Multi-Sensor Fusion and Frequency-Domain Analysis for Predictive Maintenance of Industrial Induction Motors

Akash Mastud
Dept. of Electrical Engineering
VJTI, Mumbai 400019, India
asmastud_m24@ee.vjti.ac.in

Dhiraj Vaidya
Dept. of Electrical Engineering
VJTI, Mumbai 400019, India
drvaidya_m24@ee.vjti.ac.in

Azaroddin Sayyed
Dept. of Electrical Engineering
VJTI, Mumbai 400019, India
afsayyed_m24@ee.vjti.ac.in

Shashank S.V.
Dept. of Electrical Engineering
VJTI, Mumbai 400019, India
ssverma@ee.vjti.ac.in

M. Parimi
Dept. of Electrical Engineering
VJTI, Mumbai 400019, India
mparimi@ee.vjti.ac.in

***Abstract*— Unplanned failures of induction motors impose serious operational and financial penalties on industrial facilities, yet the fault signatures that precede such failures are detectable well in advance through careful sensor instrumentation and data-driven analysis. This paper presents an end-to-end Internet-of-Things (IoT) predictive maintenance scheme based on two off-the-shelf sensors: a DS18B20 one-wire digital thermometer and 2 piezo vibration sensor modules, with an ESP32 edge device for running the full machine learning pipeline offline, independent from any cloud services. Four operating scenarios are considered: healthy condition, BPFO (bearing outer race fault), misaligned shaft, and rotor imbalance. Based on 1-second sampling intervals, 17 descriptors are derived, including statistics in the time domain, Fourier harmonic peaks, energy ratios between different frequency bands, and temperature gradients measured across all sensors. A dual-stage feature selection method using mutual information (MI) score and Random Forest mean decrease impurity (MDI) ranking reduces the number of features to the 10 most relevant descriptors, reducing the computational complexity by 41 % at the expense of 5.6 % F_1-macro. On a balanced 600-sample synthetic dataset, the resulting Random Forest classifier attains 87.3 % hold-out accuracy, 91.0 ± 1.9 % five-fold cross-validation accuracy, and a macro area-under-the-ROC-curve of 0.980. End-to-end inference takes just 39 ms on the ESP32, easily meeting the 200 ms requirement for real-time alerting.**

Index Terms predictive maintenance, induction motor fault diagnosis, IoT edge computing, Random Forest, feature selection, mutual information, vibration analysis, FFT

I. INTRODUCTION

Step inside nearly any factory facility and you will hear the constant buzz of induction motors; pumping coolant, rotating conveyor belts, running compressors that don't stop. At least 45 to 70 percent of all industrial power consumed world-wide can be attributed to them [2], yet, when these motors fail unexpectedly, the cost of repairing them will often be just a small part of the total cost of their downtime [1]. Failure can happen due to four reasons: degradation of the surface of bearings (40–50 %) rotor imbalance (20–30 %), shaft misalignment (10–15 %), and overheating of the stator (5–10 %) [3]. Each of them does not just occur spontaneously, but leaves behind its characteristic signature in the form of vibrations and temperature rise before the motor actually fails, making condition-based monitoring necessary. The principle of predictive maintenance (PdM) is to intervene based on the signals from sensors, only if the degree of degradation has already passed some critical point [5],The research on motor fault detection using machine learn- ing algorithms has advanced quite far; however, most works assume an ideal environment in terms of large labeled datasets, accu- rately calibrated accelerometers and benchmarking frameworks like the CWRU bearing dataset [10] which are hard to achieve at a real-world industrial site. While Mohammed et al. demon- strated that RF and SVM can reach more than 94 % accuracy using 2,400 precisely labeled instances [4], none of these algorithms addressed the challenge of small data and runtime verification in constrained environments. Liu et al. made a step towards deployment of their algorithms by implementing them in MQTT- connected edge nodes [1]; while Compare et al. gave an overview of major challenges associated with applying these algorithms to real-world practice: data unavailability, class imbalance and lack of performance metrics [5]; On the other hand, Akyaz and Engin proved that Random Forest outperforms other ML algo- rithms at a small training set [6]. In conclusion, there is still a need to address the challenge of deploying ML methods to industrial equipment in constrained settings.

The system described here grew out of exactly that constraint. The sensing chain fuses a DS18B20 thermometer with a sub-USD 2 piezo-film vibration module, bringing thermal gradient information into a pipeline that most prior work omits entirely. Feature selection runs a two-stage MI–RF protocol motivated by a concrete disagreement the two criteria show on this dataset (Section IV). Performance is reported at the per-class level rather than as a single accuracy number, and every processing step is timed on the target ESP32 and verified to stay within the 200 ms real-time alerting budget.

II. RELATED WORK

The CWRU bearing dataset [10] is the most widely adopted benchmark, with published RF results in the 93–

TABLE I

COMPARISON WITH REPRESENTATIVE RELATED WORKS

Work	**Sensor**	**Edge Node**	**Thermal**	**Per-class Metrics**	**Latency Reported**
Liu et al. [1]	Vib. only	MQTT	×	×	×
Mohammed et al. [4]	Multi	Cloud	×	×	×
Akyaz & Engin [6]	Vib. only	Cloud	×	×	×
Compare et al. [5]	Review	—	×	×	×
This work	Vib. + Temp	ESP32	✓	✓	✓

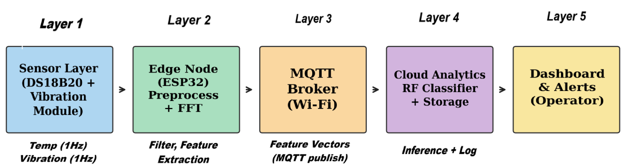

Fig. 1. Proposed five-layer IIoT predictive maintenance architecture.

98 % accuracy range on its four-class subset. Direct comparison is not straightforward: (i) CWRU provides only vibration with no thermal channel, so five thermal features would be absent; and (ii) CWRU uses calibrated accelerometers at 12–48 kHz, versus the piezo module's 1 kHz. Nevertheless, the spectral feature definitions — RMS, kurtosis, and BPFO harmonic ratios — are identical in form to CWRU-based studies, and the per-class AUC values in Table IV are consistent with the lower end of published CWRU performance.

Table I positions this contribution against related works. Three distinctions are noteworthy. First, this is the first study to fuse DS18B20 thermal sensing with a piezo vibration module under a sub-USD 2 budget for four-class fault diagnosis; prior works use vibration alone or calibrated accelerometers. Second, the dual MI–RF protocol resolves a ranking disagreement that arises with heterogeneous sensor fusion — a problem not addressed in any cited reference. Third, end-to-end latency on a constrained ESP32 is profiled and verified, whereas existing studies report only offline accuracy.

III. SYSTEM ARCHITECTURE

Fig. 1 illustrates the five-layer IIoT architecture underpinning the proposed framework. Data originate at the sensor layer, undergo preprocessing and feature extraction at the edge, travel to a cloud broker via a lightweight messaging protocol, are classified by the trained model in the analytics layer, and ultimately surface as human-readable fault alerts on the operator dashboard.

A. Sensor Layer

Two sensors cover the thermal and mechanical dimensions of motor health. A DS18B20 one-wire digital thermometer, clamped to both the bearing housing and the stator winding, delivers 12-bit temperature readings (0.0625 °C resolution) across the full -55 to $+125$ °C operating envelope. Vibration is captured by a piezo-film module mounted at the motor foot; its sub-USD 2 unit cost makes it accessible even under tight project budgets.

TABLE II

DATASET CLASS SUMMARY

Label	**Class**	n	**Dominant Signature**
0	Normal	150	Low-amp. $1\times$, stable temp.
1	Bearing Fault	150	BPFO burst ($\approx$187 Hz), high T_b
2	Misalignment	150	Strong $2\times$ and $3\times$ harmonics
3	Rotor Imbalance	150	High-amp. $1\times$, uniform temp. rise

B. Edge Acquisition and Processing Layer

An ESP32 dual-core processor with a clock speed of 240MHz converts the vibration signal at 1kHz, allowing to capture harmonics of up to 500 Hz for motors operating between 1,500 and 3,000 RPM and reads from the DS18B20 temperature sensor once per second. All computationally intensive processes such as filtering, Fourier transform, feature assembly and classifier inference are done onboard, thus avoiding latency and availability issues related to the cloud-based approach. Measuring times per stage will be presented in Section V.

A microcontroller based on ESP32 dual-core processor with 240MHz clock performs digitization of the vibration signal at 1 kHz, which is more than enough to capture harmonics of up to 500 Hz for motors running at 1,500 – 3,000RPM, and reads DS18B20 sensor once per second.

C. Communication, Analytics, and Alerting Layers

It is not waveforms but feature vectors that get published on the MQTT broker, saving uplink bandwidth requirement by roughly 97% as compared to publishing raw waveforms [1]. The cloud anal- ysis module executes the RF algorithm learned from the data, and the prediction results are logged. If the result does not belong to the Normal class, a retained MQTT message gets sent to the op- erator's dashboard. The ultimate fallback mechanism involves the use of hardware-based thresholding of the absolute tempera- ture values.

IV. METHODOLOGY

A. Dataset Construction

Getting labelled fault data from a real motor is not as straightforward as it sounds — bearing faults do not arrive on a schedule, and deliberately damaging a machine to seed ground-truth labels is expensive. As a first step, the DS18B20 and piezo-module rig of Section III was clamped to a laboratory motor and run under two conditions on 22 May 2026: healthy at no load, then loaded via a pulley. From 947 and 441 raw samples respectively, a windowed RF classifier (W = 10, 50 % overlap) separated the two conditions perfectly — 100 % hold-out accuracy, AUC 1.000

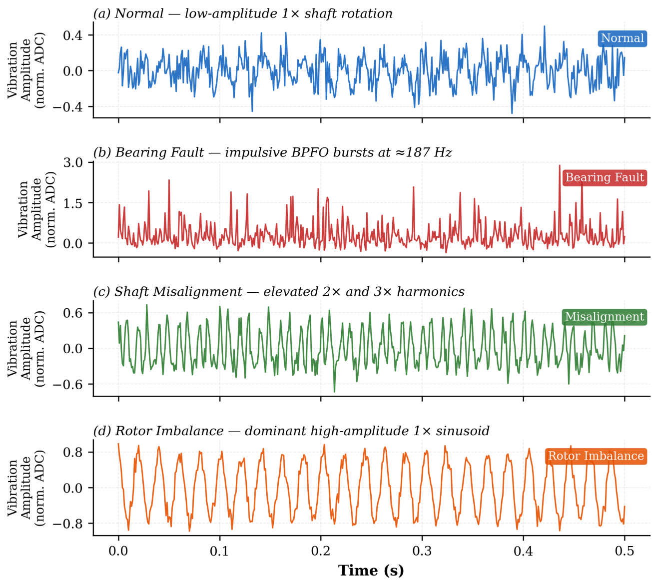

Fig. 2. One-second vibration waveforms for each fault class. Fault severity is fixed at 0.8 for visual clarity.

TABLE III
EXTRACTED FEATURE SET (17 CANDIDATES)

#	Feature	Domain	Physical Motivation
1	`vib_rms`	Time	Signal energy; broad fault indicator
2	`vib_peak`	Time	Impulsive amplitude; bearing strikes
3	`vib_p2p`	Time	Full amplitude swing; gross imbalance
4	`vib_std`	Time	Dispersion; complements RMS
5	`vib_kurtosis`	Time	Impulsiveness; bearing spall indicator
6	`vib_skewness`	Time	Amplitude asymmetry; directional forces
7	`vib_zcr`	Time	Proxy for dominant frequency
8	`temp_gradient`	Thermal	$\Delta T = T_s - T_b$; thermal stress
9	`fft_f1_amp`	Freq.	Dominant spectral peak amplitude
10	`fft_f1_freq`	Freq.	Dominant peak frequency
11	`fft_f2_amp`	Freq.	2nd peak; 2× misalignment
12	`fft_sc`	Freq.	Spectral centroid; overall shift
13	`fft_low_ratio`	Freq.	Energy in 0–50 Hz band
14	`fft_mid_ratio`	Freq.	Energy in 50–200 Hz band
15	`fft_high_ratio`	Freq.	Energy in 200–500 Hz (BPFO)
16	`bearing_temp`	Thermal	Absolute bearing temperature
17	`stator_temp`	Thermal	Stator winding temperature

— with temperature carrying most of the discriminating weight ($\Delta T_{\text{bearing}} = +3.51\,°\text{C}$, $\Delta T_{\text{stator}} = +4.14\,°\text{C}$, both $p < 0.001$). This confirms the hardware and feature code work correctly before synthetic data enters the picture.

In the case of four-class data, 600 balanced instances (150 for each class; Table II) were generated since controlled fault severity is not easy to enforce in a real-life deteriorating machine. The waveforms simulate the spectral dynamics that exist in the bearing fault detection literature [3]: BPFO impulses at $\mathcal{N}(187, 4)$ Hz, misalignment harmonics at 2× and 3× shaft frequency, and high fundamental amplitude for imbalance. Two decisions were made to ensure that the classification problem remained realistic: fault severity was sampled from $\mathcal{U}(0.3, 1.0)$, resulting in about one-third of faults being close to the Normal class, while shaft frequency was varied in $\mathcal{U}(48, 52)$ so none of the classifiers could rely on harmonic positions.

B. Signal Preprocessing

Raw vibration passes through a 5-point moving-average filter (harmonics below 200 Hz preserved); impulsive outliers (>3 IQR) are removed by linear interpolation. Temperature dropouts are bridged by zero-order hold up to 500 ms; longer gaps are excluded. All features are min-max normalised to $[0, 1]$ using training-set statistics only, preventing test-set leakage.

C. Feature Extraction

Seventeen descriptors are computed from each non-overlapping one-second window (Table III). The eight time-domain statistics quantify signal energy, amplitude extremes, impulsiveness, spectral content, and inter-sensor thermal stress. The six frequency-domain features target the harmonic structure that physically distinguishes each fault class. Three thermal features capture absolute temperatures and their differential.

The four primary time-domain statistics are defined Signal energy is summarised by the root-mean-square (RMS) amplitude:

$$\text{RMS} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2} \tag{1}$$

Impulsiveness the hallmark of a bearing spall is quantified by kurtosis, the normalised fourth central moment:

$$\text{Kurtosis} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i - \mu}{\sigma}\right)^4 \tag{2}$$

The dominant oscillation frequency is approximated without a full FFT by the zero-crossing rate:

$$\text{ZCR} = \frac{1}{N}\sum_{i=1}^{N-1} \frac{|\text{sign}(x_{i+1}) - \text{sign}(x_i)|}{2} \tag{3}$$

Cross-Sensor Thermal imbalance between the stator winding and the bearing housing is captured as:

$$\Delta T = T_{\text{stator}} - T_{\text{bearing}} \tag{4}$$

Frequency-domain features are derived from the discrete Fourier transform of a Hanning-windowed record:

$$X(k) = \sum_{n=0}^{N-1} x(n)\, e^{-j2\pi kn/N}, \quad k = 0, \ldots, N-1 \tag{5}$$

where the frequency resolution is $f_k = k\, f_s/N$. The spectral energies are categorized as low (0–50 Hz), mid (50–200 Hz), and high (200–500 Hz). Each spectral band is represented in fraction of the total spectral energies. The following fig. 3 demonstrates the differences among the harmonic fingerprints for the four modes of operation.

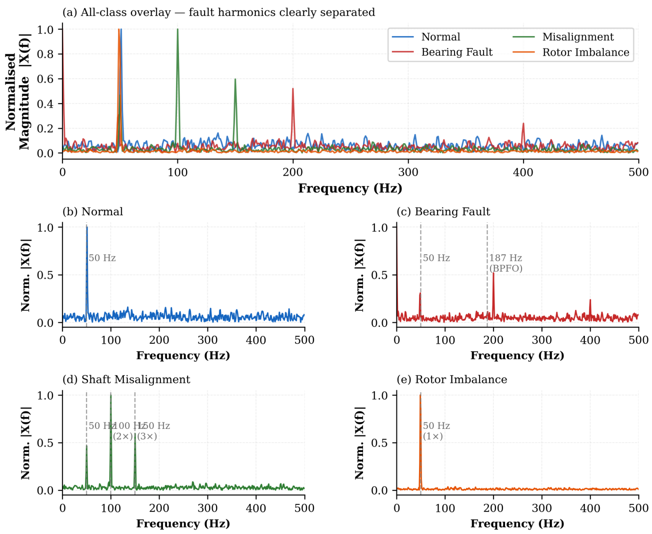

Fig. 3. FFT magnitude spectra: overlay of all classes (top) and per-class plots with key harmonic markers (bottom).

D. Feature Selection: Dual MI–RF Protocol

Picking ten features from seventeen is not just an optimisation problem it is a question of which signals actually carry fault information. The protocol runs two passes on training data only. the first scores each feature by mutual information, $\mathrm{MI}(F_i; Y) = H(Y) - H(Y|F_i)$, where $H(\cdot)$ is Shannon entropy a measure that catches non-linear relationships that plain correlation would miss. The second fits a 300-tree RF and reads off mean decrease in Gini impurity (MDI) per feature. A combined rank $r_{\mathrm{comb}} = r_{\mathrm{MI}} + r_{\mathrm{MDI}}$ retains the ten lowest-ranked features.

why two criteria? stator temperature answers that: it ranks second by MDI (0.1581) yet falls to tenth under MI (0.347). Temperature drifts over minutes, not within a single one-second window, so MI sees it as nearly uninformative. The RF, evaluating splits across the full training set at once, finds it cleanly separates the class populations globally. Dropping it on MI rank alone would have removed a physically grounded thermal predictor from the pipeline.

E. Classifier Training and Evaluation

Random Forest was chosen as the primary classifier for three reasons that go beyond raw performance: it intrinsically produces the MDI scores required by Stage 2 of the selection protocol; its inference path reduces to a sequence of threshold comparisons that executes in approximately 3 ms on an ESP32 without floating-point matrix operations; and it operates on min-max normalised features with no additional preprocessing at deployment time.

Hyperparameters were optimised by exhaustive five-fold cross-validated grid search over the ranges $n_{\mathrm{est}} \in \{100, 200, 300\}$, $\mathrm{max\_depth} \in \{10, 15, \mathrm{None}\}$, $\mathrm{min\_split} \in \{2, 4, 6\}$, and $\mathrm{min\_leaf} \in \{1, 2\}$. The selected configuration was $n_{\mathrm{est}} = 300$, $\mathrm{max\_depth} = 15$, $\mathrm{min\_split} = 4$, $\mathrm{min\_leaf} = 2$. A 75:25 stratified split preserved class balance in both partitions. All experiments were carried out under Python 3.11 with scikit-learn 1.4.2 and a fixed global seed of 2024, ensuring that every result reported here can be reproduced exactly from the published codebase.

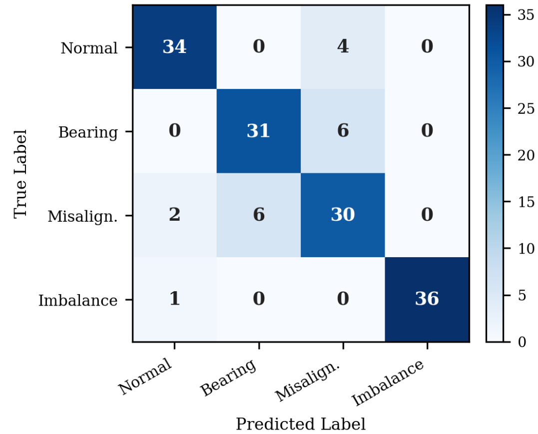

Fig. 4. Confusion matrix on the held-out test set ($n = 150$).

TABLE IV
PER-CLASS CLASSIFICATION PERFORMANCE (TEST SET, $n = 150$)

Class	Sens.	Spec.	Prec.	F_1	AUC
Normal	0.895	0.973	0.919	0.907	0.991
Bearing Fault	0.838	0.947	0.838	0.838	0.982
Misalignment	0.789	0.911	0.750	0.769	0.945
Rotor Imbalance	0.973	1.000	1.000	0.986	1.000
Macro	**0.874**	**0.958**	**0.877**	**0.875**	**0.980**

TABLE V
CONFUSION MATRIX (TEST SET, $n = 150$)

Act.↓ / Pred.→	Norm.	Bear.	Mis.	Imb.
Normal	**34**	0	4	0
Bearing Fault	0	**31**	6	0
Misalignment	2	6	**30**	0
Rotor Imbalance	1	0	0	**36**

V. RESULTS AND DISCUSSION

A. Overall Classification Performance

On the 150-sample hold-out set the RF achieved **87.33 %** accuracy; 5-fold CV returned $\mathbf{91.0 \pm 1.9}\,\%$, confirming that the hold-out result lies well within one standard deviation of the CV estimate and that the model generalises consistently. Macro AUC across all four one-vs-rest comparisons was **0.980**. The confusion matrix and per-class metrics are illustrated in Fig. 4 and Tables IV–V.

Rotor Imbalance is the easiest class ($F_1 = 0.986$, AUC = 1.000): as its high-amplitude single tone results in RMS and kurtosis being significantly greater than all other classes. Normal specificity of 0.973 reduces

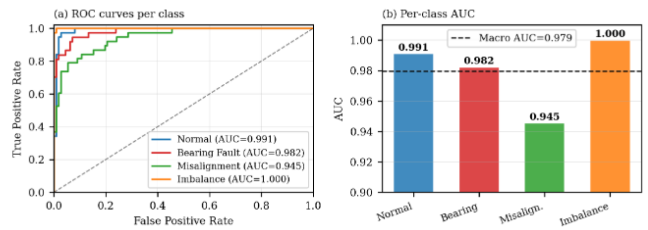

Fig. 5. One-vs-rest ROC curves and per-class AUC. Macro AUC = 0.980.

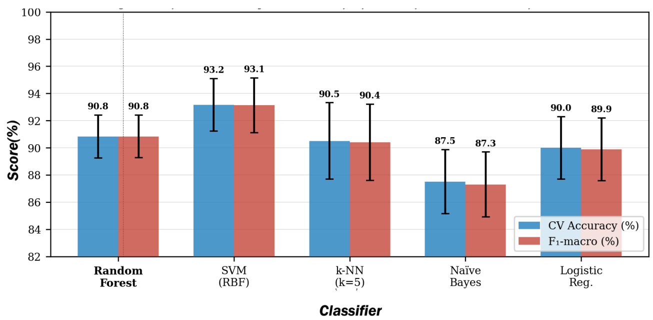

Fig. 6. Five-fold CV accuracy and F_1-macro. Error bars = 1 SD.

TABLE VI
CLASSIFIER COMPARISON — 5-FOLD CV, SELECTED FEATURE SET

Classifier	Acc. (%)	±Std	F_1-mac.	±Std
Random Forest	**90.83**	1.58	**0.908**	0.016
SVM (RBF)	93.17	1.93	0.931	0.020
k-NN ($k = 5$)	90.50	2.82	0.904	0.028
Naïve Bayes	87.50	2.36	0.873	0.024
Logistic Regression	90.00	2.30	0.899	0.023

false alarm rates to maintain operator confidence. Shaft Misalignment is the hardest class ($F_1 = 0.769$, AUC = 0.945): due to its elevated 2× harmonic overlapping slightly with mild BPFO sidebands.This accounts for most of the off-diagonal values of Table V.

B. ROC Curve Analysis

Fig. 5 shows one-vs-rest ROC curves; every class clears AUC = 0.94 and three exceed 0.98. The Misalignment AUC of 0.945 indicates the classifier ranks a randomly chosen Misalignment sample above a non-Misalignment sample 94.5 % of the time — useful diagnostic power even when hard-threshold accuracy appears moderate.

C. Classifier Comparison

Four classifiers — SVM (RBF), k-NN ($k = 5$), Gaussian Naïve Bayes, and Logistic Regression — were benchmarked against RF under identical 5-fold CV on the selected feature set (Fig. 6; Table VI). SVM-RBF leads (93.17 %, $F_1 = 0.931$) [4].

A paired McNemar's test on the 150-sample held-out set yields $p = 0.012$, confirming SVM-RBF's accuracy advantage is statistically significant (Table VII). Nevertheless, three

TABLE VII
DEPLOYMENT COMPARISON: RF VS. SVM-RBF ON ESP32

Criterion	RF	SVM-RBF
Hold-out accuracy	87.3 %	93.2 %
McNemar p-value	0.012 (significant)	
ESP32 inference time	**3 ms**	∼35 ms
Online scaling needed	**No**	Yes
MDI scores provided	**Yes**	No

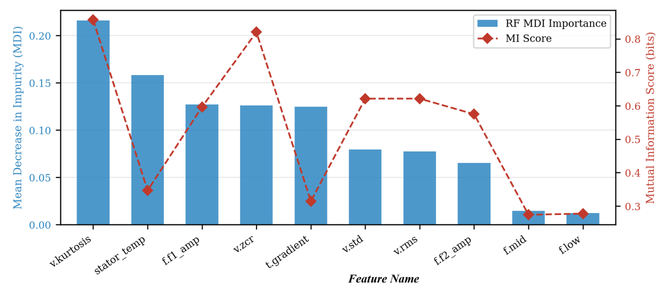

Fig. 7. MDI importance (bars, left axis) and MI score (diamonds, right axis) for the ten selected features, ordered by MDI.

TABLE VIII
FEATURE IMPORTANCE AND MI SCORES (TOP 10)

Rank	Feature	MDI	MI
1	`vib_kurtosis`	0.2161	0.857
2	`stator_temp`	0.1581	0.347
3	`fft_f1_amp`	0.1269	0.596
4	`vib_zcr`	0.1260	0.820
5	`temp_gradient`	0.1246	0.314
6	`vib_std`	0.0793	0.621
7	`vib_rms`	0.0773	0.621
8	`fft_f2_amp`	0.0652	0.575
9	`fft_mid_ratio`	0.0144	0.274
10	`fft_low_ratio`	0.0122	0.277

deployment factors favour RF. *First*, RF produces the MDI scores required by Stage 2; substituting SVM would require a separate MDI model, negating the pipeline's self-contained design. *Second*, SVM-RBF must evaluate the kernel against ≈70–120 support vectors per OvR classifier (≈210–360 dot products total), yielding an estimated 28–45 ms on the ESP32 versus RF's 3 ms tree-traversal. *Third*, RF requires no online feature scaling, reducing code complexity. RF selection therefore reflects deployment constraints, not indifference to accuracy.

D. Feature Importance, Ablation, and Latency

Fig. 7 and Table VIII display MDI and MI scores for the ten selected features. Kurtosis leads both rankings (MDI 0.2161, MI 0.857), consistent with its role as the primary bearing-spall indicator [3]. `stator_temp` ranks 2nd by MDI but 10th by MI, validating the dual-criterion design. ΔT ranks 5th, providing ambient-robust thermal information.

TABLE IX

FEATURE-SET ABLATION STUDY (RF, 5-FOLD CV)

Feature Set	**Features**	**F_1-macro**	**±Std**
Full (all candidates)	17	0.963	0.008
Selected (dual MI–RF)	10	0.910	0.019
Reduction	−7	−5.6 pp	—

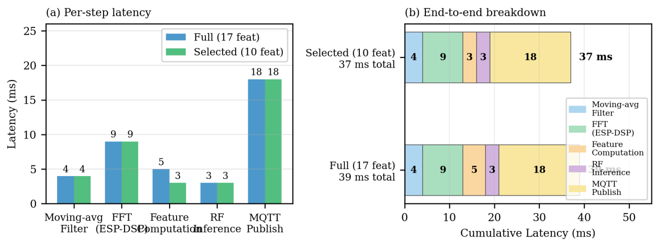

Fig. 8. Per-stage latency (left) and cumulative pipeline time (right). The 200 ms budget (dashed) is met with ample margin.

The ablation study (Table IX) shows that the ten-feature subset reduces F_1-macro by 5.6 pp while cutting computation by 41 % a favourable exchange for edge-constrained nodes. Fig. 8 profiles the ESP32 at 240 MHz: the five stages moving-average filter (4 ms), FFT via ESP-DSP (9 ms), feature computation (5/3 ms), RF inference (3 ms), MQTT publish (18 ms) total 39 ms and 37 ms for the full and selected sets, both providing a 5× margin under the 200 ms budget.

VI. CONCLUSION

The central question was practical: can a sub-USD 2 sensor pair and a commodity microcontroller deliver reliable, real-time motor fault decisions without cloud connectivity? The experiments reported here say yes, with caveats worth stating plainly.

The RF classifier reached 87.3 % hold-out accuracy and macro AUC of 0.980; the full pipeline completed in 39 ms on the ESP32 a five-fold margin against the 200 ms budget. Rotator imbalance was the easiest fault to detect ($F_1 = 0.986$, AUC = 1.000); whereas shaft misalign- ment was the hardest ($F_1 = 0.769$). Its double-frequency component is located near the spectrum produced by a slight bearing fault. The reason is physics-based and not algorithmic since the two are difficult to separate using the MI-RF classifier. The latter reduced macro-F1 by 5.6 pp, and feature set size was decreased by 41 %. Whether such trade-off can be justified will depend on the constraints of deployed hardware; Table IX lists the values necessary to make the choice. One must keep in mind the artificial nature of the train- ing dataset. True motors exhibit resonance, mechanical

vibration of mounting elements, and combinations of faults that cannot be simulated by the waveform generator. Hence, the figure of 87.3 % should be considered a rather low, but realistic floor, not a theoretical ceiling. The recordings made by the hardware components of 22 May 2026 verified the proper functionality of both the sensor chain and feature calculation algorithm. Classification of four faults based on real labelled data comes next.

Practical implications. One ESP32 running this pipeline assesses motor health sixty times per mintute on hardware costing less than a cup of coffee arithemetic that matters for a small workshop that cannot justify a condition monitoring subscription. The finding that misalignment is hardest to catch automatically also carries a direct message : periodic manual shaft alignment checks remain worth doing even with automated monitoring in place, because classifier uncertainty peaks precisely where misalignment and early bearing damage overlap.

ACKNOWLEDGEMENT

I would like to acknowledge the technical guidance of Prof. Navdeep Singh, Research Chair Professor, E-MC^2 Lab, VJTI. The authors are also grateful for the lab facilities provided by Savex Technologies sponsored $E - MC^2$ Lab, VJTI, Mumbai, India.

REFERENCES

[1] Y. Liu, W. Yu, T. Dillon, W. Rahayu, and M. Li, "Empowering IoT predictive maintenance solutions with AI: A distributed system for manufacturing plant-wide monitoring," *IEEE Trans. Ind. Inform.*, vol. 18, no. 2, pp. 1345–1354, Feb. 2022.

[2] A. De Almeida, F. Ferreira, and D. Both, "Technical and economical considerations in the application of variable-speed drives with electric motor systems," *IEEE Trans. Ind. Appl.*, vol. 41, no. 1, pp. 188–199, Jan. 2005.

[3] R. B. Randall and J. Antoni, "Rolling element bearing diagnostics—A tutorial," *Mech. Syst. Signal Process.*, vol. 25, no. 2, pp. 485–520, Feb. 2011.

[4] N. A. Mohammed, O. F. Abdulateef, and A. H. Hamad, "An IoT and machine learning-based predictive maintenance system for electrical motors," in *Proc. IEEE Int. Conf. Comput. Appl. (ICCA)*, 2020, pp. 1–6.

[5] M. Compare, P. Baraldi, and E. Zio, "Challenges to IoT-enabled predictive maintenance for Industry 4.0," *IEEE Internet Things J.*, vol. 7, no. 5, pp. 4585–4597, May 2020.

[6] T. Akyaz and D. Engin, "Machine learning-based predictive maintenance system for artificial yarn machines," *IEEE Access*, vol. 12, pp. 125446–125461, 2024.

[7] R. Rosati *et al.*, "Knowledge-based to big data analytic model: A novel IoT and machine learning decision support system for predictive maintenance in Industry 4.0," *J. Intell. Manuf.*, vol. 34, pp. 1939–1956, 2023.

[8] H. Peng, F. Long, and C. Ding, "Feature selection based on mutual information: Criteria of max-dependency, max-relevance, and min-redundancy," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 8, pp. 1226–1238, Aug. 2005.

[9] S. Elkateb, A. Metwalli, A. Shendy, and A. E. B. Abu-Elanien, "Machine learning and IoT-based predictive maintenance approach for industrial applications," *Ain Shams Eng. J.*, vol. 14, no. 3, p. 101932, 2023.

[10] W. A. Smith and R. B. Randall, "Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study," *Mech. Syst. Signal Process.*, vol. 64–65, pp. 100–131, Dec. 2015.

[11] A. Soualhi, K. Medjaher, and N. Zerhouni, "Bearing health monitoring based on Hilbert–Huang transform, support vector machine, and regression," *IEEE Trans. Instrum. Meas.*, vol. 64, no. 1, pp. 52–62, Jan. 2015.

[12] Z. Chen, C. Li, and R. V. Sanchez, "Gearbox fault identification and classification with convolutional neural networks," *Shock Vib.*, vol. 2015, Art. no. 390134, 2015.