Machine Learning-Based Forecasting Active Power Loss in Distribution Systems

Waseem Haider[1,*], *Seema Batool*[2], *Federica Milazzo*[1], and *Quang P. Ha*[1,*]

[1]Faculty of Engineering and IT, University of Technology Sydney, Sydney, Australia
[2]Department of Development and Environmental Studies, Paris-Saclay University, Paris, France

Abstract. This paper presents an ensemble learning approach to predict the active power losses during the allocation and sizing of distributed generation (DG) units in power distribution networks. The forecast model incorporates the Gradient Boosting Machine Regression (GBMR) to estimate DG location, bus voltages, DG size, and active losses without conventional power flow calculations. The results demonstrate that the suggested estimations of power losses and DG sizing are effective, practical, and adaptable for power system management. The accuracy of the proposed model has been validated using key performance metrics and tested on the standard IEEE 33 bus system. In the case of fixed load, the GBMR outperforms other machine learning techniques with the R-squared 0.9997, with a very low mean absolute percentage error (MAPE) (0.2216%) and a root mean square error (RMSE) of 1.0673 in predicting active power losses. This approach is promising in enabling grid operators to effectively manage DG unit integration of distributed energy resources from precise and reliable estimates of the power loss.

Keywords: Distributed Generation, Active Power Loss, Forecasting, Gradient Boosting Machines Regression.

1 Introduction

Towards the sustainable energy goal, the deployment of distributed energy resources (DERs) is transforming traditional power distribution systems into active distribution networks [1]. Modern electricity distribution systems have, however, encountered severe challenges in integrating to large-scale power grids. This is mainly because of their intermittent nature, which may lead to power quality issues at the consumer end, such as undervoltage, overvoltage, equipment overloading, and control system malfunctions. One of the key criteria for evaluating the efficiency and economy of a power system is the line losses, which indicate the proportion of electrical energy lost due to components such as resistors and inductors during transmission. A higher line loss rate can reduce the overall performance of the power grid. As this loss directly affects both the stability and safety of the system, its minimization is required for optimizing grid operations and economic benefits.

To estimate and examine losses in distributed energy systems, the computational intelligence methodology has been increasingly applied in addition to theoretical calculations. In [2], an association rule method was proposed to extract the characteristics of the network loss sequence and used a forecasting approach for losses with the integration of distributed power. Using artificial neural networks (ANN), a voltage magnitude and line-loading monitoring scheme was introduced in [3]. In [4], an ANN model was utilized to forecast the turbine's output power over short, medium, and long-term periods. Wind speed and turbine output power are used as inputs, and the output layer predicts the wind turbine's power output. Grey correlation analysis and neural networks have been proposed in [5] for predicting 10kV line losses.

Recently, deep learning has been proposed for power loss estimation. The deep neural network (DNN) in [6] utilized mutiple hidden layers to capture the nonlinear relationship in predicting line losses for large-scale photovoltaic and electric heating losses in low-votage distribution areas. Recurrent neural networks with long short-term memory has been applied to identify the faults leading to power line losses [7]. Machine learning with ensemble techniques has also been suggested. For example, an enhanced random forest approach was developed in [8] to calculate and analyze the theoretical line loss of microgrids. In [9], two machine learning models, namely the light gradient boosting machine and K-nearest neighbors were compared in anticipating solar energy generation for microgrid applications.

It is known that the integration of DERs up to a necessary level can improve the network performance in terms of bus voltage magnitude, address environmental concerns, and reduce line current as well as energy losses. Indeed, efforts devoted to effectively mitigate the power line loss problem also involve the development of strategies to maximize the penetration of active power with sizing and loca-

*Email: waseem.haider1@uts.edu.au, quang.ha@uts.edu.au

tion of DERs [10,11]. For example, the placement of photovoltaic sources and wind turbines presented in [12] by using a heuristic method could achieve minimal voltage deviation index and annual energy losses. The distribution system reported in [13] applied optimal allocation of power capacity and placement of solar photovoltaic (SPV)-based DG, resulting in reduced costs of DERs, decreased CO2 emissions, and enhanced penetration level. A genetic algorithm was adopted for the IEEE-34 system in [14] to minimize active power losses subject to voltage and harmonic constraints. More recently, algorithms based on improved clustering and isolated forest have been developed in [15] for analysis of line loss of distribution power systems.

Given all techniques for enhancing grid efficiency in achieving optimal energy management of distributed power systems, to improve their effectiveness it is essential to accurately forecast reduction of active power losses. This is because precise prediction of power loss is of crucial importance for ensuring reliable power delivery at reduced operational costs. From accurate estimation, utilities can identify weak points in the system, enabling targeted maintenance and infrastructure improvements. This proactive approach also helps minimize fuel consumption and lower emissions, leading to higher economic and more environmentally sustainable performance.

This study addresses the prediction of the active power losses of distribution networks using a machine learning (ML)-based technique. As power networks are subject to intermittent operations, we choose to obtain a forecasting model from weak predictive models generated from past data with gradient boosting machines [16]. This ensemble learning algorithm, Gradient Boosting Machine Regression (GBMR), utilizes gradient boosting for regression from decision trees as base learners to estimate the power losses. The model effectiveness is demonstrated through performance metrics commonly-used in machine learning for evaluation of accuracy and reliability testing on the IEEE 33-bus distribution system. The merits of the proposed model are confirmed through a comparative analysis with other ensemble learning techniques. This approach allows grid operators to effectively manage DG unit integration by providing accurate estimates of power losses with respect to loading conditions, DER size and location.

The paper is organized as follows. After the introduction in Section I, Section II provides the problem formulation and constraints. The proposed forecast model is presented in Section III. Simulation results are included in Section IV along with a discussion. Finally, a conclusion is drawn in Section V.

2 Problem Formulation

Given the increasing complexity of modern electricity networks, power loss prediction is particularly important due to the decentralized nature of generation and the variable output of renewable energy sources. Accurate forecasting of power losses helps optimize the system performance, reduce energy waste, and ensure the stable integration of DG sources into the grid. As distributed generation systems often involve multiple interconnected sources with varying capacities and loads, the task of power loss prediction becomes more challenging and critical for maintaining grid reliability, minimizing operational costs, and improving the overall economic feasibility of the system.

2.1 Power Loss Function and Variables

In this study, our aim is to predict the power losses P_{loss} based on three input variables: load (L), size of active power injection (P), and location (Loc) of in the IEEE-33 bus system used for experimentation. Data for the function $P_{\text{loss}}(L,P,Loc)$ collected or generated from the system are then divided into two parts, the training dataset, D_{train}, and the testing dataset, D_{test}. The prediction of the power loss function $\hat{P}_{\text{loss}}$ can be obtained by

$$\hat{P}_{\text{loss}} = \frac{1}{n} \sum_{(L_k,P_k,Loc_k)\in D_{\text{test}}} \hat{P}_{\text{loss}}(L_k,P_k,Loc_k), \tag{1}$$

where $\hat{P}_{\text{loss}}(L_k,P_k,Loc_k)$ is the predicted power loss for each sample, k is an index that represents the individual samples in a dataset, and n is a number of samples in the testing dataset D_{test}.

2.2 System Description

The IEEE-33 bus radial system includes 33 nodes, 37 lines, 32 loads, 32 voltage- and reactive power-controlled (PQ) buses, one feeder, and one slack bus. The system functions with 32 closed and 5 open switches, with power supplied at bus 1, to maintain a steady voltage of 12.66 kV. The loads are treated as constant, with a total active power of 3715 kW and reactive power of 2300 kVAr. The system variables involved are subjected to the following constraints:

A) Voltage constraints: Let V_{min} and V_{max} be the minimum and maximum bus voltage limits, with values respectively of 0.95 pu and 1.05 pu in this study. We have:

$$V_{min} \le V_i \le V_{max}, \tag{2}$$

where V_i is the voltage at the ith node.

B) Line current constraints: The current flowing in each bus I_i should be at or below its maximum capacity rating to prevent overloading feeder lines considering the thermal effect and to maintain the grid operation. Therefore,

$$I_{min} \le I_i \le I_{max}, \tag{3}$$

where I_{min} and I_{max} are the minimum and maximum acceptable current I_i limits.

C) DGs constraints: The capacity of DG units at each bus i is also subject to both minimum and maximum generation limits, which are represented as follows:

$$P^{DG}_{min} \le P^{DG}_i \le P^{DG}_{max}, \tag{4}$$

where P^{DG}_{min} and P^{DG}_{max} are the lowest and highest acceptable active power.

 https://doi.org/10.1051/e3sconf/202562604003

3 GBMR for Active Power Loss Forecasting

In this study, the goal is to predict the active power losses of the system by considering three input variables: fixed load (L), the active power of the injected DG units (P), and their location (Loc). The Gradient Boosting machine regression (GBMR) and Decision tree regression are used for training and testing with our datasets. Here, two cases were considered: Case 1 for fixed load and Case 2 for variable load. Key performance metrics collectively offer a comprehensive evaluation of the accuracy and reliability of the proposed model across different scenarios. After penetration of DG units of various sizes at various locations in the IEEE-33 bus power system, the power loss and voltage data are obtained from the load flow calculations using the Newton-Raphson method, and saved in CSV format to be used for the purpose.

3.1 Prediction Model Development

The ensemble learning technique used, GBMR, provides the best prediction from a series of decision trees or weaker predictive models generated from the system data, as illustrated in Fig. 1. Each tree aims to rectify the error made by the previous trees [16]. The model initializes with a constant prediction, usually, the mean of the target values, and iteratively fits new trees to the residuals of the predictions. GBMR is used for its high accuracy and flexibility on regression tasks. The final prediction $\hat{y}$ is the sum of all predictions with contribution of all trees, scaled by a learning rate α:

$$\hat{y} = \hat{y}_0 + \sum_{m=1}^{M} \alpha \cdot f_m(x), \tag{5}$$

where $f_m(x)$ is the prediction from m-th tree from M total number of trees in the model.

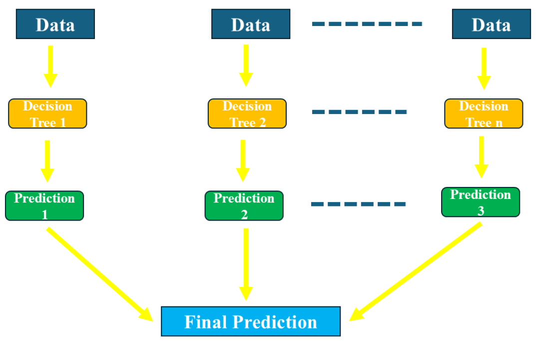

Figure 1: Diagram of gradient boost trees for regression

The following steps are involved in the prediction process:

- Data Validation Check:

$$\text{If exists NaN or Inf in data, then clean the data.}$$

- Data Organization: Define input features X and output labels Y:

$$X = \begin{bmatrix} x_{1,1} & x_{1,2} & \dots & x_{1,N} \\ x_{2,1} & x_{2,2} & \dots & x_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ x_{P,1} & x_{P,2} & \dots & x_{P,N} \end{bmatrix}, \quad Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_P \end{bmatrix}, \tag{6}$$

where the rows correspond to P observations, and N is the number of features.

- Model Development: After splitting system data into the training and testing sets, D_{train} and D_{test}, utilize the dataset D_{train} to train a predictive model to learn the relationship between the output (power losses) and the input variables. Here, training parameters for GBMR are shown in Table 1.

Table 1: Training parameters for gradient boosting machine model

Parameters	Value
Number of Trees M ($numTrees$)	100
Method	LSBoost
Number of Learning Cycles	100
Learner Type	Decision Tree
Max Number of Splits	100
Dataset Division	75% training, 25% testing

- Model Training: Train the ensemble model with least-square Gradient Boosting:

$$\begin{aligned} &\text{GBMModel} = \text{fitrensemble}(X_{\text{train}}, Y_{\text{train}}, \\ &\text{Method} =' LSBoost', \text{NumLearningCycles} = 100) \end{aligned} \tag{7}$$

- Prediction: Predict outputs for the test data:

$$Y_{\text{pred}} = \text{predict}(\text{GBMModel}, X_{\text{test}}) \tag{8}$$

- Hyperparameter Tuning: The key hyperparameters of gradient-boosted regression models include the learning rate, number of trees, and maximum depth, which determine model complexity and performance. Adjust the hyperparameters to optimize performance, using the cross-validation technique.

- Validation: Set aside a portion of D_{train} for validation to ensure the model generality.

$$\begin{aligned} \text{valid\_idx} = \sim (&\text{isNAN}(Y_{\text{test}}) \vee \text{isInf}(Y_{\text{test}}) \vee \\ &\text{isNAN}(Y_{\text{pred}}) \vee \text{isInf}(Y_{\text{pred}})) \end{aligned} \tag{9}$$

where the logical OR ($\vee$) combines the conditions, and logical NOT ($\sim$) inverts the boolean array to mark valid indices.

- Objective Function: The predicted power loss $\hat{P}_{\text{loss}}$ (1) is used as the objective function specifically computed from predictions made on D_{test} to assess how well the model predicts power losses on unseen data.

$$Y_{\text{test valid}} = Y_{\text{test}}[\text{valid\_idx}], \quad Y_{\text{pred valid}} = Y_{\text{pred}}[\text{valid\_idx}] \tag{10}$$

3.2 Performance Evaluation Metrics

The accuracy of the proposed prediction model is evaluated by using mean absolute error (MAE), mean squared error (MSE), mean absolute percentage error (MAPE), and root mean squared error RMSE. The coefficient of determination (R^2) is also used to evaluate the reliability of the estimation.

Metric MAE measures the average magnitude of the errors in a set of predictions, without considering their direction:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|, \tag{11}$$

where n represents samples, y_i represents actual value, and $\hat{y}$ represents its prediction. For minimization of the prediction errors, commonly used are MSE and RMSE:

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y})^2, \tag{12}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}. \tag{13}$$

To provide a percentage-based assessment of how well the estimates match the actual values, MAPE is used:

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100. \tag{14}$$

The formula for R^2 is given by:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2}. \tag{15}$$

The pseudo-code for the proposed prediction model is presented in Algorithm 1.

Algorithm 1 Power flow and proposed forecasting methodology

Input: Read system data
Initialization:
Add DG units & Perform load flow analysis
Calculate:
DG_size, DG_location, Load_value, Power_loss & Voltage
Data Collection:
Collect and store data for:
DG_size, DG_location, Load_value, Power_loss & Voltage
Data Analysis:
Perform feature selection
Model Training & Testing:
Fit the collected data into a machine-learning model as in Eqs. (5) & (10)
Output:
Print model performance metrics as in Eqs. (11-15)

3.3 Comparison with other ML-based techniques

To highlight the merits of the proposed prediction model, we conduct a comparative analysis with other ML-based ensemble learning techniques. The decision tree regression (DTR), a non-linear predictive modeling technique commonly used in statistics and machine learning, is applied here. DTR is suitable for regression tasks where the link between the input features and the desired outcome output are complex and not easily captured by linear models, widely applied when interpretability is important with a medium-sized dataset. The dataset is recursively split based on feature values. Once the tree is built, predictions are made based on the average target value of the samples that reach each leaf node. If a sample x reaches leaf L, the prediction $\hat{y}(x)$ is given by [17]:

$$\hat{y}(x) = \frac{1}{n_L}\sum_{j=1}^{n_L} y_j, \tag{16}$$

where n_L is the number of samples in leaf L and y_j are the target values of those samples.

The second ML-based approach used for benchmarking with our proposed prediction model is the support vector regression (SVR) based on support vector machines, a powerful technique for classification and regression analysis [18]. SVR owing to its ability to address nonlinearities and independence from input dimensionality, is advantageous in high-dimensional spaces over bagging or ridge regressions. Its training process can be mathematically described as minimizing the following objective function:

$$\min \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}(L(y_i, f(x_i))), \tag{17}$$

where w is a weight vector, C is the regularization parameter over n training samples, and $L(y_i, f(x_i))$ is the epsilon-insensitive loss defined as,

$$L(y_i, f(x_i)) = \begin{cases} 0 & \text{if } |y_i - f(x_i)| \leq \varepsilon \\ |y_i - f(x_i)| - \varepsilon & \text{otherwise.} \end{cases} \tag{18}$$

After training, predictions $\hat{x}$ for the test data $x \in X_{\text{test}}$ are made using the following prediction function:

$$\hat{x} = w^T\phi(x) + b, \tag{19}$$

where $\phi(x)$ is the radial basis kernel function which signifies the input data features and b is a bias term.

4 Results and Discussion

For predicting power losses, load is often assumed to be constant and the DG output adjustable during DG allocation. However, in practice, both load and DG output are subject to continuous fluctuations. This variability complicates the calculation of losses and other parameters using power flow-based techniques, resulting in a somewhat cumbersome and time-consuming process. In this study, we considered two cases, namely Case 1 for active power injection at fixed load (FL) (100%) and Case 2 for power injection with normalized load variation (NLV) (50-100%). The power loss is influenced by factors such as DG size, location, load, and voltage profile. We tested the estimation model for the IEEE 33-bus distribution system, using the proposed gradient boost machine for regression to estimate the power loss.

4.1 Power Loss Forecasting Results

The predicted results of the proposed model are compared with to actual values and with DTR and SVR models are presented respectively in Fig. 2 for Case 1 and Fig. 3 for Case 2. The comparison results are summarized in Table 2.

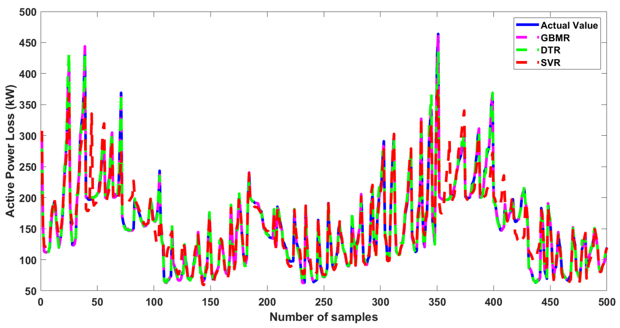

Figure 2: Active power losses prediction with Fixed Load

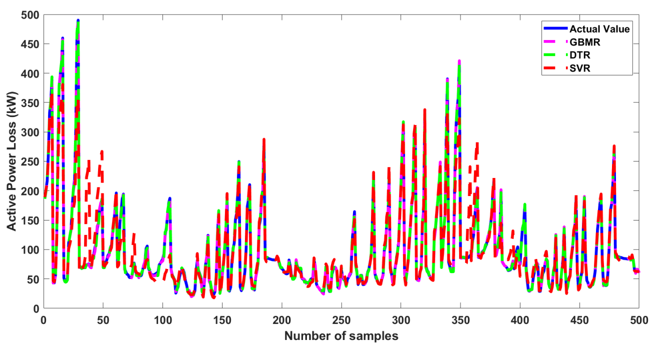

Figure 3: Active power losses prediction with Variable Load

In evaluating the performance of various models, GBMR emerges as the top performer across key performance indicators. It achieves a MAE of 0.3148, an MSE of 1.1391, a MAPE of 0.2216, and an RMSE of 1.0673, along with an impressive R-squared value of 0.9997 in the case of fixed load. This indicates an excellent fit to the data compared to the other models assessed. In Case 2 for active power injection using normalized load variation, the results indicate that GBMR also outperforms the other models, achieving the lowest MAE of 1.5864, an MSE of 5.6449, and a RMSE of 2.3759. Additionally, its R-squared value of 0.999 reflects a strong fit to the data, suggesting consistent predictions for the active power loss for the IEEE-33 system in consideration.

Figure 4 shows the comparison of performance metrics for both cases. DTR follows closely for Case 1 and Case 2, with solid performance but not as strong as with GBMR. Albeit being methodologically competitive, SVR, in contrast, displays higher errors in prediction and a lower R-squared value for both cases, making it less suitable for predicting active power losses with the obtained dataset for the IEEE-33 system mentioned in this study. Although SVR can model complex relationships in general, it comes with the trade-off in meeting a high accuracy and explanatory capability requirements.

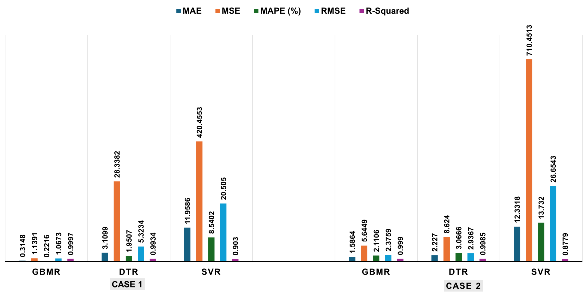

Figure 4: Performance evaluation results of MAE, MSE, MAPE, RMSE, and R-Squared using different model

4.2 Discussion

Since machine learning techniques depend on key performance indicators, such as data quality, feature selection, and algorithm choice, the accuracy of forecasting models is significantly influenced by the relevance and cleanliness of the training data, the selection of the right features as well as the suitable algorithms tailored to the specific task. Overall, these elements are crucial for developing effective and reliable models for accurately forecasting active power losses for management of DG unit penetration into distribution power networks.

Notably, our learning-based prediction model can determine the power losses of those networks without the need for complex and time-consuming load flow techniques. In terms of data processing, tree-based structures allow to capture complicated correlations and interactions between attributes, enabling intuitive interpretation with insightful features, making GBMR and DTR useful for this application, whereby GBMR appears more favourable for higher precision, particularly with larger datasets.

Our future work will focus on more advanced models for a complex network to accurately determine real-time DG sizes for predicting power losses. This is crucial to support more informed decision-making in energy management and resource optimization for efficient energy distribution regardless of dynamic conditions.Additionally, we also aim to improve model robustness in dealing with diverse data types and more independence on feature selection. This may involve exploring better techniques for hyperparameter tuning or the use of a deep learning approach.

5 Conclusion

This paper has presented an estimation-based method for forecasting active power losses in distribution power systems with DG units injected into the network. Without conventional power flow calculations, the proposed model utilizes a machine learning algorithm based on gradient boosting machine regression and tested across two distinct cases on the IEEE-33 bus radial power system. The predictive model merits are also confirmed through a comprehensive comparison analysis with two other ensemble learn-

Table 2: Comparative Analysis of Performance Metrics of Different Models

Metrics	Models	MAE	MSE	MAPE (%)	R-Squared	RMSE
Case 1	**GBMR**	0.3148	1.1391	0.2216	0.9997	1.0673
	DTR	3.1099	28.3382	1.9507	0.9934	5.3234
	SVR	11.9586	420.4553	8.5402	0.9030	20.505
Case 2	**GBMR**	1.5864	5.6449	2.1106	0.9990	2.3759
	DTR	2.227	8.624	3.0666	0.9985	2.9367
	SVR	12.3318	710.4513	13.732	0.8779	26.6543

ing models, namely the decision tree regression and support vector regression. Promising results obtained are of significant contribution to the effective management of smart grids towards energy sustainability.

References

[1] D. A. Copp, T. A. Nguyen, R. H. Byrne, and B. R. Chalamala, "Optimal sizing of distributed energy resources for planning 100% renewable electric power systems," *Energy*, vol. 239, p. 122436, 2022.

[2] G. Zheng, G. Chen, R. Deng, J. Yi, L. Huang, and J. Zhang, "The prediction method of distribution network loss under distributed power supply access," in *2023 3rd Int. Conf. New Energy and Power Engineering (ICNEPE)*, pp. 472–476, Nov. 2023.

[3] J. H. Menke, N. Bornhorst, and M. Braun, "Distribution system monitoring for smart power grids with distributed generation using artificial neural networks," *Int. J. Electr. Power Energy Syst.*, vol. 113, pp. 472–480, Dec. 2019.

[4] A. Sen, C. Andic, E. Aydin, M. Purlu, and B. Turkay, "Forecasting of wind turbine output power with artificial neural network in Izmir, Türkiye," in *2023 14th Int. Conf. Electrical and Electronics Engineering (ELECO)*, pp. 1–5, Nov. 2023.

[5] X. Ma, C. Liang, X. Dong, Y. Li, and R. Xu, "A line loss prediction method based on neural network," in *2022 3rd Int. Conf. Advanced Electrical and Energy Systems (AEES)*, pp. 249–254, Sep. 2022.

[6] J. Zhang, L. Wang, Y. Geng, M. Ren, J. Ma, and Y. Niu, "Line loss prediction of low voltage distributions considering mass PV and electric heating," in *2023 6th Int. Conf. Energy, Electrical and Power Engineering (CEEPE)*, pp. 1041–1046, May 2023.

[7] P. K. Shukla and K. Deepa, "Deep learning techniques for transmission line fault classification – A comparative study . *Ain Shams Engineering Journal*, Vol. 15, No. 2, February 2024, (p. 102427).

[8] L. Huang, G. Zhou, J. Zhang, Y. Zeng, and L. Li, "Calculation method of theoretical line loss in low-voltage grids based on improved random forest algorithm," *Energies*, vol. 16, no. 7, p. 2971, Jul. 2023.

[9] P. Suanpang and P. Jamjuntr, "Machine learning models for solar power generation forecasting in microgrid application: Implications for smart cities," *Sustainability*, vol. 16, no. 16, p. 6087, Aug. 2024.

[10] A. Hussain, S. Shah, and S. Arif, "Heuristic optimisation-based sizing and siting of DGs for enhancing resiliency of autonomous microgrid networks," *IET Smart Grid*, v.2, no.2, pp.269–282, 2019.

[11] W. Haider and Q. Ha, "Maximum Power Penetration of Distributed Energy Resources with Sizing and Location," in *The 10th IEEE Int. Conf. .Sustainable Technology and Engineering* (i-COSTE 2024), 18-20 Dec. 2024, Perth, Australia. To appear.

[12] M. Purlu and B. E. Turkay, "Optimal allocation of renewable distributed generations using heuristic methods to minimize annual energy losses and voltage deviation index," *IEEE Access*, vol. 10, pp. 21455–21474, 2022.

[13] E. D. Melaku, E. S. Bayu, C. Roy, A. Ali, and B. Khan, "Distribution network forecasting and expansion planning with optimal location and sizing of solar photovoltaic-based distributed generation," *Computers and Electr. Eng.*, vol. 110, p. 108862, Jul. 2023.

[14] J. Fu, Y. Han, W. Li, Y. Feng, A. S. Zalhaf, S. Zhou, P. Yang, and C. Wang, "A novel optimization strategy for line loss reduction in distribution networks with large penetration of distributed generation," *Int. J. Electr. Power Energy Syst.*, vol. 150, p. 109112, Aug. 2023.

[15] J. Li, S. Li, W. Zhao, J. Li, K. Zhang, and Z. Jiang, "Distribution network line loss analysis method based on improved clustering algorithm and isolated forest algorithm," *Sci. Rep.*, vol. 14, no. 1, p. 19554, 2024.

[16] A.Natekin and A.Knoll, "Gradient boosting machines, a tutorial," *Front. Neurorobotics*, vol. 7, p. 21, 2013.

[17] S. Suthaharan and S. Suthaharan, "Decision tree learning," in *Machine Learning Models and Algorithms for Big Data Classification*, S. Suthaharan, Ed., pp. 237–269, 2016.

[18] H. Drucker, C. J. C. Burges, L. Kaufman, A. Smola and V. Vapnik, "Support vector regression machines", in *Proc. the 9th Int. Conf. Neural Information Processing Systems*, 1996, pp. 155–161.