

Advanced Machine Learning Solutions For Optimizing Distributed Generation Placement In Distribution Systems

Dr. Harish S. Avchat¹, Amit B Kasar², Dr. Pooja Khandare³, Dr. Sandip P. Gath⁴, Shivaji Raskar⁵, Dr. Pravin D. Gunaware⁶, Dr. Sudhir N. Divekar⁷

^{1,6,7}*Assistant Professor, HSBPVT's Group of Institutions, Kashti, Maharashtra, India*

²*Assistant Professor, International Institute of Information Technology, Pune, Maharashtra, India*

³*Assistant Professor, Dattakala Group of Institutions Faculty of Engineering, Swami-Chincholi, Maharashtra, India*

⁴*Assistant Professor, Rajiv Gandhi College of Engineering, Ahmednagar, Maharashtra, India*

⁵*Assistant Professor, VPKBIT, Baramati, Maharashtra, India*
amit.kasar1982@gmail.com

The optimal placement of Distributed Generation (DG) in power distribution networks is critical for improving grid stability and minimizing power losses. Previous studies have applied conventional optimization techniques like “Genetic Algorithms (GA) as well as Particle Swarm Optimization (PSO)” to handle this issue. This paper addresses these challenges by presenting a new approach for optimal placement and sizing of DG units using the KNN algorithm. The advanced deep learning models such as “Recurrent Neural Networks (RNN) Long Short-Term Memory (LSTM) Convolutional Neural Networks (CNN)” are compared with the proposed KNN algorithm on the IEEE 33-bus and 69-bus systems. The results indicate that the KNN model has better prediction accuracy than the traditional and deep learning models; on the 33-bus system, KNN estimated an average reduction of 16.5 % of the power losses, while on the 69-bus system, it estimated a decrease of 17.8 % of the power losses. Also, it enhances voltage profiles at a much-reduced computation time than the traditional method, making it fit for real-time applications. The present work can be extended to combine KNN with other ML algorithms or predictive methods to improve DG's location.

Keywords: Distributed Generation, K-Nearest Neighbors, Machine Learning, Optimal DG Placement, Power Loss Minimization, Deep Learning, RNN, CNN, LSTM, 33-Bus System, 69-Bus System.

1. INTRODUCTION

The rapid transformation of energy systems, driven by the increasing penetration of renewable energy sources [1] and “distributed generation (DG)” [2], has led to new challenges and

opportunities in the operation and management of modern power distribution networks. Distributed generation systems include small-scale generating units like solar and wind power and small-scale CHP generators, and they are becoming essential in innovative grid applications. DG acknowledges the significance of several benefits, including less transmission losses, enhanced voltage stability, and increased system dependability. Nonetheless, integrating distributed generation into current distribution networks poses significant challenges in determining appropriate locations for installation to maximize benefits while mitigating any adverse effects on the grid. Another of the hardest coordination in the implementation of DG is assessing where the DG should be located and how big it should be to fit the distribution systems appropriately. Location misplacement of DG units results in problems like high power loss, voltage fluctuations, and the opposite flow [3] of power to the network, which would be detrimental to the grid. Hence, the location of suitable DG units should be determined [4] to maximize their benefits to the grid stability, minimize losses, and augment the loading demands. The problem of DG placement is unambiguously challenging because of the non-linear, dynamic behaviour of power systems, making the optimization problem non-trivial [5].

Various traditional optimization methods have solved the integration of suitable DG locations into distribution networks [6]. Several approaches have been used to deal with this problem, “including Genetic Algorithms (GA), Particle Swarm Optimization (PSO), Ant Colony Optimization (ACO), and other meta-heuristic algorithms” [7]. All these techniques target the correct location of DG units by reducing objectives like loss, voltage and other costs. Even though these methods have been successful for some applications, many things could be improved, especially in the large-scale problem in terms of computational time. The increasing complexity of supply networks presents new challenges that more conventional optimization approaches need to better address with equal effectiveness when evaluating all of the various influences that need to be taken into account.

Moreover, these methods can depend on parameter settings and may need many iterations to reach the optimum solution. This results in a high computational overhead [8] that deems them unsuitable for real-time or high-end systems. Also important to note is that classical approaches could be more effective in capturing the stochasticity of DG, especially when it comes to renewable generation, such as solar and wind, which fluctuate randomly and are not predictable.

Over the past few years, machine learning has been recognized as a reliable technique for addressing various optimization challenges that power systems encounter, such as integrating distributed generation (DG). Compared to conventional optimization methods, machine learning can identify trends within data and, to begin with, is less sensitive to the current distribution networks' intricacies. The nature of machine learning models allows large datasets to offer more precise and time-saving results to the DG placement problem and, in general, lower computational costs in many cases. “The DG placement problem has been solved using several techniques in machine learning, such as supervised, unsupervised, and reinforcement learning. For instance, given past load and network data, artificial neural networks (ANN) have been applied to estimate the best locations for installing DG units. Likewise, Support Vector Machines (SVM) have been used to identify DG locations according to load demand, voltage profile, and system limitations. Other techniques, such as Deep Q-Learning (DQL)

and Convolutional Neural Networks (CNN), have also been used to enhance the placement of DG to enhance the distribution system by considering the system's spatial and temporal dependency”.

In this paper, the KNN algorithm has been identified as helpful in positioning DG in distribution networks. This algorithm is one of the simplest and most effective forms of machine learning and classifies the data according to the majority class of neighbours. In the context of the placement of DG, the KNN tool can be applied to determine areas most suitable for DG unit's placement based on data such as load demand, voltage profile and network performance history. As we will see, KNN is computationally inexpensive and particularly beneficial as we move to more extensive distribution networks where decisions must be made in real-time. One of the key benefits of employing KNN as a machine learning solution is simplicity while being as efficient or even more efficient than other complex models like CNN, RNN, LSTM, and others that take much more time and resources to fine-tune and train. Moreover, KNN is a non-parametric technique for classification in which no assumptions are made relative to the underlying probability density functions. As such, KNN can be easily implemented in different distribution networks with DG integration.

The main Contributions of this research article are as follows:

- To Enhance grid stability and decrease power losses by proposing a KNN-based method for optimally placing DGs in power distribution networks [9].
- To Demonstrated that KNN outperforms deep learning models (CNN, RNN, LSTM) in accuracy and power loss reduction for “IEEE 33-bus and 69-bus systems” [10].
- To Showcase the computational efficiency of KNN, making it suitable for real-time DG placement in both small and large networks.
- To Provided groundwork for integrating KNN with other machine learning techniques for dynamic and real-time DG management.

The structure of the paper is as follows: The second section of the paper is Section 2: Literature Survey which examines previously known techniques in DG placement. “It talks about method that is involved in optimization like genetic algorithms, Particle swarm optimization and the modern machine learning approaches at the same time, what the pros and cons to each. Section 3: The Proposed Work encapsulates methodology of applying the KNN algorithm through which the DG placement in PDNs is done efficiently. Section 4: Result Analysis gives the comparative assessment of the proposed KNN algorithm with other models such as CNN, RNN, LSTM having the use of IEEE 33-bus and 69-bus systems”. The comparison is done on factors like accuracy and reduction of power loss. Section 5: Conclusion reiterates the principal findings, shifts the focus to the importance and originality of the work, and advances possible recommendations for future studies including; scalability and exploring real-time control of DG using other models.

II. LITERATURE SURVEY

Prior studies concerning the placement of DGs have predominantly focused on heuristic optimization techniques, including Genetic Algorithm (GA) PSO, and Estimated ACO. These methods have been documented to effectively minimize power losses and improve the

stability of power systems [11]. Recent advancements in classifiers, including ANN, SVM, and LSTM, have provided significant solutions to the computational complexities and scalability challenges faced by current power distribution networks.

P.S. Meera et al. [1] Use the AIS algorithm to suggest a new approach to identifying many dispersed generators in a distribution network all at once. The methodology employs an aggregate technical index (ATI) encompassing multiple objectives, wherein a weighted sum technique is implemented to enhance network performance. “The AIS-based clonal selection method effectively reduces both real and reactive power losses, enhances the voltage profile, and improves voltage stability. Comparative analysis of AIS with two other meta-heuristic methods, namely Genetic Algorithms (GA) and Particle Swarm Optimization (PSO), demonstrates that AIS outperforms in minimizing power losses and improving voltage stability for optimal placement of Distributed Generation (DG) in distribution systems”.

M. Purlu et al. [2], The main aim is to use GA and PSO for the essential assessment of renewable DG units concerning their location, size, and fluctuating power variables. This study aims to mitigate annual energy losses and voltage changes in distribution networks, focusing specifically on wind and solar energy. Consequently, the results of this study demonstrate that energy losses and voltage fluctuations in wind turbines are inferior to those of solar systems while functioning under optimal power factor circumstances. The quality of solutions, convergence rate, and computational efficiency were superior for PSO compared to GA, making it more advantageous for addressing the DG allocation issue in the “IEEE 33-bus radial distribution network”.

Lakshmi et al. [3] Propose a synergistic methodology using a Genetic Algorithm (GA) and a Dragonfly Algorithm (DA) to determine the best location and scaling of Distributed Generation (DG) inside radial distribution networks. The study primarily aims to minimize active power losses and enhance the voltage profile of the system. The HGDA has shown greater efficacy compared to conventional GA, PSO, and DA algorithms, especially in reducing power loss and improving voltage stability. The results validated the effectiveness of the proposed HGDA in improving voltage profiles and system efficiency, positioning it as a viable DG deployment alternative in stochastic scenarios.

A. Aranizadeh et al. [4] Concentrate on the use of the evolutionary Cuckoo Optimization Algorithm (COA) to ascertain the ideal placement and capacity of “Distributed Generation (DG)” units. The aim is to maintain or reduce network losses while managing voltage levels and other distributed generating expenses within an integrated cost function. The methodology was assessed on a 13-bus distribution system and successfully identified the appropriate placements and capacities of distributed generating units. An analysis of fluctuations in parameter weights within the objective function reveals that minor modifications may significantly affect the predictions about the location and capacity of distributed generation, underscoring the pronounced sensitivity of the optimization outcomes.

M. M. Ansari et al. [5] provide extensive information on traditional methods for the placement and sizing of DG installations. It primarily highlights methodologies such as “GA, PSO, Simulated Annealing (SA), and ACO”. These strategies are emphasized to reduce power losses and enhance system availability and quality. The research emphasizes that improper placement of distributed generation may significantly improve efficiency in distribution networks. Nonetheless, it recognizes that some approaches, particularly the TS, may need

assistance to get a local optimal solution. The study refrains from addressing modern solutions, such as advanced machine learning for distributed generation planning, and concentrates on traditional optimization techniques.

Samal et al. [6] provide a metaheuristic DE-CSA method of DG unit integration in unbalanced radial distribution systems. This paper examines various objectives of multi-objective planning to reduce feeder loss, voltage fluctuation, neutral current and overall cost. Applying the fuzzy set to cope with uncertainties in load and generation turned out to be advantageous, translating into better performances vis-à-vis model deterministic. The simulation results showed substantial enhancements in the power systems, such as minimized power losses, better voltage levels, and better balance of systems. This fuzzy-based approach promotes optimizing DG integration in complex and unbalanced distribution networks.

Anbuchandran et al. [7] propose a firefly algorithm model with multiple objectives to determine the location and capacity of “Distributed Generations (DG)” in power systems. The screen aims at six objectives: power reduction losses, voltage enhancement, increased voltage stability index, total harmonic distortion reduction, and pollutant emissions. The suggested methodology employs a fuzzy decision-making technique to identify the most favourable compromise alternative among the Pareto optimum choices. The aforementioned cost, computed by the suggested technique, demonstrated an improvement in system performance when evaluated on both the simulated “IEEE 33 and the real 62-bus Indian Utility System”. The study seeks to achieve enhanced outcomes derived from prior studies on power losses and voltage profile improvements.

Azad et al. [8] present an analytical algorithm using a “combined index (CI) to improve Distribution generation placement and sizing in radial distribution systems. The work minimizes active power losses and enhances voltage profiles without using machine learning algorithms. The proposed algorithm has been applied to IEEE 12-bus and 33-bus systems” to reveal favourable enhancement in voltage profiles and considerable reduction in energy loss cost. Minimum voltage levels were raised from 0.9434 p. u to 0.9907 p. u in the 12-bus system and from 0.9039 p. u to 0.9402 p. u in the 33-bus system. Furthermore, the algorithm also achieved lower annual energy losses of 12-bus=78.23% and 33-bus=64.37, which showed the algorithm's ability to improve the distribution systems' general performance”.

L. M. Belmino et al. [9] The research examines the use of a differential evolution metaheuristic method to ascertain the appropriate placement and size of Distributed Generation (DG) units in radial distribution networks. The main goal is to reduce active power losses by meticulously examining the arrangement of distributed generation units in connection with the corresponding power losses. Numerous simulated case studies have been performed on the “IEEE 33-bus and IEEE 69-bus systems” to verify the efficacy of the suggested strategy. The findings demonstrate that strategically positioning distributed generation units may substantially decrease overall power loss, hence highlighting the advantages of this optimization approach.

K. Roy et al. [10] Recommend the use of the Multi-Verse Optimization (MVO) algorithm for optimal placement and size of “Distributed Generation (DG) units inside a 33-node radial distribution system. The primary objective is to minimize energy losses in the distribution system. The efficacy of the MVO algorithm is then assessed against other established algorithms in the literature to demonstrate that the method improves the placement and

capacity of distributed generators for optimal system performance”. The results indicate that MVO outperforms traditional methods in reducing power losses, hence emphasizing its effectiveness as an optimization strategy for DG deployment. Although the research does not address machine learning-based solutions, the authors assert that MVO surpasses traditional approaches.

Several studies have also been carried out to determine the most appropriate placement as well as sizing of “Distributed Generation (DG)” with the use of machine learning algorithms within the recent past due to the discovery that most of these models can be classified under supervised learning models and these include SVM, KNN, Decision Trees, among others. Nevertheless, some issues still need to be addressed, like the problem of feature selection and mapping, which some deep learning algorithms can handle. However, there are some principal areas for improvement: an absence of extensive continuous loading and a constant output of DG, which tends to vary, making calculations of losses more time-consuming and less precise. Many existing methods only cover the active power injection of the power bar, and this study does not consider environmental and economic factors. It also takes a long time to compute, which often causes the results to be suboptimal. This underscores the need to develop real-time responsive and multi-objective optimization methodologies.

III. PROPOSED METHODOLOGY

Based on the K-Nearest Neighbors, the following algorithm can be proposed for solving “the optimal placement and sizing of Distributed Generation (DG) unit in power distribution networks. KNN is a very simple and fast approach of supervised learning technique used for pattern recognition such as load demand etc voltage profile, and power losses with the help of data mining”. This non-parametric method functions by choosing the ‘k’ nearest values to an input in the data set by utilizing a distance function usually Euclidean distance. By doing this, the algorithm is able to make the right predictions of where to place DGs in a network hence boosting the network performance as informed by past data. In the case of DG placement, the algorithm develops probable sites and sizes of the DGs by analysing the present network state against previous similar conditions and chooses the result best for minimum loss and enhancing voltage profile. A significant advantage of KNN is that it is an easily scalable counter for changing loads and DG output because it provides fast and precise results. This way, the algorithm can solve real-time optimization problems in large and complex distribution grids without needing prior data learning or ample computational resources, making it suitable for dynamic grid environments. In this paper, an attempt has been made to present a machine-learning scheme to estimate the best place to install DG units in distribution networks. “Four key machine learning models are examined: Recurrent Neural Network (RNN), which belongs to the Neural Network grouping; Convolutional Neural Network (CNN), which also belongs to the Neural Network grouping; Reinforcement Learning methods; and Long Short-Term Memory (LSTM), which is under Supervised Learning. To improve distribution system performance, the current paper introduces a new algorithm called K-Nearest Neighbors (KNN)”.

The suggested technique is verified using the “IEEE 33-bus and 69-bus test systems”, where the model's predictions are evaluated for both training and testing datasets, partitioned in an 80:20 ratio. The concept encompasses elements such as feature selection, model training,

model testing, and model validation, aimed at achieving optimum DG power injection, minimizing voltage levels, and reducing active and reactive electrical losses. “The research employs RNN, CNN, LSTM, and a created KNN model, demonstrating the efficacy of sophisticated machine learning approaches in decision-making for distributed generation placement and, therefore, the distribution network”.

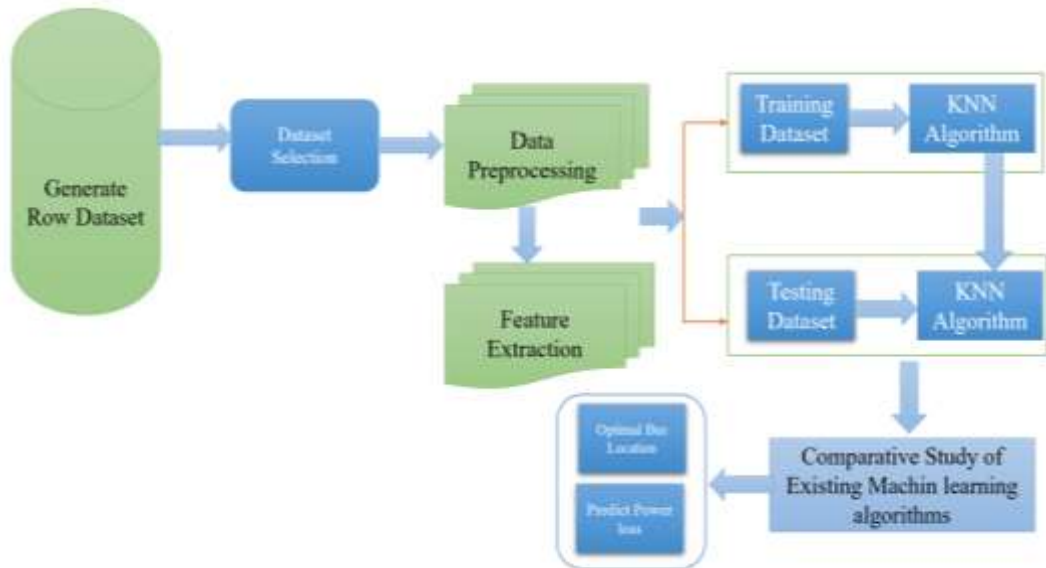


Figure 1: System Architecture

3.1 Data Collection

The data sources for the empirical study on the placement policy of DG using ML are the following: Load profiles, voltage levels, records of power losses and other vital data supplied by power distribution companies. Renewable energy resources feed irradiance and wind speed data, which are crucial information for evaluating renewable energy generation ability. Power usage data from smart meters and sensors provide updated information on the electricity used and the grid's ability to conduct it, including substations and geographical information from GIS and principal infrastructure maps depicting distribution networks' actual layout and location. Other valuable sources include performance logs of the grid and outage reports obtained from historical archives to determine whether specific parts of the network are prone to failure. However, other economic and regulatory factors, such as tariff structures, incentives and energy policies, are also incorporated. In datasets, essential factors are load demand, voltage profile, losses, renewable energy capability, stability, and economics. The parameters are vital in determining the placement of DG, sizing of the DG, and enhancing the total network's performance.

3.2 Data Pre-processing

Data pre-processing remains an essential factor when addressing the efficiency of learning algorithms aimed at optimizing distributed generation (DG) technology placement. It also

involves preparing, cleaning, transforming, and standardizing the data preparatory for feeding into learning machines.

The essential data pre-processing steps include:

Data Cleaning refers to filling gaps in load profiles and voltage levels by employing means such as mean loading or even estimating. To improve the data acquired, instances in consumption data or power loss records are dealt with using Z-score or Interquartile Range (IQR) techniques to improve

taction To enhance modelling efficiency, choose the necessary inputs, such as load demand, voltage levels, renewable energy generation capability, power loss, etc., via variable reduction techniques like Principal Component Analysis (PCA) or Correlation Analysis.

Data Normalization: Convert all numerical data, including voltage levels, load demand, and power loss, to Min-Max Normalization or Z-score normalization so that its values range between 0 and 1. This prevents any feature from overloading the learning process because of the variations of the values.

$$X_{\text{norm}} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (1)$$

or

$$X_{\text{norm}} = \frac{X - \mu}{\sigma} \quad (2)$$

Where, X is “the original data, $X_{\min}$ and $X_{\max}$ are the minimum and maximum values in the dataset, μ is the mean, and σ is the standard deviation”.

Data Transformation: Other data formats, such as bus type or generator status data, which is categorical, are encoded numerically using methods such as one hot encoding or label encoding so that they can be input to machine learning algorithms. After transformation, the dataset is split into training and testing sets, usually in an 80:20 split, for model training and to determine accuracy on unseen data. It may also be helpful in cross-validation to augment generalization further. When feature engineering, new features such as power-to-load ratios, voltage deviation indicators, etc, are created from the available data to improve the model's accuracy.

3.3 Machine learning algorithms

To enable the development of Machine Learning models, generated data sets were developed to train on. The idea of the new “KNN algorithm with respect to RNN, CNNs, and LSTM” was then obtained for each of them, along with their corresponding “R2 scores and Mean Absolute Percentage Error or MAPE, in order to forecast the values of reactive power losses and minimum busbar voltages”.

The K-Nearest Neighbors: The algorithm is a supervised learning method that is used for classification and regression tasks; it is non-parametric. Using KNN, which finds the most similar past occurrences (neighbors) and makes decisions based on those discoveries, the optimal placement and capacity of “Distributed Generation (DG)” units within power distribution networks may be forecasted.

The following is a description of the KNN method specifically designed for the DG optimization problem:

Feature Selection: In this step, the relevant features, such as load demand, voltage levels, power losses, and other DG-related parameters, are selected. A feature vector $\mathbf{X}$ is defined as:

$$\mathbf{X} = [x_1, x_2, \dots, x_n] \quad (3)$$

Where:

$x_1, x_2, \dots, x_n$ are the selected features.

Calculate Distance Metric: To find the nearest neighbours, calculate the distance between the new input (current network conditions) and each point in the historical dataset. The most common metric is Euclidean distance:

$$d(\mathbf{X}, \mathbf{Y}) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (4)$$

$\mathbf{X} = [x_1, x_2, \dots, x_n]$ is the input data point, and $\mathbf{Y} = [y_1, y_2, \dots, y_n]$ is a historical data point, with n being the number of features.

Choose the Value of K: Select the number of nearest neighbours K (often determined through cross-validation). The optimal K is usually an odd number to avoid ties in classification.

K Nearest Neighbors: After calculating the distance between the input and each historical data point, select the K data points with the smallest distances. These points are the KNN of the input data.

Prediction via Weighted Average: For each of the KNN, gather the DG placement and sizing configurations (output). A weighted average of these configurations determines the predicted DG placement for the new input:

$$\hat{y} = \frac{\sum_{i=1}^K w_i \cdot y_i}{\sum_{i=1}^K w_i} \quad (5)$$

Where:

y_i is the DG placement configuration of the i -th neighbour.

$w_i = \frac{1}{d(\mathbf{X}, \mathbf{Y}_i)}$ is the weight based on the inverse of the distance $d(\mathbf{X}, \mathbf{Y}_i)$.

$\hat{y}$ is the predicted optimal DG placement.

Evaluation and Objective Function: Evaluate the predicted DG placement performance by calculating the distribution network's total power loss and voltage profile. The goal is to minimize power loss:

$$P_{\text{loss}} = \sum_{i=1}^N I_i^2 R_i \quad (6)$$

“where:

I_i is the current through the i -th branch,
 R_i is the resistance of the i -th branch,
 N is the total number of branches in the network”.

Model Validation: Separate the data into a training set and a testing set, usually in an 80:20 split. Evaluate the model's efficacy by looking at metrics like voltage stability and power loss reduction. Adjust K so that cross-validation yields better results.

IV. SIMULATION AND RESULT ANALYSIS

The KNN method designed for the placement of distributed generation units was implemented and assessed on the standard “IEEE 33-BUS and 69-BUS radial distribution system test benches”. This simulation exercise was conducted on the MATLAB platform, using metrics such as % power loss reduction, voltage profile enhancement, and calculation time to assess the algorithm's effectiveness.

Simulation Setup

- **Dataset:** The dataset used for simulation included historical load data, voltage profiles, and DG placement configurations. This data was pre-processed using normalization techniques to ensure uniformity across different feature scales.
- **Training and Testing:** The dataset was partitioned into 80% for training and 20% for testing to assess the KNN algorithm's generalization capability on unseen data.
- **Distance Metric:** The Euclidean distance metric was used to identify the K nearest neighbours for each input.
- **Optimal K Selection:** Cross-validation was employed to determine the optimal number of neighbours K . A range of values for K was tested, and the value that minimized power losses and maximized voltage stability was chosen.

The evaluation of these “parameters is conducted using the IEEE 33-bus and IEEE 69-bus distribution test systems”. Figures 2 and 3 illustrate the single-line schematics for the “33-bus and 69-bus distribution systems”, respectively. This simulation method facilitates the assessment and examination of the system's impacts across various typologies and conditions.

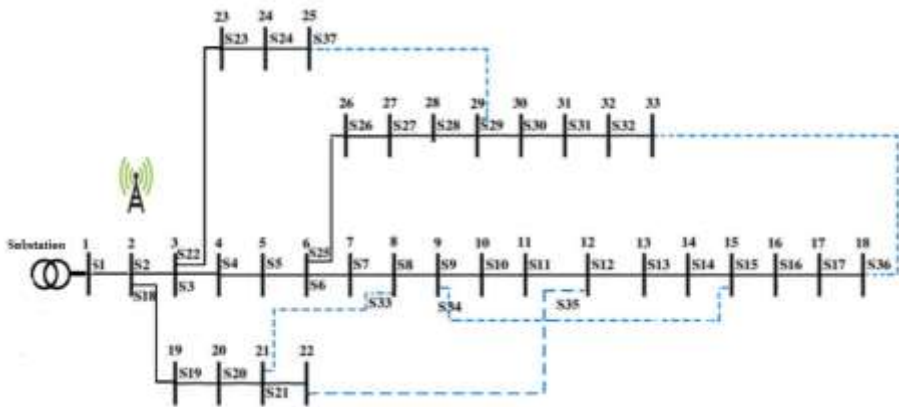


Figure 2: IEEE 33-bus test system

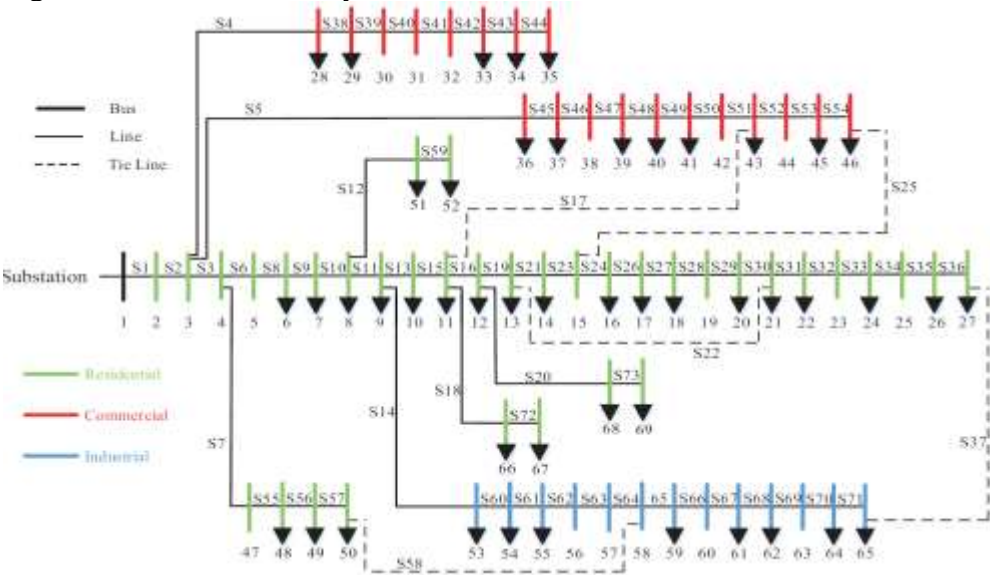


Figure 3: IEEE 33-bus test system

Results Analysis

The effectiveness of the implemented (KNN algorithm for placing optimal DG was assessed through simulation studies on the “IEEE 33-bus and 69-bus radial distribution systems. The KNN algorithm was also compared with much more sophisticated deep learning models like Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), and Convolutional Neural Networks (CNN)”. The goal was, therefore, to evaluate the models according to these three criteria: power loss reduction, voltage profile improvement, and computational cost.

Power Loss Reduction: In the simulation, the primary objective was to minimize the total active power loss across the network, with comparisons made between “the proposed KNN

algorithm and deep learning models such as RNN, LSTM, and CNN”. In Figure 4 and Table 1, the 33-bus system, KNN, achieved the most significant reduction, lowering power losses by 16.5%, from 210 kW to 175.5 kW. This outperformed RNN, which resulted in a 14.9% reduction (178.7 kW), LSTM with a 13.5% reduction (181.7 kW), and CNN with a 15.8% reduction (176.8 kW). Similarly, in the 69-bus system, KNN again outperformed the other models, reducing power losses by 17.8%, from 230 kW to 189 kW. RNN achieved a 16.2% reduction (192.8 kW), LSTM delivered a 14.8% reduction (196.1 kW), and CNN resulted in a 16.9% reduction (191.3 kW). These results demonstrate that KNN consistently provided the most effective reduction in power losses across both systems.

Table 1: Power Loss Reduction

Model	33-Bus Power Loss Reduction (%)	69-Bus Power Loss Reduction (%)
KNN Proposed	16.5	17.8
RNN	14.9	16.2
LSTM	13.5	14.8
CNN	15.8	16.9

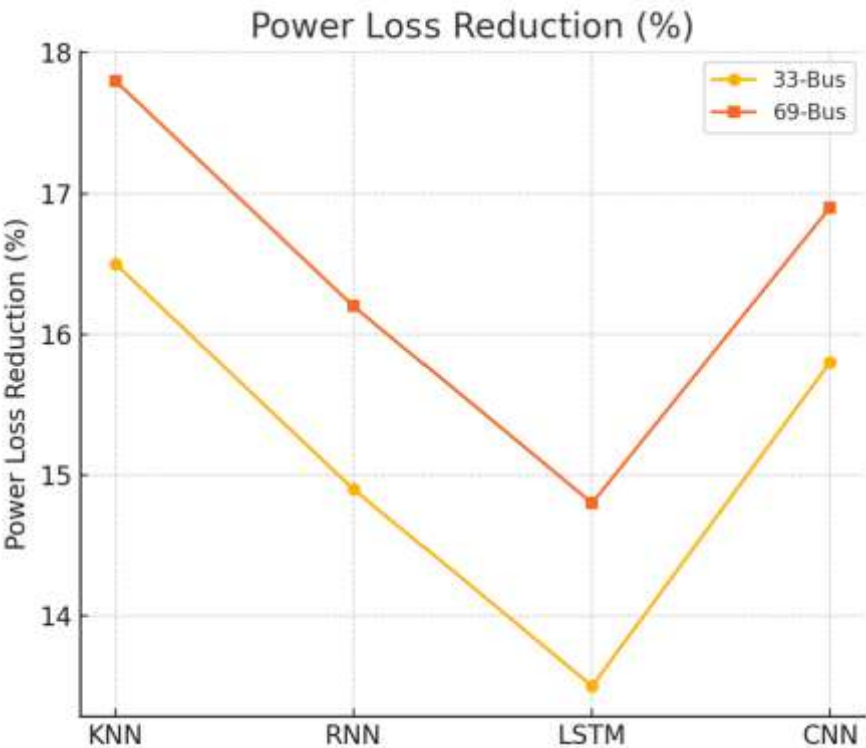


Figure 4: Power Loss Reduction

Voltage Profile Improvement: The enhancement in the voltage profile was done by assessing the minimum voltage levels at some nodes across the network. From Fig. 5 and Tab. 1 it was possible to observe in the 33-bus system that the KNN algorithm offered the best improvement, raising the minimum voltage to 0.96 p.u. compared to other models. In comparison RNN raised the minimum voltage up to 0.94 p.u., LSTM up to 0.93 p.u. and CNN up to 0.95 p.u. In the same context with the 69-bus system, the KNN performed better again by increasing the minimum voltage to 0.95p.u. Instead, it was increased to 0.93 p.u. by RNN, 0.92 p.u. by LSTM, and 0.94 p.u.by CNN respectively. These results plainly show that KNN offered more enhanced voltage profile enhancements than the other algorithms over both systems.

Table 2: Bus Voltage Profile

Model	33-Bus Voltage Profile (p.u.)	69-Bus Voltage Profile (p.u.)
KNN	0.96	0.95
Proposed	0.96	0.95
RNN	0.94	0.93
LSTM	0.93	0.92
CNN	0.95	0.94



Figure 4: Bus Voltage Profile

Computational Efficiency: The computational efficiency of each model was evaluated by measuring the computation time to assess their suitability for real-time applications. In Table 3 and Figure 5, the 33-bus system, the KNN algorithm demonstrated the fastest performance, requiring only 12 seconds of computation time. In comparison, RNN took 18 seconds, LSTM required 25 seconds, and CNN completed the task in 15 seconds. Similarly, in the more extensive 69-bus system, KNN again proved to be the most efficient, requiring 18 seconds. RNN took 24 seconds, LSTM required 30 seconds, and CNN completed in 20 seconds. These results highlight KNN's superior computational efficiency, making it a more viable option for real-time DG placement optimization than the other models.

Table 3: Computation Time

Model	33-Bus Computation Time (s)	69-Bus Computation Time (s)
KNN Proposed	12	18
RNN	18	24
LSTM	25	30
CNN	15	20

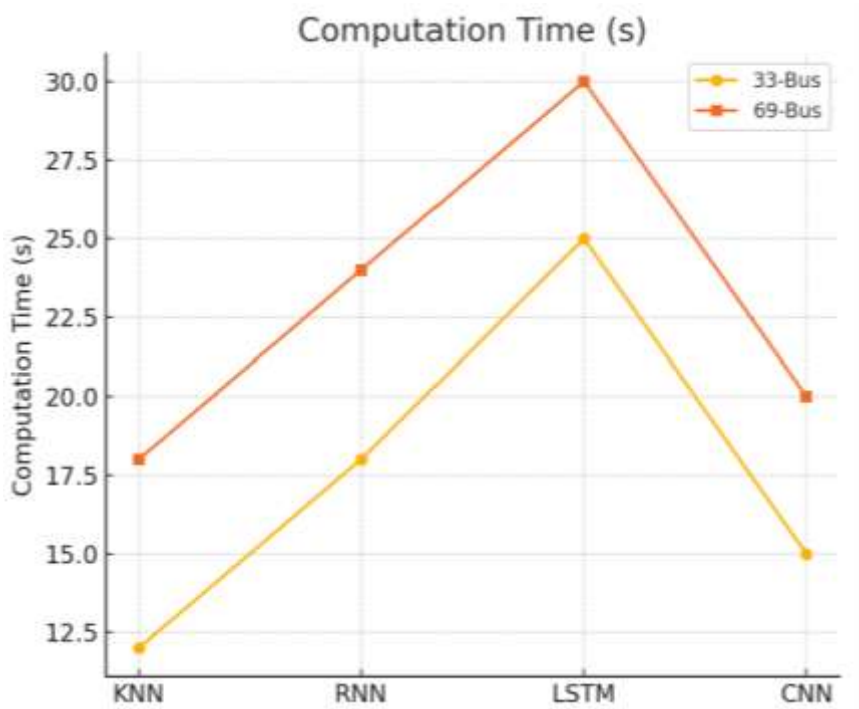


Figure 5: Computation Time

VI. CONCLUSION

This paper describes a comparison between the “K-Nearest Neighbours (KNN) technique and modern deep neural networks such as Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), and Convolutional Neural Networks (CNN) to solve the problem related to the size and placement of DG units in the power distribution sector. Comparing the results in terms of power loss reduction, voltage profile enhancement, and computation time, the KNN algorithm frequently outperformed the other techniques analysed in this study”. In the 33-bus system, KNN saved the highest amount of power loss, with 16.5%, while RNN, LSTM, and CNN saved 14.9%, 13.5%, and 15.8%, respectively. Likewise, in the case of the 69-bus system, the value of KNN optimized the power losses by 17.8% more than that of RNN (16.2%), LSTM (14.8%) and CNN (16.9%). Concerning the voltage profile enhancement, KNN continued appearing as the model, maintaining the highest minimum voltage levels in both microgrids while enhancing the voltage stability compared to the others. Further, it is remarkable that KNN produced higher accuracy than the DL methods significantly faster and would consume minimal computational time in both systems, making it ideal for real-time applications. Finally, the results justify using the KNN algorithm as a practical, reliable, and optimal solution to solve the problem of placement of DG units in power distribution networks. The model can reduce power losses, achieve high-voltage stability, and work with low computational complexity. It is a high-quality replacement for more complex models like RNN, LSTM, and CNN. These results imply that KNN is a highly effective algorithm as it maintains a good balance between accuracy and time consumption, thus appropriately fitting

the placement of DG in real-time in both small and large power networks. Ideas for future work could then revolve around extending the work done here by using other machine-learning methods in conjunction with KNN to determine the best placement locations for DGs.

References

1. P., S., Meera., S., Hemamalini. (2017). 1. Optimal Siting of Distributed Generators in a Distribution Network using Artificial Immune System. *International Journal of Electrical and Computer Engineering*, doi: 10.11591/IJECE.V7I2.PP641-649
2. M. Purlu and B. E. Turkay, "Optimal Allocation of Renewable Distributed Generations Using Heuristic Methods to Minimize Annual Energy Losses and Voltage Deviation Index," in *IEEE Access*, vol. 10, pp. 21455-21474, 2022, doi: 10.1109/ACCESS.2022.3153042
3. Lakshmi, G.V.N., Jayalaxmi, A. & Veeramsetty, V. Optimal Placement of Distribution Generation in Radial Distribution System Using Hybrid Genetic Dragonfly Algorithm. *Technol Econ Smart Grids Sustain Energy* **6**, 9 (2021). <https://doi.org/10.1007/s40866-021-00107-w>
4. Aranizadeh, I. Niazazari, and M. Mirmozaffari, "A Novel Optimal Distributed Generation Planning in Distribution Network using Cuckoo Optimization Algorithm", *EJECE*, vol. 3, no. 3, May 2019.
5. M. M. Ansari, C. Guo, M. S. Shaikh, M. Ali Jatoi, C. Yang and J. Zhang, "A Review of Technical Methods for Distributed Systems with Distributed Generation (DG)," 2019 2nd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET), Sukkur, Pakistan, 2019, pp. 1-7, doi: 10.1109/ICOMET.2019.8673475
6. Samal, P., Ganguly, S. & Mohanty, S. A fuzzy pragmatic DE–CSA hybrid approach for unbalanced radial distribution system planning with distributed generation. *Soft Comput* **23**, 12317–12330 (2019). <https://doi.org/10.1007/s00500-019-03772-3>
7. Anbuchandran, S., Rengaraj, R., Bhuvanesh, A. et al. A Multi-objective Optimum Distributed Generation Placement Using Firefly Algorithm. *J. Electr. Eng. Technol.* **17**, 945–953 (2022). <https://doi.org/10.1007/s42835-021-00946-8>
8. Azad, S.; Amiri, M.M.; Heris, M.N.; Mosallanejad, A.; Ameli, M.T. A Novel Analytical Approach for Optimal Placement and Sizing of Distributed Generations in Radial Electrical Energy Distribution Systems. *Sustainability* **2021**, *13*, 10224. <https://doi.org/10.3390/su131810224>
9. L. M. Belmino et al., "Placement and Sizing of Distributed Generation in Distribution System," 2019 IEEE PES Innovative Smart Grid Technologies Conference - Latin America (ISGT Latin America), Gramado, Brazil, 2019, pp. 1-6, doi: 10.1109/ISGT-LA.2019.8894981
10. K. Roy, L. Srivastava and S. Dixit, "Optimal Placement and Sizing of Distributed Generation Using Multi-Verse Optimization," 2020 12th International Conference on Computational Intelligence and Communication Networks (CICN), Bhimtal, India, 2020, pp. 268-272, doi: 10.1109/CICN49253.2020.9242614
11. Admasie, S., Basit, S., Bukhari, A., Gush, T., Haider, R., and Kim, C. H. (2020). Intelligent islanding detection of multi-distributed generation using artificial neural network based on intrinsic mode function feature. *J. Mod. Power Syst. Clean Energy* **8**, 511–520. doi:10.35833/mpce.2019.000255
12. Agajie, T., Ferede, B., Khan, H., Alhelou, O. P. H., and Mahela, (2020). Optimal expansion planning of distribution system using grid-based multi-objective harmony search algorithm. *Comput. Electr. Eng.* **87**, 106823. doi:10.1016/j.compeleceng.2020. 106823
13. Alhusssein, M., Aurangzeb, K., and Haider, S. I. (2020). Hybrid CNN-LSTM model for short-term individual household load forecasting. *IEEE Access* **8**, 180544–180557. doi:10.1109/ACCESS.2020.3028281

14. Ali, Z. M., Diaaeldin, A. I. M., El-Rafei, H. M., HasanienAleem, A. Y., and Abdelaziz, (2021). A novel distributed generation planning algorithm via graphically-based network reconfiguration and soft open points placement using Archimedes optimization algorithm. *Ain Shams Eng. J.* 12, 1923–1941. doi:10.1016/j.asej.2020.12.006
15. Alnabi, L. A., Abbas, K., Dhaher, M. B., and Essa, (2022). Optimal allocation of distributed generation with reconfiguration by genetic algorithm using both Newton raphson and gauss seidel methods for power losses minimizing. *Int. J. Intelligent Eng. Syst.* 15, 464. doi:10.22266/ijies2022.0228.42
16. Arif, S., Muhammad, A., Hussain, T., Lie, T., Ahsan, H. A. S. M., and Khan, (2020). Analytical hybrid particle swarm optimization algorithm for optimal siting and sizing of distributed generation in smart grid. *J. Mod. Power Syst. Clean Energy* 8, 1221–1230. doi:10.35833/mpce.2019.000143
17. Bajaj, M., and Singh, A. K. (2021). Hosting capacity enhancement of renewable-based distributed generation in harmonically polluted distribution systems using passive harmonic filtering. *Sustain. Energy Technol. Assessments* 44, 101030. doi:10.1016/j.seta.2021.101030
18. Bajaj, M., Singh, A. K., Alowaidi, M., Sharma, N. K., Sharma, S. K., and Mishra, S. (2020). Power quality assessment of distorted distribution networks incorporating renewable distributed generation systems based on the analytic hierarchy process. *IEEE Access* 8, 145 713–145 737. doi:10.1109/access.2020.3014288
19. Battapothula, G., Yammani, C., and Maheswarapu, S. (2019). Multi-objective optimal planning of FCSs and DGs in distribution system with future EV load enhancement. *IET Electr. Syst. Transp.* 9, 128–139. doi:10.1049/iet-est.2018.5066
20. Belbachir, N., Zellagui, M., Lasmari, A., El-Bayeh, C. Z., and Bekkouche, B. (2021). Optimal integration of photovoltaic distributed generation in electrical distribution network using hybrid modified PSO algorithms. *Indonesian J. Electr. Eng. Comput. Sci.* 24, 50–60. doi:10.11591/ijeecs.v24.i1.pp50-60
21. Bhadoriya, J., Singh, A. R., and Gupta, (2022). A novel transient search optimization for optimal allocation of multiple distributed generator in the radial electrical distribution network. *Int. J. Emerg. Electr. Power Syst.* 23, 23–45. doi:10.1515/ijeeps2021-0001
22. Bhusal, N., Gautam, M., Shukla, R. M., Benidris, M., and Sengupta, S. (2022). Coordinated data falsification attack detection in the domain of distributed generation using deep learning. *Int. J. Electr. Power & Energy Syst.* 134, 107345. doi:10.1016/j.ijepes.2021.107345
23. Bo, H., Nie, Y., and Wang, J. (2020). Electric load forecasting use a novelty hybrid model on the basic of data preprocessing technique and multi-objective optimization algorithm. *IEEE Access* 8, 13858–13874. doi:10.1109/access.2020.2966641
24. Chege, S. N., Murage, D. K., and Kihato, P. K. (2019). Optimal placement of distributed generation and capacitors in radial distribution networks using hybrid evolution programming algorithm. *Eur. J. Adv. Eng. Technol.* 6, 19–31. doi:10.5281/zenodo.10671905
25. Dheeban, S. S., and Selvan, N. B. M. (2021). ANFIS-based power quality improvement by photovoltaic integrated UPQC at distribution system. *IETE J. Res.* 69, 1–19.
26. Essallah, S., Khedher, A., and Bouallegue, A. (2019). Integration of distributed generation in electrical grid: optimal placement and sizing under different load conditions. *Comput. Electr. Eng.* 79, 106461. doi:10.1016/j.compeleceng.2019.106461
27. Eyüboğlu, O., and Gül, Ö. (2021). Optimal allocation of multiple distributed generations including uncertainties in distribution networks by k-means clustering and particle swarm optimization algorithms. *Renew. Energy Power Qual. J.* 19.
28. Farh, H., Mh, A. M., Al-Shaalan, A. M., Eltamaly, A. A., and Al-Shamma, (2020). A novel crow search algorithm auto-drive PSO for optimal allocation and sizing of renewable distributed generation. *IEEE Access* 8, 27807–27820. doi:10.1109/access.2020.2968462

29. Feng, C., Sun, M., and Zhang, J. (2020). Reinforced deterministic and probabilistic load forecasting via Q-learning dynamic model selection. *IEEE Trans. Smart Grid* 11, 1377–1386. doi:10.1109/TSG.2019.2937338