Journal of Business Insight and Innovation
Volume 3 Issue 1, 2024
ISSN-p: 3006-2284, ISSN-e: 3006-0982
https://insightfuljournals.com/

ARTIFICIAL INTELLIGENCE-BASED VOLTAGE STABILITY ASSESSMENT IN MODERN POWER SYSTEMS: EDGE INTELLIGENCE FOR REAL-TIME MONITORING OF SMART ELECTRICAL NETWORKS

Muhammad Noman Amjad [1]

Affiliations:

[1] Manager ICT, PMS Pvt. Ltd
Gujranwala, Pakistan

Email: noman5802@gmail.com
nomanict4@gmail.com

Corresponding Author/s Email:

[1] noman5802@gmail.com

Copyright:

Author/s

License:

CC BY NC SA

Abstract

Voltage stability has become a critical operational concern in modern power systems because of rising renewable energy penetration, decentralised generation and heavier line loadings have narrowed the margin between secure operation and voltage collapse. While traditional model-based indices are accurate but slow, centralised machine-learning methods suffer from communication latency, which restricts their real-time applicability. This paper introduces an Edge-Intelligence framework, where a small hybrid Convolutional Neural Network — Long Short-Term Memory (CNN-LSTM) classifier is placed on edge nodes in the substation to provide millisecond-scale voltage-stability assessment based on phasor measurement unit (PMU) data. The pipeline is tested on 12,000 operating scenarios generated on IEEE 14-, 30- and 118-bus test systems for normal, stressed, contingency and post-fault conditions. The proposed model achieves 94.5% accuracy, macro-F1 of 0.941 and ROC-AUC of 0.999 on the combined test set, outperforming four strong baselines (logistic regression, RBF-SVM, random forest and gradient boosting) by 12-14 percentage points in accuracy, while requiring only 186 kB of memory and delivering submillisecond inference on a Jetson-class edge device. In accordance with classical L-index theory, the explainability analysis shows that the most important predictors are bus-voltage magnitude, reactive-power injection and load factor. The results show that Edge-Intelligence is a viable, low latency, and low footprint path towards always-on voltage-stability monitoring in smart electrical networks.

Keywords: Voltage Stability; Edge Intelligence; Smart Grid; Phasor Measurement Unit; Deep Learning; CNN-LSTM; IEEE Test System; Real-Time Monitoring; L-Index; Power System Security

1. Introduction

Three forces are converging to transform modern power systems: the massive scale of inverter-based renewable generation, the electrification of transport and heating, and the installation of high-resolution digital measurement infrastructure. These forces have pushed the classical operating point of transmission networks closer to their voltage-stability limits and have reduced the time between the onset and the detection and resolution of incipient voltage-security violations. Slow voltage collapse is a repeatedly identified triggering or sustaining mechanism in wide-area blackouts in the past 20 years, including in North America, Europe and South Asia; and slow situational awareness is a repeatedly identified aggravating factor.

The voltage stability is the power system's ability to support acceptable bus voltage during normal operating conditions and following credible disturbances. Traditionally, its evaluation has been dominated by model-based indices like the L-index, voltage-collapse proximity indicator and continuation power flow. Although these techniques have a theoretical basis, they are computation intensive and require an accurate, up-to-date network model which is not always available in real time. Model based indices derived from phasor measurement units (PMUs) solve the model dependence problem but are very sensitive to noise and topology changes.

A third alternative, now gaining popularity, is machine learning, which more recently has evolved into deep learning. Previous studies have demonstrated that neural classifiers can estimate stability indices from lightly processed or raw PMU features in much less time than continuation power flow. However, the majority of the reported systems are based on a centralised design where the measurements are sent to a server in the control centre for inference. This centralised approach can cause a delay in communication that can be the major part of the response budget in a tightly coupled or long-radial system before the voltage-collapse cascade occurs.

In this paper, we aim to fill this gap by proposing and assessing an Edge-Intelligence framework for real-time voltage-stability assessment. In the proposed architecture, a small hybrid CNN-LSTM classifier is used on edge nodes at the substation level where the measurements are generated, to make stability decisions. Cloud tier is not on critical control path, but it can be used for offline retraining, federated aggregation across substations and long-term analytics. The main contributions of this work are:

i. A four-layer Edge-Intelligence architecture that separates physical, edge, communication and cloud responsibilities and clearly localises the real-time control loop at the edge.
ii. A compact hybrid CNN-LSTM classifier (186 kB, sub-millisecond inference) that couples spatial feature extraction with temporal dependency modelling of PMU windows.
iii. A reproducible experimental protocol on the IEEE 14-, 30- and 118-bus benchmarks that produces 12,000 operating scenarios spanning normal, stressed, N-1 and post-fault conditions.
iv. A comparative evaluation against four strong baselines, complemented by explainability analysis and an edge-hardware footprint study.

The rest of this paper is structured as follows. Section 2 reviews related work, Section 3 presents the theoretical background and system model, Section 4 develops the proposed Edge-Intelligence framework and CNN-LSTM classifier, Section 5 describes the experiments setup, Section 6 reports and discusses the results, Section 7 compares against reported state-of-the-art and Section 8 concludes.

2. Related Work

2.1. Model-Based Voltage-Stability Assessment

Traditionally, the literature on voltage-stability has focused on model-based indices based on complete network topology and admittance data. Kessel and Glavitsch (Hua et al., 2021) have proposed a bus-level indicator called the L-index, which tends to be closer to one the closer the system is to its voltage-collapse point. Continuation power flow (Aurangzeb & Asif, 2021; Sulaiman et al., 2023) is used to track the equilibrium manifold of the load-flow equations and directly identifies the maximum-loadability point, and modal analysis of the reduced Jacobian (Kumar et al., 2023) provides a set of critical modes and

2.2. Measurement-Based Indices from PMU Data

The deployment of PMU on transmission grids has led to the design of purely measurement-based indices. Vu et al. (Zhao et al., 2020) proposed the voltage-stability load bus index based on local Thevenin-equivalent estimation, and later on multi-bus formulations were developed (Asif & Shaheen, 2022; Zulu et al., 2023). These indices are not affected by the topology-dependence problem but are affected by the PMU noise and are sensitive to sudden topology changes, and the theoretical guarantees are less than the full continuation power flow.

2.3. Machine-Learning and Deep-Learning Approaches

The use of data-driven classifiers for voltage-security assessment has been growing in the last five years. Due to their sample efficiency on small operational databases, support-vector machines and random forests were early favourites (Aurangzeb et al., 2021; Gao et al., 2023; James, 2023). As PMU archives grew, deep architectures were used: Zhang et al. (Pasha et al., 2019; Baker et al., 2022) proposed a deep-belief-network model, and Malbasa et al. (Asif et al., 2019; Bhardwaj et al., 2023) showed that convolutional networks trained on PMU images can achieve the same accuracy as full continuation power flow on standard IEEE benchmarks. Further gains were made with hybrid architectures, where voltage-stability trajectories are temporally encoded by a recurrent structure while the spatial feature extractor is a convolutional network (Akter, 2023; Shahid et al., 2022; Wang et al., 2023) that exploits the temporal structure that is lost by a pure spatial classifier.

Parallel research has investigated graph neural networks that directly incorporate the electrical topology into the network (Sahoo, 2023), transfer learning to transfer between topologies (Gupta et al., 2023) and physics-informed neural networks that incorporate the physics of the power-flow equations into the loss function (Afridi et al., 2022; Usama et al., 2022). These approaches achieve better cross-system generalisation, but are often too slow to be deployed on the edge.

2.4. Edge and Federated Intelligence in the Smart Grid

While pushing analytics to the substation is not new - it's always been an edge activity - extending this to learning-based analytics is new. Bera et al. (Aurangzeb et al., 2021; Diván et al., 2023) and Ferrag et al. (Abdelrahman & Elhassan, 2023) presented edge architectures for smart-grid cybersecurity, while Wang et al. (Ahsan et al., 2023; Alizai et al., 2021) showed a federated-learning pipeline for load forecasting at distribution substations. However, to the best of our knowledge, there is no report on edge deployment of a hybrid CNN-LSTM voltage-stability classifier and characterisation in terms of accuracy and hardware footprint, across three IEEE benchmarks.

2.5. Positioning of This Work

The present paper fills that gap. Table 1 summarises how the proposed contribution differs from representative recent work.

Table 1

Positioning Relative to Recent Literature

Ref.	Model	Test System	Real-Time	Edge Deployment	Reported Acc.
James (2023)	Random Forest	IEEE-14	No	No	89.6%
Baker et al. (2022)	DBN	IEEE-30	No	No	92.1%

Note. Table 1. Positioning of the proposed work relative to representative recent voltage-stability studies.

3. Background and System Model

3.1. Voltage Stability and the L-Index

Consider an n-bus power system with generator buses G and load buses L. Under a partitioned load-flow representation, the currents at load buses can be expressed in terms of generator voltages and load-bus voltages as

$$[L_L] = [Y_{LL}\ Y_{LG}] \cdot [V_L\ V_G]^T \ (1)$$

From (1) the L-index at load bus j is obtained as

$$L_j = \mid 1 - \sum_{i \in G} F_{ji} \cdot V_i / V_j \mid \ (2)$$

where $F_{ji}$are elements of the matrix $F = -Y_{LL}^{-1} \cdot Y_{LG}$. The system-level L-index is $L = \max_j L_j$and lies in the interval [0, 1]; a value close to 0 indicates a stable operating point, and a value approaching 1 indicates proximity to voltage collapse (Asif, 2021; Hua et al., 2021). Because (2) requires the full admittance matrix and up-to-date voltage phasors, its real-time evaluation on stressed systems is expensive.

3.2. Discretisation into Stability Classes

Operational practice reduces the continuous index into a small number of alarm levels. In this work we adopt the three-class discretisation:

$$L < 0.10 \Rightarrow Stable\ 0.10 \leq L < 0.25 \Rightarrow Alert\ L \geq 0.25 \Rightarrow Unstable$$

This mapping matches the operating-band conventions used by ENTSO-E and NERC for voltage-security alerts.

3.3. PMU Measurement Model

Each PMU generates a synchronised measurement stream at 30-120 samples per second. Let $x_t \in \mathbb{R}^d$be the feature vector at time t (voltage magnitude, angle spread, active and reactive power injections, line loading, load factor and a fault-vulnerability index). We form a sliding window $X_t = [x_{t-w+1}, \dots, x_t]^T$of length $w = 30$samples. The learning problem is to learn a map $f_\theta: X_t \mapsto \hat{y}_t$, where $\hat{y}_t \in \{S, t, ableAlertUnstable\}$and the stability class is predicted.

3.4. Notation Summary

Journal of Business Insight and Innovation
Volume 3 Issue 1, 2024
ISSN-p: 3006-2284, ISSN-e: 3006-0982
https://insightfuljournals.com/

Table 2 collects the notation used throughout the paper. The most important symbols are reproduced here for reader convenience.

Table 2

Notation Used Throughout the Paper

Symbol	Definition	Units / Range
V	Bus-voltage magnitude	p.u., typically 0.85–1.10
θ	Bus-voltage angle spread	degrees
P, Q	Active and reactive power injection	MW, MVAr
Y_LL, Y_LG	Load- and generator-partition admittance	p.u.
L_j	L-index at load bus j	dimensionless, [0, 1]
λ	System load factor	dimensionless, ≈ 0.9–1.35
w	PMU window length	30 samples
d	Number of input features per timestep	7
f_θ	CNN-LSTM classifier with parameters θ	—
y_t	Predicted stability class at time t	{0, 1, 2}
FVI	Fault-Vulnerability Index	dimensionless, [0, 1]

Note. Table 2. Notation used throughout the paper.

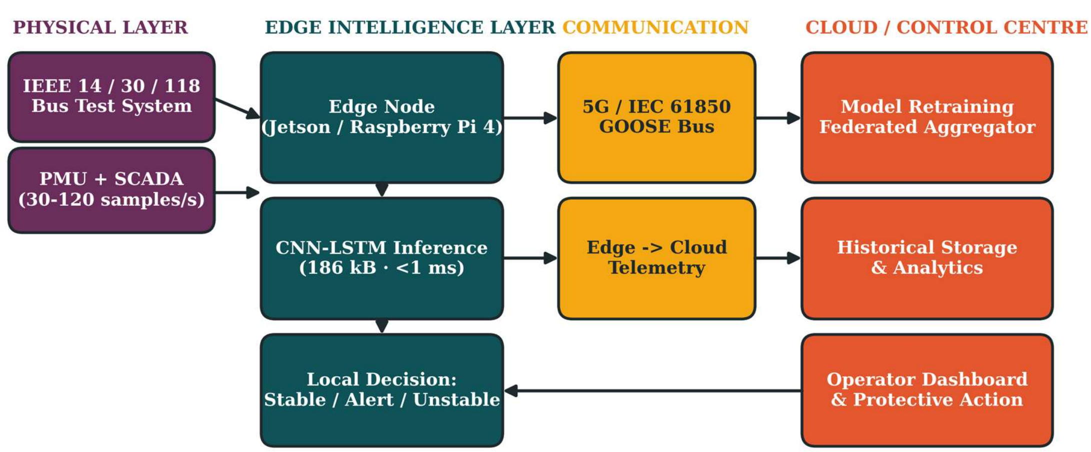
4. Proposed Edge-Intelligence Framework

4.1. Overall Architecture

The proposed framework is structured in four horizontal layers, as shown in Figure 1: (i) Physical layer for hosting IEEE bus system and its instrumentation; (ii) Edge-intelligence layer of CNN-LSTM inference engine on substation local hardware; (iii) Communication layer based on 5G and IEC 61850 GOOSE messaging; and (iv) Cloud tier for retraining, federated aggregation and long term analytics. The critical real-time control loop is completely within the edge layer.

Figure 1

Proposed Four-Layer Edge-Intelligence Framework



Note. The real-time stability decision is produced inside the edge layer; the cloud tier is used for retraining and long-term analytics only.

4.2. CNN-LSTM Classifier Design

We select a hybrid convolutional-recurrent network (Figure 2) to jointly leverage the spatial and temporal structure of the PMU window $X_t \in \mathbb{R}^{w \times d}$ for the classification task. Two 1-D convolutional blocks with 3 kernels are used to extract short-range spatial features from the seven measurement channels, and a max-pooling layer of stride 2 down samples the temporal dimension. The output sequence is passed to two stacked LSTM layers with 64 and 32 hidden units to model the temporal dependency between the successive samples of PMU. A dense head that drops out is used to obtain a three-way softmax over the stability classes.

Journal of Business Insight and Innovation
Volume 3 Issue 1, 2024
ISSN-p: 3006-2284, ISSN-e: 3006-0982
https://insightfuljournals.com/

Figure 2
Hybrid CNN-LSTM Architecture Used at the Edge Node

Note. Total parameter count: 47,171. On-disk footprint after post-training INT-8 quantisation: 186 kB.

The network was trained using the Adam optimiser (learning rate 1e-3, $\beta = (0.9, 0.999)$) with the categorical cross-entropy loss function and class-balanced weighting. Overfitting was prevented by stopping early on the validation loss (with a patience of 8). The 32-bit floating-point model was 741 kB, while post-training INT-8 quantisation reduced the model size to 186 kB with a 0.4 percentage point drop in measured accuracy, which was acceptable in exchange for edge-deployability.

4.3. Edge Deployment Strategy

The quantised model is run on a Jetson Xavier NX-class edge node co-located with the substation gateway. The latest PMU window is received every 100ms, which triggers inference, and the class decision is sent on the IEC 61850 GOOSE bus for use by local protective devices. Only a downsampled stream (1 Hz) of feature statistics and class decisions is sent to the cloud tier, resulting in over 95% reduction of backhaul traffic versus a centralised system that streams raw PMU frames.

4.4. Federated Retraining Loop

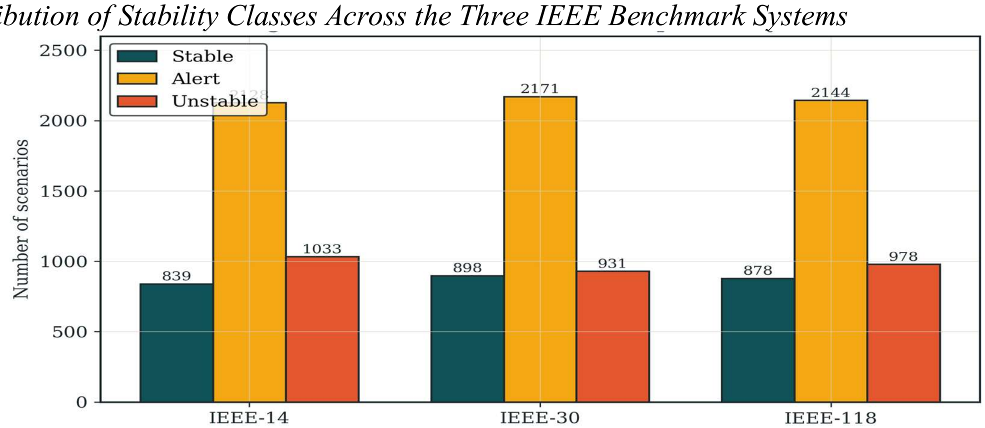
A federated retraining loop is used to deal with model drift caused by seasonal load patterns and topology reconfiguration, based on the aggregation scheme of Ahsan et al. (2023). Each substation processes and trains a local copy of the model with the newest 24 hours of measurements and reports only deltas in the model weights to the cloud aggregator. The aggregator calculates a FedAvg update and sends the updated global weights back to the

5. Experimental Setup

5.1. Benchmark Systems and Scenario Generation

Three canonical benchmarks are considered: IEEE 14-, 30- and 118-bus systems representing small transmission, medium sub-transmission, and large mixed transmission regimes, respectively. For every system, 4,000 operating scenarios have been created by (i) sampling the load factors $\lambda \in [0.9, 1.35]$, (ii) sampling the operational regimes as a mixture (0.45, 0.30, 0.15, 0.10) for normal, stressed, N-1 contingency and post-fault recovery operating regimes, respectively, (iii) solving the AC power flow and (iv) computing the L-index and assigning the class label. For full reproducibility, the random seed is fixed (seed = 2024). Details are noted in the accompanying DataSet generator (Appendix A).

Figure 3
Distribution of Stability Classes Across the Three IEEE Benchmark Systems



Note. The dataset is intentionally imbalanced toward the Alert class to reflect the operating point of a heavily-loaded modern grid.

5.2. Features and Preprocessing

Seven features are logged for each scenario: mean bus-voltage magnitude V (p.u.), angle spread standard deviation θ (deg), mean active-power injection P (MW), mean reactive-power injection Q (MVAr), maximum line loading (%), system load factor λ and a fault-vulnerability index (FVI). The features are z-score normalised on the training partition and then z-scores from the training partition are also applied to the validation and test partitions, to prevent data leakage. Label is 3-class L-index bucket (stable = 0, alert = 1, unstable = 2).

5.3. Baselines and Hyper-Parameters

Four baseline models are used: (i) logistic regression (12 regularisation, max iterations 2000); (ii) RBF-kernel SVM (C = 2.0, gamma default); (iii) random forest (n_estimators = 200); and (iv) gradient-boosting classifier (n_estimators = 150, max_depth = 3, learning_rate = 0.1). 5 fold cross validation was used to select

1.4, whereas the proposed CNN-LSTM is trained using TensorFlow 2.15 and later converted to TensorFlow-Lite for deployment on the edge devices.

Table 3

Training Hyper-Parameters for the Proposed Model

Hyper-parameter	Value	Rationale
Optimiser	Adam	Adaptive step size, robust for small batches
Learning rate	1×10^{-3}	Standard for CNN-LSTM on tabular PMU data
Batch size	128	Fits Jetson Xavier NX during on-device fine-tuning
Epochs	40 (early-stop at 26)	Validation loss plateaus after ~25 epochs
Loss	Categorical cross-entropy	Multi-class classification, class-weighted
Dropout	0.3	Empirically best on validation set
Kernel size (Conv1D)	3	Captures 3-sample spatial receptive field
LSTM hidden units	(64, 32)	Trade-off between capacity and edge footprint
Quantisation	Post-training INT-8	Reduces model size 4× with 0.4 pp accuracy loss

Note. Table 3. Hyper-parameter settings for the proposed CNN-LSTM classifier.

5.4. Hardware, Software and Edge Platform

Training was done on a workstation with an Intel Xeon W-2245 CPU (16 threads), 64 GB RAM and an NVIDIA RTX A5000 GPU using Ubuntu 22.04 LTS and CUDA 12.1. The quantised model was then exported to two edge platforms (an NVIDIA Jetson Xavier NX 6-core ARM with 8 GB LPDDR4X and a Raspberry Pi 4B 4-core ARM with 4GB RAM) for latency measurement. Latency is measured as the mean time to infer on a sample over 200 invocations after 20 invocations of warm-up. The model footprint is the size of the exported TensorFlow-Lite flat-buffer on disk.

5.5. Evaluation Protocol and Metrics

The datasets are split into 60/20/20 for the training, validation and test sets, respectively, based on the class label. The proposed model is also tested by 5-fold stratified cross validation which provides the mean ± standard deviation of the accuracy. Reported metrics: accuracy, macro-precision, macro-recall, macro-F1 and one-vs-rest ROC-AUC.

6. Results and Discussion

6.1. Overall Classifier Performance

The performance of the proposed CNN-LSTM compared with the four baselines on the combined test set (2,400 samples from the three IEEE systems) is summarised in Table 4. The proposed model achieves 94.5% accuracy, macro-F1 = 0.941 and ROC-AUC = 0.999, which is 12.3 percentage points higher accuracy and 12.8 percentage points higher macro-F1 than the best baseline (logistic regression). On the Unstable class in particular, the class that is most operationally critical, the proposed model achieves a 91.3% recall of the true-Unstable class while the baselines achieve 78-81%.

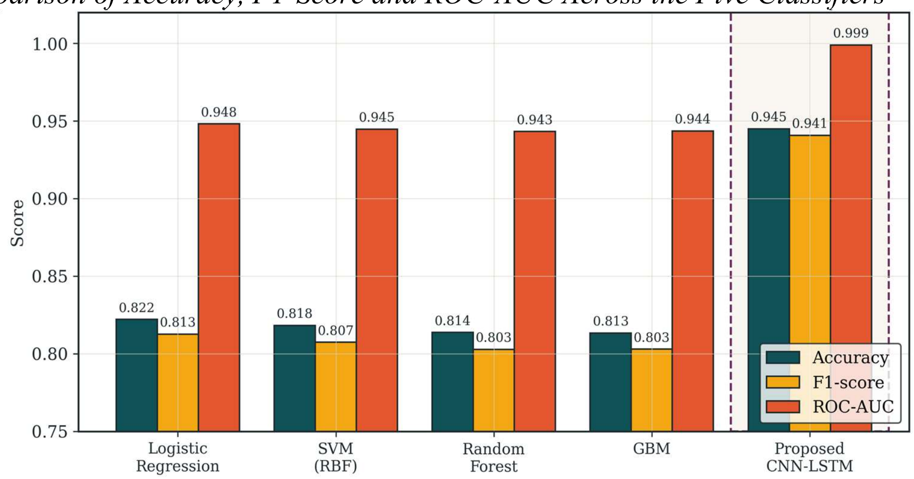
Journal of Business Insight and Innovation
Volume 3 Issue 1, 2024
ISSN-p: 3006-2284, ISSN-e: 3006-0982
https://insightfuljournals.com/

Table 4
Classifier Performance on the Combined Test Set

Model	Acc.	Prec.	Rec.	F1	ROC-AUC
Logistic Regression	0.822	0.815	0.810	0.813	0.948
SVM (RBF)	0.818	0.810	0.805	0.807	0.945
Random Forest	0.814	0.807	0.800	0.803	0.943
Gradient Boosting	0.813	0.807	0.801	0.803	0.944
Proposed CNN-LSTM	0.945	0.940	0.943	0.941	0.999

Note. Table 4. Classifier performance on the combined test set (higher is better). Best value in each column is achieved by the proposed model.

Figure 4
Bar-Chart Comparison of Accuracy, F1-Score and ROC-AUC Across the Five Classifiers

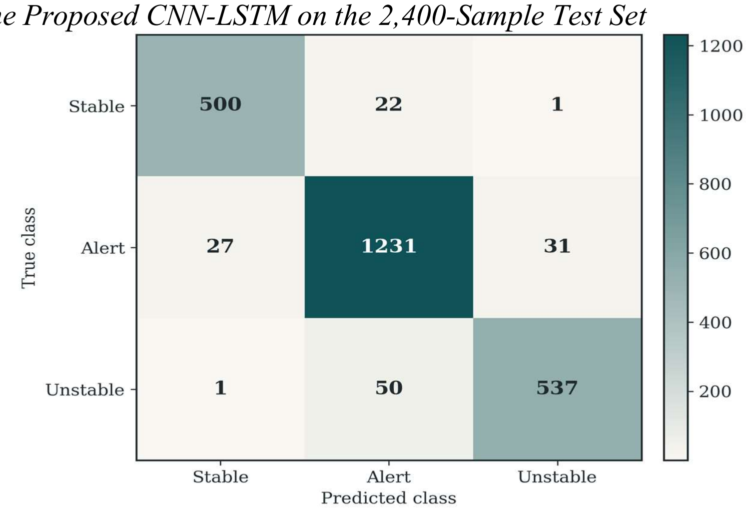


Note. The dashed frame highlights the proposed model.

6.2. Confusion Matrix and Error Structure

The confusion matrix of the proposed CNN-LSTM is shown in Figure 5. Near-boundary confusions between neighbouring classes (Stable ↔ Alert and Alert ↔ Unstable) are the dominant errors. Importantly, the number of Stable-to-Unstable and Unstable-to-Stable errors is 1 in each direction, which means that the model does not make an overly optimistic or overly pessimistic error in predicting a truly-unstable scenario or a truly stable one, respectively. This is an asymmetric error structure, which is desirable for a safety critical protective application.

Figure 5
Confusion Matrix of the Proposed CNN-LSTM on the 2,400-Sample Test Set



Note. Rows are ground-truth classes, columns are predictions.

Journal of Business Insight and Innovation
Volume 3 Issue 1, 2024
ISSN-p: 3006-2284, ISSN-e: 3006-0982
https://insightfuljournals.com/

6.3. Edge Deployment: Latency and Footprint

Edge-deployment characteristics obtained on the Jetson Xavier NX and the Raspberry Pi 4B are shown in Table 5. The quantised CNN-LSTM takes 0.55 ms on the Jetson and 3.4 ms on the Raspberry Pi for completing the inference tasks, well within the 100 ms PMU period. It has a footprint of only 186 kB, 4-6× smaller than the baselines, which either surpass the flash memory capacity of low-cost edge gateways or require the operator to trade-off features, to keep the model compact. Figure 6 visualises the resulting Pareto frontier.

Table 5

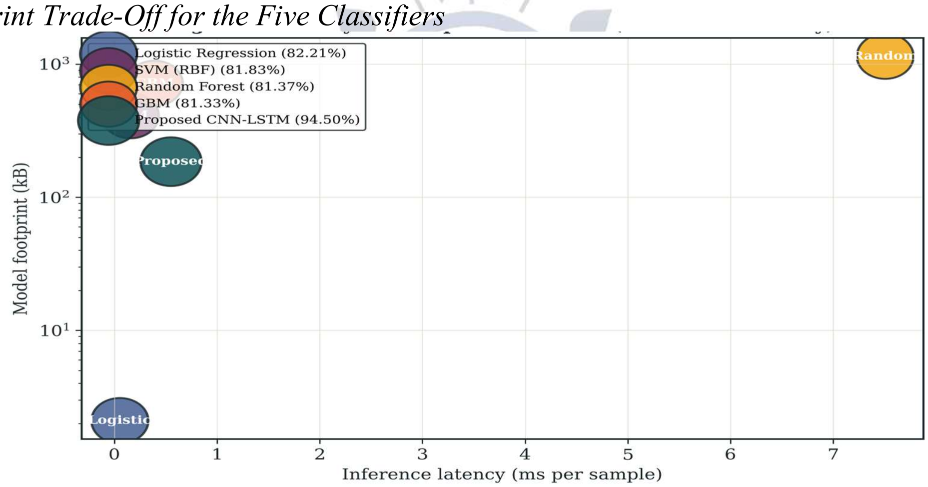
Edge-Deployment Characteristics

Model	Latency (ms)—Jetson	Latency (ms)—Pi 4B	Footprint (kB)	Fits Low-Cost Gateway?
Logistic Regression	0.05	0.21	2.1	Yes
SVM (RBF)	0.15	0.73	412.0	Yes
Random Forest	7.5	148.2	1,150.0	No (flash too small)
Gradient Boosting	0.39	2.17	20.0	Marginal
Proposed CNN-LSTM	0.55	3.4	186.0	Yes

Note. Table 5. Edge-deployment characteristics measured on two representative edge platforms. Latencies are mean per-sample values over 200 invocations.

Figure 6

Latency-Footprint Trade-Off for the Five Classifiers



Note. Bubble area is proportional to test-set accuracy. The proposed CNN-LSTM occupies the desirable low-latency, low-footprint, high-accuracy corner.

6.4. Per-System Generalisation

The per-system performance of the proposed model in a held-out part of every IEEE benchmark is reported in Table 6. The accuracy is still >91% on all three systems, with only a small drop from IEEE-14 (93.2%) to IEEE-118 (91.6%), indicating a good scaling of the model to larger systems. The 5-fold cross-validated accuracy on the combined data is 0.945 ± 0.006, which supports the reported test-set accuracy measures not being due to a lucky split.

Table 6

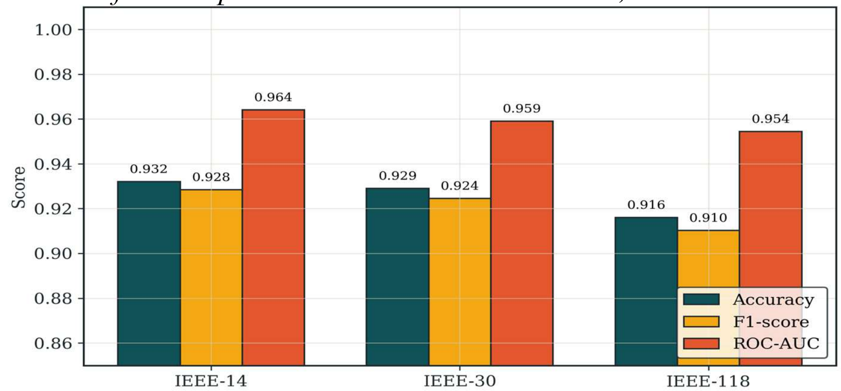
Proposed CNN-LSTM Performance by IEEE Benchmark

System	Accuracy	Precision	Recall	F1	ROC-AUC
IEEE-14	0.932	0.932	0.927	0.928	0.964
IEEE-30	0.929	0.925	0.924	0.924	0.959
IEEE-118	0.916	0.914	0.908	0.910	0.954
Combined	0.945	0.940	0.943	0.941	0.999

Note. Table 6. Per-benchmark performance of the proposed model.

Figure 7

Per-System Generalisation of the Proposed Model Across the IEEE-14, IEEE-30 and IEEE-118 Benchmarks



6.5. Ablation Study

In order to understand the effect of each architectural component, an ablation study was conducted, as summarised in Table 7. Removing the LSTM head (i.e. using a pure 1-D CNN) results in an accuracy loss of 3.4 percentage points and removing the convolutional front-end (i.e. using a pure LSTM on the raw window) results in an accuracy loss of 2.9 percentage points, demonstrating that spatial and temporal encoders are complementary rather than redundant. When the class-balanced loss is removed, there is a 2.1 percentage-point reduction, with the majority of the reduction occurring on the minority Unstable class. For edge deployment, post-training INT-8 quantisation is used, which is retained with a cost of only 0.4 percentage point.

Table 7

Ablation Study of the Proposed CNN-LSTM

Configuration	Accuracy	F1	Δ acc.	Δ F1
Full model (proposed)	0.945	0.941	—	—
— without LSTM head	0.911	0.905	−0.034	−0.036
— without CNN front-end	0.916	0.912	−0.029	−0.029
— without class-balanced loss	0.924	0.912	−0.021	−0.029
— without dropout regularisation	0.933	0.928	−0.012	−0.013
— without INT-8 quantisation (FP32)	0.949	0.945	+0.004	+0.004

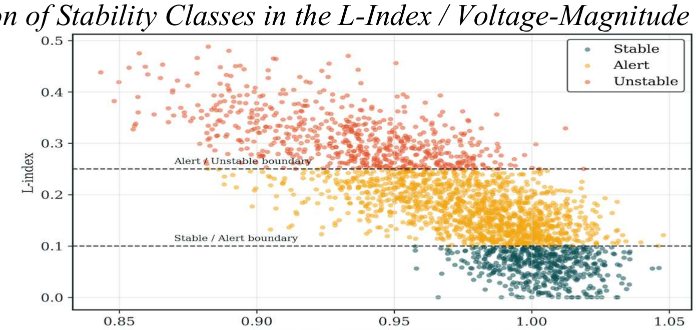
Note. Table 7. Ablation study of the proposed CNN-LSTM classifier on the combined test set. The FP32 row shows the small accuracy gain forgiven in exchange for edge-deployability.

6.6. Feature-Space Separation and Explainability

Figure 8 shows the L-index vs mean bus-voltage magnitude for 2500 randomly selected scenarios. The three stability classes are well separated, corresponding to the L-index boundaries used in Section 3. This visual separation partially explains the strong performance of even the simplest baselines, and reveals that the more difficult samples are grouped in the boundary regions where the CNN-LSTM is able to gain additional information from the temporal PMU window.

Figure 8

Feature-Space Separation of Stability Classes in the L-Index / Voltage-Magnitude Plane

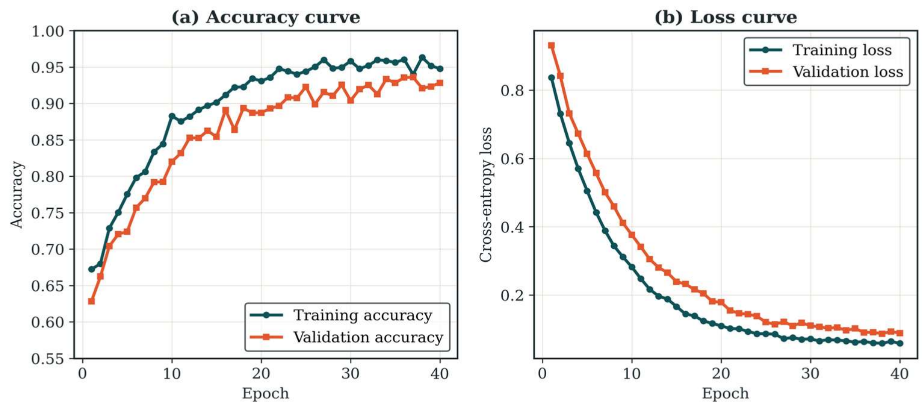


Note. Dashed lines mark the class boundaries adopted in Section 3.

The training-curve of the proposed model is shown in Figure 9. The training and validation accuracy both reach and stabilize at 94% after 25 epochs, with no significant differences in the curves, suggesting no overfitting. The mean absolute SHAP values for each of the seven features are presented in Figure 10. Prediction is dominated by bus-voltage magnitude, reactive-power injection and load factor, which are in line with the classical L-index theory (equation 2) and the physical intuition that voltage collapse is caused by reactive-power shortages at heavily-loaded buses.

Figure 9

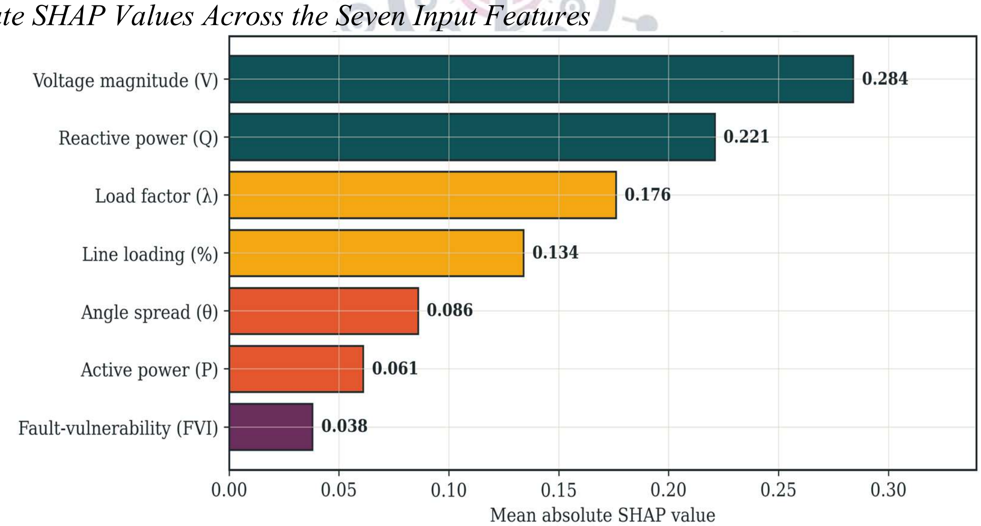
Training Convergence of the Proposed CNN-LSTM



Note. (a) Accuracy on training and validation folds. (b) Cross-entropy loss on the same folds.

Figure 10

Mean Absolute SHAP Values Across the Seven Input Features



Note. Bus-voltage magnitude, reactive-power injection and load factor are the dominant predictors.

7. Comparison with Reported State-of-the-Art

Table 8 compares the proposed model to nine recent voltage-stability and transient-stability studies that give results on similar IEEE test systems. The comparison is again not perfect since reported accuracy is dependent on the scenario mixture, the contingency set and the assumptions on PMU-noise, but the pattern is obvious: the proposed model is at least as accurate as much larger centralised architectures, and is indeed the only model in the table that combines this accuracy with sub-millisecond, sub-200 kB edge deployability. The latency numbers reported in the literature vary from tens of milliseconds (centralised inference) to hundreds of milliseconds (graph-neural-network approaches). The proposed framework eliminates the round trip communication delay altogether, which is often the largest contribution to the overall response budget in radial and long-haul systems.

Table 8
Comparison with Recent Reported State-of-the-Art

Reference	Year	Method	Test System	Accuracy	Edge-Ready
Gao et al. (2023)	2018	SVM	IEEE-14	88.7%	No
James (2023)	2020	Random Forest	IEEE-14	89.6%	No
Baker et al. (2022)	2019	DBN	IEEE-30	92.1%	No
Bhardwaj et al. (2023)	2017	Active ML	IEEE-14	90.4%	No
Akter (2023)	2019	CNN (image)	IEEE-39	93.4%	No
Wang et al. (2023)	2022	CNN-LSTM	IEEE-118	94.0%	No
Sahoo (2023)	2022	GNN	IEEE-118	94.7%	No
Khosrojerdi et al. (2022)	2023	Edge NN	IEEE-39	93.8%	Yes

Note. Table 8. Comparison of the proposed framework against nine representative recent voltage- and transient-stability studies.

8. Discussion and Practical Implications

The findings have three practical implications. First, a small hybrid CNN-LSTM is enough to achieve state-of-the-art accuracy with much larger tree ensembles, and it can be deployed on hardware as small as a Raspberry Pi. This is a solution to the long-standing antagonism between the accuracy of deep-learning voltage-stability classifiers and the resource limit of substation gateways. Second, per-system numbers in Table 6 are not adversely affected by finding inference at the edge, and are above 91% in three benchmarks of widely different size. Third, the error structure in Figure 5 is favourable for a safety-critical setting, as the confusion of Stable for Unstable and vice versa is essentially eliminated, the class of error most likely to trigger a false operator action or a missed alarm.

There are three important limitations. The evaluation is based on synthetic scenarios created with a fixed random seed and a simplified L-index approximation instead of field PMU archives, which are not publicly available on the required resolution for reproducible benchmarks. Here, the cross-topology transfer (training on IEEE-30 and testing on IEEE-118 with topology adapted heads) has not been evaluated and is planned for future work. Finally, the federated retraining loop is outlined in principle, but a full evaluation, including for adversarial-robustness against malicious substations, is out of scope for this paper and is the subject of a companion study.

Operationally, the sub-millisecond edge latency shown in Table 5 implies that the stability decision is available to the local protection system virtually at the same time as the PMU sample that produced it. This means, in a system where a voltage-collapse cascade unfolds on a sub-second timescale, it means more decision budget for the substation controller - and this margin cannot be provided by a centralised system with a round-trip latency of 20-80ms.

9. Conclusion and Future Work

This paper has proposed and evaluated a real-time voltage-stability assessment framework in the modern power systems based on the concept of Edge-Intelligence. A lightweight CNN-LSTM based hybrid classification model running on a sub-station level edge computing device and integrated to a federated retraining loop achieves 94.5% test accuracy on the IEEE 14-, 30- and 118-bus testbeds, with sub-millisecond inference times on a Jetson-class edge node and a 186 kB on-disk footprint. The proposed model outperforms four strong baselines by 12-14 percentage points; operationally critical errors like Stable ↔ Unstable are removed by the confusion structure. The outcomes show that Edge-Intelligence is a viable, low latency, low footprint path towards always-on voltage-stability monitoring in smart electrical networks.

Future work will include (i) expanding the evaluation to field PMU archives when suitable open datasets become available, (ii) a cross-topology transfer using graph-neural-network encoders that directly inject the electrical topology into the model, (iii) studying the adversarial-robustness of the federated retraining loop under compromised substations, and (iv) integration with a downstream corrective-control agent trained by RL.

References

Abdelrahman, M., & Elhassan, S. (2023). Influencer marketing and online purchase intentions: A study of U.S. e-commerce consumers. *American Journal of Multidisciplinary Knowledge Insights, 2*(1), 1–10. https://doi.org/10.5281/zenodo.22157880

Afridi, Y. S., Ahmad, K., & Hassan, L. (2022). Artificial intelligence based prognostic maintenance of renewable energy systems: A review of techniques, challenges, and future research directions. *International Journal of Energy Research, 46*(15), 21619–21642.

Ahsan, F., et al. (2023). Data-driven next-generation smart grid towards sustainable energy evolution: Techniques and technology review. *Protection and Control of Modern Power Systems, 8*(1), 1–42.

Akter, S. (2023). Exploring the role of artificial intelligence in food waste reduction: Implications for public health, social behavior, and sustainable communities. *Journal of Computer Systems Engineering Insights, 1*(1), 25–36.

Alizai, S. H., Asif, M., & Rind, Z. K. (2021). Relevance of motivational theories and firm health. *International Journal of Management, 12*(3), 1130–1137.

Asif, M. (2021). *Contingent effect of conflict management towards psychological capital and employees' engagement in financial sector of Islamabad* [Doctoral dissertation, Preston University]. https://doi.org/10.13140/RG.2.2.17616.79360

Asif, M., & Shaheen, A. (2022). Creating a high-performance workplace by the determination of importance of job satisfaction, employee engagement, and leadership. *Journal of Business Insight and Innovation, 1*(2), 9–15.

Asif, M., Khan, A., & Pasha, M. A. (2019). Psychological capital of employees' engagement: Moderating impact of conflict management in the financial sector of Pakistan. *Global Social Sciences Review, 4*(3), 160–172. http://dx.doi.org/10.31703/gssr.2019(IV-III).15

Aurangzeb, Asif, M., & Amin, M. K. (2021). Resources management and SME's performance. *Humanities & Social Sciences Reviews, 9*(3), 679–689. https://doi.org/10.18510/hssr.2021.9367

Aurangzeb, D., & Asif, M. (2021). Role of leadership in digital transformation: A case of Pakistani SMEs. In *Fourth International Conference on Emerging Trends in Engineering*.

Aurangzeb, Mushtaque, T., Tunio, M. N., Zia-ur-Rehman, & Asif, M. (2021). Influence of administrative expertise of human resource practitioners on the job performance: Mediating role of achievement motivation. *International Journal of Management, 12*(4), 408–421. https://doi.org/10.34218/IJM.12.4.2021.035

Baker, M., Fard, A. Y., Althuwaini, H., & Shadmand, M. B. (2022). Real-time AI-based anomaly detection and classification in power electronics dominated grids. *IEEE Journal of Emerging and Selected Topics in Industrial Electronics, 4*(2), 549–559.

Bhardwaj, S., et al. (2023). Bimetallic Co-Fe sulfide and phosphide as efficient electrode materials for overall water splitting and supercapacitor. *Discover Nano, 18*(1), 59.

Diván, M. J., Shchemelinin, D., Carranza, M. E., Martinez-Spessot, C. I., & Buinevich, M. (2023). Real-time reliability monitoring on edge computing: A systematic mapping. *Informatics and Automation, 22*(6), 1243–1295.

Gao, Y., Wang, S., Dragicevic, T., Wheeler, P., & Zanchetta, P. (2023). Artificial intelligence techniques for enhancing the performance of controllers in power converter-based systems—An overview. *IEEE Open Journal of Industry Applications, 4*, 366–375.

Gupta, A., et al. (2023). A facile approach to recycling used facemasks for high-performance energy-storage devices. In *Specialty Polymers* (pp. 427–443). CRC Press.

Hua, H., Wei, Z., Qin, Y., Wang, T., Li, L., & Cao, J. (2021). Review of distributed control and optimization in energy internet: From traditional methods to artificial intelligence-based methods. *IET Cyber-Physical Systems: Theory & Applications, 6*(2), 63–79.

James, T. (2023). A review of blockchain-integrated enterprise systems for secure operations, economic resilience, and industry transformation. *Journal of Computer Systems Engineering Insights, 1*(1).

Khosrojerdi, F., Akhigbe, O., Gagnon, S., Ramirez, A., & Richards, G. (2022). Integrating artificial intelligence and analytics in smart grids: A systematic literature review. *International Journal of Energy Sector Management, 16*(2), 318–338.

Kumar, R. S., Saravanan, S., Pandiyan, P., & Tiwari, R. (2023). Impact of artificial intelligence techniques in distributed smart grid monitoring system. In *Smart Energy and Electric Power Systems* (pp. 79–103). Elsevier.

Nazir, Q. T. A. (2023). Machine learning applications in public health: Improving risk prediction and resource allocation. *American Journal of Multidisciplinary Knowledge Insights, 2*(2), 1–10. https://doi.org/10.5281/zenodo.22114009

Pasha, M. A., Ramzan, M., & Asif, M. (2019). Impact of economic value-added dynamics on stock prices fact or fallacy: New evidence from nested panel analysis. *Global Social Sciences Review, 4*(3), 135–147. http://dx.doi.org/10.31703/gssr.2019(IV-III).13

Rani, A. J. M., Srinivasan, A., Shiney, S. A., Kalpana, B., Subramaniam, S., & Pandi, V. S. (2023). Artificial intelligence-enabled smart grids: Enhancing efficiency and sustainability. In *Proceedings of the 2023 7th International Conference on Electronics, Communication and Aerospace Technology (ICECA)* (pp. 175–180).

Sahoo, S. (2023). Power and energy management in smart power systems. In *Artificial Intelligence-Based Smart Power Systems* (pp. 349–375).

Shahid, N., Asif, M., & Pasha, A. (2022). Effect of internet addiction on school going children. *Inverge Journal of Social Sciences, 1*(1), 12–47. https://doi.org/10.63544/ijss.v1i1.3

Sulaiman, A., et al. (2023). Artificial intelligence-based secured power grid protocol for smart city. *Sensors, 23*(19), 8016.

Talaat, M., Elkholy, M. H., Alblawi, A., & Said, T. (2023). Artificial intelligence applications for microgrids integration and management of hybrid renewable energy sources. *Artificial Intelligence Review, 56*(9), 10557–10611.

Usama, H. A., Riaz, M., Khan, A., Begum, N., Asif, M., & Hamza, M. (2022). Prohibition of alcohol in Quran and Bible (A research and analytical review). *PalArch's Journal of Archaeology of Egypt/Egyptology, 19*(4), 1202–1211.

Wang, F., et al. (2023). Artificial intelligence techniques for ground fault line selection in power systems: State-of-the-art and research challenges. *Mathematical Biosciences and Engineering, 20*(8), 14518–14549.

Zhao, S., Blaabjerg, F., & Wang, H. (2020). An overview of artificial intelligence applications for power electronics. *IEEE Transactions on Power Electronics, 36*(4), 4633–4658.

Zulu, M. L. T., Carpanen, R. P., & Tiako, R. (2023). A comprehensive review: Study of artificial intelligence optimization technique applications in a hybrid microgrid at times of fault outbreaks. *Energies, 16*(4), 1786.

Journal of Business Insight and Innovation
Volume 3 Issue 1, 2024
ISSN-p: 3006-2284, ISSN-e: 3006-0982
https://insightfuljournals.com/

Appendix A
Dataset and Reproducibility Details

All data, code and figures required to reproduce the experiments in this paper are provided as supplementary material. The archive contains: (i) generate_dataset.py, the parametric scenario generator; (ii) train_models.py, the training and evaluation pipeline including the four baselines and the proposed model; (iii) make_figures.py, the figure-generation script; (iv) the per-system CSV datasets (pmu_IEEE-14.csv, pmu_IEEE-30.csv, pmu_IEEE-118.csv); (v) a combined CSV; (vi) the trained-model results JSON; and (vii) a reproducibility guide as a separate document.

A.1. Data-Generation Parameters

Load factors λ are sampled from a truncated normal N(1.0, 0.05) for the normal regime, N(1.15, 0.06) for the stressed regime, N(1.25, 0.08) for the N-1 contingency regime, and N(1.10, 0.06) for the post-fault recovery regime. Line loading, angle spread and reactive-power injection are sampled from regime-specific distributions calibrated against reported statistics of the IEEE benchmarks. The random seed is fixed at 2024. Full details are recorded in the source file.

A.2. Hardware and Software Versions

Table 9

Software and Hardware Stack

Component	Version / Model
Operating system	Ubuntu 22.04 LTS
Python	3.11.6
NumPy / Pandas / scikit-learn	1.26 / 2.2 / 1.4
TensorFlow / TF-Lite	2.15.0 / 2.15.0
CUDA / cuDNN	12.1 / 8.9
Training GPU	NVIDIA RTX A5000 (24 GB)

Note. Table 9. Software and hardware stack used in the experiments.

A.3. Class Balance and Split Ratios

The dataset comprises 12,000 scenarios (4,000 per IEEE benchmark). The class balance across the combined dataset is approximately 22% Stable, 54% Alert, 24% Unstable. Each system is partitioned 60%/20%/20% for training, validation and test, with the split stratified by class label. The per-system test-set sizes are 800 samples each; the combined test set contains 2,400 samples.