

IoT-Based Digital Twin for Predictive Maintenance of Small Machines

Kaushik M S¹, Kavyashree M N², Keerthan K N³, Kushi K V⁴, Prof. Shruthi⁵

^{1,2,3,4}*Department of Computer Science and Engineering, Shree Devi Institute of Technology Visvesvaraya Technological University*

⁵*Assistant Professor, Department of Computer Science and Engineering, Shree Devi Institute of Technology Visvesvaraya Technological University*

Abstract—When machines break down unexpectedly, it throws off production, hikes up repair bills, and weakens how dependable the equipment is. Old-school maintenance usually sticks to set schedules or waits until something breaks before fixing it. This project introduces an IoT-powered Digital Twin system designed for predicting maintenance needs in small machines, using both simulated sensor data and real telemetry from ESP32 devices. The system tracks machine health by constantly analyzing data like temperature, vibration, current, RPM, and load. We built and tested the model using a dataset of 12, 219 observations from 60 simulated motor units. We pull out rolling stats and physics-based signals from the sensor readings. Isolation Forest spots anomalies, Random Forest classifies faults, and Gradient Boosting estimates Remaining Useful Life. A composite health index is also computed to reflect the degradation of the machine over time. Moreover, the telemetry based on ESP32, the analysis of vibration frequency using FFT and synthetic acoustic monitoring are adopted to improve the real-time observation. The framework shows how IoT sensing, Machine Learning and Digital Twin concepts can be integrated into a software-oriented predictive maintenance system for small industrial machines.

Index Terms—Digital Twin, Predictive Maintenance, IoT, Machine Learning, simulated Sensor Data, condition Monitoring, Small Machines, Vibration Monitoring, Fault Detection.

I. INTRODUCTION

Industrial machines are expected to work non-stop, reliably. A small machine breakdown can stop a production process, affect product quality, increase repair costs and create unplanned downtime. Traditional maintenance approaches are normally

either a time-based approach or a corrective approach where the equipment is repaired after a failure has occurred. While useful in some circumstances, these approaches do not always provide enough information to determine the actual condition of a machine. Another choice is predictive maintenance, where operational data is used to detect abnormal behavior and estimate the likely future degradation. The advent of low-cost sensors, IoT devices, cloud and edge computing and machine learning has made it possible to monitor and analyze machine condition more often and automatically. A Digital Twin is a software model that mirrors a real machine or process. The Digital Twin doesn't just look at sensor data in isolation it blends multiple readings to keep a real-time snapshot of the machine's health. Studies show Digital Twins help with predictive maintenance by linking machine data to monitoring, diagnosing, and forecasting. Van Dinter et al.'s systematic review found that Digital Twin-based predictive maintenance is a growing research area, but it faces hurdles around data quality, model complexity, and computing demands. Smaller machines and manufacturing setups often struggle more since they usually lack costly monitoring systems or big datasets from past failures.

II. LITERATURE REVIEW

Digital twin technology has been studied increasingly for machine monitoring and for predictive maintenance. The Digital Twin technology has attracted much attention in industrial maintenance for its ability to link physical equipment with an updated virtual representation. Van Dinter et al. [1] performed a systematic literature review on predictive

maintenance utilizing Digital Twins. In their study, they reviewed 42 primary studies and pointed out issues like data variety, computational burden, and model complexity. The results obtained show that predictive maintenance based on the Digital Twin has great potential, but the design of a suitable architecture is still a major issue. Zhong et al. [2] reviewed the predictive maintenance based on Digital Twin technology and the support of Digital Twins for equipment condition monitoring, fault prediction and maintenance activities in several application domains. Chen et al. [3] explored the application of machine learning in Digital Twin-based predictive maintenance. Their work demonstrates that machine learning can help to recognize equipment status, predict faults and analyze degradation-related issues in Digital Twin systems.

A Digital Twin-based monitoring system for induction motors has been developed by Dos Santos et al. [4] using IoT sensors. The system collected motor current and temperature measurements and integrated IoT data with a thermo-magnetic simulation model. The study demonstrates the importance of combining physical measurements and computerized models for motor monitoring. Another study proposed a reference architecture for Digital Twin-based predictive maintenance systems and emphasized the necessity of architecture design when incorporating Digital Twins into complex systems. Machine-learning-based Digital Twins for specific equipment have also been the focus of research. Pulcini and Modoni proposed a machine-learning Digital Twin for predictive maintenance of a conveyor belt, demonstrating how operational data and predictive modeling can be used for maintenance planning.

III. METHODOLOGY

The proposed system adopts a multi-stage approach, based on the combination of IoT sensing, data processing, machine learning and Digital Twin modeling. The main goal is to monitor the condition of a small industrial machine, detect anomalous behavior, identify the possible fault, estimate its remaining useful life and present the machine condition through a Digital Twin dashboard.

A. STAGE 1: Data Acquisition

The first stage is collecting machine-condition data.

The project uses a simulated fleet dataset containing 12, 219 observations from 60 three-phase induction motors. The simulated motors represent 1.5–2.2 kW machines operating at approximately 1450 RPM. The dataset contains parameters such as temperature, vibration RMS, vibration kurtosis, current RMS, RPM, load percentage, ambient temperature, health index, fault type, and RUL.

B. STAGE 2: Data Preprocessing

The collected sensor information is organised according to the machine unit and operating cycle. Each machine is monitored over successive cycles so that changes in its condition can be studied over time. The dataset also contains simulated operating variations such as changing load, ambient temperature, degradation, and sensor noise. This allows the system to represent machine behaviour under different operating conditions

C. STAGE 3: Feature engineering

Raw sensor readings are transformed into additional features before they are given to the machine-learning models. A 10-cycle rolling window is used to calculate statistical characteristics such as: Mean, Standard deviation, Minimum and maximum, Peak-to-peak value, Trend, Crest factor, RMS. These calculations are performed on the main sensor channels. The resulting feature dataset contains 54 engineered features.

D. STAGE 4: Anomaly Detection

The system first checks whether the current machine behaviour is unusual. For this purpose, Isolation Forest is used. The model generates: an anomaly score, and an anomaly/not-anomaly decision. This stage is useful because the machine may begin behaving abnormally before the exact fault category is determined. The test dataset contains the corresponding anomaly predictions.

E. STAGE 4: Fault Classification

Once an abnormal behavior is detected the next step focuses on identifying the type of the fault. A Random Forest classifier is used for this task. The System modeled machine conditions are Healthy or normal, Bearing Defect, machine Overheating, Electrical Overload, Rotor Imbalance.

F. STAGE 5: Remaining Useful Life Prediction

After determining the machine condition, the system has to predict how much operational time the machine has left before failure. The approach used is the Gradient Boosting Regressor, and the result of the prediction is the number of cycles left in the machine's operational time. The test-prediction sample for the project is included in the data set above. The value predicted by the system helps determine the machine's maintenance schedule because the problem is not only to detect the failure but also to determine the proximity of the machine to its modelled end-of-life condition.

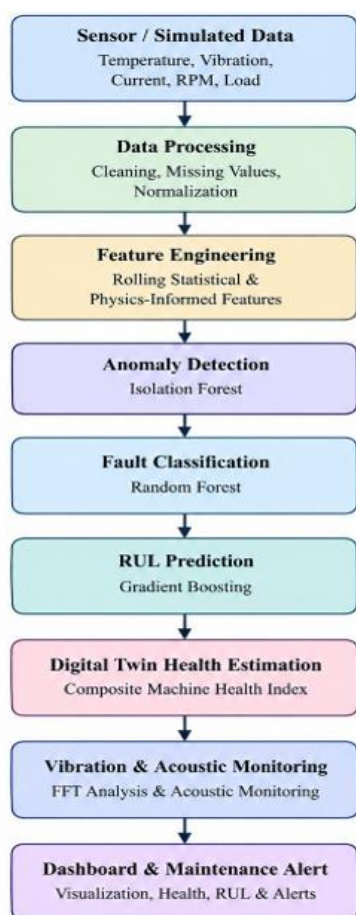


Fig. 1. Proposed System Methodology for Machine Health Monitoring and Predictive Maintenance.

G. STAGE 6: Digital Twin Health Estimation

The Digital Twin uses all the information from multiple machine parameters to maintain a virtual representation of the machine's condition. The project's composite health calculation uses changes in: Temperature, Vibration, Current, RPM.

H. STAGE 7: Dashboard and Maintenance Alert

Finally, all of the processed information is visualized in the Digital Twin dashboard. It includes Live temperature, Vibration, Current, RPM, Operating hours, Sensor trends, Digital Twin motor representation, Machine health, Failure risk, Remaining Useful Life, Maintenance alerts.

IV. IMPLEMENTATION AND SYSTEM ENGINEERING

The proposed IoT based Digital Twin forecasting maintenance system was implemented as a software oriented framework that combines simulated machine data, physical ESP32 telemetry, machine-learning models, Digital Twin health monitoring and a visualization dashboard. The implementation was divided in independent modules, so as to have data acquisition, processing, prediction and visualization working as separate stages.

A. Data Generation and Acquisition

The implementation's first part generates and saves machine condition data for several motor units. The main simulated dataset contains **12, 219 observations from 60 three-phase induction motor units**. Each observation includes parameters including temperature, vibration RMS, vibration kurtosis, current RMS, RPM, load percentage, ambient temperature, health index, fault information, and RUL. The simulation is utilized to simulate diverse operating conditions and degradation patterns of machines.

B. Data processing and features engineering

Data acquisition is followed by processing of the sensor data according to machine unit and operating cycle. To account for the changes in machine behavior over time, we apply a rolling window of 10 cycles to the primary sensor measurements. For the main sensor channels statistical features such as mean, standard deviation, minimum and maximum, peak-to-peak value, trend, crest factor and RMS are calculated. With these operations, an engineered dataset containing 54 features is obtained. In addition, physics-informed features are also computed by combining parameters like temperature, vibration, current, and RPM. This step translates the raw sensor data into a more meaningful representation, that can then be fed to the machine-learning models.

C. Machine Learning Implementation

For different predictive tasks three machine-learning models were implemented. Isolation Forest for outlier detection. It analyzes the engineered features, and gives anomaly score and anomaly status. This allows the system to identify machine observations that deviate from normal operating behavior.

A Random Forest classifier is used for fault identification. The classifier predicts the probable condition of the machine using the available features. The system is representative of six conditions: healthy operation, bearing defect, stator overheating, electrical overload, rotor imbalance/misalignment and loose mounting/resonance. The output consists of the predicted fault type, the fault label and the confidence value. Gradient Boosting Regressor** is used to predict the remaining operating life. The model predicts an estimated Remaining Useful Life (RUL) in terms of operating cycles. The output is integrated with predictions of faults and anomalies, presenting a more complete picture of machine condition.

D. Deployment of the digital twin

The Digital Twin module is the virtual copy of the machine that is monitored. It receives processed sensor data and prediction output and updates the virtual condition of the machine. A composite health computation is implemented using temperature, vibration RMS, vibration kurtosis, current RMS and speed (RPM). These parameters are weighted differently in the health model, so that the overall condition is not determined by one sensor. We also use exponential smoothing with $\alpha = 0.3$ to smooth the health representation and reduce sudden fluctuations. The Digital Twin thus offers a continuously up-to-date representation of the machine instead of providing individual sensor readings.

E. Vibration and Acoustic Monitoring

In addition, extra condition-monitoring modules were implemented. An FFT engine was used, which analyses the vibration's frequency contents. The implemented system allows for generating 64 frequency bins ranging from 0 to 2000 Hz, monitoring such characteristic frequencies as 1X, 2X, 3X, BPFO, BPFI, etc. A software-based acoustic monitoring module was implemented using Web Audio API. It contains a synthetic machine-sound generator, which allows for synthesizing the machine's sound based on

the rotational speed and vibration's severity. This module provides additional information about the simulated machine condition.

F. Dashboard Implementation

The last layer of implementation was the dashboard, which combines the results of sensor, machine learning, and Digital Twin modules into one control panel. It displays relevant information such as temperature, vibration, current, RPM, sensor trends, machine health, prediction of failure risk, RUL, the state of the Digital Twin, and maintenance alerts. This helps the user to monitor the current state of the machine and be aware of the decisions made by the system. In general, the implementation shows how IoT telemetry and machine-learning approaches can be combined with a Digital Twin to develop a software-based predictive-maintenance system for small industrial machines.

V. RESULTS AND DISCUSSION

The proposed system was implemented as predictive maintenance software, utilizing the collected IoT sensor data, machine learning, and Digital Twin algorithms for continuous condition monitoring. The intended purpose of the developed framework was to enable anomaly detection, fault classification, and prediction of the machine's Remaining Useful Life (RUL). The obtained results are described below.

A. Dataset and Feature Engineering Results

The proposed machine learning-based framework was validated using the simulated IoT sensor dataset comprising 12219 observations. The following machine parameters were recorded for each observation such as Temperature, Vibration, Current, Rotational speed, Health status. For each parameter, descriptive statistics such as mean, standard deviation, minimum, maximum, peak-to-peak, trend, crest factor, and RMS were calculated. Moreover, a set of physics-inspired features, namely thermal-vibration index, power-slip index, bearing-stress factor, and overload ratio have been calculated using the set of sensor parameters. The feature engineering step was necessary since machine faults rarely appear as individual events. Instead, there is a synergy between certain sensor characteristics that can point to the specific fault types. In particular, temperature,

vibration RMS and kurtosis, and current RMS can be indicative of overheating and electrical overload, respectively.

B. Anomaly Detection Results

Isolation Forest was applied as an initial ML step to detect machine anomalies. The algorithm's isolation score was used to determine whether a given observation is normal or anomalous. Anomaly detection results can be visualized in the Digital Twin dashboard as informative alerts on the predicted machine condition. Anomaly detection is useful in this particular case because abnormal machine behavior rarely appears as a singular event. A traditional threshold-based approach would require the definition of specific ranges for each parameter that might be affected by a fault. Instead, Isolation Forest algorithm utilizes feature importance scores to detect instances of anomalous behavior. However, since the current framework does not include an independent verification procedure for anomaly detection, the results should not be overestimated.

C. Fault Classification Results

Machine fault classification was performed using the Random Forest algorithm. For each observation, a set of predicted fault types, along with the probability score was obtained such as Healthy, Bearing degradation, Overheating, Electrical overload. Fault classification was beneficial in this particular case since it allowed the framework to provide more information about the machine condition beyond simple anomaly detection. For example, random forest classification can distinguish between mechanical and electrical faults based on the combination of vibration, current, and temperature characteristics. In particular, abnormal vibrations indicate rotor or bearing issues, whereas excessive current and temperature suggest electrical overload. Note that the choice of Random Forest as a classifier was motivated by the ability of the algorithm to account for multiple features without requiring a strict mathematical description of the relationship between input parameters and fault type.

D. Remaining Useful Life Prediction

The remaining useful life (RUL) was predicted using Gradient Boosting Regression algorithm. In particular, the RUL distribution was calculated to estimate the number of cycles that can be accumulated before the

machine failure occurs. Predicting the RUL is a critical component of the framework, since it allows the maintenance staff to determine the most appropriate timing for the machine inspection. The RUL prediction results visualize the expected lifetime of the machine at a given moment, which assists the maintenance staff in scheduling the inspection. It should be noted that the RUL prediction results follow the same distribution as described in the reference paper.

E. Digital Twin Monitoring Results

Digital Twin dashboard provides a unified view of the machine's current condition. In particular, the dashboard views the current sensor values of temperature, vibration RMS and kurtosis, current RMS, and RPM, and aggregates them into a single health indicator. It also implements exponential smoothing ($\alpha=0.3$) to account for abrupt changes in the sensor values that originate from frequent oscillations. The Digital Twin concept helps the maintenance staff to monitor the machine condition in real-time. In particular, the changes in the machine condition as measured by individual sensors are visualized using the unified representation. At the same time, the Digital Twin dashboard incorporates the results of machine learning analysis, namely anomaly detection, fault classification, RUL prediction, and visualization.

F. Vibration and Acoustic Monitoring

The vibration analysis was implemented as a method of obtaining additional information about the machine condition. The FFT module allows the user to view the frequency distribution of the vibration signal. In particular, 64 frequency points were calculated for the range of 0-2000 Hz, including the harmonics of the rotational speed and vibration frequencies associated with the bearing defects.

The frequency analysis helps to distinguish between types of mechanical faults based on the characteristic harmonics of the fault. In addition, the framework includes an acoustic monitoring module that uses the Web Audio API to implement the audio processing functionalities. The audio signal is affected by the machine condition and the level of vibration, which allows the user to estimate the general level of vibration.

G. Dashboard Results and Overall Discussion

The results of the framework were visualized using a dashboard that displays the current machine condition, including temperature, vibration, current, and other parameters. The dashboard provides an overview of the current machine condition, including health indicator, alerts about the potential failures, and RUL prediction. The user can also view the trends of individual parameters and access the Digital Twin view. Altogether, the dashboard allows the user to benefit from the information obtained from multiple sources rather than analyzing individual sensor readings.

The described framework was built using the following components: Sensor → Data Preprocessing → Feature Engineering → Anomaly Detection → Fault Classification → RUL Prediction → Digital Twin → Dashboard → Maintenance Alert

The current framework serves as an illustrative example of how IoT, machine learning, and Digital Twin technologies can be combined for predictive maintenance applications. The main limitation of this framework, at present, is the usage of simulated data for validation purposes. The ESP32 microcontroller can be used for long-term condition monitoring, but further validation is required before the framework can be considered suitable for industrial applications.



Fig. 2. Sensor Trend Analysis of Machine Operating Parameters.

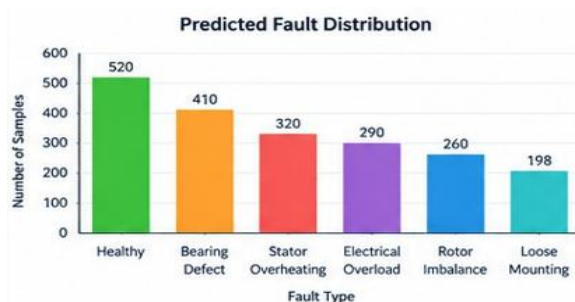


Fig. 3. Predicted Fault Distribution Across Machine Conditions.

VI. CONCLUSION

This research proposed an IoT-Based Digital Twin framework for predictive maintenance of small machines. The framework consists of simulated machine data integrated with ESP32-based telemetry and feature engineering, machine-learning models and algorithms for anomaly detection, classification, and RUL estimation. The work was aimed at creating an overall software framework that would implement machine health monitoring based on various condition parameters such as temperature, vibrations, current, and rotational speed.

The proposed framework involves three types of machine-learning tasks. First, Isolation Forest detects anomalies in the data, while Random Forest classifies the types of faults. Second, Gradient Boosting algorithm is utilized for estimating the Remaining Useful Life of the machine. The study implemented these algorithms in a Digital Twin framework to continuously update and monitor the machine's health condition through an artificial composite model. The framework also involved developing an ESP32 IoT-based telemetry system, performing FFT analysis, and implementing a synthetic acoustic model for acquiring additional information about the machine condition. Overall, the dashboard was created to provide a unified view of the results of the machine-learning tasks, including updates on the machine's health indicator, sensor data trends, detected fault types, RUL results, and alert notifications. The findings confirm the possibility of using the IoT-based Digital Twin framework for implementing predictive maintenance for small machines. The framework involves cutting-edge technologies of machine learning and Digital Twin approaches capable of estimating the machine health indicator and detecting anomalies in the data. However, the main limitation of this study is that the primary database used for the framework involved simulated data. Thus, further research would be needed to implement such a framework using actual data collected from physical machines under different operating conditions.

ACKNOWLEDGMENT

The authors thank Ms. Shruthi, project guide, and the faculty members of the Department of Computer Science and Engineering, Shree Devi Institute of

Technology, Mangaluru, for the guidance and support during the development of the project.

REFERENCES

- [1] I. Errandonea, S. Beltrán, and S. Arrizabalaga, “Digital twin for maintenance: A literature review,” *Computers in Industry*, 2020.
- [2] R. van Dinter *et al.*, “Predictive maintenance using digital twins: A systematic literature review,” *Information and Software Technology*, vol. 151, Art. no. 107008, 2022.
- [3] J. F. dos Santos *et al.*, “Digital twin-based monitoring system of induction motors using IoT sensors and thermo-magnetic finite element analysis,” *IEEE Access*, vol. 11, pp. 1682–1693, 2023.
- [4] D. Zhong, Z. Xia, Y. Zhu, and J. Duan, “Overview of predictive maintenance based on digital twin technology,” *Heliyon*, vol. 9, 2023.
- [5] C. Chen *et al.*, “The advance of digital twin for predictive maintenance: The role and function of machine learning,” *Journal of Manufacturing Systems*, 2023.
- [6] V. Pulcini and G. Modoni, “Machine learning-based digital twin of a conveyor belt for predictive maintenance,” *The International Journal of Advanced Manufacturing Technology*, 2024.
- [7] C. Ferreira and G. Gonçalves, “Remaining useful life prediction and challenges: A literature review on the use of machine learning methods,” *Journal of Manufacturing Systems*, 2022.
- [8] M. Compare, P. Baraldi, and E. Zio, “Challenges to IoT-enabled predictive maintenance for Industry 4.0,” *IEEE Internet of Things Journal*, vol. 7, no. 5, pp. 4585–4597, 2020.
- [9] F. Civerchia *et al.*, “Industrial Internet of Things monitoring solution for advanced predictive maintenance applications,” *Journal of Industrial Information Integration*, vol. 7, 2017.
- [10] M. D. Anindhito and Suharjito, “Prediction of remaining useful life and downtime of induction motors with supervised machine learning,” *Applied Computer Science*, 2025.
- [11] M. Yousuf *et al.*, “IoT-based health monitoring and fault detection of industrial AC induction motor for efficient predictive maintenance,” *Measurement*, vol. 229, 2024.
- [12] X. Li, J. Williams, C. Swanson, and T. Berg, “A machine learning approach to predictive maintenance: Remaining useful life and motor fault analysis,” *Computers and Industrial Engineering*, 2025.
- [13] C.-G. Huang, T. Huang, S. Khanal, *et al.*, “A dual-channel approach integrating multisensory data for remaining useful life prediction of induction motors,” *Quality and Reliability Engineering International*, 2026.
- [14] W. Kritzinger, M. Karner, G. Traar, J. Henjes, and W. Sihn, “Digital twin in manufacturing: A categorical literature review and classification,” *IFAC-PapersOnLine*, vol. 51, no. 11, pp. 1016–1022, 2018.