Journal of Quality Engineering and Production Optimization

Volume 10, Issue 1, Winter & Spring 2025

http://jqepo.shahed.ac.ir

Research Paper

DOI: 10.22070/jqepo.2025.20698.1303

Smart Power Generation in Combined Cycle Power Plants: Machine Learning Models for Power Prediction and Neural Network-Driven Input Optimization

Hani Gilani [1], Hadi Sahebi [1*]

[1] *School of Industrial Engineering, Iran University of Science & Technology, Tehran, Iran*

**** Corresponding Author:** Hadi Sahebi (Email: hadi_sahebi@ iust.ac.ir)*

Abstract *–Fossil fuel-based power generation remains a core component of many energy systems. However, growing electricity demand influenced by government policies, economic development, and societal behavior can challenge the stability of power supply networks. This study focuses on improving the productivity of combined-cycle power plants using machine learning techniques to better match electricity generation with consumption while minimizing inefficiencies. Several algorithms, including Bagging decision tree, Random Forest, Gradient Boosting, and XGBoost, are implemented to predict energy output. A parameter tuning process is conducted to enhance model accuracy using negative root mean square error as the evaluation metric. The results indicate that Gradient Boosting achieves the highest accuracy, with a mean absolute error of 2.215449. Additionally, a neural network model is developed to optimize key operational parameters such as temperature, ambient pressure, and humidity. This model integrates a Quasi-Newton-based activation function to improve adaptability and precision. The identified optimal values are a temperature of 19.38°C, exhaust vacuum of 25.43 cm Hg, ambient pressure of 1021.4 millibars, relative humidity of 60.8 percent, and an energy output of 462.09 megawatts. Temperature is recognized as the most significant influencing factor. Using real-world data from the Zavareh Combined Cycle Power Plant, the study offers practical insights for enhancing energy sector performance.*

Keywords*– Artificial Intelligence, Disruption, Electricity Generation, Neural Network, Supply Chain Management.*

I. INTRODUCTION

The energy sector is a vital component of society, often described as a supply chain that spans from power generation in plants to consumption by end-users. It encompasses various operations, including generation, transmission, distribution, and consumption. Efficient management of these processes is essential to meet end-user energy demands effectively. Therefore, energy security in societies is closely linked to the efficient functioning of this supply chain. The energy supply chain can include networks for electricity, gas, petroleum, and crude oil, all of which are essential for meeting people's daily energy needs (Attari et al., 2022; Gilani et al., 2025).

Manuscript Received: 25-Oct-2024 & Revised:15-Jan-2025 & Accepted: 20-Mar-2025
ISSN: 2423-3781

Among these networks, electricity plays a major role in the energy mix of many societies, powering lighting, cooling, and various electrical devices essential to every household. The electricity supply network generally consists of power generation plants, transmission centers, distributors, and retailers. The efficient operation of each of these facilities is crucial for meeting the electricity demands of the affected area. Within the electricity supply network, power generation technologies include fossil fuel-based systems, renewable energy sources, and nuclear technologies (Nagurney and Matsypura, 2007; Aghajani et al., 2025). Among these, fossil-fuel-based systems, particularly combined-cycle power plants (CCPPs), are commonly used in developing countries.

High energy consumption can significantly disrupt the balance of the electricity supply network in societies. This imbalance may stem from government policies, economic strategies, or, in some cases, social behaviors (Alipoor et al., 2024). The inability to meet electricity demand and ensure a stable supply in CCPPs can lead to imbalances in the electricity supply network, a challenge observed in some developing countries (Zhang et al., 2019; Eskandari et al., 2024). Addressing this imbalance in the electricity supply network requires enhancing productivity to maximize the network's capacity and transition to a more responsive system. CCPPs, in particular, hold significant potential for productivity improvements, as many currently operate with low efficiency due to thermal waste (Darmawan et al., 2025; Ardestani et al., 2025; Gilani et al., 2023).

Artificial Intelligence (AI) has become an indispensable part of modern life, revolutionizing how humans interact with technology and address daily challenges. From healthcare to education, transportation, communication, and energy, AI is transforming industries by offering innovative solutions, enhancing efficiency, and personalizing experiences (Vijayalakshmi et al., 2023; Gilani and Sahebi, 2024). This potential can also be harnessed in the energy sector to efficiently develop innovative solutions for addressing disruptions.

AI is revolutionizing the energy sector, particularly in CCPPs, by enhancing productivity and operational efficiency through advanced machine learning (ML) and neural network technologies. ML algorithms are employed to predict electricity outputs with high accuracy, enabling power plants to align generation with demand and reduce inefficiencies (Kaewprapha et al., 2022; Fattahi et al., 2020). Simultaneously, neural networks optimize critical operational parameters such as temperature, humidity, and ambient pressure, ensuring the plant operates at peak performance under varying conditions. These AI-driven solutions not only improve energy conversion efficiency and reduce emissions but also enhance the plant's adaptability to dynamic environmental and operational challenges. Furthermore, this application can be extended to manage disruptions, such as energy supply chain imbalances, by providing real-time insights and predictive analytics to balance supply and demand, mitigate risks, and maintain grid stability. By integrating AI, CCPPs can achieve greater reliability, sustainability, and resilience in an increasingly complex energy landscape.

The present study offers the following key contributions:

- To enable precise performance assessment of CCPPs using real-time data, this study deploys Internet of Things (IoT) sensors throughout the plant to continuously monitor and analyze critical operational parameters, including temperature, exhaust vacuum, ambient pressure, and relative humidity.
- Given the strong dependence of the proposed machine learning algorithms on input parameters and their sensitivity to parameter variations, this study introduces a parameter tuning procedure designed to optimize algorithm performance. The effectiveness of this approach is evaluated using the negative root mean square error (NRMSE) as the performance metric.
- To accurately predict electricity output in a CCPP, this study employs a suite of advanced machine learning algorithms, including Bagging decision tree, Random Forest, Gradient Boosting, and XGBoost. The performance of these models is systematically compared to identify the most effective approach for optimizing power generation in response to fluctuating demand.
- This study proposes a novel method to optimize key input parameters of CCPPs, including temperature, exhaust vacuum, ambient pressure, and relative humidity, by developing a neural network model designed to

improve operational efficiency. The model incorporates a Quasi-Newton-based activation function that enhances its adaptability and predictive accuracy under dynamic operating conditions.

The paper is structured as follows: the second section presents a literature review, the third section outlines the methodology, the fourth section discusses the results, and the fifth section provides the conclusion.

II. LITERATURE REVIEW

Today, most studies on predictive models in electricity networks focus on renewable energy-based systems. In this context, the first subsection highlights the most prominent research in this field.

A. Application of machine learning methods in renewable energy-based systems

Recent studies on machine learning (ML) models for predicting electricity generation from photovoltaic (PV) systems highlight the effectiveness of various techniques. Bayesian regularized neural networks achieved the highest accuracy with R^2 of 99.99% (Chahboun and Maaroufi, 2021). Hybrid models, such as those combining an improved bird swarm algorithm with extreme learning machines, also demonstrated strong performance across different weather conditions (Wu et al., 2021). Support Vector Machines (SVM), artificial neural networks (ANN), and deep learning models were widely used, with SVM showing the best results in some cases, achieving R^2=0.9921 (Ağbulut et al., 2020). Other models, such as ensemble learning approaches and recurrent neural networks, were also effective, particularly in short-term PV power forecasting (Yan et al., 2023; Dairi et al., 2020; Essam et al., 2022).

In the domain of wind power prediction, hybrid and ensemble models have been dominant. The combination of three-level decomposition, kernel extreme learning machines, and improved Grey Wolf Optimization achieved R^2=0.9922, significantly improving forecasting accuracy (Han and Tong, 2020). Other methods, including deep residual networks and graph neural networks, successfully captured spatiotemporal correlations in wind turbine data (Li, 2022). Extreme learning machines (ELM) and recurrent neural networks (RNN) were also popular, particularly for short-term forecasting, due to their ability to handle fluctuations in wind power (An et al., 2022; Özen and Deniz, 2022). Meanwhile, for hydropower generation, Gaussian process regression (GPR) and ensemble models, such as random forests, proved to be effective, with GPR achieving a correlation coefficient of 0.92 (Ekanayake et al., 2021; Mohammad-Rafiei et al., 2025).

The integration of AI into CCPPs has gained significant attention in recent years due to its potential to enhance productivity, operational efficiency, and sustainability. CCPPs, which utilize both gas and steam turbines to maximize energy output, are complex systems that require precise control and optimization of multiple parameters. AI, particularly through ML and neural networks, has shown significant promise in addressing these challenges by enabling predictive analytics, real-time optimization, and proactive disruption management. The literature on CCPPs has been reviewed using two methodologies: traditional machine learning techniques and neural networks.

B. Application of traditional machine learning methods in CCPP-based systems

In the literature, neural networks are widely used to predict the performance of energy systems because they can model complex, nonlinear relationships in operational data. The predictive performance of machine learning models varies across algorithms, making it crucial to determine the most suitable model for a given task. In a study by Rasheed (2023), the performance of Deep Learning was evaluated against Random Forest, Gradient Boosting, and Multi-Layer Perceptron models using RMSE and MAE as metrics. The results indicated that Deep Learning achieved superior accuracy, with a 35% reduction in MAE and a 48% decrease in RMSE compared to the other approaches. Bistline and Merrick (2020) developed an open-source capacity planning model for power plants to assist stakeholders in evaluating investment, profitability, plant lifespan, and environmental impact. They identified data quality issues in emerging markets, which affected reliability. To address this, they employed Linear Regression and K-Nearest Neighbors (KNN) to estimate unknown parameters by learning from a known market (USA) and applying the models to an unknown

region (Canada). Their study concluded that KNN provided more accurate predictions than Linear Regression.

Hundi and Shahsavari (2020) developed a machine learning-based model to predict CCPP power generation without relying on thermodynamic equations. Using Support Vector Machines, Random Forest Regression, Linear Regression, and Multi-Layer Perceptron (MLP), they trained models on five years of data, with 75% for training and the rest for testing. Key input variables included humidity, exhaust vacuum pressure, atmospheric pressure, and ambient temperature. They reported that the Random Forest model delivered the highest performance, attaining an R^2 of 95.9%, a MAE of 2.4 MW, and a RMSE of 3.5 MW. Similarly, Tufekci (2014) introduced a machine learning-based method for predicting the full-load electric power output of a combined cycle power plant, offering an alternative to conventional thermodynamic modeling. Using six years of data, she analyzed 15 regression algorithms with four input variables: air humidity, ambient temperature, exhaust pressure, and atmospheric pressure. The study found that the Bagging Rep Tree (BREP) meta-learning algorithm was the most accurate, achieving an RMSE of 3.787 and an MAE of 2.818 (Danesh et al., 2021).

C. Application of neural networks in CCPP-based systems

Hydropower, supplying 60% of global renewable energy, requires accurate capacity prediction. Condemi et al. (2021) modeled hydropower production in Northern Italy using meteorological and climatic data, along with reservoir water volume. They applied ANN algorithms (MLP, ELP, SVR) with Gaussian, Polynomial, and Linear kernels, incorporating Principal Component Analysis (PCA) for feature reduction. Results showed MLP as the most accurate, achieving an RMSE of 0.2593 and MAE of 0.2128 TWh. Ali et al. (2021) developed a power prediction model for a Waste Heat Recovery (WHR) system at a cement plant in Pakistan using a Feed Forward Back-Propagation Neural Network (FF-BPNN). The training process utilized steam pressure, inlet steam temperature, and mass flow rate as input variables. A thermodynamic analysis using the Engineering Equation Solver (EES) estimated the system's thermal efficiency at 19.75% and power output at 10.06 MW. The BPNN model accurately predicted power with a mean squared error of 0.283 using 10 hidden neurons. Rashid et al. (2015) used Particle Swarm Optimization (PSO) to train a Feed Forward Neural Network (FFNN) for predicting the average hourly power output of a CCPP. The model incorporated ambient temperature, atmospheric pressure, relative humidity, and vacuum as key input parameters. Trained with 10 hidden neurons over 50 iterations, the model achieved a Mean Squared Error (MSE) of 0.0001019 for training and 0.0055 for testing data.

Adams et al. (2020) developed a machine learning model combining deep learning and least-squares Support Vector Machine (SVM) to predict Sulphur Oxides (SOx) and Nitrogen Oxides (NOx) emissions from coal combustion in a power plant. Using KEPCO's commercially available data, the Deep Neural Network (DNN) improved prediction accuracy by 39.24% for SOx and 26.58% for NOx, while the Least Squares Support Vector Machine (LSSVM) further increased accuracy by 74.01% for SOx and 26.51% for Nox. Akdemir (2016) developed a Multi-Layer Perceptron (MLP) model for predicting hourly electric power generation in a CCPP using an Artificial Neural Network (ANN). The dataset, which included four variables and 9,685 features, was analyzed using a two-fold cross-validation to ensure reliability, and the model achieved a Mean Squared Error (MSE) of 3.176 and an R^2 of 0.96675.

Recent advancements in machine learning have significantly enhanced the predictive capabilities and operational efficiency of CCPPs. Yi et al. (2023) introduced a hybrid model integrating transformer encoders with deep neural networks (DNNs) to accurately forecast power generation, demonstrating the potential of attention mechanisms in handling complex temporal dependencies. Asghar et al. (2023) focused on sustainable CCPP operations, employing AI-based prediction models to align energy output with environmental and operational goals. Similarly, Ntantis and Xezonakis (2024) proposed an innovative machine learning approach aimed at optimizing electric power prediction, contributing to improved control strategies under varying load conditions. Saeed et al. (2023) presented a hybrid framework combining recurrent neural networks (RNNs) with the Waterwheel Plant Algorithm for optimizing power output predictions, achieving enhanced accuracy and convergence. Zhang et al. (2024) further advanced this line of research by integrating machine learning techniques with the Hunger Games Search (HGS) algorithm, offering a robust

solution for electricity load forecasting in CCPPs. Collectively, these studies underscore the growing importance of intelligent algorithms in managing and optimizing power plant performance under dynamic and uncertain operating environments.

Despite recent advancements, several critical research gaps remain unaddressed in the context of CCPPs:

- Limited focus on energy supply chain disruptions: While many studies emphasize predictive modeling and optimization within power plants, there is a noticeable lack of research addressing the broader implications of supply chain disruptions—particularly regarding how fluctuations in electricity output affect the reliability and resilience of energy delivery systems.
- Insufficient attention to electricity market imbalances: Imbalances between electricity supply and demand can lead to significant consequences, including increased operational costs, grid instability, and adverse effects on social welfare. Existing literature does not adequately explore strategies to mitigate these imbalances through intelligent, plant-level interventions.
- Underutilization of AI for productivity enhancement: Although artificial intelligence techniques like machine learning and deep learning are gaining traction in power output prediction, their full potential remains underexploited. Specifically, many models lack integration with real-time operational data, optimization of operational inputs, and the prediction of power generation, limiting their adaptability and precision in dynamic environments.
- Need for comprehensive input optimization through IoT integration and intelligent modeling: Advanced AI models depend on accurate and timely input data to function effectively. While some studies have used static datasets, few have integrated real-time data acquisition from Internet of Things (IoT) sensors to continuously monitor critical operational parameters, such as temperature, ambient pressure, exhaust vacuum, and relative humidity. Moreover, the optimization of these inputs remains an underexplored area. This study addresses this gap by employing a neural network model that incorporates a Quasi-Newton-based activation function, enabling more effective tuning of inputs. This approach significantly enhances the model's learning efficiency, adaptability to changing plant conditions, and overall predictive performance.

III. PROBLEM AND METHODOLOGY

The electric grid, or electric supply chain, consists of several critical facilities and operations. First, electricity is generated in CCPPs. The produced electricity is routed to transmission substations for conversion to high-voltage levels suitable for long-distance transmission. It is then delivered to distribution substations, where the voltage is reduced for local supply to end-users. Finally, power lines deliver electricity to end consumers. In the first stage, IoT sensors are deployed across a CCPP to collect and transmit real-time operational data. These sensors report critical operational parameters, including temperature, exhaust vacuum, ambient pressure, and relative humidity, within the CCPP. Figure 1 illustrates the electric grid and the strategic placement of IoT sensors for monitoring and reporting critical operational parameters.

Imbalances in the electric supply chain can be addressed by enhancing productivity in CCPPs. To achieve this, we identify key operational parameters in these plants and leverage AI technologies, including machine learning and neural networks. Ensemble-based methods such as Bagging, Random Forest, Gradient Boosting, and XGBoost have worked well when handling high-dimensional, multivariate data and overfitting through aggregating, while at the same time providing high interpretability of feature importance. The Neural Network (multi-layer perceptron) was incorporated to leverage its capacity to approximate highly nonlinear transformations between the input and output variables, which best applies to modeling CCPPs thermodynamic and operating complexity. Through comparison of a number of algorithms with complementary strengths, the study gives a closer evaluation and makes it possible to determine the optimum model for credible power prediction in CCPPs under real operating conditions.

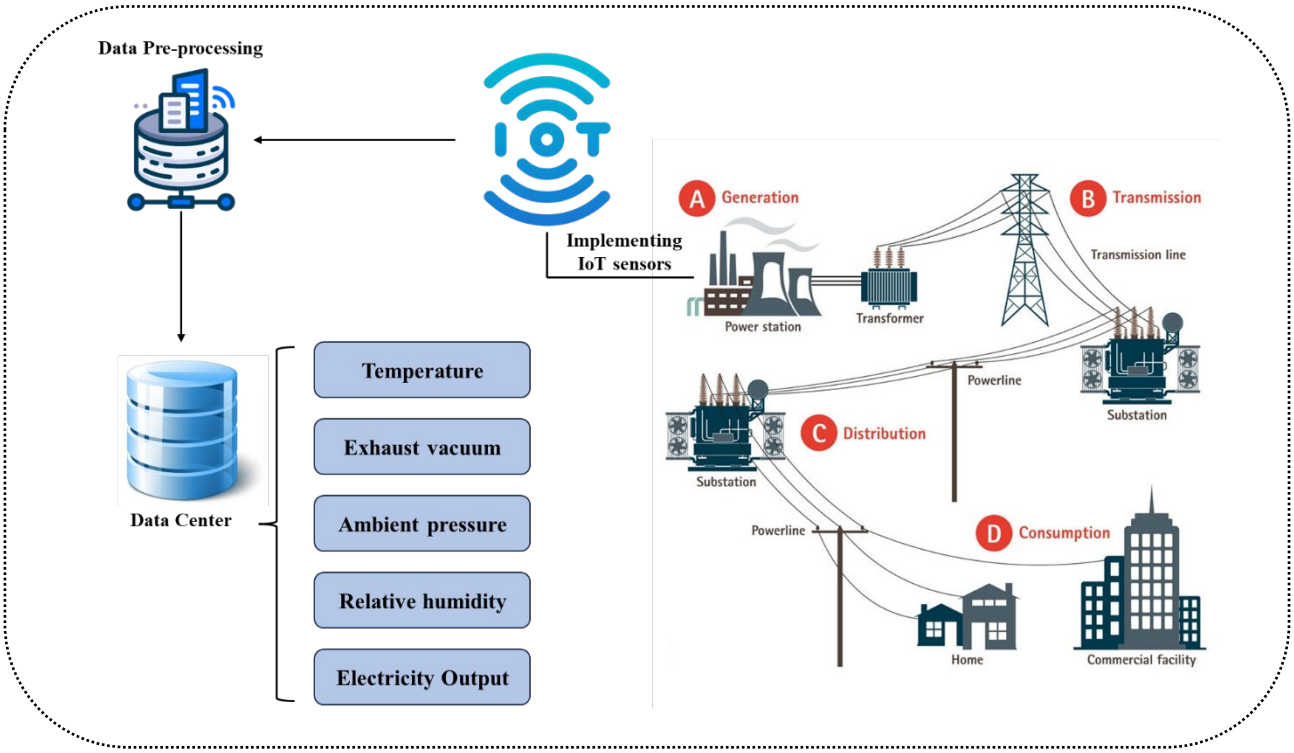


Fig. 1. Electricity grid network and data collection procedure

In this study, power generation prediction for CCPPs is investigated through the application of various machine learning algorithms, including Bagging decision trees, Random Forest, Gradient Boosting, XGBoost, and a Neural Network based on a multi-layer perceptron (MLP). The predictive modeling is performed using a dataset that incorporates key input features known to influence CCPP performance, namely, temperature, exhaust vacuum, ambient pressure, and relative humidity. Beyond prediction, the study also aims to enhance operational productivity by optimizing these critical input parameters. This is achieved through a neural network model employing backpropagation within a multi-layer perceptron architecture, integrated with a Quasi-Newton-based activation function to improve convergence and adaptability. Such optimization not only supports more accurate output prediction but also increases thermodynamic and operational efficiency in CCPPs by enabling real-time adjustments that reduce energy losses and improve overall system performance. The following section provides a detailed overview of the machine learning models and neural network techniques utilized in the analysis.

A. Bagging Decision Tree Algorithm

Bagging decision tree algorithm (Bootstrap Aggregating) is an ensemble learning method that improves the stability and accuracy of decision trees by reducing variance (Sutton, 2005). It works by training multiple decision trees on different bootstrap samples of the dataset and averaging their predictions (for regression) or using majority voting (for classification).

Algorithm Steps:

1. Generate B bootstrap samples (D_1, D_2, …, D_B) from the original dataset D.
2. Train a decision tree $h_b(x)$ on each bootstrap sample.

Aggregate predictions for regression by averaging: $\hat{y} = \frac{1}{B}\sum_{b=1}^{B} h_b(x)$

Bagging decision tree algorithm reduces overfitting and improves predictive performance, especially for high-variance models like decision trees.

B. Random Forest algorithm

The Random Forest algorithm is an extension of the Bagging Decision Tree algorithm that further enhances performance by introducing feature randomness (Breiman, 2001). It builds multiple decision trees using bootstrapped datasets and randomly selects a subset of features for each tree split, reducing correlation between trees and improving generalization.

Algorithm Steps:

1. Generate B bootstrap samples (D_1, D_2, …, D_B) from the original dataset D.
2. For each tree, at each split, randomly select a subset of m features from the total M features (where $m < M$).
3. Train a decision tree $h_b(x)$ on each bootstrap sample using the selected features.
4. Aggregate predictions for regression by using averaging:

$$\hat{y} = \frac{1}{B}\sum_{b=1}^{B} h_b(x) \tag{1}$$

Random Forest reduces overfitting, handles high-dimensional data, and maintains strong accuracy even with missing values. Its ensemble approach enhances generalization, making it robust to noise and effective for both classification and regression. Additionally, it provides feature importance insights for better interpretability.

C. Gradient Boosting algorithm

Gradient Boosting is an ensemble learning method that builds a strong predictive model by sequentially training decision trees, where each tree corrects the errors of its predecessor (Friedman, 2001). It optimizes a loss function by minimizing residual errors using gradient descent.

Algorithm Steps:

1. Start with a weak learner, typically a single decision tree:

$$F_0(x) = \underset{\gamma}{\operatorname{argmin}} \sum_{i=1}^{N} L(y_i, \gamma) \tag{2}$$

Where $L(y_i, \gamma)$ is the loss function.
For $m = 1$ to M (number of trees):

2. Compute residuals (negative gradient of loss function):

$$r_i^{(m)} = -\frac{\partial L(y_i, F_{m-1}(x_i))}{\partial F_{m-1}(x_i)} \tag{3}$$

Fit a new decision tree $h_m(x)$ to the residuals.
Update the model:

$$F_m(x) = F_{m-1}(x) + \lambda h_m(x) \tag{4}$$

Where λ is the learning rate.

3. Final Prediction:

$$\hat{y} = F_M(x) \tag{5}$$

Gradient Boosting iteratively improves predictions by correcting residual errors using gradient descent. It supports various loss functions, making it highly flexible for different tasks. While powerful, it requires careful tuning of parameters like learning rate, subsampling, and tree depth to prevent overfitting.

D. Extreme Gradient Boosting algorithm

Extreme Gradient Boosting (XGBoost) is an optimized version of the Gradient Boosting algorithm designed for efficiency, scalability, and performance (Chen and Guestrin, 2016). It introduces regularization techniques, handles missing data, and leverages parallel processing for faster computations.

Algorithm Steps:

1. XGBoost minimizes a regularized loss function:

$$\mathcal{L} = \sum_{i=1}^{N} L(y_i, \hat{y}_i) + \sum_{m=1}^{M} \Omega(f_m) \tag{6}$$

Where $L(y_i, \hat{y}_i)$ is the loss function (e.g., squared error for regression, log-loss for classification), and $\Omega(f_m)$ is a regularization term to prevent overfitting.

2. Each new tree is added to correct the previous model's residuals:

$$F_m(x) = F_{m-1}(x) + f_m(x) \tag{7}$$

The new tree $f_m(x)$ is trained using the gradient of the loss function:

$$g_i = -\frac{\partial L(y_i, \hat{y}_i)}{\partial \hat{y}_i}, \quad h_i = -\frac{\partial^2 L(y_i, \hat{y}_i)}{\partial \hat{y}_i^2} \tag{8}$$

3. XGBoost calculates the optimal split using the gain function:

$$Gain = \frac{1}{2}\left[\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{(G_L + G_R)^2}{H_L + H_R + \lambda}\right] - \gamma \tag{9}$$

Where G and H are the sums of first and second-order gradients, λ is the regularization term, and γ is a pruning parameter to prevent unnecessary splits.

4. Regularization in XGBoost includes $L1$ (α) and $L2$ (λ) penalties to control model complexity and prevent overfitting. Additionally, tree pruning removes splits with low information gain using a threshold parameter (γ), ensuring a more efficient and interpretable model.

XGBoost offers strong regularization to prevent overfitting, efficiently handles missing data, and leverages parallel processing for faster training. Its tree pruning mechanism improves model interpretability, while optimized split selection enhances predictive performance.

E. Neural Network with multi-layer perceptron for prediction and optimization

A Multi-Layer Perceptron (MLP) is a type of feedforward artificial neural network composed of several interconnected layers of neurons (Rumelhart et al., 1986). It is widely used for prediction and optimization tasks by learning complex patterns in data through weight adjustments using backpropagation.

1. Architecture of MLP: an MLP consists of:

Input Layer: Receives features $X = [x_1, x_2, \dots, x_n]$
Hidden Layers: Perform transformations using activation functions.
Output Layer: Produces final predictions.
Each neuron in a layer receives inputs from the previous layer and applies the transformation:

$$z_j^{(l)} = \sum_{i=1}^{N} w_{ij}^{(l)} a_j^{(l-1)} + b_j^{(l)} \tag{10}$$

where:
$z_j^{(l)}$ is the weighted sum of inputs for neuron j in layer l
$w_{ij}^{(l)}$ are the weights between neurons.
$b_j^{(l)}$ is the bias term.
$a_j^{(l)}$ is the activation function output.
The activation function introduces non-linearity:

$$a_j^{(l)} = f(z_j^{(l)}) \tag{11}$$

Common activation functions include:
Linear: $f(x) = x$
ReLU: $f(x) = \max(0, x)$
Sigmoid: $f(x) = \frac{1}{1+e^{-x}}$
Tanh: $f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$

2. Forward Propagation for Prediction: In the forward propagation step, the input features are passed through the network layers to compute the predicted output. The activation functions in each layer introduce non-linearity, enabling the network to model complex relationships.

For a given input X, the forward propagation is performed as follows:

- Compute activations for the first hidden layer:

$$Z_1 = W_1 X + b_1 \tag{12}$$

$$A_1 = \tanh(Z_1) \tag{13}$$

- Compute activations for the second hidden layer (linear activation):

$$Z_2 = W_2 A_1 + b_2 \tag{14}$$

$$A_2 = Z_2 \tag{15}$$

(since the activation function is linear)

- Compute activations for the output layer

$$Z_3 = W_3 A_2 + b_3 \tag{16}$$

$$\hat{Y} = Z_3 \tag{17}$$

- Compute the Normalized Squared Error (NSE) Loss with L2 Regularization:

$$NSE = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{y_i - \hat{y}_i}{y_i}\right)^2 \tag{18}$$

$$L_{Total} = NSE + \lambda \sum \|W\|^2 \tag{19}$$

Where:
y_i is the actual target value,
$\hat{y}_i$ is the predicted value,
λ is the regularization weight,
And $\|W\|^2$ represents the sum of squared parameters for all layers to prevent overfitting.

This loss function ensures that the model minimizes prediction errors while incorporating L2 regularization to improve generalization and stability.

3. Backpropagation for Optimization

To optimize the network, a loss function $L(y_i, \hat{y}_i)$ is minimized, such as Mean Squared Error (MSE) for regression:

$$L_{Total} = NSE + \lambda \sum \|W\|^2 \tag{20}$$

Instead of using standard gradient descent, the Quasi-Newton method is applied, which improves convergence by approximating the Hessian matrix without computing second-order derivatives explicitly (Robitaille et al., 1996).

- Compute the gradient of the loss function:

$$g_k = \nabla L(w_k) \tag{21}$$

- Approximate the inverse Hessian matrix B_k^{-1} iteratively:

$$B_{k+1} = B_k + \Delta B_k \tag{22}$$

Where ΔB_k is updated based on observed gradient differences.

- Update the weights using:

$$w_{k+1} = w_k - B_k^{-1} g_k \tag{23}$$

- Common Quasi-Newton updates include:

BFGS update:

$$B_{k+1} = B_k + \frac{y_k y_k^T}{y_k^T s_k} - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} \tag{24}$$

DFP update:

$$B_{k+1} = B_k + \frac{y_k y_k^T}{y_k^T s_k} - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} \tag{25}$$

Where $s_k = w_{k+1} - w_k$ and $y_k = g_{k+1} - g_k$ represent weight and gradient differences, respectively.

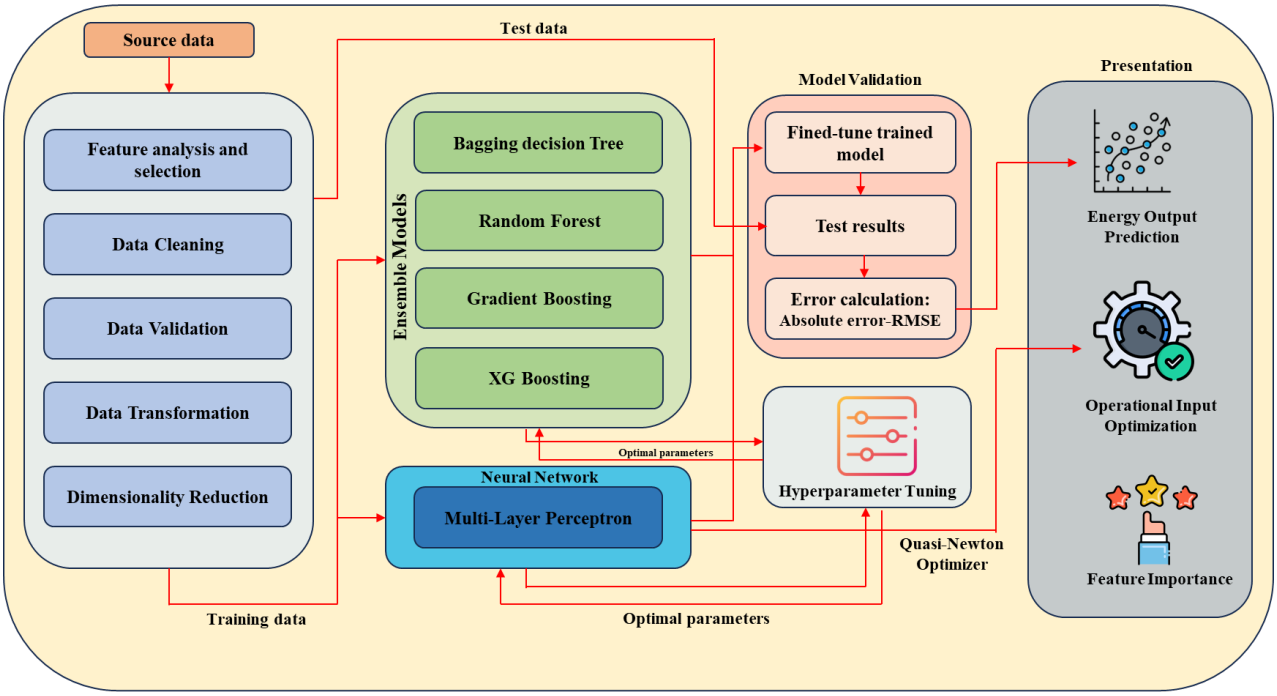


Fig. 2. Graphical representation of the proposed methodology and the implemented models

This approach improves convergence speed compared to standard gradient descent and is especially effective in optimizing neural networks for prediction and complex decision-making tasks.

As previously mentioned, the proposed AI methods, including machine learning algorithms and neural networks, are applied to the collected data from IoT sensors deployed in CCPPs. These approaches are designed to forecast electricity generation for future demand management while optimizing operational parameters to boost overall productivity. Fig. 2 provides a graphical representation of the proposed methodology and implemented models. The case study and numerical results are presented in the following section.

IV. CASE STUDY AND MODEL CONFIGURATION

This section applies the proposed models to real data from the Zavareh Combined Cycle Power Plant, and outlines the dataset, key features, and preprocessing steps. The configuration and tuning of each machine learning model are described. The goal is to evaluate model performance and identify the most accurate approach for predicting power output and optimizing operations.

A. Case Description

The Zavareh Combined Cycle Power Plant, located in Isfahan, Iran, is a key facility contributing to the region's electricity generation. This power plant operates using a combined cycle configuration, which integrates gas and steam turbines to enhance efficiency and reduce fuel consumption. By utilizing waste heat from the gas turbine to produce steam for the steam turbine, the plant achieves higher energy conversion efficiency compared to conventional thermal power plants. As part of Iran's energy infrastructure, Zavareh plays a crucial role in meeting the growing electricity demands of both residential and industrial sectors in the region.

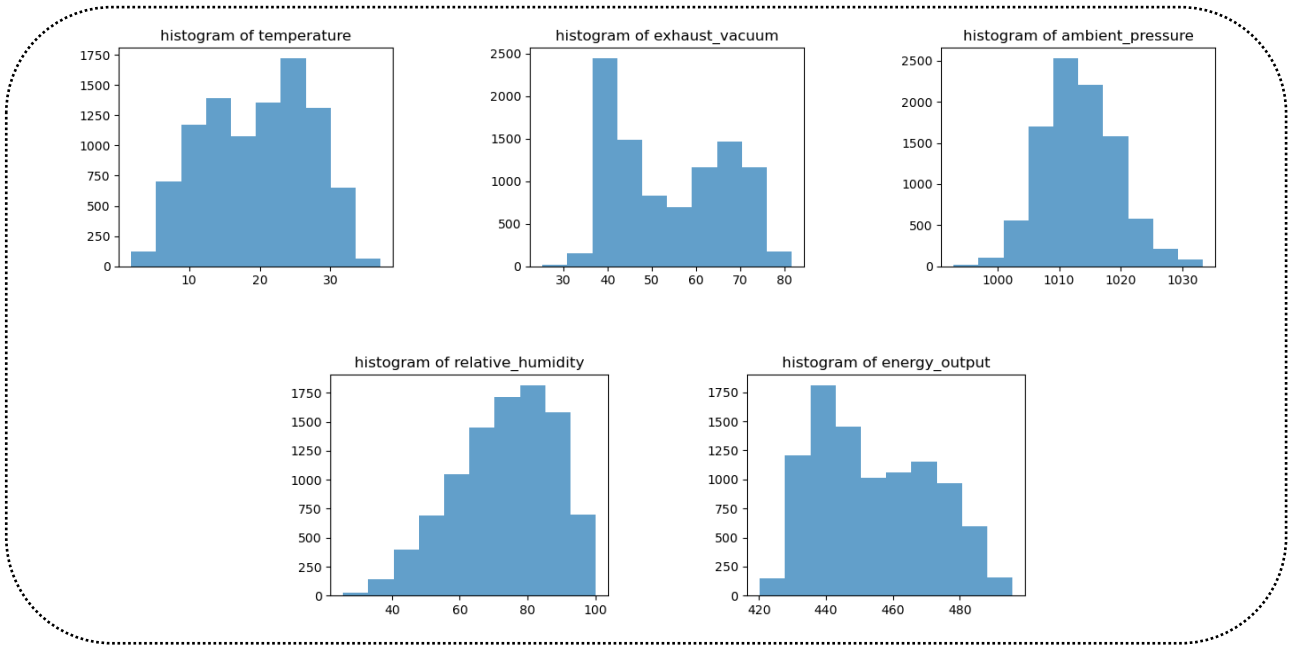


Fig. 3. Data understanding by histogram of variables

To further enhance operational efficiency and reliability, the Zavareh power plant has incorporated IoT-based monitoring systems, enabling real-time data collection on critical parameters such as temperature, pressure, fuel consumption, and power output. These data-driven insights allow for predictive maintenance, fault detection, and performance optimization through advanced AI techniques. In this case study, machine learning algorithms and neural networks are employed to analyze sensor data, predict electricity output for demand planning, and optimize key operational factors, ultimately improving productivity and sustainability in power generation.

B. Data Collection and Preprocessing

As discussed in Section 3, real operational data from the Zavareh CCPP is collected using IoT sensors, capturing key parameters such as temperature, exhaust vacuum, ambient pressure, relative humidity, and electricity output. Data monitoring is conducted over a one-month period, resulting in a dataset of 9,568 recorded data points for analysis after data manipulation and cleaning.

The data description is summarized in Table I, detailing the data count, mean, standard deviation, minimum, maximum, and quartiles. To enhance data understanding, we present histograms of the variables in Fig. 3, where the horizontal axis represents variable values and the vertical axis indicates frequency. This visualization provides insights into the distribution of each variable.

In the next step, we analyze the Pearson correlation in Fig. 4 to examine relationships between variables. This analysis helps identify how each independent variable influences the response variable. Despite being useful for detecting linear relationships, it may not capture complex nonlinear dependencies.

Table I. Data description

	Temperature	Exhaust vacuum	Ambient pressure	Relative humidity	Energy output
Count	9568	9568	9568	9568	9568
Mean	19.65123	54.3058	1013.259	73.30898	454.365
Std	7.452473	12.70789	5.938784	14.60027	17.06699
Min	1.81	25.36	992.89	25.56	420.26
25%	13.51	41.74	1009.1	63.3275	439.75
50%	20.345	52.08	1012.94	74.975	451.55
75%	25.72	66.54	1017.26	84.83	468.43
Max	37.11	81.56	1033.3	100.16	495.76

However, it provides an initial insight for data analysts to better understand variable interactions. Fig. 4 demonstrates that the temperature and exhaust vacuum exhibit a strong relationship with the target variable, which is the energy output.

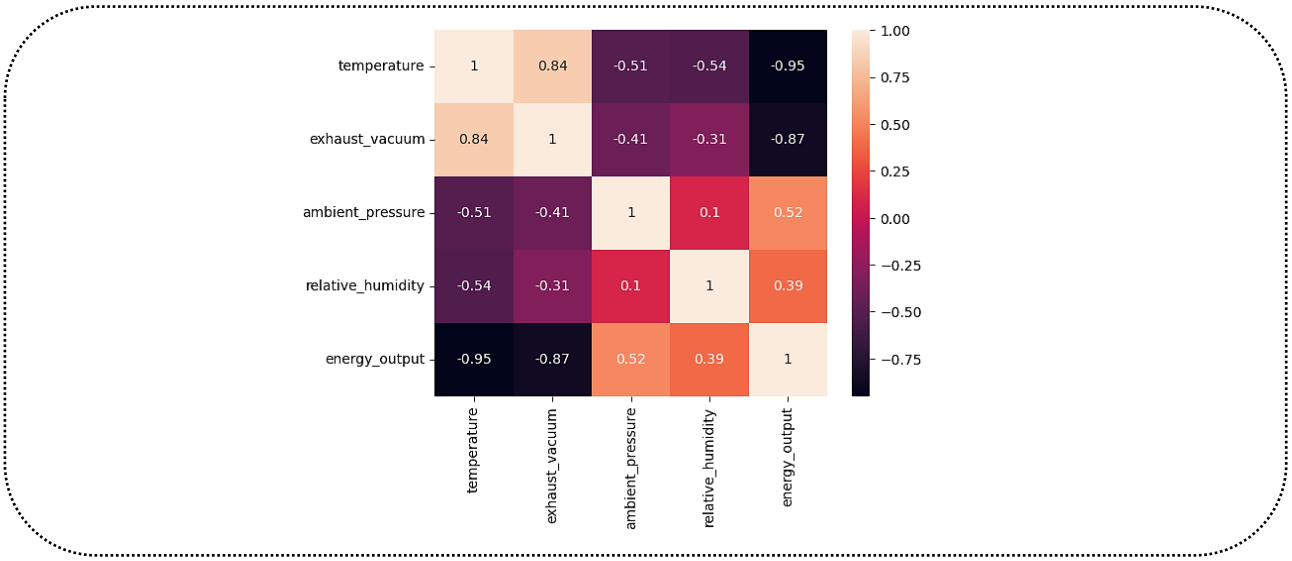


Fig. 4. Pearson correlation among data variables

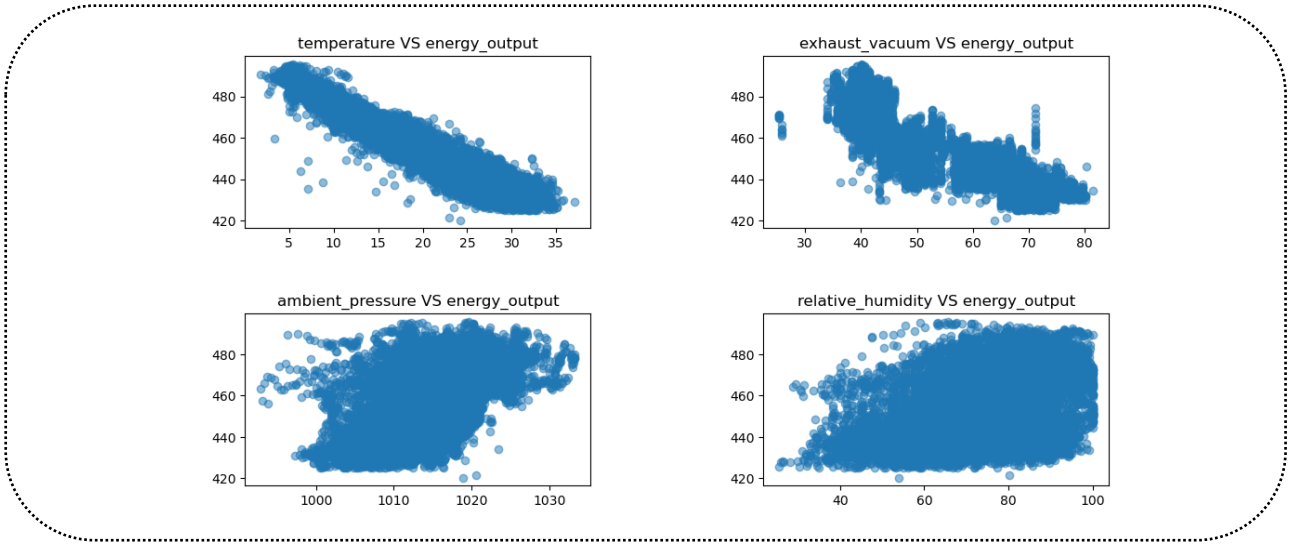


Fig. 5. Response variable versus independent variables

Subsequently, the relationship between the independent variables and the target variable is examined. Fig. 5 presents scatter plots of energy output against temperature, exhaust vacuum, ambient pressure, and relative humidity. Consistent with the correlation analysis, temperature and exhaust vacuum represent a strong relationship with energy output.

C. Model setting and hyperparameter tuning

In this subsection, the proposed machine learning models are implemented to predict energy output using the selected features: temperature, exhaust vacuum, ambient pressure, and relative humidity. The dataset, consisting of 9,568 recorded data points, is split into 70% for training (6,697 data points) and 30% for testing (2,871 data points). This approach allows the model to learn patterns from the training data and evaluate its performance on the unseen test data, ensuring its ability to generalize.

The first implemented model is a Bagging Decision Tree, built using Bootstrap sampling. Like other decision trees, it has specific hyperparameters, including max depth, minimum sample leaf, and the number of trees. To optimize these parameters, the k-fold cross-validation method is applied. Specifically, three values are evaluated for max depth (2, 3, 5), three for minimum sample leaf (5, 10, 15), and three for the number of trees (150, 300, 500). This results in 27 hyperparameter combinations, each of which is tested using 5-fold cross-validation. The negative root mean squared error (RMSE) results for each case in the CV model are presented in Table II. The case with the highest negative RMSE indicates the optimal set of hyperparameters (No. 21), resulting in the best model performance.

Table II. Parameter tuning for the Bagging decision tree algorithm

No.	Max dept	Min samples leaf	N estimators	-RMSE	No.	Max dept	Min samples leaf	N estimators	-RMSE
1	2	5	150	-5.99498	15	3	10	500	-4.88796
2	2	5	300	-5.99414	16	3	15	150	-4.88691
3	2	5	500	-5.98417	17	3	15	300	-4.88709
4	2	10	150	-5.99976	18	3	15	500	-4.88386
5	2	10	300	-5.99792	19	5	5	150	-4.2176

Continue Table II. Parameter tuning for the Bagging decision tree algorithm

No.	Max dept	Min samples leaf	N estimators	-RMSE	No.	Max dept	Min samples leaf	N estimators	-RMSE
6	2	10	500	-5.99578	20	5	5	300	-4.21711
7	2	15	150	-5.98194	21	5	5	500	-4.21676
8	2	15	300	-5.99751	22	5	10	150	-4.22304
9	2	15	500	-5.98805	23	5	10	300	-4.21896
10	3	5	150	-4.88565	24	5	10	500	-4.21785
11	3	5	300	-4.88316	25	5	15	150	-4.22288
12	3	5	500	-4.88518	26	5	15	300	-4.22122
13	3	10	150	-4.88013	27	5	15	500	-4.22176
14	3	10	300	-4.8852					

The second implemented model is Random Forest, which differs from Bagging Decision Tree primarily in the randomness of feature selection during training. In addition to the hyperparameters used in Bagging Decision Tree, Random Forest introduces an additional hyperparameter: maximum participating features. To optimize these parameters, the k-fold cross-validation method is applied. Specifically, three values are evaluated for max features (2, 3, 4), three for max depth (2, 3, 5), and three for minimum sample leaf (5, 10, 15), resulting in 27 hyperparameter combinations. Consistent with the Bagging decision tree algorithm hyperparameter tuning from the previous stage, the number of trees is set to 500. Each combination is tested using 5-fold cross-validation, and negative root mean squared error (RMSE) results for each case in the cross-validation model are presented in Table III. The case with the highest negative RMSE (No. 16) represents the optimal set of hyperparameters, leading to the best model performance.

Table III. Parameter tuning for the Random Forest algorithm

No.	Max features	Max depth	Min samples leaf	-RMSE	No.	Max features	Max depth	Min samples leaf	-RMSE
1	2	2	5	-5.7919	15	3	3	15	-4.72413
2	2	2	10	-5.7919	16	3	5	5	-4.17973
3	2	2	15	-5.79192	17	3	5	10	-4.18022
4	2	3	5	-4.90602	18	3	5	15	-4.18184
5	2	3	10	-4.90596	19	4	2	5	-6.00018
6	2	3	15	-4.90563	20	4	2	10	-6.00018
7	2	5	5	-4.21619	21	4	2	15	-6.00018
8	2	5	10	-4.21633	22	4	3	5	-4.88731
9	2	5	15	-4.21986	23	4	3	10	-4.88731
10	3	2	5	-5.51114	24	4	3	15	-4.88731
11	3	2	10	-5.51114	25	4	5	5	-4.21675
12	3	2	15	-5.51114	26	4	5	10	-4.2177
13	3	3	5	-4.72465	27	4	5	15	-4.22067
14	3	3	10	-4.72443					

The Gradient Boosting algorithm also requires hyperparameter optimization, which is performed using the k-fold

cross-validation method. Since Gradient Boosting is a decision tree-based algorithm, key hyperparameters include the learning rate, the fraction of samples used for fitting individual base learners (subsample), and max depth.

Table IV. Parameter tuning for the Gradient boosting algorithm

No.	Learning rate	Subsample	Max depth	-RMSE	No.	Learning rate	Subsample	Max depth	-RMSE
1	0.1	0.6	1	-4.13792	15	0.15	0.8	5	-3.28603
2	0.1	0.6	3	-3.54109	16	0.15	1	1	-4.1674
3	0.1	0.6	5	-3.34404	17	0.15	1	3	-3.41422
4	0.1	0.8	1	-4.15182	18	0.15	1	5	-3.26048
5	0.1	0.8	3	-3.50968	19	0.2	0.6	1	-4.16764
6	0.1	0.8	5	-3.30108	20	0.2	0.6	3	-3.47223
7	0.1	1	1	-4.16499	21	0.2	0.6	5	-3.37267
8	0.1	1	3	-3.48256	22	0.2	0.8	1	-4.17355
9	0.1	1	5	-3.30251	23	0.2	0.8	3	-3.42835
10	0.15	0.6	1	-4.14494	24	0.2	0.8	5	-3.3016
11	0.15	0.6	3	-3.50131	25	0.2	1	1	-4.18041
12	0.15	0.6	5	-3.34325	26	0.2	1	3	-3.38752
13	0.15	0.8	1	-4.15831	27	0.2	1	5	-3.276
14	0.15	0.8	3	-3.44807					

For this study, we set the number of trees to 500 and the minimum sample leaf to 15. Specifically, we evaluate three values for max depth (2, 3, 5), three for subsample (0.6, 0.8, 1), and three for learning rate (0.1, 0.15, 0.2), resulting in 27 hyperparameter combinations. Each combination is tested using 5-fold cross-validation. The negative root mean squared error (RMSE) results for each case in the cross-validation model are presented in Table IV. The case with the highest negative RMSE (No. 18) represents the optimal set of hyperparameters, demonstrating the best model performance.

The XGBoost algorithm is the next model to be optimized. This algorithm includes a wide range of hyperparameters, such as the number of trees, maximum depth, learning rate, subsample ratio of training instances, subsample ratio of columns for tree construction, L1 regularization (Lasso), and L2 regularization (Ridge) on weights.

Table V. Parameter tuning for the XG-boosting algorithm

No.	Max depth	Subsample	L1 regularization	-RMSE	No.	Max depth	Subsample	L1 regularization	-RMSE
1	2	0.6	0.1	-4.13816	15	3	0.8	0.2	-3.97609
2	2	0.6	0.15	-4.14242	16	3	1	0.1	-4.03684
3	2	0.6	0.2	-4.13673	17	3	1	0.15	-4.02958
4	2	0.8	0.1	-4.14088	18	3	1	0.2	-4.02709

Continue Table V. Parameter tuning for the XG-boosting algorithm

No.	Max depth	Subsample	L1 regularization	-RMSE	No.	Max depth	Subsample	L1 regularization	-RMSE
5	2	0.8	0.15	-4.14339	19	5	0.6	0.1	-3.69605
6	2	0.8	0.2	-4.14316	20	5	0.6	0.15	-3.70467
7	2	1	0.1	-4.2362	21	5	0.6	0.2	-3.71135
8	2	1	0.15	-4.23081	22	5	0.8	0.1	-3.70952
9	2	1	0.2	-4.23302	23	5	0.8	0.15	-3.70161
10	3	0.6	0.1	-3.95567	24	5	0.8	0.2	-3.69963
11	3	0.6	0.15	-3.96956	25	5	1	0.1	-3.72983
12	3	0.6	0.2	-3.96979	26	5	1	0.15	-3.72161
13	3	0.8	0.1	-3.98416	27	5	1	0.2	-3.71675
14	3	0.8	0.15	-3.97665					

For this study, we set the number of trees to 500, the learning rate to 0.1, the subsample ratio of columns for tree construction to 0.5, and L2 regularization term to 0.1. Then, the max depth, the subsample ratio of training instances, and the L1 regularization term are fine-tuned using 5-fold cross-validation. The negative root mean squared error (RMSE) results for each case in the cross-validation model are presented in Table V. The case with the highest negative RMSE (No. 19) represents the optimal set of hyperparameters, resulting in the best model performance.

The proposed neural network model, as described in Section 3, is designed to predict energy output and optimize operational parameters in the CCPP. Initially, the input features are standardized using the mean and standard deviation. The first perceptron layer consists of four input features and three neurons, utilizing the tangent hyperbolic activation function. The second perceptron layer has three inputs and a single neuron, using a linear activation function. Finally, the output is rescaled to its original scale to facilitate accurate interpretation. Fig. 6 presents a schematic representation of this neural network. The loss function and regularization term are defined according to the formulations presented in Section 3.

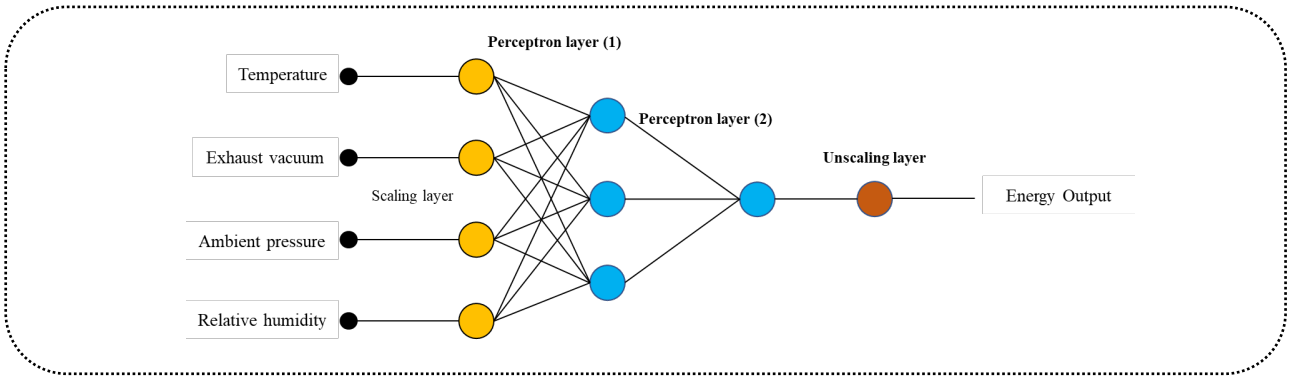


Fig. 6. The proposed neural network structure

Now that the optimal hyperparameters for each proposed model have been identified, the models are implemented on the dataset. In the following section, these models are employed to predict energy output for demand planning and to optimize the operational parameters of the CCPP.

V. EXPERIMENTAL DISCUSSION

In this section, the results of applying the proposed models to the dataset are presented. First, the models are trained on the training data using their optimized parameters and then applied to the test data for prediction. Next, the proposed neural network is utilized to optimize the operational parameters. Finally, the feature importance across all proposed models

A. Model evaluation

The proposed models, introduced in Section 3, are configured with various hyperparameters, which were optimized in Section 4. After training these models on the training dataset, the test dataset is used to predict energy output for each model.

To ensure a comprehensive evaluation and minimize reporting bias, the performance of the models on the test dataset is analyzed from three perspectives:

- Actual vs. Predicted Values: A direct comparison to assess prediction accuracy.
- Absolute Errors and Their Statistical Characteristics: To provide insights into the distribution and magnitude of errors.
- R-Squared Value Evaluation: To measure how well each model explains the variance in the data.

These evaluations allow for a well-rounded assessment of model performance from different analytical viewpoints.

Fig. 7 presents a comparison of actual versus predicted values for each of the five proposed models. In these plots, the horizontal axis represents the actual values, and the vertical axis represents the predicted values. The closer the data points cloud around the diagonal reference line, the more accurate the model's predictions on unseen data.

As shown in Fig. 7, the Bagging Decision Tree model exhibits the widest data spread, indicating the lowest prediction accuracy on unseen data. The Random Forest and Neural Network models follow, and are ranked among the least accurate models. In contrast, the Gradient Boosting model demonstrates the highest accuracy, with its data cloud closest to the diagonal reference line, indicating superior predictive performance.

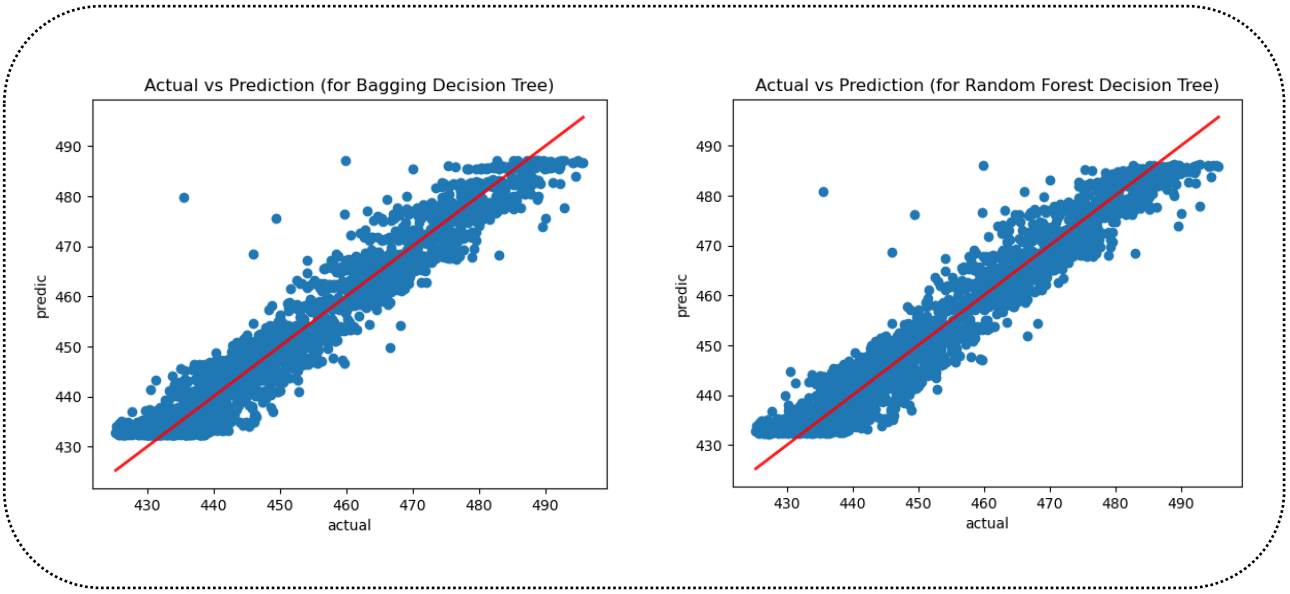
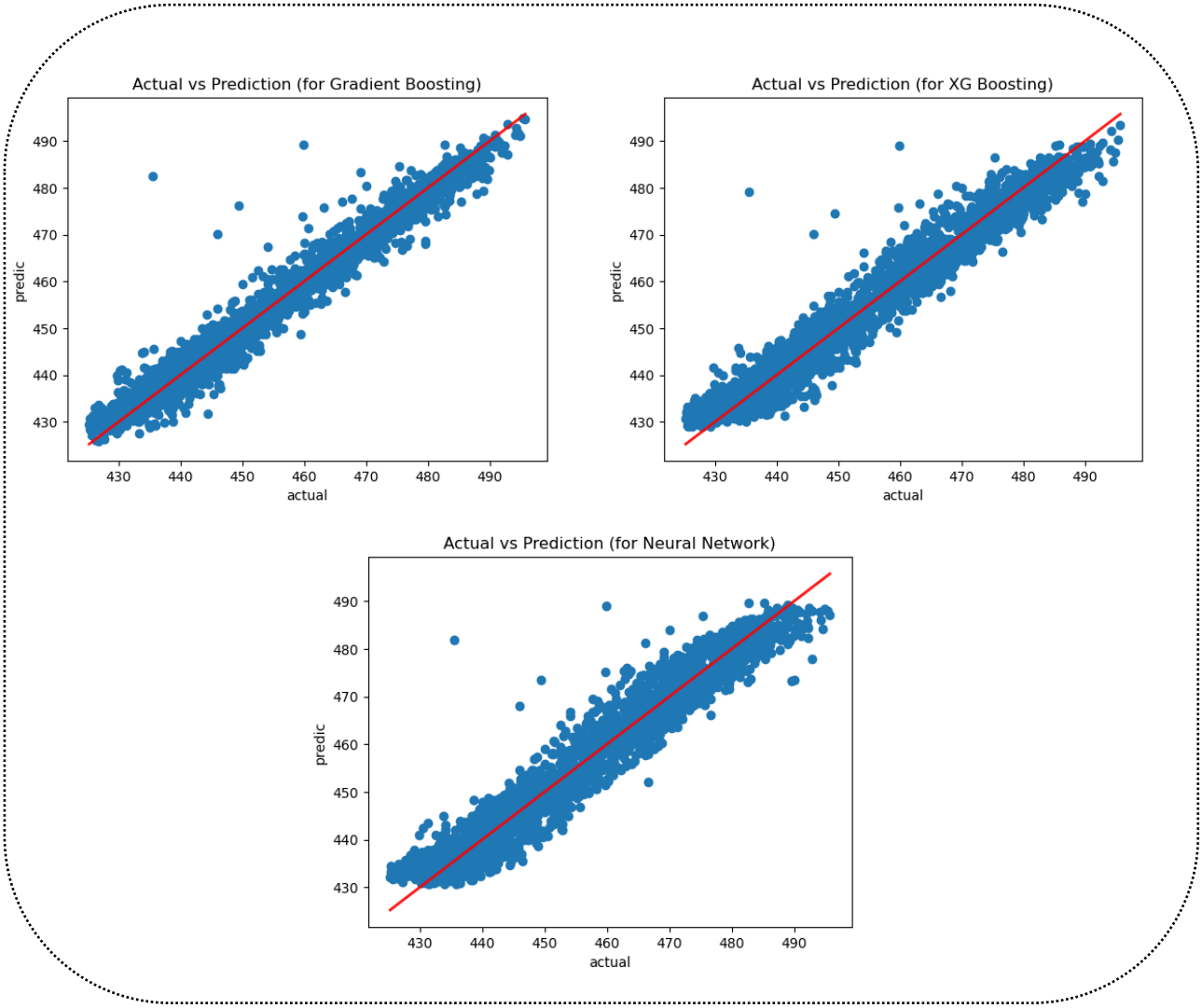


Fig. 7. Actual versus prediction chart for the proposed models



Continue Fig. 7. Actual versus prediction chart for the proposed models

In the second analysis, absolute errors are computed as $|y_{actual} - y_{predicted}|$, and key statistical measures, including mean, median, standard deviation, interquartile range, minimum, and maximum, are evaluated and compared across the proposed models. Table VI presents these statistical measures for each model separately. The results indicate that the Gradient Boosting model delivers the best performance, with significantly lower mean and standard deviation compared to the other models. Following closely, the XGBoost model ranks second, exhibiting the lowest absolute error metrics among the remaining models.

In the third analysis, the R-squared value is calculated for each model to evaluate how well it explains the variance in the data. The results, presented in Table VI, indicate that the Neural Network model achieves the highest explainability, followed closely by the Gradient Boosting model, which ranks second with a slight performance difference.

In summary, the Gradient Boosting model consistently delivers superior performance across the majority of comparative analyses conducted in this study. Given its high predictive accuracy and robustness, it is recommended that future demand planning in CCPPs adopt the Gradient Boosting algorithm, configured with the tuned parameters identified in this study. Implementing this approach can help decision-makers to generate more precise energy output

forecasts, thereby improving operational planning, enhancing resource allocation, and supporting more stable and efficient power generation.

Table VI- R2 and absolute error statistical measures in each proposed model

Model	Mean of absolute error	Median of absolute error	SD of absolute error	IQR of absolute error	Min of absolute error	Max of absolute error	R^2
Bagging decision tree	3.232435	2.674517	2.660329	3.382061	0.000263	44.25493	0.938
Random Forest	3.201816	2.680826	2.64876	3.313089	0.00032	45.35575	0.938
Gradient Boosting	2.215449	1.699907	2.21167	2.241848	0.000402	46.90702	0.975
XG Boosting	2.796226	2.304817	2.414523	2.829543	0.002137	43.6368	0.924
Neural Network	3.260177	2.865112	2.607343	3.245578	0.000529	46.28578	0.985

B. Optimization of Operational Parameters for Enhanced Productivity

Optimizing operational parameters under favorable conditions and leveraging historical data can help decision-makers enhance energy output, thereby mitigating energy supply chain imbalances. Neural networks facilitate this optimization through a backward optimization procedure, as in Section 3. To identify the conditions that improve productivity, it is essential to reassess the data understanding stage for deeper insights.

To explore these relationships, Fig. 8 presents a scatter plot of exhaust vacuum vs. energy output along with a histogram of energy output distribution. The results clearly indicate that the CCPP achieves higher energy output when exhaust vacuum levels are lower. Additionally, for the plant to operate efficiently in terms of productivity, the generation rate must exceed 450 MW. These two key conditions will be incorporated into the optimization procedure within the proposed neural network model.

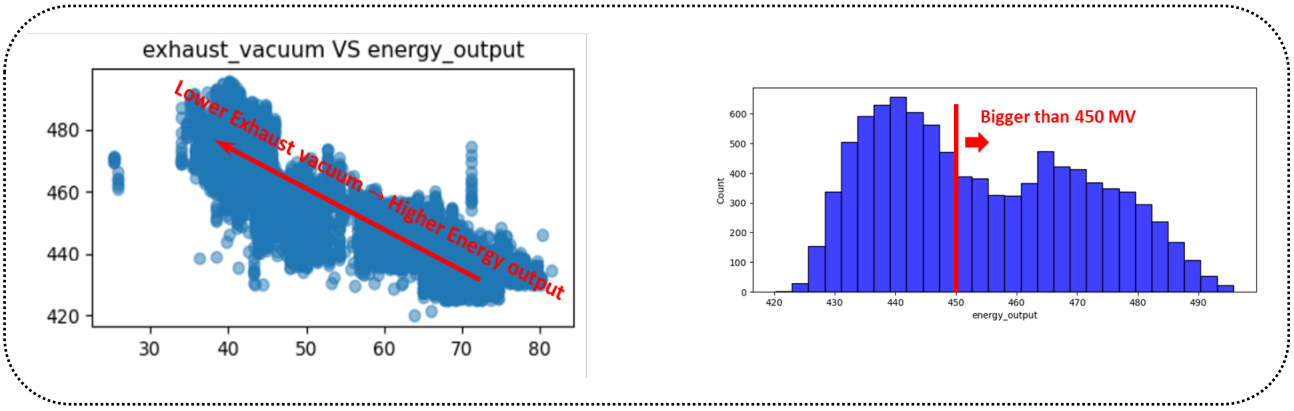


Fig. 8. Conditions for Neural network optimization by data understanding

Next, the proposed Quasi-Newton method, introduced in Section 3, is implemented on the trained neural network model, incorporating the two specified optimization conditions. These conditions, along with the optimization results from the backward neural network process, are presented in Table VII, where the optimal values for the operational inputs are determined. The optimal values for temperature, ambient pressure, and relative humidity under the proposed conditions are presented in Table VII. By applying these optimized configurations, energy sector decision-makers can not only enhance the productivity of the CCPP but also improve thermal efficiency, reduce fuel consumption, and minimize operational costs. Furthermore, such data-driven optimization enables plants to operate closer to their ideal thermodynamic conditions, ensuring more stable output under fluctuating environmental and demand conditions. This

approach ultimately supports long-term sustainability goals by improving resource utilization and reducing unnecessary emissions associated with inefficient operation.

Table VII. Results of the optimization procedure in the proposed neural network by backpropagation

Variable Name	Condition		Optimum values
Temperature	None	Optimization by Quasi-Newton optimizer	19.3846 Celsius degrees
Exhaust vacuum	Minimize		25.432 cm Hg
Ambient pressure	None		1021.4 millibar
Relative humidity	None		60.8 %
Energy	Greater than or equal to 450		462.086 MW

C. Feature Importance

In the final section, we analyze the feature importance across the five proposed models within the dataset. This evaluation helps decision-makers identify which features play a crucial role in predicting the response variable. Fig. 9 illustrates the importance of temperature, exhaust vacuum, ambient pressure, and relative humidity in each model. The feature importance is determined by measuring the change in the loss function when a specific feature is excluded from the training dataset, and the model is retrained using the permuted dataset.

The results indicate that temperature is the most critical feature for predicting energy output in a CCPP, consistently accounting for over 70% importance across all models. Following temperature, exhaust vacuum emerges as the second most influential feature, with an importance ranging between 5% and 25% across the proposed models. Ambient pressure and relative humidity rank third and fourth, respectively, varying in importance depending on the model. Since Gradient Boosting was selected as the best-performing model, it is evident that ambient pressure holds the third rank, while relative humidity ranks fourth.

This ranking highlights the dominant influence of thermal conditions on CCPP performance, as temperature and exhaust vacuum directly affect turbine efficiency and steam cycle stability. The findings suggest that operational strategies should prioritize the continuous monitoring and fine-tuning of these key variables, particularly under changing environmental conditions. Moreover, integrating this feature-importance insight with real-time sensor networks and optimization frameworks, such as the Quasi-Newton-based neural network approach proposed in this study, could allow operators to anticipate performance fluctuations and proactively adjust settings to sustain optimal output and efficiency.

VI. CONCLUSION

The energy sector operates as a supply chain, extending from power generation to end-user consumption. In developing countries, fossil fuel-based systems, particularly CCPPs, are widely used. However, supply-demand imbalances in these plants can disrupt the electricity network. Enhancing productivity is essential to maximizing capacity and creating a more responsive system to ensure a stable electricity supply. This study proposes an AI-driven approach to enhance CCPP productivity and address electricity grid imbalances as a long-term challenge. IoT sensors collect real-time data on temperature, exhaust vacuum, ambient pressure, relative humidity, and energy output. In the first phase, machine learning models, including Bagging decision tree, Random Forest, gradient boosting, XG-Boost, and neural networks, predict energy output, aiding decision-makers in demand planning and mitigating disruptions. In the second phase, a Quasi-Newton optimizer within a neural network optimizes operational parameters, improving efficiency and productivity while helping to balance the electricity grid during supply-demand imbalances.

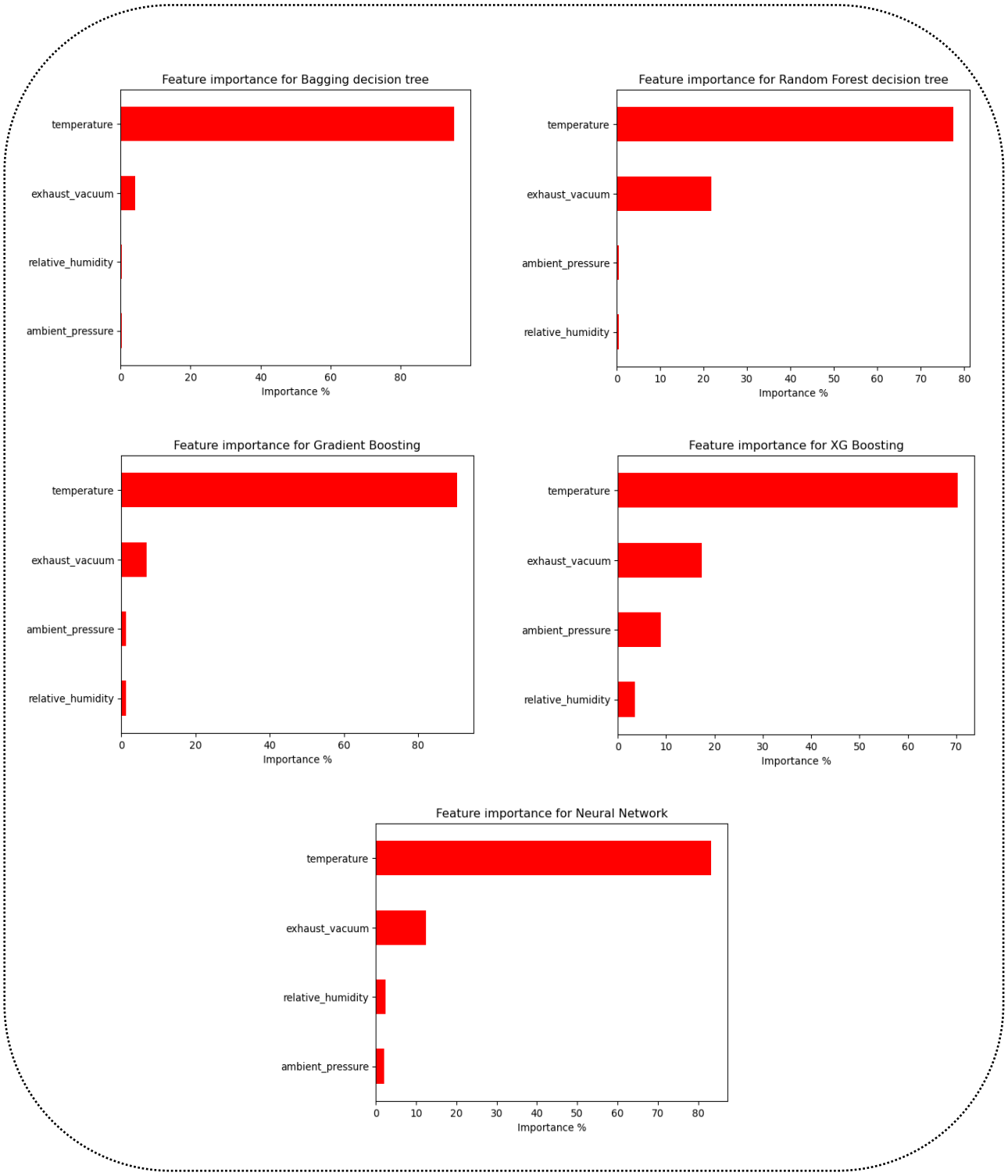


Fig. 9. Feature importance across the proposed models

The initial results demonstrate that the proposed AI models, implemented on the Zavareh Combined Cycle Power Plant dataset, successfully tuned hyperparameters and identified the optimal structure for each model. A comparison of models using actual vs. predicted scatter plots reveals that the gradient boosting algorithm delivers the highest accuracy in energy output prediction, with a mean absolute error of 2.215449. Adopting this approach enables decision-makers to

produce more accurate energy output forecasts, strengthening operational planning, optimizing resource allocation, and ensuring more stable and efficient power generation.

The Quasi-Newton optimizer, integrated into a backward-propagated neural network, identified optimal operational parameters: temperature of 19.38 °C, exhaust vacuum of 25.43 cm Hg, ambient pressure of 1021.4 mbar, relative humidity of 60.8%, and energy output of 462.09 MW. Applying these configurations can enhance CCPP productivity, improve thermal efficiency, reduce fuel consumption, and lower operational costs. Operating closer to ideal thermodynamic conditions also enables greater output stability under fluctuating environmental and demand conditions, supporting long-term sustainability by reducing emissions from inefficient operation.

Feature importance analysis indicates that temperature is the dominant predictor of energy output, followed by exhaust vacuum, ambient pressure, and relative humidity. These findings highlight the importance of continuously monitoring and fine-tuning critical variables. When combined with real-time sensor networks and advanced optimization frameworks, this approach can enable operators to anticipate performance fluctuations and proactively adjust operational settings, ensuring sustained efficiency and stability.

This study has certain limitations that present opportunities for future improvement. Future research can integrate deep learning and reinforcement learning to enhance both prediction accuracy and interaction with the power plant environment. Additionally, time series forecasting methods can be employed for demand planning, and their integration with supervised learning models could further improve prediction precision and operational efficiency.

REFERENCES

Adams, D., Oh, D.H., Kim, D.W., Lee, C.H. & Oh, M. (2020). Prediction of SOx–NOx emission from a coal-fired CFB power plant with machine learning: Plant data learned by deep neural network and least square support vector machine. *Journal of Cleaner Production*, 270, 122310.

Ağbulut, Ü., Gürel, A.E., Ergün, A. & Ceylan, İ. (2020). Performance assessment of a V-trough photovoltaic system and prediction of power output with different machine learning algorithms. *Journal of Cleaner Production*, *268*, 122269.

Aghajani, D., Gilani, H., Sahebi, H. & Vilko, J. (2025). A Robust Design of Hydrogen Supply Chain with Integrated Sustainable Pricing Policy under Government Intervention: Circular Economy-Driven Management. *Energy*, 137660.

Akdemir, B. (2016). Prediction of hourly generated electric power using artificial neural network for combined cycle power plant. Int. J. Electr. *Energy*, *4*(2), 91-95.

Ali, A., Kamal, K., Ratlamwala, T.A.H., Fahad Sheikh, M. and Arsalan, M. (2021). Power prediction of waste heat recovery system for a cement plant using back propagation neural network and its thermodynamic modeling. *International Journal of Energy Research*, *45*(6), 9162-9178.

Alipoor, F., Gilani, H., Sahebi, H. & Ghannadpour, S.F. (2024). Renewable energy-powered water desalination and treatment network under wind power and water demand uncertainty: A possibilistic chance-constrained programming. *Energy Strategy Reviews*, *55*, 101511.

An, G., Chen, L., Tan, J., Jiang, Z., Li, Z. & Sun, H. (2022). Ultra-short-term wind power prediction based on PVMD-ESMA-DELM. *Energy Reports*, *8*, 8574-8588.

Ardestani, R., Mehdizadeh, E. & Etebari, F. (2024). A Mixed-Integer Nonlinear Programming Model for Solving Integrated Oil and Gas Supply Chain Problem by Considering Enhanced Oil Recovery Methods. *Journal of Quality Engineering and Production Optimization*, *9*(1), 1-28.

Asghar, A., Ratlamwala, T.A.H., Kamal, K., Alkahtani, M., Mohammad, E. & Mathavan, S. (2023). Sustainable operations of a combined cycle power plant using artificial intelligence based power prediction. *Heliyon*, *9*(9).

Attari, M.Y.N., Ala, A. & Khalilpourshiraz, Z. (2022). The electric power supply chain network design and emission reduction policy: a comprehensive review. *Environmental science and pollution research*, *29*(37), 55541-55567.

Bistline, J.E. & Merrick, J.H. (2020). Parameterizing open-source energy models: Statistical learning to estimate unknown power plant attributes. *Applied Energy*, *269*, 114941.

Breiman, L. (2001). Random forests. *Machine learning, 45*, 5-32.

Chahboun, S. & Maaroufi, M. (2021). Principal component analysis and machine learning approaches for photovoltaic power prediction: A comparative study. *Applied Sciences*, *11*(17), 7943.

Chen, T. & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd international conference on knowledge discovery and data mining* (pp. 785-794).

Condemi, C., Casillas-Perez, D., Mastroeni, L., Jiménez-Fernández, S. & Salcedo-Sanz, S. (2021). Hydro-power production capacity prediction based on machine learning regression techniques. *Knowledge-Based Systems*, *222*, 107012.

Dairi, A., Harrou, F., Sun, Y. & Khadraoui, S. (2020). Short-term forecasting of photovoltaic solar power production using variational auto-encoder driven deep learning approach. *Applied Sciences*, *10*(23), 8400.

Danesh, T., Ouaret, R. & Floquet, P. (2021, May). Towards Robust process design. The sensitivity analysis using machine learning methods. In *ENBIS 2021 Spring Meeting on Data Science in Process Industries*, *60*, 126-140.

Darmawan, A., Tokimatsu, K. & Aziz, M. (2025). Integrated Gasification Combined-Cycle Power Plants Utilizing Biomass Waste: Energy-Efficient Recovery. In *Decarbonizing Power Generation Sectors Using Biomass and Hydrogen-Based Fuels: A Roadmap to Sustainable Energy Transformation* (pp. 109-136). Singapore: Springer Nature Singapore.

Ekanayake, P., Wickramasinghe, L., Jayasinghe, J.J.W. & Rathnayake, U. (2021). Regression-Based Prediction of Power Generation at Samanalawewa Hydropower Plant in Sri Lanka Using Machine Learning. *Mathematical Problems in Engineering*, *2021*(1), 4913824.

Eskandari, M., Gilani, H., Sahebi, H. & Sharifi, S. (2024). Design and planning of global sustainable bio-hydrogen supply chain with uncertainty: A transportation-oriented robust model. *Chemical Engineering Science*, *283*, 119365.

Essam, Y., Ahmed, A.N., Ramli, R., Chau, K.W., Idris Ibrahim, M.S., Sherif, M., Sefelnasr, A. & El-Shafie, A. (2022). Investigating photovoltaic solar power output forecasting using machine learning algorithms. *Engineering Applications of Computational Fluid Mechanics*, *16*(1), 2002-2034.

Fattahi, P., Shakeri Kebria, Z. & Setak, M. (2020). Planning of Multiple Energy Hubs and Scheduling of Preventive Maintenance Equipment Under Uncertainty and Energy Storage. *Journal of Quality Engineering and Production Optimization*, *5*(1), 119-136.

Friedman, J.H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics*, 1189-1232.

Gilani, H. & Sahebi, H. (2024). Optimizing sustainable multiple biomass-to-biofuel conversion network with integrated water resource management utilizing data-driven robust planning. *Energy Conversion and Management: X*, *24*, 100727.

Gilani, H., Sahebi, H. & Van Woensel, T. (2025). Resilient coordination of test sampling and requirement supply in the lab test supply chain: soft worst-case distributionally robust optimization. *Computers & Industrial Engineering*, 111316.

Gilani, H., Sahebi, H., Jabalameli, M.S. & Hosseininezhad, S.J. (2023). Flexible robust optimization planning of wastewater–energy nexus-based biofuel network design under cultivation rate and climate change uncertainty. *Industrial & Engineering Chemistry Research*, *62*(47), 20300-20325.

Han, Y. & Tong, X. (2020). Multi-step short-term wind power prediction based on three-level decomposition and improved grey wolf optimization. *IEEE Access*, *8*, 67124-67136.

Hundi, P. & Shahsavari, R., (2020). Comparative studies among machine learning models for performance estimation and health monitoring of thermal power plants. *Applied Energy*, *265*, 114775.

Kaewprapha, P., Prempaneerach, P., Singh, V., Tinikul, T. & Intarangsi, N. (2022, March). Machine Learning approaches for estimating the efficiency of combined cycle power plant. In 2022 *International Electrical Engineering Congress (iEECON),*(pp. 1-4). IEEE.

Li, H. (2022). Short-term wind power prediction via spatial temporal analysis and deep residual networks. *Frontiers in energy research*, *10*, 920407.

Mohammad-Rafiei, A., Mardan, E. & Kamran Rad, R. (2025). Flow-shop scheduling problem with tool changes and tool wear to minimize energy consumption. *Journal of Quality Engineering and Production Optimization, 9*(2), 19-46.

Nagurney, A. & Matsypura, D. (2007). A supply chain network perspective for electric power generation, supply, transmission, and consumption. In *Optimisation, econometric and financial analysis* (pp. 3-27). Berlin, Heidelberg: Springer Berlin Heidelberg.

Ntantis, E.L. & Xezonakis, V. (2024). Optimization of electric power prediction of a combined cycle power plant using innovative machine learning technique. *Optimal Control Applications and Methods, 45*(5), 2218-2230.

Özen, C. & Deniz, A. (2022). A comprehensive country-based day-ahead wind power generation forecast model by coupling numerical weather prediction data and CatBoost with feature selection methods for Turkey. *Wind Engineering*, *46*(5), 1359-1388.

Tüfekci, P. (2014). Prediction of full load electrical power output of a base load operated combined cycle power plant using machine learning methods. *International Journal of Electrical Power & Energy Systems*, *60*, 126-140.

Rasheed, A.A. (2023). Improving prediction efficiency by revolutionary machine learning models. *Materials today: proceedings, 81*, 577-583.

Rashid, M., Kamal, K., Zafar, T., Sheikh, Z., Shah, A. & Mathavan, S. (2015, December). Energy prediction of a combined cycle power plant using a particle swarm optimization trained feedforward neural network. In 2015 International Conference on Mechanical Engineering, Automation and Control Systems (MEACS), 1-5. IEEE.

Robitaille, B., Marcos, B., Veillette, M. & Payre, G. (1996). Modified quasi-Newton methods for training neural networks. *Computers & chemical engineering, 20*(9), 1133-1140.

Rumelhart, D.E., Hinton, G.E. & Williams, R.J. (1986). Learning representations by back-propagating errors. *nature*, *323*(6088), 533-536.

Saeed, M.A., El-Kenawy, E.S.M., Ibrahim, A., Abdelhamid, A.A., Eid, M.M., Karim, F.K., Khafaga, D.S. & Abualigah, L. (2023). Electrical power output prediction of combined cycle power plants using a recurrent neural network optimized by waterwheel plant algorithm. *Frontiers in Energy Research*, *11*, 1234624.

Sutton, C.D. (2005). Classification and regression trees, bagging, and boosting. *Handbook of statistics*, *24*, 303-329.

Vijayalakshmi, S., Balusamy, B. & Dhanaraj, R.K. eds. (2023). AI-Powered IoT in the Energy Industry: Digital Technology and Sustainable Energy Systems. Springer International Publishing.

Wu, D., Kan, J., Lin, H.C. & Li, S. (2021). Hybrid Improved Bird Swarm Algorithm with Extreme Learning Machine for Short-Term Power Prediction in Photovoltaic Power Generation System. *Computational Intelligence and Neuroscience*, *2021*(1), 6638436.

Yan, M., Guo, W., Hu, Y., Xu, F., Chen, J., Du, Q., Zheng, H. & Qin, T. (2023). Improved hybrid sparrow search algorithm for an extreme learning machine neural network for short-term photovoltaic power prediction in 5G energy-routing base stations. *IET Renewable Power Generation, 17*(2), 336-348.

Yi, Q., Xiong, H. & Wang, D. (2023). Predicting power generation from a combined cycle power plant using transformer encoders with DNN. *Electronics, 12*(11), 2431.

Zhang, H., Zhao, H., Li, Z. & Hu, E. (2019). Optimization potentials for the waste heat recovery of a gas-steam combined cycle power plant based on absorption heat pump. *Journal of Thermal Science, 28*, 283-293.

Zhang, J., Zhang, M., Yang, J. & Zheng, X. (2024). Prediction of electricity load generated by combined cycle power plants using integration of machine learning methods and HGS algorithm. *Computers and Electrical Engineering, 120*, 109644.