A Lightweight Sequential Convolutional Neural Network for Smart Grid Stability Analysis

***Hiren K. Mewada*[1,*], *L Syam Sundar*[2], *Bimal Patel*[3], *Miral Desai*[4], *Adil Khan*[1],**

[1]Electrical Engineering Department, Prince Mohammad Bin Fahd University, Al Khobar, Saudi Arabia

[2]Mechanical Engineering Department, Prince Mohammad Bin Fahd University, Al Khobar, Saudi Arabia

[3]K D Patel Department of Information Technology, Chandubhai S Patel Institute of Technology, Faculty of Technology and Engineering, Charotar University of Science and Technology, Changa, Gujarat, India.

[4]V T Patel Department of Electronics and Communication Engineering, Chandubhai S Patel Institute of Technology, Faculty of Technology and Engineering, Charotar University of Science and Technology, Changa, Gujarat, India

*Correspondence: hmewada@pmu.edu.sa

Received 20 June, 2024; Revised 2 July, 2024; Accepted 2 July, 2024

Available online 2 July, 2024 at www.atlas-tjes.org, doi: 10.22545/2024/00259

Abstract: *A Smart grid stability analysis is essential for ensuring modern power systems' reliable and secure operation. This approach helps identify potential instabilities and disturbances that can lead to blackouts or equipment failures. By analyzing the stability of the grid, operators can take proactive measures to maintain a stable and resilient power infrastructure. Monitoring smart grid data from various sources and analyzing how to control the stability of the grid are challenging tasks. A convolutional neural network (CNN) can effectively capture spatial dependencies and patterns from grid data and can help in enabling accurate prediction and classification of stability-related events in a power system. However, developing a CNN that has fewer learnable parameters and provides high accuracy is challenging. This paper presents a sequential CNN architecture to detect the stability of the Decentral Smart Grid Control (DSGC) system. A mathematical model of the 4-node start architecture of a smart grid was presented. Later, 12 parameter-based grid datasets from the UCI repository were used to validate the proposed network. The proposed CNN accepts sequential data to capture temporal dependencies in the data. The sequential process in a single dimension offers fewer learnable parameters, making the network more compact and computationally efficient. The proposed 11-layered CNN has a total of 12.7K learnable parameters. The detailed analysis of the proposed CNN using ambiguity and the t-SNE score suggested that the model can identify discriminative features for classifying data into stable and unstable classes. A comparison analysis of the quantitative parameters revealed that the model performed well, with 98.82%, 98.55%, 98.88%, and 98.77% accuracy, precision, recall, and F1 score, respectively.*

Keywords: Smart grid, Sustainability, Renewable energy, Convolution neural network, Stability.

1 Introduction

A smart grid is a revolutionary technology that promises to revolutionize the management and distribution of electrical electricity. The utilization of cutting-edge sensing, communication, and control systems has made it possible to create a power grid that is more reliable, efficient, and kind to the environment (Bashir et al., 2021). It facilitates bidirectional energy distribution and transportation, encouraging all of its customers to make decisions related to energy (England & Alouani, 2020). A framework of the ecosystem using energy incentives was presented in (Van Zyl-Bulitta, 2019) for the supply of electrical infrastructure systems. The author suggested that Energy co-prosumption (ECP) can facilitate transformative changes and introduce a commons-based co-prosumer model instead of an individual-based prosumer one. Regarding the operation of a smart grid, one of the most essential factors to consider is stability, which can be defined as the capacity of the grid to maintain a balanced and secure operation despite various operating conditions and disturbances.

Due to the intricate integration of renewable energy sources, energy storage systems, distributed generation, and intelligent devices, the usual methods of stability assessment and control employed in traditional power systems must be modified and improved. To address this problem, academics, engineers, and researchers have utilized predictive analytics and machine learning techniques to create models and algorithms that can identify and prevent stability difficulties in advance.

Predicting the stability of a smart grid involves analyzing a vast amount of real-time data, including information on power generation, consumption patterns, weather conditions, and grid topology. By applying data analytics and machine learning algorithms to these data, it becomes possible to identify patterns, correlations, and anomalies that can help anticipate potential stability concerns before they escalate into major disruptions.

Historical grid data can be used to train machine learning algorithms to discern patterns of both stable and unstable grid behavior. These algorithms can subsequently be employed to forecast the stability of a grid under various scenarios and settings. As an illustration, they have the ability to foresee the consequences of varying renewable energy supplies on the stability of the power system or anticipate the outcomes of abrupt shifts in demand for electricity. To ensure grid stability and prevent cascading failures, operators can proactively address stability issues by making precise predictions. This can involve adjusting generation or load patterns, implementing energy storage devices, or activating control mechanisms.

Furthermore, predictive analytics can assist in optimizing grid operations by providing insights into the optimal placement and sizing of renewable energy sources, energy storage systems, and other grid assets. By considering stability predictions alongside economic factors, environmental considerations, and regulatory constraints, decision-makers can make more informed choices to ensure the long-term stability and sustainability of smart grids.

There are many challenges associated with smart grids, such as the hybridization of many cutting-edge technologies and functional characterization of the peer-to-peer microgrid energy market (Gould et al., 2023). In general, there are two fundamental categories: socioeconomic difficulties and technological obstacles. Some of the socioeconomic challenges include stakeholder management, insufficient information, an absence of rules, and large capital expenditures (Kotsiopoulos et al., 2021). Technological challenges include storage limitations, the absence of policies, vulnerabilities in cybersecurity when connecting grids to cyber-physical systems, inadequate grid infrastructure to accommodate future needs for storing intermittent energy production, complex data management from multiple grid components, grid stability issues related to energy-sharing, power oscillation, system complacency, and power reservations. This paper focuses on the primary technical problem of predicting the stability of a synchronous generator, as it directly impacts the dependable transmission of power. A repository for machine learning at the University of California, Irvine (UCI) (Arzamasov et al., 2018) was used to obtain the datasets for the Electricity Grid. The main contributions of this work include the following:

- A comprehensive analysis of the efficacy of various machine learning algorithms in smart grid stability prediction was performed.
- A mathematical model used for a 4-node star network to augment the dataset is presented in the paper.
- There is a need for a CNN architecture that yields high accuracy and has fewer learnable parameters. Therefore, a new lightweight convolutional neural network is proposed to forecast smart grid power stability using 12 different attributes.
- A comparative analysis of the proposed model with those in previous literature using quantitative parameters is presented in the paper.

The overall paper is organized as follows: The second section presents a literature review focusing on the application of machine learning algorithms used for stability analysis. The proposed convolutional neural network is presented in section 3, and a comprehensive analysis of the obtained results and discussion are presented in section 4. Finally, the conclusion and future work are presented in section 5.

2 Literature Study

This Prior research has focused primarily on traditional centralized power systems that experience little frequency variation (Bano et al., 2020). Decentralized power networks that are connected to renewable energy sources experience significant swings across several time scales, such as seasonal, intraday, and short-term fluctuations (Schäfer et al., 2016). Therefore, this study focuses on previous studies based on decentralized smart grid systems. According to Mohsen et al. (Mohsen et al., 2023), evaluating information gathered from a smart grid is time-consuming. Machine learning and deep neural networks are two examples of AI that have transformed the energy production and distribution process. These neural networks are currently viable in applications such as image processing (Desai & Mewada, 2023) (Bharali, 2024) and speech recognition (Mewada et al., 2023) and can be used to anticipate smart grid stabilities. It was also crucial in predicting smart grid stability (Hong et al., 2020). Moosavirad et al. (Moosavirad et al., 2024) studied the influence of sociocultural policies which can reduce electricity consumption emphasizing educational policies.

Yin et al. (Yin et al., 2019) created a KRR-XGBoost model to predict the stability of distributed power systems. This model also offers valuable insights for designing these systems and optimizing costs. The data input components encompassed the grid stability index and stability predictor (i.e., stable or unstable) and the elements that influence the grid's stability. Different data-level resampling approaches have been employed to address the problem of data imbalance. Later, an XGBoost-based machine learning model was used for stability detection. The findings indicated that a dataset with equal proportions of data samples across different classes had superior performance compared to a dataset with unequal proportions in terms of the performance of classifiers. When XGBoost was combined with random oversampling, the accuracy of the forecast increased from 94.7% to 96.8%. A support vector machine was used to detect faults in the power transmission line in (Shingade & Shah, 2023). The model succeeded with 95% accurate detection using current measurement as a feature set.

An extreme learning machine optimized using a memetic algorithm to predict stability was presented in (Mishra et al., 2022). The model was validated using UCI datasets and achieved 99% accuracy. However, they used the original 10,000 observation samples in the network. One study (Abu Al-Haija et al., 2021) examined seven machine learning architectures, which included a decision tree classifier (DTC), support vector machine (SVM), logistic regression (LoR), linear regression (LR), linear discriminant classifier (LDC), k-nearest neighbor (kNN) and naïve Bayes classifier (NBC). All classifiers were tested on the original UCI datasets of the smart grid (Arzamasov et al., 2018), and the authors observed that optimizing the SVM with 30 epochs achieved 99.93% accuracy. A cascade of feature extraction and multiple classifiers was presented in (Önder et al., 2023). They observed that the hybridization of a supervised attribute filter and fuzzy C-means clustering-based feature weighting with a bagged tree classifier performed best among different machine learning algorithms.

Recurrent neural networks (RNNs) have been suggested as a means to reliably forecast smart grid stability. However, RNNs are hindered by the absence of a control mechanism and the challenge of vanishing gradient or explosion problems (Massaoudi et al., 2021). Therefore, the gated RNN was proposed (Massaoudi et al., 2021) for overcoming the vanishing gradient problem. The authors estimated the hyperparameters of the gated RNN using a stimulated annealing optimization algorithm. The validation of the structure on their simulated dataset resulted in a 97.9% prediction rate. Similarly, an LSTM network was proposed in (Alazab et al., 2020) that offers 97% stability analysis. Nevertheless, recurrent neural networks (RNNSs) frequently fail to achieve desirable levels of precision (Mohsen et al., 2023).

Multilayer perceptron networks or feed-forward ANN networks offer efficient methods for addressing the issue of low accuracy. Moldovan and Salomie (Moldovan & Salomie, 2019) normalized the datasets, and three feature selection algorithms were used to extract features from the dataset. Four different classifiers were tested with three sets of features. They concluded that particle swarm optimization-based feature selection with a multilayer perceptron classifier performed well, with 93.8% accuracy. A 19-layered feedforward neural network (FNN) was proposed by Darbandi et al. (Darbandi et al., 2020). The suggested approach utilizes the conjugate gradient technique in conjunction with the Fletcher–Reeves update algorithm to train the hidden and output layers. The network was tested on the IEEE 39-bus system and achieved a maximum accuracy of 97.8% on a single-phase connection.

Feedforward neural networks cannot retain information from previous inputs or outputs. Each input is treated autonomously, without considering the sequence or past data. In addition, its architecture is fixed and consists of a predefined number of layers and neurons. An inflexible framework can pose difficulties in adjusting the network to meet the intricacy of a specific challenge. Therefore, larger and deeper networks are needed for large datasets with greater complexity. In contrast, parameter sharing in CNNs reduces the number of parameters, resulting in decreased computations. Backward learning makes it more robust for sequential data. This study presents a CNN for grid stability prediction. A 4-node star architecture is considered for a decentralized smart grid system. The following sections present the proposed CNN and its experimental results analysis in comparison with past machine learning, FNN and RNNs.

3 Proposed Method

3.1 Mathematical Modelling of Star Architecture Network-based Smart Grids

This study considers a four-node star architecture network-based power grid. To enable bidirectional communication between nodes, the consumer nodes are linked to the central node (generator node). This allows them to operate at lower power levels. The star topology has the benefit of autonomous networks, which is a major advantage. If one consumer node goes down or makes a mistake, it will not impact the others, and the network will continue to function normally. Star and bus topologies were found to be popular in the literature survey (Omar et al., 2022). Figure 1 shows a network with a four-node star topology that was used for dataset collection. The original dataset consists of 10000 samples with 13 attributes. However, deep learning-based networks need large training datasets to learn them better. Therefore, the mathematical model presented in (Omar et al., 2022) was used for data augmentation. The mathematical model of decentralized smart grid control can be formulated by considering the key components and control strategies involved in the decentralized control of a smart grid. The model incorporates the power generation, consumption, and control mechanisms of the grid.

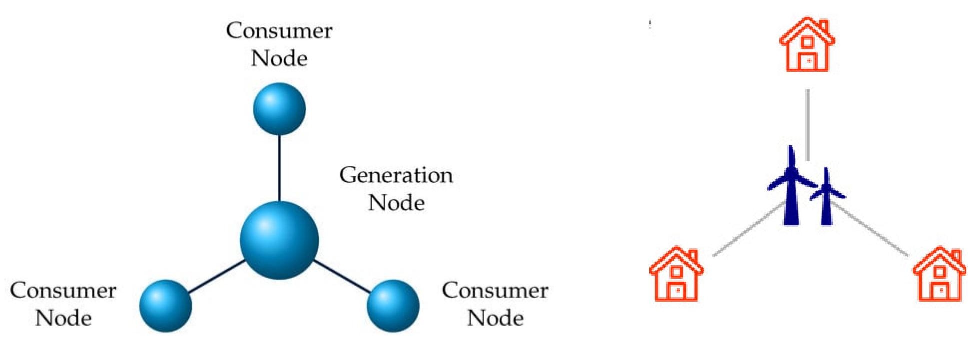


Figure 1. Four-node star architecture of a smart grid

Decentralized smart grid control systems using a star architecture can be modeled using various mathematical approaches. The following are a few types of mathematical models commonly used for such systems:

Linear State-Space Model: This model represents the smart grid control system as a set of linear differential equations. It describes the dynamics of the system by defining the state variables, inputs, outputs, and their interrelationships. The state-space model allows for the analysis and design of control strategies, such as feedback control, using techniques such as linear control theory.

Agent-Based Model: In an agent-based model, the smart grid control system is represented as a collection of autonomous agents, each representing a component or entity in the system. Agents have their behaviors, decision-making capabilities, and interactions with other agents. This modeling approach captures the decentralized nature of the control system and enables the study of emergent system behaviors.

Optimal Power Flow (OPF) Model: OPF models are optimization-based mathematical models used for power system operation and control. They aim to optimize the power flows in the grid while satisfying various constraints, such as generator limits, transmission line capacities, and demand requirements. Decentralized SGCSs can utilize distributed optimization algorithms to solve OPF problems and achieve efficient and reliable operation.

Game Theory Model: Game theory provides a mathematical framework for analyzing decision-making and interactions among multiple players in a system. In a smart grid control system, different entities, such as generators, consumers, and storage devices, can be modeled as players. Game theory models can capture the decentralized decision-making and strategic interactions among these players, leading to the analysis of equilibrium outcomes, such as Nash equilibria.

Stochastic Model: Stochastic models incorporate uncertainties, such as renewable energy generation, demand fluctuations, and market prices, into the smart grid control system. These models utilize probabilistic methods, such as Markov chains, stochastic differential equations, or Monte Carlo simulations, to capture the random nature of system variables. Stochastic models enable the assessment of system performance, reliability, and risk analysis.

To develop a mathematical equation for a smart power grid considering the time required to adjust power generation and consumption, as well as the power generated by nodes and their price elasticity coefficient, we can utilize a dynamic model. We consider the following parameters:

- $P_i(t)$: Power generation or consumption at node i at time t, where i = 1, 2, ..., N.
- T_i: Time constant representing the time required for node i to adjust its power generation or consumption.
- $P_{demand}(t)$: Total power demand in the grid at time t.
- ε_i: Price elasticity coefficient at node i, representing the responsiveness of power generation or consumption to changes in price.
- $P_{price}(t)$: Price of electricity in the grid at time t.
- $\Delta f(t)$: Deviation of the system frequency from the nominal frequency at time t.
- K_f: Frequency control gain.

The power generation or consumption at each node can be modeled as a first-order dynamic system, considering the time required for adjustment:

$$dP_i(t)/dt = (1/T_i) * (P_i(t) - P_{demand}(t)) \quad (1)$$

This equation represents the rate of change in power generation or consumption at node i with respect to time. It is proportional to the difference between the current power output ($P_i(t)$) and the total power demand ($P_{demand}(t)$) with a time constant of T_i. The price elasticity coefficient can be incorporated into the model to reflect the responsiveness of power generation or consumption to changes in price. The power adjustment equation is modified as follows:

$$dP_i(t)/dt = (1/T_i) * (P_i(t) - P_{demand}(t) + \varepsilon_i * (P_{price}(t) - P_{price}(t-1))) \quad (2)$$

In this equation, the additional term $\varepsilon_i * (P_{price}(t) - P_{price}(t-1))$ captures the influence of price elasticity. It represents the change in power generation or consumption at node i due to the difference between the current price ($P_{price}(t)$) and the previous price ($P_{price}(t-1)$) multiplied by the price elasticity coefficient ε_i.

By solving these dynamic equations for each node in the smart power grid, considering the time constants, power demand, and price elasticity coefficients, we can model the power adjustment and response within the grid. Modeling the grid as an oscillator, where the power generator node is considered a synchronous generator and the consumer is a synchronous motor. For each synchronous generator I (producer) or motor i, the swing equation can be written as follows:

$$M_i * d^2(\Delta\delta_i(t))/dt^2 + D_i * d\Delta\delta_i(t)/dt + \kappa_i * \Delta\delta_i(t) = \Delta P_i(t - \tau) - \Delta P_{load}(t - \tau) \quad (3)$$

In this equation, $\Delta\delta_i(t)$ represents the rotor angle deviation of synchronous generator i at time t, M_i is the moment of inertia of generator i, D_i is the damping coefficient, κ_i is the friction coefficient, $\Delta P_i(t)$ represents the power imbalance of generator i (the difference between the generated power and the power demand), and $\Delta P_{load}(t)$ represents the total power imbalance due to consumer loads. The above equation has many solutions. Therefore, to determine the stability, the eigenvalues of the equation are calculated considering linear stability. In the specific instance when the maximum eigenvalue has a negative value, the system is deemed stable. Conversely, if the maximum eigenvalue is positive, the system is regarded as unstable (Franović et al., 2023).

3.2 Proposed Sequential CNN Network

The main objective of the work is to propose a convolutional neural network to predict stability from the observed data. In most of the literature, an SVM and a 4-layer ANN network with varying numbers of neurons in each layer were tested for prediction. Because the grid data are time series data, a 1D CNN network that operates on one-dimensional sequences is proposed in comparison to 2D CNNs used for image processing. Normally, a 1D sequential CNN begins with an input layer that accepts sequential data. The second layer is a convolutional layer. A convolutional layer that applies filters over the input sequence to extract local patterns or features.

$$y_i^p = \sum_{n=1}^{N} \sum_{m=0}^{M-1} k_m^p X_{i+m}^n + b^p \quad (4)$$

where y is the input data, f is the feature value, k is the kernel, p indicates the kernel number and b is the bias function. N represents the number of channels in the data, and M represents the size of the kernel. To prevent the loss of information at the edges of the input, it is common to apply zero-padding to both sides of the input. Hence, the dimensions of the feature map F are determined by the padding width and the stride length. Another layer is the activation layer. A nonlinear activation function introduces nonlinearity and allows the network to learn complex relationships in the data. This nonlinear ReLU function is expressed as

$$f(y) = \mathrm{MAX}(0, y) \quad (5)$$

Later, the pooling layer is used in the CNN. A pooling layer reduces the dimension of feature maps from the convolutional layers. The pooling operation can be expressed as

$$V_i^p = F\left(y_{(i+0)}^n, y_{(i+1)}^n, y_{(i+2)}^n \ldots . y_{(i+M-1)}^n\right) = \frac{1}{|y_i|} \sum_{y\epsilon\, y_i} y \quad (6)$$

In the proposed network, global average pooling is used. As it averages the data, it does not have any learnable parameters. A fully connected layer that makes predictions based on the extracted features and an output layer with a softmax activation function for classification tasks.

$$Softmax_i = \frac{e^{-y_i}}{\sum_{m=0}^{M} e^{-y_i}} \quad (7)$$

During the training phase, the parameters (weights and biases) of the 1D CNN are learned by minimizing a loss function via techniques such as backpropagation and gradient descent. The model is trained to minimize the difference between the predicted output and the true output for a given input sequence. The width of the time series is determined by the values of two variables: the number of features, denoted K; and the length of the series, denoted N. The convolutional filters possess a width equivalent to the width of the time series; however, their lengths may vary. The filters are specifically designed for convolutive operation in a unidirectional manner, starting at the beginning of the time series and moving toward its endpoint.
The overall pseudocode of the proposed network is presented in Algorithm 1.

Algorithm-1 1D sequential CNN pseudocode

Procedure Training with 1DCNN
Input :
Input training dataset power, time and price elasticity
Output:
Updated weights k, and bias b for equation 4 using training dataset
Forward Pass:
Convolve network (equation 4)
Compute activation function (equation 4)
Normalizes all the activations of a single layer
Convolve network (equation 4)
Compute activation function (equation 4)
Normalizes all the activations of a single layer
Apply global average pooling (v) (equation 6)
Calculate the softmax function (equation 7)
Predict the class
Backward Pass:
Compute the gradient $\frac{\delta l}{\delta k_{ij}^{n}} = \frac{\delta l}{\delta V_j^{n}} \times y_i^{n}$
Compute Loss function $L = -\log\left(\frac{\exp(y_i)}{\sum_{i-1}^{N}(y_i)}\right)$
End Procedure

4 Results and Discussions

For network validation, a dataset from the UCI repository (Arzamasov et al., 2018) is used. This dataset is simulated using a 4-node star architecture. It has 10,000 observations with 12 properties. The real-time dataset consists of 10,000 observations, which can be expanded to 60,000 by generating a 3-factorial. The dataset consists of 12 primary classification features and 2 dependent variables. The dataset contains features that provide values of time required to adjust power generation and consumption for grid users, e.g., τ_1, τ_2, τ_3, τ_4, and power generated or consumed by the nodes, e.g., power generated by grid +P1 and consumed by the users -p2, -p3, and -p4 and their price elasticity coefficient, e.g., γ1, γ2, γ3, and γ4. Using these parameters, the grid stability is analyzed and reported using their eigenvalues, i.e., stab, and categorical values, i.e., stable or unstable. The power adjustment time τ ranges from 0.5 to 10. The power P ranges from -2 to -0.5 for consumers and γ ranges between 0.05 and 1.

Figure 2 displays a heatmap illustrating the correlation values. An insignificant correlation between the input values is evident, except for a moderate correlation between $p1$, $p2$, $p3$, and $p4$. This correlation is predicted, as the equation $p1=p2+p3+p4$ was predetermined at the beginning of the simulation.

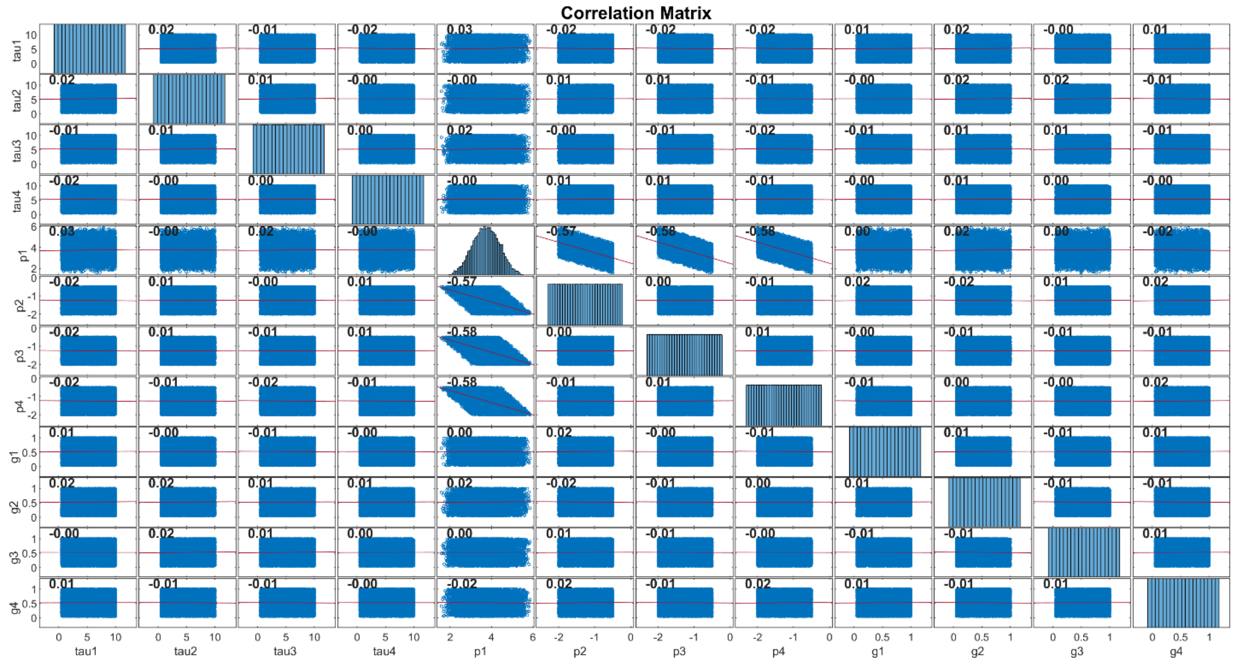


Figure 2 Correlations between the input parameters

The proposed 1D sequential CNN consists of a total of 11 layers. The dataset has 12 features; therefore, the input layer is initialized with 12 channels. The convolutional layers use 32 and 64 kernels to calculate intermediate features. All layer details are presented in Table 1, along with their learnable parameters. The proposed network has a total of 12.7 K learnable parameters.

Table 1 Layer details of the proposed CNN

Layer Name	Activation	Number of Learnable
Sequence Input	12 (C) x 1 (B) x 1 (T)	0
1D Convolution	32 (C) x 1 (B) x 1 (T)	2112
ReLU	32 (C) x 1 (B) x 1 (T)	0
Layer Normalization	32 (C) x 1 (B) x 1 (T)	64
1D Convolution	64 (C) x 1 (B) x 1 (T)	10304
ReLU	64 (C) x 1 (B) x 1 (T)	0
Layer Normalization	64 (C) x 1 (B) x 1 (T)	128
1D Global Avg. Pooliing	64 (C) x 1 (B)	0
Fully Connected Layer	2 (C) x 1 (B)	130
Softmax	2 (C) x 1 (B)	0
Classification layer	2 (C) x 1 (B)	0
Total		12.7 K

A complete flow chart for processing the data for stability classification is shown in Figure 3. Here, the dataset is divided into three parts. The training and validation datasets are used to train the network, and the test dataset is used to evaluate the performance of the network. The dataset comprises 60,000 samples, 80% of which were allocated for training, 10% for validation and 10% for testing.

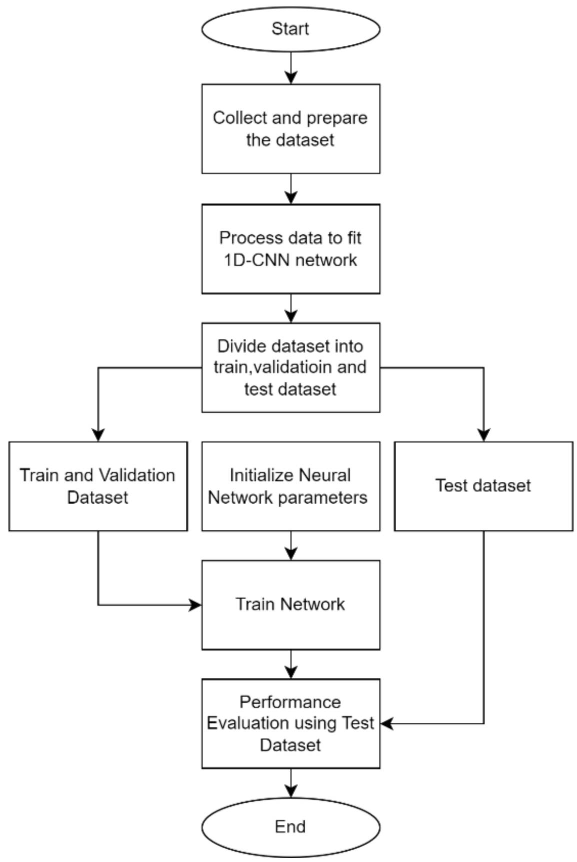


Figure 3 Flowchart of the proposed method

The model was tested on an augmented dataset with 60000 observations. The model is validated using quantitative parameters, i.e., accuracy, precision, recall, and F1 score. Let true positives (TPs) represent the correct prediction of stable cases, true negatives (TNs) represent the correct prediction of unstable cases, false positives (FPs) represent the incorrect prediction of unstable cases, and false negatives (FNs) represent the incorrect prediction of stable cases. Using these parameters, the accuracy, precision, recall, and F1 score are obtained via the following equations.

$$Accuracy(\%) = \frac{TP+TN}{TP+FP+TN+FN} \times 100 \tag{8}$$

$$Precision(\%) = \frac{TP}{TP+FP} \times 100 \tag{9}$$

$$Recall(\%) = \frac{TP}{TP+FN} \times 100 \tag{10}$$

$$F1 - Score(\%) = \frac{2 \times Precision \times Recall}{Precision+Recall} \times 100 \tag{11}$$

Initially, the model was tested using different optimizers. The ‘ADAM’ optimizer provides fast learning, stochastic gradient descent with momentum (SGDM) has better generalization capability, and RMSPROP reduces computational effort in the training phase of the CNN. The network was trained for 15 epochs. Table 2 presents the experimental results obtained using a test dataset for these optimizers. The results showed that the ADAM optimizer performed better than did the SGDM and RMSPROP algorithms. Figure 4 shows the accuracy and loss curve for both the training and validation datasets over the number of epochs. The validation accuracy of the ADAM optimizer was 98.77%. The alignment of accuracy and loss for the training and validation datasets showed that there was no overfitting of the network. Table 2 compares the accuracy of the test dataset using different optimizers.

Table 2 Performance results of the optimizer in training the CNN network

Optimizer	Accuracy	Precision	Recall	F1Score
SGDM	89.57	91.58	88.69	90.11
RMSPROP	90.12	92.25	89.29	90.74
ADAM	98.82	98.55	98.88	98.77

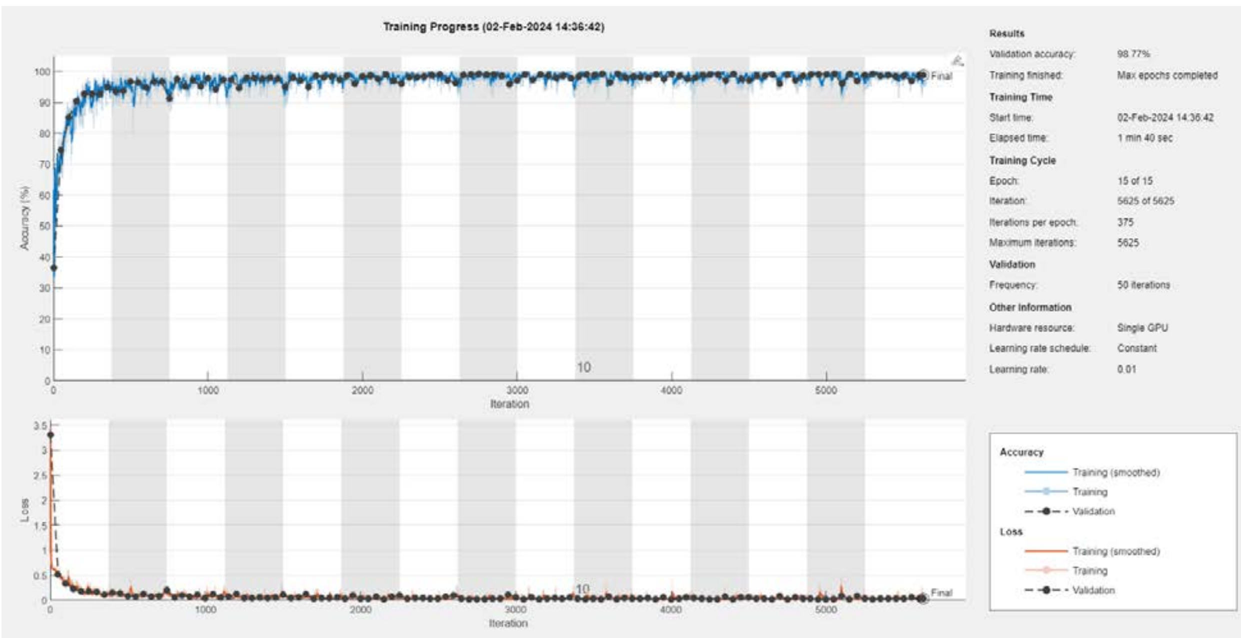
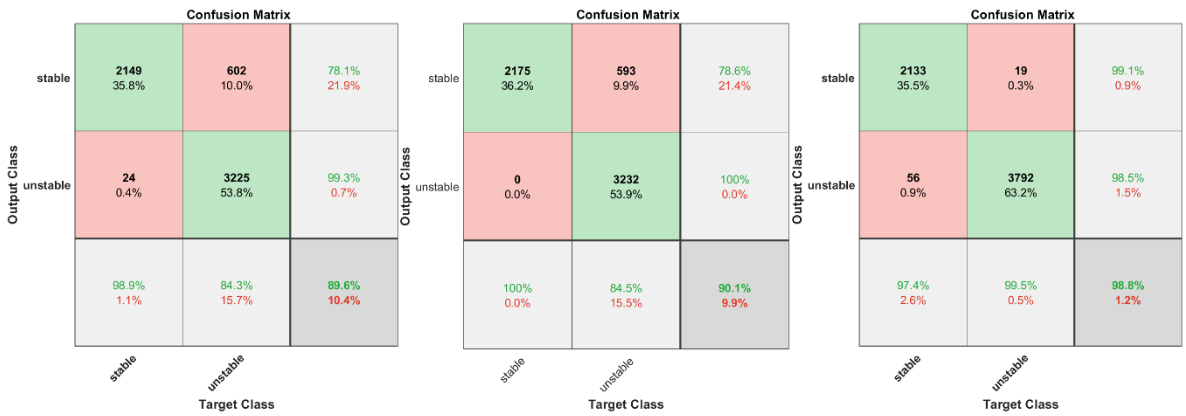


Figure 4 Accuracy and loss of the training and validation datasets over a number of epochs using the ADMA optimizer

Figure 5 shows the confusion matrix on the test dataset obtained using the three optimizers. The testing accuracy of the SGDM and RMSPROP was 90%, whereas the ADAM optimizer outperformed the other algorithms, providing the best accuracy of 98.8%. The first column of Figure 5(c) shows that 2133 stable observations and 3792 unstable observations were classified correctly out of a total of 6000 test datasets. Out of 2152 stable observations, 97.4% of the predictions were correct, and only 0.9% were stable observations misclassified by the network. Overall, only 0.3% of the unstable observations were misclassified as stable by the proposed network.

CNNs often learn high-dimensional feature representations of input data. Applying t-SNE to these feature representations can help visualize how the network has grouped or separated different classes or categories in the data. It can provide insights into the discriminative power of the learned features and help identify any potential issues or biases in the network's representations.

Figure 6 shows the t-SNE plot obtained from the feature set of the softmax layer. There are very few overlaps of the features in the training set, which suggests that the model has succeeded in extracting strong features to aid in better classification.



(a) (b) (c)

Figure 5 Confusion matrix for the test dataset: (a) SGDM optimizer, (b) RMSPROP optimizer, and (c) ADAM optimizer

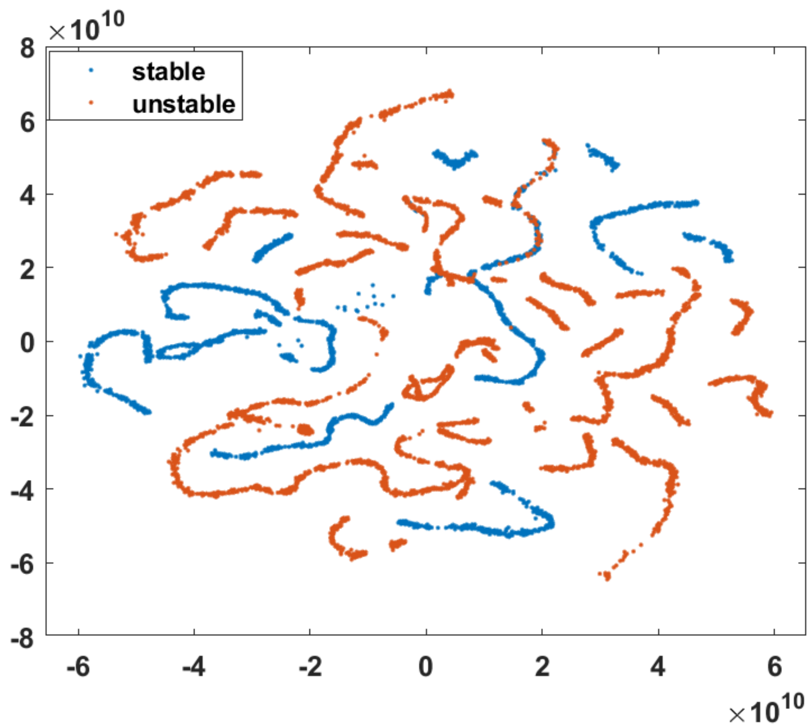


Figure 6 TSNE plots of the trained network at the softmax layer

The receiver operating characteristic (ROC) curve representing the performance of the network is plotted in Figure 7. The ROC curve provides insights into the trade-off between the true positive rate and the false positive rate at various classification thresholds. This approach helps visualize the model's performance across a range of operating points and can assist in selecting the optimal classification threshold based on the specific requirements of the application. A higher AUC-ROC indicates better discrimination between the classes. The proposed network achieved an AUC =1 for both the stable and unstable categories.

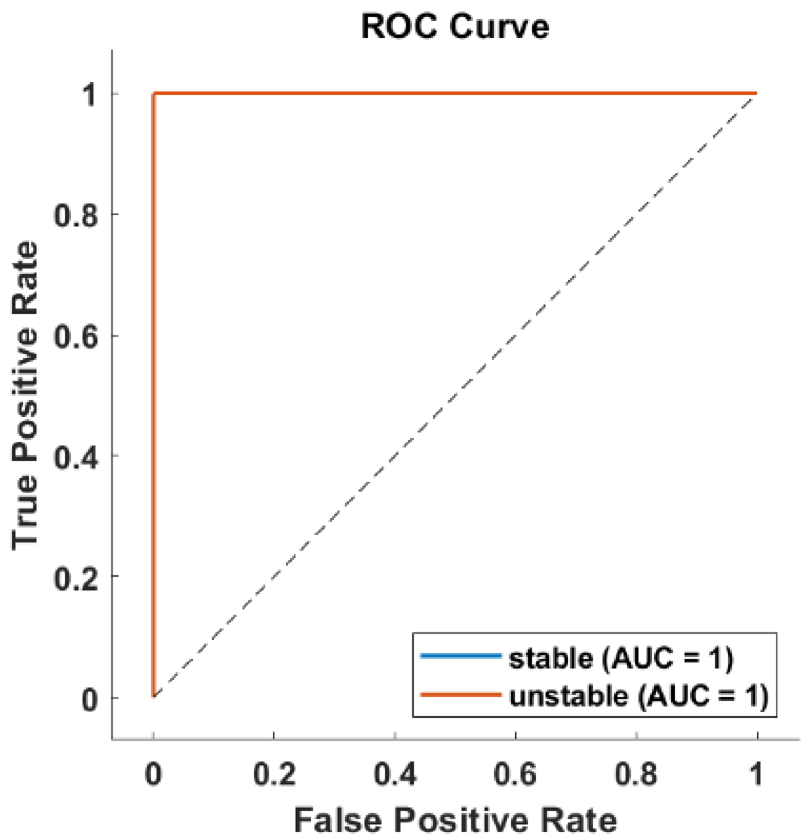


Figure 7 ROC Curve

The proposed network is further analyzed by observing the probability obtained at the softmax layer for classification. The ambiguity of a classification is quantified as the ratio between the second-highest probability and the largest probability. An ambiguity near 1 indicates that the network has a high level of uncertainty regarding the classification of a certain class. The presence of two classes with highly comparable observations may result in uncertainty since the network is unable to discern the distinctions between them. Table 3 lists the top 10 observation IDs for which the network showed similarities between the stable and unstable classes. Table 3 shows that observation number 5180 from the test dataset represents a stable grid. However, the network perceives this observation as unstable.

Table 3 Ambiguity among the classes at the softmax layer

Observation ID	Ambiguity	Likeliest	Second	True Class
5180	0.998268	**'unstable'**	'stable'	**'stable'**
2519	0.997432	**'stable'**	'unstable'	**'unstable'**
1222	0.995666	**'unstable'**	'stable'	**'stable'**
294	0.993712	**'unstable'**	'stable'	**'stable'**
3889	0.9937	'stable'	'unstable'	'stable'
1243	0.988337	'unstable'	'stable'	'unstable'
5209	0.970164	'stable'	'unstable'	'stable'
5358	0.956865	**'unstable'**	'stable'	**'stable'**
460	0.95252	**'unstable'**	'stable'	**'stable'**
5954	0.935641	**'unstable'**	'stable'	**'stable'**

Grid stability detection is a highly important subject that has been extensively studied in this field. Finally, a comparison of the results obtained in this study with those in the literature is presented in Table 4. Paper (Chen et al., 2019) introduces an XGBoost method for predicting transient stability in power systems. They also modeled the system considering the generator and consumer as a dynamic generator. The authors removed redundant features using correlation filtering, and the XGBoost model was proposed to predict patient status. The XGBoost model achieved 97.82% classification accuracy. The authors did not present all the parameters. However, their model requires hyperparameter optimization. In (Bashir et al., 2021), the authors tested various machine learning algorithms, including naïve Bayes, KNN, logistic regression, SVM and neural network algorithms. They observed that the neural network performed best, with 98% training accuracy and a 97.80% F1 score. However, the details of ANNs are not presented well. In addition, the error rate of their ANN is 4%, whereas that of the proposed network is 1.2%. Dhingra and Tomar (Dhingra et al., 2022) also performed a comparative analysis of various machine learning algorithms, including the extra tree classifier, CatBoost classifier, random forest classifier, light gradient boosting machine, gradient boosting classifier, extreme gradient boosting, K neighbors classifier, logistic regression, decision tree classifier and naïve Bayes. All the CatBoost classifiers performed well, with 95.06% stability detected in the dataset.

Breviglieri et al. (Breviglieri et al., 2021) proposed a deep CNN to predict the stability of a grid. Initially, they presented a five-layer architecture, and using a trial-and-error method, they tried to find the best combination of neurons in each layer. In addition, they tested different optimizers in their network. The results showed that the accuracy of the ADAM optimizer was limited to 95.35 for 20 epochs. Another model with 288-288-24-12-1 neurons trained with 50 epochs using the NADAM performed well, with 98.82%, 98.55%, 98.88%, and 98.77% accuracy, precision, recall and F1 score, respectively. In (Gauli et al., 2023), the authors also used a similar four-layer structure with a 24-24-12-1 architecture. They trained their network for 50 epochs to obtain a better classification. However, a detailed analysis is not presented in the paper. A multilayer perceptron-based feed-forward ANN was used in (Alsirhani et al., 2023). Initially, the feature dimensions were reduced via principal component analysis, and the MLP network trained for more than 20 epochs was subsequently used for classification. The accuracy of these methods was limited to 95.35%. Mohsen et al. (Mohsen et al., 2023) developed an ANN network similar to that in (Gauli et al., 2023) with a five-layered ANN composed of 288-288-24-12-1. Their network has 100K learnable parameters, and the model accuracy was 97.82%. The proposed model has an 11-layer architecture and only 12.7 K learnable parameters. The training times were 1 minute and 40 sec, in contrast to 50 minutes (Breviglieri et al., 2021). Thus, in comparison, the model achieved similar or better classification accuracy.

Table 4 Comparison of the experimental results with those of existing models.

Model	Accuracy	Precision	Recall	F1Score
ANN (Chahal et al., 2022)	97.27	96.79	95.67	96.22
ANN (Gauli et al., 2023)	98.66	NA	NA	NA
XGBoost (Chen et al., 2019)	97.82	NA	NA	NA
ANN (Bashir et al., 2021)	98.00	98.30	97.60	97.80
CNN + ADAM (Breviglieri et al., 2021)	95.35	97.56	95.10	96.46
CNN + NADAM (Breviglieri et al., 2021)	97.52	98.67	98.86	98.24
MLP-ELM (Alsirhani et al., 2023)	95.8	90	88	89
ANN based on MLP (Mohsen et al., 2023)	97.82	97.64	98.01	NA

CatBoost classifier	95.06	95.12	97.25	96.17
Proposed CNN	98.82	98.55	98.88	98.77

5. Conclusion

Predictive stability evaluation is expected to be more important for ensuring that a smart grid stays resilient and runs efficiently as it gets better. The reliability and security of an energy network depend on the ability to predict the smart grid's stability via data analytics and machine learning techniques. First, the study offered a variety of mathematical models to assess the dependent characteristics that contribute to smart grid stability. A lightweight sequential CNN for grid stability forecasting was subsequently presented in this paper. The model's lower prediction error rate was revealed by rigorous experimental analysis using ambiguity analysis and the t-SNE score. According to the results of the experiments, there are only 12.7K learnable parameters in the suggested network. The network was trained using 15 epochs in 1 minute and 40 seconds. With a 98.82% test accuracy, 98.55% precision, 98.88% recall rate, and 98.77% F1 score, the suggested network performed admirably in the experiments compared to earlier techniques. A comparative analysis of several mathematical models can be performed in the future. Furthermore, it is necessary to verify the network with respect to the dependent parameter variation

Authors' Contribution: This paper was written collaboratively by the authors. All authors have read and agreed to the published version of the manuscript.

Funding Statement: There is no funding for this manuscript.

Conflicts of Interest: The authors declare that there is no conflict of interest regarding the publication of this paper.

References

Abu Al-Haija, Q., Smadi, A., & Allehyani, M. (2021). Meticulously Intelligent Identification System for Smart Grid Network Stability to Optimize Risk Management. *Energies*, *14*(21), 6935. https://doi.org/10.3390/en14216935

Alazab, M., Khan, S., Krishnan, S. S. R., Pham, Q.-V., Reddy, M. P. K., & Gadekallu, T. R. (2020). A Multidirectional LSTM Model for Predicting the Stability of a Smart Grid. *IEEE Access*, *8*, 85454–85463. https://doi.org/10.1109/ACCESS.2020.2991067

Alsirhani, A., Mujib Alshahrani, M., Abukwaik, A., Taloba, A. I., Abd El-Aziz, R. M., & Salem, M. (2023). A novel approach to predicting the stability of the smart grid utilizing MLP-ELM technique. *Alexandria Engineering Journal*, *74*, 495–508. https://doi.org/10.1016/j.aej.2023.05.063

Arzamasov, V., Bohm, K., & Jochem, P. (2018). Towards Concise Models of Grid Stability. *2018 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, 1–6. https://doi.org/10.1109/SmartGridComm.2018.8587498

Bano, H., Tahir, A., Ali, I., Khan, R. J. U. H., Haseeb, A., & Javaid, N. (2020). Electricity Load and Price Forecasting Using Enhanced Machine Learning Techniques. In L. Barolli, F. Xhafa, & O. K. Hussain (Eds.), *Innovative Mobile and Internet Services in Ubiquitous Computing* (Vol. 994, pp. 255–267). Springer International Publishing. https://doi.org/10.1007/978-3-030-22263-5_25

Bashir, A. K., Khan, S., Prabadevi, B., Deepa, N., Alnumay, W. S., Gadekallu, T. R., & Maddikunta, P. K. R. (2021). Comparative analysis of machine learning algorithms for prediction of smart grid stability [†]. *International Transactions on Electrical Energy Systems*, *31*(9). https://doi.org/10.1002/2050-7038.12706

Bharali, D. (2024). An Improved CNN model for Identifying Tomato Leaf Diseases. *Transdisciplinary Journal of Engineering & Science*, *15*. https://doi.org/10.22545/2024/00253

Breviglieri, P., Erdem, T., & Eken, S. (2021). Predicting Smart Grid Stability with Optimized Deep Models. *SN Computer Science*, *2*(2), 73. https://doi.org/10.1007/s42979-021-00463-5

Chahal, A., Gulia, P., Gill, N. S., & Chatterjee, J. M. (2022). Performance Analysis of an Optimized ANN Model to Predict the Stability of Smart Grid. *Complexity*, *2022*, 1–13. https://doi.org/10.1155/2022/7319010

Chen, M., Liu, Q., Chen, S., Liu, Y., Zhang, C.-H., & Liu, R. (2019). XGBoost-Based Algorithm Interpretation and Application on Post-Fault Transient Stability Status Prediction of Power System. *IEEE Access*, *7*, 13149–13158. https://doi.org/10.1109/ACCESS.2019.2893448

Darbandi, F., Jafari, A., Karimipour, H., Dehghantanha, A., Derakhshan, F., & Raymond Choo, K. (2020). Real-time stability assessment in smart cyber-physical grids: A deep learning approach. *IET Smart Grid*, *3*(4), 454–461. https://doi.org/10.1049/iet-stg.2019.0191

Desai, M., & Mewada, H. (2023). A novel approach for yoga pose estimation based on in-depth analysis of human body joint detection accuracy. *PeerJ Computer Science*, *9*, e1152. https://doi.org/10.7717/peerj-cs.1152

Dhingra, B., Yadav, S., & Tomar, A. (2022). A Comparative Study of Various Machine Learning Algorithms to Detect Power Plant Stability. *2022 2nd International Conference on Innovative Sustainable Computational Technologies (CISCT)*, 1–6. https://doi.org/10.1109/CISCT55310.2022.10046497

England, B. S., & Alouani, A. T. (2020). Real time voltage stability prediction of smart grid areas using smart meters data and improved Thevenin estimates. *International Journal of Electrical Power & Energy Systems*, *122*, 106189. https://doi.org/10.1016/j.ijepes.2020.106189

Franović, B., Baressi Šegota, S., Anđelić, N., & Car, Z. (2023). Decentralized Smart Grid Stability Modeling with Machine Learning. *Energies*, *16*(22), 7562. https://doi.org/10.3390/en16227562

Gauli, M. K., Phoungthong, K., Techato, K., & Gyawali, S. (2023). Predicting the stability of smart grid for improving the sustainability using distributed generation technology. *E-Prime - Advances in Electrical Engineering, Electronics and Energy*, *5*, 100185. https://doi.org/10.1016/j.prime.2023.100185

Gould, Z., Day, S., Reichard, G., Shealy, T., & Saad, W. (2023). A Transdisciplinary Blueprint for Energy Markets Modelled after Mycorrhizal Networks. *Transdisciplinary Journal of Engineering & Science*, *13*. https://doi.org/10.22545/2023/00226

Hong, Y., Zhou, Y., Li, Q., Xu, W., & Zheng, X. (2020). A Deep Learning Method for Short-Term Residential Load Forecasting in Smart Grid. *IEEE Access*, *8*, 55785–55797. https://doi.org/10.1109/ACCESS.2020.2981817

Kotsiopoulos, T., Sarigiannidis, P., Ioannidis, D., & Tzovaras, D. (2021). Machine Learning and Deep Learning in smart manufacturing: The Smart Grid paradigm. *Computer Science Review*, *40*, 100341. https://doi.org/10.1016/j.cosrev.2020.100341

Massaoudi, M., Abu-Rub, H., Refaat, S. S., Chihi, I., & Oueslati, F. S. (2021). Accurate Smart-Grid Stability Forecasting Based on Deep Learning: Point and Interval Estimation Method. *2021 IEEE Kansas Power and Energy Conference (KPEC)*, 1–6. https://doi.org/10.1109/KPEC51835.2021.9446196

Mewada, H., Al-Asad, J. F., Almalki, F. A., Khan, A. H., Almujally, N. A., El-Nakla, S., & Naith, Q. (2023). Gaussian-Filtered High-Frequency-Feature Trained Optimized BiLSTM Network for Spoofed-Speech Classification. *Sensors*, *23*(14), 6637. https://doi.org/10.3390/s23146637

Mishra, M., Nayak, J., Naik, B., & Patnaik, B. (2022). Enhanced Memetic Algorithm-Based Extreme Learning Machine Model for Smart Grid Stability Prediction. *International Transactions on Electrical Energy Systems*, *2022*, 1–10. https://doi.org/10.1155/2022/8038753

Mohsen, S., Bajaj, M., Kotb, H., Pushkarna, M., Alphonse, S., & Ghoneim, S. S. M. (2023). Efficient Artificial Neural Network for Smart Grid Stability Prediction. *International Transactions on Electrical Energy Systems*, *2023*, 1–13. https://doi.org/10.1155/2023/9974409

Moldovan, D., & Salomie, I. (2019). Detection of Sources of Instability in Smart Grids Using Machine Learning Techniques. *2019 IEEE 15th International Conference on Intelligent Computer Communication and Processing (ICCP)*, 175–182. https://doi.org/10.1109/ICCP48234.2019.8959649

Moosavirad, S. H., Torabi, A., & Mirhosseini, M. (2024). Transdisciplinary approach to reduce electricity consumption using system dynamics. *Transdisciplinary Journal of Engineering & Science*, *15*. https://doi.org/10.22545/2024/00249

Omar, M. B., Ibrahim, R., Mantri, R., Chaudhary, J., Ram Selvaraj, K., & Bingi, K. (2022). Smart Grid Stability Prediction Model Using Neural Networks to Handle Missing Inputs. *Sensors*, *22*(12), 4342. https://doi.org/10.3390/s22124342

Önder, M., Dogan, M. U., & Polat, K. (2023). Classification of smart grid stability prediction using cascade machine learning methods and the internet of things in smart grid. *Neural Computing and Applications*, *35*(24), 17851–17869. https://doi.org/10.1007/s00521-023-08605-x

Schäfer, B., Grabow, C., Auer, S., Kurths, J., Witthaut, D., & Timme, M. (2016). Taming instabilities in power grid networks by decentralized control. *The European Physical Journal Special Topics*, *225*(3), 569–582. https://doi.org/10.1140/epjst/e2015-50136-y

Shingade, G., & Shah, S. (2023). Fault Detection And Location Based Svm For Three Phase Transmission Lines Applying Positive Sequence Fault Components. *Transdisciplinary Journal of Engineering & Science*, *14*. https://doi.org/10.22545/2023/00234

Van Zyl-Bulitta, V. H. (2019). Framing Commons for Society-Technology for Electric Infrastructure Supply Systems. *Transdisciplinary Journal of Engineering & Science*, *10*. https://doi.org/10.22545/2019/0113

Yin, D., Yang, Y., Yang, M., Yang, Z., Li, C., & Li, L. (2019). A New Distributed Power System for Stability Prediction And Analysis. *2019 IEEE 10th International Conference on Software Engineering and Service Science (ICSESS)*, 1–4. https://doi.org/10.1109/ICSESS47205.2019.9040711

About the Authors

Hiren K Mewada is an Assistant Research Professor at Prince Mohammad Bin Fahd University, Al Khobar, Kingdom of Saudi Arabia, and held other previous position as an Associate Professor in the Electronics Engineering Department at Charotar University of Science and Technology, Gujarat, India. He has more than 18 years of academic and research experience. Hiren has completed his Ph. D. and M. Tech. in Electronics Engineering from S.V. National Institute of Technology – Surat, Gujarat, India. His research interest lies in the area of Signal/Image Processing ranging from Software simulation to hardware implementation including FPGA and ARM processor-based design. He has focused on the implementation of signal processing and Image segmentation algorithms using FPGA or ARM cortex A8 processor; and Video analysis and summarization applications. He is a life member of IETE and ISTE. During his academic career, he has taught various courses at UG/PG/Ph.D. level and has authored one book and four chapters in different book series. He has completed several research projects from different agencies including the Board of Research in Nuclear Science (BRNS), the Government of India, the Gujarat Council of Science and Technology (GUJCOST), the Government of Gujarat, Solution with Innovation, USA. He has published more than 100 papers including peer-reviewed international journals and conferences. He supervised five PH D students. He has guided more than 33 PG students and 50 UG students.

Dr. Syam Sundar Lingala obtained his doctorate degree in the field of Mechanical Engineering (Energy Systems) from the Jawaharlal Nehru Technological University-Hyderabad, Hyderabad, India in 2010 and Post-Doctoral Research Fellow from the University of Aveiro, Portugal, M. Tech in Thermal Engineering from the Jawaharlal Nehru Technological University-Hyderabad, Hyderabad, India in 2003 and B.E. in Mechanical Engineering from the Andhra University, India in 1998. His research interests include energy, thermal power plants, modeling, and optimization of forced convection heat transfer from tube banks and micro-channel heat sinks, thermal system optimization using entropy generation minimization and genetic algorithm, mixed convection, and Nanofluids. Dr. Syam Sundar Lingala has guided 02 PhDs, and 04 Master thesis and 02 PhDs are ongoing. Dr. Syam Sundar Lingala is in the top 100 100,000 or in the top 2% of the most cited researchers in the world throughout their career and in their scientific field, according to a Stanford University study signed by the team led by John P.A. Ioannidis. He has more than 100 publications in refereed International Journals and 10 book chapters.

Bimal Patel holds a Doctor of Philosophy (PhD) degree in Computer Science Engineering from CHARUSAT University. He received an M. Tech degree in Information Technology from the prestigious L.D. College of Engineering, Ahmedabad, Gujarat Technological University in 2013 with a gold medal. He also received a BE degree in Information Technology from CITC, Changa, Gujarat University, India in 2000. He has guided more than 4 students at Master's level and published more than 17 papers in reputed journals and International Conferences. His research interest includes the Internet of Things, Green Communication, Smart Cities Applications, MANET, wireless sensor networks, and denial of service attacks.

Miral M Desai is Assistant Professor at Charotar University of Science & Technology, Changa, Gujarat, India. He has more than 10 years of academic and research experience. He has completed his Ph. D. from Charotar University of Science & Technology, Changa, Gujarat, India and M. Tech. in Embedded Systems from Institute of Technology, Nirma University, Ahmedabad, Gujarat, India. His Research interest lies in the area of Machine Learning, Deep Learning, Computer vision algorithm, Embedded System Design, Embedded Linux based applications. He is a member of IEEE. He has published 17 papers including peer reviewed international journals and conferences. He is a life member of ISTE.

Mr. Adil Khan did his Ph.D. in Computer Science and Information Technology from UNIVERSITI MALAYSIA SARAWAK, Malaysia. His research is in image segmentation and classification using machine and deep learning. In his Ph.D. thesis, he proposed statistical, traditional machine, and deep learning-based models to segment and classify skin lesions from dermoscopic images. He did MS in Electrical Engineering from KFUPM, Saudi Arabia, where area of research was signal processing. His BS is in Telecommunication Engineering from COMSATS University, Lahore, Pakistan. He secured the Institute and Campus Gold Medal in undergraduate degree.