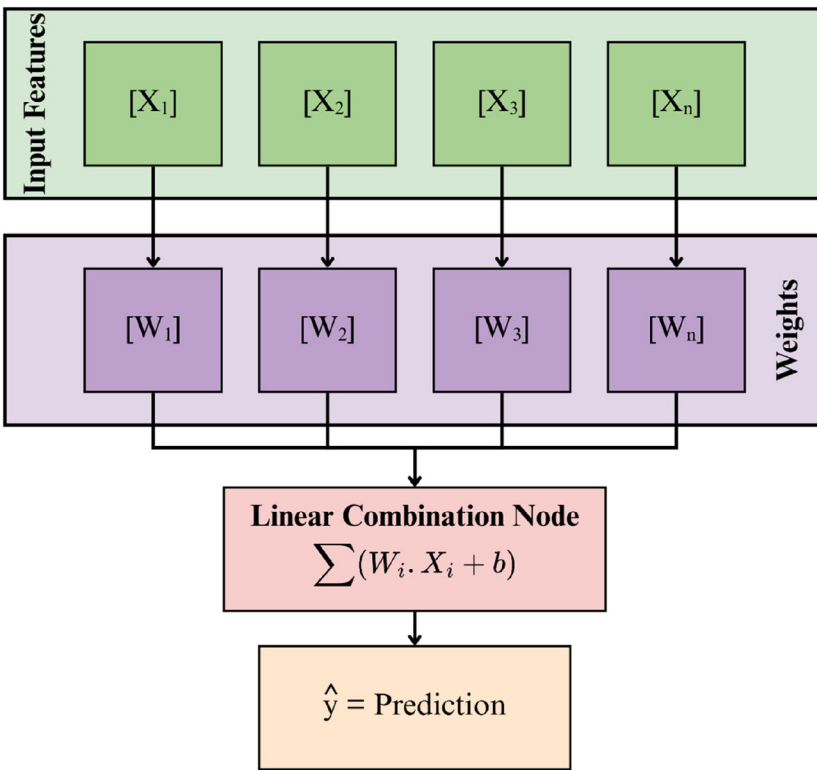


Fig 4 | General mechanism of linear regression

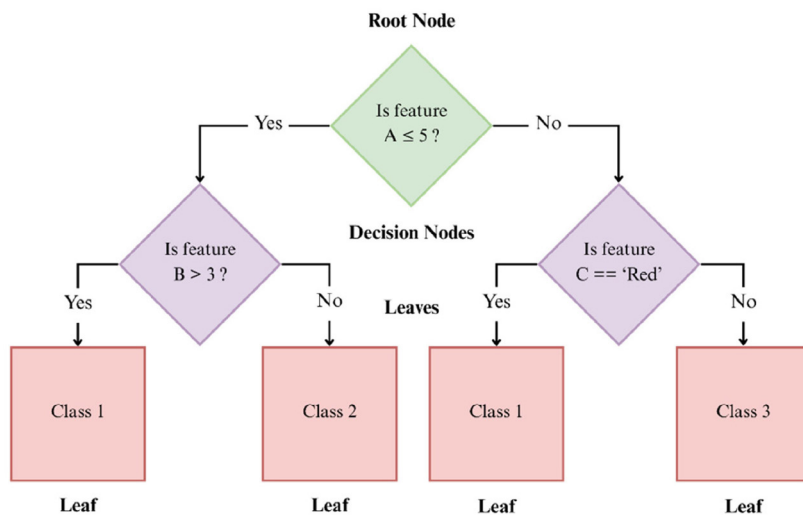


Fig 5 | General mechanism of DT

Decision Tree (DT)

As DTs partition the dataset based on feature importances, DTs become suitable for capturing nonlinear dependencies in energy load data (Figure 5). DTs are good with categorical as well as numeric features, but are prone to overfitting, more so for deep tree structure. Likewise, Hambali et al.¹⁹ evaluated various DT algorithms for electric power load forecasting and found that the REPTree technique outperformed others, highlighting its effectiveness in handling nonlinear relationships in load data. Likewise, Jahan et al.²⁰ designed a DT model using weather and power load data as inputs for short-term demand forecasting, demonstrating the model’s capability to capture complex dependencies in energy consumption patterns.

Support Vector Machine (SVM)

SVMs are effective for short-term load forecasting as they find an optimal hyperplane for classification or regression (Figure 6). Furthermore, Tian and Noore²¹ proposed an SVM-based approach for short-term load forecasting, demonstrating improved generalization and lower prediction error compared to neural network methods. On the other side, Li and Fang²² applied SVMs to power system load forecasting, effectively enhancing forecasting accuracy. The model is based on:

$$\min_{w,b,\xi} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i \tag{2}$$

subject to constraints ensuring that predictions fall within an acceptable error margin. However, SVMs can be computationally expensive for large datasets.

Random Forest (RF)

RF is an ensemble of DTs that serves the purposes of improving forecasting accuracy and reducing overfitting (Figure 7). It merges the outputs from multiple trees, effectively improving robustness, and is a good candidate for energy load forecasting in dynamic ecosystems. Some researchers work on it like Dudek,²³ who applied RF to short-term electricity load forecasting, demonstrating its ability to handle nonstationary and seasonal time series data effectively. Meanwhile, Gao et al.²⁴ proposed an improved RF regression algorithm for ultra-short-term electricity load forecasting, achieving higher accuracy compared to traditional methods (Table 4).

DL Models

The growing popularity of DL models in energy load forecasting stems from their ability to capture long-term dependencies and model complex nonlinear relationships in energy consumption data. Notable models in this category include the artificial neural network (ANN), LSTM network, and GRU. Unlike traditional models, these architectures are capable of automatically extracting hierarchical features from raw time series data, making them particularly effective in scenarios with highly dynamic and temporally correlated consumption patterns.

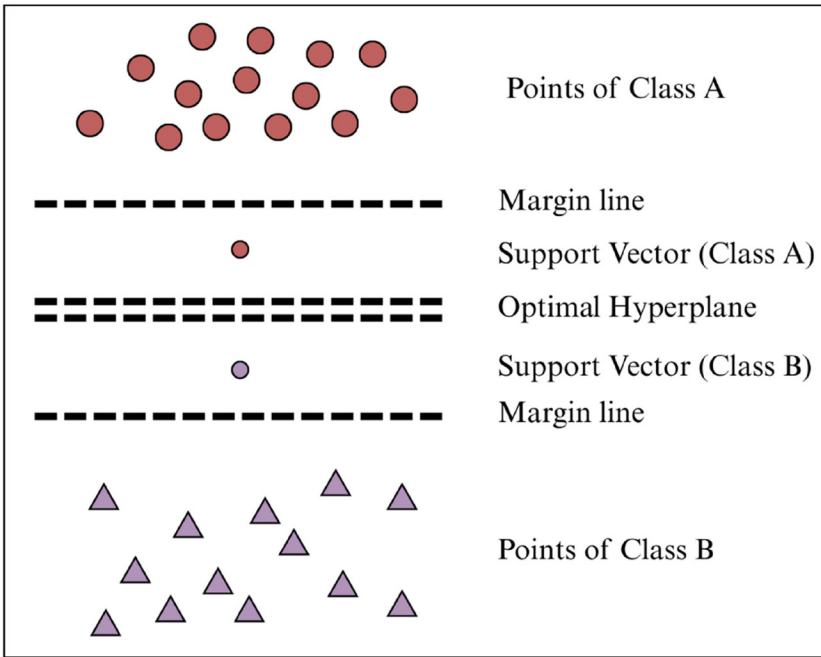


Fig 6 | General mechanism of SVM

Artificial Neural Networks (ANNs)

ANNs consist of multiple layers of neurons that learn energy consumption patterns from historical data (Figure 8). Numerous researchers work on this model, including, Oreshkin et al.²⁵ who introduced the N-BEATS neural network architecture, which demonstrated superior performance in mid-term electricity load forecasting across various European datasets. Similarly, Gao et al.²⁶ proposed an ensemble deep random vector functional link (edRVFL) network, which achieved high accuracy in short-term load forecasting tasks by leveraging deep representation learning and ensemble strategies. The weight updates are as follows:

$$w_{ij} = w_{ij} + \eta \frac{\partial E}{\partial w_{ij}} \tag{3}$$

where w_{ij} are weights, η is the learning rate, and E is the error function. ANNs perform well for medium- to long-term load forecasting but require large datasets and high computational power.

LSTM Networks

Energy load data are captured with temporal dependencies by LSTM networks, a type of recurrent neural network (RNN). However, due to vanishing gradients in traditional RNNs, they overcome vanishing gradients through memory cells that store information for long periods (Figure 9). However, the DA-LSTM framework introduced by Bayram et al.²⁷ is a dynamic drift-adaptive learning model designed for interval load forecasting. This approach enhances forecasting accuracy by adapting to changes in consumption patterns without requiring predefined drift thresholds. Meanwhile, a hybrid LSTM model was developed by Lu et al.²⁸ incorporating online correction mechanisms. Their model integrates temporal and nontemporal features and employs an online correction strategy to adjust to real-time data distribution shifts, thereby improving day-ahead load forecasting accuracy. The cell state update is given by:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \tag{4}$$

where C_t is the cell state, f_t is the forget gate, and i_t is the input gate. LSTMs are highly effective for energy load forecasting due to their ability to recognize seasonality and trends.

Gated Recurrent Units (GRUs)

The main difference between GRUs and LSTMs is that the former has a simpler architecture and therefore computational complexity (Figure 10). The short-term dependency between the periods is balanced with the long-term dependency and therefore, it is a decent indication for real-time energy load forecasting. For example, Emshagin et al.²⁹ developed customized LSTM and GRU models for short-term household electricity consumption prediction. Their study found that while both models effectively captured consumption patterns, LSTM slightly outperformed GRU in forecasting accuracy. Likewise, Dong and Grumbach³⁰ introduced a hybrid long-term load forecasting method utilizing

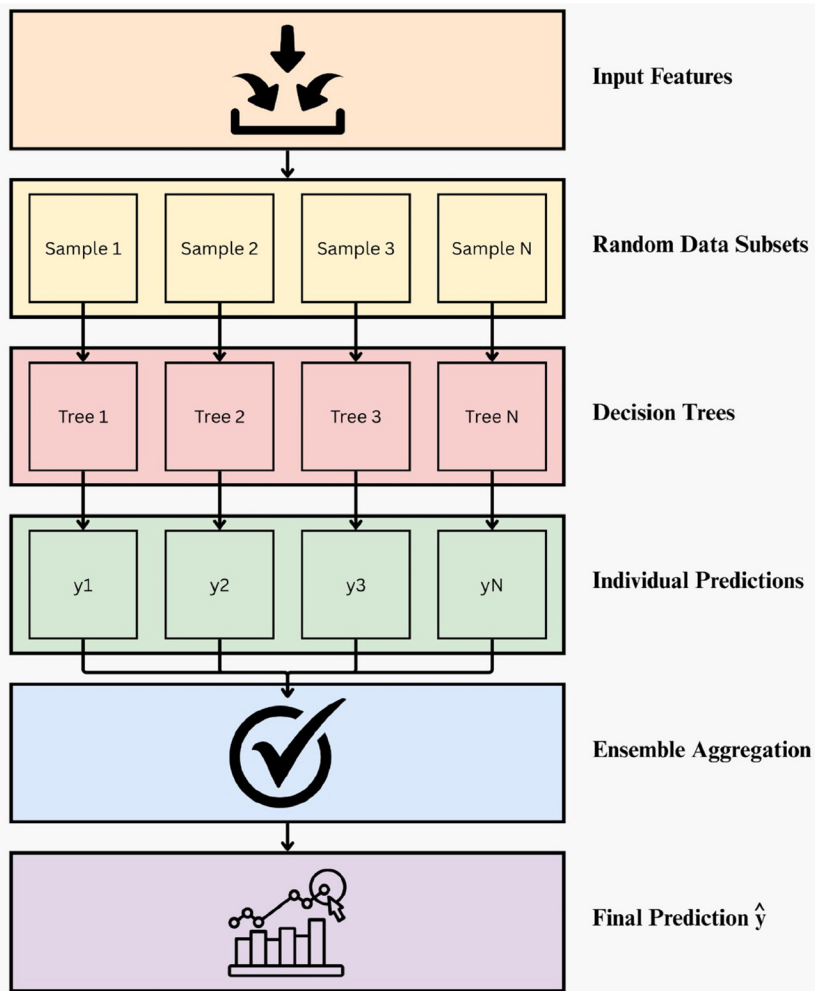


Fig 7 | General mechanism of RF

Model	Strengths	Weaknesses	Best For	Complexity	References
Linear regression	Simple, interpretable	Cannot model nonlinear patterns	Baseline forecasting	Low	17,18
DTs	Handles categorical and numerical data well	Prone to overfitting	Medium-term forecasting	Medium	19,20
SVM	Robust to outliers	Computationally expensive	Short-term forecasting	High	21,22
RF	High accuracy, reduces overfitting	Less interpretable	Complex energy load patterns	Medium-High	23,24

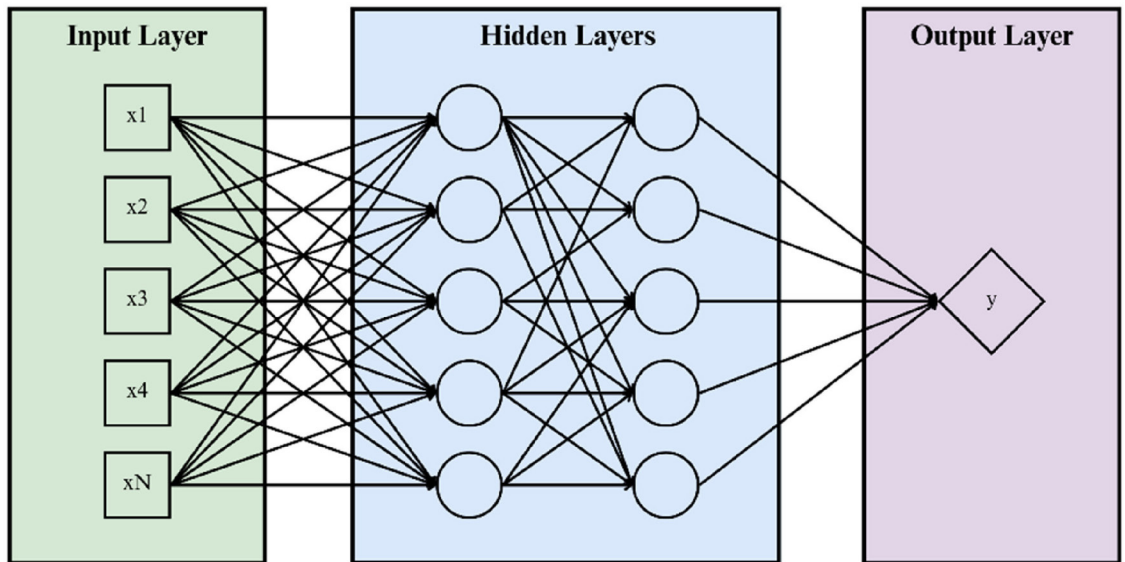


Fig 8 | General mechanism of ANN

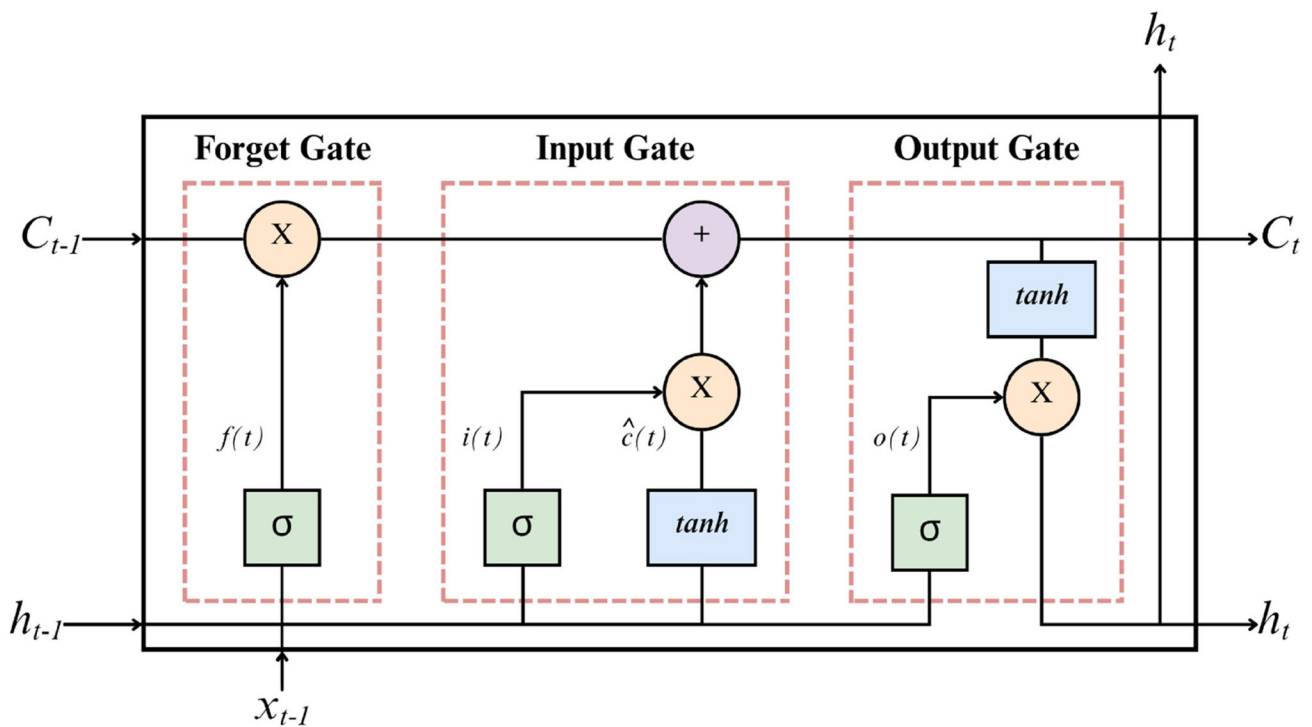


Fig 9 | General mechanism of LSTM

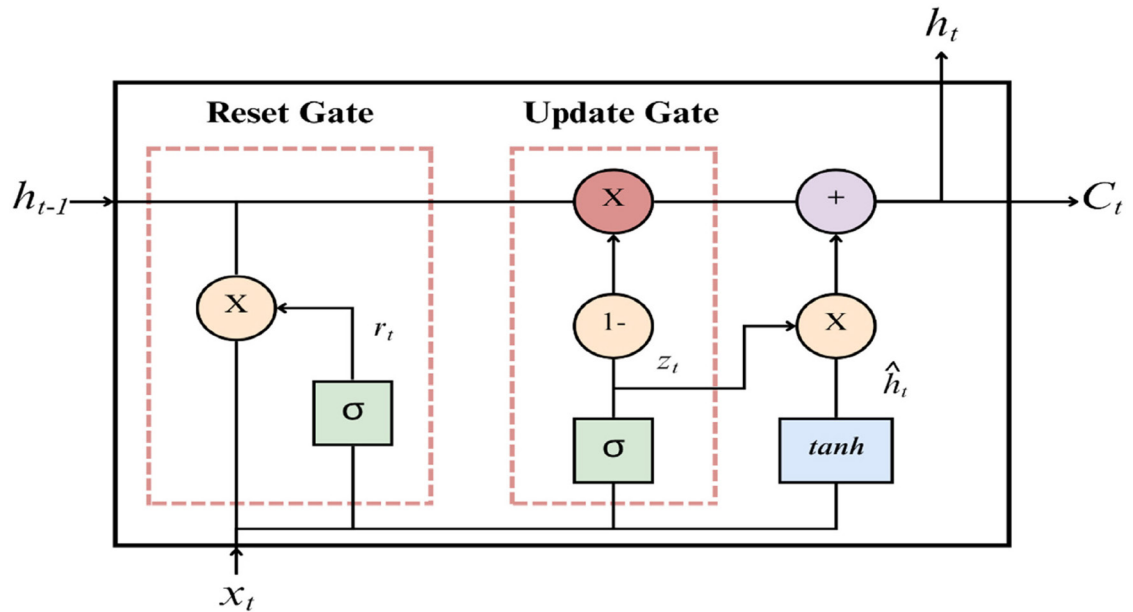


Fig 10 | General mechanism of GRU

Model	Strengths	Weaknesses	Best For	Complexity	References
ANN	Captures nonlinear relationships	Requires large datasets	Medium-term forecasting	High	25,26
LSTM	Models long-term dependencies	Computationally expensive	Long-term forecasting	Very High	27,28
GRU	Simplifies LSTM with similar accuracy	May not capture complex dependencies	Real-time forecasting	High	29,30

both LSTM and GRU networks. Their approach successfully integrated top-down, bottom-up, and sequential information, demonstrating the practicality of GRUs in capturing temporal dependencies in energy consumption data (Table 5).

Ensemble Approaches

Ensemble learning combines the predictive strengths of multiple models to improve forecasting accuracy and reduce variance in predictions. This approach is particularly effective in addressing the limitations of individual models by aggregating their outputs. The common ensemble methods used for energy load forecasting include boosting, bagging, and hybrid models. Boosting techniques sequentially train models to correct errors of previous ones, while bagging methods reduce variance by averaging predictions across independently trained models. Hybrid models, on the other hand, integrate different algorithmic paradigms—such as combining traditional and DL models—to leverage the advantages of both.

Boosting Methods

Boosting methods, such as gradient boosting machines (GBMs), XGBoost, and AdaBoost, iteratively improve weak learners by focusing on difficult-to-predict instances (Figure 11). A bagging model named XASXG proposed by, Ding et al.³¹ which integrates autoregressive integrated moving average (ARIMA),

support vector regression (SVR), and XGBoost. This model effectively captures the nonstationary and nonlinear characteristics of electricity load time series, resulting in superior forecasting accuracy across various Chinese cities. Similarly, Rubattu et al.³² developed a predictive model utilizing probabilistic LightGBM combined with temporal hierarchies and advanced feature engineering. Their approach achieved high accuracy in the BigDeal Challenge 2022, demonstrating the efficacy of boosting methods in load and peak forecasting tasks. The loss function for boosting is:

$$F_m(x) = F_{m-1}(x) + \gamma h_m(x) \tag{5}$$

where $F_m(x)$ is the current model, $h_m(x)$ is the new weak learner, and γ is a learning rate. These methods enhance accuracy but may suffer from overfitting if not carefully tuned.

Bagging Methods

To reduce variance and improve stability, bagging techniques such as RF and bootstrap aggregation perform a number of versions of the training dataset (Figure 12). Thus, these models are good at handling fluctuating energy demand patterns. A cluster-based bootstrapping method was introduced by Dube et al.³³ for interval prediction of electricity demand. Their approach utilizes residual bootstrapping combined with unsupervised clustering to generate accurate interval forecasts, particularly

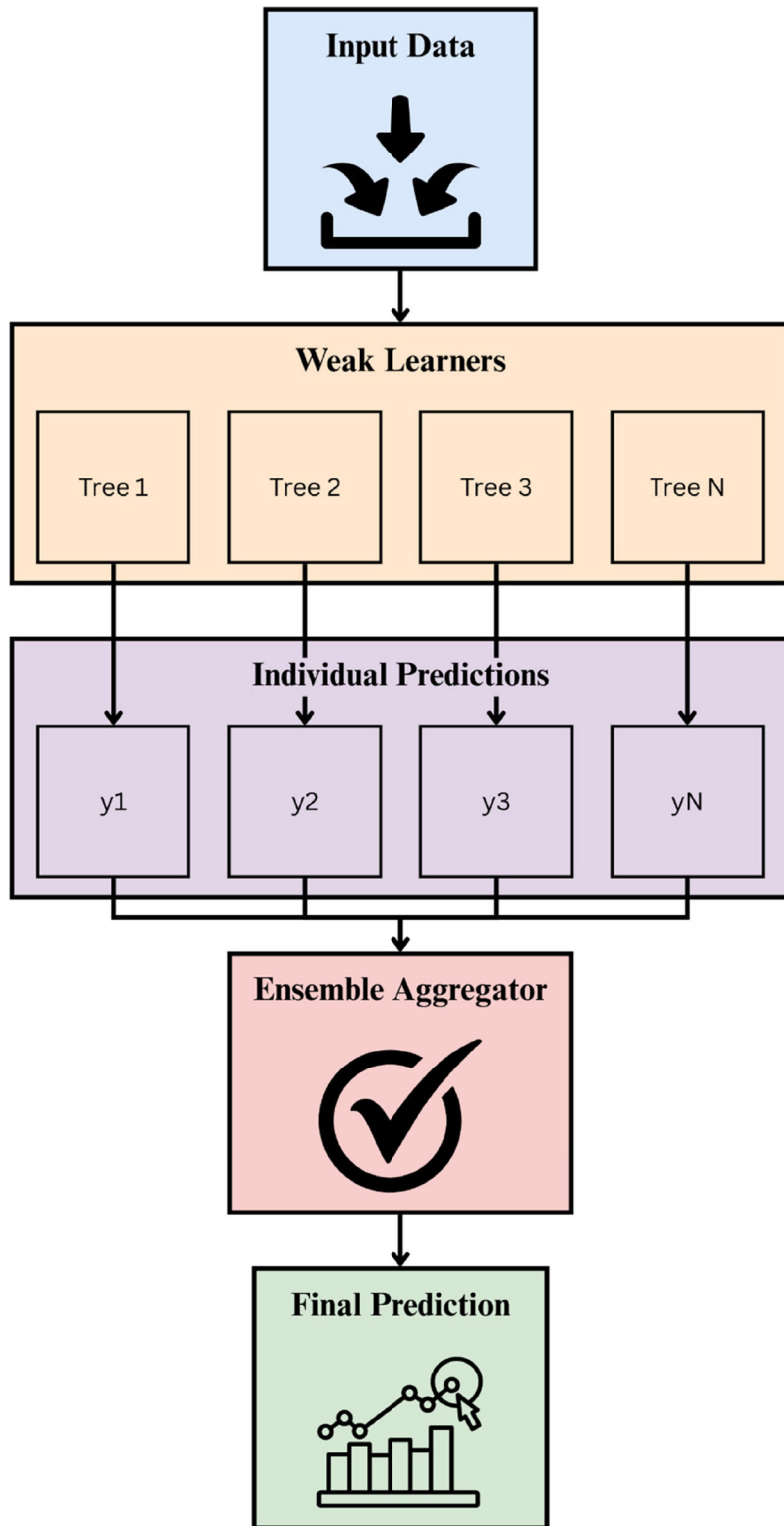


Fig 11 | General mechanism of boosting method

beneficial in microgrid settings with high demand variability. Additionally, the BaggingSHAP model, which was proposed by Olumba et al.,³⁴ combines bagging with SHapley Additive exPlanations (SHAP) for enhanced interpretability in energy demand forecasting. The model achieved an R^2 score of 0.9947, indicating superior predictive performance.

Hybrid Models

Traditional forecasting techniques are combined with the DL approaches using hybrid models to optimize the forecasting performance (Figure 13). As an example, RF and LSTM can be used together by an RF-LSTM model to select features with RF, apply LSTM to temporal pattern recognition while remaining both interpretable and accurate. Meanwhile, Zhang et al.³⁵ proposed a hybrid model combining RF and LSTM to improve short-term electricity load forecasting, demonstrating reduced prediction errors compared to standalone models. Similarly, Ma et al.³⁶ developed a hybrid RF-LSTM approach that leveraged RF for selecting the most relevant input features and LSTM for modeling sequential dependencies, achieving superior results in multistep energy load forecasting tasks (Table 6).

Ensemble methods provide a robust solution for energy load forecasting, particularly when dealing with fluctuating consumption patterns and large-scale smart grid data. They enhance forecasting accuracy by reducing bias and variance while leveraging the strengths of multiple models.

Comparison of Predictive Techniques and Common Evaluation Metrics

The effectiveness of predictive techniques in energy load forecasting largely depends on the accuracy, reliability, and efficiency of the ML models employed. Selecting the appropriate technique requires evaluating which model performs best based on specific performance metrics, while also considering the trade-offs associated with each approach. This section reviews commonly used ML models for energy load forecasting and provides a comparative analysis of their strengths and limitations with respect to various evaluation metrics.

Performance Metrics Used for Evaluation

There have been diverse performance metrics used to assess the predictive accuracy and reliability of the ML models used in energy load forecasting. This is useful in determining how well the model will minimize errors, and if it can be applied to average real-world problems. Mean squared error (MSE), RMSE, MAE, MAPE, and R^2 are the most common performance metrics.

Mean Squared Error (MSE)

MSE measures the average squared difference between predicted and actual values, with a focus on penalizing larger errors more than smaller ones. According to, Chai and Draxler³⁷ MSE is a commonly preferred metric in regression tasks due to its mathematical simplicity and sensitivity to large deviations. Likewise, Willmott and Matsuura³⁸ compared MSE with MAE, emphasizing MSE's utility when highlighting significant prediction errors is crucial in model evaluation. It is calculated as:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \tag{6}$$

where y_i is the actual value, \hat{y}_i is the predicted value, and n is the total number of data points. MSE is sensitive to

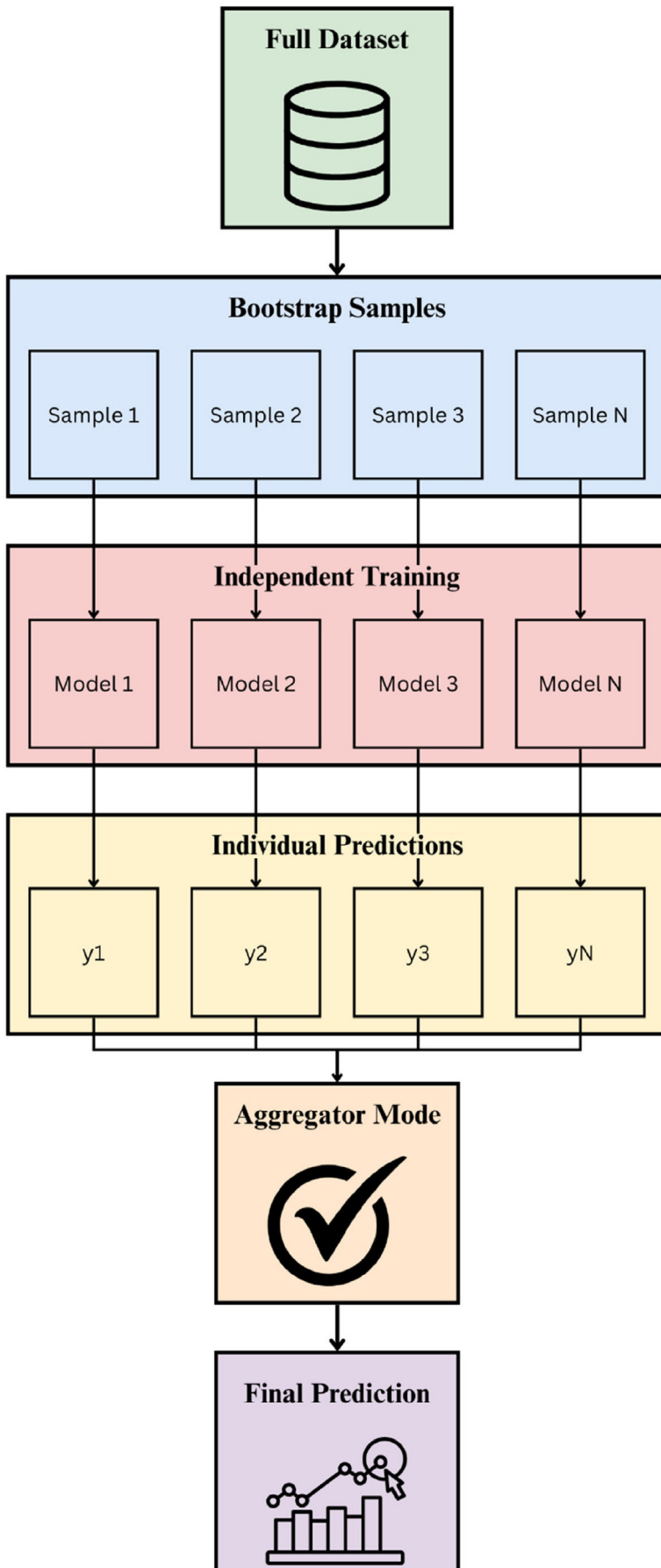


Fig 12 | General mechanism of bagging method

outliers, as it squares the error terms, which can lead to a disproportionate influence of large errors.

Root Mean Squared Error (RMSE)

RMSE is the square root of the MSE and provides a more interpretable error metric in the original units of the data. As highlighted by, Hyndman and Koehler³⁹ RMSE is particularly useful when large prediction errors are undesirable, offering a balance between interpretability and sensitivity. Meanwhile, Hong et al.⁴⁰ emphasized the importance of RMSE in energy forecasting competitions due to its ability to reflect forecasting accuracy effectively across different models and datasets. It is calculated as:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \tag{7}$$

RMSE is widely used in energy load forecasting because it reflects the magnitude of the error in a way that is more understandable to practitioners. However, like MSE, it is also sensitive to large errors.

Mean Absolute Error (MAE)

MAE measures the average magnitude of errors in a set of predictions, without considering their direction. However, Willmott et al.³⁸ emphasized the interpretability and robustness of MAE in environmental and climate modeling contexts, showing its reliability across diverse datasets. However, Hyndman and Koehler³⁹ discussed MAE as a reliable baseline metric in forecast accuracy comparisons, particularly useful when comparing models across datasets with different scales. It is given by:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \tag{8}$$

MAE is less sensitive to outliers compared to MSE and RMSE, making it suitable for applications where large deviations are less critical. It provides a straightforward measure of how far the predictions are from the actual values on average.

Mean Absolute Percentage Error (MAPE)

MAPE calculates the error as a percentage of the actual value, making it easier to interpret and compare across different datasets. As noted by Makridakis,⁴¹ MAPE remains one of the most commonly used metrics in forecasting competitions due to its intuitive nature and interpretability. However Evans,⁴² also pointed out that MAPE’s sensitivity to low actual values may distort its usefulness, especially in datasets with high variability or occasional zero values. It is calculated as:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100 \tag{9}$$

MAPE is useful when comparing forecast accuracy across different models and datasets, but it can be problematic when actual values are close to zero, leading to inflated error percentages.

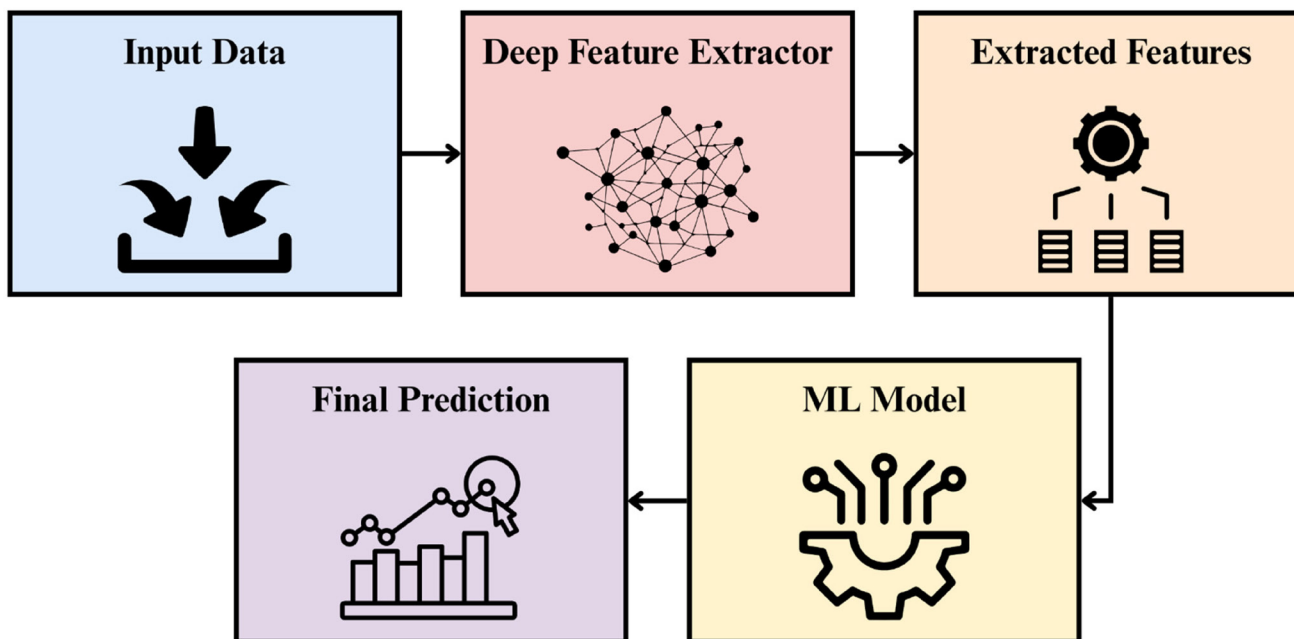


Fig 13 | General mechanism of hybrid models

Table 6 | Comparison of ensemble approaches for energy load forecasting

Model	Strengths	Weaknesses	Best For	Complexity	References
Boosting (XGBoost, AdaBoost)	Improves weak learners iteratively	May overfit	High-variance energy patterns	High	31,32
Bagging (RF)	Reduces variance, robust	Less interpretable	Medium-term forecasting	Medium	33,34
Hybrid (RF-LSTM)	Leverages best of both models	High computational cost	Long-term forecasting	Very High	35,36

Table 7 | Comparison of performance metrics for energy load forecasting

Metric	Strengths	Weaknesses	Best Used For	Sensitivity to Outliers	References
MSE	Penalizes large errors, easy to compute	Sensitive to outliers, hard to interpret	Measuring overall error magnitude	High	37,38
RMSE	Intuitive, in original units of data	Sensitive to large errors	Performance comparison across models	High	39,40
MAE	Less sensitive to outliers	Does not penalize large errors enough	When outliers are not as critical	Low	38,39
MAPE	Easily interpretable, percentage-based	Issues with values close to zero	Comparing models across datasets	High	41,42
R^2	Indicates how well the model fits data	Does not guarantee prediction accuracy	Evaluating goodness-of-fit	Medium	43,44

R-Squared (R^2)

R^2 measures the proportion of the variance in the dependent variable that is predictable from the independent variables. According to Draper and Smith,⁴³ R^2 is widely used in regression analysis as it succinctly summarizes the explanatory power of a model. However, as noted by Alexander,⁴⁴ R^2 alone can be misleading in evaluating model performance, especially in cases of overfitting or when used with nonlinear models, thus warranting the use of adjusted R^2 or complementary metrics. It is calculated as:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \tag{10}$$

Where \bar{y} is the mean of the observed values. R^2 provides a measure of the goodness-of-fit, indicating how well the model explains the variation in the data. A higher R^2 value indicates a better model fit, though it does not always correlate with improved prediction accuracy, particularly in complex datasets (Table 7).

Strengths and Limitations of Different ML Models

Typically, when choosing an ML model for energy load forecasting, one needs to evaluate the strengths and weaknesses of each model in terms of the nature of the energy load forecasting task. The performance may vary from model to model according to the complexity of the data available and the length of the

forecasting horizon as well as the amount of historical data available.

Traditional ML Models

Popular traditional ML models include linear regression, DTs, SVMs, and RF which are always popular for their interpretability and ease of implementation. These models are best applicable for reasonably simple, short-term forecasting tasks with relatively simpler and easy-to-understand data. But they are not very useful with long-term dependencies or high non-linearity in the process data. For instance, Hippert et al.⁴⁵ and Wang et al.⁴⁶ demonstrated the effectiveness of linear regression in short-term load forecasting due to its simplicity and speed. DTs, as shown by Liu et al.⁴⁷ and, Singh et al.⁴⁸ are useful for capturing nonlinear relationships in structured energy data. Meanwhile, RF have been proven robust in handling noisy data and enhancing forecasting accuracy, as shown in the works of Lahouar et al.⁴⁹ and, Deb et al.⁵⁰ particularly in short-term electricity demand forecasting (Table 8).

DL Models

Indeed, energy load forecasting is a good application for DL models, especially LSTM networks, GRU, and even ANN. These models are capable of capturing the time series and the nonlinear patterns within a large amount of data. This type of model is great in finding long-term trends and fine relationships between time series data, which are perfect for predicting purposes when historical consumption data is used as a fundamental in forecasting future energy demand.

LSTM and GRU are structures of rather diverse DL models, being parts of the RNN, which aims to deal more efficiently with sequential data compared to classic feed-forward networks, e.g., ANN. GRUs are known to be not as good as LSTMs at capturing long-term dependencies in time series data; however, GRUs are computationally lighter and more efficient at performing similar tasks as LSTMs. Several studies

have validated this: Kuster et al.⁵¹ and Suganthi and Samuel⁵² demonstrated the utility of ANN in modeling complex energy consumption behaviors due to its flexibility and learning capabilities. LSTMs, as applied by Marino et al.⁵³ and, Kong et al.⁵⁴ have shown exceptional performance in capturing long-term dependencies in electricity demand prediction. Meanwhile, GRUs have been successfully employed by Liu et al.⁵⁵ and, Liang et al.⁵⁶ where they offered competitive accuracy with reduced computational overhead compared to LSTMs, making them suitable for real-time forecasting applications (Table 9).

Ensemble Approaches

Ensemble methods like boosting (e.g., XGBoost) and bagging (e.g., RF) combine multiple models to improve forecasting accuracy and reduce variance. These methods are highly effective in energy load forecasting, particularly for datasets with noisy or fluctuating patterns. However, several researchers have demonstrated the effectiveness of these ensemble approaches. For boosting, Hong et al.⁵⁷ applied gradient boosted regression trees (GBRTs) to improve short-term load forecasting accuracy in large-scale smart grid applications, while Taieb and Hyndman⁵⁸ evaluated boosting strategies for probabilistic forecasting in energy systems. For bagging, Lahouar and Slama⁵⁹ showed that RFs provided high accuracy and robustness in hourly electricity demand prediction and Zhou et al.⁶⁰ confirmed their efficiency in handling feature-rich energy datasets and capturing nonlinear relationships (Table 10).

Selecting the right ML technique for energy load forecasting depends on the specific characteristics of the dataset and the forecasting horizon. While traditional models are efficient and easy to interpret, DL and ensemble approaches offer superior accuracy in complex scenarios. Performance metrics such as MSE, RMSE, MAE, and R^2 play a vital role in comparing and selecting the best model for a given task.

Table 8 | Comparison of traditional ML models for energy load forecasting

Model	Strengths	Limitations	Best Use Case	Computational Complexity	References
Linear regression	Simple, interpretable	Limited to linear relationships	Baseline forecasting, short-term forecasts	Low	45,46
DT	Handles nonlinear relationships	Prone to overfitting, poor generalization	Small to medium-sized datasets	Medium	47,48
RF	Robust, reduces overfitting	Less interpretable, computationally intensive	Complex forecasting tasks	High	49,50

Table 9 | Comparison of DL models for energy load forecasting

Model	Strengths	Limitations	Best Use Case	Computational Complexity	References
ANN	Captures nonlinear relationships, flexible	Needs large datasets, prone to overfitting	Medium-term forecasting, large datasets	High	51,52
LSTM	Long-term dependency handling, robust for time series	Computationally expensive, complex tuning	Long-term forecasting, capturing seasonality	Very High	53,54
GRU	Efficient, faster training, captures temporal dependencies	May not capture long-range dependencies as well as LSTM	Short to medium-term forecasting, large datasets	High	55,56

Table 10 | Comparison of ensemble approaches for energy load forecasting

Model	Strengths	Limitations	Best Use Case	Computational Complexity	References
Boosting	Improves weak learners, high accuracy	Overfitting, sensitive to noisy data	Complex, noisy datasets, high-variance forecasting	High	57,58
Bagging	Reduces overfitting, stable	Less interpretable, computationally expensive	Large datasets, medium-term forecasting	Medium-High	59

Challenges and Future Directions

In smart grids, energy load forecasting is extremely important for effective energy distribution and real-time usage management. Despite advancements, several challenges remain unresolved, particularly those related to improving prediction accuracy, model scalability, and general applicability across diverse systems. These challenges not only hinder current progress but also highlight key directions for future research and innovation in this field (Table 9).

Data Availability and Quality Issues

The availability and quality of data are considered to be among the main issues in energy load forecasting. Most of the effectiveness of ML models lies in the richness and reliability of the dataset that they are trained on. Often, the energy load data is sparse, noisy, and incomplete, and hence it results in inaccurate prediction.⁶¹ Recovery Digital Data, a source for historically obsolete utility consumption data, is often ill-suited for the months and years before a system implemented this type of data. In addition, data from various sources might not be standardized, and thus, it becomes challenging to concatenate heterogeneous datasets to train predictive models.⁶²

There is another important issue: the quality of data. Typically, energy load data contains missing values, outliers, and errors that can do a lot of harm to model performance. Although the data may always be available, it does not always mean it covers all the possible conditions that affect energy consumption, such as weather events, special holidays, or unusual spikes of demand.⁶³ Further, data heterogeneity, or where different sources collect data in various formats at different scales and intervals, also creates an issue. All of these factors contribute to the difficulty in developing such generalizable, accurate models. In future research, these aspects should be improved concerning data acquisition, data format standardization, and robust data handling methods for missing and noisy data.⁶⁴

Computational Complexity and Model Interpretability

Energy load forecasting is another challenge and calls for computing complexity due to the usage of advanced ML models such as DL. Only with models like LSTMs networks and GRUs, we get good accuracy, but at a high cost of computation in terms of training and inference.⁶⁵ Such models can be quite expensive in terms of processing power and time, as well as storage, in particular when large and high-dimensional sets of data are involved. Therefore, these models cannot be

deployed in real time or on a vast scale, thus making them unfeasible for some utilities or organizations regarding the resources required and the cost associated with such deployment.⁶⁶

In addition, DL models are also complex, and this constitutes a potential problem with model interpretability. Accurate prediction of energy consumption is not the only thing important in energy load forecasting; we also need to know why we predicted such things.⁶⁷ Smart grids are particularly important given the need among stakeholders, grid operators, and regulators, as well as end users, to interpret how and why particular energy usage patterns are predicted. Often, such complex black box models, such as deep neural networks, cannot provide users with deep insight into factors that would either enable them to trust the predicted value or act on that particular prediction.⁶⁸ Future research should aim to strike a balance between model complexity and interpretability, possibly through the integration of explainable artificial intelligence (XAI) techniques that can shed light on the decision-making process of these models without compromising their predictive power.⁶⁹

Potential Improvements and Emerging Trends

Despite them, there are some trends and potential for improving the effectiveness of energy load forecasting models. The area of incorporation of external factors is what is being proposed: weather conditions, socioeconomic factors, and real-time events. However, it turns out that these variables can have a major effect on energy demand without affecting forecasting models that rely heavily on historical load data.⁷⁰ Introducing the real-time data from the sensors, weather forecasting models, and social media trends, ML models can come up with more accurate and timely predictions as they inherently consider both predictable and unpredictable events.

Furthermore, it is also important that attempts for the integration of hybrid models, those that combine the strength of multiple ML technologies, should also be moving forward. Hybrid approaches, such as combining DL models with traditional ML models like RF or SVMs, can help improve model accuracy and robustness by leveraging the strengths of each individual model.⁷¹ In other words, DL models are great at processing huge amounts of data with intricate interrelationships, while quaint models have increased ability to interpret and run faster on a computer. By joining the use of these approaches, researchers can yet develop more accurate, simple, and readable forecasting systems.⁷²

Another promising direction is the application of federated learning, where ML models are trained

Challenge	Description	Aligned Future Direction
Data availability and quality	Incomplete, noisy, or nonstandardized datasets hinder model accuracy and generalizability	Data standardization, robust preprocessing, integration of real-time sources
Data heterogeneity	Varying formats, scales, and intervals from different sources make it hard to train unified models	Federated learning and unified data representation methods
Computational complexity	High computational cost for training and deployment, especially for DL models	Lightweight hybrid models, hardware-aware optimization
Model interpretability	Deep models act as black boxes, reducing trust and usability for decision-making	XAI integration, model simplification
Scalability across regions/systems	Models often trained on a small scale or local data may not generalize well	Federated learning, transfer learning
Forecasting under uncertainty	Limited adaptability to changing patterns due to events or policy shifts	RL, real-time adaptive models
Benchmarking and evaluation practices	Lack of standardized datasets and inconsistent use of evaluation metrics	Common benchmarking protocols, open-access repositories

across multiple decentralized devices or nodes without transferring raw data. This is useful for collaborative learning across regions or utility companies within the context of smart grids while maintaining the privacy of data and minimizing the storage to be centralized.⁷³ Furthermore, in circumstances where data sharing is restricted due to regulatory limitations, federated learning could prove very useful.

Finally, another emerging trend is to incorporate RL techniques into energy load forecasting. This is because RL models can continuously learn and adapt to an environment that is changing, and that is precisely the case for dynamic, real-time forecasting scenarios.⁶⁵ As energy consumption patterns change due to factors like economic shifts, population growth, or new energy policies, RL models could help adjust forecasting strategies on the fly, improving the accuracy of long-term predictions.

Limitations in Benchmarking and Evaluation Practices

Another key challenge in energy load forecasting research lies in the inconsistency and limitations of benchmarking practices. A significant number of studies rely on private or domain-specific datasets, which limits the comparability of results across different models and research groups. This lack of standard publicly available datasets leads to fragmented progress and makes it difficult to reproduce results or validate the effectiveness of newly proposed methods.⁶⁶ Moreover, the evaluation metrics used in various studies are not always uniform. While metrics like MAE, RMSE, and MAPE are common, their application without proper context—such as normalization, peak-hour analysis, or seasonal decomposition—can obscure the practical implications of model performance.⁷⁴ Some studies also report results on short forecasting horizons only, ignoring the long-term performance stability of models. Therefore, there is a pressing need for the development of standardized benchmarking protocols, unified datasets, and comprehensive evaluation frameworks that consider multiple forecasting horizons, variable load types, and model robustness under different operational scenarios (Table 11).

Conclusion

Energy load forecasting is an important task to improve the energy distribution and to increase the operational efficiency of smart grids. It has been shown that a wide array of ML techniques can be effectively applied in order to boost the forecasting accuracy to a point where better decisions can be made by energy providers, while at the same time enhancing the pursuit of the sustainability goal of energy consumption efficiency. Different aspects of the forecasting task, for instance, the length of the forecast horizon, the complexity of the data and computational resources, show that traditional ML models, DL architectures like LSTMs and GRUs, and ensemble methods all have different benefits.

Yet, there are several challenges that have not yet been overcome in order to fully unleash these models' potential. Nevertheless, despite these issues related to the data availability and quality, including sparse and noisy datasets, there has not been much progress in developing accurate and robust forecasting systems. As for the computational complexity of such models, especially of DL, which is quite advanced, it raises doubts about the possibility of their real-time implementation in a resource constrained environment. Furthermore, model interpretability is an open issue for the smart grid system stakeholders since the models need to explain and convince stakeholders of the predictions.

This study has several limitations. First, the review considered only peer-reviewed papers written in English, so relevant work published in other languages or in industry reports may have been missed. Second, most of the evidence is drawn from small-scale case studies focused on temperate-region grids, which may limit the generalizability of the findings to larger or more climate-diverse systems. Third, the analysis groups models by broad families rather than by specific architectural choices, which can hide finer factors that affect performance. Finally, the study does not quantify the economic impact of forecast errors, leaving the practical value of the reported accuracy gains open to further verification.

References

- 1 Cui X, Zhu J, Jia L, Wang J, Wu Y. A novel heat load prediction model of district heating system based on hybrid whale optimization algorithm (WOA) and CNN-LSTM with attention mechanism. *Energy*. 2024;312:133536. <https://doi.org/10.1016/j.energy.2024.133536>
- 2 Shrestha A, Mahmood A. Review of deep learning algorithms and architectures. *IEEE Access*. 2019;7:53040–65. <https://doi.org/10.1109/ACCESS.2019.2912200>
- 3 Mekonnen Y, Namuduri S, Burton L, Sarwat A, Bhansali S. Review—machine learning techniques in wireless sensor network based precision agriculture. *J Electrochem Soc*. 2020;167(3):037522. <https://doi.org/10.1149/2.0222003JES>
- 4 Agupugo CP, Ajayi AO, Nwanevu C, Oladipo SS. Policy and regulatory framework supporting renewable energy microgrids and energy storage systems. *Eng Sci Technol J*. 2024;5(8):2589–615. <https://doi.org/10.51594/estj.v5i8.1460>
- 5 Nuthakki S, Kulkarni CS, Kathiriyi S, Nuthakki Y. Artificial intelligence applications in natural gas industry: a literature review. *Int J Eng Adv Technol*. 2024;13(3):64–70. <https://doi.org/10.35940/ijeat.c4383.13030224>
- 6 Nuthakki S, Buddiga SKP, Koganti S. Exploring deep learning models for image recognition: a comparative review. *Signal Image Process An Int J*. 2024;15(3):1–10. <https://doi.org/10.5121/sipij.2024.15301>
- 7 Patil S, Awan KH, Arakeri G, Seneviratne CJ, Muddur N, Malik S, et al. Machine learning and its potential applications to the genomic study of head and neck cancer—a systematic review. *J Oral Pathol Med*. 2019;48(9):773–79. <https://doi.org/10.1111/jop.12854>
- 8 Nallore SS, Velumani A, Reddimasu V. Machine learning for early detection of neurodegenerative diseases. *Int J Res Publ Rev*. 2023;4(11):1894–909. <https://doi.org/10.55248/gengpi.4.1123.113123>
- 9 Gou J, Yu B, Maybank SJ, Tao D. Knowledge distillation: a survey. *Int J Comput Vis*. 2021;129(6):1789–819. <https://doi.org/10.1007/s11263-021-01453-z>
- 10 Manzoor MF, Farooq MS, Haseeb M, Farooq U, Khalid S, Abid A. Exploring the landscape of intrinsic plagiarism detection: benchmarks, techniques, evolution, and challenges. *IEEE Access*. 2023;11:140519–45. <https://doi.org/10.1109/ACCESS.2023.3338855>
- 11 Ishaq M, Abid A, Farooq MS, Manzoor MF, Farooq U, Abid K, et al. Advances in database systems education: methods, tools, curricula, and way forward. *Educ Inf Technol (Dordr)*. 2023;28(3):2681–725. <https://doi.org/10.1007/s10639-022-11293-0>
- 12 Dridi S, Machine V, Tree D, Forest R, Regression L. S l - a s l r ; 2021.
- 13 Alzubaidi L, Zhang J, Humaidi AJ, Al-Dujaili A, Duan Y, Al-Shamma O, et al. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *J Big Data*. 2021;8:53. <https://doi.org/10.1186/s40537-021-00444-8>
- 14 Sarker IH. Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions. *SN Comput Sci*. 2021;2(6):1–20. <https://doi.org/10.1007/s42979-021-00815-1>
- 15 Söderberg MJ, Meurling A. Feature selection in short-term load forecasting. *DIVA*; 2019.
- 16 Lahouar A, Ben Hadj Slama J. Random forests model for one day ahead load forecasting. *IREC 2015 6th Sixth international renewable energy congress*; 2015. <https://doi.org/10.1109/IREC.2015.7110975>
- 17 Sarker MT, Alam MJ, Ramasamy G, Uddin MN. Energy demand forecasting of remote areas using linear regression and inverse matrix analysis. *Int J Electr Comput Eng*. 2024;14(1):129–39. <https://doi.org/10.11591/ijece.v14i1.pp129-139>
- 18 Mahmud MA. Isolated area load forecasting using linear regression analysis: practical approach. *Energy Power Eng*. 2011;3(4):547–50. <https://doi.org/10.4236/epe.2011.34067>
- 19 Hambali M, Akinyemi A, Oladunjoye J, Yusuf N. Electric power load forecast using decision tree algorithms. *Comput Inf Syst Dev Informatics Allied Res J*. 2016;7(4):29–42.
- 20 Jahan IS, Snasel V, Misak S. Intelligent systems for power load forecasting: a study review. *Energies*. 2020;13(22):1–12. <https://doi.org/10.3390/en13226105>
- 21 Tian L, Noore A. A novel approach for short-term load forecasting using support vector machines. *Int J Neural Syst*. 2004;14(5):329–35. <https://doi.org/10.1142/S0129065704002078>
- 22 Li EYY, Fang T. Study of support vector machines for short-term load forecasting. *Proc. CSEE* 2003. https://doi.org/10.1142/9789812796769_0063
- 23 Dudek G. Short-term load forecasting using random forests. *Adv Intell Syst Comput*. 2015;323:821–28. https://doi.org/10.1007/978-3-319-11310-4_71
- 24 Gao J, Wang K, Kang X, Li H, Chen S. Ultra-short-term electricity load forecasting based on improved random forest algorithm. *AIP Adv*. 2023;13(6):065208. <https://doi.org/10.1063/5.0153550>
- 25 Oreshkin BN, Dudek G, Pelka P, Turkina E. N-BEATS neural network for mid-term electricity load forecasting. *Appl Energy*. 2021;293:1–33. <https://doi.org/10.1016/j.apenergy.2021.116918>
- 26 Gao R, Du L, Suganthan PN, Zhou Q, Yuen KF. Random vector functional link neural network based ensemble deep learning for short-term load forecasting. *Expert Syst Appl*. 2022;206:1–10. <https://doi.org/10.1016/j.eswa.2022.117784>
- 27 Bayram F, Aupke P, Ahmed BS, Kassler A, Theocharis A, Forsman J. DA-LSTM: a dynamic drift-adaptive learning framework for interval load forecasting with LSTM networks. *Eng Appl Artif Intell*. 2023;123:0–3. <https://doi.org/10.1016/j.engappai.2023.106480>
- 28 Lu N, Ouyang Q, Li Y, Zou C. Electrical load forecasting model using hybrid LSTM neural networks with online correction. *arXiv:2403.03898*; 2024:1–9.
- 29 Emshagin S, Halim WK, Kashef R. Short-term prediction of household electricity consumption using customized LSTM and GRU models. *arXiv:2212.08757*; 2022:11–30.
- 30 Dong M, Grumbach L. A hybrid distribution feeder long-term load forecasting method based on sequence prediction. *IEEE Trans Smart Grid*. 2020;11(1):470–82. <https://doi.org/10.1109/TSG.2019.2924183>
- 31 Ding Y, Wu D, He Y, Luo X, Deng S. Highly-accurate electricity load estimation via knowledge aggregation. *arXiv:2212.13913*; 2022:1–6.
- 32 Rubattu N, Maroni G, Corani G. Electricity load and peak forecasting: feature engineering, probabilistic LightGBM and temporal hierarchies. *Lect Notes Comput Sci (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*. 2023;14343:276–92. https://doi.org/10.1007/978-3-031-49896-1_18
- 33 Dube R, Gautam N, Banerjee A. Learning for interval prediction of electricity demand: a cluster-based bootstrapping approach. *arXiv:2309.01336*; 2023. <https://doi.org/10.2139/ssrn.4720434>
- 34 Olumba WC, Monday HN, Nneji GU, Agbonifo D, David GM, Umana ES, et al. BaggingSHAP: a novel ensemble approach for high-accuracy energy demand forecasting. 2024;198:14–28.
- 35 Zhang C, Peng Y, Chen H. A hybrid model based on random forest and LSTM for short-term electricity load forecasting. *IEEE Access*. 2020;8:119566–76.
- 36 Ma W, Liu X, Wang Y, Zhao Y. Hybrid model of random forest and LSTM for electricity load forecasting. *Energy Rep*. 2021;7:196–203.
- 37 Chai T, Draxler RR. Root mean square error (RMSE) or mean absolute error (MAE)? Arguments against avoiding RMSE in the literature. *Geosci Model Dev*. 2014;7(3):1247–50. <https://doi.org/10.5194/gmd-7-1247-2014>
- 38 Willmott CJ, Matsuura K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Clim Res*. 2005;30(1):79–82. <https://doi.org/10.3354/cr030079>
- 39 Hyndman RJ, Koehler AB. Another look at measures of forecast accuracy. *Int J Forecast*. 2005;22:679–88. <https://doi.org/10.1016/j.ijforecast.2006.03.001>
- 40 Hong T, Pinson P, Fan S. Global energy forecasting competition 2012. *Int J Forecast*. 2014;30(2):357–63. <https://doi.org/10.1016/j.ijforecast.2013.07.001>
- 41 Makridakis S. The M3-competition: results, conclusions and implications. *Int J Forecast*. 2000;16:451–76. [https://doi.org/10.1016/S0169-2070\(00\)00057-1](https://doi.org/10.1016/S0169-2070(00)00057-1)
- 42 Evans MG. On the asymmetry of g. *Psychol Rep*. 1999;85(3 Pt 2):1059–69. <https://doi.org/10.2466/pr0.1999.85.3f.1059>
- 43 Draper NR, Smith H. Applied regression analysis. 3rd edn. Wiley; 1998. <https://doi.org/10.1002/9781118625590>
- 44 Alexander C. Market models: a guide to financial data analysis. John Wiley Sons; 2001.
- 45 Hippert RC, Pedreira HS, Souza CE. Neural networks for short-term load forecasting: a review and evaluation. *IEEE Trans Power Syst*. 2001;16(1):44–55. <https://doi.org/10.1109/59.910780>

- 46 Wang C, Qin J, Chen Y. A short-term load forecasting model based on wavelet transform and a hybrid intelligent algorithm. *Sustainability*. 2019;11(3):685.
- 47 Liu J, Sun Y, Wu H, Zhang Z. A decision tree-based short-term load forecasting model for electricity markets. *Energy*. 2017;135:739–49.
- 48 Singh V, Pal SP, Sharma R. A decision tree based energy forecasting model using smart meter data. *Procedia Comput Sci*. 2020;167:2450–59.
- 49 Lahouar JBH, Slama A. Day-ahead load forecast using random forests and expert input selection. *Energy Convers Manag*. 2015;103:1040–51. <https://doi.org/10.1016/j.enconman.2015.07.041>
- 50 Deb KW, Zhang C, Yang F, Lee J, Shah SE. A review on time series forecasting techniques for building energy consumption. *Renew Sustain Energy Rev*. 2017;74:902–24. <https://doi.org/10.1016/j.rser.2017.02.085>
- 51 Kuster C, Rezgui Y, Mourshed M. Electrical load forecasting models: a critical systematic review. *Sustain Cities Soc*. 2017. <https://doi.org/10.1016/j.scs.2017.08.009>
- 52 Suganthi L, Samuel AA. Energy models for demand forecasting—a review. *Renew Sustain Energy Rev*. 2012;16:1223–40. <https://doi.org/10.1016/j.rser.2011.08.014>
- 53 Marino D, Amarasinghe K, Manic M. Building energy load forecasting using deep neural networks. *IEEE IECON*; 2016. <https://doi.org/10.1109/IECON.2016.7793413>
- 54 Kong W, Dong ZY, Jia Y, Hill DJ, Xu Y, Zhang Y. Short-term residential load forecasting based on LSTM recurrent neural network. *IEEE Transactions on Smart Grid*; 2017. p. 841–51. <https://doi.org/10.1109/TSG.2017.2753802>
- 55 Liu Y, Zhang Z, Lin X, Zhang Z. A hybrid model based on GRU for short-term load forecasting. *Appl Sci*. 2021.
- 56 Liang X, Hu Y, Li C, Xie Y. A novel GRU-based approach for energy load forecasting under multivariate temporal features. *Energy*. 2022.
- 57 Hong A, Pinson T, Fan S, Zareipour H, Troccoli P. Probabilistic energy forecasting: Global Energy Forecasting Competition 2014 and beyond. *Int J Forecast*. 2016;32:896–913. <https://doi.org/10.1016/j.ijforecast.2016.02.001>
- 58 Taieb SB, Hyndman RJ. A gradient boosting approach to the Kaggle load forecasting competition. *Int J Forecast*. 2014.
- 59 Lahouar A, Slama JBH. Hourly forecasting of Tunisian electricity demand using random forest. *Int J Electr Power Energy Syst*. 2015.
- 60 Zhou K, Yang S, Shen C. A review of electric load classification in smart grid environment. *Renew Sustain Energy Rev*. 2013;24:103–10. <https://doi.org/10.1016/j.rser.2013.03.023>
- 61 Tymoshenko K, Bonadiman D, Moschitti A. Learning to rank non-factoid answers: comment selection in web forums. *Int Conf Inf Knowl Manag Proc*. 2016:2049–52. <https://doi.org/10.1145/2983323.2983906>
- 62 Wrembel. *Data integration revitalized: from data warehouse through data lake to data mesh*. Springer; 2023. https://doi.org/10.1007/978-3-031-39847-6_1
- 63 Kelechi AH, Alsharif MH, Okpe J, Ezra PJ, Iorshase K, Atayero AA, et al. Artificial intelligence: an energy efficiency tool for enhanced high performance computing. *Symmetry (Basel)*. 2020;12(6):1029. <https://doi.org/10.3390/SYM12061029>
- 64 Brooks C, Thompson C. Predictive modelling in teaching and learning. In: *Handbook of learning analytics*; 2022. p. 29–37. <https://doi.org/10.18608/hla22.003>
- 65 Shiri FM, Perumal T, Mustapha N, Mohamed R. A comprehensive overview and comparative analysis on deep learning models: CNN, RNN, LSTM, GRU. *arXiv:2305.17473*; 2023.
- 66 Gao S, Huang Y, Zhang S, Han J, Wang G. Short-term runoff prediction with GRU and LSTM networks without requiring time step optimization during sample generation. *J Hydrol*. 2020;589:125188. <https://doi.org/10.1016/j.jhydrol.2020.125188>
- 67 Bommasani R, et al. *Language models are unsupervised multitask learners*; 2021.
- 68 Van Houdt G, Mosquera C, Nápoles G. A review on the long short-term memory model. *Artif Intell Rev*. 2020;53(8):5929–55. <https://doi.org/10.1007/s10462-020-09838-1>
- 69 Abdul-Jabbar SS, Farhan AK. *Data analytics and techniques*. Aro Sci J Koya Univ. 2022;10(2):45–55. <https://doi.org/10.14500/aro.10975>
- 70 Halverson LR, Graham CR, Spring KJ, Drysdale JS, Henrie CR. A thematic analysis of the most highly cited scholarship in the first decade of blended learning research. *Internet High Educ*. 2014;20:20–34. <https://doi.org/10.1016/j.iheduc.2013.09.004>
- 71 Stas M, Van Orshoven J, Dong Q, Heremans S, Zhang B. A comparison of machine learning algorithms for regional wheat yield prediction using NDVI time series of SPOT-VGT. 2016 5th International Conference on Agro-Geoinformatics; 2016. <https://doi.org/10.1109/Agro-Geoinformatics.2016.7577625>
- 72 Mendoza-Bernal J, González-Vidal A, Skarmeta AF. A convolutional neural network approach for image-based anomaly detection in smart agriculture. *Expert Syst Appl*. 2024;247:123210. <https://doi.org/10.1016/j.eswa.2024.123210>
- 73 Kumar V, Grag ML. Predictive analytics: a review of trends and techniques. *Int J Comput Appl*. 2018;182(1):31–7. <https://doi.org/10.5120/ijca2018917434>
- 74 Cui M. District heating load prediction algorithm based on bidirectional long short-term memory network model. *Energy*. 2022;254:124283. <https://doi.org/10.1016/j.energy.2022.124283>