

# Artificial Intelligence Techniques for Energy Forecasting

Sergey Kokin<sup>1</sup>, Begmurod Saidmurodov<sup>1</sup>, Stepan Dmitriev<sup>1</sup>  
<sup>1</sup>Department of Automated Electrical Systems, Ural Federal University, Russia

| Article Info                                                                                                                                                                                                                                         | ABSTRACT                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p><b>Article history:</b></p> <p>Received Feb 25, 2024<br/>Revised Apr 11, 2024<br/>Accepted May 15, 2024</p> <hr/> <p><b>Keywords:</b></p> <p>Energy forecasting<br/>Artificial intelligence<br/>Machine learning<br/>Survey<br/>Deep learning</p> | <p>Forecasting power consumption is a crucial aspect of managing cities and regions. Accurate forecasts ensure smooth and uninterrupted operation of consumer and industrial units. While the traditional methods of forecasting have been useful, the advent of big data has enabled the use machine learning techniques. In this paper, we discuss the applications of machine learning to power forecasting. We describe the technical details of the existing methods and highlight their strengths and weaknesses. We find that there is no single method that fits all scenarios. The optimal method depends on various factors including the characteristics of the data, the size of the data, and the forecasting horizon.</p> <p><i>This is an open access article under the <a href="#">CC BY-SA</a> license.</i></p> <div></div> |
| <p><b>Corresponding Author:</b></p> <p>Begmurod Saidmurodov<br/>Department of Automated Electrical Systems, Ural Federal University<br/>Mira 19, Yekaterinburg, Russia<br/>Email: begmurod1999@mail.ru</p>                                           |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |

## 1. INTRODUCTION

The effective management of power consumption is a crucial responsibility for both public and private utilities, as electricity serves as the foundation for nearly all aspects of modern society. Accurate forecasting of energy demand is a complex undertaking that involves numerous variables with a high degree of stochasticity. While traditional autoregressive and statistical models have been useful in this regard, they fail to leverage the availability of big data, which has emerged in recent times. The increased capacity to gather and store data allows for more comprehensive analysis and deeper insights.

Enter machine learning and deep learning algorithms, which harness the power of big data to develop accurate models that learn from data. Intelligent systems have been increasingly utilized for power demand forecasting, and machine learning-based forecasting algorithms have become more prevalent than traditional models in the forecasting literature. Furthermore, hybrid methods that combine statistical and deep learning techniques have also demonstrated promising results.

Given the significance and popularity of power demand forecasting, the amount of literature dedicated to this topic has grown substantially over the past decade. To navigate the existing body of research, this paper proposes a survey of the current state of the field. The primary goal is to analyze the existing literature related to power demand forecasting using machine learning techniques. Specifically, the major machine learning algorithms employed in power demand forecasting will be identified, and the technical details, achievements, advantages, and disadvantages of each approach will be discussed. Additionally, guidelines for comparing and selecting the optimal machine learning algorithms will be provided.

The paper is structured as follows: Section 2 provides an overview of the existing machine learning methods employed for power demand forecasting. Section 3 delves into the details of each method. Finally, Section 4 concludes the paper.

## 2. ARTIFICIAL INTELLIGENCE AND ENERGY FORECASTING

A wide range of machine learning approaches have been employed in power demand forecasting, including regression models, support vector machines (SVMs), artificial neural networks (ANNs), deep learning techniques, and ensemble learning models. Each method possesses unique strengths and weaknesses, and their performance varies depending on the characteristics of the specific problem at hand. While regression models tend to perform better in scenarios with limited data availability, deep learning models excel in problems where large amounts of data are available. Therefore, a proper understanding of each method is necessary to identify the optimal approach for a given problem.

Regression models are among the most commonly used tools in power demand forecasting. They provide a quantitative measure of the relationship between the dependent variable (power demand) and one or more independent variables (such as temperature, time of day, and so on). The two key types are linear and non-linear regression models. Linear regression models are the simplest, assuming a straight-line relationship between the variables. In contrast, non-linear regression models accommodate more complex relationships and are generally more accurate, although computationally more intensive. Multiple regression models, which involve more than one independent variable, are often used for power demand forecasting due to the multiple factors influencing power demand.

SVMs have been widely utilized in power demand forecasting due to their unique capabilities in addressing non-linear, high-dimensional, and sparse data issues. SVMs operate by mapping input vectors to a higher-dimensional space, where a hyperplane can be used to perform linear regression or classification. This characteristic allows SVMs to capture complex relationships in the data, making them a powerful tool in this domain.

ANNs are a subset of machine learning models loosely inspired by the human brain. They are capable of learning from non-linear and complex data, making them well-suited for power demand forecasting. Feed-forward neural networks (FFNNs) and radial basis function networks (RBFNs) are the two most popular types used in this field. FFNNs are well-suited to model complex non-linear relationships, while RBFNs offer advantages in terms of speed and simplicity of training. The major drawback of ANNs is their "black box" nature, which can make the interpretation of results challenging.

Deep learning is a subfield of machine learning that focuses on algorithms inspired by the structure and function of the brain, called artificial neural networks. Deep learning was originally conceived as an extension of ANNs with multiple layers of neurons for better feature extraction. However, it has evolved into many different architectures. Specifically, Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks have been particularly effective in power demand forecasting. RNNs have loops that allow information to be carried across nodes while processing sequential data, giving them a crucial advantage in power demand forecasting, as they can use historical data (previous power demands) to influence the network's forecasts. However, RNNs have difficulties handling long sequences due to the "vanishing gradient" problem. To address this issue, researchers developed LSTM, a type of RNN that overcomes this problem by introducing a mechanism called "gates." These gates allow the network to selectively remember or forget information, thereby effectively handling longer sequences. LSTM networks have shown exceptional performance in predicting power demand due to their ability to learn from long-term dependencies, which is a common scenario in power demand data.

Ensemble methods combine predictions from multiple machine learning algorithms to deliver superior predictive performance. The basic principle behind the ensemble approach is to leverage the different strengths of various algorithms, thereby mitigating their individual weaknesses. Bagging and boosting are two popular ensemble methods used in power demand forecasting. Bagging improves prediction accuracy by reducing the variance of the estimate, while boosting reduces bias.

Each of these machine learning techniques offers unique advantages and challenges. The choice of method typically depends on the specifics of the power demand forecasting problem at hand, such as the nature of the data, the forecasting horizon (short, medium, or long-term), and the computational resources available. It is also worth noting that hybrid models, which combine two or more of these methods, are becoming increasingly popular, as they offer the potential to capture the strengths of multiple methods while mitigating their weaknesses. Machine learning, with its capability to handle large, complex datasets and uncover hidden patterns, offers a powerful set of tools for power demand forecasting and is set to play an increasingly important role in this area.

## 3. REVIEW OF ARTIFICIAL INTELLIGENCE IN ENERGY FORECASTING

### 3.1. Regression

Regression techniques are widely adopted in the field of machine learning for forecasting applications, including predictions of power demand. These methods aim to estimate a continuous target variable (in this case, power demand) based on one or more input features. This section delves into the details of some commonly used regression-based models, namely linear regression, polynomial regression, and ridge regression. In linear regression, the dependent variable is modeled as a linear combination of the independent variables. The fundamental equation of a simple linear regression model is:

$$Y = \beta_0 + \beta_1 X + \epsilon \quad (1)$$

where:  $Y$  is the dependent variable,  $X$  is the independent variable,  $\beta_0$  is the constant,  $\beta_1$  is the slope of the regression line, and  $\epsilon$  is the error term.

Linear regression models have been compared with neural networks for photovoltaic power forecasting tasks, considering different input methods [4]. Researchers in [12] employed linear regression to forecast power demand in Spain, utilizing time series data of previous electricity consumption and other factors such as weather data. Their model demonstrated competitive performance for short-term forecasting.

Polynomial regression is an extension of linear regression, where the model incorporates not only linear terms but also polynomial functions of the independent variables. A second-degree polynomial regression model can be expressed as:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_k X_k + \epsilon \quad (2)$$

Polynomial regression was explored in [21], where the authors forecasted power demand in Alice Springs, a geographically solar energy-rich area in Australia, using various environmental parameters. In another study [26], polynomial regression was applied for long-term load forecasting in Taiwan's electricity system, incorporating factors such as economic indicators, population growth, and technological advancements. The polynomial regression model outperformed linear regression in terms of accuracy for long-term forecasting.

Ridge regression is a technique employed to address multicollinearity (high correlation among independent variables) in multiple linear regression data. It introduces a penalty term to the sum of squared coefficients in the cost function. The key equation for Ridge Regression is:

$$\min \|Y - X\beta\|_2 + \lambda \|\beta\|_2 \quad (3)$$

where  $\lambda$  is the hyper parameter that controls the magnitude of the penalty term,  $\|\cdot\|$  denotes the Euclidean norm, and  $\beta$  are the parameters to be learned.

In [20], the authors utilized ridge regression to forecast the intermittency of Photovoltaic (PV) energy in smart grids. Another study by researchers in [3] applied ridge regression to short-term load forecasting in the Saudi electricity system, incorporating weather data, previous load patterns, and calendar variables. Ridge regression improved the model's performance in the presence of highly correlated input variable.

These regression-based models have been widely adopted in power demand forecasting due to their interpretability and ability to capture linear and non-linear relationships between the input features and the target variable. However, their performance might be limited in scenarios where the underlying relationships are highly complex or involve intricate patterns. In such cases, more advanced machine learning techniques, such as neural networks or ensemble methods, may prove more suitable.

### 3.2. Support Vector Machines

Support Vector Machines (SVMs) are a category of machine learning algorithms utilized for both classification and regression problems. Their primary objective is to identify the optimal hyperplane that most effectively separates the data in a high-dimensional feature space, making them particularly suitable for intricate non-linear challenges.

When employed for regression (Support Vector Regression, or SVR), the SVM aims to fit the best line within a predefined or specified boundary termed an  $\epsilon$ -tube to approximate the actual data points. The key equation for Support Vector Regression can be expressed as follows:

$$\min_{w,b,\xi,\xi^*} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \quad (4)$$

$$y_i - w^T \phi(x_i) - b \leq \epsilon + \xi_i \quad (5)$$

$$w^T \phi(x_i) + b - y_i \leq \epsilon + \xi_i^* \quad (6)$$

where:  $y_i$  are the target values,  $x_i$  are the input feature vectors,  $w$  is the normal vector to the hyperplane,  $b$  is the bias,  $\phi(x_i)$  is the function mapping  $x_i$  to a higher dimensional space,  $\epsilon$  defines the margin of tolerance where no penalty is given to errors,  $\xi_i$  and  $\xi_i^*$  are nonnegative slack variables that allow for deviations outside the  $\epsilon$ -margin,  $C$  is a regularization parameter which determines the trade-off between the flatness of the hyperplane and tolerance to deviations larger than  $\epsilon$ .

SVMs have been extensively applied in various research studies. An approach for photovoltaic power forecasting based on an enhanced support vector machine with improved ant colony optimization was proposed in [25], demonstrating superior results. In [34], the authors utilized a single SVM model to predict daily power demand in China, considering historical power consumption data, weather information, and day-of-week information. Similarly, the researchers in [7] employed a standard SVM model to forecast hourly electricity load in Taiwan, incorporating input data on weather conditions, day of the week, and time of day. Their SVM model outperformed several competing forecasting models.

In [18], the authors combined SVM, WLS, and PSO for short-term load forecasting in Brazil, where PSO was used to optimize the SVM parameters. The combined model exhibited better performance than a standalone SVM model. A study by researchers in Iran [5] employed a hybrid approach of wavelet transform (WT) and SVM for short-term load forecasting. The WT decomposed the load time series into sub-series, which were then forecasted using SVMs.

The authors in [9] utilized Least Squares SVM to construct a time series model for short-term wind power forecasting, achieving accurate results. Least Squares SVM was also employed in the study by [40], where an LS-SVM model was used for day-ahead electricity load forecasting in Belgium, incorporating past load data, temperature data, and calendar variables. The LS-SVM outperformed traditional ARIMA models in this study.

Conversely, Kernel SVM was employed in [30] for benchmark analysis of day-ahead solar power forecasting techniques using weather predictions. Similarly, Kernel SVM was utilized in [15], where the researchers implemented a Kernel SVM model for forecasting half-hourly electricity demand in Australia, considering predictors such as past demand data, temperature data, and time variables. The Kernel SVM model could capture complex non-linear relationships in the data. STOPHERE such as neural networks or ensemble methods, may prove more suitable.

### 3.3. Artificial Neural Networks

Artificial Neural Networks (ANNs) are a class of machine learning models that mimic the structure and functionality of biological neural networks found in the human brain. These models are composed of interconnected nodes or "neurons" organized into layers that transmit and process information. ANNs excel at identifying intricate patterns and relationships within data, making them widely applicable for tasks such as power demand forecasting.

The most basic form of an ANN is the Feed-Forward Neural Network (FFNN), also referred to as the Multi-layer Perceptron (MLP). This architecture consists of an input layer, one or more hidden layers, and an output layer. Each node in a layer is connected to all nodes in the subsequent layer, with each connection having an associated weight. A node calculates the weighted sum of its inputs, applies a bias, and then passes the result through an activation function to generate its output.

The equation for a neuron's output in an FFNN can be represented as follows

$$y = f\left(\sum_{j=1}^n w_j x_j + b\right) \quad (7)$$

where  $y$  is the output of the neuron,  $f$  is the activation function,  $w_j$  are the associated weights,  $x_j$  are the inputs to the neuron, and  $b$  is the bias.

Training an ANN involves adjusting the weights and biases of the network to minimize the discrepancy between the network's predictions and the actual values. This process is typically carried out using techniques such as backpropagation and optimization algorithms like stochastic gradient descent.

ANNs have been widely employed for power demand forecasting applications. One active area of research has been forecasting solar power generation [27]. Another clean energy source, wind power, has also garnered significant attention from forecasting researchers [24]. MLP has been used as the primary model for

short-term load forecasting in Brazil, utilizing historical load data, temperature data, and calendar variables as inputs. Additionally, researchers have successfully developed probabilistic wind power forecasting models using radial basis function neural networks [28], demonstrating superior performance compared to benchmark models.

### 3.4. Deep Learning

Deep learning is a subset of machine learning that utilizes artificial neural networks (ANN) with multiple layers, designed to replicate the structure and function of the human brain. These models are proficient in identifying patterns within extensive and intricate datasets. Over the years, the number of layers in these models has expanded, reaching thousands in the latest versions. The primary feature of deep learning models is their capability to learn features for tasks such as classification and regression. In power demand forecasting, these models can handle the complexity and variability of power demand data, leading to precise forecasts [6].

One of the prominent deep learning models used in forecasting is the Long Short-Term Memory (LSTM) network. LSTMs, a type of Recurrent Neural Network (RNN), are designed to overcome the long-term dependency issue, enabling them to remember inputs over long durations. Traditional RNNs use past inputs to calculate future inputs through a hidden state mechanism, but repeated matrix multiplication during back-propagation often leads to the vanishing gradient problem. LSTM was developed to solve this problem.

The LSTM architecture, introduced in 1998 [14], includes a cell state and three gates: input, forget, and output gates. The cell state runs through the entire chain with minimal linear interactions, while the gates manage the addition and removal of information from the cell state.

The key equations of the LSTM cell are the following: Forget gate:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (8)$$

Input gate:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (9)$$

Cell state:

$$C_t = f_t * C_{t-1} + i_t \cdot \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (10)$$

Output gate:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (11)$$

$$h_t = o_t \cdot \tanh(C_t) \quad (12)$$

where  $x_t$  is the input at time step  $t$ ,  $h_t$  is the output at time step  $t$ ,  $C_t$  is the cell state at time step  $t$ ,  $f_t$ ,  $i_t$ , and  $o_t$  are the forget gate, input gate, and output gate, respectively,  $W$  and  $b$  are weights and biases, respectively,  $\sigma$  is the sigmoid function, and  $\tanh$  is the hyperbolic tangent function.

Numerous studies have applied LSTM to power demand forecasting. The research in [33] introduced a day-ahead photovoltaic (PV) power forecasting method using an LSTM-RNN model with time correlation modification under a partial daily pattern prediction framework. Another innovative approach for long-term solar photovoltaic power forecasting was presented in [29], where an LSTM with the Nadam optimizer was used, showing promising results in India. The study in [16] utilized LSTM for short-term residential power demand forecasting, incorporating historical power demand data, weather data, and time-related features. The LSTM model outperformed traditional machine learning models. In [31], a CNN-LSTM model was employed to forecast power demand in China, demonstrating the effectiveness of combining time series data with text data for enhanced forecasting performance. Similarly, [32] proposed an LSTM-CNN model that first extracts temporal features using LSTM and then spatial features using a convolutional neural network (CNN). The results showed that this hybrid model outperformed single prediction models.

### 3.5. Convolutional Neural Networks

The Convolutional neural networks (CNNs), a subset of deep learning models, are predominantly employed in image processing, but they also find utility in time series analysis, such as power demand

forecasting. These networks consist of an input layer, several hidden layers (including convolutional, pooling, and fully connected layers), and an output layer.

The convolution operation in the first layer of a CNN is defined as follows:

$$z^{[l]} = w^{[l]} * a^{[l-1]} + b^{[l]} \quad (13)$$

$$a^{[l]} = \sigma(z^{[l]}) \quad (14)$$

where  $w^{[l]}$  is the kernel or filter in the  $l^{\text{th}}$  layer,  $a^{[l-1]}$  is the output activations from the previous layer,  $b^{[l]}$  is the bias in the  $l^{\text{th}}$  layer,  $*$  represents the convolution operation,  $\sigma$  is the activation function,  $z^{[l]}$  and  $a^{[l]}$  are the pre-activation and post-activation values, respectively, in the  $l^{\text{th}}$  layer.

Following the convolutional layers, pooling layers are frequently utilized to diminish the spatial dimensions of the representation. This reduction lowers the number of parameters and computational load within the network, thereby aiding in the prevention of overfitting. The pooling layer functions independently across each depth slice of the input, adjusting its spatial dimensions.

CNNs are notably effective for power demand forecasting, particularly when integrated with Long Short-Term Memory (LSTM) networks. For instance, a hybrid technique employing a combination of Variational Mode Decomposition (VMD), CNN, and Gated Recurrent Unit (GRU) was implemented for short-term wind power forecasting, accounting for spatiotemporal characteristics [34]. Another approach merged CNN and LSTM to predict wind power generation [19]. Researchers in [17] introduced a hybrid model featuring both CNN and LSTM to forecast stable power generation. In this model, CNN classified weather conditions, while LSTM captured power generation patterns based on these conditions. Similarly, the study in [2] applied a CNN-LSTM architecture to predict power demand in Rabat, Morocco, yielding promising outcomes.

### 3.6. Ensemble Learning

Ensemble learning techniques combine multiple machine learning models to enhance prediction accuracy. The core principle is that a collection of weak learners can collectively form a robust model. Among the various ensemble methods, the most prevalent are bagging, boosting, and stacking.

Bagging involves training the same algorithm multiple times using different subsets sampled from the training data, with the Random Forest algorithm being the most notable example. In bagging, the final decision is made by averaging the outputs for regression tasks or through majority voting for classification tasks. Boosting, on the other hand, trains models sequentially, with each new model correcting the errors of its predecessor. The final decision is derived from a weighted sum of all models' decisions, with Gradient Boosting Machine (GBM) being a well-known boosting algorithm. Stacking combines multiple classification or regression models using a meta-classifier or meta-regressor. The base models are trained on the full training set, and their outputs are used as inputs for the meta-model. Some ensemble methods also incorporate feature selection.

Recently, ensemble learning methods have gained traction in power demand forecasting. For instance, a calibrated ensemble of model chains was introduced in [22] for probabilistic photovoltaic power forecasting. An adaptive stochastic configuration network ensemble for short-term photovoltaic power forecasting was proposed in [11]. In [10], Random Forest was employed for multivariable response surface in short-term load forecasting, integrating the Random Forest model with the mean generating function model to yield promising results. An interpretable probabilistic model using natural gradient boosting for short-term solar power forecasting was presented in [23], featuring a two-stage probabilistic forecasting framework that delivers accurate, reliable, and transparent point forecasts and prediction intervals. A stacking approach for forecasting photovoltaic (PV) power generation was proposed in [1], utilizing a stacked ensemble algorithm (Stack-ETR) to predict PV output power one day ahead. This method employed three machine learning algorithms—random forest regressor (RFR), extreme gradient boosting (XGBoost), and adaptive boosting (AdaBoost)—as base models.

## 4. COMPARATIVE ANALYSIS

Selecting a machine learning algorithm for power forecasting depends on multiple factors such as the characteristics of the data, available computational resources, the need for model interpretability, and the desired level of accuracy. Each algorithm comes with its own set of strengths and weaknesses.

Table 1. Sampling methods used in the study.

| Algorithm              | Strength                                                                                                                                                             | Weakness                                                                                                                                                              |
|------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Linear Regression      | Easy to comprehend and interpret, quick in both training and prediction, and performs effectively with numerical and well-prepared datasets.                         | Assumes a linear relationship between input and output, is sensitive to outliers, and struggles with irrelevant features.                                             |
| Support Vector Machine | Performs well in high-dimensional environments, capable of capturing non-linear relationships, and resistant to overfitting, especially in high-dimensional contexts | Sensitive to kernel parameter choices, computationally heavy, thus less ideal for large datasets, and difficult to interpret                                          |
| Deep Learning          | Capable of modeling complex non-linear relationships, automatically extracts features in deep learning, scalable, and efficient with large-scale datasets.           | Needs substantial data, computationally expensive, model parameters are hard to interpret, and prone to overfitting without proper regularization and network design. |
| Ensemble Methods       | Can enhance accuracy and stability by integrating multiple models, mitigates overfitting, and handles unbalanced and missing data.                                   | Lack of model interpretability, computationally intensive, and requires careful tuning.                                                                               |

It is important to remember that there is no universal algorithm for power forecasting. The most appropriate model typically depends on the specific problem, including data volume and quality, available computational resources, and particular forecasting goals. Therefore, it is often advisable to experiment with multiple models and choose the one that performs best for the given problem.

## 5. CONCLUSION

This paper explored different machine learning techniques for power consumption forecasting. We provided technical details and discussed the advantages and disadvantages of each method. There is no universally superior method; the best choice depends on the specifics of the dataset. Deep learning models are often preferred when large datasets are available, while linear regression might be more suitable for smaller datasets. Ensemble methods have gained traction recently because they integrate several models, thereby reducing individual model weaknesses and enhancing their strengths.

## REFERENCES

- [1] Zhang, H., & Zhu, T. (2022). Stacking model for photovoltaic-power-generation prediction. *Sustainability*, 14(9), 5669..
- [2] Elizabeth Michael, N., Mishra, M., Hasan, S., & Al-Durra, A. (2022). Short-term solar power predicting model based on multi-step CNN stacked LSTM technique. *Energies*, 15(6), 2150.
- [3] Hammad, M. A., Jereb, B., Rosi, B., & Dragan, D. (2020). Methods and models for electric load forecasting: a comprehensive review. *Logist. Sustain. Transp.*, 11(1), 51-76..
- [4] Huang, C. J., & Kuo, P. H. (2019). Multiple-input deep convolutional neural network model for short-term photovoltaic power forecasting. *IEEE access*, 7, 74822-74834.
- [5] Chitsaz, H., Shaker, H., Zareipour, H., Wood, D., & Amjady, N. (2015). Short-term electricity load forecasting of buildings in microgrids. *Energy and Buildings*, 99, 50-60..
- [6] Wazirali, R., Yaghoubi, E., Abujazar, M. S. S., Ahmad, R., & Vakili, A. H. (2023). State-of-the-art review on energy and load forecasting in microgrids using artificial neural networks, machine learning, and deep learning techniques. *Electric power systems research*, 225, 109792..
- [7] Han, D., Chan, L., & Zhu, N. (2007). Flood forecasting using support vector machines. *Journal of hydroinformatics*, 9(4), 267-276.
- [8] Khan, F., Enzmann, F., & Kersten, M. (2016). Multi-phase classification by a least-squares support vector machine approach in tomography images of geological samples. *Solid Earth*, 7(2), 481-492.
- [9] Ding, M., Zhou, H., Xie, H., Wu, M., Liu, K. Z., Nakanishi, Y., & Yokoyama, R. (2021). A time series model based on hybrid-kernel least-squares support vector machine for short-term wind power forecasting. *ISA transactions*, 108, 58-68.
- [10] Magalhães, B., Bento, P., Pombo, J., Calado, M. D. R., & Mariano, S. (2024). Short-Term Load Forecasting Based on Optimized Random Forest and Optimal Feature Selection. *Energies*, 17(8), 1926.
- [11] Zhou, X., Ao, Y., Wang, X., Guo, X., & Dai, W. (2023). Learning with privileged information for short-term photovoltaic power forecasting using stochastic configuration network. *Information Sciences*, 619, 834-848..
- [12] Hernandez, L., Baladron, C., Aguiar, J. M., Carro, B., Sanchez-Esguevillas, A. J., Lloret, J., & Massana, J. (2014). A survey on electric power demand forecasting: Future trends in smart grids, microgrids and smart buildings. *IEEE Communications Surveys & Tutorials*, 16(3), 1460-1495..
- [13] Baliyan, A., Gaurav, K., & Mishra, S. K. (2015). A review of short term load forecasting using artificial neural network models. *Procedia Computer Science*, 48, 121-125..
- [14] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735-1780.
- [15] Mokilane, P. (2018). Density forecasting for long-term electricity demand in South Africa using quantile regression. *South African Journal of Economic and Management Sciences*, 21(1), 1-14..
- [16] Jiao, R., Zhang, T., Jiang, Y., & He, H. (2018). Short-term non-residential load forecasting based on multiple sequences LSTM recurrent neural network. *IEEE Access*, 6, 59438-59448..
- [17] Al-Ali, E. M., Hajji, Y., Said, Y., Hleili, M., Alanzi, A. M., Laatar, A. H., & Atri, M. (2023). Solar energy production forecasting based on a hybrid CNN-LSTM-transformer model. *Mathematics*, 11(3), 676..

- [18] Moradzadeh, A., Zakeri, S., Shoaran, M., Mohammadi-Ivatloo, B., & Mohammadi, F. (2020). Short-term load forecasting of microgrid via hybrid support vector regression and long short-term memory algorithms. *Sustainability*, 12(17), 7076.
- [19] Agga, A., Abbou, A., Labbadi, M., El Houm, Y., & Ali, I. H. O. (2022). CNN-LSTM: An efficient hybrid deep learning architecture for predicting short-term photovoltaic power production. *Electric Power Systems Research*, 208, 107908.
- [20] Kim, H., Park, S., Park, H. J., Son, H. G., & Kim, S. (2023). Solar radiation forecasting based on the hybrid CNN-CatBoost model. *IEEE Access*, 11, 13492-13500..
- [21] Mahmud, K., Azam, S., Karim, A., Zobaed, S., Shanmugam, B., & Mathur, D. (2021). Machine learning based PV power generation forecasting in alic springs. *IEEE Access*, 9, 46117-46128.
- [22] Mayer, M. J., & Yang, D. (2023). Pairing ensemble numerical weather prediction with ensemble physical model chain for probabilistic photovoltaic power forecasting. *Renewable and Sustainable Energy Reviews*, 175, 113171.
- [23] Wang, J., Li, P., Ran, R., Che, Y., & Zhou, Y. (2018). A short-term photovoltaic power prediction model based on the gradient boost decision tree. *Applied Sciences*, 8(5), 689..
- [24] Zhao, L., Nazir, M. S., Nazir, H. M. J., & Abdalla, A. N. (2022). A review on proliferation of artificial intelligence in wind energy forecasting and instrumentation management. *Environmental Science and Pollution Research*, 29(29), 43690-43709..
- [25] Gao, B., Yang, H., Lin, H. C., Wang, Z., Zhang, W., & Li, H. (2022). A hybrid improved whale optimization algorithm with support vector machine for short-term photovoltaic power prediction. *Applied Artificial Intelligence*, 36(1), 2014187..
- [26] Banday, U. J., & Aneja, R. (2020). Renewable and non-renewable energy consumption, economic growth and carbon emission in BRICS: evidence from bootstrap panel causality. *International Journal of Energy Sector Management*, 14(1), 248-260..
- [27] El-Amarty, N., Marzouq, M., El Fadili, H., Bennani, S. D., & Ruano, A. (2023). A comprehensive review of solar irradiation estimation and forecasting using artificial neural networks: data, models and trends. *Environmental Science and Pollution Research*, 30(3), 5407-5439..
- [28] Lu, H. J., & Chang, G. W. (2018, August). Wind power forecast by using improved radial basis function neural network. In 2018 IEEE Power & Energy Society General Meeting (PESGM) (pp. 1-5). IEEE.
- [29] Salman, D., Direkoglu, C., Kusaf, M., & Fahrioglu, M. (2024). Hybrid deep learning models for time series forecasting of solar power. *Neural Computing and Applications*, 1-18.
- [30] Nespoli, A., Ogliari, E., Leva, S., Massi Pavan, A., Mellit, A., Lughi, V., & Dolara, A. (2019). Day-ahead photovoltaic forecasting: A comparison of the most effective techniques. *Energies*, 12(9), 1621.
- [31] Kong, Z., Zhang, C., Lv, H., Xiong, F., & Fu, Z. (2020). Multimodal feature extraction and fusion deep neural networks for short-term load forecasting. *IEEE access*, 8, 185373-185383.
- [32] Wang, Y., Chen, Y., Liu, H., Ma, X., Su, X., & Liu, Q. (2021, March). Day-ahead photovoltaic power forecasting using convolutional-LSTM networks. In 2021 3rd Asia Energy and Electrical Engineering Symposium (AEEES) (pp. 917-921). IEEE.
- [33] Akhter, M. N., Mekhilef, S., Mokhlis, H., Almohaimed, Z. M., Muhammad, M. A., Khairuddin, A. S. M., ... & Hussain, M. M. (2022). An hour-ahead PV power forecasting method based on an RNN-LSTM model for three different PV plants. *Energies*, 15(6), 2243.
- [34] Chen, N., Sun, H., Zhang, Q., & Li, S. (2022). A short-term wind speed forecasting model based on EMD/CEEMD and ARIMA-SVM algorithms. *Applied Sciences*, 12(12), 6085.