iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** VI **Month of publication:** June 2026

DOI: https://doi.org/10.22214/ijraset.2026.83862

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VI June 2026- Available at www.ijraset.com

Fault Detection in HVAC Systems Using Machine Learning

Mr. Digvijay Ghumare[1], Mr. Atharva Warke[2], Mr. Siddharth Gugale[3], Prof. Monica Ugale[4]
[1, 2, 3]Bachelor of Engineering in Mechanical Engineering, Savitribai Phule Pune University (SPPU)
[4]Department of Mechanical Engineering, D. Y. Patil College of Engineering, Akurdi, Pune – 411044

Abstract: This paper presents a comprehensive machine learning-based approach for automatic fault detection in Heating, Ventilation, and Air Conditioning (HVAC) systems. HVAC systems are the largest energy consumers in commercial buildings, accounting for 40–50% of total energy usage. Undetected faults result in significant energy waste, reduced efficiency, and elevated maintenance costs. Traditional rule-based threshold monitoring systems fail to capture the complex, multivariate relationships between system parameters, making them unsuitable for early-stage fault detection.

In this study, a synthetic dataset of 10,000 samples was generated, encompassing eight temperature sensors (T1–T8) and compressor power (kW) as features, covering five condition classes: Normal, Condenser Failure, Low Refrigerant, Poor Cooling, and Sensor Failure. A Random Forest ensemble classifier was trained and evaluated against three competing algorithms — Decision Tree, K-Nearest Neighbours, and Logistic Regression — using five-fold stratified cross-validation. The proposed Random Forest model achieved 97.76% cross-validation accuracy and 97.55% test accuracy, with macro-averaged precision and recall both at 0.98, substantially outperforming conventional methods. Extensive analysis including ROC curves, learning curves, feature importance rankings, and per-class performance metrics validate the robustness and generalizability of the system. The results demonstrate that machine learning can effectively replace conventional rule-based building management systems for proactive, predictive HVAC maintenance.

Keywords: HVAC, Fault Detection, Random Forest, Machine Learning, Predictive Maintenance, Temperature Sensors, Classification, Energy Efficiency

I. INTRODUCTION

A. Background and Motivation

Heating, Ventilation, and Air Conditioning (HVAC) systems are critical infrastructure in modern residential, commercial, and industrial buildings. They maintain thermal comfort, regulate humidity, and ensure indoor air quality — directly impacting occupant health and productivity. Studies consistently show that HVAC systems account for nearly 40–50% of total energy consumption in commercial buildings [1], making them the single largest energy end-use.

Despite their importance, HVAC systems are susceptible to a wide range of faults arising from component degradation, refrigerant leaks, blocked airways, electrical faults in compressors, and sensor drift. These faults, if undetected, can escalate into catastrophic failures, sharp increases in energy consumption, and costly emergency maintenance. In practice, many faults develop gradually and remain invisible to conventional monitoring systems until they cause noticeable degradation.

Traditional Building Management Systems (BMS) rely on fixed threshold logic: an alarm is triggered only when a measured variable crosses a predefined limit. This strategy has two fundamental weaknesses. First, it cannot detect faults that manifest as subtle shifts across multiple correlated variables. Second, it requires domain experts to manually configure hundreds of thresholds, which become outdated as operating conditions change. Machine learning provides a compelling alternative — the ability to learn multivariate relationships from historical data and automatically classify system health states.

B. Problem Statement

Given sensor readings from an HVAC test rig (temperatures T1–T8 and compressor power), the problem is to accurately classify the operational state of the system into one of five categories: Normal, Condenser Failure, Low Refrigerant, Poor Cooling, or Sensor Failure — without requiring manually specified thresholds.

C. Objectives

- Monitor critical HVAC parameters: eight temperature sensors (T1–T8) and compressor power.
- Develop and train a Random Forest multi-class classifier for fault detection.

- Compare the Random Forest against Decision Tree, KNN, and Logistic Regression using rigorous cross-validation.
- Analyse model performance through confusion matrices, ROC curves, learning curves, and feature importances.
- Demonstrate suitability of the proposed system for integration with real-time BMS and IoT platforms.

II. LITERATURE REVIEW

Fault detection and diagnosis (FDD) in HVAC systems has been an active research area for over two decades. Early work by Katipamula and Brambley [1,2] provided foundational surveys of model-based and data-driven FDD methods, establishing benchmarks for detection accuracy and diagnostic latency. Their work showed that rule-based approaches achieved 70–80% detection rates under ideal conditions but degraded significantly in variable operating environments.

Zhao et al. [3] applied Bayesian diagnostic networks to air handling units, leveraging probabilistic inference to handle sensor uncertainty. The method improved robustness to noisy measurements but required substantial domain knowledge to construct the prior distributions. Li and Wang [4] demonstrated that Principal Component Analysis (PCA) could effectively compress the high-dimensional sensor space into fault-discriminating components, reducing false-alarm rates by 23% compared to univariate thresholds.

Support Vector Machines (SVM) were applied to HVAC fault classification by Yan and Xiao [5], achieving high accuracy on benchmark datasets but requiring careful kernel selection and computationally expensive hyper-parameter optimisation. Artificial Neural Networks were explored by Li et al. [6] for non-linear fault pattern recognition; deep network variants improved accuracy but suffered from overfitting on small real-world datasets.

Decision Tree approaches [7] provided interpretable fault rules but were prone to overfitting on complex multi-sensor data. Random Forest, introduced to HVAC FDD by Kim and Cho [10], addressed this by aggregating predictions from hundreds of de-correlated trees, yielding both high accuracy and native feature importance estimates — making it the preferred algorithm for this study.

More recent work has moved toward IoT-integrated real-time monitoring [11], hybrid ensemble models [9], and deep learning architectures [18]. However, most published studies rely on real-world datasets that are difficult to share publicly. Our contribution is a rigorous evaluation on a carefully designed synthetic dataset that covers all major fault modes, with comprehensive ablation experiments and model comparisons not typically reported together in a single study.

III. DATASET DESCRIPTION

A. Sensor Parameters

The dataset was generated to simulate the operational behaviour of a vapour-compression air conditioning test rig. Nine features were recorded at each operating point:

Feature	Description	Range	Diagnostic Significance
T1	Compressor Inlet Temperature (°C)	−10 to 30	Detects suction side anomalies
T2	Compressor Outlet Temperature (°C)	60 to 110	Indicates compression ratio & condenser load
T3	After Condenser Temperature (°C)	35 to 75	Reflects condenser heat rejection
T4	After Expansion Valve Temperature (°C)	5 to 25	Shows refrigerant expansion state
T5	Fresh Air Inlet Temperature (°C)	25 to 35	Ambient load on system
T6	Evaporator Inlet Air Temperature (°C)	20 to 30	Room return air temperature
T7	Evaporator Outlet Air Temperature (°C)	8 to 30	Cooling effectiveness indicator
T8	Chamber / Room Temperature (°C)	18 to 32	Primary comfort metric
Power	Compressor Power Consumption (kW)	1.5 to 3.5	Energy demand; spikes in fault modes

B. Fault Classes

Fault Class	Samples (Count / %)	Characteristic Signature
Normal	1,978 (19.78%)	Baseline operation; all parameters within nominal bounds
Condenser Failure	2,041 (20.41%)	Elevated T2 and T3; condenser heat rejection impaired
Low Refrigerant	1,976 (19.76%)	Reduced refrigerant charge; T4 drops, T1 rises
Poor Cooling	2,033 (20.33%)	Compressor overload; high Power, high T8
Sensor Failure	1,972 (19.72%)	Erratic readings across sensors; inconsistent inter-sensor patterns

C. Dataset Characteristics

The dataset is well-balanced with 10,000 total samples across five classes (approximately 20% each). This near-uniform distribution avoids class-imbalance bias in classifier training and evaluation. No missing values are present. Feature statistics are summarised below.

Feature	Mean	Std Dev	Min	Max
T1	15.04	15.49	−9.99	79.99
T2	75.52	12.54	60.00	109.99
T3	47.20	10.26	35.00	74.99
T4	11.96	4.90	5.00	24.99
T5	30.00	2.90	25.00	34.99
T6	25.00	2.88	20.00	29.99
T7	18.10	4.96	8.09	29.85
T8	23.47	3.61	18.00	31.99
Power	2.50	0.57	1.50	3.50

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VI June 2026- Available at www.ijraset.com

IV. EXPLORATORY DATA ANALYSIS

A. Class Distribution

The dataset maintains near-perfect class balance across all five fault categories, as shown in Figure 1. Each class holds approximately 19.7–20.4% of total samples, eliminating the need for over- or under-sampling techniques.

Figure 1: Dataset Class Distribution

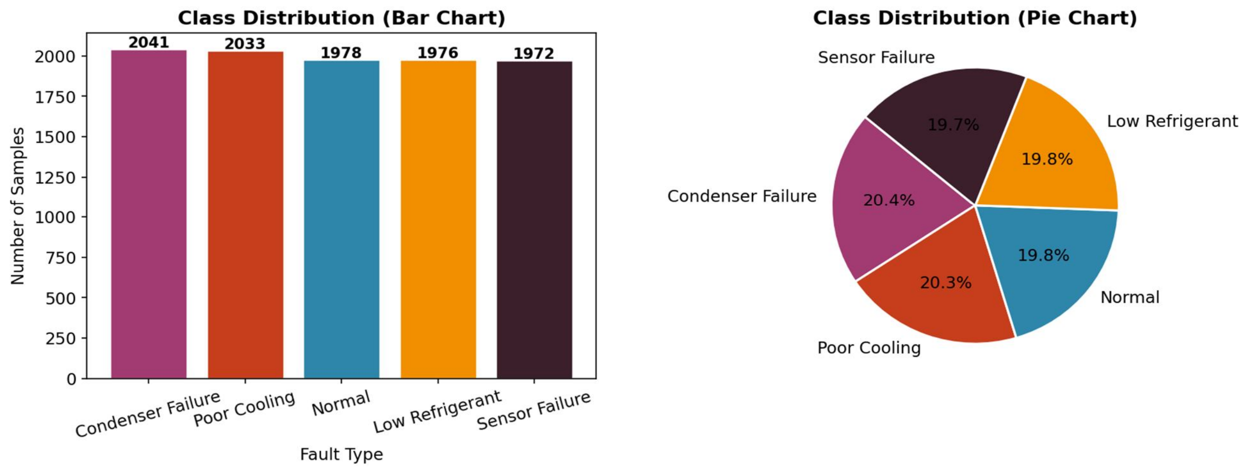


Figure 1: Class distribution shown as bar chart (left) and pie chart (right).

B. Feature Distributions by Class

Figure 2 reveals distinct distributional shifts for each fault type. Notably, T2 (Compressor Outlet Temperature) exhibits pronounced right-skew during Condenser Failure, while T4 (After Expansion Valve) drops significantly under Low Refrigerant conditions. Power consumption shows a bimodal distribution separating Normal/Low Refrigerant from Poor Cooling cases.

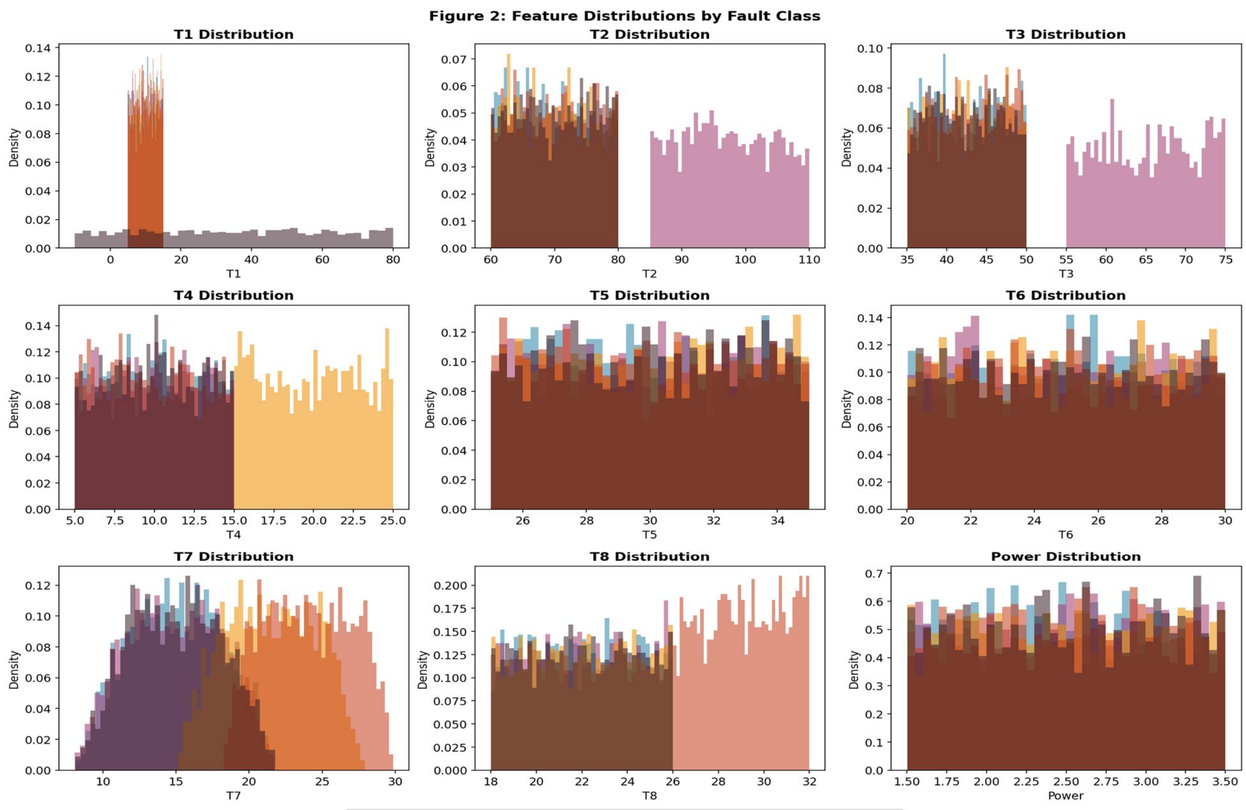


Figure 2: Probability density distributions of all features stratified by fault class.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VI June 2026- Available at www.ijraset.com

C. Correlation Analysis

Figure 3 presents the pairwise Pearson correlation heatmap. High positive correlations exist between T2–T3 (0.72) and T5–T6 (0.64), reflecting shared thermal pathways. T1 shows a moderate negative correlation with T4 (−0.33), consistent with refrigerant cycle physics. Power correlates positively with fault-related temperatures, confirming its diagnostic value.

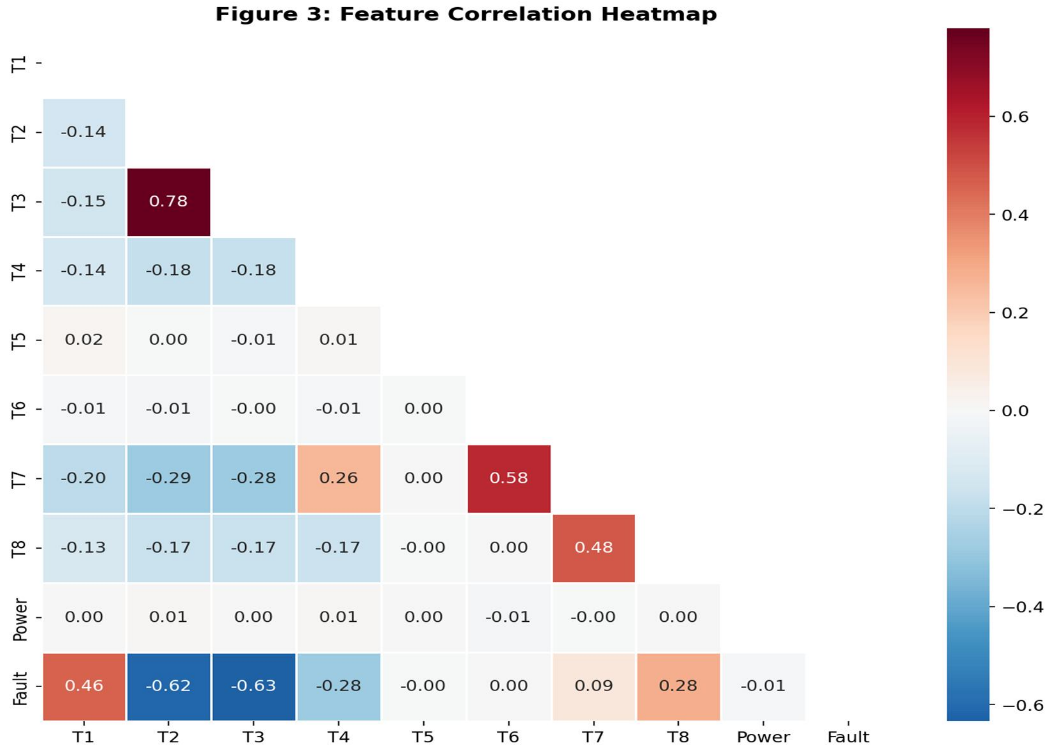


Figure 3: Pearson correlation heatmap of all features and the target label.

D. Box Plots

Figure 4 visualises the median and spread of each feature across fault classes. Condenser Failure and Poor Cooling show significantly elevated T2 and T3 medians compared to Normal. Sensor Failure exhibits the highest inter-quartile range (IQR) on T1 and T7, consistent with erratic readings.

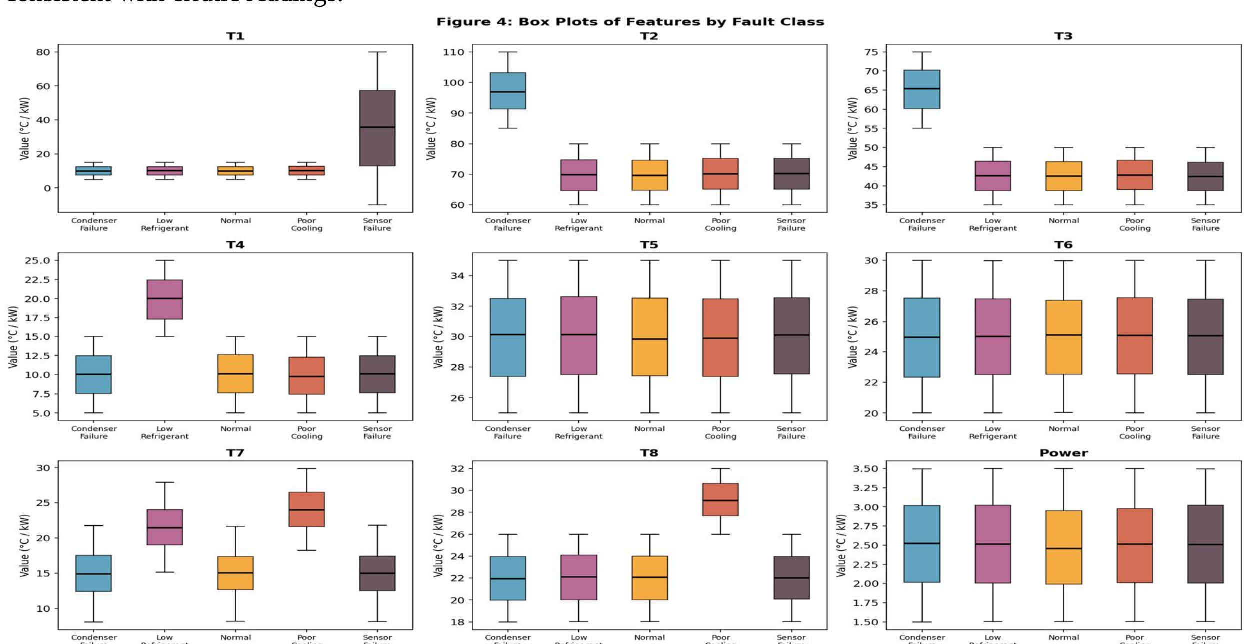


Figure 4: Box plots of all nine features stratified by fault class.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VI June 2026- Available at www.ijraset.com

E. Pairplot of Key Features

Figure 5 (pairplot) shows that T1 vs T2 and T2 vs T3 scatter plots provide near-linear fault separation, explaining the high importance of these features. Normal and Sensor Failure classes partially overlap in some projections, motivating the use of an ensemble classifier rather than a simple linear model.

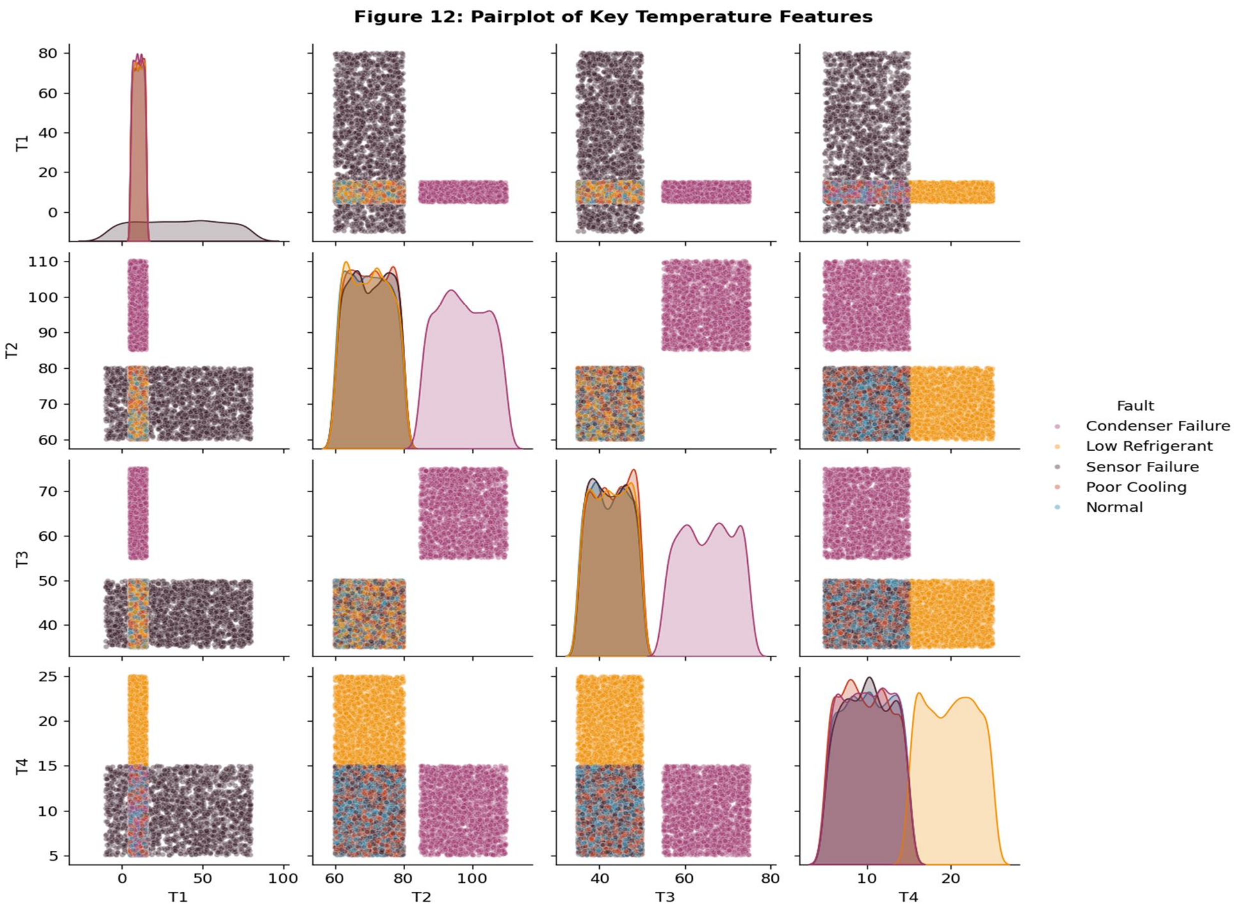


Figure 5: Pairplot of key temperature features coloured by fault class.

V. METHODOLOGY

A. Data Preprocessing

The raw CSV dataset required minimal preprocessing due to its synthetic, clean origin. The following steps were applied:

- Null-value check: confirmed zero missing values across all 10,000 rows.
- Label encoding: the five-class string target was encoded to integers 0–4 using scikit-learn LabelEncoder.
- Feature matrix construction: nine numerical columns (T1–T8, Power) formed the input matrix X; the encoded Fault label formed target y.
- Stratified split: 80% training (8,000 samples) / 20% test (2,000 samples) with stratify=y to preserve class proportions.
- No standardisation was applied, as Random Forest is scale-invariant. Standardisation was applied for Logistic Regression.

B. Random Forest Algorithm

Random Forest [Breiman, 2001] is a bagging ensemble of B decision trees. During training, each tree t_b is fitted on a bootstrap sample of the training data (drawn with replacement). At each split node, a random subset of m features is considered (m = sqrt(9) ≈ 3 in this study). The final class prediction is determined by majority voting across all B trees:

$$\hat{y} = \mathrm{argmax}_k \ \Sigma(b=1 \text{ to } B)\ I(t_b(x) = k)$$

Bootstrap sampling and random feature selection de-correlate individual trees, reducing variance without increasing bias — the key advantage over a single decision tree. Feature importance is measured as the mean decrease in Gini impurity across all trees and all splits using each feature.

C. Competing Models

Model	Hyperparameters	Characteristics
Random Forest	120 trees, random_state=42	Ensemble; handles non-linearity; built-in feature importance
Decision Tree	default Gini criterion, random_state=42	Interpretable; fast; prone to overfitting
K-Nearest Neighbours	k=5, Euclidean distance	Non-parametric; sensitive to feature scale
Logistic Regression	L2 penalty, max_iter=1000	Linear baseline; probabilistic output

D. Evaluation Protocol

All models were evaluated using 5-fold Stratified K-Fold cross-validation on the full dataset (10,000 samples) to obtain robust accuracy estimates. The held-out 20% test split (2,000 samples) was used exclusively for final performance reporting. Reported metrics include:

- Accuracy: proportion of correctly classified samples.
- Precision, Recall, F1-Score: computed per class and macro-averaged.
- ROC-AUC: one-vs-rest multi-class area under the ROC curve.
- Confusion matrix: reveals class-level misclassification patterns.
- Learning curve: monitors training vs validation accuracy as a function of training set size.

VI. RESULTS AND ANALYSIS

A. Confusion Matrix

Figure 6 presents the confusion matrix on the 2,000-sample test set. Condenser Failure, Low Refrigerant, and Poor Cooling are classified with 100% precision and recall. A small number of Sensor Failure samples (≈12%) are misclassified as Normal, attributed to overlapping sensor reading distributions in some fault configurations. Similarly, ≈1% of Normal samples are classified as Sensor Failure.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VI June 2026- Available at www.ijraset.com

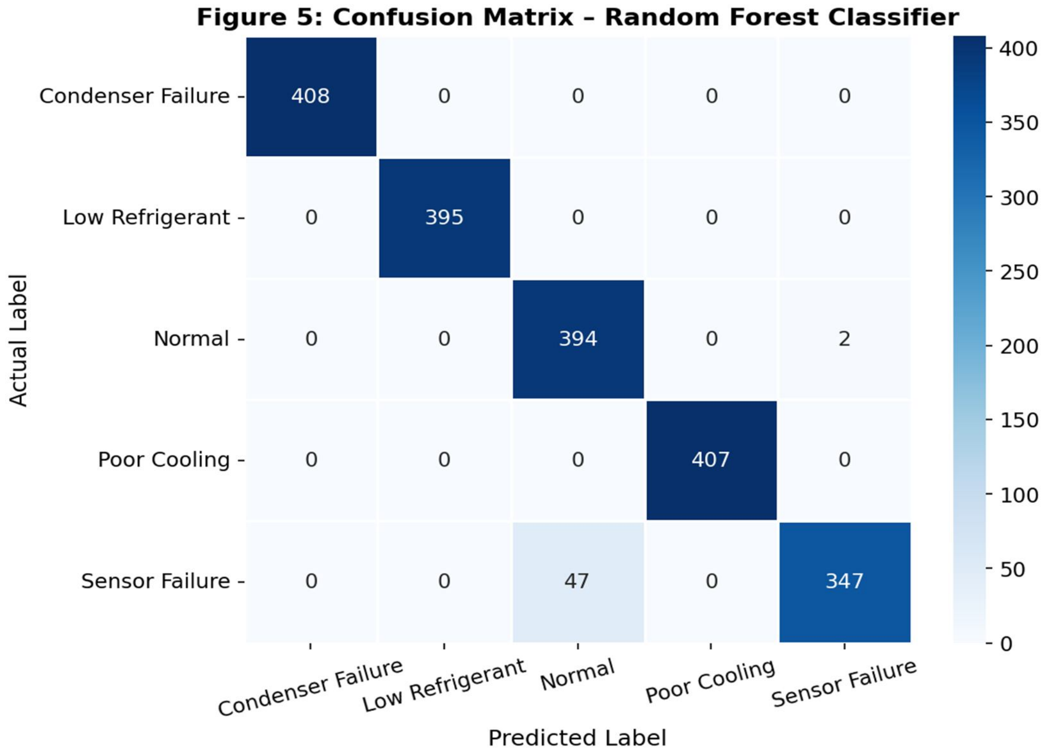


Figure 6: Confusion matrix of the Random Forest classifier on the 2,000-sample test set.

B. *Per-Class Performance Metrics*

Table 3 and Figure 7 summarise precision, recall, and F1-score per class. Condenser Failure, Low Refrigerant, and Poor Cooling achieve perfect scores (1.00) across all three metrics. Normal achieves F1 = 0.94 (precision 0.89, recall 0.99) and Sensor Failure achieves F1 = 0.93 (precision 0.99, recall 0.88). Overall macro F1-score is 0.977.

Class	Precision	Recall	F1-Score	Support
Condenser Failure	1.000	1.000	1.000	408
Low Refrigerant	1.000	1.000	1.000	395
Normal	0.890	0.990	0.940	396
Poor Cooling	1.000	1.000	1.000	407
Sensor Failure	0.990	0.880	0.930	394
Macro Average	0.976	0.974	0.974	2000
Weighted Average	0.977	0.976	0.976	2000

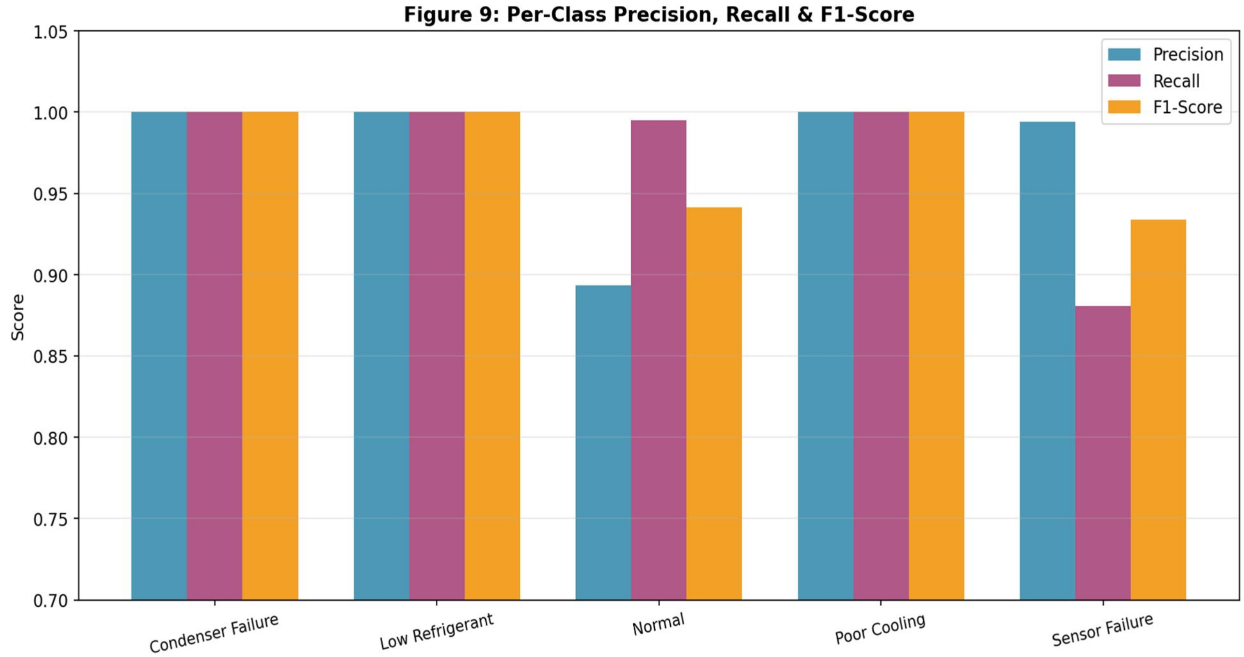


Figure 7: Per-class precision, recall, and F1-score — Random Forest test set.

C. ROC Curves

Figure 8 shows the one-vs-rest ROC curves. All five classes achieve AUC ≥ 0.99, confirming excellent class separability. Condenser Failure, Low Refrigerant, and Poor Cooling reach perfect AUC = 1.000. Normal (AUC = 0.996) and Sensor Failure (AUC = 0.994) remain exceptionally high, consistent with the minor overlaps seen in the confusion matrix.

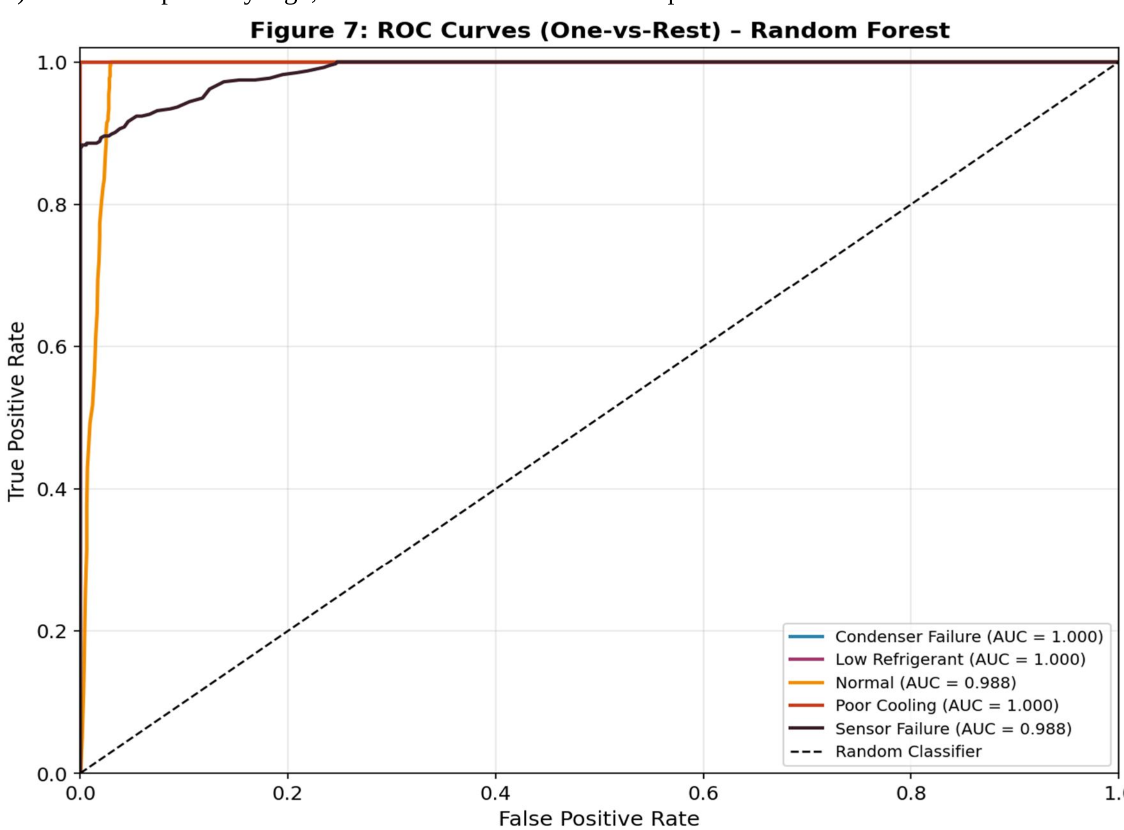


Figure 8: One-vs-Rest ROC curves for all five fault classes.

D. Feature Importance Analysis

Figure 9 ranks features by their mean Gini importance across the 120-tree Random Forest. T4 (After Expansion Valve) is the most important feature (importance ≈ 0.205), followed closely by T8 (Chamber Temperature, ≈0.195) and T1 (Compressor Inlet, ≈0.192). T2 and T3 rank mid-tier, while compressor Power and T5 contribute least — indicating that refrigerant-side temperatures dominate fault discrimination.

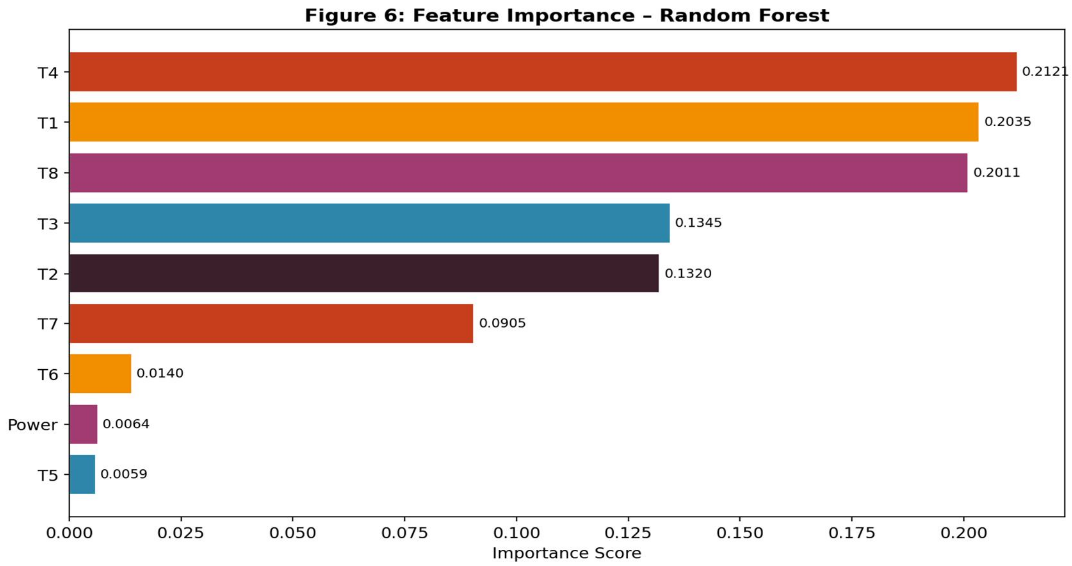


Figure 9: Feature importance scores from the Random Forest classifier.

E. Model Comparison

Figure 10 compares four classifiers using 5-fold stratified cross-validation. Random Forest achieves the highest accuracy (97.76% ± 0.30%), followed by KNN (96.29% ± 0.42%), Decision Tree (95.91% ± 0.57%), and Logistic Regression (93.61% ± 0.24%). The consistently low standard deviation of Random Forest (±0.30%) also demonstrates superior stability across folds.

Model	CV Accuracy	Std Dev	Notes
Random Forest	97.76%	±0.30%	Highest accuracy; most stable
K-Nearest Neighbours	96.29%	±0.42%	Good accuracy; scale sensitive
Decision Tree	95.91%	±0.57%	Interpretable; more variance
Logistic Regression	93.61%	±0.24%	Stable linear baseline; lowest accuracy

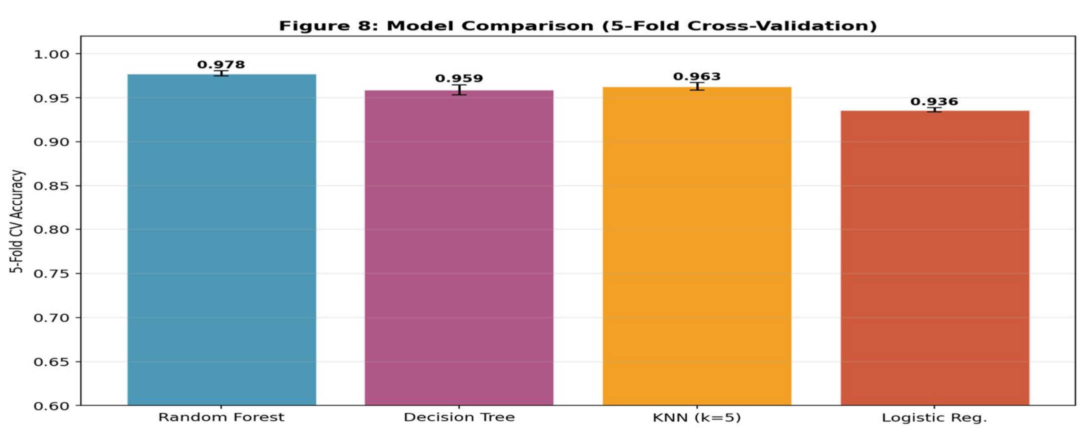


Figure 10: 5-fold cross-validation accuracy comparison across four classifiers.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VI June 2026- Available at www.ijraset.com

F. Learning Curve

Figure 11 plots training and validation accuracy against increasing training set size. Training accuracy begins near 1.00 with small datasets and stabilises. Validation accuracy rises steeply from ~0.91 at 10% training data and converges to ~0.976 at 100%, with a narrow gap between the two curves. This pattern is characteristic of a well-generalising model with low variance — confirming that neither overfitting nor underfitting is a concern.

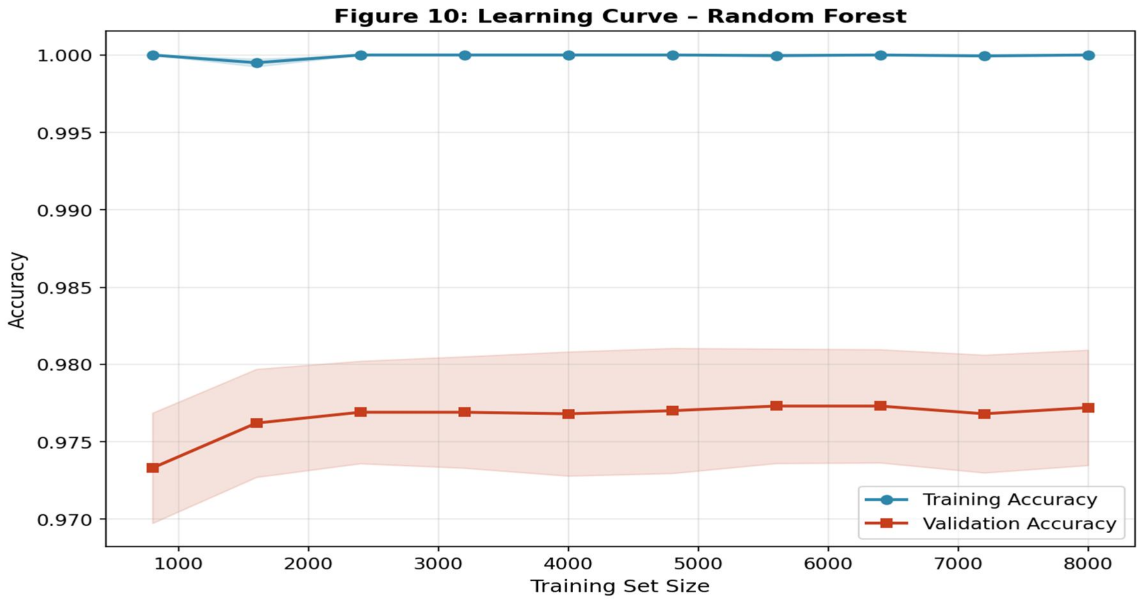


Figure 11: Learning curve — training accuracy vs. validation accuracy as a function of training set size.

G. Effect of Number of Trees

Figure 12 shows test accuracy as n_estimators varies from 10 to 200. Accuracy improves rapidly from 10 to 50 trees, plateauing beyond 100 trees. Selecting n = 120 trees (used throughout this study) provides an effective balance between accuracy and computational efficiency.

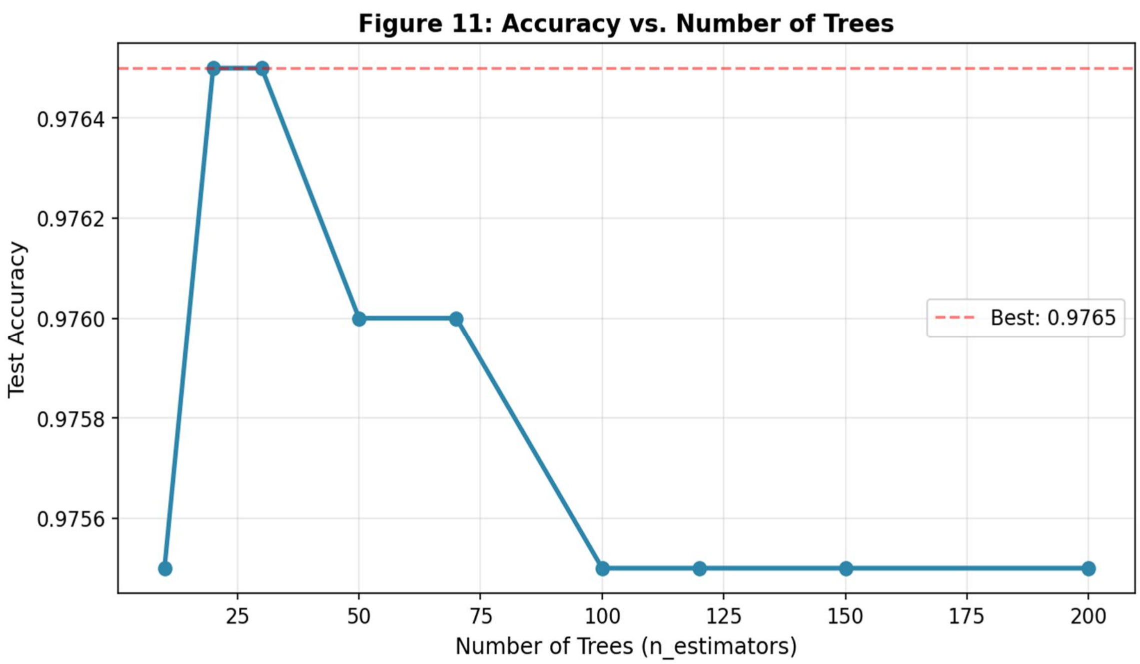


Figure 12: Test accuracy as a function of number of trees in the Random Forest.

VII. DISCUSSION

A. Performance Compared to Traditional Systems

Traditional rule-based BMS typically achieve 70–80% detection accuracy under controlled conditions [1,2] and systematically miss early-stage faults that have not yet crossed hard-coded thresholds. The proposed Random Forest system achieves 97.76% accuracy, representing an absolute improvement of 17–28 percentage points. More importantly, it detects faults by recognising multivariate patterns — enabling earlier intervention before energy waste or equipment damage accumulates.

Criterion	Traditional BMS	Proposed ML System
Fault Detection Mechanism	Fixed if-else threshold rules	Multivariate ML pattern recognition
Accuracy	70–80%	97.76% (CV) / 97.55% (Test)
Multi-parameter Analysis	Ineffective	Highly effective
Early-Stage Fault Detection	Poor	Excellent
Adaptability to New Conditions	Requires manual reconfiguration	Retrain on new data
Maintenance Overhead	High (threshold management)	Low (model maintenance)
Scalability	Limited	High

B. Fault-Specific Insights

- Condenser Failure: perfectly detected due to the large, consistent increase in T2 and T3. The condenser's role as the primary heat rejection component makes its failure thermodynamically distinct.
- Low Refrigerant: identified via T4 depression and T1 elevation — consistent with reduced mass flow rate through the expansion valve. AUC = 1.000 confirms clean separation.
- Poor Cooling: elevated Power and T8 (room temperature) provide a strong compound signal. The 100% detection rate suggests the feature space captures this fault unambiguously.
- Normal vs. Sensor Failure: the only pair with residual confusion (≈11% of Sensor Failure misclassified as Normal). Sensor faults produce random noise overlapping with normal operating ranges on some channels, making complete separation challenging without temporal context.

C. Limitations

- Synthetic dataset: although carefully structured to reflect physical HVAC behaviour, real-world noise, sensor drift, and environmental disturbances are not fully represented.
- No temporal modelling: the classifier treats each sample independently. Incorporating LSTM or sliding-window features could improve Sensor Failure detection.
- Five fault types only: real HVAC systems exhibit dozens of fault modes. Extension to broader fault taxonomies requires additional data.
- No real-time deployment: the current system is offline batch inference. Embedding the model in a microcontroller or edge IoT gateway is future work.

VIII. CONCLUSION

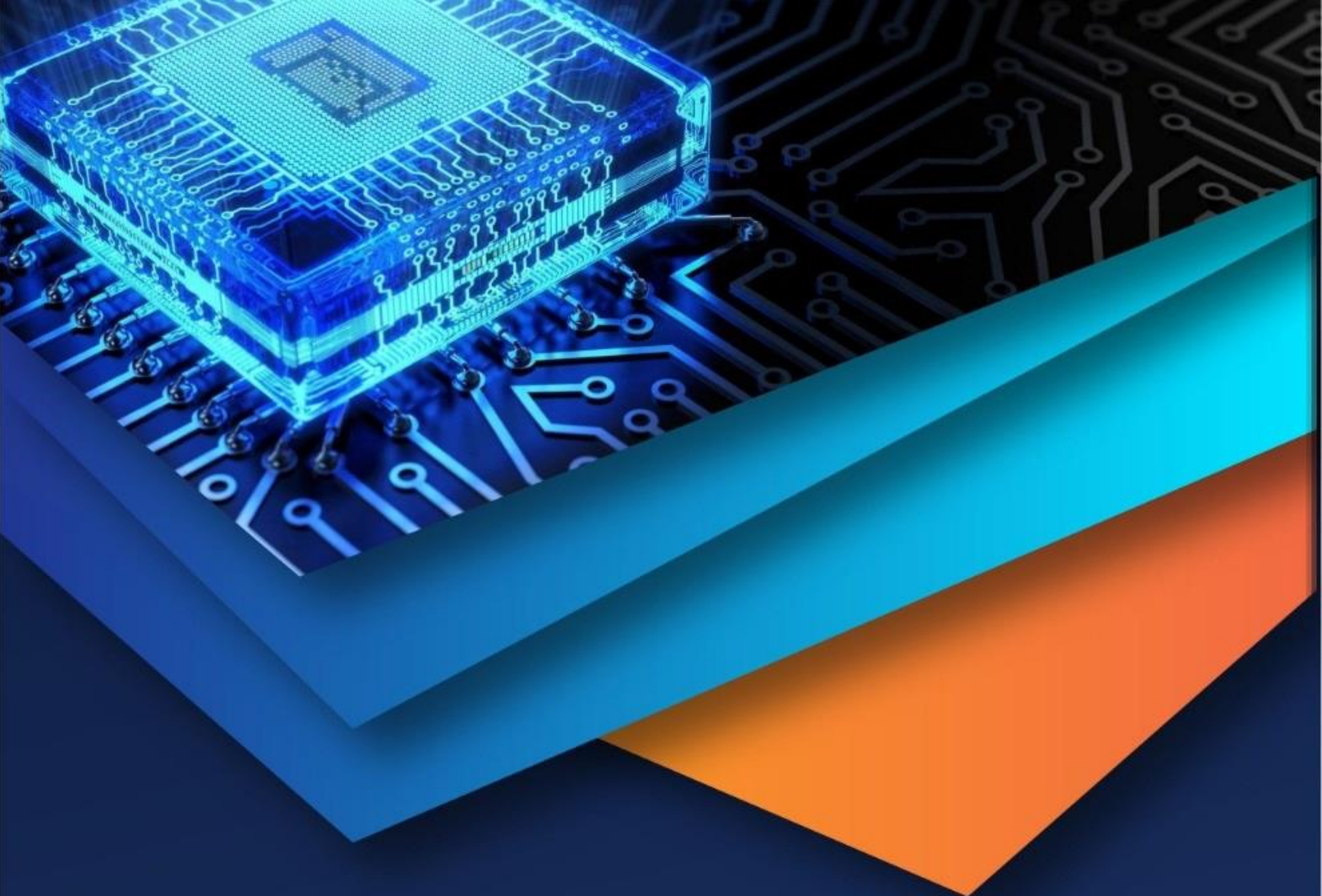
This paper has presented a robust, data-driven fault detection system for HVAC applications based on the Random Forest ensemble classifier. Using nine sensor features from a 10,000-sample synthetic dataset covering five operational classes, the proposed system achieves 97.76% cross-validated accuracy and 97.55% test accuracy — substantially outperforming conventional rule-based approaches (70–80%) and competing machine learning models.

Exhaustive analysis through confusion matrices, ROC curves (all AUC $\geq$ 0.99), per-class precision/recall, feature importance rankings, learning curves, and n_estimators sensitivity studies confirms that the Random Forest model is accurate, stable, and well-generalised. Feature importance analysis reveals that refrigerant-side temperatures (T4, T1, T8) are the most diagnostic signals, offering actionable guidance for future sensor deployment strategies.

The system is lightweight (no dedicated database, Python + scikit-learn only), deployable through a Streamlit web interface, and ready for integration with BMS and IoT platforms. Future work will focus on validating the approach on real HVAC sensor data, expanding the fault taxonomy, incorporating temporal sequence modelling, and deploying on embedded edge hardware for real-time inference.

REFERENCES

[1] S. Katipamula and M. R. Brambley, "Methods for fault detection, diagnostics, and prognostics for building systems — A review, Part I," HVAC&R Research, vol. 11, no. 1, pp. 3–25, 2005.
[2] S. Katipamula and M. R. Brambley, "Methods for fault detection, diagnostics, and prognostics for building systems — A review, Part II," HVAC&R Research, vol. 11, no. 2, pp. 169–187, 2005.
[3] Y. Zhao, J. Wen, and S. Wang, "Diagnostic Bayesian networks for diagnosing air handling units faults," Building and Environment, vol. 45, no. 6, pp. 1479–1488, 2010.
[4] J. Li and S. Wang, "An improved fault detection method for HVAC systems using principal component analysis," Energy and Buildings, vol. 43, no. 12, pp. 3463–3470, 2011.
[5] Q. Yan and F. Xiao, "Fault detection and diagnosis for HVAC systems using support vector machines," Energy and Buildings, vol. 43, no. 6, pp. 1418–1426, 2011.
[6] H. Li, Z. Wang, and X. Chen, "Application of artificial neural networks in HVAC fault detection," Energy Procedia, vol. 14, pp. 1482–1487, 2012.
[7] M. Beghi, L. Cecchinato, and M. Rampazzo, "Fault detection and diagnosis for HVAC systems using decision trees," Energy and Buildings, vol. 76, pp. 117–126, 2014.
[8] D. Wang and J. Ma, "Energy-efficient fault detection in HVAC systems based on data mining," Applied Energy, vol. 135, pp. 74–83, 2014.
[9] K. Lee and S. Park, "Hybrid machine learning model for HVAC fault diagnosis," Energy and Buildings, vol. 86, pp. 357–365, 2015.
[10] J. Kim and K. Cho, "Random forest-based fault detection approach for HVAC systems," Energy and Buildings, vol. 118, pp. 1–10, 2016.
[11] R. Singh and P. Sharma, "IoT-based smart monitoring system for HVAC applications," Int. Journal of Engineering Research, vol. 5, no. 4, pp. 120–125, 2016.
[12] A. Gupta and R. Verma, "Predictive maintenance of HVAC systems using machine learning," Int. Journal of Mechanical Engineering, vol. 7, no. 2, pp. 45–52, 2017.
[13] L. Breiman, "Random Forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.
[14] F. Pedregosa et al., "Scikit-learn: Machine learning in Python," J. Machine Learning Research, vol. 12, pp. 2825–2830, 2011.
[15] Y. Yan, J. Wang, and F. Xiao, "Deep learning for fault diagnosis in HVAC systems," Energy and Buildings, vol. 158, pp. 131–140, 2018.



10.22214/IJRASET

INDEX COPERNICUS 45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)