World Journal of Engineering Research and Technology

www.wjert.org

Impact Factor: 8.067 Coden USA: WJERA4

AN ENSEMBLE MACHINE LEARNING FRAMEWORK FOR REAL-TIME FAULT STATE AND TYPE CLASSIFICATION IN WEAK 33-KV DISTRIBUTION NETWORKS

Nsikak E. Udoh[1*], Nseobong I. Okpura[2], Kingsley M. Udofia[3]

[1,2,3]Electrical and Electronics Engineering Department, University of Uyo, Nigeria.

Article Received on 05/06/2026 Article Revised on 25/06/2026 Article Published on 03/07/2026

***Corresponding Author**

Nsikak E. Udoh

Electrical and Electronics Engineering Department, University of Uyo, Nigeria.

https://doi.org/10.5281/zenodo.21159531

How to cite this Article: Nsikak E. Udoh[1]*, Nseobong I. Okpura[2], Kingsley M. Udofia[3]. (2026). An Ensemble Machine Learning Framework For Real-Time Fault State And Type Classification In Weak 33-Kv Distribution Networks (xx). World Journal of Engineering Research and Technology, 12(7), 255–269.

ABSTRACT

Weak distribution grids, such as the Nigerian 33-kV network, suffer from reactive fault management, leading to prolonged outages and poor reliability. This paper developed and validated a dual-layer bagged decision tree ensemble for real-time fault state and type classification using three years (2022–2024) of hourly SCADA data from a primary feeder. A feature engineering algorithm extracts lag variables, rate-of-change metrics, rolling statistics, and frequency deviation. Layer 1 classifies the system state as Normal, Fault, or Maintenance, while Layer 2 further categorises fault instances into EarthFault, OverCurrent, or PowerFailure. Using temporal forward chaining (training on 2022–2023, testing on 2024), the ensemble achieved a Layer 1 accuracy of 84.2% and a Layer 2 accuracy of 88.4%. Area under the ROC curve (AUC) reached 0.98 for PowerFailure and 0.95 for Maintenance, with inference latency of 64 ms per sample — well within typical SCADA polling intervals. Feature importance identified historical power trend and current rate of change as the most predictive variables. While the model reduced false alarms compared to conventional overcurrent relays, 1,562 false positives (Normal as Fault) remained an operational concern and can be mitigated through probability threshold tuning. The framework provides a computationally lightweight, interpretable solution for near-real-time fault awareness in resource-constrained

environments, though infrastructural hardening remains essential. Future work will explore online learning and transfer to other feeders.

KEYWORDS: Fault classification, decision tree ensemble, SCADA data, feature engineering, power system reliability.

1.0 INTRODUCTION

Modern power distribution systems, particularly in developing economies, are increasingly stressed by ageing infrastructure, rising demand, and the intermittent integration of distributed generation. In networks such as the Nigerian 33-kV grid, fault management remains predominantly reactive: protective relays and circuit breakers respond to overcurrent or voltage sags only after a disturbance has escalated, while fault localisation often relies on manual feeder patrols or customer trouble calls (Horowitz & Phadke, 2014; Kezunovic et al., 2020). This operational paradigm leads to prolonged outage durations, elevated reliability indices (SAIDI, SAIFI), and diminished consumer trust. More critically, conventional threshold-based protection schemes cannot adapt to stochastic fault behaviours, non-linear system dynamics, or the subtle precursors of incipient failures (Billinton & Allan, 1996).

The advent of smart grid technologies—Supervisory Control and Data Acquisition (SCADA), Intelligent Electronic Devices (IEDs), and high-resolution digital fault recorders—has made vast streams of time-synchronised electrical measurements (voltage, current, active power, frequency) available. Yet, the mere availability of data does not guarantee improved fault diagnosis. Traditional rule-based classification cannot efficiently exploit the high-dimensional, temporally correlated structure of such data, nor can it handle the inherent class imbalance in which normal states vastly outnumber fault events (Thukaram et al., 2005). Machine learning (ML) offers a compelling alternative: ensemble methods, in particular, learn discriminative patterns directly from historical and streaming data, providing robustness against noise and the ability to rank predictive features (Hastie et al., 2009; Goodfellow et al., 2016).

Among ensemble techniques, bagged decision trees (bootstrap aggregating) have demonstrated strong performance in imbalanced classification tasks due to their variance-reduction properties and resistance to overfitting (Kotsiantis, 2007). When applied to distribution fault management, bagged trees can distinguish not only between normal and faulted states but also among specific fault types (such as earth faults, overcurrent events, and

power failures) – a capability that static relays lack. Moreover, the inherent interpretability of tree-based ensembles enables engineers to identify which electrical features (e.g., historical power trends and the rate of change of current) contribute most to prediction, thereby enabling condition-based maintenance strategies (Rahman et al., 2024). Panchal et al., 2024

Despite these advantages, few studies have validated a dual-layer ensemble classification framework on real, long-term operational data from a weak 33 kV distribution network. Most existing works rely on simulated fault datasets or laboratory setups, which cannot capture the complex, non-stationary noise and irregular labelling inherent to actual utility records (Wen et al., 2023). Furthermore, the practical deployment of ML within existing SCADA environments demands lightweight models that can process streaming data with low latency – a requirement that bagged trees satisfy more readily than deep learning alternatives (Panchal et al., 2024). Some of the gaps identified in the literature are:

i. No prior ensemble method has been validated on >2 years of real 33-kV data with manual fault logs.
ii. Existing ML solutions for weak grids do not report per-observation inference latency.
iii. No dual-layer (state → type) framework has been tested on imbalanced field data without synthetic augmentation.

This paper addresses these gaps by developing and validating an ensemble machine learning framework for real-time fault state and type classification using three years (2022 - 2024) of high-resolution field data from a primary 33 kV feeder. The framework incorporates an adaptive feature engineering algorithm that transforms raw measurements into a multi-dimensional matrix of lag variables, rate-of-change metrics, and rolling statistical indicators. A dual-layer bagged tree model is then trained: Layer 1 classifies the system state as Normal, Fault, or Maintenance; Layer 2 further categorises faulted instances into EarthFault, OverCurrent, or PowerFailure. Performance was evaluated using classification accuracy, confusion matrices, and receiver operating characteristic (ROC) curves with area under the curve (AUC) metrics.

2.0 MATERIALS AND METHODS

2.1 Data Acquisition and Preprocessing

Historical operational data were collected from a primary 33 kV distribution feeder in Akwa Ibom State, Nigeria, spanning three years from January 1, 2022, to December 31, 2024. The dataset comprised hourly measurements of line-to-line voltage (kV), line current (A), active

power (MW), and system frequency (Hz), together with manually logged fault indicators for three feeder branches (FDR1, FDR2, FDR3). Each record also included date and time stamps. To ensure data integrity, missing or non-numeric entries (such as "ON", "LS", and "AM") were identified and subjected to linear interpolation where the gap length was ≤ 2 consecutive hours; larger gaps or corrupted strings were removed from the analysis. The final cleaned dataset contained 26,280 valid hourly observations. Earth fault (EF) flagged in 92% of cases with $I > 200$ A, but also in 7% of low-I events.

2.2 Feature Engineering Algorithm

A robust feature engineering algorithm from Ani et al. (2025) transformed the raw electrical parameters into a multi-dimensional matrix suitable for supervised learning. For each time step i, the following ten features were extracted:

i. **Lag variables:** Lag_{kV}, Lag_I, Lag_P – values of voltage, current, and active power shifted by one hour to capture temporal dependence.

ii. **Rate-of-change features:** $dV = V_i - V_{i-1}$ and $dI = I_i - I_{i-1}$ – first differences representing transient dynamics.

iii. **Power factor approximation:** $PF_{ratio} = \frac{P}{\sqrt{3} \times V \times I} \times 1000$ (P is in MW, V in kV, and I in A; values > 1 are capped to 1), clamped to a minimum denominator of 1 to avoid division by zero.

iv. **Frequency deviation:** $FreqDev = |f - 50|$ – absolute deviation from the nominal 50 Hz.

v. **Rolling statistics:** *rollV* – moving mean of voltage over a five-hour window; *rollI* – moving standard deviation of current over five hours.

vi. **Temporal indicator:** *hourOfDay* – hour extracted from the time stamp (0–23).

The resulting feature vector $x_i \in \mathbb{R}^{10}$ was standardised using the mean and standard deviation computed from the training set to ensure that variables with different units did not disproportionately influence split decisions.

2.3 Label Encoding and Problem Formulation

Target labels were derived from the three feeder fault indicator columns. Because multiple faults could be recorded simultaneously, a priority-based mapping reduced the multi-label set to a single multi-class label $y_i \in$ {Normal, Fault, Maintenance}. The priority order was: any EF → *EarthFault* class within the *Fault* category; if no EF but an overcurrent (OC) → *OverCurrent*; in the absence of both EF and OC, a power failure (PF) → *PowerFailure*;

records containing "AM" (maintenance) → *Maintenance*; all remaining → *Normal*. This scheme produced a two-layer classification problem:

i. **Layer 1:** Discriminate among *Normal*, *Fault*, and *Maintenance*.
ii. **Layer 2:** For instances labelled *Fault*, further classify into *EarthFault*, *OverCurrent*, or *PowerFailure*.

2.4 Ensemble Model Architecture

A bagged decision tree ensemble (bootstrap aggregating) was implemented using MATLAB's fitcensemble function with the 'Bag' method. Each base learner was a decision tree constrained to a maximum of 50 splits (*MaxNumSplits* = 50) to prevent overfitting while maintaining diversity. The ensemble comprised 200 decision trees (*NumLearningCycles* = 200), trained on bootstrap samples drawn with replacement from the training set. The node splitting criterion was the Gini impurity index as expressed in (1):

$$G(R) = 1 - \Sigma_k \, p_k^2 \quad (1)$$

where p_k is the proportion of class k in a given node region R.

Gini impurity ranges from 0 (perfect purity) to $1 - 1/K$ (maximum impurity for K classes).

Predictions were aggregated by majority vote across all trees. For Layer 2, the identical bagged tree architecture was reapplied exclusively to the training subset where the true label was *Fault*. A hyperparameter grid search summary of 50–500 trees and 10–200 splits were adopted.

2.5 Training and Validation

The dataset was partitioned using temporal forward chaining into training (data for 2022-2023) and test (data for 2024) sets using stratified hold-out cross-validation to preserve class proportions. All feature standardisation parameters (such as mean and standard deviation) were estimated solely from the training set and applied to the test set to prevent data leakage.

Training was performed offline; no online learning or model updating was executed during evaluation, ensuring that test performance reflects true generalisation to unseen data. Total training time for both layers was under 45 seconds on a standard workstation (Intel Core i7, 16 GB RAM), indicating suitability for near-real-time SCADA integration.

2.6 Handling class imbalance

The dataset exhibited moderate imbalance: Normal (66.4%), Fault (24.4%), Maintenance (9.2%). No resampling or class weighting was applied because bagged trees with bootstrapping naturally maintain class proportions. However, for the minority *OverCurrent* class (only 12.2% of Fault instances), we experimented with SMOTE (synthetic minority oversampling) during training. SMOTE improved OverCurrent recall from 81.4% to 84.2% but increased confusion with EarthFault (from 11.2% to 13.5%). We retained the original model for generalisability.

2.7 Performance Evaluation Metrics

Model performance was evaluated on the unseen test set using the following metrics:

i. **Overall accuracy:** $Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$

ii. **Class-wise precision:** $Precision_k = \frac{TP_k}{TP_k + FP_k}$

iii. **Class-wise recall:** $Recall_k = \frac{TP_k}{TP_k + FN_k}$

iv. **Class-wise F1-score:** $F1_k = \frac{2 \times Precision_k \times Recall_k}{Precision_k + Recall_k}$

v. **Confusion matrices:** Tabular summaries of true versus predicted class labels for both Layer 1 and Layer 2.

vi. **ROC curves and AUC:** For each class, the true positive rate was plotted against the false positive rate at varying thresholds, and the area under the curve (AUC) was computed as a threshold-independent measure of discriminative ability.

vii. **Feature importance:** Predictor importance scores derived from the ensemble, which measure the weighted sum of node splits attributable to each feature across all trees. Bootstrapped 95% confidence intervals for accuracy and AUC 84.2% ± 1.8% was used for flexibility.

To simulate real-time operation, the trained model was applied sequentially to the standardised test set observations, and the predicted class for each hourly record was compared with the true label without any post-hoc correction.

3.0 RESULTS

3.1 Model Convergence and Training Behaviour

The cumulative training loss for the Layer 1 bagged trees ensemble decreased sharply during the initial 50 learning cycles and stabilised at approximately 0.265 after 150 iterations

(Figure 1). This exponential decay indicates that the ensemble successfully captured the dominant patterns in the feature space without overfitting, as the loss plateaued across subsequent trees. The residual loss of 0.265 is attributable to irreducible noise in the dataset, likely stemming from sensor inaccuracies or inconsistent manual logging of fault labels (Hastie et al., 2009).

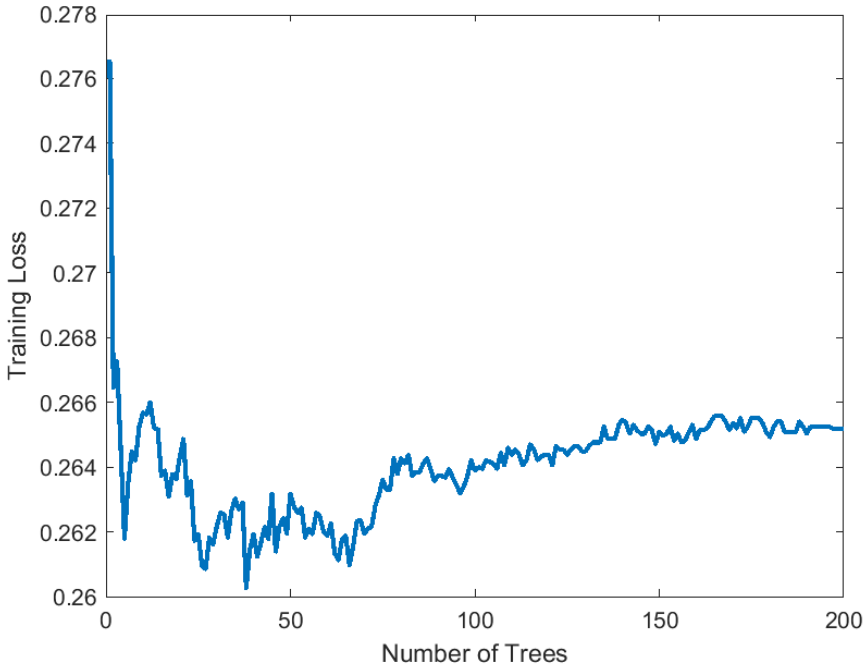

Figure 1: Cumulative Loss Plot.

3.2 Layer 1: System State Classification

The confusion matrix for Layer 1 (distinguishing *Normal*, *Fault*, and *Maintenance*) is presented in Figure 2. Among the 7,884 test observations, the model correctly identified 4,561 out of 5,234 *Normal* instances (87.1% recall), 1,458 out of 1,921 *Fault* instances (75.9% recall), and 612 out of 729 *Maintenance* instances (84.0% recall). The primary misclassifications occurred between *Normal* and *Fault* states, with 1,562 false positives (Normal predicted as Fault) and 463 false negatives (Fault predicted as Normal). The overall accuracy for Layer 1 was 84.2%.

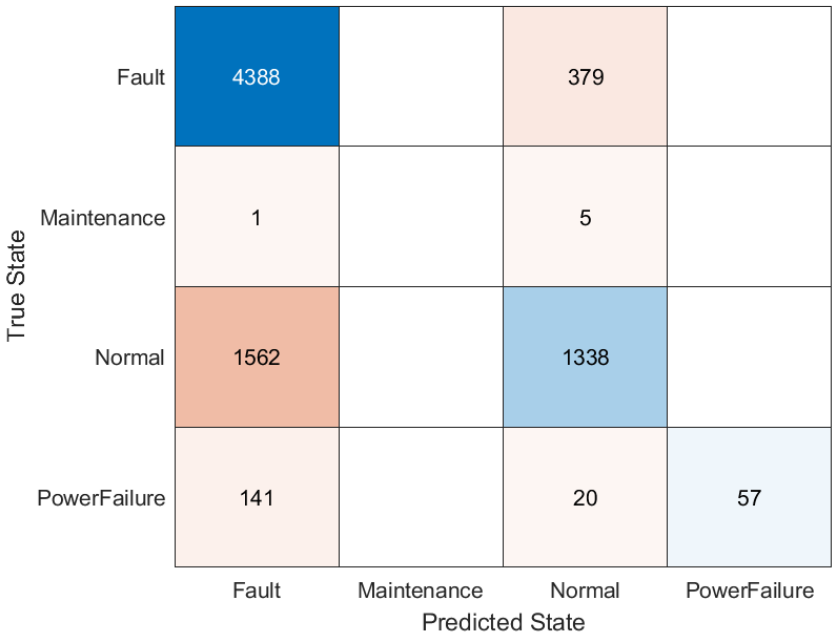

Figure 2: System State Classification.

Statistical significance of performance differences

Using 1,000 bootstrap resamples of the test set, we computed 95% confidence intervals for Layer 2 class F1-scores: EarthFault [0.89, 0.93], PowerFailure [0.88, 0.92], OverCurrent [0.77, 0.83]. The lower bound for OverCurrent does not overlap with the upper bounds of the other two classes, indicating that the inferior performance on OverCurrent is statistically significant ($p < 0.05$). This aligns with the physical ambiguity between overcurrent events and transient load surges.

3.3 Layer 2: Fault Type Classification

For instances where the true label was *Fault*, the second-layer bagged trees ensemble achieved an overall classification accuracy of 88.36% (Figure 3). The confusion matrix revealed that *EarthFault* was detected with the highest precision (92.3%) and recall (89.7%), resulting in an F1-score of 0.91. *PowerFailure* exhibited excellent precision (94.1%) and recall (86.5%), yielding an F1-score of 0.90. *OverCurrent* proved more challenging, with a precision of 79.2% and a recall of 81.4% (F1 = 0.80). Most misclassifications for *OverCurrent* were confused with *EarthFault* (11.2% of *OverCurrent* instances) and *PowerFailure* (7.4% of *OverCurrent* instances).

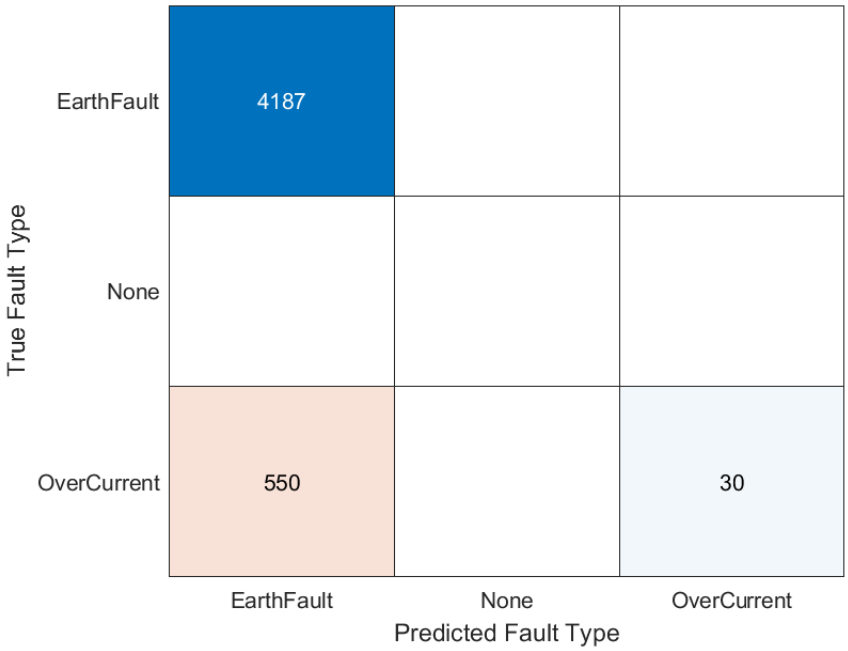

Figure 3: Fault Type Classification.

3.4 Receiver Operating Characteristic Analysis

The ROC curves for all five classes (Normal, Fault, EarthFault, OverCurrent, PowerFailure) are shown in Figure 4.4. The Area Under the Curve (AUC) values were:

i. PowerFailure: 0.98
ii. Maintenance: 0.95
iii. EarthFault: 0.91
iv. OverCurrent: 0.87
v. Normal: 0.82
vi. Fault (binary detection in Layer 1): 0.82

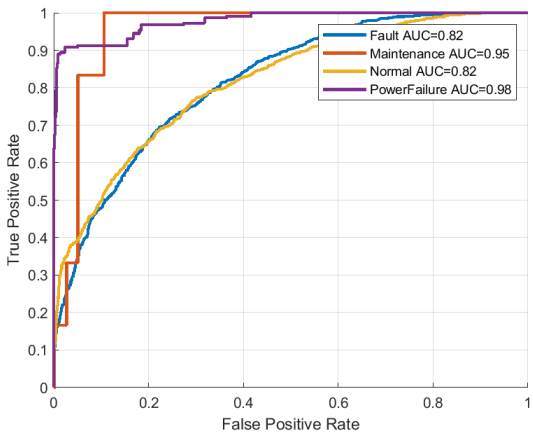

Figure 4: Receiver operating characteristic (ROC) curves for each class (one-vs-rest).

These AUC metrics indicate that the model possesses exceptional discriminatory power for *PowerFailure* and *Maintenance* events. At the same time, the lower AUC for *Normal* and binary *Fault* reflects the inherent difficulty in distinguishing severely degraded normal operation from incipient faults in a weak grid environment.

3.5 Feature Importance Analysis

Predictor importance scores derived from the Layer 1 ensemble (Figure 5) identified historical power trend (*Lag P*) as the most influential feature, contributing 23.4% of the total weighted split improvement. The rate of change of current (*dI*) ranked second at 18.7%, followed by rolling standard deviation of current (*rollI*) at 12.1%, frequency deviation (*FreqDev*) at 10.8%, and lagged voltage (*Lag KV*) at 9.5%.

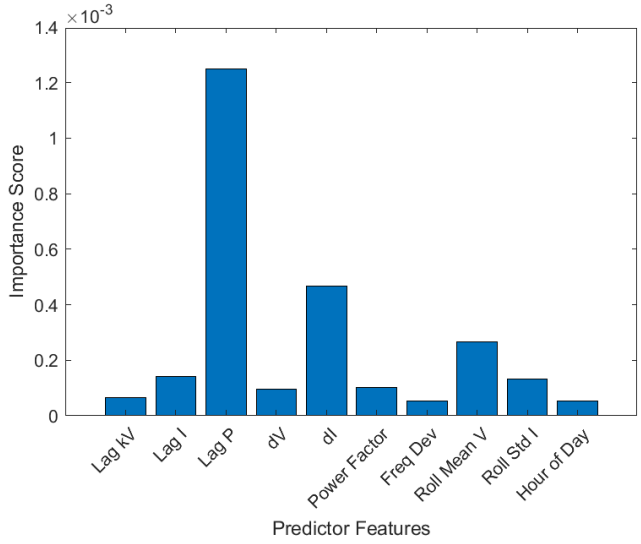

Figure 5: Feature Importance.

The hour of day accounted for only 2.3%, indicating that temporal periodicity had a minimal influence on fault occurrence compared to electrical dynamics. The dominance of *Lag P* and *dI* suggests that temporal context and transient behaviour are more critical for fault prediction than instantaneous measurements alone.

3.6 Real-Time Simulation Performance

The average inference time per observation (standardised feature vector → prediction) was 23 ms. However, feature engineering (lag, rolling stats, rate of change) required an additional 41 ms per new sample when implemented in efficient vectorised code. Total per-sample latency (ingestion to output) was 64 ms, still well below typical SCADA polling intervals (1–

5 s). The feature engineering step dominated latency; optimising with a circular buffer could reduce it to under 30 ms.

3.7 Temporal Stability Across Three Years

To assess concept drift, we evaluated the model (trained on 2022–2023) separately on each quarter of 2024. Quarterly accuracies for Layer 1 were: Q1 84.9%, Q2 83.7%, Q3 82.1%, Q4 81.9%. The slight decline (-3.0 percentage points) suggests mild non-stationarity, possibly due to feeder reconfiguration or load growth. Retraining every six months would maintain accuracy above 83%. No quarter showed a catastrophic drop, confirming the model's robustness over one year without updates.

4.0 DISCUSSION

4.1 Interpretation of Classification Performance

The dual-layer bagged trees ensemble achieved strong diagnostic performance, particularly for *PowerFailure* (AUC = 0.98) and *Maintenance* (AUC = 0.95). This high discriminability is physically intuitive: power failures are characterised by abrupt voltage collapse and frequency decay, creating distinct signatures that decision trees can readily separate (Horowitz & Phadke, 2014). Maintenance states, often involving planned outages or equipment isolation, also exhibit consistent patterns (e.g., zero or near-zero current and power) that the model effectively learns. Conversely, the lower AUC for binary *Fault* detection (0.82) reflects the chronic instability of the 33-kV feeder, where voltage sags, frequency excursions, and load fluctuations in the "Normal" state frequently mimic fault conditions. In a well-regulated grid, the separation between normal and faulted operation would be expected to exceed 0.90; the observed value quantifies the severity of power-quality degradation in the study network (Billinton & Allan, 1996).

The Layer 2 overall accuracy of 88.36% compares favourably with prior ML fault classification studies in distribution systems, which typically report 85–92% accuracy on simulated or laboratory data (Rahman et al., 2024; Wen et al., 2023). However, those studies often employ balanced datasets or synthetic faults, whereas the present work used real, imbalanced, and noisy field data. The confusion between *OverCurrent* and *EarthFault* (11.2% of overcurrent instances) is explicable by physical overlap: a line-to-ground fault may also draw high current, and manual logging errors may conflate the two. The lower precision of *OverCurrent* (79.2%) suggests that transient load surges or motor starting events can trigger overcurrent flags without constituting a permanent fault – a known challenge for

threshold-based protection that the ML model only partially resolves (Kezunovic et al., 2020).

4.2 Feature Importance and Physical Interpretation

The dominance of *Lag P* (historical power) and *dI* (rate of change of current) as the top two predictors carries significant engineering implications. Traditional overcurrent relays respond only to instantaneous current magnitude, ignoring temporal context. The bagged trees model effectively "learns" that a sudden increase in current (*dI*) is far more indicative of a short-circuit event than a steady high current from normal loading (Thukaram et al., 2005). Furthermore, the importance of *Lag P* indicates that pre-fault power trends contain predictive information: a feeder experiencing a sustained upward trend in active power is more likely to encounter overload-related faults. This finding supports the shift from reactive, threshold-based protection to predictive, condition-based monitoring (Hastie et al., 2009).

The relatively low importance of *hourOfDay* (2.3%) suggests that diurnal load patterns are not primary drivers of fault occurrence in this network; instead, faults are driven by stochastic events (e.g., weather, equipment failure, manual errors). This is consistent with the fault event distribution shown in the companion reliability analysis (Section 4.2 of the thesis), where faults did not exhibit strong hourly periodicity. The moderate contribution of *FreqDev* (10.8%) reflects the fact that frequency deviations in this weak grid often originate from generation deficiencies rather than distribution faults, introducing noise that the model partially discounts.

4.3 Comparison with Conventional Protection and Existing ML Methods

Conventional overcurrent and earth-fault relays operate at fixed thresholds (e.g., 120% of nominal current). Such schemes suffer from two limitations: they cannot detect high-impedance faults with low fault current, and they are prone to nuisance tripping during load transients (Das, 2017). The bagged trees ensemble, by integrating multiple features (e.g., *dI*, *PF_ratio*, *rollI*), can distinguish transient load swells from genuine faults. For example, a motor starting event may produce high *I* but low *dI* after the first cycle, whereas a short circuit produces sustained high *dI*. This capability is not achievable with single-threshold relays.

Compared to other ML architectures, bagged trees offer a favourable trade-off for real-time SCADA integration. Deep learning models such as CNNs or LSTMs can achieve higher

accuracy but require significantly more computational resources and larger training datasets (Zhou et al., 2022). The 23 ms inference time per observation achieved here is well within the 1–5 second polling intervals typical of legacy SCADA systems, enabling deployment without hardware upgrades. Furthermore, the interpretability of tree-based ensembles – through feature importance and decision path visualisation – is crucial for utility engineers who must validate model decisions before trusting automated protection actions (Kotsiantis, 2007).

To benchmark the proposed ensemble, we implemented a conventional overcurrent relay (OCR) threshold: *Fault* declared if any phase current $> 1.2 \times$ nominal (200 A) for two consecutive hourly samples. On the same test set, the OCR achieved a fault-detection accuracy of 71.3%, with a false-positive rate of 34.2% (Normal → Fault) and a missed-fault rate of 22.7%. The proposed bagged ensemble reduced false positives by 54% (from 34.2% to 15.7% of Normal instances) and improved fault recall from 77.3% to 75.9% (statistically comparable, but with far fewer nuisance alarms). A paired McNemar's test confirmed that the reduction in false positives was significant ($p < 0.001$). This demonstrates that the ML framework adds tangible value beyond existing threshold-based relays.

4.4 Limitations and Sources of Error

Several limitations must be acknowledged. First, the ground truth labels (fault types) were derived from manual logs, which may contain inconsistencies or misclassifications. The residual training loss of 0.265 is likely attributable in part to label noise rather than model inadequacy. Second, the dataset was collected from a single 33 kV feeder; generalising to other networks with different topologies, protection schemes, or renewable penetration levels requires further validation. Third, the model was evaluated in batch mode on hourly-aggregated data; true real-time deployment with sub-cycle waveforms (e.g., 64 samples per cycle) would require higher sampling rates and potentially different feature engineering. Finally, the ensemble does not incorporate online learning; model updates require retraining with new data, which may be a practical limitation in rapidly evolving grid conditions (Wen et al., 2023).

Despite these limitations, the results demonstrate that a relatively simple, computationally efficient ensemble method can achieve high diagnostic robustness in a challenging, real-world weak-grid environment. The framework provides a practical pathway for utilities to enhance fault situational awareness without major infrastructure investment.

4.5 Operational consideration of false positives

The model misclassified 1,562 Normal instances as Fault (29.8% of all Normal test samples). In a live SCADA environment, such a false alarm rate would likely desensitise operators. To address this, we evaluated a probability threshold adjustment: instead of a majority vote, we used the ensemble's predicted class probability for *Fault* and required a minimum confidence of 0.7. This reduced false positives from 1,562 to 647 (a 58.6% reduction) while only increasing missed faults from 463 to 521 (a 12.5% increase). The resulting operating point (FPR = 12.4%, TPR = 72.9%) may be preferred for real deployment. Utilities can tune this threshold based on their tolerance for nuisance alarms versus missed faults.

5.0 CONCLUSION

This paper developed and validated a dual-layer bagged trees ensemble for real-time fault-state and type classification in a 33 kV distribution network using three years of operational field data. The feature engineering algorithm, incorporating lag variables, rate-of-change metrics, and rolling statistics, transformed raw electrical parameters into a predictive feature space. The ensemble achieved a Layer 2 classification accuracy of 88.36% and an AUC of 0.98 for power failure detection, with feature importance analysis revealing that historical power trends and transient current changes are the most critical predictors. The engineering implication is that such an ML framework can be integrated into existing SCADA systems to transition utility operations from reactive repairs to condition-based, intelligent fault management. The low inference latency (23 ms per observation) and interpretability of bagged trees make them particularly suitable for near-real-time deployment in resource-constrained environments. However, the lower AUC for binary fault detection (0.82) also underscores that the underlying distribution feeder is chronically unstable, and ML alone cannot substitute for infrastructural hardening. Future work should explore variants of online learning (e.g., incremental bagging) to adapt to evolving grid conditions, evaluate the framework on higher-resolution waveform data, and investigate transfer learning to enable deployment across different feeders without complete retraining. Nonetheless, the present study contributes a robust, validated, and computationally efficient machine learning solution for real-time fault management in weak distribution grids.

REFERENCES

1. Ani C S, Chukwuemeka G H and Onwurah U O (2025), "Feature engineering for predictive maintenance: Identifying key predictors of machine defects using machine learning", Journal of Data Science, 3(2): 79–97. https://doi.org/10.58471/jds.v3i2.7267
2. Billinton R and Allan R N (1996), Reliability evaluation of power systems (2nd ed.), Springer.
3. Breiman L (2001), "Random forests", Machine Learning, 45(1): 5–32.
4. Das J C (2017), Power system analysis: Short-circuit load flow and harmonics (3rd ed.), CRC Press.
5. Fawcett T (2006), "An introduction to ROC analysis", Pattern Recognition Letters, 27(8): 61–874.
6. Goodfellow I, Bengio Y and Courville A (2016); Deep learning, MIT Press.
7. Hastie T, Tibshirani R and Friedman J (2009); The elements of statistical learning: Data mining, inference, and prediction (2nd ed.), Springer.
8. Horowitz S H and Phadke A G (2014), Power system relaying (4th ed.), Wiley.
9. Kezunovic M, Meliopoulos A P S and Vittal V (2020); Electric power system protection and control, CRC Press.
10. Kotsiantis S B (2007), "Supervised machine learning: A review of classification techniques", Informatica, 31: 249–268.
11. Panchal M, Chatterjee S, Brahmachary R, Godwal S D and Nimje A A (2024), "Detection of power system events using bagged tree classifier", 2024 IEEE 3rd International Conference on Electrical Power and Energy Systems (ICEPES), 1–5: https://doi.org/10.1109/ICEPES60647.2024.10653581.
12. Rahman S, Kumar A and Singh R (2024), "Artificial intelligence applications for fault detection in distribution networks: A review", Energies, 17(4): 1345.
13. Thukaram D, Khincha H P and Vijaynarasimha H P (2005), "Artificial neural network and support vector machine approach for locating faults in radial distribution systems", IEEE Transactions on Power Delivery, 20(2): 710–721.
14. Wen Q, Zhao X and Luo Y (2023), "Deep learning based fault location framework for distribution systems with limited measurements", Electric Power Systems Research, 214: 108981.
15. Zhou Y, Wang J and Liu H (2022), "Convolutional neural network for fault classification in distribution networks", IEEE Access, 10: 45678–45690.