Advanced Engineering Informatics 65 (2025) 103388

Contents lists available at ScienceDirect

Advanced Engineering Informatics

journal homepage: www.elsevier.com/locate/aei

An edge-cloud IIoT framework for predictive maintenance in manufacturing systems

Nivethitha Somu *, Nirupam Sannagowdara Dasappa

Energy Research Institute @ NTU (ERI@N), Nanyang Technological University, 637141 Singapore

ARTICLE INFO

Keywords:
Big data analytics
Edge computing
Cloud computing
Predictive maintenance
Fault detection
Industrial Internet of Things

ABSTRACT

Despite significant research efforts on Industrial Internet of Things (IIoT) based Predictive Maintenance (PdM) systems, challenges related to the availability of real-time machine operational data, reliable computing-deployment architecture, and implementation in real-time manufacturing environments continue to be major concerns. Hence, this work presents Intelligent PdM (IntelliPdM), an end-to-end IIoT predictive maintenance framework implemented on an edge-cloud platform that processes the real-time heterogeneous data streams (IoT sensors and cameras) and provides intelligent decisions on faults, failures, and maintenance schedules via. endpoints and interactive web user interface (dashboards, alerts/recommendations, and analytics. SmartHome, a synthetic data generation framework was properly configured to generate synthetic data based on limited real-time operational data or open-source benchmark machine health datasets covering all possible industrial fault scenarios. Experimental validations using synthetic data, generated from real-time machine health data collected from a testbed setup at a research center in Western Europe, along with on-site implementation in a large manufacturing unit in Singapore, effectively demonstrate the efficiency of IntelliPdM in delivering accurate and reliable fault diagnostics. Over a 12-months real-time implementation, IntelliPdM demonstrated (i) an accuracy of 93–95%, (ii) 25–30% reduction in maintenance costs, (iii) 70–75% decrease in equipment breakdowns, (iv) 35–45% reduction in downtime, (v) 20–25% increase in production, and (vi) 10x return on investment.

1. Introduction

Intelligent manufacturing strategies and principles have a great impact on the high quality and efficient economy by improving the quality and management of the manufacturing process [1]. According to Dowlatshahi, nearly 79.6 % of the major breakdowns such as minor stoppages, equipment idling, and sudden shutdowns are caused by unplanned interruptions [2,3]. A fault in a manufacturing asset causes a huge loss in terms of production, downtime, efforts in identifying the root cause of the fault (malfunctioning of the equipment, operator error, or environmental factors), cost of repair, and deterioration of equipment. The recursive revenues in the manufacturing industry are feasible through the implementation of intelligent and actionable strategies for planning, monitoring, and maintenance of manufacturing assets [4,5]. The major objective of any industrial maintenance strategy is to ensure the high availability of manufacturing equipment and sustainable operational management at low maintenance costs [6–8]. The state-of-the-art maintenance strategies in the manufacturing industry can be classified into three categories [9], (i) **run-to-failure (R2F):** a simple approach that is carried out only when failure occurs; incurs high cost and production loss, i.e., suspension of the production line and activities, proper adjustment of the production facilities, and repair of the faulty equipment, (ii) **preventive maintenance:** detects and prevents faults by scheduled inspections; inevitable unnecessary actions lead to increased operational costs, and (iii) **predictive maintenance:** analyze the historical operational data of the equipment using intelligent techniques and predict equipment failure, i.e., provide insights on when the equipment fails and at what failure probability, thereby helps in

Abbreviations: IntelliPdM, Intelligent predictive maintenance; API, Application programming interface; IIoT, Industrial Internet of Things; LSTM, Long short term memory; SVM, Support vector machine; XGBoost, Extreme gradient boosting; AdaBoost, Adaptive boosting; MLP, Multi layer perceptron; kNN, K–nearest neighbour; CNN-LSTM, Convolutional neural network – LSTM; RNN, Recurrent neural network; GNB, Gaussian Naïve Bayes; DNN, Deep neural network; DBN, Deep belief network; ANN, Artificial neural network; SmoteNC, Synthetic Minority Oversampling Technique-Nominal Continuous; ctGAN, Conditional Tabular Generative Adversarial Network; BLR, Binary Logistic Regression.

* Corresponding author at: Energy Research Institute @ NTU (ERI@N), Nanyang Technological University, Singapore 637141.
E-mail addresses: nivethitha.somu@ntu.edu.sg (N. Somu), nirupam.dasappa@ntu.edu.sg (N.S. Dasappa).

https://doi.org/10.1016/j.aei.2025.103388
Received 21 October 2024; Received in revised form 13 January 2025; Accepted 18 April 2025
Available online 22 April 2025
1474-0346/

performing maintenance actions at right time to guarantee extended operation lifetime and reduced downtime & maintenance costs. Fig. 1 presents the P-F curve to demonstrate the transition phases between the point of potential failure and functional failure. The maintenance guidelines of (i) and (ii) are designed based on the normal lifespan of the machines and the decision on when to repair or replace the components of a machinery relies mainly on the domain knowledge of the maintenance manager [10]. Therefore, the most effective approach is to design a predictive maintenance system that can monitor the health condition of the machines and evaluate the need for maintenance using intelligent algorithms to detect anomalies, raise alarms, and send maintenance alerts well in advance to reduce planned and unplanned downtime.

In recent years, industrial manufacturing equipment has become more complex and technologically advanced, making them more difficult to control [12]. Recent advancements in Information and Communication Technology (ICT) have transformed the manufacturing industry towards Industry 4.0, wherein the system and processes are connected through IIoT technologies, physical manufacturing systems are represented as 'digital twin', real-time machine health data acquisition, processing, and intelligent decision-making are handled using artificial intelligence, big data analytics, cloud, and edge computing [13–15]. IIoT enables the design of intelligent predictive maintenance systems to achieve smart manufacturing with sustainable benefits, however open challenges with respect to high dimensional IIoT data collection, management, processing, and real-time analytics remain unresolved. Extensive research in IIoT based predictive maintenance systems has always focussed on the analytics engine which has evolved from model based approaches to artificial intelligence based approaches for making intelligent and supportive decisions on maintenance management, i.e., resource allocation, maintenance schedule, cost-effective maintenance strategy, etc. [16]. However, IIoT based predictive maintenance system involves the development of an end-to-end intelligent framework that involves sensor modality and placement strategy, data acquisition and ingestion, data annotation, accurate analysis and prediction of machine conditions, computational needs, and real-time processing and inference. Most of the state-of-the-art predictive maintenance architectures have relied on cloud technologies for data management, processing, and storage of machine health data. Studies in [10] highlight the challenges of cloud based IoT PdM architectures in providing real-time decisions and propose an integrated architecture for fault detection in IoT enabled smart manufacturing implemented on edge computing. The motivation for this work stems from the critical need to resolve substantial in the manufacturing industry, where unplanned equipment failures result in significant economic losses, increased downtime, and inefficiencies in maintenance operations. Conventional approaches are either expensive or ineffective, while advanced PdM systems struggle due to lack of benchmark datasets, difficulties in managing heterogeneous and high-dimensional IIoT data, and challenges in deploying real-time analytics in resource-constrained environments. With growing complexity of modern manufacturing equipment and rapid advancements in Industry 4.0 technologies there is a need for an intelligent, end-to-end maintenance solution. Therefore, this research is motivated by the goal of developing a comprehensive PdM framework that integrates efficient data acquisition, real-time analytics, and intelligent decision-making to enhance operational efficiency, minimize downtime, and support sustainable manufacturing practices.

Table 1 presents a comprehensive view of the recent research works in IIoT based predictive maintenance systems. The major challenges that hinder well-established deployment and application of IIoT based predictive maintenance systems for smart manufacturing spaces are, (i) **benchmark datasets:** unavailability of open-source repositories that contain machine health operational data for a wide range of machine-specific faults in different manufacturing spaces, (ii) **real-time operational data:** difficult to gather real-time data covering all possible faults that can occur in different machinery in the manufacturing space, (iii) **data format:** issues in handling structured (numeric and text) and unstructured (audio, image, video) data formats, (iv) **specific faults:** models designed for the identification of specific faults in the machinery; e.g., extensive research on identifying bearing faults of rotating machines, ignoring equally significant faults such as misalignment, looseness, unbalance, etc., (v) **model validation:** validation of the PdM models using benchmark datasets which do not consider the industrial conditions while dataset generation and benchmarking, and (vi) **cloud/edge deployment:** IoT based PdM systems deployed either on cloud or edge based architectures is unsuitable for resource constrained environments.

The research gaps and challenges in the state-of-the-art IIoT based PdM architectures highlight several key challenges that hinder their successful deployment and application in smart manufacturing environments which can be formulated as, (i) **RQ1:** What can be a 'one-stop solution' for data challenges related to real-time operational data and open-source datasets?, (ii) **RQ2:** What is the effective strategy to handle heterogeneous data from multiple sensors?, (iii) **RQ3:** How to effectively consider multi-fault aspects of fault diagnosis in manufacturing assets?, (iv) **RQ4:** How to demonstrate the effectiveness of the trained model in a real-time industrial environment?, and (v) **RQ5:** What are the potential

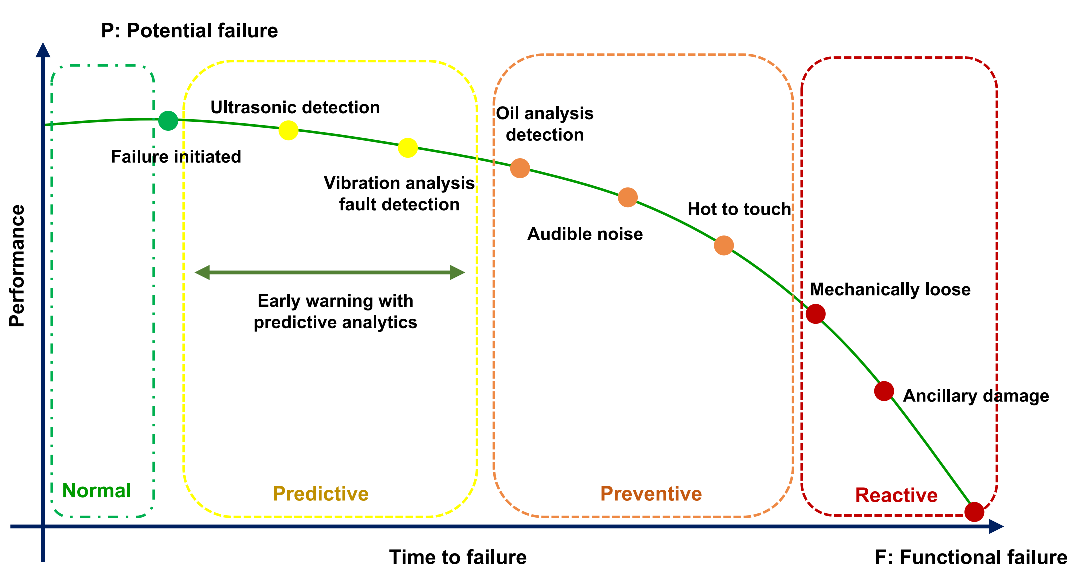

Fig. 1. Predictive maintenance P-F curve [11].

Table 1
Recent research works.

Approach	Dataset	Metrics	Application
Autoencoders; LSTM [17]	Real time data – Rolling mill machine	Recall: 73.48—88.18 %; Precision: 71.47–88.09 %; F1-score: 72.54—88.12 %; Specificity: 74.99—91.23 %	Cyber physical production system
SVM; random forest; XGBoost; gradient boosting; AdaBoost; MLP [18]	Real time data – Manufacturing plant, Turkey	R^2 : 0.982; MAE: 51.97; MAPE: 3.27; RMSE: 147.19	Production lines in manufacturing
Attribute attentioned LSTM [9]	Real time data – Aircraft manufacturing cooperation	Accuracy: 84.60 %, Precision: 89.79 %, Recall: 94.43 %, F1-score: 89 %	Manufacturing system
Random forest; decision trees; kNN[19]	Real time data – Algerian oil and gas company	MAE: 0.571	Predictive maintenance – Generic
LSTM autoencoder; CNN-LSTM [20]	Real time data – Offshore natural gas treatment plant; Sea water injection centrifugal pump	LSTM autoencoder – Precision: 83.61 – 90.81 %; Recall: 83.68 – 90.82 %; Specificity: 80.33 – 88.96 % CNN-LSTM – Precision: 99.04–100 %; Recall: 99.91–100 %; Specificity: 99.50–100 %	Oil and gas production system
Jaya algorithm – sea lion optimization based feature selection; SVM; RNN [21]	Aircraft engine and Li-ion battery dataset	Aircraft – MAE: 1.1259; RMSE: 5.0895 Li-ion – MAE: 47.221; RMSE: 196.02	Sustainable manufacturing
CNN-LSTM [22]	NASA's C-MAPSS dataset	MAE: 9.86; RMSE: 12.76; MAPE: 13.21	Cyber physical manufacturing system
Neural network, decision trees, random forest, k-NN, SVM, GNB [23]	PHM data challenge – vibration data of industrial oil-injection screw compressor	Accuracy: 98.51 %, Sensitivity: 99.78 %; Specificity: 98.19 %	Manufacturing asset
Genetic algorithm based resource scheduling; two-class logistic regression [24]	Real time dataset – Fidan Boylu Uz [25]	Accuracy: 94.50 %; Precision: 94.60 %; Recall: 93.30 %; F1-score: 93.90 %	Manufacturing asset
Balanced K-Star [26]	AI4I 2020 predictive maintenance dataset	Accuracy: 98.77 %; Precision: 98.75 %; Recall: 98.75 %; F1-score: 98.75 %	Manufacturing asset
Multi-Scale Dilation Attention CNN; probabilistic beetle swarm-butterfly optimization; DNN; DBN [27]	Datasets – 1. Aircraft engine 2. NASA's C-MAPSS, 3. Genesis demonstrator data 4. Ultrasonic flowmeter diagnostics	Dataset 1 – Accuracy: 95.91 %; Precision: 96 %; F1-score: 96 % Dataset 2 – Accuracy: 97 %; Precision: 43.75 %; F1-score: 60.31 % Dataset 3 – Accuracy: 96.55 %; Precision: 2.83 %; F1-score: 5.51 % Dataset 4 – Accuracy: 97.12 %; Precision: 94.39 %; F1-score: 95.74 %	Manufacturing asset
Geometric area analysis and trapezoidal area estimation based feature extraction and random forest [28]	Real time data – Robotic part loading system for coordinate measuring machine	MAE: 0.089; MSE: 0.018; R^2 : 0.868	Predictive maintenance – Generic
kNN-LSTM; knowledge graph [29]	Real time data – spot-welding robot of new energy automotive workshop	Accuracy: 91.09 %; Precision: 93.88 %; Recall: 96.30 %; F1-score: 90.91 %	Manufacturing asset
Edge computing-assisted stacked sparse autoencoder [10]	Real time data – Global manufacturing system, Australia	Test cases on equipment functional failure	Manufacturing asset
ANN [30]	Real time data – Talgø MøreTre AS	Accuracy: 98 %	Production lines in manufacturing
CNN-Bidirectional LSTM [31]	MIMII dataset	Recall: 0.89–0.96; Precision: 0.89–0.96	Manufacturing asset
Stacked ensemble learning; CNN [32]	NASA's turbofan engine degradation simulation dataset	Accuracy: 93.93–95.72 %	Predictive maintenance – Generic
Full consistently optimization method; multi-criteria decision making; ANN [33]	Data from ISyE lab, Ghent university	MSE: 0.0216; RMSE: 0.132	Manufacturing assets
AdaBoost [34]	Real time data – Circular knitting machines	Precision: 92 %; Recall: 92 %; F1-score: 91 %	Manufacturing assets
Deep transfer spiking neural network [35]	Real-time data – Rolling bearing test rig	Average precision: 98.12 %	Manufacturing assets
Deep convolutional neural Network [36]	Real-time data – Rolling bearing test rig	Accuracy: 98.9 %	Manufacturing assets

deployment options for resource constrained computing environment? To address these challenges and research questions, this work presents IntelliPdM, an IoT based predictive maintenance system for smart manufacturing spaces implemented on edge-cloud computing architecture. The novelty and contributions of this work are highlighted as,

1. IntelliPdM, an end-to-end predictive maintenance framework for smart manufacturing spaces with (i) **Data acquisition and ingestion layer:** collects and preprocess the machine operational data from heterogeneous data sources; ingest pre-processed data into appropriate data pipelines (sensor, audio, video/image) which are then transferred over secured File Transfer Protocol (FTP) to be staged and stored in the data store, (ii) **Predictive modeling layer:** build models to be deployed at edge or cloud (based on the computing capabilities of the edge), and (iii) **Application layer:** intelligent decisions provided by the model deployed at the edge/cloud are stored in the data store which is accessed through the exposed API endpoints by various components in the application

layer and presented to the appropriate personnel as alerts/recommendations via. dashboards.
2. A separate data pipeline was implemented for handling and ingesting structured and unstructured data into appropriate data stores (structured: MongoDB; unstructured: Amazon S3).
3. A synthetic data generation framework to generate synthetic data based on (i) open-source benchmark datasets or real-time datasets, in case of data sparsity (limited data samples) and (ii) scenarios based on real-time industrial conditions ('normal,' fault') encoded in configurable scenario file and software codes (Python script and libraries).
4. For the experimental investigations, real-time operational machine health data was collected from a testbench setup of manufacturing machinery over six months at 30-minute granularity. This data covers all possible faults in the critical components of the machinery, i.e., induction motor, rotating equipment, etc. Later, the real-time data from the testbench is used by the synthetic data generation framework to ensure adequate representation of data samples across various classes. The synthetic data was then used to train and evaluate the AI models using standard quality metrics.
5. A set of artificial intelligence models such as SVM, isolation forest, decision trees, random forest, ANN, 1D CNN, and LSTM were trained using the real-time operational data obtained from the synthetic data generation framework. Model building and training were done on the cloud and the API endpoints of the best-performing fault detection and classification models were exposed as API endpoints and deployed at the edge infrastructure. The trained models for unstructured data-based fault detection and classification techniques can be opted for deployment in the cloud infrastructure.
6. End-to-end implementation of IntelliPdM, predictive maintenance framework was done using Apache Kafka, Apache Spark, MongoDB, Amazon S3, React, Django, docker containers, AWS Elastic Kubernetes Service (EKS), and Fargate services on edge-cloud computing infrastructure.
7. IntelliPdM was implemented in a large-scale manufacturing unit in Singapore for one year and its performance was well-demonstrated in terms of accuracy, increase in production, reduction in maintenance cost, breakdowns, and downtime.

The rest of the paper is organized as: Section 2 presents the functionalities and working of IntelliPdM, the proposed IIoT based predictive maintenance system implemented on edge-cloud computing infrastructure. Section 3 details the real-time operational data, synthetic data generation framework, implementation details of IntelliPdM, and validation in a real-time manufacturing industry environment. Section 4 concludes the paper with future directions.

2. Intelligent predictive maintenance framework

Fig. 2 presents a detailed architecture of IntelliPdM, an end-to-end intelligent framework that embeds big data analytics, edge, and cloud computing technologies for predictive maintenance and condition monitoring of manufacturing assets. The proposed framework comprises three layers, namely the data acquisition and ingestion layer, the predictive modelling layer, and the application layer to provide accurate decisions and recommendations on fault detection and classification in the different machinery located in any manufacturing system. The functionality of each layer is detailed as,

- **Data acquisition and ingestion layer:** Common equipment and components in the manufacturing industry include pumps, motors, compressors, bearings, conveyors, heat exchangers, valves and actuators, heat exchangers, switchgear, and transformers for which the implementation of predictive maintenance strategies are crucial to ensure efficient functioning of equipment and prevent downtime. The data acquisition layer gathers real-time operational data from IoT sensors strategically placed on machinery and from cameras. When real-time data is unavailable, it utilizes benchmark PdM datasets and synthetic data generation framework detailed in Subsection 3.1. To facilitate predictive maintenance, the operational data for the 'normal' and 'faulty' conditions are generally collected from sensors (vibration sensors, temperature sensors, pressure sensors, and acoustic sensors) and cameras (high resolution or infrared thermal imaging cameras). Edge computing devices also known as gateways or edge nodes are installed close to the equipment to collect, preprocess, and analyze data locally to reduce latency and bandwidth usage (Solution to RQ5).

To enforce data security, a secure FTP server with a suitable communication protocol and security measures acts as an intermediary between the sensors, gateway, and database. The real-time operational data provided by the sensors and cameras is pre-processed using data cleaning and transformation techniques (such as noise removal, handling missing values, data normalization, data synchronization, and

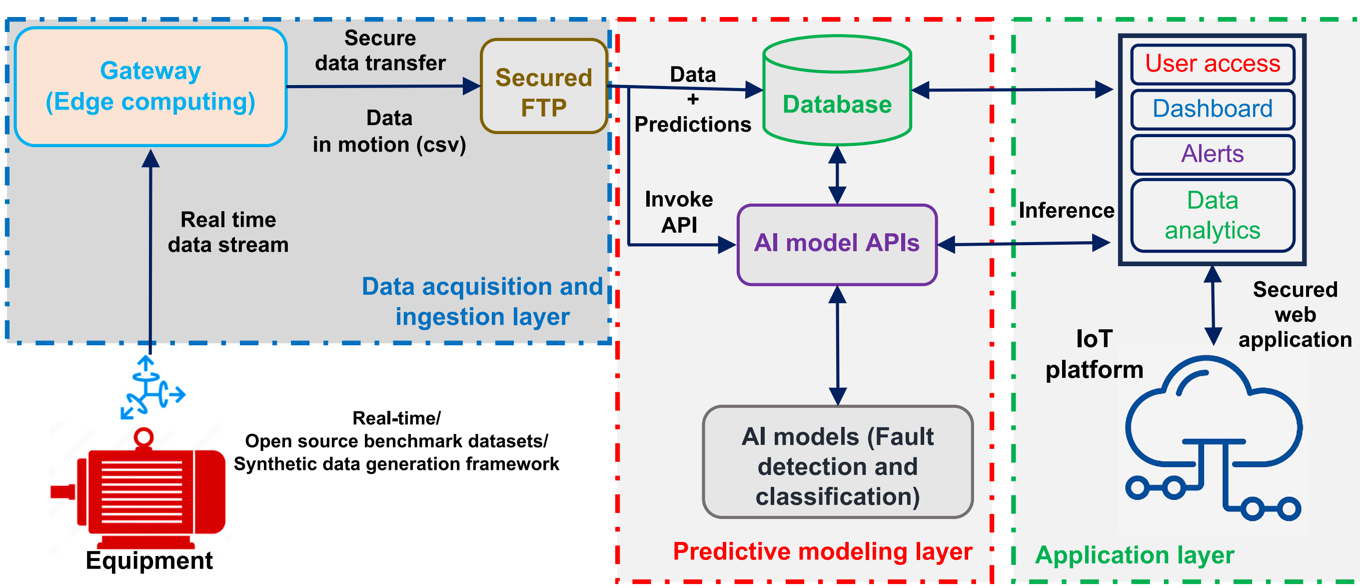

Fig. 2. Detailed architecture of IntelliPdM.

video/image segmentation) to prepare the data for further analysis. Subsequently, the processed data is ingested into appropriate data pipelines and stored in suitable databases for further processing. Separate data pipelines (sensor, video, and audio) and data stores are used for structured (MongoDB) and unstructured (Amazon S3) data (Solution to RQ2). Feature extraction or selection techniques are then applied to identify the informative features indicative of the machine's health and performance.

- **Predictive modelling layer:** The pre-processed data (real-time and historical) and decisions by the edge computing model are stored in the appropriate data store for efficient
- data organization and management. A set of machine learning (sensor data) and deep learning or computer vision (video/image) models are trained using historical data with 'normal' and 'faulty' conditions to detect potential faults, anomalies, and degradations in the manufacturing equipment. The best-performing models assessed and selected based on the different quality metrics were deployed with respect to the nature of the data and predictive objectives. The trained models can either be deployed on the gateway (edge) or cloud which is completely based on the compute power of the gateway. The trained predictive models are deployed using APIs thereby enabling the external systems or dashboard applications, to interact with the predictive models and real-time querying of the models to obtain predictions, recommendations, and insights (Solution to RQ5).
- **Application layer:** It consists of self-developed software which includes a dashboard, data analytics, alert generation, and reporting functionalities to assist the user with real-time monitoring, fault reporting, and predictive maintenance of the manufacturing assets. The dashboard has been designed and implemented to,

a. Visualize the real-time equipment data from various plants.
b. Monitor the real-time sensor and equipment data from various plants.
c. Generate alerts/alarms when equipment is running in abnormal conditions.
d. Generate real-time reports and graphs on the running conditions of the equipment and sensor data based on user queries.

2.1. Implementation details

The technologies used to implement IntelliPdM on an edge-cloud computing platform are, (i) **Python:** build IntelliPdM and AI-ML algorithms, (ii) **Kafka message broker:** handle live streaming data, (iii) **Apache Spark:** ingest live data streams into data store, (iv) **MongoDB and Amazon S3:** stage and store real-time operational data and predictions, and (iv) **Power BI:** visualization of anomalies and faulty patterns. The real-time machine health data ('normal,'faulty') of different manufacturing assets were collected from the vibration sensors strategically positioned at key locations, including the blower, drive end, and non-drive end. Recommendations from the vibration domain expert on the required number of sensors for each equipment, sensor position, identification of outlier frequencies (noise), and data annotations were considered. The acquired data from the sensors are streamlined through appropriate data pipelines, i.e., based on sensor and equipment type to the gateway (edge device). A step-by-step procedure of edge device registration and edge provisioning is detailed as,

1. **Device registration:** register the equipment and sensors with the necessary details of the specifications.
2. **Resource allocation:** allocate resources to the data pipelines for the registered equipment and sensors.
3. **Upstream testing:** ensure proper allocation of resources and update the details of the data pipelines in the configuration file.
4. **Utility package:** contains configuration files (information on available resources, pipeline, and API endpoints) and code files (read data from sensors, edge communication, upstream testing, update configuration).

On successful registration of the equipment and sensors, the real-time operational data are sent to the respective data pipelines. The Kafka message broker creates a topic for each data pipeline, wherein each topic has multiple partitions. The Spark streaming jobs process the operational data from the different data pipelines and store them in the data store (MongoDB – structured data and Amazon S3 – unstructured data). The predictive modeling can be carried out either on the edge device or cloud, the decision is completely based on the compute power of the edge device. To enforce data security, the pre-processed data and the prediction results are transferred to the cloud data store through a secured file transfer protocol with appropriate security mechanisms and configurations. The interaction between the major components involved in data acquisition, ingestion, and storage occurs, (i) operational data provided by the sensors are published to the pre-defined topics on the Kafka server, (ii) a consumer at the Kafka server subscribes to the appropriate topics where Kafka publishes the operational data, and (iii) Spark streaming jobs provides the pre-processed data and predicted results from the edge device are staged and stored in the cloud data store using secured FTP. This entire process is executed using, (i) **IoT utility:** provides information on the available equipment and sensors to the upstream utility; enables the edge device to establish a secure connection with the server; collects the operational data from the sensors and transforms them into compatible format (CSV or JSON), and (ii) **upstream utility:** publish the data into appropriate data pipelines topics; edge server consumes the data, preprocess, stages, and stores in the cloud data store using spark streaming jobs and secured FTP. The pre-processed data from the data store are further processed by the predictive modeling layer and application layer to provide recommendations/alerts on equipment health and condition monitoring (usage, pattern, and anomalies), predicted equipment failures, and optimized maintenance scheduling and planning. The core application components are containerized, and the container orchestration is managed by Kubernetes. For code and data safety, the core components of the application are deployed as a docker image and executed on the docker container. The predictive maintenance application architecture was deployed using AWS Elastic Kubernetes Service (EKS) and Fargate services.

3. Industrial case study

3.1. Dataset

The real time machine health data for the experimental analysis was obtained from a predictive maintenance testbed at an energy research and development centre in Western Europe to generate fault scenarios for the objective and comparative evaluation of PdM solutions. A 3-phase induction motor operating at 400Volt and 50 Hz was used to collect the operational data of the five faults namely, belt loosening, bearing conditions, shaft alignment, load coupling, and voltage & current fluctuations. Table 2 presents the fault detection scenarios and specifications.

The operational data of the 3-phase induction motor were obtained using the vibration sensors (Murata wireless vibration sensor unit – LBAC0ZZ1LZ) strategically positioned based on the operating conditions of the induction motor. The vibration sensor used in this study is a single axial sensor, therefore six sensors, i.e., two sensors at the drive end, two sensors at the non-drive end, and two sensors on the load (blower) were used to collect the vibration data from the induction motor. The hardware setup required to collect the vibration data includes, (i) **multiple vibration sensors:** to monitor and collect vibration data, (ii) **gateway:** to collect and process the data from sensors, (iii) **4G LTE router:** to provide an internet connection to PC, (iv) **PC:** setup with necessary

Table 2
Fault detection scenarios.

Detection scenarios	Affected part	Failure observed	Speed (%)	Duration (h/speed)
Fan motor belt stretched	Original new belt	Fan motor belt stretched	10–100 %	0.5 h/speed
	Slack belt		20 %, 40 %, 60 %, 80 %, 100 %	
	Old belt 1			
	Old belt 2			
Fan motor belt misaligned	Original new belt	30 % misalignment	25 %, 50 %, 100 %	0.5 h/speed
	Original new belt	60 % misalignment		
	No belt			
Motor imbalance	Motor	Motor with motor mounting unscrewed	25 %, 50 %, 100 %	0.5 h/speed
Stoppage of fan unit	Fan	Partial stoppage of fan unit with rotation	25 %, 50 %, 100 %	Risk of damages to the motor. Control it with the follow-up of the electric consumption
		Partial stoppage of fan unit with no rotation		
Motor mechanical fault	Rotor bar	Rotor bar broken	25 %, 50 %, 100 %	0.5 h/speed
		Rotor bar misaligned		
	Phase problem	Phase problem connection		
	Electric problem	Electric problem – overload	25 %, 50 %, 100 %	Risk of damages to the motor. Control it with the follow-up of the electric consumption

software (big data, edge, and cloud technologies) to analyze and store the pre-processed data. The operational data was collected over 6 months at 30-minute granularity and the interval is customizable according to the business requirements. The data format of the vibration sensor is given in Table 3. The FFT mode of the vibration sensor is enabled to collect the operational data in the data format. Freq_n and Acc_n correspond to the frequency and acceleration of top-5 peaks in the frequency spectrum obtained by Fast Fourier Transform (FFT) analysis of the measured time-domain vibration data. The RMS and kurtosis of the vibration signals are the key features for anomaly detection in manufacturing assets.

The operational data from the testbed collected for six months at 30 min granularity contains insufficient samples for different classes of 'faulty' data. Therefore, SMARTHome – synthetic data generation framework in [37] designed using Python scripts and necessary packages was used to generate synthetic data for predictive maintenance in manufacturing assets (Solution to RQ1). Appropriate modifications were made in the SMARTHome framework to generate real-time operational data for a wide range of faults in different manufacturing machinery. The use of a synthetic generation module addresses the challenges in the benchmark datasets, (i) benchmark datasets contain 'normal' and 'faulty' data samples generated for specific machinery specifications, (ii) real-time faulty scenarios cannot be mimicked for data collection period spanning over months, (iii) operational schedule and faulty scenarios differ with respect to the nature of the machinery, (iv) difficult to collect data for 'high-risk' faults, and (v) high cost and time involved in setting up the testbed environment. Different scenarios modeled based on the faults occurring in the manufacturing machinery are the inputs to the synthetic data generation module to generate operational data similar to the real-time manufacturing environment. The synthetic data generation module generates real-time operational data based on the logic in the primary data sources such as benchmark PdM datasets, limited real-time data, or well-defined use cases and scenarios (Solution to RQ3). Therefore, this work uses the limited real-time operational data generated from the testbed as an input logic to the synthetic data generation framework to generate the required number of real-time operational data samples for each class of 'faulty' and 'normal' data. Table 4 presents the modules in the SMARTHome, a synthetic data generation framework specifically designed to provide synthetic data based on the limited real-time operational data for different faults in the machinery. The data and more details on the testbed (layout, specifications, and configuration) used in this study will be made available upon reasonable request. The reliability of the synthetic operational data was validated by the domain experts in industrial systems to ensure a realistic simulation of machine behaviours, operational conditions and failure patterns in real-world scenarios. The operational data is further subjected to important data preprocessing steps to ensure high-quality data, thereby enhancing the robustness of intelligent models, and enabling accurate fault detection and diagnosis (Table 5).

3.2. Experimental analysis

IntelliPdM, an end-to-end intelligent framework for predictive maintenance and condition monitoring of the manufacturing assets was designed and implemented using essential Python libraries, Apache Spark, Apache Kafka, MongoDB, PowerBI, and AWS cloud services. The challenges in the multi-dimensional operational data generated from the synthetic data generation module (high frequency data and consistent frequency peaks) were handled through the application of appropriate data preprocessing techniques. The presence of outlier frequencies in the data was verified with the vibration analysis domain expert and then addressed by transposing the data and removing the high frequency outliers generally caused by sensor malfunctions, measurement errors, or any systematic issues. The Z-score normalization technique was implemented using StandardScaler from Scikit-learn to transform the features in the dataset to have a mean of '0′ and standard deviation of '1′. Further, principal component analysis was implemented using sklearn. decomposition.PCA (Principal Component Analysis) to identify the informative feature subset to improve model performance, generalization, and interpretability with reduced computational complexity. The data annotation of each sample generated in the testbed and synthetic data generation module was carried out using a supervised approach based on the scenarios ('normal' and 'faulty'). In addition, the machinery specification which is highly correlated with the fault type is used as an additional feature for labeling the fault type, i.e., fault classification. Table 6 presents the machine specifications of the different machinery faults.

The predictive maintenance application consists of two phases, namely (i) **fault detection:** detect the faults/anomalies in the manufacturing asset with probability, e.g., "96 % Probability that current data contains anomaly", and (ii) **fault classification:** classify the detected fault and return alert with fault type and probability, e.g., "Belt misalignment": 95 %, "Belt looseness": 2 %, "No Fault": 3 %. A set of machine learning and deep learning models like SVM, isolation forest, random forest, ANN, autoencoder, 1D-CNN, and LSTM were modelled and evaluated with the train-valid-test (70–20-10) split of the pre-

Table 3
Data format of the operational data.

Timestamp	Freq_1 (Hz)	Acc_1(m/s2)	Freq_2 (Hz)	Acc_2(m/s2)	Freq_3 (Hz)	Acc_3(m/s2)	Freq_4 (Hz)	Acc_4(m/s2)	Freq_5 (Hz)	Acc_5(m/s2)	Acc_RMS(m/s2)	Kuttosis	Temperature (°C)	Battery voltage (V)	Status
07/07/2022 0:00	12	0.3	37	0.22	62	0.22	87	0.22	125	0.15	0.13	1.69	20.65	3.49	Normal
07/07/2022 0:30	12	0.3	37	0.22	62	0.22	87	0.15	100	0.15	0.13	1.21	20.65	3.5	Normal
07/07/2022 01:00	12	0.22	37	0.22	62	0.22	87	0.07	100	0.07	0.12	1.43	20.65	3.5	Normal
07/07/2022 01:30	12	0.3	37	0.22	62	0.22	75	0.22	150	0.22	0.15	12.76	20.65	3.5	Normal
07/07/2022 23:30	12	0.3	37	0.22	62	0.22	75	0.22	112	0.22	0.15	13.3	20.65	3.5	Normal

Table 4
SMARTHome modules.

Module	Functionalities
Temperature module	Temperature of the various components of the machinery
Vibration module	Vibrations and frequencies associated with different faults
Noise module	High-frequency noise of defects
Operational module	Speed, motor current, and power consumption
Displace and misalignment module	Shaft and coupling misalignment measurements
Historical module	Load cycles, machine usage, replacements, historical faults/repairs
Load/Force/Lubrication module	Strain/stress and force/load in components; lubrication levels

Table 5
Data preprocessing.

Task	Technique used
Missing value	Mean based imputation
Scaling	Z-score standardization
Outlier detection	Fast Fourier Transform (FFT); Remove high-frequency components
Feature extraction	Time-Domain Features (e.g., RMS, Kurtosis, Skewness, etc.)

Table 6
Additional inputs for labeling faulty data.

S. No.	Machine faults	Additional data inputs	Features
1	Bearing defects	Bearing specification	no. of bearing balls, diameter of ball, pitch, shaft rate, and contact angle
2	Breakage of fan blades	Fan specification	RPM
3	Stoppage of fan unit		
4	Fan motor belt stretched	Belt specification	Pich diameter, and belt length
5	Fan motor belt misaligned		
6	Motor imbalance	Motor specification	no. of poles, frequency of electric power grid, synchronous speed, slip frequency, no. of stator winding slot, and no. of rotor bars
7	Phase problems		
8	Open or short windings of rotor		
9	Eccentric rotor		
10	Broken rotor bar		
11	Gearbox	Gear specification	no. of teeth on pinion
12	Overloading	Temperature	

processed dataset for fault detection and classification. The model fine-tuning through hyperparameter optimization was done using the halving grid search optimization technique (sklearn.model_selection.HalvingGridSearchCV).

Table 7 provides a comparative analysis of the considered machine learning and deep learning models over the state-of-the-art PdM frameworks. The best performing models are containerized and deployed as a docker image using a docker container on the edge device, wherein the container orchestration is managed by Kubernetes. The functionality of the containerized model is exposed as API endpoints via. RESTful APIs receive real-time operational data from the machinery, perform inference, and return decisions on machinery health or maintenance needs (Solution to RQ4). To minimize the overall computation, low-compute models are deployed at the edge device and whenever more confidence is required, data will be sent to the cloud for further processing. The predictive maintenance use cases supported by IntelliPdM are provided in Table 8 along with the necessary inputs and

Table 7
Comparative analysis – State-of-the-art PdM frameworks.

Authors	Models	Accuracy	Precision	Recall	F1-Score
Chen et al. [38]	CatBoost	0.9521	0.3680	0.3102	0.3359
	SmoteNC + CatBoost	0.9553	0.3800	0.7905	0.5141
	ctGAN + CatBoost	0.9030	0.4200	0.8104	0.5532
	SmoteNC + ctGAN + CatBoost	0.8756	0.4300	0.8893	0.5857
Mota et al. [39]	Gradient Boosting	0.9400	0.38	0.83	0.50
	SVM	0.9055	0.27	0.88	0.41
Matzka et al. [39]	Bagged Decision Trees	0.9855	0.8500	0.9756	0.9100
Iantovics et al. [40]	BLR	0.9750	0.9800	0.2900	0.4500
Torcianti et al. [41]	RUSBoost Trees	0.9321	0.3200	0.9000	0.4700
Ghasemkhani et al. [42]	Balanced K-Star	0.9890	0.9900	0.9850	0.9870
	Average	0.9430	0.7070	0.6700	0.6030
Considered models for the study	SVM	0.9200	0.7350	0.6200	0.6730
	Isolation Forest	0.8750	0.6900	0.7000	0.6950
	Random Forest	**0.9901**	**0.9876**	**0.9930**	**0.9901**
	XGBoost	0.9800	0.8800	0.9000	0.8890
	Autoencoder	0.8900	0.6700	0.7300	0.6970
	ANN	0.9400	0.8300	0.8100	0.8200
	1D-CNN	0.9650	0.8600	0.8900	0.8740

Table 8
Use cases in IntelliPdM – Input and output.

Use case	Input	Output
Anomaly detection	Sensor data	Detection of abnormal patterns or deviations from normal behaviour, e.g., identifying abnormal vibration patterns in a motor that indicate potential bearing wear
Failure prediction	Sensor data, maintenance history	Prediction of machinery failures or degradation before they occur, e.g., forecasting that a specific machine will fail within the next 7 days due to a critical component reaching a wear threshold
Health monitoring	Sensor data, machinery specifications, historical data	Assessment of machinery health based on multiple indicators, e.g., assigning a health score to a compressor based on temperature, pressure, vibration, and lubricant condition
Optimal maintenance scheduling	Predictive maintenance models, maintenance logs	Optimization of maintenance schedules to minimize downtime and maximize machinery availability, e.g., generating an optimized maintenance schedule that prioritizes critical machinery based on predicted failure risks
Remaining useful life estimation	Sensor data, historical maintenance records	Estimation of the remaining useful life of machinery, e.g., predicting that a pump has 30 days of remaining useful life before maintenance is required

outputs.

The real-time implementation of the IntelliPdM framework was carried out in a large manufacturing unit located in Singapore. The IoT sensors were strategically positioned and installed at different machinery with the help of vibration domain experts. The sensors continuously record the real-time operational data such as vibration, temperature, pressure, speed, velocity, humidity, etc. at per second granularity. The recorded sensor values remain stable under 'normal' conditions and exhibit spikes or significant changes in case of anomalies, i.e., faults. The spikes in the recorded sensor values, i.e., frequency peaks vary with respect to the type of fault. For example, the frequency peaks of shaft misalignment are different from the bearing faults. The API endpoints of the deployed lightweight model receive the real-time operational data through appropriate data pipelines, process them, and send the decisions over secured FTP to the data lake and core application component in the application layer. Over 12 months, IntelliPdM achieves (i) accuracy of 93–95 %, (ii) 25–30 % reduction in maintenance costs, (iii) 70–75 % elimination of breakdowns, (iv) 35–45 % reduction in downtime, (v) 20–25 % increase in production, and (vi) 10x return on investment (Solution to RQ4).

The execution time for the IntelliPdM pipeline depends on the specific stage of processing and the deployment configuration. For the reported experiments, the data ingestion and preprocessing stage took approximately 5–15 s per batch of 10,000 records. Model inference for fault detection and classification at the edge was completed in under 500 ms per record, ensuring low-latency predictions suitable for real-time operations. The framework's scalability was tested by increasing the dataset size incrementally, from 10 GB to 100 GB. Leveraging Apache Spark and Kafka for distributed processing, IntelliPdM demonstrated linear scalability with a negligible increase in processing time per record. The system maintained consistent performance across dataset growth, with resource utilization optimized by dynamic load balancing in the edge-cloud architecture. These results confirm the framework's ability to handle high-dimensional, high-frequency data generation typical of real-world industrial environments.

The solutions offered by the IntelliPdM for the research questions formulated based on the challenges in the state-of-the-art IIoT based predictive maintenance framework for manufacturing systems are, (i) **RQ1:** What can be a 'one-stop solution' for data challenges related to real-time operational data and open-source datasets? SMARTHome, a synthetic data generation framework was properly configured to generate real-time operational machine health data for various manufacturing machinery, (ii) **RQ2:** What is the effective strategy to handle heterogeneous data from multiple sensors? A separate data pipeline is designed to ingest structured and unstructured data into the appropriate data store, (iii) **RQ3:** How to effectively consider multi-fault aspects of fault diagnosis in manufacturing assets? The real-time operational machine health data for multi-faults in different machinery was generated by the synthetic data generation framework based on the well-defined faulty scenarios or use cases, (iv) **RQ4:** How to demonstrate the effectiveness of the trained model in a real-time industrial environment? The best performing model deployed on the edge/cloud is exposed as API to provide real-time inference in a real time industrial environment and the effectiveness of the IntelliPdM was assessed in terms of accuracy, maintenance cost, breakdowns, downtime, production, and return on investment, and (v) **RQ5:** What are the potential deployment options for resource-constrained computing environment? Edge-cloud deployment architectures are best suited for resource-constrained environments, wherein sensitive data are processed locally on the edge; insights from the locally processed data, and complex data (audio, video, and images) are processed and staged on the cloud computing infrastructure ensuring efficient use of bandwidth, security and privacy, near real-time data analysis and decision-making & minimal cost and latency.

4. Conclusions

The proposed IntelliPdM framework demonstrates a robust and scalable solution for PdM in smart manufacturing environments, leveraging advanced AI techniques, scalable edge-cloud infrastructure, and a robust synthetic data generation framework. Key results include deployment in the large-scale manufacturing unit, resulting in high fault detection accuracy, minimized breakdowns and reduced costs and downtime contributing to a more streamlined and efficient operation. The integration of real-time data acquisition and synthetic data generation ensured comprehensive fault representation for robust AI model training, while the flexibility to deploy models at the edge or cloud highlights the adaptability to real-world applications. The IntelliPdM framework has broad real-world applications across industries prioritizing operational efficiency, safety, and cost reduction. It is highly effective in smart manufacturing for monitoring machinery health and reducing downtime, in oil and gas or power utilities for ensuring equipment reliability, and in transportation for optimizing fleet maintenance and safety. Additionally, it supports critical equipment maintenance in healthcare, public utility management in smart cities, and enhanced uptime in retail and warehousing operations.

Despite its promising outcomes, IntelliPdM has limitations on (i) effectiveness on synthetic data generation framework depends on the accuracy of initial real-time data and scenarios/assumptions encoded in the framework, (ii) computational demand of training and deploying advanced AI models for unstructured data pose challenges for resource-constrained environments, limiting the framework's accessibility and scalability, and (iii) adaptability to rapidly evolving industries requires significant customization and domain-specific tuning, increasing complexity and deployment time.

Future efforts on IntelliPdM will focus on (i) exploring advanced techniques like GANs and domain-adaptation algorithms to improve data quality and model robustness, (ii) optimizing AI models for resource-constrained environments by exploring lightweight algorithms, on-device processing, and advanced compression techniques, (iii) expanding framework's adaptability through modular components, domain-agnostic APIs, and improved cybersecurity measures. Additionally, emerging technologies like blockchain and federated learning will be integrated, with pilot implementations planned in domains such as renewable energy, smart cities, and healthcare to validate and extend its applicability.

CRediT authorship contribution statement

Nivethitha Somu: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Nirupam Sannagowdara Dasappa:** Supervision, Resources, Project administration, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Nirupam Sannagowdara Dasappa reports financial support was provided by ENGIE Lab Singapore Pte Ltd.

Acknowledgments

This work was funded by the ENGIE Lab Singapore Pte Ltd (Ref. REQ0175585). The authors thank the Energy Research Institute at Nanyang Technological University (ERI@N), Singapore for their consistent support.

Appendix

A. Murata vibration sensor specification

Part Number		LBACOZZ1LZ
Size (Typ)		37 mm 24 mm 38 mm(H)
Weight		45 g
Communication Interface	Radio (Japan Config)	Proprietary SubGHz Radio (ARIB STD-T108 Compliant)
		Center Frequency: 923.7 MHz-926.9 MHz
		Output Power: 20mW
	Range	200 m (Line of Sight) (*1)
Indicator	LED	Red 1Pc, Green 1Pc
Configuration interface	Push button	1Pc
	AMR switch	1Pc
Measurement parameter		Peak Frequency (top 5), Peak acceleration (top 5), RMS, Kurtosis, Temperature
Sensor	Detection frequency	12.5 ~ 10,000 Hz
	Acceleration	Absolute Max Rating ± 30GResolution: 28mG
	Temperature	0–85 °C Resolution: 0.1 °C 25 °C
Measurement accuracy (Typ)	Vibration	±3dB
	Temperature	±1°C 25 °C
Power source (*2)	Battery	1/2AA Lithium Battery 1pc
Battery life (*3)		5 years (*4)
Current consumption		30 mA(Tx mode), SuA (sleep mode) IP65
Water resistance	IP Grade	IP65
Operating range	Body (Ambient)	−10–60 °C −90 % RH (non-condensing)
	Sensor (Surface)	−10–85 °C < 90 % RH (non-condensing)
Accessories		Battery Cover 1pc, Battery 1pc
Installation method		Magnet or Screw

Data availability

The data that has been used is confidential.

References

[1] Kusiak A. Smart manufacturing must embrace big data. Nature 2017 544:7648 2017;544:23–5. https://doi.org/10.1038/544023a.

[2] S. Dowlatshahi, The role of industrial maintenance in the maquiladora industry: An empirical analysis, Int. J. Prod. Econ. 114 (2008) 298–307, https://doi.org/10.1016/J.IJPE.2008.02.009.

[3] B. Luo, H. Wang, H. Liu, B. Li, F. Peng, Early fault detection of machine tools based on deep learning and dynamic identification, IEEE Trans. Ind. Electron. 66 (2019) 509–518, https://doi.org/10.1109/TIE.2018.2807414.

[4] J. Lee, B. Bagheri, H.A. Kao, A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems, Manuf. Lett. 3 (2015) 18–23, https://doi.org/10.1016/J.MFGLET.2014.12.001.

[5] A. Rohani, M.H. Abbaspour-Fard, S. Abdolahpour, Prediction of tractor repair and maintenance costs using Artificial Neural Network, Expert Syst. Appl. 38 (2011) 8999–9007, https://doi.org/10.1016/J.ESWA.2011.01.118.

[6] J. Lee, H.A. Kao, S. Yang, Service Innovation and Smart Analytics for Industry 4.0 and Big Data Environment, Procedia CIRP 16 (2014) 3–8, https://doi.org/10.1016/J.PROCIR.2014.02.001.

[7] T. Stock, G. Seliger, Opportunities of Sustainable Manufacturing in Industry 4.0, Procedia CIRP 40 (2016) 536–541, https://doi.org/10.1016/J.PROCIR.2016.01.129.

[8] K. Komonen, A cost model of industrial maintenance for profitability analysis and benchmarking, Int. J. Prod. Econ. 79 (2002) 15–31, https://doi.org/10.1016/S0925-5273(00)00187-0.

[9] Y. Jiang, P. Dai, P. Fang, R.Y. Zhong, X. Zhao, X. Cao, A2-LSTM for predictive maintenance of industrial equipment based on machine learning, Comput. Ind. Eng. 172 (2022) 108560, https://doi.org/10.1016/J.CIE.2022.108560.

[10] W. Yu, Y. Liu, T. Dillon, W. Rahayu, Edge computing-assisted IoT framework with an Autoencoder for fault detection in manufacturing predictive maintenance, IEEE Trans. Industr. Inform. 19 (2023) 5701–5710, https://doi.org/10.1109/TII.2022.3178732.

[11] HanAra Software. Predict and Prevent: Improving Your Maintenance Strategy n.d. https://www.hanarasoft.com/preventive-maintenance-vs-predictive-maintenance/ (accessed March 12, 2024).

[12] I. Alsyouf, Maintenance practices in Swedish industries: Survey results, Int. J. Prod. Econ. 121 (2009) 212–223, https://doi.org/10.1016/J.IJPE.2009.05.005.

[13] B. He, K.J. Bai, Digital twin-based sustainable intelligent manufacturing: a review, Adv. Manuf. 9 (2021) 1–21, https://doi.org/10.1007/S40436-020-00302-5/FIGURES/4.

[14] Y. Lu, C. Liu, K.I.K. Wang, H. Huang, X. Xu, Digital Twin-driven smart manufacturing: Connotation, reference model, applications and research issues, Robot Comput. Integr. Manuf. 61 (2020) 101837, https://doi.org/10.1016/J.RCIM.2019.101837.

[15] S. Jeschke, C. Brecher, T. Meisen, D. Özdemir, T. Eschert, Industrial internet of things and cyber, Manuf. Syst. (2017) 3–19, https://doi.org/10.1007/978-3-319-42559-7_1.

[16] J. Lee, F. Wu, W. Zhao, M. Ghaffari, L. Liao, D. Siegel, Prognostics and health management design for rotary machinery systems—Reviews, methodology and applications, Mech. Syst. Signal Process. 42 (2014) 314–334, https://doi.org/10.1016/J.YMSSP.2013.06.004.

[17] Bampoula X, Siaterlis G, Nikolakis N, Alexopoulos K. A Deep Learning model for predictive maintenance in cyber-physical production systems using LSTM Autoencoders. Sensors 21 (2021) 972 https://doi.org/10.3390/S21030972.

[18] S. Ayvaz, K. Alpay, Predictive maintenance system for production lines in manufacturing: A machine learning approach using IoT data in real-time, Expert Syst. Appl. 173 (2021) 114598, https://doi.org/10.1016/J.ESWA.2021.114598.

[19] A. Ouadah, L. Zemmouchi-Ghomari, N. Salhi, Selecting an appropriate supervised machine learning algorithm for predictive maintenance, Int. J. Adv. Manuf. Technol. 119 (2022) 4277–4301, https://doi.org/10.1007/S00170-021-08551-9/FIGURES/34.

[20] J. Figueroa Barraza, L. Guarda Bräuning, R. Benites Perez, C.B. Morais, M. R. Martins, E.L. Droguett, Deep learning health state prognostics of physical assets in the Oil and Gas industry. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability 236 (4) (2022) 598–616.

[21] Abidi MH, Mohammed MK, Alkhalefah H. Predictive Maintenance Planning for Industry 4.0 Using Machine Learning for Sustainable Manufacturing. Sustainability 14 (2022) 3387-2022. https://doi.org/10.3390/SU14063387.

[22] M. Shcherbakov, C. Sai, A hybrid deep learning framework for intelligent predictive maintenance of cyber-physical systems, ACM Trans. Cyber-Phys. Syst. 6 (2022) 1–22, https://doi.org/10.1145/3486252.

[23] N. Ringler, D. Knittel, J.-C. Ponsart, M. Nouari, A. Yakob, D. Romani, Machine learning based real time predictive maintenance at the edge for manufacturing systems: A practical example, in: 2023 IEEE IAS Global Conference on Emerging Technologies (GlobConET), IEEE, 2023, pp. 1–6, https://doi.org/10.1109/GlobConET56651.2023.10150033.

[24] Y.K. Teoh, S.S. Gill, A.K. Parlikad, IoT and fog-computing-based predictive maintenance model for effective asset management in Industry 4.0 using machine learning, IEEE Internet Things J 10 (2023) 2087–2094, https://doi.org/10.1109/JIOT.2021.3050441.

[25] Predictive Maintenance Modelling Guide Data Sets | Azure AI Gallery n.d. https://gallery.azure.ai/Experiment/Predictive-Maintenance-Implementation-Guide-Data-Sets-1 (accessed March 4, 2024).

[26] Ghasemkhani B, Aktas O, Birant D. Balanced K-Star: An Explainable machine learning method for internet-of-things-enabled predictive maintenance in manufacturing. Machines 2023, Vol 11, Page 322 2023;11:322. https://doi.org/10.3390/MACHINES11030322.

[27] P. Murugiah, A. Muthuramalingam, S. Anandamurugan, A design of predictive manufacturing system in IoT-assisted Industry 4.0 using heuristic-derived deep learning, Int. J. Commun Syst 36 (2023) e5432.

[28] X. Wang, M. Liu, C. Liu, L. Ling, X. Zhang, Data-driven and Knowledge-based predictive maintenance method for industrial robots for the production stability of intelligent manufacturing, Expert Syst. Appl. 234 (2023) 121136, https://doi.org/10.1016/J.ESWA.2023.121136.

[29] R. Rosati, L. Romeo, G. Cecchini, F. Tonetto, P. Viti, A. Mancini, et al., From knowledge-based to big data analytic model: a novel IoT and machine learning based decision support system for predictive maintenance in Industry 4.0, J. Intell. Manuf. 34 (2023) 107–121, https://doi.org/10.1007/S10845-022-01960-X/FIGURES/8.

[30] J.M. Fordal, P. Schjølberg, H. Helgetun, T.Ø. Skjermo, Y. Wang, C. Wang, Application of sensor data based predictive maintenance and artificial neural networks to enable Industry 4.0, Adv. Manuf. 11 (2023) 248–263, https://doi.org/10.1007/s40436-022-00433-x.

[31] V. Justus, G.R. Kanagachidambaresan, Machine learning based fault-oriented predictive maintenance in industry 4.0, Internat. J. Syst. Assur. Eng. Manag. 15 (2024) 462–474, https://doi.org/10.1007/S13198-022-01777-0/FIGURES/14.

[32] B.A. Ture, A. Akbulut, A.H. Zaim, C. Catal, Stacking-based ensemble learning for remaining useful life estimation, Soft Comput. 28 (2024) 1337–1349, https://doi.org/10.1007/S00500-023-08322-6/TABLES/6.

[33] P. Khazaelpour, S.H. Zolfani, FUCOM-optimization based predictive maintenance strategy using expert elicitation and Artificial Neural Network, Expert Syst. Appl. 238 (2024) 121322, https://doi.org/10.1016/j.eswa.2023.121322.

[34] S. Elkateb, A. Métwalli, A. Shendy, A.E.B. Abu-Elanien, Machine learning and IoT – Based predictive maintenance approach for industrial applications, Alex. Eng. J. 88 (2024) 298–309, https://doi.org/10.1016/j.aej.2023.12.065.

[35] X. Chen, X. Li, S. Yu, Y. Lei, N. Li, B. Yang, Dynamic Vision Enabled Contactless Cross-Domain Machine Fault Diagnosis with Neuromorphic Computing, IEEE/CAA J. Autom. Sin. 11 (2024) 788–790, https://doi.org/10.1109/JAS.2023.124107.

[36] X. Li, S. Yu, Y. Lei, N. Li, B. Yang, Intelligent Machinery Fault Diagnosis With Event-Based Camera, IEEE Trans. Industr. Inform. 20 (2024) 380–389, https://doi.org/10.1109/TII.2023.3262854.

[37] N.S. Dasappa, G.K. Kumar, N. Somu, Multi-sensor data fusion framework for energy optimization in smart homes, Renew. Sustain. Energy Rev. 193 (2024) 114235, https://doi.org/10.1016/J.RSER.2023.114235.

[38] Chen CH, Tsung CK, Yu SS. Designing a hybrid equipment-failure diagnosis mechanism under mixed-type data with limited failure samples, Appl. Sci. 12 (2022) 9286-2022. https://doi.org/10.3390/APP12189286.

[39] Matzka S. Explainable artificial intelligence for predictive maintenance applications, in: Proceedings - 2020 3rd International Conference on Artificial Intelligence for Industries, AI4I 2020, 2020, pp. 69–74. https://doi.org/10.1109/AI4I49448.2020.00023.

[40] L.B. Iantovics, C. Enăchescu, Method for data quality assessment of synthetic industrial data, Sensors 22 (2022) 1608. https://doi.org/10.3390/S22041608.

[41] A. Torcianti, S. Matzka, Explainable artificial intelligence for predictive maintenance applications using a local surrogate model, in: 4th International Conference on Artificial Intelligence for Industries (AI4I), 2021, pp. 86–88. https://doi.org/10.1109/AI4I51902.2021.00029.

[42] B. Ghasemkhani, O. Aktas, D. Birant, Balanced K-star: An explainable machine learning method for internet-of-things-enabled predictive maintenance in manufacturing, Machines 11 (2023) 322. https://doi.org/10.3390/MACHINES11030322.