SSRG International Journal of Electrical and Electronics Engineering
ISSN: 2348-8379/ https://doi.org/10.14445/23488379/IJEEE-V11I12P103

Volume 11 Issue 12, 26-35, December 2024

Original Article

Synchrophasor Driven Voltage Stability Assessment Using Adaptive Deep Learning Based Tools on Temporal Ensembling and Data Augmentation

Kunal Samad[1], Arunkumar Patil[2], Katkar Siddhant Satyapal[3], Santosh Diggikar[4], N. Mohan[5]

[1,2,3,4]*Department of Electrical Engineering, Central University of Karnataka, Karnataka, India.*
[5]*Department of Electrical and Electronics Engineering, SJCE, JSS Science and Technology University, Karnataka, India.*

[2]*Corresponding Author : arunkumarpatil@cuk.ac.in*

Received: 03 October 2024 Revised: 03 November 2024 Accepted: 02 December 2024 Published: 31 December 2024

Abstract - *This research explores the application of Phasor Measurement Units (PMUs) and deep learning methodologies to predict voltage stability by enhancing the efficient operation of the power grid. The study addresses the challenges of topological changes and contingency labeling in power system networks. The methodology incorporates generative artificial intelligence and semi-supervised learning by significantly improving the predictive accuracy of the various learning models. Least Squares Generative Adversarial Networks (LSGANs) strategically augment the training dataset, expanding both the feature space and temporal domain. The enriched dataset enables more accurate classification and enhances the model's reliability under previously unseen dynamic time-series signatures. The Recurrent Neural Network (RNN) based Temporal Ensembling improves the model's ability to determine voltage stability by clustering time signatures based on signal transitions and temporal dynamics. The deep learning model is applied to PMU time series data, which undergoes systematic evaluation criteria using various data preparation stages. In addition, the study explores multiple network topologies for the model's adaptability, testing across diverse time windows and signal conditions. Also, Hyperparameter tuning of the nominated model optimizes the performance through cross-validation, and the best configurations for the best settings were ranked based on the test scores. The findings underscore the potential of artificial intelligence and machine learning models to enhance power system stability. Such forecasts can support proactive decision-making, improve power system operations, and lay the foundation for future advancements in wide-area power system monitoring and control.*

Keywords - *Voltage stability, Phasor Measurement Units (PMUs), Deep Learning, Least Square Generative Adversarial Networks (LSGANs), Recurrent Neural Network (RNNs), Temporal ensembling, Data augmentation.*

1. Introduction

Voltage instability denotes the incapacity of a power system to maintain stable voltage magnitudes at all bus nodes after perturbations. This instability can emanate from various factors, including abrupt load fluctuations, equipment malfunctions, and faults within the system.

Understanding voltage and overall power system stability is essential for ensuring a reliable power supply and preventing catastrophic failures within the grid. To mitigate these risks, grid operators must monitor system conditions and be prepared for potential instabilities. They employ various techniques, such as stability analysis, dynamic simulations, and voltage stability assessments, to anticipate issues before they escalate into significant outages or equipment failures. Traditionally, bus voltage data from Phasor Measurement Units (PMUs) in conjugation with traditional data provides essential insights into a power system's health. However, relying solely on bus voltage data might lack a comprehensive stability analysis because voltage levels influence voltage stability and depend on factors like reactive power flow, load variations, and network configuration.

Voltage stability involves dynamic interactions, where changes in load or disturbances impact both active and reactive power flow. Generally, event detection systems are reactive and identify anomalies or disturbances after the occurrence. Voltage instability often develops gradually and is influenced by cumulative stresses like rising load demand, reactive power deficiencies, or system faults. When traditional event detection systems recognize an unstable event, there may be limited response time. Proactive prediction, on the other hand, leverages models trained to recognize early warning signs based on historical patterns in the power grid data. Predicting the onset of voltage instability proactively, rather than merely detecting it in real-time, is essential for effective grid management because it allows operators to intervene before a critical failure occurs. These models can detect subtle trends that precede voltage instability, such as gradual voltage drops, phase angle shifts, or unusual reactive power patterns. With predictive insights, operators gain crucial lead time to adjust system parameters, redistribute loads, or activate compensatory measures, thus stabilizing the grid before it reaches a critical threshold.

This paper explores the potential of deep learning models for voltage stability forecasting using time-series data obtained from PMU. The analysis of this data uncovers many patterns and anomalies that may indicate emerging instability. Our approach leverages generative artificial intelligence and semi-supervised clustering techniques to extract meaningful features from voltage measurements. This facilitates understanding the complex relationships between voltage patterns and stability margins, even when relying primarily on bus voltage data. The Deep learning models, particularly those designed for time series analysis such as Recurrent Neural Network (RNNs) can learn from complex, non-linear patterns and dependencies that may not be obvious in conventional analysis.

By training for both stable and unstable scenarios, these models can identify early warning signs of voltage instability. To validate the findings, the model must undergo various performance checks under ambiguous circumstances occurring in a large power system. Due to the semi-supervised nature of the objective, the model needs to be verified by projecting diverse operational conditions, such as new faults and topological changes. Also, the models should be immune from undesirable data quality issues. Examining existing studies and theoretical perspectives relevant to this topic is essential. The following literature review explores previous work in this area, highlighting key findings, gaps, and insights that will guide the approach of this research.

2. Literature Review

To build a solid foundation for addressing the identified research problem, the paper [1] introduces a novel hybrid approach that leverages a combination of randomized learning algorithms for enhancing post-fault short-term stability assessments of voltage. This system integrates technologies such as Random Vector Functional Link networks (RVFL) and Extreme Learning Machines (ELM) applied to the New England 39-bus system, achieving a performance increase of 27.5%-27.3% over standalone methods.

A paper [2] demonstrates that LSTM networks can be utilized following semi-supervised clustering with Constraint-partitioning (k-means) to forecast temporal dependencies beneficial for STVS. This study deployed an LSTM-based model for STVS to evaluate an IEEE 39-bus system, achieving superior prediction accuracy and dynamic computational efficiency of time series.

Another paper [3] discusses a hybrid real-time method for assessing short-term voltage stability by using RNNs for both spatial and temporal data. The effectiveness of a hybrid model was evaluated on both the IEEE 9 Bus and New England 39 Bus systems. In paper [4], SVC and SVG hybrid reactive power compensation was enhanced for below and above 25hz reactive power change using physics-informed based DL. Predictive control algorithms used dynamic PMU voltage measurement at regional nodes. Similarly, a transferable deep learning-based model that utilizes physics-informed topological features is constructed using PMU data-the model evaluated on the IEEE 39-bus system for STVS assessment in the paper [5].

A PMU measurements-based deep transfer learning was also introduced in the paper [6], showing improved adaptability to topological changes and maintaining high performance under varying conditions. The model was evaluated on the data quality of various hyperparameters on the IEEE 39 bus test system. In the paper [7], topology-aware voltage dynamic features are based on the principle of deriving features that capture the essential dynamics. CNN model is used as a classifier to fit the rule-based features and STVS status.

Another study in [8] explored the application of Variational Autoencoders (VAE) in voltage stability assessment, demonstrating enhanced accuracy across various IEEE standard systems by utilizing VAE to regulate latent variables and extract significant low-dimensional data representations. Further research [9] introduced a rule of disagreement-based learning model for voltage stability monitoring, achieving over 94.03% and demonstrating adaptability to network conditions or topology alteration with a transfer learning technique.

Additionally, a method to integrate STVS evaluation with the reconfiguration of the distribution network was developed [10], employing a custom CNN to optimize distribution network voltage stability, confirmed by case studies on both modified 69-bus networks. Another advancement, like in [11], involved a stability-constrained based Deep Reinforcement Learning (DRL) method for voltage control dynamically, which reduced transient control costs and shortened response times while ensuring voltage stability.

Research [12] addressed small transient rotor angle stability utilizing CNN for transient stability classification and LSTM to monitor oscillatory responses in stable conditions, with extensive simulations verifying effectiveness across multiple IEEE test systems. Also, introducing a transfer learning for STVS assessment showed high classification accuracy and adaptability to new grid topologies, effectively capturing topological, temporal patterns post-disturbance and PMU errors [13].

A Temporal Convolutional Neural (TCN) was proposed for real-time STVS assessment [14], achieving high temporal sequence forecasting of power system parameters in an IEEE 9-bus system. A similar convolution framework, like a Graph Convolution Neural (GCN) based DRL framework, was developed [15], showing superior performance on the IEEE-39 bus system by enhancing the DRL agent's ability to recognize nodal spatial correlations like a graph-structured nature of power system data.

Furthermore, a contingency control scheme based on DRL was developed to optimize control actions under emergency conditions in a test-bed New England 39-bus system. Considering fault as a feature to reinforce the learning model to devise an optimal controlling scheme [16]. Lastly, an adaptive online learning algorithm with momentum for voltage stability assessment was introduced [17], showing high accuracy on the NETS-NYPS 68-bus system.

The investigation [18] analyzed various Artificial Neural Network (ANN) strategies for steady-state stability of power system networks, pinpointing the Feedforward (FFNN) and Cascade Forward Neural Network (CFNN) as the top performers. Utilizing the New Voltage Stability Pointer (NVSP) to refine input data, trained via a Levenberg-Marquardt (LM) method, and evaluated using regression algorithms of IEEE 30-bus and Nigerian power grid.

A related study [19] introduced a deep transfer learning approach for transient stability assessment, employing a CNN (ImageNet library). The model adapted to assess the correlation between disturbance intensity and transient stability is well suited for spatio-temporal analysis of power grid data. The methodology was validated using simulations on IEEE 39 and IEEE 118-bus models. In the research [20], profound learning-supported Stochastic Distribution Network Reconfiguration (SDNR) technique was introduced to enhance voltage stability. This method utilized a CNN to estimate voltage stability indices from network reconfiguration decisions, integrating these predictions into branch reduction algorithms to optimize the radial topology, thereby diminishing power losses and boosting voltage stability on IEEE network models.

Therefore, the optimal methodology for voltage stability strongly correlates with the data type, resolution and depth of system information. The literature highlights different learning techniques in AI and how the objective strongly depends on the type of extracted features. Extracted features like space, time, and contingency types were combined with deep learning models for forecasting, classification and determining optimal operation schemes. The literature carefully highlighted all types of neural networks developed in the last five years, emphasising synchrophasor data for overall voltage and power system stability.

2.1. Overview of the Paper

Following the literature survey, this paper delves into several crucial aspects of enhancing voltage stability through advanced deep learning. Initially, it discusses data preparation, emphasizing the importance of high-resolution data collection, understanding event impacts on the grid, and simulating diverse fault scenarios.

The exploration continues by analyzing LSGANs for synthesizing augmented data to enrich training and train models on rare and variable scenarios. Subsequently, the paper explores the benefits of temporal ensembling using Recurrent Neural Networks (RNNs) to enhance prediction accuracy and stability in semi-supervised settings by clustering similar time series signatures into clusters.

Nevertheless, this creates a base for temporal-based forecasting and classification. Further, the nominated models were evaluated based on multiple parameter configurations like data preparation, temporal and spatial augmentation, and Signal Noise Ratio (SNR). Consequently, the hyperparameter tuning of the best model was validated and ranked based on the best score of respective parameter settings. Finally, the conclusion summarizes the findings and emphasizes the effectiveness of data implementing artificial intelligence in improving voltage stability predictions while suggesting future research directions.

3. Database Generation

A comprehensive database capturing post-fault time signatures using the IEEE standard 6-Bus System was generated to enhance the understanding of dynamic response within a power system. It is commonly referenced in power system studies, particularly when examining voltage stability and contingency analysis. This database is a valuable resource for analyzing the system’s behavior following disturbances, particularly transmission line outages and variations in load and generation ratios.

Several key factors, such as data quality, resolution, and time steps in data collection, were considered, which overall impacted the accuracy, requiring a balance between detail and practical limitations. The cause of an event may not map uniquely to physical phenomena, as multiple events can stem from common or varied underlying causes. Also, the same events may impact different grid locations, depending on specific grid characteristics.

Time-domain simulations are crucial for understanding post-disturbance system behaviors. Fault simulations (line-to-line, line-to-ground, three-phase) evaluate system responses, and incorporating diverse fault cases in datasets reveals system vulnerability and resilience. In [21, 22], the emphasis is on considering diverse load scenarios in datasets, such as collecting transient and short-term operation data to improve voltage stability models' accuracy. Managing uncertainties in modern power grids requires advanced data analytics to determine recommended settings for real-time monitoring and autonomous control systems. Improving model generalization and developing systems autonomously responding to changes ensures continuous adaptation to evolving grid conditions. Figure 1 is a block diagram for clustering time series data using RNNs in power system analysis. It starts with an input data block for PMU measurements and time series data, followed by a data preprocessing step. The pre-processed data is then fed into an RNN encoding block, where features are extracted to capture temporal dependencies. These features are used in the clustering block to group similar data points. Cluster centers are dynamically updated based on the clustering output. A convergence check ensures the process iteratively refines until stable clusters are formed. The system includes blocks for model training, performance testing, and parameter optimization. This diagram efficiently outlines the workflow for applying advanced machine-learning techniques to enhance power system analysis.

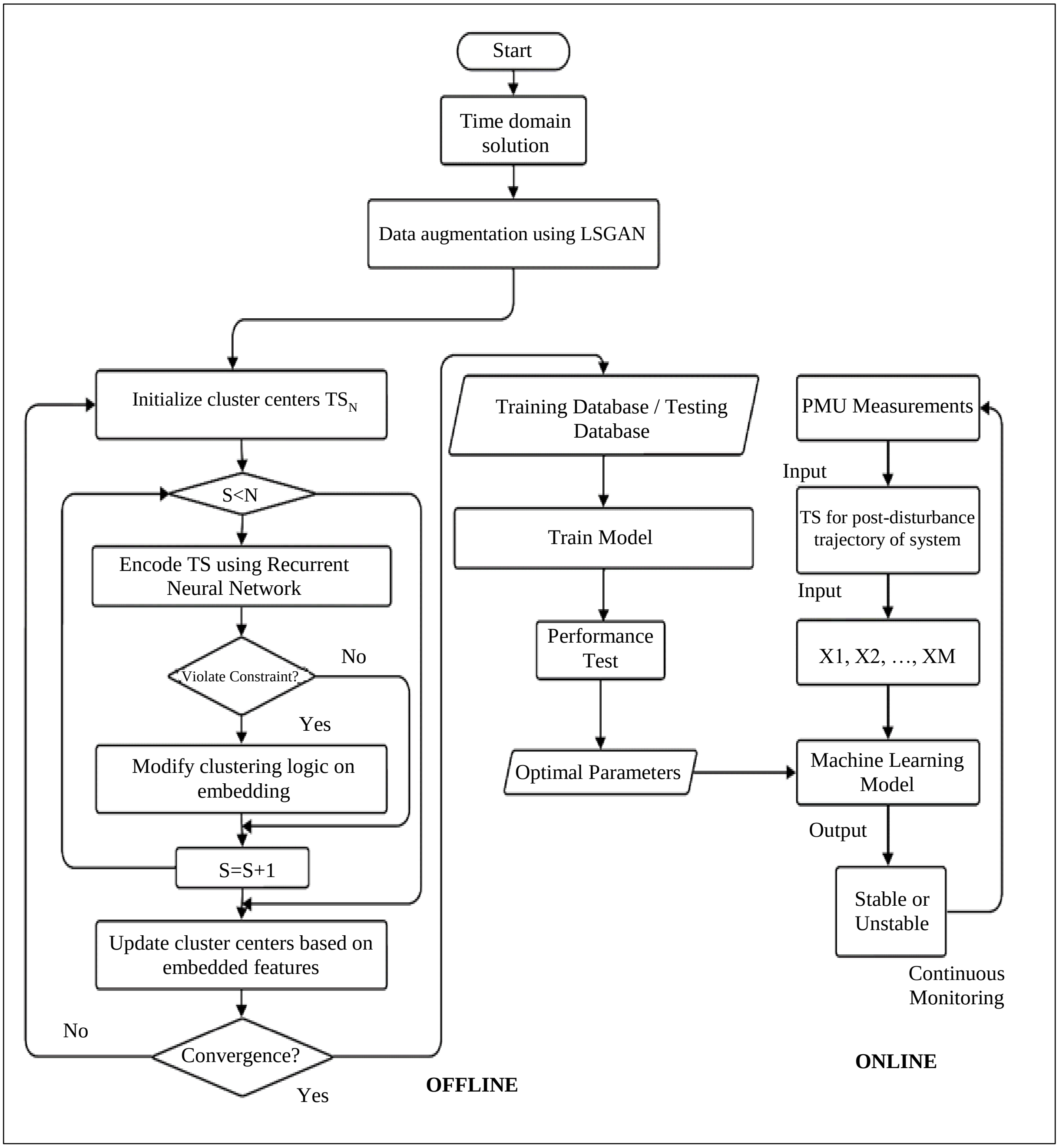

Fig. 1 Dynamic clustering of electric grid data using Recurrent Neural Network.

4. Least Square Generative Adversarial Networks (LSGAN)

LSGAN is a powerful class of neural networks used to generate synthetic data. It can be particularly useful in creating diverse fault scenarios for training models in voltage stability analysis of power networks. By augmenting PMU data with synthetic samples, LSGANs can help address data scarcity and variability. An LSGAN consists of two main components: a Generator (G) task, which generates new data instances from random noise; The generator tries to make the synthetic data as close as possible to the real dataset, attempting to fool the discriminator. Discriminator’s (D) task is to distinguish between instances of the actual and generated data. It is essentially a binary classifier.

$$[\![\min]\!]\_G \; [\![\max]\!]\_D \; V(D,G) = E\_(x \sim Pdata)\; [\log D(X)] - E\_(Z \sim P\_Z\,(Z))\; [\log(1-D(G(z)))] \quad (1)$$

In LSGAN training, the generator and discriminator improve through competition, represented by a min-max game value function. In standard GANs, the discriminator improves at distinguishing real from generated samples, but this can cause the generator's gradient to vanish if the discriminator becomes too effective. LSGANs solve this by replacing cross-entropy loss with least squares loss, providing a more stable gradient for the generator. The loss of the least squares penalizes samples far from the decision boundary more heavily, improving gradient stability and reducing training stalls. This approach forces generated samples near the real data distribution boundary, reducing mode collapse. Using a quadratic cost for misclassified samples, LSGANs provide a stronger gradient signal for the generator, effectively addressing the vanishing gradient problem and enabling continuous, effective training on complex datasets.

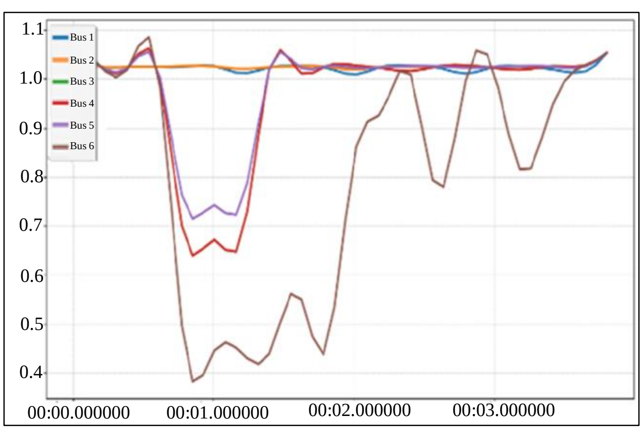

Fig. 2 Augmented data based on the original data set using LSGAN closely depicts probable power system fault type

Figure 2 represents augmented data (time-augmentation) generated using a Least-Squares Generative Adversarial Network (LSGAN) based on an original power system dataset. Each line corresponds to a bus voltage response during a simulated fault, with variations in dips and oscillations showing probable fault types in the system. The pronounced voltage dips, especially in buses 5 and 6, highlight the LSGAN's capability to simulate severe fault conditions.

5. Temporal Ensembling

This technique is primarily used in semi-supervised learning with neural networks. This technique leverages the consistency of predictions by averaging a model's predictions for multiple training epochs over time, aiding in regularization and reducing overfitting for clustering time series data, such as PMU data for voltage stability.

$$\text{Ensemble}^{(t+1)} = \beta * \text{Ensemble}^{(t+1)} + (1-\beta)\text{Prediction}^{(t)} \quad (2)$$

This formula updates the ensemble prediction at each time step (t + 1) by combining the previous and current predictions. The decay factor determines how much weight is given to past predictions versus the current prediction. This is crucial for capturing the temporal dependencies in PMU data.

$$L = \frac{1}{N}\sum_{N}^{i=1} \omega_i * (\text{Ensemble}_i - \text{PMU}_i)^2 \quad (3)$$

The loss (L) calculated between the ensemble predictions and the actual PMU measurements guides the weight assigned to each measurement based on its reliability or variance, ensuring that more reliable measurements have a greater influence on the learning process.

$$S = \lambda \sum_{t=2}^{T} (\text{Ensemble}^{(t)} - \text{Ensemble}^{(t-1)})^2 \quad (4)$$

This smoothness constraint (S) penalizes large fluctuations between consecutive ensemble predictions, helping to ensure that the prediction evolves smoothly over time. The regularization parameter controls the extent of this smoothing, which is important for reflecting the physical continuity in the electrical system.

$$C = \frac{1}{N}\sum_{i=1}^{N} (\text{Prediction}_i^{(t)} - \text{Ensemble}_i^{(t)})^2 \quad (5)$$

The dynamic consistency cost (C) encourages the current predictions to converge to ensemble predictions. This promotes consistency across time steps, essential for the stable clustering of PMU data. Combining these components into a unified learning objective with the loss function, smoothness constraint, and dynamic consistency cost improves the prediction accuracy and system response.

6. Results and Discussion

Evaluating model performance is crucial in data science and machine learning, especially when forecasting voltage stability in power systems through advanced methodologies. The primary evaluation metrics are Accuracy, Silhouette Coefficient, Mean Squared Error (MSE), Mean Absolute

Error (MAE), Root Mean Squared Error (RMSE), and Signal-to-Noise Ratio (SNR) Performance. These are essential metrics for assessing various aspects of model effectiveness, ensuring reliable outcomes across diverse applications.

- Accuracy: Accuracy measures true results (both true positives and true negatives) within the total number of cases examined, highlighting the model's overall effectiveness. It is particularly relevant in balanced datasets where the cost of different types of errors is similar.
- Silhouette Coefficient: Used predominantly in unsupervised learning to assess clustering quality. This metric gauges how closely an object aligns with its cluster compared to others. A high Silhouette Coefficient indicates strong intra-cluster similarity and clear differentiation from neighboring clusters, making it essential for evaluating temporal ensembling techniques.
- Mean Absolute Error (MAE): These metrics quantify average absolute differences between predicted and actual values (MAE). They are vital for regression analysis, providing direct error magnitude and model accuracy measures.
- Root Mean Squared Error (RMSE): This metric offers a measure of accuracy that gives relatively high weight to large errors by taking the square root of the average squared differences between predictions and actual observations.
- Signal-to-Noise Ratio (SNR) Performance: Evaluate the model's robustness under various noise levels, maintaining accuracy across different signal-to-noise conditions. This is critical for models operating in dynamic and potentially noisy power system environments.

6.1. Evaluation of the Models

Deep learning models generally outperform traditional machine learning models across key performance metrics. Models like RNN networks achieve exceptionally low MSE and high accuracy for non-linear data structures like large power grids. In contrast, traditional models like Linear Regression and Decision Trees show higher error rates and lower overall accuracy, particularly in tasks involving sequential or high-dimensional data. So, comparing model performance intends to establish that deep learning models are more suitable.

Table 1 presents the performance of Deep Learning Models like GRU-LSTM and Bi-LSTMT. These models effectively handle both short-term and long-term dependencies, with high accuracy rates nearing 99%. LSTM and GRU Models offer low MSE and high accuracy (above 98%), which is excellent for sequential data processing. CNN & CNN-LSTM stands out for low MSE and high accuracy (above 99%), demonstrating the superior capability to extract complex data features.

Table 1. Performance metrics comparison of AI models

Model	MSE	MAE	RMSE	Accuracy
LSTM	0.0001	0.010	0.012	0.987
GRU	0.0001	0.012	0.013	0.986
Bi-GRU	0.0001	0.011	0.013	0.986
Bi-LSTM	7.649	0.007	0.008	0.991
CNN	3.410	0.004	0.005	0.994
CNN-LSTM	8.215	0.006	0.009	0.990
GRU-LSTM	**0.0000**	**0.001**	**0.001**	**0.998**

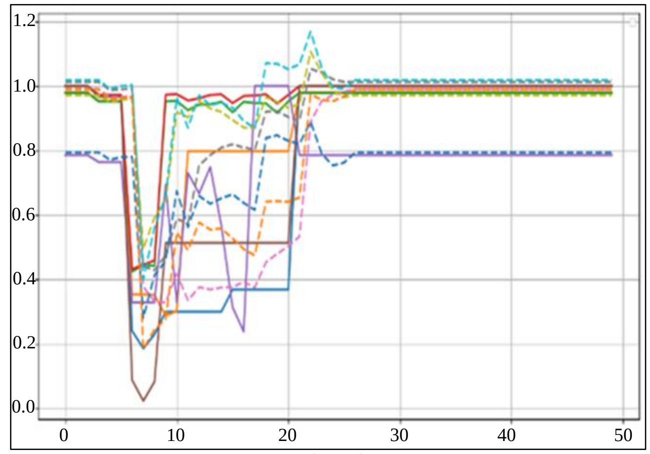

Fig. 3 It shows the overall trend of the forecast across the whole dataset

Figure 3 shows the plot for a six-bus power system's actual and predicted voltage responses following fault events. The x-axis represents the time step, and the y-axis represents the time series value. The plot highlights areas where predictions diverge from actual values, offering insights into model performance and fault detection across the buses.

6.2. Comparison of Clustering Performance

Clustering methods play a pivotal role in categorizing and understanding different states of voltage stability based on temporal data when using PMUs for voltage stability analysis. Temporal ensembling, compared to methods like K-Nearest Neighbors (KNN), DBSCAN, and other clustering techniques, often demonstrates superior performance due to its specific suitability for handling the nuances of temporal data.

Table 2. Comparison of silhouette scores for different clustering methods

Methods	K-Means	DTW	Temporal Ensembling
Silhouette score	0.52	0.59	**0.658**

Table 2 presents the Silhouette Scores of three clustering methods, K-Means, Dynamic Time Warping (DTW), and Temporal Ensembling, for clustering time series data. For K-Means, the Silhouette Score of 0.52 indicates moderate clustering quality. Dynamic Time Warping (DTW) scores

0.59, suggesting better performance than K-Means by handling time series data's temporal dynamics more effectively. Temporal Ensembling scores 0.658, the highest among the three, demonstrating superior clustering quality by leveraging the consistency of predictions over multiple training epochs. Overall, RNN-based temporal ensembling provides the most accurate and reliable clustering results.

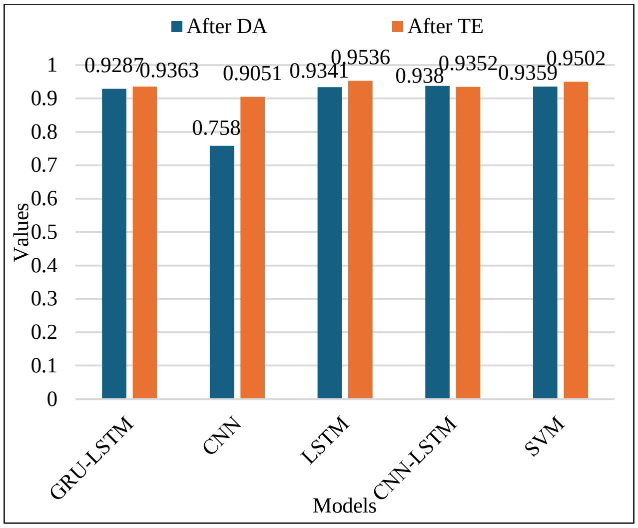

Fig. 4 Comparison of model performance based on data preparation level

Figure 4 illustrates the accuracy of various machine learning models for voltage stability analysis in power systems, focusing on two stages: Data Augmentation (DA) and Temporal Ensembling (TE) with Recurrent Neural Networks (RNNs). Key observations: The analysis shows that the GRU-LSTM hybrid model achieves the highest accuracy for augmented datasets, making it the best choice for stability analysis.

The CNN-LSTM model also performs well but is slightly less accurate than the GRU-LSTM hybrid. Also, significant accuracy gains in GRU-LSTM and LSTM models post-TE, highlighting their proficiency in handling temporal data. CNN shows substantial initial improvement but slightly decreased after TE, indicating a possible misalignment with temporal processing needs. CNN-LSTM benefits from both stages, demonstrating its hybrid spatial and temporal data handling capability.

6.3. Performance under Different Topologies

This analysis evaluates the model's performance by augmenting additional features rather than increasing the time stamps to emulate topological changes. This approach allows the model to gain exposure to new structural scenarios, enriching its ability to generalize across varied grid configurations. Traditional datasets typically provide static topologies, limiting the model's adaptability to real-world changes in power system structures. The model learns patterns associated with different topologies by adding more features (representing virtual buses).

Table 3. Model performance under various topologies

Model	Accuracy	RMSE	MAE	MSE
Original	0.98411	0.0158	0.01402	0.000252
T1	0.98852	0.0114	0.00951	0.000131
T2	0.99746	0.00253	0.00161	0.000006
T3	0.99817	0.00182	0.00151	0.000003
T4	0.99805	0.00194	0.00143	0.000003

Table 3 shows a clear improvement in metrics with increased augmentation. The "Original" configuration achieved an accuracy of 0.9841, while augmented models (T1, T2, T3) reached progressively higher accuracies, with "T3" achieving 0.99817 and the lowest error metrics (RMSE: 0.00182, MAE: 0.00151, MSE: 0.000003). This approach highlights the benefits of augmentation, as it enhances the model's performance with horizontal (space) augmentation. Incorporating topological diversity in training data to ensure reliable performance is key for any learning model.

6.4. Performance under Noise (SNRs/with another Model)

This study uses Signal-to-Noise Ratio (SNR) to examine deviations from the ideal fundamental frequency in noise-free PMU bus voltage time series data, which is standardized in per unit (PU) values. Using PU voltage levels ensures consistency across buses and enhances the comparability of SNR calculations across different system conditions. Signal Definition and PU Normalization: In the noise-free PU voltage data, variations from the expected 50Hz or 60Hz sinusoidal waveform represent "noise" relative to the ideal signal. Using PU voltage measurements provides a consistent baseline, facilitating accurate assessment of deviations without bias due to differing voltage magnitudes across the buses.

SNR Calculation and Interpretation: With the ideal PU voltage signal as a reference, SNR calculations quantify the strength of the fundamental frequency compared to observed deviations expressed in PU. Higher SNR values suggest stability and adherence to the ideal waveform, while lower values indicate system dynamics or potential disturbances. In evaluating the resilience of various models against noise and varying signal-to-noise ratios (SNRs), the primary focus is on the accuracy metric, which directly reflects each model's ability to predict outcomes despite noise interference correctly. In Table 4, the accuracy of deep learning models is compared under 10, 30, and 50 --dB SNR. This measure is crucial in noisy environments where maintaining prediction reliability is essential for voltage stability assessments.

Table 4. Comparison of model performance under different SNR

SNR (dB)	GRU-LSTM	CNN	LSTM	CNN-LSTM	SVM
Noise-free	0.9363	0.9051	0.9536	0.9352	0.9502
10dB	0.9287	0.7582	0.9343	0.9389	0.9359
30dB	0.9337	0.9352	0.9346	0.9377	0.9359
50dB	0.9265	0.8333	0.9347	0.9367	0.9359

The GRU-LSTM hybrid model is the most robust and accurate choice for predicting time series data across a spectrum of noise conditions. While the CNN and LSTM models have their strengths at specific noise levels, the GRU-LSTM hybrid offers the best overall performance, making it the preferred model for tasks involving noisy time series prediction. As the SNR decreases, the challenge of distinguishing signal from noise grows, affecting all models but impacting the GRU-LSTM hybrid the least, thereby underscoring its superior noise-handling capabilities.

6.5. Performance under Different Time Window

Assessment of the performance of GRU-LSTM, CNN-LSTM, CNN, and SVM models across varying sequence lengths (10, 30, 50) is essential to understanding how each model processes temporal dependencies and extracts features. This understanding is crucial because the structural nuances of each model influence their capability to handle different observation time windows (OTWs). The following detailed analysis breaks down the impact of each model's architecture on its performance across these varied temporal scales.

Table 5. Comparison of nominated models' performance under different OTW

Model	(OTW)	MSE	MAE	RMSE
GRU-LSTM	3	6.672	0.006	0.008
	9	8.522	0.006	0.009
	15	8.630	0.006	0.009
CNN-LSTM	3	2.847	0.004	0.005
	9	1.799	0.003	0.004
	15	1.0327	0.0025	0.003
CNN	3	0.0001	0.009	0.010
	9	3.341	0.004	0.005
	15	2.251	0.003	0.004

In Table 5, The GRU-LSTM exhibits increasing error metrics as the sequence length increases from 3 to 15. This trend suggests that while the model can capture both short-term and long-term dependencies, it might suffer from overfitting or computational complexity when dealing with longer sequences. The consistency in MAE across different OTWs indicates that the model's average error magnitude remains stable. However, the increasing MSE and RMSE suggest growing variance in the error distribution or outlier effects as the sequence lengthens. Contrasting with GRU-LSTM, the CNN-LSTM shows improved performance as the sequence length increases, with all three metrics decreasing. This model effectively combines CNN's ability to extract spatial or feature-based information efficiently with LSTM's proficiency in capturing temporal dependencies, making it highly effective for longer sequences. The CNN performs best at the shortest OTW, indicating its strength in handling spatial features over short sequences.

6.6. Model Architecture and Hyperparameter Tuning

Hyperparameter tuning is conducted to optimize the GRU-LSTM model's performance for sequence prediction tasks. The primary objective was to identify the most practical combination of units, epochs, batch size, and dropout rate to maximize the model's predictive accuracy while minimizing overfitting. Table 6 summarizes the results of our experiments, showcasing the mean test score and standard deviation for various configurations of the GRU-LSTM model. The results of different configurations of the GRU-LSTM model by varying the hyperparameters are as follows.

Table 6. Hyperparameter tuning results for GRU-LSTM model (ranked from high to low score)

Units	Epoch	Batch Size	Dropout Rate	Standard Test Score	Mean Test Score
75	10	10	0.1	-0.00027	0.000028
50	10	10	0.1	-0.00038	0.000139
50	20	20	0.2	-0.00043	0.000091
25	10	10	0.1	-0.00147	0.000397
25	10	0.3	0.2	-0.00187	0.000138

The tuning results are presented in Table 6, where each row represents a unique combination of hyperparameters and the corresponding mean and standard deviation of test scores. The mean test score indicates the average performance across multiple runs, while the standard deviation reflects the stability and consistency of each configuration. The results reveal notable insights into the accuracy of each hyperparameter of the respective model.

The tuning results indicate that configurations with 50 units ranked consistently higher than those with 25 units, suggesting that more units improve the model's ability to interpret temporal patterns effectively. The top configurations used 50 units, achieving mean test scores of -0.000383 and -0.000433, respectively. Increasing the number of epochs from 10 to 20 generally resulted in a minimal change in mean test score, implying that ten epochs may suffice for convergence.

Batch size also impacted performance, with a smaller batch size of 16 outperforming larger sizes in configurations with 50 units, likely due to the increased number of gradient updates per epoch. Regarding regularization, lower dropout rates, particularly 0.1, were more favorable, as seen in the highest-ranked configurations. Increasing the dropout rate to 0.2 or 0.3 led to higher mean test scores, suggesting that lighter regularization may be better suited for this model.

Finally, regarding consistency, the configuration with 50 units, 20 epochs, 0.1 dropouts, and a batch size 16 demonstrated the lowest standard deviation (0.000091), indicating the most stable performance. Higher dropout rates were generally associated with greater variability, as shown in the configurations ranked 3 to 5. The results from hyperparameter tuning suggest that the best configuration for the GRU-LSTM model is 50 units, ten epochs, a dropout rate of 0.1, and a batch size 16. This configuration achieved the lowest mean test score, indicating superior predictive performance while maintaining a relatively low standard deviation.

7. Conclusion

This study tackles the challenge of forecasting voltage instability within power systems by employing sophisticated deep-learning methods. It innovatively applies generative AI to address the inherent limitation of sparse datasets in these dynamic systems. The work also highlights the capability of RNNs in time-series clustering. This research validates the hypothesis that prior data augmentation followed by temporal ensembling enhances model generalization and predictive outcomes of the GRU-LSTM models. The methodological core of this investigation involved a detailed evaluation of multiple deep learning architectures. These models were assessed using diverse data preparation to determine the forecasting efficacy of power system dynamics under varied operational scenarios. The model's effectiveness was evaluated using the IEEE 6 Bus System for various deep learning models, especially the hybrid GRU-LSTM and CNN-LSTM methods, including various parameter settings of the nominated deep learning model. The best hyperparameter settings were ranked based on the test score. These tests confirm that the deep learning models perform well under semi-supervised conditions and produce less error than standard supervised machine learning models. Future research will benefit from integrating emerging AI trends like transfer and quantum machine learning by enhancing the power system's predictive accuracy and operational resilience. These advancements will be crucial in developing sophisticated contingency plans and ensuring optimal responses to dynamic grid conditions by maintaining the security of the power system.

Conflicts of Interest

The time series data was generated using MiPOWER software by conventional data preparation requirements. The IEEE Standard 6-Bus system is referenced for database generation in power system studies. The sorting of data generation, resolution, and authenticity of the data may differ compared with standards at the industry level. The model is assumed to work on any data with minor adjustments of data structure before the learning model training.

References

[1] Chao Ren et al., "A Hybrid Randomized Learning System for Temporal-Adaptive Voltage Stability Assessment of Power Systems," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 6, pp. 3672-3684, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[2] Meng Zhang et al., "Deep Learning for Short-Term Voltage Stability Assessment of Power Systems," *IEEE Access*, vol. 9, pp. 29711-29718, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[3] Lipeng Zhu, David J. Hill, and Chao Lu, "Intelligent Short-Term Voltage Stability Assessment via Spatial Attention Rectified RNN Learning," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 10, pp. 7005-7016, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[4] Shu Liu et al., "Improved Model Predictive Dynamic Voltage Cooperative Control Technology Based on PMU," *Frontiers in Energy Research*, vol. 10, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[5] Manish K. Singh et al., "Physic-Informed Transfer Learning for Voltage Stability Margin Prediction," *2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Rhodes Island, Greece, pp. 1-5, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[6] Yang Li et al., "PMU Measurements-Based Short-Term Voltage Stability Assessment of Power Systems via Deep Transfer Learning," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[7] Mohamad Khairuzzaman Mohamad Zamani, Ismail Musirin, and Saiful Izwan Suliman, "Convolutional Neural Network for Voltage Stability Prediction in Power System Operation," *International Journal of Emerging Trends in Engineering Research*, vol. 8, no. 1.1, pp. 205-212, 2020. [CrossRef] [Publisher Link]

[8] Haosen Yang et al., "Deep Learning Architecture for Voltage Stability Evaluation in Smart Grid Based on Variational Autoencoders," *arXiv Preprint*, 2018. [Google Scholar]

[9] Tong Wu, Ying-Jun Angela Zhang, and He Wen, "Voltage Stability Monitoring Based on Disagreement-Based Deep Learning in a Time-Varying Environment," *IEEE Transactions on Power Systems*, vol. 36, no. 1, pp. 28-38, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[10] Wanjun Huang, Weiye Zheng, and David J. Hill, "Distribution Network Reconfiguration for Short-Term Voltage Stability Enhancement: An Efficient Deep Learning Approach," *IEEE Transactions on Smart Grid*, vol. 12, no. 6, pp. 5385-5395, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[11] Yuanyuan Shi et al., "Stability Constrained Reinforcement Learning for Real-Time Voltage Control," *2022 American Control Conference (ACC)*, Atlanta, GA, USA, pp. 2715-2721, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[12] Syafiq Kamarul Azman et al., "A Unified Online Deep Learning Prediction Model for Small Signal and Transient Stability," *IEEE Transactions on Power Systems*, vol. 35, no. 6, pp. 4585-4598, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[13] Zijian Feng et al., "Transferable Deep Learning Power System Short-Term Voltage Stability Assessment with Physics-Informed Topological Feature Engineering," *arXiv Preprint*, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[14] Ananta Adhikari, Sumate Naetiladdanon, and Anawach Sangswang, "Real-Time Short-Term Voltage Stability Assessment Using Temporal Convolutional Neural Network," *2021 IEEE PES Innovative Smart Grid Technologies - Asia (ISGT Asia)*, Brisbane, Australia, pp. 1-5, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[15] Ramij R. Hossain, Qiuhua Huang, and Renke Huang, "Graph Convolutional Network-Based Topology Embedded Deep Reinforcement Learning for Voltage Stability Control," *IEEE Transactions on Power Systems*, vol. 36, no. 5, pp. 4848-4851, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[16] C.X. Jiang et al., "Power System Emergency Control to Improve Short-Term Voltage Stability Using Deep Reinforcement Learning Algorithm," *2019 IEEE 3rd International Electrical and Energy Conference (CIEEC)*, Beijing, China, pp. 1872-1877, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[17] Zhijie Nie et al., "Adaptive Online Learning with Momentum for Contingency-Based Voltage Stability Assessment," *2019 IEEE Power & Energy Society General Meeting (PESGM)*, Atlanta, GA, USA, pp. 1-5, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[18] Tayo Uthman Badrudeen, Nnamdi I. Nwulu, and Saheed Lekan Gbadamosi, "Neural Network Based Approach for Steady-State Stability Assessment of Power Systems," *Sustainability*, vol. 15, no. 2, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[19] Jongju Kim et al., "Transient Stability Assessment Using Deep Transfer Learning," *IEEE Access*, vol. 11, pp. 116622-116637, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[20] Wanjun Huang, and Changhong Zhao, "Deep-Learning-Aided Voltage-Stability-Enhancing Stochastic Distribution Network Reconfiguration," *IEEE Transactions on Power Systems*, vol. 39, no. 2, pp. 2827-2836, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[21] Guru Mohan Baleboina, and R. Mageshvaran, "A Survey on Voltage Stability Indices for Power System Transmission and Distribution Systems," *Frontiers in Energy Research*, vol. 11, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[22] Yan Chen et al., "Transient Voltage Stability Assessment and Margin Calculation Based on Disturbance Signal Energy Feature Learning," *Frontiers in Energy Research*, vol. 12, 2024.