



Power Forecasting Using ANN and ELM: A Comparative Evaluation of Machine Learning Approaches

Rouibah Brahim^{1*}, Labdaoui Ahlam¹, Aggoun Hamza¹, İnan Güler²

¹ Applied Mathematics and Modelling Laboratory, Department of Mathematics, Faculty of Exact Sciences, University of Frères Mentouri Constantine 1, Constantine 25000, Algeria

² Department of Electrical and Electronics Engineering, Gazi University, Ankara 06500, Turkey

Corresponding Author Email: brahim.rouibah@doc.umc.edu.dz

Copyright: ©2025 The authors. This article is published by IIETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/mmep.120101>

ABSTRACT

Received: 6 October 2024

Revised: 3 December 2024

Accepted: 10 December 2024

Available online: 25 January 2025

Keywords:

combined cycle power plant, computational efficiency, extreme learning machine (ELM), machine learning models, mean squared error (MSE), power output prediction, predictive accuracy, R^2 (coefficient of determination)

Accurate predictions of power output in Combined Cycle Power Plants (CCPPs) are crucial for improving operational efficiency and enhancing performance monitoring. This paper compares two prominent machine learning models, artificial neural networks and extreme learning machines, for the prediction of hourly electrical power output. The analysis is based on a publicly available CCPP dataset containing 9,568 instances with key parameters like ambient temperature, atmospheric pressure, relative humidity, and exhaust vacuum. The performances of the models were compared based on standard regression metrics. The result showed that the extreme learning machine (ELM) outperformed artificial neural network (ANN) with mean squared error (MSE) of 0.26, mean absolute error (MAE) of 0.41, root mean squared error (RMSE) of 0.51, and R^2 of 0.98 when both models yielded a good prediction result, against the ANN model with an MSE of 19.33, MAE of 3.52, RMSE of 4.40, and R^2 of 0.85. Overfitting when dealing with small datasets and necessity of preprocessing for fine-tuning performance of ANN were the potential drawbacks highlighted by the paper. Results indicate that the use of ELM is quite viable and capable for estimation with excellent accuracy and, as a consequence, may have pragmatic implications for performance optimization studies concerning CCPP and also find broad applicability in energy management studies.

1. INTRODUCTION

The analysis of a dynamic system requires multiple hypotheses for accounting for the randomness of the outcome. These theories can be more applicable in practical, real-time studies of chaotic systems [1, 2]. To begin with, many nonlinear equations must be solved, requiring considerable computational resources. Due to this limitation, machine learning and neural network techniques have gained popularity as solutions to thermodynamic-based approaches that show paradoxical features and offer efficiency advantages over engineering considerations [3-5]. These advances determine complex correlations and connections between crucial input and output parameters [6].

Power prediction in CCPPs is essential for optimizing plant performance, improving energy efficiency, and reducing operational costs. Accurate forecasting of power output allows better decision-making, efficient energy distribution, and improved plant management. Currently, machine learning models ANN and ELM have gained much attention due to their capability in modeling complicated nonlinear relationships between the input variables and power output, particularly in predictive tasks for these power generation systems.

The current study will investigate the performances of ANN

and ELM models for power output prediction in CCPP. More precisely, predictive accuracy, computational efficiency, and applicability of ANN and ELM techniques will be assessed and compared by using real-time operational data from a CCPP operating in Turkey. By conducting this comparison, we aim to identify which model offers the most reliable and efficient solution for power prediction, contributing to improved plant performance and more informed operational decision-making.

A CCPP shows a known thermodynamic system. The efficiency of a PP at total capacity depends on many variables, including weather, system interactions, and coupling, limiting the development of an appropriate mathematical model [7]. The CCPP uses steam and gas turbines to produce 50 per cent more electrical power using the same fuel as a classic simple cycle plant [8]. However, estimating output power during full load is essential to improving plant productivity and financial results [9]. In addition, electric power production performance should be evaluated, particularly regarding materials prevention requirements and enhanced durability [10]. Therefore, accurate power generation forecasting is essential to improve the performance of PP and the environment [11]. Recently, researchers have used different machine learning (ML) methods for predicting the output power at the maximum load of CCPP [12-15].

The process of teaching computers the capacity to learn through data and experience, much like a human brain, refers to machine learning [16]. The primary purpose of machine learning is to develop models that can learn from previous data to improve over time, recognize layers, and make predictions about future problems. Moreover, numerous research on the combined cycle power plant (CCPP) have employed neural network approaches to investigate problems in this field, despite these techniques being particularly beneficial for solving industrial problems [17]. Moreover, these neural network methodologies were constructed and compared regarding their ability to analyze the system for random input and output patterns [18].

The functional mechanism of the neurological system of humans inspired the creation of artificial neural networks (ANN) [19, 20]; this method is frequently employed to solve a variety of problems in engineering [21]. ANN is also considered a specialized system for signal processing containing multiple interconnected layers linked with weight vectors [22]. For example, Xezonakis and Ntantis [23] used an ANN to increase the accuracy of measurement data on a thermal power plant model; Arferiandi et al. [24] used an ANN technique, and the CCPP heat rate was predicted to support maintenance personnel in monitoring the CCPP efficiency; Farajollahi et al. [25] developed coupled artificial neural network and genetic algorithm for optimization and modelling of hybrid geothermal-solar energy plant; Karaçor et al. [26] employed ANN for forecasting the life performance of natural gas combined cycle power plant, the given literature suggests that ANN can provide adequate solutions for problems in industry.

ELM is a new training technique for the single-hidden-layer feedforward neural networks [27, 28]. After randomly identifying input-hidden weights and hidden biases, ELM calculates the hidden-output layer weights by employing the generalized Moore-Penrose inverse of the hidden layer's output matrix. ELM offers improved generalization capabilities and a faster training process than traditional gradient-based algorithms, making it more suitable for real-world applications. Deepika et al. [29] employed ELM to forecast a boiler output; Markowska-Kaczmar and Kosturek [30] compared ELM and classical neural methods from the usability perspective; Huang et al. [31] proposed optimum prediction intervals generated by FA-ELM of wind power generation; Sarira et al. [32] developed an extreme learning machine method for regular modelling of power plant efficiency; Liu [33] used ELM for fault diagnosis of Combined Cycle Power Plant, current simulations have proven that ELM is a potential approach for solving complex regression and classification problems.

Artificial neural networks (ANNs) and extreme learning machines (ELMs) have been widely explored for their efficiency in power prediction models, each offering distinct advantages. ANNs have been found to significantly enhance the efficiency of power prediction in various contexts. For example, in power distribution networks, ANNs combined with the Levenberg-Marquardt algorithm showed a 40% reduction in computational time and a 1.3 times improvement in prediction accuracy compared to traditional response surface methodology (RSM) approaches [34]. Similarly, the application of ANNs has been made in wireless power transmission systems for the estimation of some performance characteristics, hence smoothing the design process and reducing cumbersome calculations. In renewable energy,

ANN is applied to forecast solar photovoltaic power generation, performing much better than traditional methods with high accuracy and efficiency for grid management and optimization [35].

On the other side, ELMs, especially those optimized by the SSA, have been found to achieve better accuracy and stability than traditional methods in the domain of wind power generation forecasting. The approach therefore reinforces the operational efficiency and reliability of the wind power system and is a strong method for power prediction in renewable energy contexts [36]. Both ANNs and ELMs give substantial improvements in performance over the traditional approaches. ANNs are more powerful when the scenario to be forecast is complex and nonlinear, while optimized ELMs perform both fast and accurately enough for energy system applications.

Despite these developments in ANN and ELM models for power forecast estimation, direct comparison among them is still lacking in the case of CCPPs, while the majority of the literature focuses on single model applications or does not address modeling challenges such as model overfitting and the need for data preprocessing. Furthermore, although ANN and ELM models have been applied to renewable energy systems like wind and solar, there is limited exploration of their potential for power prediction in CCPPs under various operational conditions. This study bridges this gap by comparatively analyzing the performance of ANN and ELM models for predicting power output in CCPPs, overcoming the limitations of previous studies and offering new insights into predicting power generation efficiency.

The remaining parts of this document are organized in this order: The second section provides an overview of the methodologies used in this investigation. The quantitative indices, experimental results, and discussions are given in Sections 3 and 4, and the conclusion is presented in Section 5.

2. DATA PROVISION

It is recognized that in machine learning simulations, any parameter defined for prediction needs several inputs that function as significant parameters for the network. This study considers four input factors: ambient temperature (AT), exhaust vacuum (V), atmospheric pressure (AP), and relative humidity (RH), with the performance efficiency (PE) as the output.

Based on the work of reference [37], the dataset was obtained for free from a machine learning repository. A predictive network operates in two different phases for data processing: (a) the phase of training, when the correlation of inputs and targets is established and then applied to modify the simulation model, and (b) the phase of testing, where the efficiency of the improved model will be evaluated. The dataset is divided into a training set and a testing set containing 9568 instances, including 7654 and 1914 examples, which correspond to 80% and 20% of the overall total.

Table 1 presents further statistical information about the dataset, including parameters with a mean PE value of 454.4 mW and a range of [420.3 and 495.8 mW]. The highest and lowest temperatures recorded are 1.8°C and 37.1°C.

The exhaust vacuum pressure values range between 25.4 and 81.6 cmHg. Similarly, the highest and lowest atmospheric pressures recorded are 992.9 and 1033.3 mbar, respectively. The relative humidity readings in Table 1 range between a minimum of 25.6% and a maximum of 100.2%.

In addition, data correlation analysis demonstrated that output (i.e., PE) is significantly and negatively proportional to the AT. In addition, a similar tendency was found for the V parameter but with a lower association. Figure 1 shows the distribution order of inputs (a, b, c, and d) and the target

parameter (e) to facilitate data comprehension.

The typical values for (AT), (V), (AP), and (RH) are approximately 25°C, 40 cmHg, 1010 mbar, and 81%, as the graph indicates. Additionally, values of PE close to 440 are the most obtained.

Table 1. Distribution of inputs and output

Average Temperature	Exhaust Vacuum	Ambient Pressure	Relative Humidity	Net Hourly Electrical Energy Output
0-8.34	40.77	1010.84	90.01	480.48
1-23.64	58.49	1011.40	74.20	445.75
2-29.74	56.90	1007.15	41.91	438.76
3-19.07	49.69	1007.22	76.79	453.09
4-11.80	40.66	1017.13	97.20	464.43

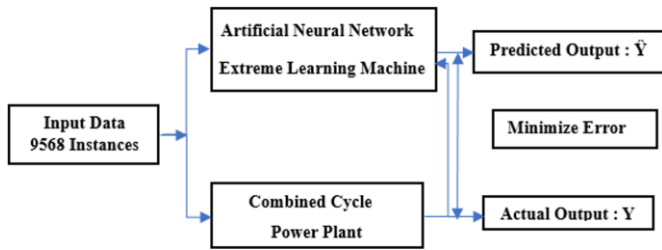


Figure 1. Predicting output by using neural networks regression technique

3. METHODS

This section describes the research techniques and principles of this work, with a particular emphasis on the mathematical formulae and models used in the ANN and ELM algorithms, to simplify the introduction of better algorithms in the following section.

3.1 Dataset and preprocessing

This study utilizes a publicly available Combined Cycle Power Plant (CCPP) dataset, which consists of 9,568 instances. Key features include ambient temperature, atmospheric pressure, relative humidity, and exhaust vacuum—critical factors in predicting power output. The target variable is the hourly electrical power output (EP) of the plant.

Preprocessing involved splitting the data into 80% training and 20% testing subsets to evaluate model performance on unseen data.

To ensure consistent scaling across features, the input data was standardized to have zero mean and unit variance using the StandardScaler from the scikit-learn library. This step is crucial for models like ANN, which are sensitive to the scale of input features. No feature engineering or imputation for missing values was required as the dataset was complete and ready for use.

3.2 Data normalization

Normalization is crucial for most machine-learning approaches because unnormalized data might produce poorly conditioned results.

To ensure the proper functioning of machine learning algorithms, standardization was applied to the input features. Each feature was transformed to have a mean of zero and unit variance using the StandardScaler from the scikit-learn library.

This step was particularly crucial for the ANN and ELM models, as these models are sensitive to the scale of the data. The output power (target variable) was not normalized since it is a continuous value and was used in its original form.

As an illustrative case in point, consider neural networks with a sigmoid activation function: for these networks, the $f(t)$ gradient will become very small if and only if t is very big. It renders the whole training process useless and inefficient. The BP algorithm, a classification method based on the ELM approach that classifies data based on their forms, was used for the classification of the demand data.

For those interested, the corresponding Python code was given to those interested in the Kaggle website [38].

3.3 Artificial neural networks

An artificial neural network typically requires a minimum of two hidden layers, each containing a varying number of neurons. These neurons, whether in linear or nonlinear relationships, perform mathematical operations to transform input into output.

In this study, a feedforward neural network with multiple hidden layers is utilized to predict the electrical output power of a CCPP. The network architecture involves several hidden layers positioned between the input and output layers. This feedforward structure ensures that information flows unidirectionally from input nodes through hidden layers to output nodes. While a single hidden layer with an optimal number of neurons can serve as a universal approximator, the use of multiple hidden layers provides substantial benefits.

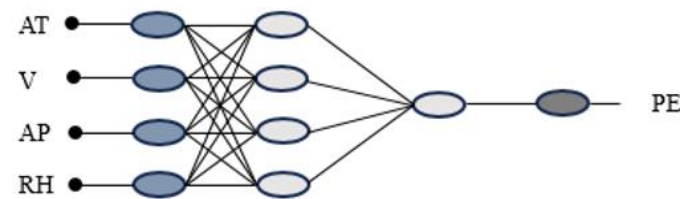


Figure 2. Distribution of inputs and output in ANN model

Figure 2 illustrates a distribution of inputs and output in ANN model. It has m input nodes and a single output node. The data vector input is represented by:

$$X = [x_1 x_2 \dots x_n]^T \quad (1)$$

where, the superscript T signifies the matrix's transposition. The initial hidden layer is made of n neurons. The connection of the input layer weight matrix to the first hidden layer is:

$$W_1 = \begin{bmatrix} w_{111} & w_{112} & \dots & w_{11m} \\ w_{121} & w_{122} & \dots & w_{12m} \\ \vdots & \vdots & \ddots & \vdots \\ w_{1n1} & w_{1n2} & \dots & w_{1nm} \end{bmatrix} \quad (2)$$

First-hidden-layer bias is defined as:

$$B_1 = [b_{11} b_{12} \dots b_{1n}]^T \quad (3)$$

The activation function in hidden layers is denoted by $F(\cdot)$. Two traditional activation functions that include nonlinearity in neural networks are the hyperbolic tangent and the sigmoid functions. The sigmoid function returns a value in the interval $[0, 1]$, which helps conduct probability computations. The sigmoid and hyperbolic tangent functions have similar structures, except for horizontal rescaling and vertical translation to $(1, 1)$. This application's hyperbolic tangent function is superior to the sigmoid function. In addition, it is mean-centering and has a more significant gradient than the sigmoid, making it more straightforward to train. The definition for the hyperbolic tangent function is:

$$(z) = \tanh(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}} \quad (4)$$

Therefore, the output of the first hidden layer is obtained by:

$$H_1 = \tanh(W_1 \times X + B_1) \quad (5)$$

The second hidden layer is composed of p neurons. The connection weight between two hidden layers is:

$$W_2 = \begin{bmatrix} w_{211} & w_{212} & \dots & w_{21m} \\ w_{221} & w_{222} & \dots & w_{22m} \\ \vdots & \vdots & \ddots & \vdots \\ w_{2p1} & w_{2p2} & \dots & w_{2pn} \end{bmatrix} \quad (6)$$

The second hidden layer bias is given as:

$$B_2 = [b_{21} b_{22} \dots b_{2p}]^T \quad (7)$$

Therefore, the output for the second hidden layer is:

$$H_2 = \tanh(W_2 \times H_1 + B_2) \quad (8)$$

The network output is:

$$y = W_3 \times H_2 \quad (9)$$

where, W_3 is the connection weight from the second hidden layer to the output and is defined by:

$$W_3 = [w_{31} w_{32} \dots w_{3p}] \quad (10)$$

In this study, a feedforward neural network with two hidden layers is employed. The input layer comprises 5 neurons representing dataset features. The first hidden layer contains 64 neurons, while the second layer has 32 neurons. The ReLU activation function is chosen for its non-linearity and ability to address the vanishing gradient issue. The output layer consists of one neuron with a linear activation function, predicting the continuous target variable of hourly power output. The model is trained using MSE loss and optimized with the Adam.

3.4 Extreme learning machine

Recently, Huang et al. [31] have proposed a technique for training single hidden layer feedforward neural networks, termed "extreme learning machine" (ELM). Then, the input weights of hidden nodes are randomly chosen. SLFN output weights are computed using the pseudoinverse function of the hidden layer output matrix.

There are two phases involved in the ELM algorithm: feature mapping and solving the output weight.

The mapping process of ELM feature: For generalized Single Layer Feedforward Networks (SLFNs), the output function for input data x is:

$$f(x) = \sum_{i=1}^L \beta_i h_i(x) = h(x) \beta \quad (11)$$

And $h(x) = [h_1(x), \dots, h_L(x)]$ represents a hidden layer's output vector, where $\beta = [1, \dots, L]^T$ is the inverse.

The hidden layer, which has L hidden nodes, is connected to the output layer by T weights. ELM feature mapping is the process of obtaining h , it takes input data from RD to feature space RL. In practical uses, h is characterized as:

$$h_i(x) = g(a_i b_i, X), a_i \in R^D, b_i \in R \quad (12)$$

An activation function as $g(a, b, x)$ satisfies theorems relative to ELM's ability to provide a universal approximation. As activation function h , any nonlinear independently continuous function (such as Gaussian, Sigmoid, etc.) could be employed. The parameter values h in ELM are generated randomly based on a continuous probability distribution.

During the second phase of calculating the output weights for ELM, given a training sample set, $(X_i, t_i)_{i=1}^n$ with $t_i = [0, \dots, 0, 1, 0, \dots, 0_m]^T$ is satisfied.

ELM uses to reduce the training error and the Frobenius norm of output weights for the x_i -class indicator. This objective function can be described as follows for each binary and classification with multi-class tasks.

$$\min_{\beta, \xi} \frac{\omega}{2} \sum_{i=1}^n \|\xi\|_2^2 + \frac{1}{2} \|\beta\|_F^2 \quad (13)$$

$$s.t. \beta h_i(x) = t_i - \xi_i, \forall i=1, 2, \dots, n \quad (14)$$

where, n represents the number of samples, and signifies the training error of the i -th sample, ω is a regularization parameter that balances the average of output weights, training error, and the Frobenius norm. The optimization problem defined in Eq. (13) will be simply resolved. In particular, the optimal can be derived analytically according to the Woodbury identity [30].

$$\beta^* = \begin{cases} \left(H^T H + \frac{1}{\omega} \right) H^T T & \text{if } L \leq k \\ H^T \left(H H^T + \frac{1}{\omega} \right)^{-1} T & \text{otherwise} \end{cases} \quad (15)$$

If not, I_n and I_L are identifying matrices, while H denotes the output matrix, which is characterized as a randomized matrix of the hidden layer, which is described in Eq. (15):

$$H = \begin{bmatrix} h(x_1) \\ \vdots \\ h(x_n) \end{bmatrix} = \begin{bmatrix} h_1(x_1) & \dots & h_L(x_1) \\ \vdots & \ddots & \vdots \\ h_1(x_n) & \dots & h_L(x_n) \end{bmatrix} \quad (16)$$

The extreme learning machine (ELM) model features a single hidden layer with 1,000 neurons using a sigmoid activation function. ELM initializes the weights and biases randomly in the input layer and computes output weights using the Moore-Penrose pseudo-inverse, which avoids iterative optimization. This configuration allows faster training while maintaining high accuracy. Both models used the same training and testing datasets to ensure a fair comparison of their performance.

4. SIMULATION AND RESULTS

4.1 Materials and methods

In this article, we developed ELM and ANN algorithms for forecasting the hourly electrical power output (EP) of a CCPP with the purpose of monitoring the performance and related efficiency, in addition to the effective utilization of its power output, quite simply this comparison will be made between the mean squared error values.

Statistical analysis needs software, and we opted for Python over R or MATLAB due to their modern and more usable in big data.

In this work, the Python 3.9 version is used, Anaconda distribution is used with the following packages: NumPy, pandas, matplotlib, ELM and TensorFlow. All the tests were done on a high-quality PC made by ASUS.

4.2 Application of the proposed ANN and ELM models in power prediction

The data is generated and simulated with five features and a target variable representing the power output, using a combination of the feature sums and some added noise to simulate realistic data.

For both models, we used an 80/20 training-test split and trained the models using the training data. The models were optimized using their respective loss functions, and the training process was terminated after 100 epochs for ANN and once the weights were computed for ELM. The data is standardized to have zero mean and unit variance, which helps improve model performance.

Table 2. Compilation of ANN

Epochs	Loss	Val_Loss
1	204973.0469	201167.7656
2	189543.8750	172472.3906
9	7557.0292	6213.2607
10	5347.1230	4402.7392
19	446.423	365.673
20	314.680	257.568

Table 3. Comparison between real values and predictions

Data Points	Real Values	Predictions
0	433.27	435.462
1	438.16	437.506
2	458.42	461.626
...
1910	438.04	432.152
1911	467.80	467.066
1912	437.14	431.761

A simple feedforward neural network is built with two hidden layers and trained using the loss function.

Table 2 contains the compilation of ANN and Table 3 shows the comparison between real values and predictions.

Both the ANN and ELM models were evaluated on the test set (20% of the data), using the following performance metrics.

The following calculations were used to calculate MSE, MAE, RMSE and R^2 for improving model prediction accuracy:

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (e_t)^2}{n}} \quad (17)$$

$$MAPE = 100\% \frac{1}{n} \sum_{t=1}^n \left| \frac{e_t}{A_t} \right| = 100\% \frac{1}{n} \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right| \quad (18)$$

$$R^2 = 1 - (RSS/TSS) \quad (19)$$

where, A_t denotes the experimental values, F_t signifies the obtained values, when n represents the total number of observations and the following figures show the learning results of the proposed models.

Figure 3 presents a graphical representation of predictions and evaluate model performance model of ANN.

Figure 4 represents the scatter plot of actual values and predictions, in addition, Figure 5 represents ANN and ELM residuals and finally the graphical representation of statistical criteria in Figure 6.

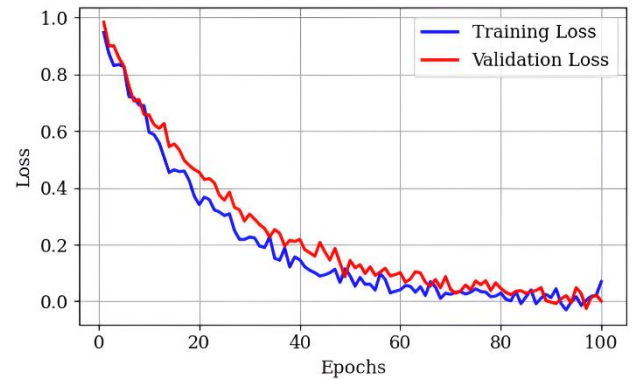


Figure 3. Graphical representation of ANN training and validation Loss

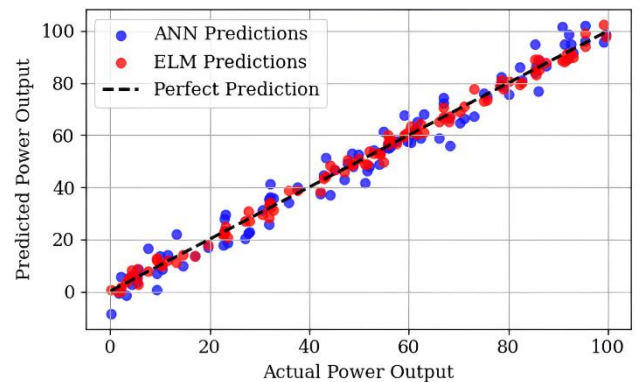


Figure 4. Scatter plot of actual values and predictions

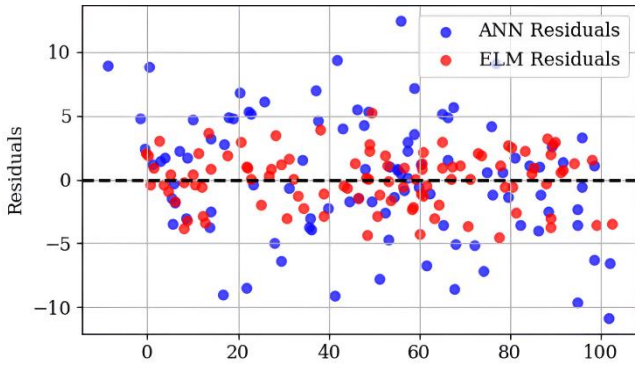


Figure 5. Scatter plot of residuals

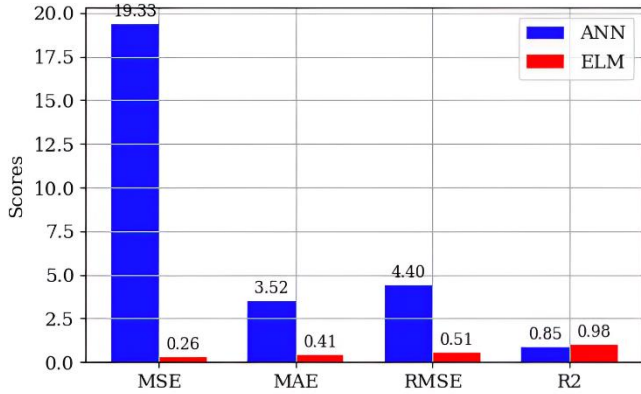


Figure 6. Graphical representation of statistical criteria

Table 4. Results of predictions

Statistical Evaluation Criteria	ANN	ELM
MAE	3.51507	0.4062
MSE	19.3280	0.2579
RMSE	4.3963	0.5078
R ²	0.85	0.98

Table 4 presents a summary of the prediction results for each of the methodologies: ANN and ELM.

From Table 4, the average absolute error is only 3.51 which is so small, which shows that the model almost corresponds to the real values and the root mean square error has a value close to zero 4.3 so the model is very good.

We found the general RMSE value of the ELM method is 0.5 plus close to zero compared to the first neural RNA technique. Therefore, ELM is a model that gives an excellent result with as little curve error as possible. The ELM model took 60 iterations (epochs) to improve its results until it reached its final RMSE value.

The performance of the present ELM and ANN approaches are depended on randomly initialized initial input weights and biases.

The regression method automatically identifies the appropriate number of hidden neurons, n , using the two approaches described above.

The suggested method demonstrates better training time results than backpropagation methods.

Figure 6 clearly shows that the RMSE of the ELM algorithm's findings is inferior to those of the backpropagation ANN algorithms and the R^2 for the ANN is 0.8502, suggesting that approximately 85% of the variance in the output can be explained by the model.

In contrast, the ELM has an R^2 of 0.9801, indicating that

about 98% of the variance is explained by the ELM model, signifying a superior fit.

Finally, the proposed ANN and ELM models improve accuracy in training and validation data sets.

As result, the findings demonstrate that the ELM model significantly outperforms the ANN model across all performance metrics. ELM achieved an MSE of 0.2579, MAE of 0.4062, RMSE of 0.5078, and an R^2 of 0.9801, while the ANN produced an MSE of 19.3280, MAE of 3.51507, RMSE of 4.3963, and an R^2 of 0.8502. These findings highlight the superior accuracy and generalization capability of the ELM model.

The advantages of ELM are mainly due to its simpler architecture and non-iterative training process, which help reduce overfitting risks and computational demands. On the other hand, the lower performance of the ANN model may be attributed to a bias-variance tradeoff issue: the network architecture might have been insufficiently complex to capture all data patterns, or it may have overfitted due to noise in the dataset. Improving the ANN, through techniques like hyperparameter tuning and regularization, could further enhance its predictive performance.

The implications of these results for CCPP operations are significant. Accurate power output prediction using machine learning models like ELM and ANN can optimize fuel consumption, improve maintenance scheduling, and increase overall plant efficiency. Among these, ELM stands out as a particularly effective tool due to its higher accuracy, faster training times, and scalability. Implementing ELM in real-world CCPP operations allows plant managers to make data-driven decisions that boost productivity and reduce operational costs.

While ELM demonstrated excellent performance in this study, it is essential to validate these results with different datasets or under various operational conditions to ensure robustness. Future work could include feature engineering, hyperparameter tuning, and cross-validation techniques to improve the performance and adaptability of both models across different industrial scenarios.

5. CONCLUSION

This paper presents the estimation of hourly power output of CCPP using two machine learning techniques, namely ANN and ELM. In training and testing, the models were used with a publicly available dataset from a CCPP that was collected over six years. The main objective was to compare the predictive performance of ANN and ELM in power output forecasting and to identify the model that gives the best predictions.

Our investigation concluded that ELM outperformed ANN in all the assessed performance metrics, indicating that ELM has higher predictive accuracy. These results obviously prove that ELM gives more accurate and reliable predictions of CCPP power output compared to ANN.

These findings evidence ELM's probable role as an efficient way of power prediction in the CCPP in real time with strong generalization capability due to the faster training in ELM, which would especially suit those applications requiring quick and accurate predictions because it optimizes plant performance-energy management.

Future work may be done by hyperparameter tuning, inclusion of more input features, and testing other advanced machine learning algorithms to see their performance for

improving the quality of predictions on different power plants. Cross-validation on more diverse datasets will give further validity to the results so that these are applicable to practical situations.

REFERENCES

- [1] Wikner, A., Harvey, J., Girvan, M., Hunt, B.R., Pomerance, A., Antonsen, T., Ott, E. (2024). Stabilizing machine learning prediction of dynamics: Novel noise-inspired regularization tested with reservoir computing. *Neural Networks*, 170: 94-110. <https://doi.org/10.1016/j.neunet.2023.10.054>
- [2] Thapar, V. (2023). Applications of machine learning to modelling and analysing dynamical systems. *arXiv preprint arXiv:2308.03763*. <https://doi.org/10.48550/arXiv.2308.03763>
- [3] Schötz, C., White, A., Gelbrecht, M., Boers, N. (2024). Machine Learning for predicting chaotic systems. *arXiv preprint arXiv:2407.20158*. <https://doi.org/10.48550/arXiv.2407.20158>
- [4] Liu, Y.B., Hong, W.X., Cao, B.Y. (2019). Machine learning for predicting thermodynamic properties of pure fluids and their mixtures. *Energy*, 188: 116091. <https://doi.org/10.1016/j.energy.2019.116091>
- [5] Nathan Kutz, J. (2023). Machine learning methods for constructing dynamic models from data. In *Machine Learning in Modeling and Simulation: Methods and Applications*, pp. 149-178. https://doi.org/10.1007/978-3-031-36644-4_4
- [6] Pourmohammad Azizi, S., Neisy, A., Ahmad Waloo, S. (2023). A dynamical systems approach to machine learning. *International Journal of Computational Methods*, 20(9): 2350007. <https://doi.org/10.1142/S021987622350007X>
- [7] Elwardany, M., Nassib, A.M., Mohamed, H.A. (2024). Exergy analysis of a gas turbine cycle power plant: A case study of power plant in Egypt. *Journal of Thermal Analysis and Calorimetry*, 149(14): 7433-7447. <https://doi.org/10.1007/s10973-024-13324-z>
- [8] 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. <https://doi.org/10.1016/j.ijepes.2014.02.027>
- [9] Khan, M.S., Xuebing, P., Yuntao, S., Bin, G., Imran, M. (2024). An optimization of efficient combined cycle power generation system for fusion power reactor. *Case Studies in Thermal Engineering*, 57: 104344. <https://doi.org/10.1016/j.csite.2024.104344>
- [10] Nagaraju, S., Bhat, C.R., Baskar, M., Rajaram, G., Rani, L.P., Bhaskar, K.V. (2024). Enhancing the accuracy of forecasting of the pattern of load usage of the electrical appliances via soft computing. In *2024 4th International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, Greater Noida, India, pp. 1900-1903. <https://doi.org/10.1109/ICACITE60783.2024.10617251>
- [11] Wang, C. (2024). Analysis of key factors influencing electrical load profiles. In *2024 3rd International Conference on Energy, Power and Electrical Technology (ICEPET)*, Chengdu, China, pp. 1352-1355. <https://doi.org/10.1109/ICEPET61938.2024.10627430>
- [12] 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. <https://doi.org/10.1002/OCA.3152>
- [13] Siddiqui, R., Anwar, H., Ullah, F., Ullah, R., Rehman, M.A., Jan, N., Zaman, F. (2021). Power prediction of combined cycle power plant (CCPP) using machine learning algorithm-based paradigm. *Wireless Communications and Mobile Computing*, 2021(1): 9966395. <https://doi.org/10.1155/2021/9966395>
- [14] Anđelić, N., Lorencin, I., Mrzljak, V., Car, Z. (2024). On the application of symbolic regression in the energy sector: Estimation of combined cycle power plant electrical power output using genetic programming algorithm. *Engineering Applications of Artificial Intelligence*, 133: 108213. <https://doi.org/10.1016/J.ENGAPPAI.2024.108213>
- [15] Rodríguez, E.Y.A., Gamboa, A.A.R., Rodríguez, E.C.A., da Silva, A.F., Rizol, P.M.S.R., Marins, F.A.S. (2022). Comparison of adaptive neuro-fuzzy inference system (ANFIS) and machine learning algorithms for electricity production forecasting. *IEEE Latin America Transactions*, 20(10): 2288-2294. <https://doi.org/10.1109/TLA.2022.9885166>
- [16] Yao, K., Zheng, Y. (2023). Fundamentals of machine learning. In *Nanophotonics and Machine Learning: Concepts, Fundamentals, and Applications*, pp. 77-112. https://doi.org/10.1007/978-3-031-20473-9_3
- [17] Degadwala, S., Vyas, D. (2024). Survey on systematic analysis of deep learning models compare to machine learning. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 10(3): 556-566. <https://doi.org/10.32628/CSEIT24103206>
- [18] Herceg, S., Ujević Andrijić, Ž., Rimac, N., Bolf, N. (2023). Development of mathematical models for industrial processes using dynamic neural networks. *Mathematics*, 11(21): 4518. <https://doi.org/10.3390/math11214518>
- [19] Hong, W.K. (2023). *Artificial Intelligence-Based Design of Reinforced Concrete Structures: Artificial Neural Networks for Engineering Applications*. Elsevier.
- [20] Rogachev, I.K., Shushnov, M.S. (2024). Neural Networks in Industry. In *Conference: Modern Problems of Telecommunications-2024*, pp. 191-193. <https://doi.org/10.55648/SPT-2024-1-191>
- [21] Esteghamati, M.Z., Bean, B., Burton, H.V., Naser, M.Z. (2024). Beyond development: Challenges in deploying machine learning models for structural engineering applications. *arXiv Preprint arXiv:2404.12544*. <https://doi.org/10.48550/arXiv.2404.12544>
- [22] Abdolrasol, M.G., Hussain, S.S., Ustun, T.S., Sarker, M.R., et al. (2021). Artificial neural networks based optimization techniques: A review. *Electronics*, 10(21): 2689. <https://doi.org/10.3390/ELECTRONICS10212689>
- [23] Xezonakis, V., Ntantis, E.L. (2023). Modelling and energy optimization of a thermal power plant using a multi-layer perception regression method. *WSEAS Transactions on Systems and Control*, 18: 243-254. <https://doi.org/10.37394/23203.2023.18.24>
- [24] Arferiandi, Y.D., Caesarendra, W., Nugraha, H. (2021).

- Heat rate prediction of combined cycle power plant using an artificial neural network (ANN) method. *Sensors*, 21(4): 1022. <https://doi.org/10.3390/S21041022>
- [25] Farajollahi, A., Baharvand, M., Takleh, H.R. (2024). Modeling and optimization of hybrid geothermal-solar energy plant using coupled artificial neural network and genetic algorithm. *Process Safety and Environmental Protection*, 186: 348-360. <https://doi.org/10.1016/j.psep.2024.04.001>
- [26] Karaçor, M., Uysal, A., Mamur, H., Şen, G., et al. (2021). Life performance prediction of natural gas combined cycle power plant with intelligent algorithms. *Sustainable Energy Technologies and Assessments*, 47: 101398. <https://doi.org/10.1016/j.seta.2021.101398>
- [27] Huang, G.B., Zhu, Q.Y., Siew, C.K. (2004). Extreme learning machine: A new learning scheme of feedforward neural networks. In 2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No.04CH37541), Budapest, Hungary, pp. 985-990. <https://doi.org/10.1109/IJCNN.2004.1380068>
- [28] Huang, G.B., Zhu, Q.Y., Siew, C.K. (2006). Extreme learning machine: Theory and applications. *Neurocomputing*, 70(1-3): 489-501. <https://doi.org/10.1016/j.neucom.2005.12.126>
- [29] Deepika, K.K., Varma, P.S., Reddy, C.R., Sekhar, O.C., Alsharef, M., Alharbi, Y., Alamri, B. (2022). Comparison of principal-component-analysis-based extreme learning machine models for boiler output forecasting. *Applied Sciences*, 12(15): 7671. <https://doi.org/10.3390/AP12157671>
- [30] Markowska-Kaczmar, U., Kosturek, M. (2021). Extreme learning machine versus classical feedforward network: Comparison from the usability perspective. *Neural Computing and Applications*, 33(22): 15121-15144. <https://doi.org/10.1007/s00521-021-06402-y>
- [31] Huang, H., Wang, H., Peng, J. (2020). Optimal prediction intervals of wind power generation based on FA-ELM. In 2020 IEEE Sustainable Power and Energy Conference (iSPEC), Chengdu, China, pp. 98-103. <https://doi.org/10.1109/iSPEC50848.2020.9350964>
- [32] Sarira, Y.I., Syafaruddin, S., Gunadin, I.C., Utamidewi, D. (2024). Modeling and control of a based extreme learning machine as distributed setpoint for the HEPP cascade system in a nickel processing plant. *Journal of Applied Data Sciences*, 5(2): 628-635. <https://doi.org/10.47738/JADS.V5I2.211>
- [33] Liu, G. (2023). The application of fault diagnosis techniques and monitoring methods in building electrical systems—based on ELM algorithm. *Journal of Measurements in Engineering*, 11(4): 388-404. <https://doi.org/10.21595/JME.2023.23357>
- [34] Bera, A., Mande, S. (2024). Machine learning based computationally efficient approach for accurate prediction of power integrity performance of power distribution networks. In 2024 28th International Symposium on VLSI Design and Test (VDAT), Vellore, India, pp. 1-5. <https://doi.org/10.1109/VDAT63601.2024.10705749>
- [35] Kumar, N., Katiyar, A. (2024). Artificial intelligence and the future of renewable energy: Solar PV power forecasting. *New Innovations in AI, Aviation, and Air Traffic Technology*. <https://doi.org/10.4018/979-8-3693-1954-3.CH001>
- [36] Zhou, L., Zhang, Y., Guo, S. (2024). Research on new energy generation power prediction method based on SSA-ELM. In 2024 6th International Conference on Communications, Information System and Computer Engineering (CISCE), Guangzhou, China, pp. 1054-1058. <https://doi.org/10.1109/CISCE62493.2024.10653171>
- [37] UCI Machine Learning Repository (2014). Combined cycle power plant. <https://archive.ics.uci.edu/dataset/294/combined+cycle+power+plant>.
- [38] Combined cycle power plant data set - UCI data. Kaggle. <https://www.kaggle.com/datasets/rinichristy/combined-cycle-power-plant-data-set-uci-data>.

NOMENCLATURE

mW	MilliWatt
°C	Celsius degree
cmHg	Millimeters of mercury
mbar	Millibar