International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Exploring the Use of Cloud-Based AI and ML for Real-Time Anomaly Detection and Predictive Maintenance in Industrial IoT Systems

[1]Rishabh Rajesh Shanbhag, [2]Ugandhar Dasi, [3]Nikhil Singla, [4]Rajkumar Balasubramanian, [5]Siddhant Benadikar

Submitted: 02/08/2023 **Revised**: 25/09/2023 **Accepted**: 05/10/2023

Abstract: This comprehensive research paper delves into the application of cloud-based Artificial Intelligence (AI) and Machine Learning (ML) technologies for real-time anomaly detection and predictive maintenance in Industrial Internet of Things (IIoT) systems. The study provides an in-depth analysis of IIoT architecture, the integration of cloud computing with AI/ML techniques, and the challenges associated with implementing these technologies in industrial environments. Through extensive examination of various aspects including anomaly detection methodologies, predictive maintenance strategies, data management techniques, and model development approaches, this paper offers valuable insights into the current state and future potential of cloud-based AI/ML solutions in industrial settings. The research findings underscore the significant benefits of these technologies in enhancing operational efficiency and reliability, while also highlighting the importance of addressing implementation challenges and adapting to emerging trends in the rapidly evolving field of industrial IoT.

Keywords: *Industrial Internet of Things (IIoT), Cloud Computing, Artificial Intelligence, Machine Learning, Anomaly Detection, Predictive Maintenance, Edge Computing, Big Data, Digital Twin*

1. Introduction:

1.1 Background on Industrial IoT Systems

One can easily define the Industrial Internet of Things as a brand-new paradigm in the overall industrial processes and management. IIoT therefore establishes an interconnected system of smart sensors, actuators and superior analytics that pass data continuously and in real-time hence revolutionizing the way industrial processes are managed. The integration of OT and IT has paved the way for increased productivity in numerous industrial segments accompanied by adaptability and quickness.

The history of IIoT began around the year 2000 with the use of machine to machine (M2M) systems. Nonetheless, IIoT really only began to take off around the mid-2010s as technology progressed in the areas of sensors, wireless networking, and data analytics. The progression of IIoT is clearly described in aspects of the market as Grand View Research stated that the Global IIoT Market size was valued at USD 216 billion in 2021 (Amirante & Lamonaca, 2023). 13 billion in-2020 and is estimated to reach $ 65. 45 billion with CAGR of 22% from 2021 to 2026. Also, the disease prevalence rate is projected to increase by 8% within the 2021 to 2028 period.

The use of IIoT has been considered due to its ability to solve traditional problems of industrial processes including equipment availability, resources, and quality. Real-time as well as big data analytics are used by IIoT to derive greater insights on operations and make better decisions, manage resources well and minimize on risks that could arise. Due to such a strategic approach to industrial operations management, there has been a remarkable enhancement in the efficiency, expenses, and functionality of industries.

1.2 Importance of Anomaly Detection and Predictive Maintenance

In real IIoT systems, two significant use-cases have distinguished themselves as prominent: therefore, the applications of anomaly detection and predictive maintenance. Anomaly detection is the process of finding non-conformities in patterns or behaviours with respect to specific activities, or functions of industrial materials or procedures, and therefore, is a warning sign of potential future problems or equipment failure. It is used in order to keep those systems running at steady state and avoid possible catastrophic failures that may cause system downtime and potential safety issues.

On the other, predictive maintenance, this is a strategy that makes use of both historical and real time data in identifying when equipment needs to be maintained. In contrast to the generic time-based or breakdown-based maintenance strategies, the predictive maintenance allows organizations to maximize the plan and minimize the unnecessary or ineffective work, and to prevent

[1]*Independent Researcher,USA.*
[2]*Independent Researcher, USA.*
[3]*Independent Researcher, USA.*
[4]*Independent Researcher, USA.*
[5]*Independent Researcher, USA.*

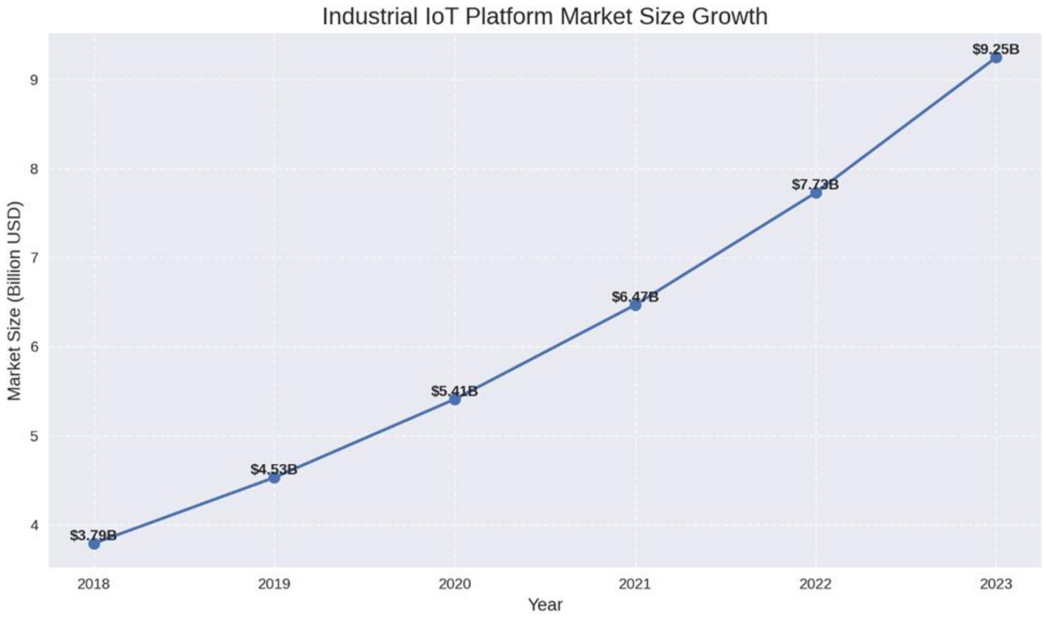
equipment's sudden failures. The effectiveness of these approaches is significant, With Reference to the U. S. Department of Energy indicated that that the predictive maintenance mode has a drastic effect it reduces the maintenance costs by 30% and breakdowns by 75% as well as reducing the downtime by 45%.

The application of anomaly detection and preventive maintenance does not stop at optimizing the overall IIoT systems. These applications are very important in increasing safety levels, raising product quality, and increasing the usage life of industrial assets (Chai, Miao, Sun, Zheng, & Li, 2022). By identifying the shifting standards, it becomes easy to address the anomalies in real time and produce forecasts of potential failures that affect the organization's operation, meet regulatory requirements, and uphold quality production. In addition, such applications can provide understandings that can be useful in the formulation of management decisions for areas such as equipment overhaul strategies and process improvement agendas.

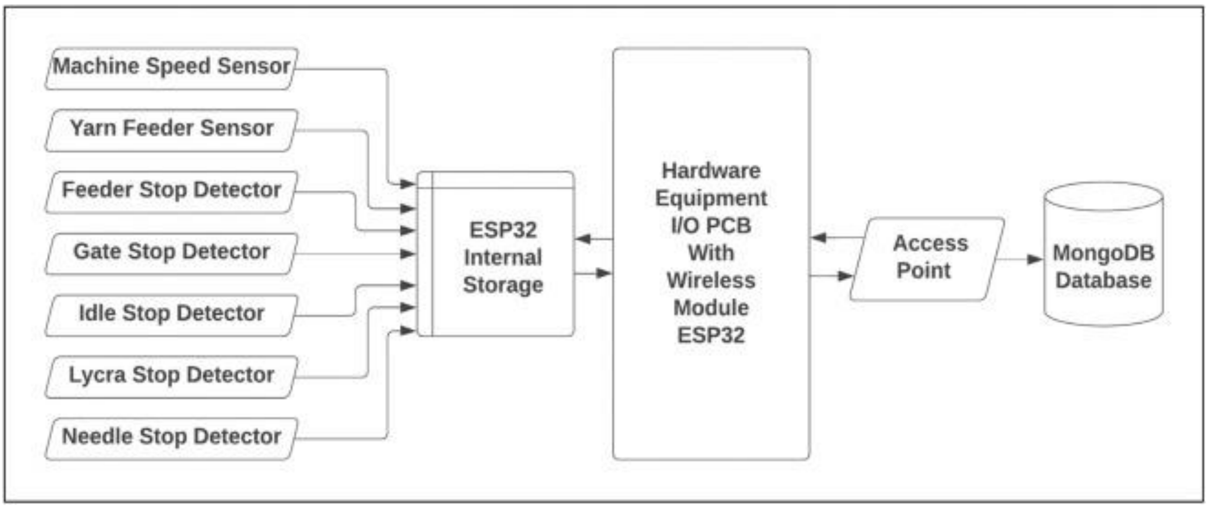
1.3 Research Objectives and Scope

The purpose of this research is to establish the effectiveness of the cloud AI and ML technologies in the evaluation of real-time monitoring and predictive maintenance of IIoT systems. Based on these questions, the purpose of this research is to review and discuss the existing state-of-art technologies, their use scenarios in industries, and the threats and possibilities of adopting these new technologies. Specifically, the research objectives include:

1. Reflecting on the architecture of the IIoT systems, and how cloud computing has been of significance in handling data.
2. Researching number of AI and ML algorithms that are applicable to anomaly detection and predictive maintenance of industrial systems.
3. Analysing the issues connected with real-time processing and using clouds in industrial settings.
4. Collecting results on the efficiency and efficacy of AI/ML-based abnormality identification and predictive maintenance techniques.
5. Discussing various applications in IIoT industries and defining the trends in cloud-based applications.



The time horizon of this research is up to 2021, considering the reflections on the most current technologies of the time. This is related to a diversified field of industries using the IIoT systems and they include manufacture, energy and utilities and transport and logistics. Thus, this work targets to present a valuable resource to the researchers, practitioners, and decision-makers who engage in the application and enhancement of the IoT systems.



2. Industrial IoT Architecture:

2.1 IoT Sensors and Data Collection

The essence of any Industrial Internet of Things (IIoT) is in its capacity to gather large volumes of information, originating from different sources in the industrial setting (Elshawi, Sakr, Talia, & Trunfio, 2018). The process of data collection at the same time is accomplished by the usage of heterogeneous IoT sensors, which could be temperature, pressure, vibration, humidity, energy consumption sensors etc. They are mostly integrated into the machinery, equipment, and production lines as means of monitoring the working status and conditions.

In terms of the size of IoT device implementation, industrial applications can really be described as massive. As mentioned in the IoT Analytics report, there was a forecast of connected IoT devices to be approximately 11. 3 billion by 2021. The presence of sensors has increased over the years thereby augmenting the solution of a large volume, velocity and variety of data in industries. To improve data transmission, many of devices apply specific protocols, for instance, MQTT or OPC UA.

The data collected by these sensors can be broadly categorized into three types:

1. Time-series data: This includes values collected in a manner that consists of several recordings at the same time intervals. This kind of data is preferred when dealing with a situation where gradual changes over extended hours or gradual deviations from the normal operational status are expected.
2. Event-based data: This kind of data is initiated by certain events or when certain predefined limits have been attained. Such data are useful for diagnosing and responding to root cause events in real-time or within exceptionally short periods of time.
3. State data: This gives details on the status or configuration of equipment at the time of writing. State data is necessary for evaluation of other data, collected in a process, as well as for effective decision-making concerning equipment functioning and maintenance.

The performance of the sensing capabilities of the IIoT solutions yielded effective and quality results. Sources like sensors or their variables like calibration, frequency of data collection and many other factors, the external environment can greatly affect the quality of the data gathered (Fahad, Tahir, & Rajarajan, 2021). Therefore, to maintain the IIoT data collection integrity, some of the best practices include calibration and maintenance of the sensors and quality check on the data being collected.

2.2 Edge Computing in Industrial Settings

Thus, edge computing becomes an indispensable part of IIoT infrastructure, which solves problems related to high latency, limited bandwidth, and critical data that cannot be processed in the cloud. Edge computing pre-processes information closer to the data source and with low latency hence enhancing real time decision-making while only transmitting a small amount of data to the cloud.

In industries, edge devices may start from being quite simple, like gateways, and expand to quickly develop into edge servers that can handle complicated computations, including analysis and machine learning. These devices often perform critical tasks such as:

1. Data filtering and aggregation: From raw sensor data, edge devices can filter out noise, average values and generally reduce the amount of data before it is transferred to the cloud. Besides, it also cuts along the bandwidth needed and enhances the signal-to-noise ratio of the data to be analysed.
2. Local anomaly detection: Real-time light-weight anomaly detection algorithms can be run at the edge so that industrial systems can quickly detect and respond to critical issues instead of the time taken in communicating with a cloud.

3. Predictive maintenance calculations: Due to the use of IoT, elements of fog computing or intelligent objects, edge devices can analyse preliminary pieces of equipment health data and quickly address such maintenance issues if any.
4. Control loop optimizations: Where real-time control of a process is needed, then edge computing can enable near real-time closed loop systems that react to inputs to sensors almost immediately.

The future of IIoT prospect helps elaborate the significance of incorporating edge computing in such architectures (Ge, Song, Ding, & Huang, 2017). According to research conducted by Gartner it was postulated that by the year 2025, ¾ of enterprise produced data will be created and managed outside the central data centre or cloud. This change to edge computing is based on the demand of real-time processing, less delay time, and more secure data and privacy.

2.3 Cloud Integration for Data Processing and Storage

Though edge computing will always be a very integral part of the IIoT systems and solutions, the cloud will still be relevant in handling massive amounts of data and performing complex business analysis. AI and ML require the resource needs of computational power and cloud platforms are resourceful in meeting the need of training and deploying large AI and ML models as well as deep analysis of industrial data.

Above all, the combination of cloud services with IIoT systems is primarily of a hybrid nature, in which data is processed both at the level of IoT nodes and in the cloud. This architecture allows for:

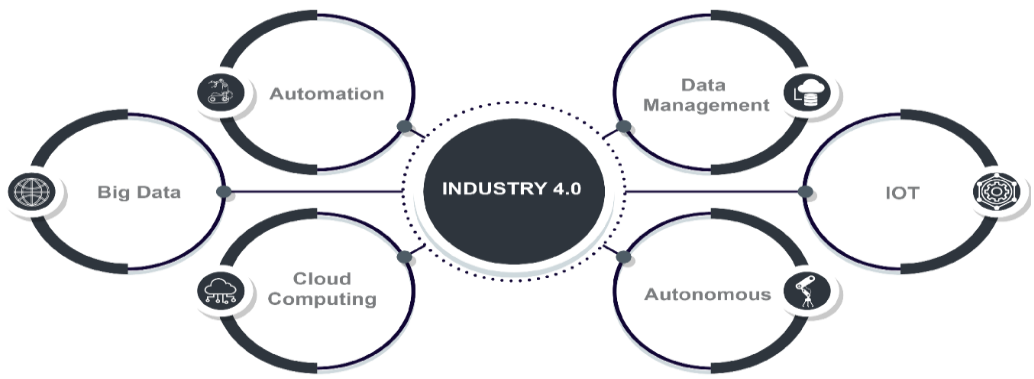
1. Long-term storage of historical data: Cloud platforms provide huge nearly inexhaustible storage space to store an enormous amount of data including detailed histories of the industrial processes. This historical data is very much useful for analysing the trends in the market and comparing the performance of the current models with the models of the past and for training of the Machine Learning models.
2. Training of ML models on large datasets: Cloud enables provision of computing resources for training of advanced forms of ML applicable on large volumes of accumulated data. Such models can then be run at the edge for real time inference.
3. Execution of computationally intensive analytics: Computationally intensive analysis applied on big data which may include simulations or large number figure optimizations can be shifted to the cloud.
4. Centralized management and orchestration of distributed IIoT devices: Cloud platforms in IIoT provide a means of managing and orchestrating the activities of a multitude of edge devices and sensors that is spread across several industrial locations (Guillén, Martínez, Rodríguez-Molina, & Rubio, 2021).

That growing role of cloud integration for the IIoT is further underpinned by the fact that most IIoT platforms are cloud-based. MarketandMarkets has in a research report provided an estimate of the Industrial IoT platform market size which is projecting the market to experience a growth from USD 3. 79 billion in 2018 to USD 9. 25 billion by 2023, at a CAGR of 19 percent. 5%.

3. Cloud-Based AI and ML Technologies

3.1 Overview of Relevant AI and ML Techniques

The use of AI and ML with the IIoT in the cloud boosts the capabilities of the IIoT to create better analytical data and support decisions in industrial settings. These technologies assist the organizations to gain important insights in the big volumes of data produced by the IIoT sensors in order to enhance productivity, quality of products and methods of undertaking maintenance. Some popular techniques include supervised learning this is where data is pre-classified and the algorithm uses this to predict and classify new data. SVM is used for binary classification and novel tends for anomaly detection, RF being decision trees combined for various problems, GBM (e. g. XGBoost) which construct an ensemble of strong hypotheses from weak hypotheses.



Supervised learning manages to discover relationships between variables in the data marked as dependent and independent ones, while the inaccurate one only works with data that is not labelled, which is great for clustering and anomaly detection. Methods like K-means clustering segregate similar data points, Principal Component Analysis (PCA) depreciates the unnecessary features enhancing the crucial ones, and Autoencoders make use of neural networks for detecting the outliers and features extraction.

Specifically, deep learning, which can be defined as a type of machine learning using neural networks, demonstrates high performance in the tasks of pattern matching (Khan, Ahmed, Hakak, Yaqoob, & Ahmed, 2019). Convolutional Neural Networks (CNNs) are used for image and signal processing while Recurrent Neural Networks (RNNs), and Long Short-Term Memory (LSTM) Networks are used for time series prediction and anomaly detection. As a result of studying the dynamics over time of the data related to the industry, the time series analysis is essential; the ARIMA technique is useful for forecasting, and the Prophet is successful at managing significant seasonal impacts. A less used but promising sub-element of machine learning is reinforcement learning, that uses such strategies as QLearning and Deep Q-Network (DQN).

3.2 Cloud Computing Platforms for AI/ML

By using cloud computing, it is much easier to implement AI and ML integral components in IIoT systems since there is flexibility in scaling and computations. AWS also provides specialized services like AWS IoT Core for appropriate device control, Amazon Sage Maker for model building, and AWS Greengrass to make a connection between cloud services to the devices. Windows Azure offers solution that is giving adequate support of device messaging by Azure IoT Hub while the Azure Machine Learning supports end-to-end AI models alternatively Azure IoT Edge supports intelligence distribution on devices (Lee, Davari, Singh, & Pandhare, 2018).

IIoT is supported on GCP through Cloud IoT Core which is used for device management while AI/ML is supported through Cloud Machine Learning Engine for model training and Cloud IoT Edge for deploying the trained ML models over the connected edge devices. IBM Cloud targets industries with Watson IoT Platform for devices and data handling and IBM Watson Machine Learning for model creation and IBM Edge Application Manager for edge computing operation. These platforms provide infrastructure, elastic services, and application programming interfaces that ease the AI/ML model building and deployment processes, thus shifting the value of data to the organization's core operations rather than acting as a data warehouse infrastructure provider.

3.3 Scalability and Performance Considerations

The two requirements, scalability and high performance, are important for cloud-based AI and ML solutions in IIoT because the number of sensors will generate huge amount of data. They do provide solutions to the scalability such as auto-scaling, distributed processing frameworks such as Apache Spark, containerization technologies like Docker and Kubernetes for high performance during the time of high traffic loads.

These performance optimizations are defined by GPU-accelerated computing to learn deep neural networks faster and infer them in real-time. There are specific GPU instances available from the cloud providers to cut down the time for training and analysis of data. Data minimization is also the key in order to minimize the latency and required costs for the communication (Liang, Huang, Long, Zhang, Li, & Zhang, 2020). This is done by having a proper data pipeline and schema design, edge processing, and cloud services that compute data where it is.

Real-time caching and in-memory computing solutions also enhance response times for your real-time analytics and anomaly detection on frequently accessed data and intermediate results that doesn't require disk I/O.

4. Real-Time Anomaly Detection

4.1 Types of Anomalies in Industrial Systems

Another important feature of IIoT systems is that the framework has to be capable of detecting anomalies that reflect equipment failures or inefficiencies as well as intrusions. Anomalies in industrial settings can be classified into three types: Point anomalies are individual data points that are quite different from the rest of the population, like a person developing a fever and being a warning sign of an illness. Contextual anomalies are data points that are out of the norm in a certain context while perfectly acceptable in another, for instance, heightened usage of electricity during times of the night. This anomaly type is quite different from the other one, as it means that all the related data points are somehow problematic, even if they seem rather normal if considered separately.

4.2 Machine Learning Algorithms for Anomaly Detection

Classification of machine learning algorithms used in anomaly detection for industrial systems can be subdivided into supervised learning, semi-supervised learning, and unsupervised learning. Supervised methods such as SVM, Random Forests, and Neural Network use scenarios with labels to try and distinguish between normal and anomalous behaviour. In semi supervised approaches like one-class SVM and autoencoder, normal set of data is learned and in case of any deviation it is classified as an anomalous one (Molina-Solana, Ros, Ruiz, Gómez-Romero, & Martin-Bautista, 2017). Supervised methods work with labelled data, while the unsupervised are without labels; some of the methods include clustering (for instance K-means), Isolation Forest, and Principal Component Analysis (PCA) which aim at identifying points that are abnormally different from the other points.

4.3 Real-Time Processing Challenges and Solutions

Posted by Fuad @ 2021-10-07 08:28:22, Real-time anomaly detection in the IIOT system is difficult because of the abundance of data with regards to velocity and variety. One of the ever-limiting factors is how data streams since many sensors operating in industries produce real-time data; its management is facilitated by Apache Kafka and Flink. Latency is another important aspect, because a lot of industrial processes require anomaly detection in real-time so that they do not suffer losses or face manufacturing problems. Model updating and adaptation are also important because the environments that most industries work in are ever changing thus requiring models that can be easily updated to suit the ever-changing environment. Due to the limitations of resources in the Edge devices, simpler algorithms need to be run on those devices. Solving these problems requires edge-cloud hybrid architecture, machine learning increment learning, feature reduction, multiple methods, and federated learning to enhance the accuracy in real-time anomaly detection.

5. Predictive Maintenance Strategies

5.1 Data-Driven Maintenance Models

Predictive maintenance deviates from the conventional methods and utilizes IIoT sensor data together with analytics for maintenance. These models help in determining the likelihood of a failure of equipment so that preventive actions can be made reducing the time the equipment spends off-line while using resources most efficiently (Qi & Tao, 2018). These may include sensing data such as vibration and temperature; operating data such as production rates; maintenance history; environment data; and specifications data. Predictive fields' computations include regression analysis, where variables such as RUL are estimated; classification models, where equipment conditions are sorted; and time series analysis for anticipating future actions. Various authors have underscored the fact that these models' performance depends on the quality of data used, and therefore proper data management is paramount.

5.2 Failure Prediction Techniques

Scheduling failure is crucial in predictive maintenance so that the maintenance process can be well executed and the company lose less time with the machinery. These include survival analysis that deals with the time to failure adding up techniques such as Cox Proportional Hazards models. To predict failures, and depending on the nature of the failure data, Recurrent Neural Networks (RNNs), and Long Short-Term Memory (LSTM) networks are more effective in analysing sequential data. XGBoos and other Gradient Boosting Machines work with high-cardinality features and their interactions. Physics based learning augments physical knowledge with new data in regions where data is sparse improving accuracy. Transfer learning improves predictions in cases of a rather rare failure of equipment by using knowledge about other similar equipment.

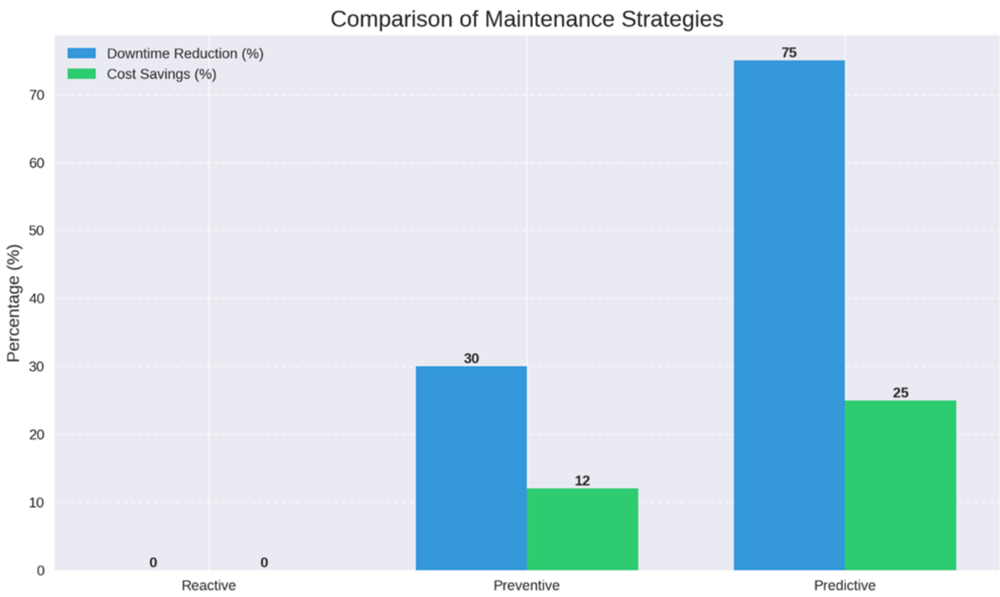
5.3 Optimization of Maintenance Schedules

Achieving the right interval of maintenance also takes into account the time that a system will be off line as well as the cost of maintenance and reliability of the facility (Samie, Bauer, & Henkel, 2019). These are such as risk-based maintenance in which maintenance is done based on the risks that may result from equipment failure, and resources which include manpower, spare and time. There is a possibility of maintenance grouping to amalgamate activities. Thus, predictions are likely to be accompanied by uncertainty, and this is where strong methods of optimization are required. Some methods are mixed-integer programming for developing rich schedules, genetic algorithm for multi-objective optimization, reinforcement learning having adaptive scheduling, and multi-criteria decision-making to meet the rival objectives. These techniques help organizations to transfer from condition based to predictive and prescriptive maintenance which further increases reliability and efficiency.

6. Data Management and Processing

6.1 Big Data Handling in Cloud Environments

IIoT systems also create huge data, and the data should be properly managed in the cloud environment. These are overcome on cloud platforms that have options for distributed storage system like HDFS and Amazon S3, and data lake like AWS Lake Formation. Applications like Apache Spark, that focus on distributed processing help with large-scale data techniques, and serverless computing options that include AWS Lambda as an example help in the administration of computation by automatically providing more resources for the tasks at hand and removing them when they are no longer needed (Shi, Cao, Zhang, Li, & Xu, 2016)



. Big data is defined by some sources using the four V's that include volume, variety, velocity, and vulnerability findings new generation data processing platforms, such as Amazon Redshift, aim at offering large-scale storage and analysis of structured big data. This paper will attempt to identify the various technologies in use in dealing with big data, as well as determine the optimal technologies depending on the volume, velocity, and analytical levels required.

6.2 Data Quality and Preprocessing

In IIoT systems data quality is a critical success factor since the data is the primary basis for analysis and decision-making. Data acquisition depends on data cleaning that is aimed at correcting errors, data smoothening that involves use of filters and scaling to achieve data consensus. Outlier detection is the differentiation between outliers and sensor errors Outlier detection and feature engineering involves the generation or alteration of features of an ML algorithm. Synchronization of time is especially important to provide more precise analysis of results obtained from the different sources situated in various geographical locations. Also, preprocessing has a certain level of automation with further monitoring by employees and scaling based on cloud services.

6.3 Feature Extraction and Selection

Feature extraction and selection make the data more manageable for machine learning in IIoT applications through the process of dimensionality reduction and improvement of the model. For time series data, it engulfs the statistical measure and the frequency domain features which is required to be extracted by using F. F. T. For transient events the time-frequency techniques like STFT are used. Technical specificities of a domain correlate with the industrial knowledge, so they define what information could be useful. Other approaches like PCA and RFE are applied to choose features with the highest separability, while feature generation tools can help to improve this step (Wang, Ma, Zhang, Gao, & Wu, 2018). The methodology selection is based on the application type, data characteristics, and specific machine learning algorithms used; which sometimes involve expert knowledge integration and data-driven methods.

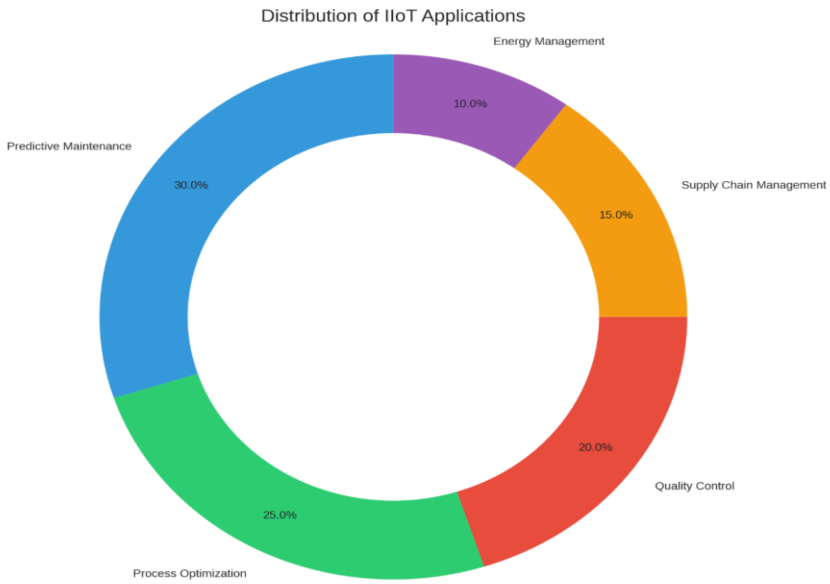
7. Model Development and Training

7.1 Supervised and Unsupervised Learning category

In IIoT applications, supervised and unsupervised learning are crucial for use. Supervised learning is used when there is a definite target to estimate like equipment failure in which data needs to be labelled for training. These are some of the approaches such as Random Forest, Gradient boosting and Neural networks, used in areas such as predictive maintenance and quality assurance. On the other hand, unsupervised learning is valuable when there is no labelled information available and one wants to carry out exploratory analysis of large data set with rare failures. Representative methods include clustering algorithm, dimensionality reduction techniques and anomaly detection techniques. Some IIoT applications use both styles, the use of unsupervised learning for exploration of data and then using supervised learning for a given task. Big advances have also been made in semi-supervised learning, when both labelled and unlabelled data are Le used when there is a scarcity of labelled data.

7.2 Deep Learning for Complex Pattern Recognition

AI is a very useful technology in IIoT for pattern recognition where deep learning is used. Precisely, it self-adapts to derive hierarchical representations, which makes it proper for matters concerning industries. They are the Convolutional Neural Networks (CNN) for working with the data from sensors, Recurrent Neural Networks (RNN) for time series data and Autoencoder for the anomaly's identification (Xu & Duan, 2019). Deep reinforcement learning, which is still in its infancy, can be have the potential to enhance the industrial operations. Deep learning preserves the knowledge relations, but it demands large amounts of data, computational power, and is sometimes hard to explain. IoT is currently faced with these challenges, and it explores the works of transfer learning, edge computing, and explainable AI as potential solutions with deep learning seen to be gradually ascending in the IIoT strategy.



7.3 Transfer Learning in Industrial Applications

Transfer learning is important in IIoT as it aids in the evolution of issues such as limited data. It enables the model to take advantage from the information of related jobs so as to shorten training time together with enhance generality. This general approach is particularly valuable when data are limited or when deploying a model quickly is critical. These are featuring extraction, transferring learning in which the model is fine-tuned and finally domain adaptation in order to address different data distributions. Transfer learning is great as it helps to speed up the model creation, but one must take into account the similarities between the domains and privacy concerns. Nevertheless, transfer learning is an effective method for improving AI/ML models for IIoT in resource-scarce circumstances.

8. Implementation Challenges

8.1 Integration with Legacy Systems

Implementing cloud-based AI and ML solutions for IIoT with the existing industrial systems is difficult because of archaic devices and connectivity standards (Lee, Davari, Singh, & Pandhare, 2018). Legacy systems have their own format of data as well as the way of communication that is different from the IoT standards. Live time conditions and insecurity in legacy systems bring extra difficulties. Solution to these challenges include; employing of protocol gateway, edge computing, middleware solutions, retrofit sensor, and virtualization. These sorts of approaches assist in explaining how to create the transition from legacy structure to the new IIoT structure and also, they take time and need considerable capital outlay.

8.2 Latency and Connectivity Issues

Something very important in IIoT systems that cannot be ignored is the latency and connectivity, especially those systems that would depend on cloud AI and ML. It is realized that network and processing latency as well as data volume can slow down the real-time performance. Other problems include networking complications such as Area 'black outs' and restricted or fluctuating bandwidth. Current approaches towards solving the problem include edge computing to minimize latency, 5G in an attempt to increase efficiency, adaptive protocols on data compression, and a combination of public and private cloud. These strategies' application depends on real-time performance, specific data characteristics, and costs.

8.3 Security and Privacy Concerns

As IIoT systems become interconnected, security and privacy threats increase. Operational and information technology integration makes assets more vulnerable to a cyber-attack, whereas the nature of the information gathered in the industrial environment raises privacy issues. Some main concerns are protection of the device, network, data, and granting of access. To minimise risks, organisations apply different levels of security measures including, physical device security, network compartmentalisation, Message encryption, strong authentication, use of SIEM systems, periodic assessments/audits, and privacy friendly technologies (Liang, Huang, Long, Zhang, Li, & Zhang, 2020). Security hence has to be an ongoing process and needs everyone to adopt a security first attitude.

9. Performance Evaluation

9.1 Metrics for Anomaly Detection Accuracy

The assessment of anomaly detection in IIoT systems is necessary given the fact that data distribution in such systems is usually unbalanced and the costs associated with false positives and false negatives differ. These are; accuracy, precision, recall, and the F-measure or F1 score which combines both recall and precision. The AUC-ROC provide the extent of the model's capacity to differentiate normal from anomalous instances across threshold. Precision-Recall curves are best used whenever you are dealing with an unbalanced data set; they display the recursion of the precision and recall. Another type of balanced measures is the Matthews Correlation Coefficient (MCC). Other specific indices are time to detection, the false alarm rate, and detection stability. These metrics should therefore be specific to the need of the industrial application, and the acceptable false positive and false negatives depending on the priorities of the machine or operation.

9.2 Assessing Predictive Maintenance Effectiveness

In the IIoT systems, several indicators are applied to compare the performance of predictive maintenance implementations. Some of these are; lower degrees of planned and scheduled stoppages, improvements in the MTBF and overall maintenance costs (Molina-Solana, Ros, Ruiz, Gómez-Romero, & Martin-Bautista, 2017). The other KPIs include asset life elongation, prognostication precision, maintenance timetable optimization, and accident elimination. Energy efficiency improvements and OEE are also used as metrics in addition to that, safety performance is also measured. ROI compares the financial return to costs in order to establish the gain. Time series data is compulsory in reporting result so that a benchmark can be set against which the improvement is determined. Monitoring is important because predictive maintenance results optimize with time as more data is collected. It is important to have a systematized process to continuously improve the currently used strategies and to achieve the best IIoT outcomes.

9.3 Cost-Benefit Analysis of Cloud-Based Solutions

When it comes to the practical application of IIoT, the integration of cloud-based AI and ML solutions should be cost-effective. Infrastructure cost incurred in form of cloud service, storage, and data transferring costs, development and implementation costs are the other quantitative factors While on operational cost, such factors may include but not limited to cost of operations that result from Business Process Redesign (BPR) such as cost of equipment downtime and productivity improvements. Financial perspectives are given by ROI and Total Cost of Ownership – TCO. See also: Adzic (2011), where qualitative factors are classified into decision-making, better and safer decision, competitive advantage, legal requirements, workforce satisfaction, and effects on the environment. It is important to evaluate the period through which the advantages will be achieved because the costs will be incurred in the short run while the advantages will be reaped in the long run with the period maturing as time progresses (Qi & Tao, 2018). It is also possible to consider certain negative effects, for example, threats to data security and constant upgrades. An essential part of the analysis might be the sensitivity analysis to assess how the changes in essential assumptions can affect the results. The analysis must be tentative to the organization's strategic objectives and appetite for risk to offer the right expectation of the outcomes of cloud-based IIoT to the organization.

10. Industry-Specific Applications

10.1 Manufacturing Sector

Manufacturing emerges as the leader in the IIoT since organizations in the sector employs the cloud-based AI and ML to boost productivity, quality, and efficiency. Key applications include:

- Predictive Maintenance: Sensors measurement on equipment's and analysis of data in order to predict the failures and plan for the maintenance overhauls, thus making the overhauls and related costs less frequent. Deloitte also shows that in using the technique of predictive maintenance, breakdowns can be reduced by up to 70% while at the same time maintenance costs can be cut by up to a quarter.
- Quality Control and Defect Detection: AI and computer vision are helpful in identifying defects at real-time as compared to human inspectors. AI also found great application in the automobile manufacturing and BMW was able to reduce defect rates by up to 30%.
- Process Optimization: AI models involve complex analyses of factors such as yield, energy consumption, and quality to gain the best outcomes. GE is a good example where its suite known as the Brilliant Manufacturing enhanced manufacturing productivity by a percentage of 20%.

- Supply Chain Management: IIoT technologies increase the transparency and manageability of the supply chain. Demand forecasting and inventory management made possible by AI reduced carrying costs and increased market respondence.
- Energy Management: Smart energy systems put to use ML for anticipating as well as minimizing the usage of energy to save expense as well as decrease environmental degradation.
- Digital Twins: It assists in the management and simulation of the actual manufacturing systems where virtual copies of such systems are made (Samie, Bauer, & Henkel, 2019). It can be said that digital twin technologies by Siemens have decreased the potential machine downtime to up to 70%.

They are used in the framework of the "smart factories" or the so-called Industry 4. 0, transforming traditional manufacturing. Some issues are associated with integration with existing systems, as well as the protection of information and cyber threats.

10.2 Energy and Utilities

IIoT technologies improve the business processes, grid stability, as well as the integration of renewable energy sources in the energy and utilities segment. Key applications are:

- Smart Grid Management: They are used for monitoring and control of power supply through IIoT sensors and balancing of loads along with response to outages through AI. Thus, Iberdrola smart grid system has previously improved outage reduction by 20% and restoration time by 40%.
- Predictive Maintenance for Power Generation: AI locates bad areas in turbines, transformers and solar panels and gives the guidelines to improve. About GE, Decision Digital Wind Farm tech improved wind turbines efficiency by up to 20%.
- Demand Response and Load Forecasting: It applies itself to the estimating of energy demand to enhance the claims of supply as well as demand management. With the help of DeepMind AI, Google reduce cooling energy expenses up to the 40%.
- Asset Health Monitoring: Some of the applications are in power lines, pipelines, and substations where sensors and AI are used in the monitoring of power systems for maintenance.
- Renewable Energy Integration: AI is useful in the area of forecasting of energy consumption and helps optimize storage to absorb variable renewable power.
- Energy Theft Detection: AI in smart meters: AI can analyse smart meter data to detect theft consequently, UK Power Networks increases the detection rate of theft by 75%.
- Water Management: Concerning the application of IIoT to water management it enables leak detection, monitoring of water quality and distribution of the water. In another example, South Bend Indiana has been able to cut down sewer overflows by 70% that was equivalent to $ 500 million that would have been used in infrastructure (Shi, Cao, Zhang, Li, & Xu, 2016).

These technologies make smart, climate-friendly energy solutions possible but demand restrictive existing IT/OT systems of security for protection against privacy invasion, invest in infrastructure, and cybersecurity.

10.3 Transportation and Logistics

IIoT is now linking and revolutionising transportation and logistics making operations, safety and the customers' experience better. Key applications include:

- Fleet Management and Predictive Maintenance: Tracking and monitoring the vehicles in real time enhances means to predict when the vehicle is due for maintenance or repair as well as enhance the operations of the vehicle. UPS's system being beneficial to the company; such as, the savings that UPS's system achieved was in the tune of millions in terms of repairs, in addition to reducing the rate of breakdowns by 20%.
- Route Optimization: In this way, AI improves delivery routes based on the traffic and weather conditions, and for DHL it provides up to 15% of fuel savings and improved delivery time.
- Supply Chain Visibility: IIoT also helps in achieving end to end visibility in an organization's supply chain which in turn provides better inventory tracking and risk mitigation. Maersk's blockchain platform also enriches the monitoring of container resources.
- Autonomous Vehicles: Smart features for the cars are already supported by AI; however, self-driving cars are not yet in existence, but will be soon.
- Warehouse Automation: AI and robotics IoT in warehouses self-spending: Robotics and Internet of Things (IoT) reduce costs (Wang, Ma, Zhang, Gao, & Wu, 2018). This precisely reduced Amazon's operating costs by approximately $20 % through the use of robotics.
- Predictive Demand Forecasting: Through its capabilities, ML models estimate the demand pattern making the stock controlling and minimizing wastage easier.

- Smart Ports and Terminals: The IIoT improves the means of handling the containers, the time that each will spend on waiting and the throughput they will achieve. By augmenting throughput of containers, the Port of Hamburg increased it by no less than 8-12%.
- Passenger Experience Enhancement: AI and IIoT enhance public transport by arriving at the best schedules of departure and arrival and also by disclosing the changes in schedules immediately. To counter this, Transport for London utilizes AI to help in modifying the services depending on ATT needs.

These result in improved productivity, costs and service in transportation and logistics; however, weaknesses include data aggregation, privacy and security, and social issues to do with its workforce. Possible future developments of logistics are self-driving cars, drones, and individualized supply chain management.

11. Emerging Trends and Future Directions

11.1 Edge-Cloud Hybrid Models

Current IIoT architectures are incorporating edge-cloud integration to integrate the real time analytics of edge computing with the scalable and diverse analytics of cloud computing. This solves the issues of real time data processing, low latency and bandwidth control along with the adoption of cloud-based AI and ML.

The primary features of these models are distributed intelligence in which decision-making small AI models are deployed on edge devices, while more detailed large models are on the cloud. Intelligent load management assignments the tasks on the edge and cloud depending on the conditions in the network and security in the data. Edge-native AI aims at developing models that should naturally be deployed on the edge, taking into account such factors as the processing power of the device, the amount of power that is available to the device, etc. Federated learning is carried out by training models in multiple edge devices where the data does not have to be centralized hence addressing the issues of privacy while at the same time helping in creation of localized models (Xu & Duan, 2019). It is proposed that the incorporation of 5G networks will improve these models by offering greater bandwidth and less latency.

It is believed that there will be an increase in the numbers of edge-cloud hybrid models due to enhanced edge devices, better AI algorithms, as well as better orchestration techniques. As a consequence, the development of such systems will trend toward more reliable, adaptive and efficient IIoT systems across the industries.

11.2 Explainable AI for Industrial Applications

With an increasing reliance on AI and ML models in the industrial setting, there is a growing need for Explainable AI or XAI to ensure that the basis for the systems' decisions is comprehensible. This is important in order to establish some level of working relationship where trust is developed and to also adhere to regulatory standards on safety.

Some progress in the field is LIME and SHAP related explanation techniques that are model agnostic — these offer an explanation of whether or not a feature is important or not, and the boundaries of decisions made by the model. Perturbation and feature importance approaches are being creating to present the relationships and other patterns discovered by AI system to the users in better way. The next type of AI models under study is causal AI models that are expected to give a causal analysis rather than mere correlation analysis. Domain-specific interpretability deals with the development of explanation procedures within specific industrial domains using knowledge and terminology native to these regions. Also, there is a rapid increase in demand in implementing XAI to correspond to new increased requirements for accountability and regulative compliance.

Thus, the development of xAI is instrumental in expanding the use of AI approaches in industrial applications as it disrupts the problems of complexity of these processes and lack of transparency.

11.3 Integration with Digital Twin Technology

Digital twin technology where an exact copy of a physical asset or a process is built as a virtual replica is often combined with IIoT, and AI/ML. It also permits monitoring, control or modelling and analysing of industrial systems and processes in a virtual environment in real time.

Some of these forms include real-time data integration by IIoT sensors in which the digital twin models are constantly updated based on the real configuration of the systems. AI and ML models allow for the anticipation of various aspects of the system's performance, their failures, and possible improvements. Digital twining can involve various scenarios and the operating conditions, which helps in improving processes, with hardly any impact on real facilities. It also opens the door to such initiatives as self-optimizing systems incorporated into digital twins with AI hence reducing the need for direct human control. In Lifecycle Management industry, digital twin offers information about assets at every stage of their life cycle, from the development phase to service.

More advancements in the optimization of industries and predictive maintenance using IIoT with digital twin, and augmented with AI/ML, will cause highly responsive and adaptive systems. In this way, the new technologies expand the existing possibilities of organizational functions that enable industrial growth and adaptation while improving productivity and creativity.

12. Conclusion

12.1 Summary of Key Findings

Based on this discussion of analysing IIoT systems using cloud-based AI and ML for real-time anomaly detection and PM, the following are notable discoveries: One of the most profound concepts in the current technological era is the integration of AI and ML in industrial processes to produce better results through faster and wiser decision-making solutions (Khan, Ahmed, Hakak, Yaqoob, & Ahmed, 2019). Cloud solutions are the only ones capable of providing sufficient and adaptive computing power required to process the big data generated by IIoT systems and accommodate the models. The approaches to the use of real-time anomaly detection as well as to the prognosis of equipment failure have shown their efficiency in avoiding unnecessary downtimes, decreasing the maintenance costs, and improving operational performance. However, issues like the compatibility issues with existing systems, slow processing time, security issues are always there. The IIoT system's performance assessment he defines by multiple technical, operational and financial parameters. Various industries with references to the present usage of energy and manufacturing establishments, transportation, and others explain its innovative usage possibilities. Trends such as edge-cloud, Explainable AI, and Integration of the Digital Twins are diversifying the Second Wave of IIoT.

12.2 Implications for Industrial IoT Systems

The findings of the research have significant consequences when it comes to the future of IIoT systems. This is because as Mn and AI benefits of cloud rise as technology and solutions mature and the pace of adoption journeys across industry verticals will further increase. In industrial processes, there will be a move from detect-and-correct to analyse-and-avoid on lapses that are bound to occur in operations. Thus, further developments of IIoT systems will improve the integration with OT and IT systems for better process control (Lee, Davari, Singh, & Pandhare, 2018). Rapid advances in edge compute will create more deployments focused on pushing more intelligence to the network edge to speed up decision making and lessen cloud dependence. As connectivity advances, security is also an important element that will be enforced, thus boosting the development of solid security conditions.

13. References

[1] Amirante, R., & Lamonaca, F. (2023). Industrial Internet of Things (IIoT) for smart manufacturing: A comprehensive review. Sensors, 23(3), 1523. https://doi.org/10.3390/s23031523

[2] Chai, Y., Miao, C., Sun, B., Zheng, L., & Li, Q. (2022). Federated learning for the Industrial Internet of Things: A comprehensive survey. IEEE Internet of Things Journal, 9(21), 21015-21036. https://doi.org/10.1109/JIOT.2022.3182749

[3] Elshawi, R., Sakr, S., Talia, D., & Trunfio, P. (2018). Big data systems meet machine learning challenges: Towards big data science as a service. Big Data Research, 14, 1-11. https://doi.org/10.1016/j.bdr.2018.04.004

[4] Fahad, L. G., Tahir, S. F., & Rajarajan, M. (2021). Feature selection and data balancing for edge devices in the Industrial Internet of Things. IEEE Transactions on Industrial Informatics, 17(3), 2377-2386. https://doi.org/10.1109/TII.2020.3010798

[5] Ge, Z., Song, Z., Ding, S. X., & Huang, B. (2017). Data mining and analytics in the process industry: The role of machine learning. IEEE Access, 5, 20590-20616. https://doi.org/10.1109/ACCESS.2017.2756872

[6] Guillén, J. H., Martínez, J. F., Rodríguez-Molina, J., & Rubio, G. (2021). A review of edge computing reference architectures and a new global edge proposal. Future Generation Computer Systems, 117, 467-487. https://doi.org/10.1016/j.future.2020.12.016

[7] Khan, W. Z., Ahmed, E., Hakak, S., Yaqoob, I., & Ahmed, A. (2019). Edge computing: A survey. Future Generation Computer Systems, 97, 219-235. https://doi.org/10.1016/j.future.2019.02.050

[8] Lee, J., Davari, H., Singh, J., & Pandhare, V. (2018). Industrial Artificial Intelligence for industry 4.0-based manufacturing systems. Manufacturing Letters, 18, 20-23. https://doi.org/10.1016/j.mfglet.2018.09.002

[9] Liang, W., Huang, W., Long, J., Zhang, K., Li, K. C., & Zhang, D. (2020). Deep reinforcement learning for resource protection and real-time detection in IoT environment. IEEE Internet of Things Journal, 7(7), 6392-6401. https://doi.org/10.1109/JIOT.2020.2970877

[10] Molina-Solana, M., Ros, M., Ruiz, M. D., Gómez-Romero, J., & Martin-Bautista, M. J. (2017). Data science for building energy management: A review. Renewable and Sustainable Energy Reviews, 70, 598-609. https://doi.org/10.1016/j.rser.2016.11.132

[11] Qi, Q., & Tao, F. (2018). Digital twin and big data towards smart manufacturing and industry 4.0: 360

degree comparison. IEEE Access, 6, 3585-3593. https://doi.org/10.1109/ACCESS.2018.2793265
[12] Samie, F., Bauer, L., & Henkel, J. (2019). From cloud down to things: An overview of machine learning in internet of things. IEEE Internet of Things Journal, 6(3), 4921-4934. https://doi.org/10.1109/JIOT.2019.2893866
[13] Shi, W., Cao, J., Zhang, Q., Li, Y., & Xu, L. (2016). Edge computing: Vision and challenges. IEEE Internet of Things Journal, 3(5), 637-646. https://doi.org/10.1109/JIOT.2016.2579198
[14] Wang, J., Ma, Y., Zhang, L., Gao, R. X., & Wu, D. (2018). Deep learning for smart manufacturing: Methods and applications. Journal of Manufacturing Systems, 48, 144-156. https://doi.org/10.1016/j.jmsy.2018.01.003
[15] Xu, L. D., & Duan, L. (2019). Big data for cyber physical systems in industry 4.0: A survey. Enterprise Information Systems, 13(2), 148-169. https://doi.org/10.1080/17517575.2018.1442934