processes

MDPI

Article

Physics-Informed Closed-Loop Python–OpenDSS Co-Simulation for Fault Detection and Localization in Inverter-Dominated Microgrids

Zukisa Emerson Nante * and Zenghui Wang

Department of Electrical and Smart Systems Engineering, University of South Africa, Florida 1709, South Africa; wangz@unisa.ac.za
* Correspondence: 37229958@mylife.unisa.ac.za

Abstract

Increasing penetration of inverter-interfaced distributed energy resources (DERs) presents new challenges for fault diagnosis in distribution networks because DER operating conditions can modify the electrical signatures available to protection and diagnostic algorithms. This paper presents a physics-informed, closed-loop Python–OpenDSS co-simulation framework for machine-learning-based fault detection, fault-type classification, and fault localization. OpenDSS provides the physics-based network solution, while Python performs automated scenario orchestration, electrical-feature extraction, dataset construction, model development, and diagnostic evaluation. The framework was evaluated using the IEEE 13-node, IEEE 34-node, and IEEE 37-node benchmark feeders, with the IEEE 13-node feeder additionally serving as the detailed PV/BESS sensitivity and binary fault-detection case study. A scenario-level data-partitioning strategy was implemented to ensure separation of physical scenarios between model development and final held-out evaluation. Under the evaluated IEEE 13-node conditions, the binary fault detector achieved 100.00% held-out balanced accuracy and 100.00% macro-F1, although the healthy-state test population was limited. For fault-type classification, the supervised models achieved held-out accuracies ranging from 83.33% to 100.00% across the benchmark feeders, whereas the highest verified fault-localization accuracies were 66.67%, 83.70%, and 61.90% for the IEEE 13-node, IEEE 34-node, and IEEE 37-node feeders, respectively. Matched DER operating cases on the IEEE 13-node feeder further demonstrated measurable changes in network-level fault signatures, with mean absolute changes in maximum phase-current magnitude of 27.78 A for PV operation and 39.41 A for combined PV–BESS operation relative to the no-DER condition. The results demonstrate that diagnostic performance is task- and feeder-dependent and that DER operating conditions measurably influence the network-level fault signatures used by the diagnostic models. The conclusions are limited to the phasor-domain network behavior represented by OpenDSS and do not extend to detailed electromagnetic-transient inverter or converter-control dynamics.

Keywords: fault detection; fault classification; fault localization; OpenDSS; machine learning; distributed energy resources; photovoltaic systems; battery energy storage systems; co-simulation; distribution networks

Check for updates

Academic Editors: Bo Yang and Franco F. Yanine

Received: 22 August 2026
Revised: 18 September 2026
Accepted: 2 October 2026
Published: 9 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

The rapid proliferation of distributed energy resources (DERs), including photovoltaic (PV) systems, battery energy storage systems (BESS), and power electronic converters, has

fundamentally reshaped the operational characteristics of modern distribution networks [1]. Microgrids have emerged as a key architectural solution for integrating renewable energy, enhancing local resilience, and supporting both grid-connected and islanded operations [2]. However, increasing penetration of inverter-interfaced DERs introduces important challenges for conventional protection schemes [1,3–5]. Compared with synchronous-generator-based systems, inverter-interfaced resources exhibit different fault-current characteristics, while bidirectional power flows and changing network configurations further influence the electrical quantities used by conventional overcurrent- and impedance-based protection methods [1,3,4,6,7].

To address these challenges, researchers have explored machine-learning (ML) and artificial-intelligence (AI) techniques for fault detection, classification, and localization [8–15]. Supervised learning algorithms such as Random Forest (RF), Support Vector Machines (SVM), Decision Trees (DT), and Artificial Neural Networks (ANN) have demonstrated strong classification performance using voltage and current measurements obtained from simulated or historical datasets [8,9]. Deep-learning approaches can further extract nonlinear representations from electrical measurements, particularly when structured time-series data are available [10]. In parallel, graph-based learning methods have been proposed to exploit the inherent topological structure of distribution networks by representing buses and electrical components as interconnected graph elements [11–13,16–19], while signal-processing-based approaches using synchronized phasor measurements and time-frequency transformations have also been investigated [9,10,20]. Hybrid and ensemble techniques additionally combine multiple learning paradigms or optimization-based feature-selection procedures [8]. Fault identification and localization methods that explicitly account for inverter-interfaced distributed generation have also been investigated [21]. Despite these advances, several methodological challenges remain. Diagnostic performance may depend on the electrical feeder characteristics, DER operating conditions, measurement placement, fault resistance, and the specific diagnostic task [8,11,13,22–25]. The generalization of graph-based fault-diagnostic models across distribution networks with different topologies has also emerged as an important consideration [26]. Changes in active-network topology and distributed-generation operating conditions can further complicate fault-location performance [27]. Moreover, when multiple measurement records originate from the same simulated physical fault scenario, conventional record-level random partitioning can place closely related observations in both training and test datasets, potentially producing optimistic estimates of generalization performance. Reproducible evaluation therefore requires not only physics-based generation of electrical fault signatures but also careful separation of physical scenarios during model development and final testing.

Physics-based distribution-system simulation provides a controlled environment in which these factors can be investigated systematically. OpenDSS is widely used for distribution network analysis and supports representations of feeders, loads, distributed generation, PV systems, energy-storage resources, and fault elements. When combined with Python-based automation, OpenDSS simulations can be programmatically configured across multiple fault and DER operating conditions, while the resulting electrical measurements can be extracted and processed within a consistent machine-learning workflow. In this arrangement, OpenDSS provides the physics-based network solution, whereas Python provides scenario orchestration, data processing, feature generation, and machine-learning evaluation. This distinction is important because the physical behavior represented in the present study is determined by the OpenDSS network and DER models rather than by Python itself. Accordingly, this paper presents a physics-informed closed-loop Python–OpenDSS co-simulation framework for fault detection, fault-type classification, and fault localization in DER-integrated distribution networks. Within the terminology of this study,

physics-informed refers to the use of electrical measurements generated from physics-based OpenDSS network solutions as inputs to the diagnostic models; the machine-learning algorithms themselves do not embed the network equations within their learning architectures. The closed-loop terminology refers to the automated computational cycle in which Python configures a physical scenario, OpenDSS solves the corresponding network state, Python retrieves and processes the resulting electrical quantities, and the diagnostic workflow evaluates the generated features before the subsequent scenario is executed. Machine-learning predictions are not fed back to actuate protection devices or alter inverter-control actions.

The experimental evaluation considers three complementary diagnostic tasks: binary fault detection, fault-type classification, and fault localization. The IEEE 13-node, IEEE 34-node, and IEEE 37-node benchmark feeders are used to examine classification and localization performance under different network configurations. The IEEE 13-node feeder also serves as the detailed DER case study, incorporating explicit PV and BESS operating conditions and a separate binary fault-detection experiment. The DER analysis compares matched operating cases to quantify how PV and BESS operation modifies the network-level phasor-domain fault signatures available to the diagnostic models. The DER representations are evaluated at the phasor-domain network level and are not intended to reproduce electromagnetic-transient phenomena such as PWM switching, phase-locked-loop dynamics, DC-link transients, or detailed inner converter-control loops.

The remainder of this paper is organized as follows. Section 2 presents the proposed co-simulation methodology. Section 3 describes the simulation framework and experimental setup, including the benchmark feeders, DER representations, fault scenarios, measurement configuration, and evaluation procedure. Section 4 presents the fault-detection, fault-type classification, fault-localization, DER-sensitivity, and cross-feeder results, together with evaluation-integrity and generalization analyses. Section 5 concludes the paper and outlines directions for future research.

2. Proposed Closed-Loop Diagnostic Methodology

2.1. Closed-Loop Python–OpenDSS Co-Simulation Framework

This section presents the physics-informed closed-loop Python–OpenDSS co-simulation methodology used for fault detection, fault-type classification, and fault localization. The framework integrates physics-based distribution-network simulation in OpenDSS with Python-based scenario orchestration, electrical-feature generation, machine-learning model development, and diagnostic evaluation. OpenDSS determines the electrical response of each configured network condition, whereas Python manages the computational workflow and processes the resulting measurements [28,29]. The methodology provides a consistent procedure for generating and evaluating fault-diagnostic datasets across different benchmark feeder configurations. Within this study, the term physics-informed refers to the use of electrical quantities generated from physics-based OpenDSS network solutions as inputs to the diagnostic models. Machine learning models do not explicitly embed power-system equations or physical constraints within their architecture, objective functions, or inference procedures. Similarly, the term closed-loop describes the automated computational cycle connecting scenario configuration, physics-based simulation, measurement retrieval, feature generation, and diagnostic evaluation. It does not denote a real-time protection-control feedback loop. The overall closed-loop Python–OpenDSS co-simulation and diagnostic workflow is illustrated in Figure 1.

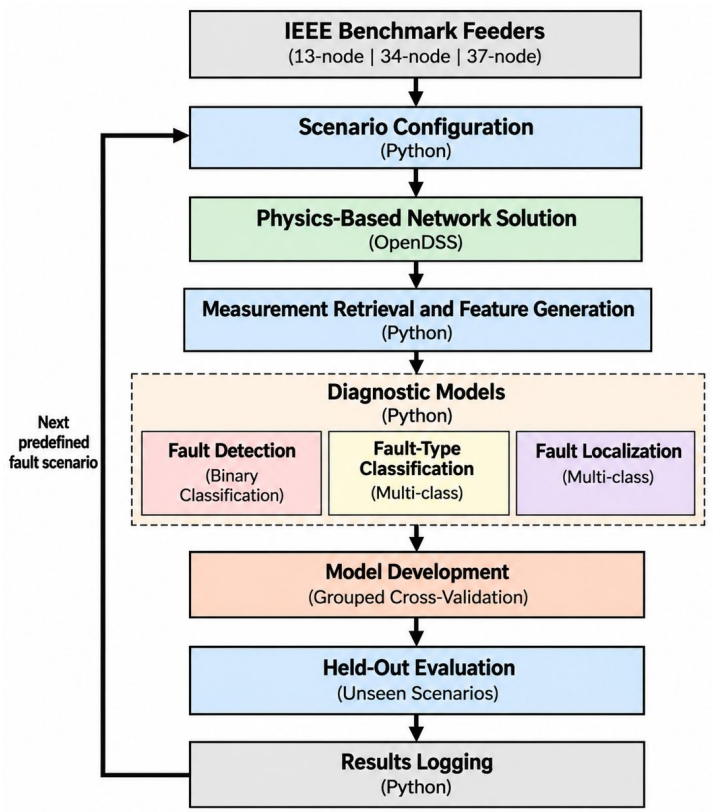


Figure 1. Proposed Python–OpenDSS closed-loop co-simulation workflow. Python configures predefined simulation scenarios, OpenDSS computes the physics-based phasor-domain network solution, Python extracts electrical features, performs machine-learning-based fault diagnosis, evaluates performance, and proceeds to the next predefined scenario.

2.2. Fault Scenario Generation

Fault scenarios are generated using OpenDSS fault elements applied at predefined candidate locations within each benchmark feeder [30]. Three representative fault categories are considered: single-line-to-ground (SLG), line-to-line (LL), and three-phase (LLL) faults. Each fault category is evaluated over a range of fault-resistance values, as defined in Equation (1)

$$R_f \in \{0.01, 0.1, 1, 10, 50\}\ \Omega, \quad (1)$$

representing conditions from low to comparatively high resistance. A physical scenario is defined by the combination of fault type, fault location, fault resistance, and DER operating condition. For the detailed IEEE 13-node DER study, three matched DER operating cases are considered: C0, representing PV OFF/BESS OFF; C1, representing PV ON/BESS OFF; and C2, representing PV ON/BESS ON. This matched structure isolates the effect of DER operation on the network-level electrical fault response while maintaining identical underlying fault configurations. Each converged scenario produces electrical measurements at multiple monitoring locations, meaning several records may originate from the same physical scenario. A unique physical-scenario identifier is retained to preserve this relationship during dataset construction and partitioning. Scenario descriptors such as fault resistance, DER operating case, fault location, and measurement-bus identifier are maintained as metadata and are not supplied to the supervised diagnostic models as predictive inputs. All simulations represent phasor-domain network solutions. Consequently, no artificial waveform sampling frequency, fault-duration window, or sampled time-domain analysis interval is imposed. The simulated responses should therefore be interpreted as steady-state phasor-domain fault conditions rather than electromagnetic-transient waveforms [30].

2.3. Electrical Measurement and Feature Representation

For each converged fault scenario, OpenDSS was used to obtain the three-phase voltage and current response at the predefined monitoring locations. The extracted electrical quantities represent the phasor-domain network solution associated with each fault condition rather than high-frequency electromagnetic-transient waveforms. Accordingly, no artificial waveform sampling frequency or fixed-duration time window was imposed during dataset generation. The extracted features therefore represent steady-state phasor quantities rather than sampled time-domain waveforms. At each predefined monitoring location, the phase-voltage magnitudes V_a, V_b, and V_c, together with the corresponding phase-current magnitudes I_a, I_b, and I_c, were extracted from the OpenDSS solution, where the subscripts a, b, and c denote phases A, B, and C, respectively. These quantities formed the basis of the diagnostic feature vectors [10].

To characterize the electrical response, additional statistical features were derived from the phase measurements. The minimum, maximum, and mean phase-voltage magnitudes were computed as

$$\begin{aligned} V_{min} &= min\ (V_a, V_b, V_c), \\ V_{max} &= max\ (V_a, V_b, V_c), \\ V_{mean} &= \frac{V_a+V_b+V_c}{3}, \end{aligned} \tag{2}$$

with the corresponding current features I_{min}, I_{max}, and I_{mean} calculated analogously from I_a, I_b, and I_c. Voltage-retention features were calculated relative to the corresponding healthy operating condition and comprised the phase-specific quantities $V_{ret,A}$, $V_{ret,B}$, and $V_{ret,C}$, together with the aggregate quantities $V_{ret,min}$, $V_{ret,max}$, and $V_{ret,mean}$. These features characterize both the magnitude distribution of the electrical response and the degree of fault-induced voltage depression. Only the resulting physics-generated electrical features were supplied as predictive inputs to the machine-learning models. Scenario descriptors and target information were explicitly excluded from the predictive feature set. In particular, the scenario identifier, fault resistance, DER operating-case label, fault location, measurement-bus identifier, fault-type label, and binary-detection label were never provided as predictive inputs. Instead, the diagnostic labels were retained solely as supervised-learning targets, whereas the remaining scenario descriptors were used only as metadata for dataset construction, scenario grouping, and post-processing analysis.

2.4. Machine-Learning-Based Fault Diagnosis

The diagnostic framework considers three complementary diagnostic tasks: binary fault detection, fault-type classification, and fault localization. Binary fault detection distinguishes healthy and faulted operating conditions and is evaluated separately using the IEEE 13-node detailed case study. Fault-type classification identifies the simulated fault category among SLG, LL, and LLL, whereas fault localization estimates the candidate network location at which the fault was applied. Four representative learning approaches are evaluated for the fault-type classification and fault-localization tasks: Support Vector Classifier (SVC), Random Forest (RF), Convolutional Neural Network (CNN), and K-Means clustering [8,10]. SVC and RF represent supervised learning algorithms with different decision mechanisms, whereas the CNN provides a deep-learning-based supervised approach. K-Means is retained as an unsupervised reference baseline for assessing the separability of the physics-generated electrical feature space.

For the supervised diagnostic tasks, the general learning objective can be represented as in Equation (3)

$$\hat{y} = f_\theta(x), \tag{3}$$

where x denotes the electrical feature vector, f_θ denotes the learned diagnostic model with model parameters θ, and $\hat{y}$ denotes the predicted diagnostic class. Depending on the task,

the target y represents either the binary healthy/faulted state, the fault category, or the candidate fault location. The binary fault-detection experiment is evaluated separately using an SVC selected exclusively from the IEEE 13-node development partition. The comparative SVC, RF, CNN, and K-Means analysis is applied to the fault-type classification and fault-localization tasks. Model-specific preprocessing, model development, and hyperparameter selection were performed exclusively using the corresponding development partitions, without using information from the final held-out test partitions.

2.5. Scenario-Level Model Development and Evaluation

A central component of the proposed methodology is the scenario-level separation of data during model development and final evaluation. Because a single physical scenario can generate measurements at multiple monitoring locations, random record-level partitioning could distribute related observations across both the model-development and held-out datasets, thereby introducing information leakage. To prevent this, dataset partitioning uses the physical-scenario identifier as the grouping variable. All measurement records originating from a given physical scenario are assigned exclusively to the same partition. The model-development partition is used for model fitting, preprocessing, and hyperparameter selection, while the final held-out partition remains isolated until model development has been completed. Where grouped cross-validation is employed, grouping is likewise performed at the physical-scenario level so that all records associated with the same physical scenario remain within the same fold.

Final held-out evaluation is performed only after model selection has been completed. Performance is assessed using accuracy, macro-averaged precision, macro-averaged recall, and macro-F1 score, with balanced accuracy additionally reported for the binary fault-detection task because of the unequal representation of healthy and faulted observations. Data-integrity audits are also performed to assess physical-scenario separation and identify potential duplication of electrical feature vectors across the model-development and held-out partitions. The methodology is applied across the IEEE 13-node, IEEE 34-node, and IEEE 37-node benchmark feeders for fault-type classification and fault localization. The IEEE 13-node feeder additionally provides the detailed PV/BESS sensitivity analysis and the separate binary fault-detection experiment. The resulting methodology therefore provides a common physics-based, scenario-level, leakage-controlled evaluation framework for comparing diagnostic performance across benchmark distribution feeders.

3. Simulation Framework and Experimental Setup

This section describes the simulation environment, benchmark microgrid models, fault configurations, and measurement settings used to evaluate the proposed closed-loop diagnostic framework. All simulations are performed using OpenDSS as the physics-based power system simulator, while Python orchestrates fault generation, measurement acquisition, feature extraction, and ML inference through a bidirectional COM interface [9]. This integrated environment ensures reproducible, physics-consistent evaluation of diagnostic performance across IEEE 13-node, IEEE 34-node, and IEEE 37-node benchmark feeders.

3.1. MG Test Systems

Three benchmark distribution feeders, the IEEE 13-node, IEEE 34-node, and IEEE 37-node test feeders, were used to evaluate the proposed fault-diagnosis framework under networks of different sizes and topological characteristics [11]. Using multiple feeders enables investigation of whether fault-classification and fault-localization performance is preserved as feeder structure, electrical distance, phase unbalance, and measurement-to-fault relationships change. The IEEE 13-node feeder also serves as the detailed DER

case-study platform, while the IEEE 34-node and IEEE 37-node feeders assess the topology-dependent generalization of the proposed diagnostic methodology. The feeder models were implemented in OpenDSS using traceable benchmark network definitions. The IEEE 13-node feeder represents a compact and highly unbalanced distribution network and was used for detailed investigation of inverter-interfaced DER effects. The IEEE 34-node feeder provides a longer radial structure with greater electrical separation between candidate fault locations, while the IEEE 37-node feeder offers a distinct unbalanced topology with different phase-configuration characteristics [31].

For the IEEE 13-node DER case study, a photovoltaic (PV) system was connected at bus 634 using the OpenDSS PVSystem element [30]. The PV plant was represented with a maximum-power-point rating of 285 kW, an inverter rating of 300 kVA, an irradiance operating point of 1.0 p.u., and unity power factor. A battery energy storage system (BESS) was additionally connected at bus 675 using the OpenDSS Storage element. The BESS was rated at 250 kW/500 kWh, initialized at 50% state of charge, and operated at a discharge level of 125 kW. The PV operating parameters were based on the OpenDSS PVSystem modelling example, whereas the BESS rating, initial state of charge, and discharge operating point were study-defined parameters adopted for the comparative DER analysis [10]. The principal PV and BESS parameters used in the IEEE 13-node case study are summarized in Table 1. Three DER operating conditions were defined to isolate the effect of inverter-interfaced resources on the simulated fault signatures: C0, with both the PV and BESS disabled; C1, with the PV enabled and the BESS disabled; and C2, with both the PV and BESS enabled. All other feeder characteristics, fault parameters, and measurement locations were maintained unchanged when comparing these operating conditions. The modified IEEE 13-node feeder configuration, including the PV and BESS locations, candidate fault locations, and measurement points used in the detailed case study, is illustrated in Figure 2.

This matched-case experimental design ensures that changes in the measured fault response can be attributed to the DER operating condition rather than to variations in fault type, fault resistance, or fault location. The DER representation adopted in this work is intentionally limited to the network-level phasor-domain modelling capabilities of OpenDSS. Accordingly, the PVSystem and Storage elements represent the electrical contribution of inverter-interfaced resources through specified operating points and their corresponding phasor-domain network response but do not reproduce detailed electromagnetic-transient (EMT) phenomena such as semiconductor switching, pulse-width modulation, phase-locked-loop dynamics, DC-link transients, or inner current-control loops [30]. Consequently, references to inverter-dominated behavior throughout this paper denote the influence of inverter-interfaced DERs on network-level voltage and current fault signatures, rather than detailed converter electromagnetic dynamics. For the IEEE 34-node and IEEE 37-node comparative studies, five candidate fault locations were selected across upstream, intermediate, and downstream sections of each feeder to provide representative topological coverage. This topology-based selection avoids restricting the localization problem to electrically adjacent buses and provides a consistent basis for evaluating how feeder structure influences the separability of fault-location signatures. The same fault categories, fault-resistance levels, DER operating-case structure, measurement philosophy, feature-extraction procedure, machine-learning models, and evaluation protocol were maintained wherever applicable to ensure consistent comparison across all benchmark feeders. This experimental design enables diagnostic performance to be evaluated across different network sizes and topological conditions while preserving a consistent physics-based simulation and machine-learning evaluation framework.

Table 1. DER representation and operating parameters for the IEEE 13-node detailed DER case study.

DER	Parameter	Value	Parameter Source
PV	OpenDSS element	PVSystem	OpenDSS model
PV	Connection bus	634	Feeder configuration
PV	Nominal voltage	0.48 kV	Feeder configuration
PV	Maximum-power-point rating, Pmpp	285 kW	OpenDSS PVSystem example
PV	Inverter rating	300 kVA	OpenDSS PVSystem example
PV	Irradiance	1.0 p.u.	Operating condition
PV	Power factor	1.0	Operating condition
BESS	OpenDSS element	Storage	OpenDSS model
BESS	Connection bus	675	Feeder configuration
BESS	Nominal voltage	4.16 kV	Feeder configuration
BESS	Rated power	250 kW	Study-defined
BESS	Energy capacity	500 kWh	Study-defined
BESS	Initial state of charge	50%	Study-defined
BESS	Discharge operating point	125 kW	Study-defined

Note: C0 denotes PV OFF/BESS OFF, C1 denotes PV ON/BESS OFF, and C2 denotes PV ON/BESS ON. The PV parameters follow the OpenDSS PVSystem modelling basis, whereas the BESS rating and operating point are study-defined parameters used for the comparative DER analysis.

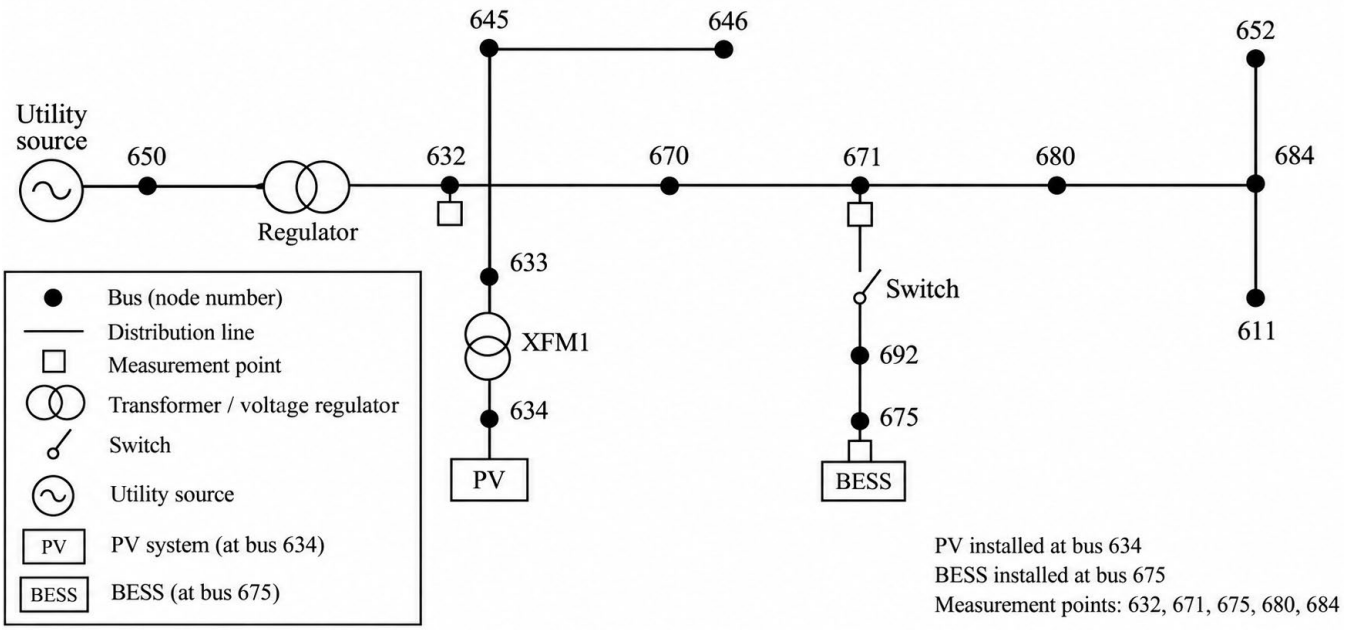


Figure 2. Modified IEEE 13-node benchmark feeder used for the detailed DER case study, showing the integration of the PV system at bus 634 and BESS at bus 675, together with the candidate fault locations and measurement points used for fault-diagnosis dataset generation.

3.2. Fault Scenario Generation and Dataset Construction

Fault scenarios were generated through systematic variation of fault type, fault location, fault resistance, and DER operating condition. Three fault categories were considered: single-line-to-ground (SLG), line-to-line (LL), and three-phase (LLL) faults. The corresponding fault configurations are illustrated in Figure 3. Fault resistance was varied over 0.01, 0.1, 1, 10, and 50 Ω, spanning low-impedance to comparatively high-resistance fault

conditions. The same fault-generation procedure was applied consistently across all benchmark feeders to enable direct comparison of diagnostic performance. For the IEEE 13-node feeder, faults were applied at buses 632, 633, 671, 675, 680, and 684. Each fault configuration was evaluated under the three DER operating conditions C0, C1, and C2. The resulting factorial design comprised 270 physical fault scenarios (6 fault locations × 3 fault types × 5 fault-resistance values × 3 DER conditions). Three predefined monitoring locations were retained for each physical scenario, producing 810 measurement records. For the IEEE 34-node feeder, five topology-distributed candidate fault locations were evaluated using the same three fault categories, five fault-resistance values, and three DER operating conditions. All 225 configured physical fault scenarios converged successfully in OpenDSS. Three predefined monitoring locations were retained for each scenario, resulting in 675 measurement records. The IEEE 37-node feeder was initially configured using the same factorial design of five fault locations, three fault types, five fault-resistance values, and three DER operating conditions, corresponding to 225 intended physical fault scenarios. However, 15 fault scenarios were reproducibly non-convergent in OpenDSS.

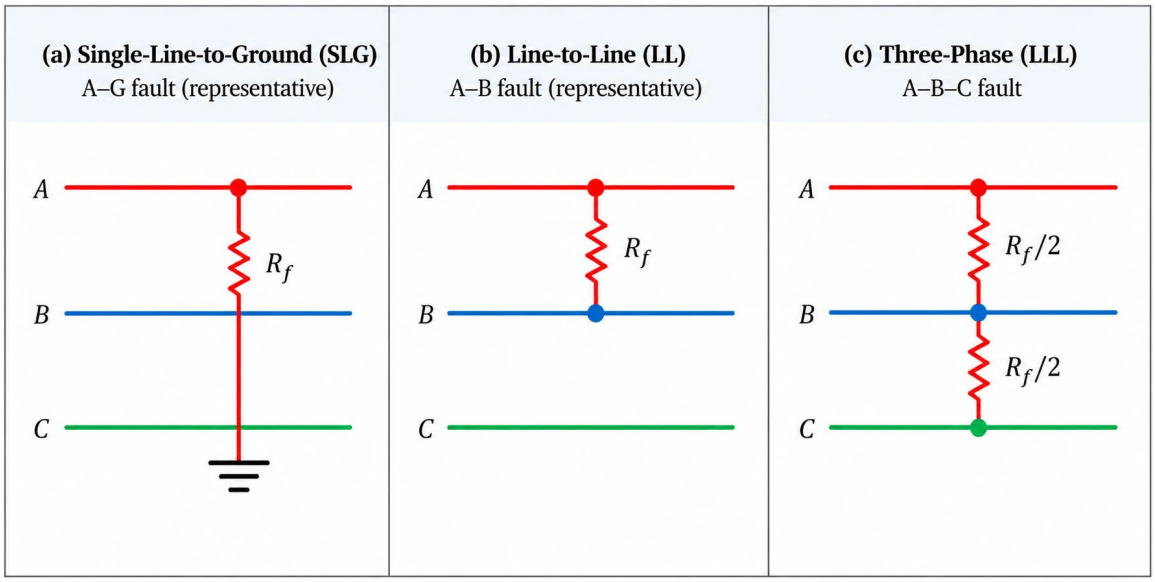


Figure 3. Fault configurations considered in the study: single-line-to-ground (SLG), line-to-line (LL), and three-phase (LLL) faults, evaluated for fault-resistance values of 0.01, 0.1, 1, 10, and 50 Ω. The colors distinguish the three phases (A, B, and C).

Repeated reconstruction of the affected scenarios confirmed that the corresponding healthy operating conditions converged, whereas the faulted solutions did not. Alternative OpenDSS fault-element formulations were also evaluated but did not restore convergence. These cases were therefore excluded before dataset partitioning and machine-learning evaluation, rather than modifying network parameters to force numerical convergence. The final IEEE 37-node dataset consequently contained 210 converged physical fault scenarios and 630 measurement records. This exclusion is explicitly reported as a numerical limitation of the OpenDSS-based simulation study. To prevent information leakage during machine-learning evaluation, every physical fault configuration was assigned a unique scenario identifier before dataset partitioning. All measurement records originating from the same physical scenario were retained within the same partition. Consequently, measurements derived from a single OpenDSS fault simulation could not appear in both the training and held-out test sets. The same scenario-level partitioning strategy was applied to the IEEE 13-node dataset to ensure consistency across all benchmark feeders. The final IEEE 34-node partition contained 180 training and 45 held-out test scenarios, while the IEEE

37-node partition contained 168 training and 42 held-out test scenarios. For the fault-type classification and fault-localization datasets, audit checks confirmed zero physical-scenario overlap and zero exact electrical-feature duplication between the corresponding model-development and held-out test partitions.

3.3. Measurement Configuration and Data Generation

For each converged fault scenario, OpenDSS was used to obtain the three-phase voltage and current response at the predefined monitoring locations. The extracted electrical quantities represent the phasor-domain network solution associated with each fault condition rather than high-frequency electromagnetic-transient waveforms. Accordingly, no artificial waveform sampling frequency or fixed-duration time window was imposed during dataset generation. Consequently, the extracted features represent steady-state phasor quantities rather than sampled time-domain waveforms. At each predefined monitoring location, the phase-voltage magnitudes $\boldsymbol{V_a}$, $\boldsymbol{V_b}$, and $\boldsymbol{V_c}$, together with the corresponding phase-current magnitudes $\boldsymbol{I_a}$, $\boldsymbol{I_b}$, and $\boldsymbol{I_c}$, were extracted from the OpenDSS solution. These measurements formed the basis for a set of derived electrical features describing voltage and current magnitude distributions and fault-induced voltage retention. The candidate feature set included the minimum, maximum, and mean phase-voltage magnitudes ($\boldsymbol{V_{\mathrm{min}}}$, $\boldsymbol{V_{\mathrm{max}}}$, and $\boldsymbol{V_{\mathrm{mean}}}$); the corresponding current quantities ($\boldsymbol{I_{\mathrm{min}}}$, $\boldsymbol{I_{\mathrm{max}}}$, and $\boldsymbol{I_{\mathrm{mean}}}$); and phase-specific and aggregate voltage-retention quantities ($\boldsymbol{V_{\mathrm{ret},A}}$, $\boldsymbol{V_{\mathrm{ret},B}}$, $\boldsymbol{V_{\mathrm{ret},C}}$, $\boldsymbol{V_{\mathrm{ret,min}}}$, $\boldsymbol{V_{\mathrm{ret,max}}}$, and $\boldsymbol{V_{\mathrm{ret,mean}}}$). Voltage-retention features were calculated relative to the corresponding healthy network operating condition, thereby providing normalized indicators of fault-induced voltage depression.

The resulting feature vectors were used as the electrical inputs to the machine-learning models. Scenario descriptors and target information were explicitly excluded from the predictive feature set. In particular, the scenario identifier, fault resistance, DER operating-case label, fault location, measurement-bus identifier, fault-type label, and binary-detection label were never provided as predictive inputs. Instead, the diagnostic labels were retained solely as targets for their respective supervised-learning tasks, whereas the remaining scenario descriptors were used only as metadata for dataset construction, scenario grouping, and post-processing analysis. Three diagnostic tasks were evaluated. Binary fault detection was evaluated using the IEEE 13-node detailed case study and formulated as a binary classification problem distinguishing healthy from faulted operating conditions. Fault-type classification was formulated as a multiclass classification problem distinguishing among SLG, LL, and LLL faults, while fault localization was formulated as a multiclass classification problem in which the objective was to identify the candidate fault location. Fault-type classification and fault localization were evaluated across the IEEE 13-node, IEEE 34-node, and IEEE 37-node benchmark feeders. The physics-generated electrical features were used by the Support Vector Classifier (SVC), Random Forest (RF), Convolutional Neural Network (CNN), and K-Means models for the fault-type classification and fault-localization studies, with model-specific preprocessing performed within the corresponding training procedures. The binary fault-detection experiment was evaluated separately using a supervised Support Vector Classifier (SVC), with model selection performed exclusively on the IEEE 13-node development partition.

To preserve the independence of the final performance assessment, model development and hyperparameter selection were performed using the training partition only. For the IEEE 34-node and IEEE 37-node studies, model selection employed grouped cross-validation at the physical-scenario level, ensuring that all records originating from the same physical scenario remained within the same fold. The IEEE 13-node binary fault-detection experiment likewise employed grouped cross-validation on the development partition, with physical-scenario identifiers used as grouping variables and the final held-out detection

partition excluded from model selection. The held-out test partitions were kept locked throughout model development and were evaluated only after the final model configurations had been selected. This procedure prevents measurements associated with the same physical scenario from contributing simultaneously to model development and final held-out evaluation.

3.4. Software Environment

The co-simulation and machine-learning framework was implemented in Python 3.12.3 on a Windows 11 platform. Distribution-network and fault calculations were performed using OpenDSS version 10.2.0.1 (64-bit), accessed programmatically from Python through the OpenDSS COM interface [28,29]. Python orchestrated simulation-scenario configuration, execution of the OpenDSS solution process, electrical-measurement extraction, dataset construction, feature processing, machine-learning model development, and diagnostic performance evaluation. OpenDSS remained the physics-based distribution-network solver throughout the study. The machine-learning implementation employed standard Python scientific-computing and machine-learning libraries. Scikit-learn was used for the Support Vector Classifier (SVC), Random Forest (RF), K-Means clustering, preprocessing, grouped cross-validation, and performance-metric calculations, while TensorFlow/Keras was used for Convolutional Neural Network (CNN) development and training. NumPy and pandas were used for numerical processing and dataset management. The experiments were executed on a computer equipped with an AMD Ryzen 7 7000-series processor, 32 GB RAM, and an NVIDIA GeForce RTX-series GPU.

3.5. Experimental Workflow and Evaluation Procedure

The complete experimental workflow was executed as an automated Python–OpenDSS co-simulation workflow. For each configured physical scenario, Python established the feeder configuration, DER operating condition, fault type, fault location, and fault resistance before invoking the OpenDSS physics-based network solution. Following convergence, the resulting three-phase voltage and current measurements were extracted at the predefined monitoring locations and processed to construct the electrical feature vectors used for diagnostic evaluation. The generated datasets were subsequently subjected to scenario-level partitioning, model-specific preprocessing, training-only model development and hyperparameter selection, and final evaluation using the held-out test partition. For the fault-type classification and fault-localization studies, the Support Vector Classifier (SVC), Random Forest (RF), Convolutional Neural Network (CNN), and K-Means models were evaluated using the physics-generated electrical feature set. The binary fault-detection task was evaluated separately on the IEEE 13-node detailed case study using the selected SVC, with healthy and faulted operating conditions forming the two diagnostic classes. In all cases, the final held-out partition remained isolated from model development and model selection.

The fault-type classification and fault-localization protocol was applied across the IEEE 13-node, IEEE 34-node, and IEEE 37-node benchmark feeders to provide a consistent basis for assessing diagnostic performance under different network configurations. Within this study, the term closed-loop co-simulation refers to the automated computational cycle in which Python configures the physical scenario, OpenDSS solves the corresponding network state, Python retrieves and processes the resulting electrical measurements, and the diagnostic workflow evaluates the generated features before the subsequent scenario is executed. Machine-learning predictions are not fed back to actuate protection devices or modify inverter controls. Figure 4 summarizes the complete experimental workflow,

from automated physics-based scenario generation in OpenDSS through feature generation, machine-learning evaluation, and final performance assessment.

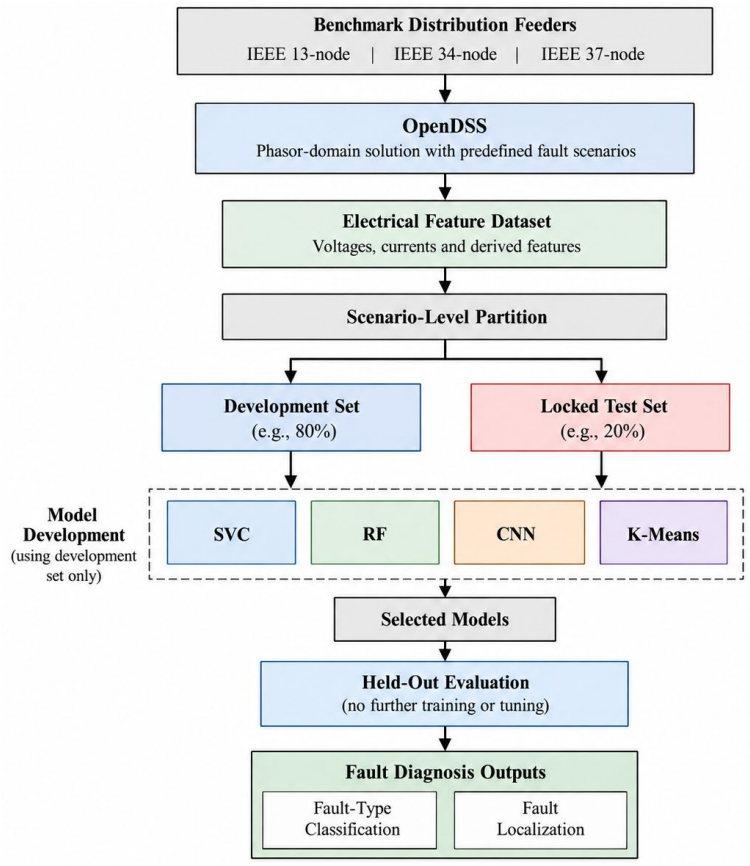


Figure 4. Automated Python–OpenDSS experimental workflow illustrating physics-based fault-scenario generation, electrical-feature extraction, scenario-level dataset partitioning, machine-learning model development, and final held-out diagnostic evaluation.

4. Results and Discussion

This section presents the results obtained from the proposed Python–OpenDSS co-simulation and machine-learning evaluation framework. The analysis considers the IEEE 13-node, IEEE 34-node, and IEEE 37-node benchmark feeders and addresses three complementary diagnostic tasks: fault detection, fault-type classification, and fault localization. Binary fault detection is evaluated using the IEEE 13-node detailed case study, while fault-type classification and fault localization are evaluated across the three benchmark feeders. Performance is assessed using accuracy, balanced accuracy (where applicable), macro-averaged precision, recall, and F1-score. In addition to diagnostic performance, the results examine the influence of DER operating conditions and feeder topology on the electrical fault signatures and the resulting machine-learning performance. All reported held-out results were obtained using scenario-level data partitions designed to prevent physical-scenario overlap between model development and final testing.

4.1. Network-Level Fault Response and DER Sensitivity

The OpenDSS simulations produced distinct network-level voltage and current responses across fault types, fault resistances, fault locations, and DER operating conditions. Consistent with the phasor-domain formulation adopted in this study, these responses represent steady-state network-level fault signatures rather than electromagnetic-transient converter waveforms. The resulting voltage and current variations provided the physics-

generated electrical features used for subsequent machine-learning evaluation. The detailed DER sensitivity analysis was performed using the IEEE 13-node feeder under the matched C0, C1, and C2 operating conditions defined in Section 3.1. Comparing the PV-enabled condition (C1) with the no-DER condition (C0) resulted in a mean absolute change in maximum phase-current magnitude of 27.78 A, while enabling both the PV and BESS (C2) resulted in a mean absolute change of 39.41 A relative to C0. The incremental difference between C2 and C1 was 11.71 A. These results demonstrate that the inclusion and operating state of inverter-interfaced DERs measurably alter the network-level current signatures observed during otherwise matched fault scenarios. Figure 5 summarizes the mean absolute changes in maximum phase-current magnitude for the C1–C0 and C2–C0 comparisons.

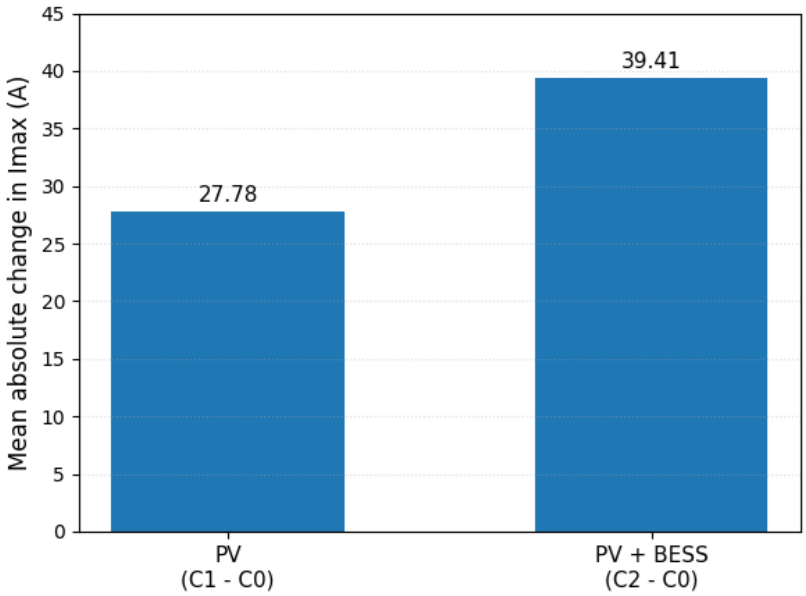


Figure 5. Mean absolute change in maximum phase-current magnitude under matched IEEE 13-node fault scenarios for PV operation (C1–C0) and combined PV–BESS operation (C2–C0), relative to the no-DER condition (C0).

The corresponding voltage-retention differences were smaller overall. The mean absolute change in voltage-retention magnitude was approximately 0.000705 p.u. for C1 relative to C0 and 0.000976 p.u. for C2 relative to C0, with maximum observed differences of approximately 0.00727 p.u. and 0.00720 p.u., respectively. The C2–C1 comparison produced a mean absolute voltage-retention difference of approximately 0.000314 p.u. These observations indicate that, for the DER operating points considered, the most pronounced DER-dependent variation occurred in the current-related fault signatures. These results should be interpreted within the modelling scope established in Section 3. The observed differences demonstrate the effect of inverter-interfaced PV and BESS operating conditions on the phasor-domain network response; they should not be interpreted as representing detailed inverter switching, phase-locked-loop dynamics, DC-link transients, or inner converter-control dynamics. Consequently, the conclusions are explicitly limited to the network-level fault behaviour represented by the OpenDSS models. These observations demonstrate that changes in DER operating conditions measurably influence the electrical quantities available to the machine-learning models, thereby motivating the subsequent evaluation of fault-classification and fault-localization performance across different benchmark feeder configurations. The influence of fault resistance on the network-level current response is further illustrated in Figure 6, which shows the mean maximum phase-current magnitude for the three simulated fault categories.

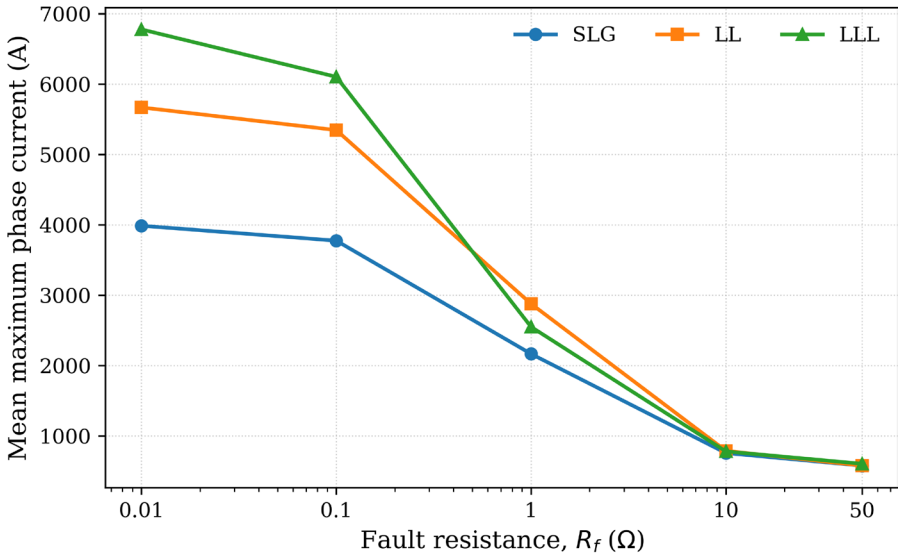


Figure 6. Mean maximum phase-current magnitude as a function of fault resistance for SLG, LL, and LLL faults in the IEEE 13-node case study, aggregated across the configured fault locations, DER operating conditions, and measurement locations.

4.2. Fault Detection and Overall Diagnostic Performance

A separate binary fault-detection task was evaluated to determine whether the physics-generated electrical measurements could distinguish between healthy and faulted operating conditions. This evaluation aligns directly with the initial diagnostic objective of the proposed framework and was conducted separately from the subsequent fault-type classification and fault-localization tasks. Binary fault detection was evaluated using the IEEE 13-node detailed case study because this feeder provides the complete DER operating-condition dataset developed in this work. For the IEEE 13-node detailed case study, healthy operating observations were generated under the same DER operating conditions used for the fault scenarios and were combined with the fault observations to construct the binary detection dataset. Scenario-level partitioning was retained so that all measurement records associated with a given physical operating scenario remained exclusively within either the training or held-out partition. The final held-out detection set comprised 273 physical scenarios and 819 measurement records, including 270 fault scenarios (810 measurement records) and 3 healthy scenarios (9 measurement records), with zero physical scenario-ID overlap between the training and held-out partitions. The final held-out binary detector achieved 100.00% accuracy, 100.00% balanced accuracy, 100.00% macro-precision, 100.00% macro-recall, and 100.00% macro-F1. Class-specific recall was likewise 100.00% for both the healthy and fault classes, with no false-positive or false-negative detections in the held-out evaluation. The corresponding confusion matrix contained 9 correctly identified healthy measurements and 810 correctly identified fault measurements, with zero misclassifications.

Because the binary detection dataset is strongly imbalanced toward fault observations, overall accuracy alone would provide an incomplete assessment of detector performance [32]. For comparison, a naive classifier that assigns every observation to the fault class would obtain approximately 98.90% overall accuracy because of the class distribution, despite achieving only 50.00% balanced accuracy and approximately 49.72% macro-F1. The simultaneous achievement of 100.00% balanced accuracy, 100.00% macro-F1, 100.00% healthy-class recall, and 100.00% fault-class recall therefore confirms that the reported detection performance is not attributable simply to majority-class prediction. Nevertheless, the detection results should be interpreted with appropriate caution because the held-out healthy class contains only three physical healthy scenarios and nine corresponding mea-

surement records. The present experiment therefore demonstrates successful separation of the evaluated healthy and faulted operating conditions within the IEEE 13-node study. Still, it does not establish detection performance across the full range of normal operating variability that may occur in practical distribution networks. Future evaluation should therefore incorporate a substantially broader set of healthy operating states, including variations in loading, DER output, switching configuration, and other non-fault disturbances that may resemble fault-induced electrical signatures. The binary detection experiment therefore establishes fault detection as a distinct first-stage diagnostic task within the proposed framework, rather than treating fault-type classification as a surrogate for detection. After detecting a faulted operating condition, the subsequent analyses evaluate the diagnostic models' ability to identify the fault category and estimate its candidate network location.

4.3. Held-Out Fault-Type Classification Performance

The held-out fault-type classification results demonstrate that the physics-generated electrical features provide strong discriminative capability for distinguishing single-line-to-ground (SLG), line-to-line (LL), and three-phase (LLL) faults across the IEEE benchmark feeders. Table 2 summarizes the classification performance of the Support Vector Classifier (SVC), Random Forest (RF), Convolutional Neural Network (CNN), and K-Means models using accuracy, macro-averaged precision, recall, and F1-score obtained from the scenario-level held-out evaluation protocol described in Section 3. For the IEEE 13-node feeder, both the SVC and RF models achieved 100% fault-type classification accuracy, together with perfect macro-precision, macro-recall, and macro-F1 scores. The CNN achieved 83.33% accuracy and a macro-F1 score of 83.10%. These results indicate that the physics-generated electrical features extracted from the detailed IEEE 13-node DER case study provided highly discriminative information for supervised fault-type classification.

Table 2. Held-out fault-type classification performance for the IEEE benchmark feeders using scenario-level evaluation.

Feeder	Model	Accuracy (%)	Macro Precision (%)	Macro Recall (%)	Macro F1 (%)
IEEE 13	SVC	100.00	100.00	100.00	100.00
IEEE 13	RF	100.00	100.00	100.00	100.00
IEEE 13	CNN	83.33	90.32	80.96	83.10
IEEE 34	SVC	91.11	91.90	91.11	91.23
IEEE 34	RF	92.59	93.11	92.59	92.62
IEEE 34	CNN	94.07	94.12	94.07	94.09
IEEE 34	K-Means	56.30	47.76	56.30	47.33
IEEE 37	SVC	91.27	93.45	89.81	90.36
IEEE 37	RF	99.21	99.28	99.07	99.16
IEEE 37	CNN	100.00	100.00	100.00	100.00
IEEE 37	K-Means	69.05	68.10	65.74	63.55

Note: A comparable held-out K-Means result is not reported for the IEEE 13-node feeder because the revised scenario-level evaluation protocol was established after completion of the detailed DER case study. Consequently, only the supervised models are included for the IEEE 13-node comparison.

For the IEEE 34-node feeder, the CNN model achieved the highest fault-type classification accuracy of 94.07%, with a macro-F1 score of 94.09%. The RF model achieved 92.59% accuracy and 92.62% macro-F1, while the SVC achieved 91.11% accuracy and 91.23% macro-F1. K-Means produced substantially lower performance, with 56.30% accuracy and

47.33% macro-F1, illustrating the performance difference between supervised classification and unsupervised clustering when applied to the same electrical feature representation. For the IEEE 37-node feeder, the CNN achieved 100% accuracy, together with perfect macro-precision, macro-recall, and macro-F1 scores on the held-out test partition. The RF model also demonstrated excellent performance, achieving 99.21% accuracy and 99.16% macro-F1, while the SVC achieved 91.27% accuracy and 90.36% macro-F1. K-Means achieved 69.05% accuracy and 63.55% macro-F1. These results indicate that the supervised learning models remained highly effective despite the different topology and electrical characteristics of the IEEE 37-node feeder.

The held-out fault-type classification accuracies across the three benchmark feeders are compared in Figure 7. Across the benchmark feeders, the results demonstrate that fault-type classification is comparatively robust to changes in feeder topology, although the relative performance of the individual learning algorithms varies with the electrical characteristics of each benchmark feeder. The supervised learning models consistently outperformed the K-Means baseline wherever comparative results were available, indicating that the physics-generated electrical features contain strong discriminative information for fault-type identification under the scenario-level held-out evaluation. K-Means is presented as an unsupervised reference and should not be interpreted as directly equivalent to the supervised models because cluster assignments are not learned from fault-type labels during model fitting.

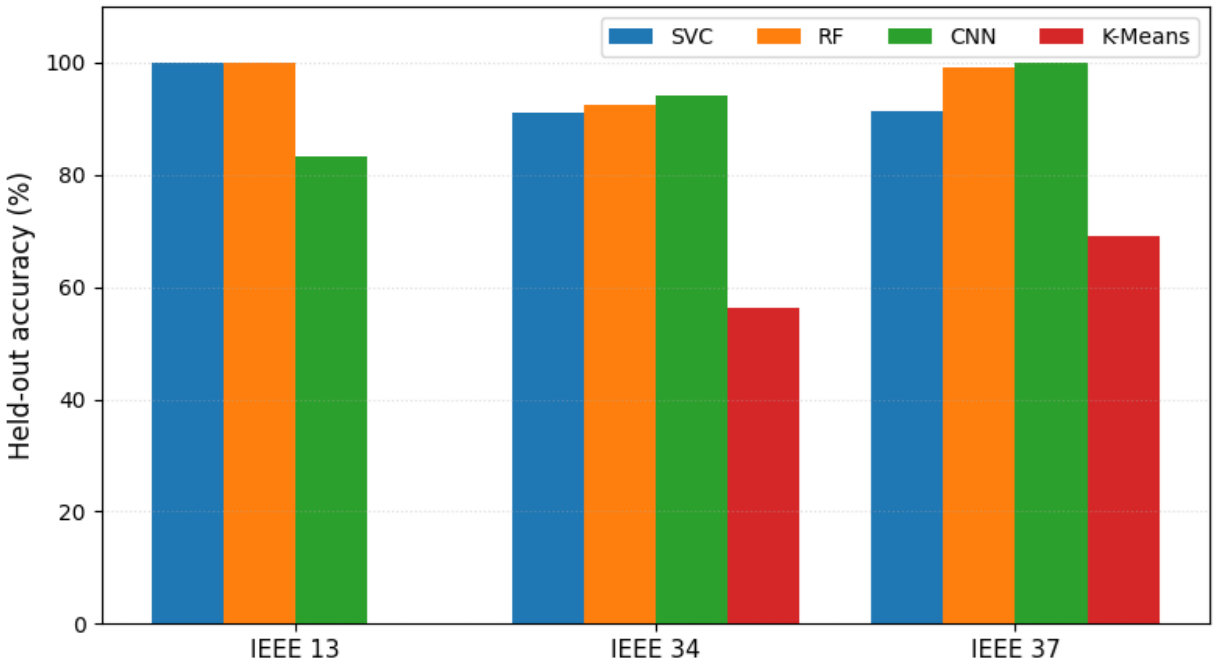


Figure 7. Held-out fault-type classification accuracy of the SVC, RF, CNN, and K-Means models across the IEEE 13-, IEEE 34-, and IEEE 37-node benchmark feeders. K-Means results are shown where verified held-out values were available.

4.4. Held-Out Fault-Localization Performance

The held-out fault-localization results demonstrate that fault-location identification is substantially more challenging than fault-type classification. Table 3 summarizes the localization performance of the Support Vector Classifier (SVC), Random Forest (RF), Convolutional Neural Network (CNN), and K-Means models across the three benchmark feeders. Localization was formulated as a multiclass classification problem in which the target corresponds to the candidate fault location defined for each feeder. For the IEEE 13-node feeder, the SVC achieved the highest verified localization accuracy of 66.67%, with

a macro-precision of 70.20%, macro-recall of 65.90%, and macro-F1 score of 64.42%. The RF achieved 65.43% accuracy and 63.40% macro-F1, while the CNN achieved 53.09% accuracy and 50.68% macro-F1. In contrast to the perfect fault-type classification achieved by the SVC and RF models on this feeder, these results demonstrate that identifying the exact fault location from the available electrical measurements is substantially more difficult than distinguishing the fault category. For the IEEE 34-node feeder, the RF achieved the strongest localization performance, with 83.70% accuracy, 84.44% macro-precision, 83.70% macro-recall, and 83.67% macro-F1. The SVC achieved 64.44% accuracy and 64.02% macro-F1, while the CNN achieved 60.00% accuracy and 60.06% macro-F1. The K-Means baseline achieved 40.74% accuracy and 30.48% macro-F1. The stronger RF performance indicates that nonlinear relationships within the physics-generated electrical features provide substantial localization information for this feeder configuration.

Table 3. Held-out fault-localization performance for the IEEE benchmark feeders using scenario-level evaluation.

Feeder	Model	Accuracy (%)	Macro Precision (%)	Macro Recall (%)	Macro F1 (%)
IEEE 13	SVC	66.67	70.20	65.90	64.42
IEEE 13	RF	65.43	65.51	65.53	63.40
IEEE 13	CNN	53.09	64.17	52.18	50.68
IEEE 34	SVC	64.44	64.86	64.44	64.02
IEEE 34	RF	83.70	84.44	83.70	83.67
IEEE 34	CNN	60.00	60.25	60.00	60.06
IEEE 34	K-Means	40.74	24.64	40.74	30.48
IEEE 37	SVC	35.71	36.50	35.46	35.64
IEEE 37	RF	61.90	62.33	61.57	61.66
IEEE 37	CNN	26.19	18.85	26.94	17.05
IEEE 37	K-Means	19.84	17.64	19.72	18.16

Note: A comparable held-out K-Means result is not reported for the IEEE 13-node feed.

For the IEEE 37-node feeder, localization performance decreased relative to the IEEE 34-node case. The RF remained the strongest supervised model, achieving 61.90% accuracy and 61.66% macro-F1. The SVC achieved 35.71% accuracy and 35.64% macro-F1, while the CNN achieved 26.19% accuracy and 17.05% macro-F1. The K-Means baseline achieved 19.84% accuracy and 18.16% macro-F1. The contrast between the near-perfect fault-type classification and the lower exact-location performance on this feeder demonstrates that successful identification of the fault category does not necessarily imply equally strong spatial discrimination of the fault location. Across the three benchmark feeders, the localization results demonstrate a stronger dependence on network topology and the spatial information contained within the available electrical features than was observed for fault-type classification. Fault-type discrimination can exploit characteristic differences in the phase-voltage and phase-current patterns, whereas exact fault-location identification additionally requires the models to distinguish electrical signatures produced at different positions within the feeder. Similar electrical signatures may therefore remain readily separable by fault category while providing less distinctive information regarding their spatial origin. The corresponding held-out fault-localization accuracies across the three benchmark feeders are compared in Figure 8.

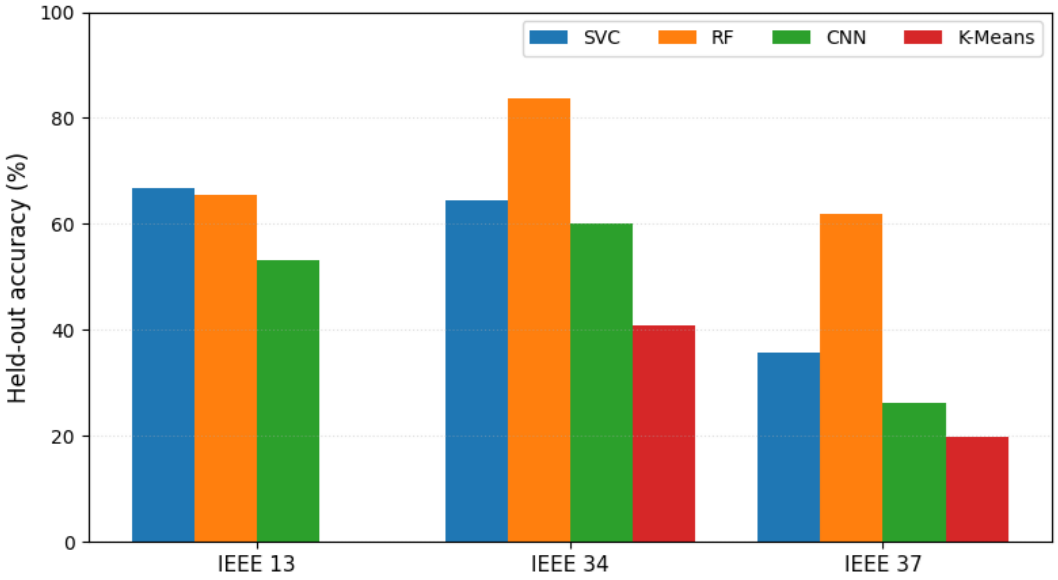


Figure 8. Held-out fault-localization accuracy of the SVC, RF, CNN, and K-Means models across the IEEE 13-, IEEE 34-, and IEEE 37-node benchmark feeders. K-Means results are shown where verified held-out values were available.

The CNN results further demonstrate that strong fault-type classification does not automatically translate into strong localization performance. In particular, the CNN achieved perfect held-out fault-type classification on the IEEE 37-node feeder but substantially lower exact-location accuracy. Under the feature representation adopted in this study, the CNN operates on a compact set of physics-generated phasor-domain electrical features rather than naturally structured spatial or time-series waveform data. Consequently, the convolutional architecture does not necessarily receive the type of structured spatial or temporal information from which convolutional models typically derive their principal advantage. These findings indicate that fault localization constitutes the more topology-sensitive component of the diagnostic problem under the measurement configuration considered in this study. They further motivate future investigation of additional spatial measurements, topology-aware features, graph-based representations, and other methods capable of explicitly encoding the electrical and structural relationships among feeder locations [11].

4.5. Cross-Feeder Performance and Diagnostic Implications

The comparative results across the IEEE 13-node, IEEE 34-node, and IEEE 37-node benchmark feeders demonstrate that diagnostic performance depends on both the diagnostic task and the electrical characteristics of the feeder. Fault-type classification remained comparatively robust across the three benchmark systems, whereas fault-localization performance exhibited substantially greater variation. This distinction indicates that the electrical information required to identify the nature of a fault is not necessarily sufficient to determine its exact location with comparable accuracy. The fault-type classification results show that the characteristic phase-voltage and phase-current responses associated with SLG, LL, and LLL faults remained distinguishable under the evaluated operating conditions across different feeder configurations. However, the relative performance of the SVC, RF, and CNN models varied across the three benchmark networks. This variation demonstrates that no single supervised learning model consistently provided the highest performance across all feeder configurations, reinforcing the importance of evaluating diagnostic algorithms on multiple benchmark networks rather than drawing conclusions from a single feeder.

The fault-localization results exhibited a stronger dependence on feeder configuration. In particular, the RF model achieved a localization accuracy of 83.70% on the IEEE 34-node

feeder, compared with 65.43% on the IEEE 13-node feeder and 61.90% on the IEEE 37-node feeder. The SVC and CNN models exhibited similarly pronounced variations across the benchmark feeders. These differences indicate that localization difficulty is not determined solely by feeder size. Instead, the distinguishability of candidate fault locations depends on the interaction between network topology, electrical parameters, measurement placement, fault position, and the resulting voltage and current signatures. These findings should also be considered together with the DER sensitivity analysis presented in Section 4.1. For the IEEE 13-node feeder, matched comparisons among the C0, C1, and C2 operating conditions demonstrated measurable changes in fault-current signatures when the PV and BESS operating states were modified. Consequently, DER operating condition represents an additional source of variation within the physics-generated electrical feature space used by the diagnostic models. The scenario-level evaluation framework explicitly incorporates this variation while ensuring that records originating from the same physical fault scenario are not distributed across both the training and held-out test partitions.

The comparison with K-Means further illustrates the distinction between supervised diagnostic learning and unsupervised structure discovery. For the IEEE 34-node and IEEE 37-node feeders, K-Means produced lower fault-type classification and fault-localization performance than the supervised models under the reported evaluation protocol. This outcome does not imply that clustering is unsuitable for distribution-system monitoring; rather, it indicates that clusters formed solely from similarity within the present electrical feature space do not necessarily correspond directly to the predefined fault-type or fault-location classes. Accordingly, K-Means is presented as an unsupervised reference rather than a directly equivalent alternative to the supervised classifiers. Overall, the cross-feeder results demonstrate that the proposed Python–OpenDSS co-simulation framework provides a consistent and reproducible computational environment for generating physics-based electrical fault data and evaluating multiple diagnostic approaches using a common scenario-level evaluation methodology. At the same time, the observed variation in fault-localization performance emphasizes that diagnostic results obtained for one feeder should not automatically be assumed to generalize to networks with different topologies, electrical characteristics, or measurement configurations. Collectively, these findings support the use of multi-feeder validation and leakage-controlled scenario-level held-out evaluation when assessing machine-learning-based fault-diagnosis methods for distribution networks.

4.6. Evaluation Integrity and Generalization Considerations

The interpretation of machine-learning performance depends not only on the reported accuracy and macro-averaged metrics but also on the independence of the training and held-out test data. Because each simulated physical fault scenario produced measurements from multiple monitoring locations, random record-level partitioning could distribute measurements originating from the same underlying scenario across both training and test sets. To prevent this source of information leakage, the revised evaluation adopted scenario-level partitioning, whereby all measurement records associated with a given physical fault scenario were assigned exclusively to either the training or held-out test partition. Data-integrity audits were performed following dataset partitioning. For the fault-type classification and fault-localization datasets, the audits confirmed zero physical scenario-ID overlap between the model-development and held-out partitions and zero exact duplication of the electrical-feature vectors across the two partitions. For the detailed IEEE 13-node dataset, additional numerical inspection identified a small number of feature signatures that became identical only after decimal rounding. Examination of these cases showed that they originated from distinct fault-resistance scenarios rather than duplicated physical scenarios, and no conflicting fault-type labels were associated with the rounded overlaps.

Accordingly, the held-out results reported in Sections 4.3 and 4.4 should be interpreted within this leakage-controlled scenario-level evaluation framework. For the IEEE 34-node and IEEE 37-node studies, model development, hyperparameter selection, and grouped cross-validation were performed exclusively within the training partition, while the held-out test partition remained isolated for the final evaluation. This procedure provides a more stringent assessment than evaluating models on records that may share the same underlying physical fault scenario with observations used during model development.

A numerical-convergence limitation was encountered for the IEEE 37-node feeder. Of the 225 intended fault scenarios, 15 scenarios were reproducibly non-convergent, resulting in a final dataset comprising 210 converged physical scenarios and 630 measurement records. The corresponding healthy network conditions converged successfully, and alternative OpenDSS fault-element formulations did not restore convergence for the affected cases. These scenarios were therefore excluded before dataset partitioning and machine-learning evaluation, rather than modifying the feeder or forcing numerical solutions solely to retain the intended sample count. Consequently, the IEEE 37-node performance metrics reported in this study apply to the converged scenario set.

The cross-feeder results should also be interpreted as within-feeder held-out evaluations rather than direct evidence of transfer learning between different feeders. The IEEE 13-node, IEEE 34-node, and IEEE 37-node studies provide different benchmark network configurations for examining the sensitivity of diagnostic performance to feeder characteristics, but the models were not trained on one feeder and subsequently evaluated unchanged on another. Accordingly, the reported results demonstrate performance under independently evaluated benchmark configurations rather than universal generalization to previously unseen distribution networks. Overall, the scenario-level partitioning strategy, explicit separation of model development from final held-out evaluation, and data-integrity audits strengthen the methodological basis for interpreting the reported diagnostic performance. At the same time, the limited healthy-state representation in the binary fault-detection experiment, the observed cross-feeder variation, and the IEEE 37-node convergence limitation defines important boundaries on the interpretation of these results. Future work should therefore investigate broader healthy and non-fault operating conditions, cross-feeder transfer learning, additional operating conditions, expanded measurement configurations, and independent benchmark datasets to further evaluate the generalizability of machine-learning-based fault diagnosis for distribution networks.

4.7. Discussion

The present study demonstrates that fault detection, fault-type classification, and fault localization constitute distinct diagnostic tasks with different levels of difficulty under the evaluated conditions. While binary fault detection achieved strong held-out performance in the IEEE 13-node detailed case study, fault-type classification remained strong across the evaluated benchmark feeders, whereas fault localization proved substantially more sensitive to feeder topology, measurement placement, and the electrical separability of candidate fault locations. These findings indicate that strong fault-type classification performance should not be interpreted as evidence of equally reliable fault localization. The results further demonstrate that inverter-interfaced DER operating conditions modify the network-level phasor-domain electrical signatures used by the diagnostic models. However, these conclusions remain bounded by the modelling fidelity of the OpenDSS phasor-domain framework employed in this study. Accordingly, the reported results represent network-level phasor-domain behaviour rather than detailed electromagnetic-transient (EMT) inverter dynamics. Finally, the adoption of scenario-level dataset partitioning, together with the explicit separation of model development from final held-out evaluation,

provides a rigorous basis for assessing diagnostic performance while minimizing information leakage. Future work should investigate broader healthy operating conditions, topology-aware feature representations, cross-feeder transfer learning, and higher-fidelity electromagnetic-transient (EMT) simulation to further evaluate the generalizability of machine-learning-based fault diagnosis.

5. Conclusions and Future Work

This paper presented a Python–OpenDSS co-simulation framework for physics-based fault-scenario generation and machine-learning-based fault detection, fault-type classification, and fault localization in DER-integrated distribution networks. OpenDSS provided a physics-based network solution, while Python performed automated scenario orchestration, electrical-feature extraction, dataset construction, model development, and diagnostic evaluation. The framework was evaluated using the IEEE 13-node, IEEE 34-node, and IEEE 37-node benchmark feeders, with the IEEE 13-node feeder additionally serving as the detailed PV/BESS sensitivity analysis and binary fault-detection case study. The binary fault-detection experiment on the IEEE 13-node feeder achieved 100.00% held-out balanced accuracy and 100.00% macro-F1, with no false-positive or false-negative detections under the evaluated operating conditions. Fault-type classification also remained strong across the benchmark feeders, with held-out accuracies ranging from 83.33% to 100.00% among the supervised models. In contrast, fault localization exhibited substantially greater variation across feeder configurations and learning algorithms, demonstrating that reliable identification of fault type does not necessarily imply equally reliable identification of the exact fault location. The localization results further indicate that diagnostic difficulty depends on the interaction among network topology, electrical characteristics, measurement placement, and fault position, rather than feeder size alone.

The detailed DER sensitivity analysis demonstrated that inverter-interfaced PV and BESS operating conditions measurably modify the network-level phasor-domain fault signatures available to the diagnostic models. Relative to the no-DER operating condition, the PV-enabled case produced a mean absolute change of 27.78 A in maximum phase-current magnitude, while the combined PV–BESS case produced a corresponding change of 39.41 A. These findings demonstrate that DER operating conditions modify the network-level phasor-domain fault response; however, they should not be interpreted as representing detailed electromagnetic-transient inverter switching, phase-locked-loop dynamics, DC-link transients, or inner converter-control dynamics. The use of scenario-level dataset partitioning, together with the explicit separation of model development from final held-out evaluation, provides a leakage-controlled basis for assessing diagnostic performance. Nevertheless, several limitations define the scope of the reported results. The binary detection study contains a limited number of healthy operating scenarios; the IEEE 37-node evaluation is restricted to the converged scenario set; and the cross-feeder analysis represents independent within-feeder evaluations rather than direct transfer between benchmark networks. These constraints should be considered when interpreting the reported diagnostic performance. Future work should investigate broader healthy and non-fault operating conditions, additional loading and DER operating states, adaptive measurement placement, topology-aware and graph-based feature representations, cross-feeder transfer learning, and higher-fidelity electromagnetic-transient (EMT) simulation. Hardware-in-the-loop and real-time implementation could further extend the framework toward practical protection-system validation. Overall, the proposed methodology provides a reproducible and extensible benchmark framework for the evaluation of machine-learning-based fault diagnosis across distribution-network configurations while explicitly accounting for DER operating conditions and evaluation integrity.

Author Contributions: Conceptualization, Z.E.N.; methodology, Z.E.N.; software, Z.E.N.; validation, Z.E.N.; formal analysis, Z.E.N.; investigation, Z.E.N.; resources, UNISA; data curation, Z.E.N.; writing—original draft preparation, Z.E.N.; writing—review and editing, Z.E.N.; visualization, Z.E.N.; supervision, Z.W.; project administration, Z.W.; funding acquisition, UNISA. All authors have read and agreed to the published version of the manuscript.

Funding: This work was partly supported by the South African National Research Foundation under Grant No. AJCR230704126719.

Data Availability Statement: The datasets generated during this study are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CNN	Convolutional Neural Network
AI	Artificial Intelligence
ML	Machine Learning
MG	Microgrid
DERs	Distributed Energy Resources
SVM	Support Vector Machine
SVC	Support Vector Classifier
COM	Component Object Model (OpenDSS interface)
BESS	Battery Energy Storage System
LL	Line-to-Line Fault
LLL	Three-Phase Fault
SLG	Single-Line-to-Ground Fault
PV	Photovoltaic
RF	Random Forest
GNN	Graph Neural Network

References

1. Lopes, F.V.; Davi, M.J.B.B.; Oleskovicz, M.; Hooshyar, A.; Dong, X.; Neto, A.A.A. Maturity Analysis of Protection Solutions for Power Systems Near Inverter-Based Resources. *IEEE Trans. Power Deliv.* **2024**, *39*, 2630–2643. [CrossRef]
2. Tu, H.; Yu, H.; Du, Y.; Eisele, S.; Lu, X.; Karsai, G.; Lukic, S. An IoT-Based Framework for Distributed Generic Microgrid Controllers. *IEEE Trans. Control Syst. Technol.* **2024**, *32*, 1692–1705. [CrossRef]
3. Chhetija, D.; Khadka, D.; Yusuf, G.; Hussain, R.Z.; Doolla, S. Fault Detection and Classification Scheme for Power Islands with Inverter Interfaced Distributed Generators. *Electr. Power Syst. Res.* **2024**, *233*, 110438. [CrossRef]
4. Chowdhury, R.; Fischer, N. Transmission Line Protection for Systems With Inverter-Based Resources—Part I: Problems. *IEEE Trans. Power Deliv.* **2021**, *36*, 2416–2425. [CrossRef]
5. Samkari, H.S.; Johnson, B.K. Time-Domain Protection Scheme for Microgrids With Aggregated Inverter-Based Distributed Energy Resources. *IEEE Access* **2023**, *11*, 13232–13242. [CrossRef]
6. Haddadi, A.; Farantatos, E.; Patel, M.; Kocar, I. Need for Load Modeling in Short Circuit Analysis of an Inverter-Based Resource-Dominated Power System. *IEEE Trans. Power Deliv.* **2023**, *38*, 1882–1890. [CrossRef]
7. Chowdhury, R.; Sun, C.; Taylor, D. Review of SIR Calculations for Distance Protection and Considerations for Inverter-Based Resources. *IEEE Trans. Power Deliv.* **2024**, *39*, 1420–1427. [CrossRef]
8. Basher, B.G.; Ghanem, A.; Abulanwar, S.; Hassan, M.K.; Rizk, M.E.M. Fault Classification and Localization in Microgrids: Leveraging Discrete Wavelet Transform and Multi-Machine Learning Techniques Considering Single Point Measurements. *Electr. Power Syst. Res.* **2024**, *231*, 110362. [CrossRef]
9. Divya, R.; Umamaheswari, S.; Stonier, A.A. Machine Learning Based Smart Intrusion and Fault Identification (SIFI) in Inverter Based Cyber-Physical Microgrids. *Expert Syst. Appl.* **2024**, *238*, 122291. [CrossRef]
10. Siddique, M.N.I.; Shafiullah, M.; Mekhilef, S.; Pota, H.; Abido, M.A. Fault Classification and Location of a PMU-Equipped Active Distribution Network Using Deep Convolution Neural Network (CNN). *Electr. Power Syst. Res.* **2024**, *229*, 110178. [CrossRef]
11. Ngo, Q.-H.; Nguyen, B.L.H.; Zhang, J.; Schoder, K.; Ginn, H.; Vu, T. Deep Graph Neural Network for Fault Detection and Identification in Distribution Systems. *Electr. Power Syst. Res.* **2025**, *247*, 111721. [CrossRef]

12. Fan, M.; Xia, J.; Zhang, H.; Zhang, X. Fault Location Method of Distribution Network Based on VGAE-GraphSAGE. *Processes* **2024**, *12*, 2179. [CrossRef]
13. Liu, J.; Huang, Y.; Chen, K.; Liu, G.; Yan, J.; Chen, S.; Xie, Y.; Yu, Y.; Huang, T. Graph Neural Networks for Fault Diagnosis in Photovoltaic-Integrated Distribution Networks with Weak Features. *Sensors* **2025**, *25*, 5691. [CrossRef] [PubMed]
14. Arsoniadis, C.G.; Nikolaidis, V.C. A Machine Learning Based Fault Location Method for Power Distribution Systems Using Wavelet Scattering Networks. *Sustain. Energy Grids Netw.* **2024**, *40*, 101551. [CrossRef]
15. Viana, E.R.; Sousa, A.S.; Rabelo, R.A.L.; de Araujo, F.H.D.; Solić, P. A Comparison Framework for Fault Classification Methods in Power System Distribution Network. In Proceedings of the 2021 6th International Conference on Smart and Sustainable Technologies (SpliTech), Split/Bol, Croatia, 8–11 September 2021; pp. 1–7. [CrossRef]
16. Hu, X.; Ma, J.; Zhang, R.; Ma, D.; Wang, Q. Ts-GSAN: A Two-Stage Graphical Spatiotemporal Attention Network Fault Localization Method for Distributed Energy Systems. *IEEE Trans. Instrum. Meas.* **2025**, *74*, 1–8. [CrossRef]
17. Huang, W.; Chen, P.; Huang, Y.; Chen, S. Fault diagnosis in active distribution networks with renewable energy using multi-task learning and graph attention networks. *Electr. Power Syst. Res.* **2027**, *263*, 113513. [CrossRef]
18. Wang, Z.; Huang, B.; Zhou, B.; Chen, J.; Wang, Y. An Enhanced Fault Localization Technique for Distribution Networks Utilizing Cost-Sensitive Graph Neural Networks. *Processes* **2024**, *12*, 2312. [CrossRef]
19. Jacob, R.A.; Paul, S.; Chowdhury, S.; Gel, Y.R.; Zhang, J. Real-time outage management in active distribution networks using reinforcement learning over graphs. *Nat. Commun.* **2024**, *15*, 4766. [CrossRef] [PubMed]
20. Jiang, K.; Wang, H.; Shahidehpour, M.; He, B. Block-Sparse Bayesian Learning Method for Fault Location in Active Distribution Networks With Limited Synchronized Measurements. *IEEE Trans. Power Syst.* **2021**, *36*, 3189–3203. [CrossRef]
21. Luo, X.; Deng, Z.; Wu, J.; Chen, K.; Jia, P. Fault Nature Identification and Location Scheme for Distribution Network Based on Active Injection from IIDG. *Int. J. Electr. Power Energy Syst.* **2024**, *156*, 109706. [CrossRef]
22. Orozco-Henao, C.; Bretas, A.S.; Herrera-Orozco, A.R.; Pulgarín-Rivera, J.D.; Dhulipala, S.; Wang, S. Towards Active Distribution Networks Fault Location: Contributions Considering DER Analytical Models and Local Measurements. *Int. J. Electr. Power Energy Syst.* **2018**, *99*, 454–464. [CrossRef]
23. Bahmanyar, A.; Jamali, S. Fault Location in Active Distribution Networks Using Non-Synchronized Measurements. *Int. J. Electr. Power Energy Syst.* **2017**, *93*, 451–458. [CrossRef]
24. Mirshekali, H.; Dashti, R.; Keshavarz, A.; Torabi, A.J.; Shaker, H.R. A Novel Fault Location Methodology for Smart Distribution Networks. *IEEE Trans. Smart Grid* **2021**, *12*, 1277–1288. [CrossRef]
25. Galvez, C.; Abur, A. Fault Location in Active Distribution Networks Containing Distributed Energy Resources (DERs). *IEEE Trans. Power Deliv.* **2021**, *36*, 3128–3139. [CrossRef]
26. Karabulut, B.; Manna, C.; Develder, C. Generalization of Graph Neural Network Models for Distribution Grid Fault Detection. In Proceedings of the 2025 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm), North York, ON, Canada, 16–29 October 2025; pp. 1–7. [CrossRef]
27. Li, Y.; Wei, X.; Lin, J.; Niu, G. A Robust Fault Location Method for Active Distribution Network Based on Self-Adaptive Switching Function. *Int. J. Electr. Power Energy Syst.* **2023**, *148*, 109007. [CrossRef]
28. Hariri, A.; Newaz, A.; Faruque, M.O. Open-Source Python-OpenDSS Interface for Hybrid Simulation of PV Impact Studies. *IET Gener. Transm. Distrib.* **2017**, *11*, 3125–3133. [CrossRef]
29. Dugan, R.C.; McDermott, T.E. An Open Source Platform for Collaborating on Smart Grid Research. In Proceedings of the 2011 IEEE Power and Energy Society General Meeting, Detroit, MI, USA, 24–28 July 2011; pp. 1–7. [CrossRef]
30. Yelem, S.; Goli, P.; Alhashem, M.; Gampa, S.R. OpenDSS and Typhoon HIL Co-Simulation for Real-Time Evaluation of a Distribution Network. In Proceedings of the 2023 North American Power Symposium (NAPS), Asheville, NC, USA, 15–17 October 2023; pp. 1–6. [CrossRef]
31. Kersting, W.H. Radial Distribution Test Feeders. *IEEE Trans. Power Syst.* **1991**, *6*, 975–985. [CrossRef]
32. Brodersen, K.H.; Ong, C.S.; Stephan, K.E.; Buhmann, J.M. The Balanced Accuracy and Its Posterior Distribution. In Proceedings of the 20th International Conference on Pattern Recognition (ICPR), Istanbul, Turkey, 23–26 August 2010; pp. 3121–3124. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.