LIGHTWEIGHT LEARNING FOR ADAPTIVE PREDICTIVE MAINTENANCE IN INDUSTRIAL IOT

*P. Shalini** [1], *Dr S Veni* [2]

ABSTRACT

Current predictive maintenance methods in IIoT struggle with complex data, limited edge device resources, and real-time adaptation. This paper proposes a novel, lightweight machine learning framework designed specifically for IIoT environments. Our framework utilizes efficient algorithms for resource-constrained devices, enabling seamless integration and potentially reducing processing time and energy consumption. Deploying the framework on edge devices allows for real-time monitoring and decision-making. Rigorous evaluation, including simulations and real-world experiments, aims to quantify the framework's benefits, such as potentially reducing downtime and maintenance costs. Additionally, the research explores techniques like knowledge distillation for further model size reduction and federated learning for collaboration between devices, potentially enhancing adaptability. This framework has the potential to revolutionize IIoT predictive maintenance by offering a cost-effective and adaptable solution for optimized operational efficiency.

Keywords: IIOT, Industrial IOT, Lightweight Learning, Federated Learning andPredictive Maintenance.

I. INTRODUCTION

The Industrial Internet of Things (IIoT) has emerged as a transformative force in industrial sectors, revolutionizing how machines, equipment, and processes are monitored, managed, and optimized. It builds upon the foundation of Distributed Control Systems (DCS) by integrating sensors, connectivity, and powerful cloud-based data analytics. This enables real-time monitoring, predictive maintenance, and advanced automation, leading to increased efficiency, reduced costs, and improved operational outcomes.

Predictive maintenance is a crucial aspect of the IIoT for several reasons. It offers significant cost reductions by allowing companies to identify potential equipment failures before they happen. This proactive approach minimizes unplanned downtime, delays the need for expensive equipment replacements and emergency repairs, resulting in lower maintenance costs. It provides valuable insights into the actual condition of equipment, allowing maintenance activities to be scheduled only when necessary. This eliminates the need for fixed-time or usage-based maintenance, optimizing the use of maintenance resources and minimizing the disruptions of production processes. It contributes to a safer work environment. Regularly monitoring equipment health and proactively addressing potential problems reduces the risk of accidents and equipment failures, creating a safer work environment for employees. Improved efficiency is another benefit of IIoT-based predictive maintenance. By minimizing downtime and maximizing equipment uptime, these systems enable organizations to meet production targets more consistently and efficiently. This allows for better planning and allocation of resources, leading to overall operational improvements.

IIoT systems collect vast amounts of data from industrial equipment and processes. Predictive maintenance leverages this data to identify patterns, trends, and anomalies, providing valuable insights into equipment performance and maintenance needs. This data-driven approach plays a vital role in optimizing asset performance, reducing costs, and improving the overall reliability of industrial operations. Devices with limited processing power benefit from lightweight machine learning models that simplify data processing and enable faster implementation. However, for massive datasets from numerous interconnected devices, more powerful machine learning and deep learning techniques are necessary to extract the most valuable

Department of Computer Science[1,2],
Karpagam Academy of Higher Education, Coimbatore, India[1,2]

* Corresponding Author

insights. Efficient algorithms are chosen, favouring simpler options like decision trees, logistic regression, or basic neural networks, due to their ease of use and minimal computational requirements. This allows for quicker training and deployment on resource-constrained devices with limited processing power. Real-time predictions are a core aspect of lightweight learning. These models are designed to provide immediate insights into potential equipment failures, enabling prompt intervention and preventative maintenance actions to minimize downtime and costly disruptions.

Furthermore, lightweight learning seamlessly integrates with existing IIoT platforms, enabling centralized monitoring and management of predictive maintenance across various industrial assets.Their ease of customization and deployment across various industries and applications showcases their scalability and flexibility, making them a powerful tool. The ever-increasing number of interconnected devices in IIoT systems generates massive amounts of data, posing a challenge for traditional analysis methods. Machine learning and deep learning emerge as powerful tools, transforming how IIoT leverages this complex data.

Machine learning techniques are revolutionizing the Industrial Internet of Things by excelling at identifying patterns and relationships within data. It can be used for various tasks in IIoT, such as anomaly detection, predictive maintenance, and process optimization. For instance, machine learning algorithms can analyze sensor data from industrial equipment to detect early signs of potential failures. This enables preventive maintenance actions to avoid costly downtime and equipment damage.

Deep learning is particularly adept at handling complex and high-dimensional data. Deep learning models can be trained on vast amounts of sensor data to perform tasks like image recognition, speech recognition, and even anomaly detection. In an IIoT setting, these models could analyze video footage from security cameras to detect safety hazards or monitor product quality on a production line. By leveraging the power of machine learning and deep learning, IIoT systems can gain valuable insights from data, enabling them to automate and optimize operations, and make data-driven decisions for improved efficiency and productivity.

II. LITERATURE SURVEY

Hasan, M.K., et. al., (2023) proposed an explainable ensemble deep learning approach for intrusion detection in IIoT systems integrating Shapley Additive Explanations (SHAP) and Local Comprehensible-Independent Clarifications (LIME) methods to provide clarity on Intrusion Detection Systems decisions. Karacayılmaz, G., et. al., (2024) proposed an expert system utilizing AI techniques to detect and prevent IIoT device attacks like denial-of-service, data manipulation, device hijacking, and physical tampering, promoting stronger security for critical infrastructure. Alalayah, K.M., et. al., (2023) proposed the Hunger Games Search Optimization with Deep Learning-Driven Intrusion Detection (HGSODLID) model for IIoT security leveraging linear normalization and HGSO for feature selection along with Sparrow Search Optimization (SSO) and a Graph Convolutional Network (GCN) for intrusion identification. Bugshan, N., et. al., (2022) proposed Federated Learning based Deep Learning service framework to address privacy concerns in IIoT. It aggregated locally trained deep learning models without sharing raw data leveraging a service-oriented architecture and differential privacy for secure execution. Mohy-eddine, M., et. al., (2023) presented an ML-based intrusion detection approach for IIoT security. It employed feature selection techniques like Pearson's Correlation Coefficient (PCC) and Isolation Forest (IF) to optimize data for a Random Forest classifier, improving attack detection Accuracy.

Ikram, S.T., et. al., (2022) proposed a two-phase IIoT traffic prediction model for anomaly detection using Multi-objective Non-dominated Sorting with Whale Optimization Approach (MNSWOA) and Ideal Point Method (IPM) for feature selection, identifying key attributes for an random forest classifier to predict normal and anomalous traffic. Xu, L., et. al., (2024) proposed a mobile communication performance analysis and prediction algorithm based on FL-GLP-Net addressing security concerns in 5G-enabled IIoT environments leveraging N-Nakagami channels to analyse Non-zero Secrecy Capacity Probability (NSCP) and utilize

XGBoost for optimal feature selection.The model integrates Graph Attention Network (GAT), LSTM and Pyramid Visual Converter (PVT) modules to handle diverse signal features, achieving high accuracy in NSCP prediction. Roopa, M.S., et. al., (2021) introduced Social Internet of Things (SIoT) for predictive maintenance in manufacturing. This ontological model predicted Remaining Useful Life (RUL) of machine elements, minimizing downtime and costs by anticipating failures in IIoT environments. Bulla, C., et. al., (2022) proposed a multi-agent system with fog computing for anomaly detection in IIoT improving Quality of Service (QoS). Their approach used multi-step prediction with a Gated Recurrent Unit (GRU) model optimized by an Artificial Bee Colony (ABC) algorithm for high accuracy anomaly detection. Li., H., et. al., (2024) proposed a lightweight privacy-preserving predictive maintenance to address privacy concerns in 6G-enabled IIoT with its vast data exchange. This method used homomorphic encryption for secure machine learning on encrypted data and leveraged inary Neural Networks (BNNs) to train privacy-preserving maintenance models.

Soliman, S., et. al., (2023) introduced an intelligent detection system to identify cyberattacks in IIoT networks tackling issues like lack of attack comprehensiveness, high feature dimensionality, outdated datasets and imbalanced datasets utilizing Singular Value Decomposition (SVD) for feature reduction and SMOTE to mitigate bias in classification. Friha, O., et. al., (2023) proposed a secure, decentralized and Differentially Private (DP) federated learning based IDS (2DF-IDS) for smart factories. This approach leveraged differential privacy and a decentralized architecture to balance security and privacy. Misbha, D.S., et. al., (2022) proposed attention-based Convolutional LSTM (Conv-LSTM) and Bidirectional Long Short-Term Memory (Bi-LSTM) network, a new attack detection system for IIoT. It extracted features from both temporal and spatial aspects of the data, fused them and classified data as normal or abnormal. The algorithm achieved a high accuracy of over 95%.Schmieg, T., et. al., (2024) surveyed deep learning techniques for time series forecasting, focusing on their role in representation learning. Among the 17 architectures reviewed, the most commonly used techniques were one-dimensional CNN, LSTM and attention-based methods. Input embedding and masking also play significant roles in some architectures.

Yang, H., et. al., (2023) proposed a deep learning based Remaining Useful Life (RUL) prediction model utilizing CNN to find important patterns in equipment data, LSTM to understand how these patterns change over time and self-attention mechanisms to focus on the most important parts of the data for prediction to improve equipment maintenance practices in IIoT environments. Chander, N., et. al., (2024) proposed Enhanced Pelican Optimization Model with an Ensemble Voting-Based Anomaly Detection (EPOA-EVAD) a new anomaly detection system for IIoT. It tackled class imbalance and selected optimal features. This method combined Gated Recurrent Unit (GRU), Bi-Directional LSTM and stacked auto encoder for anomaly detection and achieved high accuracy in dynamic IIoT environments. Isah, A., et. al., (2023) proposed digital twin temporal dependency that used LSTM to analyze temporal dependencies in time series data from IIoT systems. This approach captured long-term relationships between variables, leading to improved prediction accuracy. Smmarwar, S.K., et. al., (2023) proposed Double-Density Discrete Wavelet Transform (D3WT) for feature extraction and a combination of CNN and LSTM models for malware identification and classification with high accuracy of over 96%.

Current IIoT predictive maintenance methods face challenges in handling the complexities of data generated by industrial machines. These datasets can be large and intricate, posing difficulties for existing approaches. Additionally, these methods may not be optimized for the limited resources of edge devices typically deployed in IIoT environments. This can lead to processing delays or render the methods infeasible altogether. Furthermore, the ability to adapt to real-time changes is often limited in current solutions, hindering their effectiveness in dynamic industrial settings. This paper proposes a novel, lightweight machine learning framework specifically designed to address the limitations of existing methods in IIoT predictive maintenance. The potential contributions include:

· Attention-based LSTM network captures temporal

dependencies and prioritizes relevant features for accurate machine failure prediction, enhancing interpretability and performance.

Optimizes the framework for deployment on resource-constrained edge devices using lightweight algorithms to minimize processing time and energy consumption and enabling real-time monitoring and decision-making.

The paper is organized into five sections. The Introduction discusses how IIoT improves industrial monitoring and optimization through sensors, connectivity and cloud analytics. It highlights the role of predictive maintenance in reducing costs and downtime by detecting equipment failures early and optimizing maintenance schedules. The literature survey examines existing literature on predictive maintenance in IIoT environments, identifying the strengths and limitations of current approaches to provide context for the proposed framework. The methodology details the proposed framework covering the data preprocessing pipeline, the architecture of the attention-based LSTM model and techniques for optimizing performance on edge devices. The results and analysis section describes the experimental setup, the datasets, evaluation metrics, and implementation details demonstrating the effectiveness of the proposed approach. Finally, the conclusion summarizes the key contributions of the study and suggests directions for future research to further enhance predictive maintenance in IIoT environments.

III. METHODOLOGY

The proposed predictive maintenance model for IIoT environments is developed through a structured approach, beginning with data preprocessing. Min-Max normalization is applied to the sensor data to scale all features within the range of [0, 1], enhancing model performance by preventing any single feature from dominating. Label encoding is used to convert categorical labels into numerical values, which facilitates machine learning processing. To address class imbalance SMOTE is used generating synthetic examples of the minority class and improving the model's ability to predict these outcomes. After preprocessing, the data is fed into an attention-based LSTM network which is designed to capture temporal dependencies within the time-series sensor data. An attention layer is integrated to focus on the most relevant time steps, thereby enhancing prediction accuracy by weighting significant input sequences more heavily. Figure 1 shows overall architecture for binary classification designed to predict failure or non-failure outcomes. It employs techniques like normalization, label encoding, SMOTE, self-attention and LSTM to process input data and make accurate predictions.

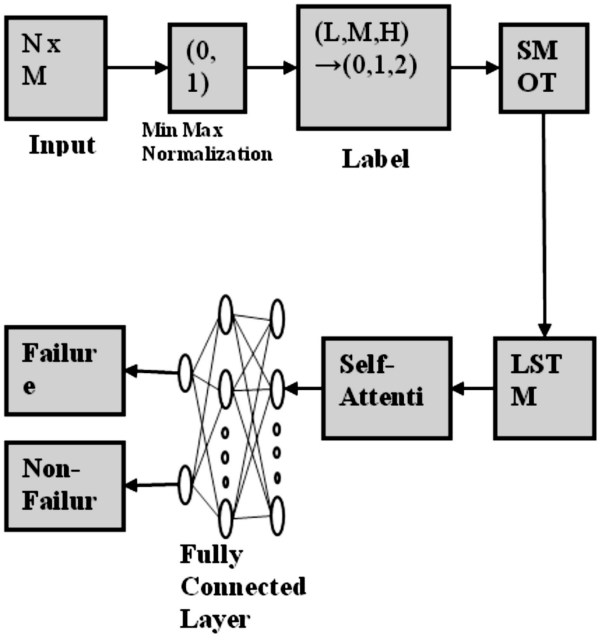


Figure 1: Overall architecture

3.1. Pre-processing

Data pre-processing is a critical step that ensures the sensor data is clean, normalized and ready for model training. The raw data from the UCI AI4I dataset comprises multiple features including air temperature, process temperature, rotational speed, torque and tool wear which require transformation before sending them into the predictive model. The first step of this methodology involves pre-processing the sensor data. The raw sensor data $V \in \mathbb{R}^{n \times m}$, where n is the number of data points which is 10,000 and m is the number of features which is 13 is processed.

3.1.1 Min-Max Normalization:

Min-max normalization is applied to ensure uniformity across different sensor readings and to standardize the features by scaling each feature to a range of [0,1]. The formula used for normalization is:

$$v' = \frac{v - v_{min}}{v_{max} - v_{min}}$$

Where v' is the normalized value, v_{min} and v_{max} are the minimum and maximum values of feature respectively.

3.1.2 Label Encoding

The dataset includes a categorical feature representing the product type containing categorical values ('L', 'M', 'H'). L (Low) is assigned a numerical value of 0, M (Medium) is assigned a numerical value of 1 and H (High) is assigned a numerical value of 2.

Label encoding is performed to map each category to a numerical value. This step converts the categorical data into a machine-readable format.

3.1.3 SMOTE

SMOTE addresses class imbalance inmachine failure data by generating synthetic samples for the minority class to balance the dataset and improve model generalization. Let $V_{min} \in \mathbb{R}^{n_{min} \times m}$ represent the minority class data points, where n_{min} is the number of minority class samples. SMOTE generates synthetic data points V_{syn} using the following interpolation formula:

$$v_{new} = v_{minority} + \lambda(v_{neighbor} - v_{minority})$$

Where v_{new} is a new synthetic point, $\lambda \in [0,1]$ and $v_{neighbor}$ is a randomly chosen sample from the nearest neighbor from the minority class $v_{minority}$.This creates synthetic data to balance the dataset.

3.2. Attention-based LSTM Model

Once the data is pre-processed, it is passed to an attention-based LSTM model for machine failure prediction.

3.2.1 LSTM Layer

The first layer of the model is the LSTM which captures temporal dependencies between the 10,000 rows of sequential data points. Each data point has 13 features, which makes the input to the LSTM layer $V' \in \mathbb{R}^{10,000 \times 13}$. This layer produces hidden states for each timestep that represent the learned temporal patterns. The LSTM processes this input and output hidden states

$$H \in \mathbb{R}^{10,000 \times 64}$$

where 64 represents the dimensionality of the hidden state. Each hidden state is a compressed representation of past sensor readings. The LSTM captures long-term dependencies, which are essential for analyzing sequential data over time.

3.2.2 Self-Attention Mechanism

The output from the LSTM is then passed into the self-attention mechanism. The attention mechanism is important because it enables the model to learn which time steps are more important for predicting machine failure. The attention mechanism computes attention weights for each time step by comparing the query, key and value vectors derived from the LSTM output.

The self-attention mechanism computes attention scores $\alpha_{t,t'}$ between every pair of time steps t and t' and uses these scores to generate weighted sums of the value vectors for each time step. This allows the model to focus on the most relevant time steps.

$$Z \in \mathbb{R}^{10,000 \times 64}$$

The output Z contains weighted feature representations that prioritize time steps based on their importance to the prediction task.

3.2.3 Fully Connected Network (FCN)

The self-attention output is passed into a fully connected layer for final classification. This layer transforms the attention-weighted outputs into a final prediction. The FCN includes a dense layer that aggregates the sequence representations and reduces the output to a single probability indicating machine failure or no failure. The fully connected network applies a final layer with a sigmoid activation function to convert the output into a probability between 0 and 1, indicating whether the machine is likely to fail or not.

$$y_{pred} = \sigma(W_{out}Z + b_{out}) \in \mathbb{R}^1$$

Where $y_{pred} \in [0,1]$ is the predicted probability of machine failure and $\sigma(\cdot)$ is the sigmoid activation function. The final output is a scalar value, with $y_{pred} > 0.5$ indicating failure $(y = 1)$ and $y_{pred} \leq 0.5$ indicating no failure $(y = 0)$.

3.2.4 Model Training and Optimization

The model is trained using the Adam optimizer with a binary cross-entropy loss function, appropriate for binary classification tasks such as predicting machine failure.

$$L = -\frac{1}{N}\sum^{N} (y_i \log(y_{pred,i}) \quad y_{pred} \quad i-th \quad + (1-y_i)\log(1-y_{pred,i}))$$

Where *N* is the number of samples, $y_{pred} \in \{0,1\}$ is the actual label for the *i - th* sample and $y_{pred,i} \in \{0,1\}$ is the predicted probability for that sample.

This methodology leverages the temporal modelling capability of LSTM layers, the interpretability of self-attention mechanisms, and the power of fully connected layers for classification to predict machine failure or no failure. Data flows through each layer, transforming from raw sensor readings into highly informative feature representations, and culminating in an accurate binary classification output.

Algorithm: Adaptive_Attention-Based_LSTM()

Input:

- Sensor Data V from the UCI AI4I dataset (10,000 samples, 13 features)

Output:

- Prediction of machine failure (1 for failure, 0 for no failure)

1. Data Preprocessing:

- Min-Max Normalization:

For each feature *v* in *V*, apply normalization :

$$v' = \frac{v - v_{min}}{v_{max} - v_{min}}$$

- Label Encoding:

Identify Categories:

Extract unique categories:

$$('L','M','H')$$

- Assign Labels:

Map each category to an integer:

$$'L' \rightarrow 0$$
$$'M' \rightarrow 1$$
$$'H' \rightarrow 2$$

- Replace Values:

For each entry in Product Type replace with corresponding label:

$$'L' \rightarrow 0, 'M' \rightarrow 1, 'H' \rightarrow 2$$

- SMOTE for Class Imbalance:

For each minority class sample, generate synthetic samples by interpolation:

$$v_{new} = v_{minority} + \lambda(v_{neighbor} - v_{minority})$$

where λ is a random value between 0 and 1, and $v_{neighbor}$ is a randomly chosen nearest neighbor of the minority class.

2. Model Architecture:

- LSTM Layer:

Input normalized data *V'* to the LSTM to capture temporal dependencies across time steps.

Generate hidden states *H* representing temporal patterns for each time step.

- Self-Attention Mechanism:

  Calculate attention weights to prioritize key time steps:

$$\alpha_{t,t'} = Attention\ (H_t, H_{t'})$$

  Compute weighted sums of value vectors for each time step, producing a focused representation Z

- Fully Connected Layer:

  Feed Z to a dense layer for binary classification.

  Apply sigmoid activation to produce a probability:

$$y_{pred} = \sigma(W_{out}Z + b_{out})$$

  Classify as failure (1) if $y_{pred} > 0.5$, else no failure (0).

3. Model Training and Optimization:

- Binary Cross-Entropy Loss:

$$L = -\frac{1}{N}\sum_{i=1}^{N}(y_i \log(y_{pred,i}) + (1 - y_i)\log(1 - y_{pred,i}))$$

- Optimizer: Use Adam optimizer for gradient descent.

  Early Stopping: Monitor validation loss to prevent over fitting, stopping if no improvement is observed over a set number of epochs.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

4.1. Dataset Description

The UCI AI4I dataset consists of 10,000 records with 14 attributes primarily focused on machine operating conditions and failure types (Table 1). Each record is uniquely identified by a Unique Data Identifier (UDI) and is categorized by Product ID and Type representing different machine models. The dataset includes numerical features such as Air temperature [K], Process temperature [K], Rotational speed [rpm], Torque [Nm] and Tool wear [min], which capture the operating parameters of industrial machines. The target variable which is Machine failure and it is binary which is 0 for no failure and 1 for failure indicating whether a machine failed during operation. Additionally, there are categorical indicators of different failure types, Tool Wear Failure (TWF), Heat Dissipation Failure (HDF), Power Failure (PWF), Overstrain Failure (OSF) and Random Failure (RNF). These failure types help identify specific causes contributing to overall machine failure. The dataset is complete with no missing values across any of the attributes making it well-suited for analysis of machine reliability and predictive maintenance applications.

Table 1: Machine failure types and variables

Failure Type	Data Type	Variable
Tool Wear Failure (TWF)	(0,1)	Tool Wear [min]
Heat Dissipation Failure (HDF)	(0,1)	Process Temperature [K]
Power Failure (PWF)	(0,1)	Torque [Nm]
Overstrain Failure (OSF)	(0,1)	Rotational Speed [rpm]
Random Failure (RNF)	(0,1)	Various (Random Causes)

4.2. Experimental Setup

The research was conducted on a computer equipped with an 11th Gen Intel Core i7 processor, 16 GB of RAM and Microsoft Windows 11 as the operating system. The framework was implemented using Python, developed within the Anaconda IDE which facilitated efficient management of project dependencies through its virtual environment capabilities.

4.3. Performance Analysis

The proposed work leverages a deep learning model to predict machine failure using sensor data. Data preprocessing involves normalizing features to ensure consistent scales and encoding categorical features. To address potential class imbalance, SMOTE is employed to

generate synthetic samples for the under represented class. The core of the model is an LSTM model, which is expert at capturing temporal dependencies within the time series data. A self-attention mechanism is incorporated to focus on the most relevant parts of the input sequence. Finally, a fully connected layer with a sigmoid activation function produces the probability of machine failure. The model is trained using binary cross-entropy loss and optimized with the Adam optimizer. Early stopping is employed to prevent over fitting and ensure optimal performance.

Seon, J., et. al., 2023 introduced Graph SAGE with Contrastive Encoder (GCE) to address the challenge of imbalanced datasets in IIoT systems. By leveraging graph-based representation and contrastive learning, GCE significantly improved classification accuracy compared to traditional methods. Venkatasubramanian, S., et. al., 2022 explored the potential of IIoT sensor data for industrial device breakdown detection. By addressing challenges like data noise and missing values, and by employing data fusion techniques, the study leveraged deep learning models to effectively identify faults. The proposed method evaluated on the CWRU dataset, demonstrated high accuracy and efficiency. Assagaf, I., et. al., 2023 explored the integration of machine learning to enhance predictive maintenance by developing optimized models using Logistic Regression (LR), K-nearest Neighbors (kNN) and Artificial Neural Networks (ANN). After data cleaning and feature scaling, the models had predicted and classified failures based on environmental features, machine characteristics and tool wear. The results had shown that the ANN model had outperformed others achieving the highest classification accuracy and consistently balanced performance across various validation methods.

Table 2: Performance Metricsobtained by the proposed work and the works of Seon, J., et. al., 2023, Venkatasubramanian, S., et. al., 2022 and Assagaf, I., et. al., 2023

Works	Sensitivity	Specificity	Accuracy	Precision	Recall	F-1 Score
Proposed Method	0.967	0.963	0.961	0.915	0.947	0.965
Seon, J., et. al., 2023	0.954	0.958	0.951	0.912	0.935	0.964
Venkatasubramanian, S., et. al., 2022	0.947	0.941	0.943	0.909	0.932	0.959
Assagaf, I., et. al., 2023	0.829	0.826	0.805	0.807	0.825	0.813

Table 2 presents the performance metrics achieved by the proposed method compared to the works of Seon, J., et. al., (2023), Venkatasubramanian, S., et. al., (2022) and Assagaf, I., et. al., (2023). The metrics include sensitivity, specificity, accuracy, precision, recall, and f-1 score. Table 2 shows that the proposed model outperforms the works of Seon, J., et. al., 2023, Venkatasubramanian, S., et. al., 2022 and Assagaf, I., et. al., 2023 with accuracy of 96% for failure detection. Attention mechanism effectively focuses on key sensor features, improving prediction accuracy. Figure 2 shows Graphical representation of performance obtained by the proposed work and the works of Seon, J., et. al., 2023, Venkatasubramanian, S., et. al., 2022 and Assagaf, I., et. al., 2023. The proposed method found to be the effective approach for this predictive task, exhibiting superior performance in all evaluation metrics.

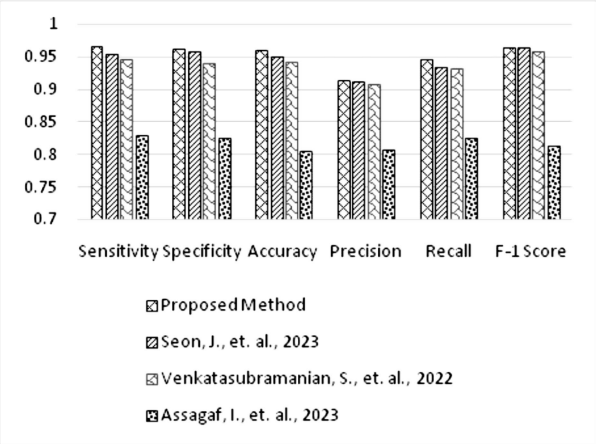


Figure 2: Graphical representation of performance obtained by the proposed work and the works of Seon, J., et. al., 2023, Venkatasubramanian, S., et. al., 2022 and Assagaf, I., et. al., 2023

Table 3 provides a comparative analysis of performance measures for the proposed method against those in the works of Seon, J., et. al., (2023), Venkatasubramanian, S., et. al., (2022), and Assagaf, I., et. al., (2023). Metrics include True Negative Rate (TNR), True Positive Rate (TPR), False Negative Rate (FNR), and False Positive Rate (FPR). This table highlights the proposed model's improved balance in true and false detection rates compared to previous models.

Table 3: Comparative analysis of performance measures obtained by proposed work and the works of Seon, J., et. al., 2023, Venkatasubramanian, S., et. al., 2022 and Assagaf, I., et. al., 2023

Works	TNR	TPR	FNR	FPR
Proposed Method	0.91	0.94	0.06	0.08
Seon, J., et. al., 2023	0.87	0.88	0.12	0.13
Venkatasubramanian, S., et. al., 2022	0.83	0.85	0.15	0.17
Assagaf, I., et. al., 2023	0.86	0.89	0.11	0.14

Figure 3 compares the performance of the proposed system with the works of Seon, J., et. al., (2023), Venkatasubramanian, S., et. al., (2022), and Assagaf, I., et. al., (2023). The proposed system consistently demonstrates higher precision across most recall values, indicating improved accuracy in positive prediction identification. This graph highlights the proposed system's superior performance in achieving high precision over a range of recall values, making it more reliable for accurate predictions.

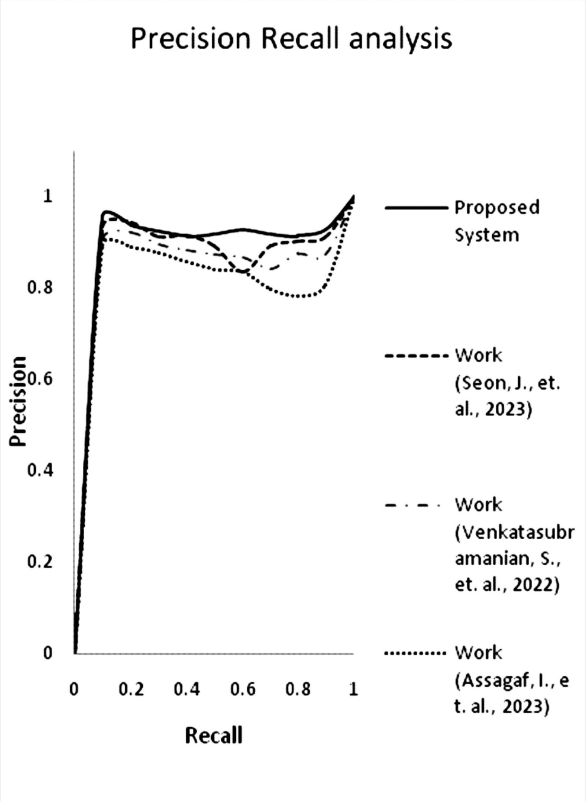
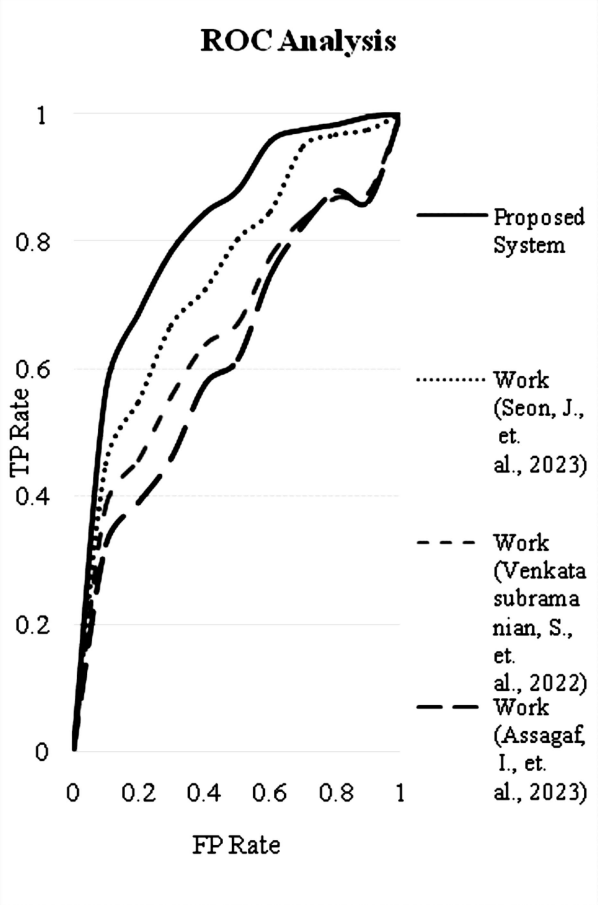


Figure 3: Precision Recall Analysis of proposed work and the works of Seon, J., et. al., 2023, Venkatasubramanian, S., et. al., 2022 and Assagaf, I., et. al., 2023

Figure 4 compares the True Positive Rate (TP Rate) against the False Positive Rate (FP Rate) for the proposed system and the works of Seon, J., et. al., (2023), Venkatasubramanian, S., et. al., (2022), and Assagaf, I., et. al., (2023). The proposed systemachieves a consistently higher TP Rate across most FP Rate values, indicating better overall classification performance. This analysis illustrates the superior capability of the proposed system in distinguishing true positives from false positives, showcasing higher reliability and accuracy in classification.



of Seon, J., et. al., 2023, Venkatasubramanian, S., et. al., 2022 and Assagaf, I., et. al., 2023

The proposed predictive maintenance model leverages an attention-based LSTM network to enhance machine failure prediction in IIoT environments by combining temporal modeling with a self-attention mechanism. This approach offers significant advantages over existing methods by capturing long-term dependencies within time-series data and focusing on critical timesteps, resulting in high accuracy, precision, recall, and F1-score. Compared to existing works the proposed work surpasses their results by effectively balancing the dataset with SMOTE and emphasizing relevant features, which improves both predictive accuracy and interpretability. Unlike the work by Assagaf et al. (2023), the proposed work is better equipped to handle class imbalance, resulting in more consistent prediction performance.

V. CONCLUSION

In conclusion, this paper presents a lightweight predictive maintenance framework for IIoT environments, addressing challenges such as data complexity, limited edge resources, and the need for real-time adaptability. The proposed work incorporates an LSTM model that effectively captures temporal dependencies and prioritizes critical features, enhancing the accuracy and interpretability of failure predictions. By employing techniques like Min-Max normalization, SMOTE for class balancing, and self-attention, the framework achieves high accuracy, precision, recall, and F1-score making it well-suited for predictive maintenance tasks. Compared to existing models, this framework demonstrates significant improvements in predictive performance while also optimizing computational efficiency for deployment on edge devices. With rigorous evaluation through simulations and real-world tests, the framework shows promise in reducing downtime, maintenance costs, and energy consumption in IIoT settings. Thus this study provides a foundational step toward more adaptive, cost-effective predictive maintenance solutions in IIoT environments.

REFERENCES

[1] Hasan, M.K., Sulaiman, R., Islam, S. and Rehman, A.U., 2023. An Explainable Ensemble Deep Learning Approach for Intrusion Detection in Industrial Internet of Things. IEEE Access.

[2] Karacayılmaz, G. and Artuner, H., 2024. A novel approach detection for IIoT attacks via artificial intelligence. Cluster Computing, pp.1-19.

[3] Alalayah, K.M., Alrayes, F.S., Alzahrani, J.S., Alaidarous, K.M., Alwayle, I.M., Mohsen, H., Ahmed, I.A. and Al Duhayyim, M., 2023. Optimal Deep Learning Based Intruder Identification in Industrial Internet of Things Environment. Comput. Syst. Sci. Eng., 46(3), pp.3121-3139.

[4] Bugshan, N., Khalil, I., Rahman, M.S., Atiquzzaman, M., Yi, X. and Badsha, S., 2022. Toward trustworthy and privacy-preserving federated deep learning service framework for industrial internet of things.

IEEE Transactions on Industrial Informatics, 19(2), pp.1535-1547.

[5] Mohy-eddine, M., Guezzaz, A., Benkirane, S. and Azrour, M., 2023. An effective intrusion detection approach based on ensemble learning for IIoT edge computing. Journal of Computer Virology and Hacking Techniques, 19(4), pp.469-481.

[6] Ikram, S.T., Priya, V., Anbarasu, B., Cheng, X., Ghalib, M.R. and Shankar, A., 2022. Prediction of IIoT traffic using a modified whale optimization approach integrated with random forest classifier. The Journal of Supercomputing, 78(8), pp.10725-10756.

[7] Xu, L., Cao, S., Li, X. and Gulliver, T.A., 2024. Analysis and Prediction of Mobile Industrial Internet of Things (IIoT) Communications Based on FL-GLP-Net. IEEE Internet of Things Journal.

[8] Roopa, M.S., Pallavi, B., Buyya, R., Venugopal, K.R., Iyengar, S.S. and Patnaik, L.M., 2021. Social Interaction-Enabled Industrial Internet of Things for Predictive Maintenance. In ICT Systems and Sustainability: Proceedings of ICT4SD 2020, Volume 1 (pp. 661-673). Springer Singapore.

[9] Bulla, C. and Birje, M.N., 2022. Anomaly detection in industrial IoT applications using deep learning approach. Artificial Intelligence in Industrial Applications: Approaches to Solve the Intrinsic Industrial Optimization Problems, pp.127-147.

[10] Li, H., Li, S. and Min, G., 2024. Lightweight privacy-preserving predictive maintenance in 6G enabled IIoT. Journal of Industrial Information Integration, 39, p.100548.

[11] Soliman, S., Oudah, W. and Aljuhani, A., 2023. Deep learning-based intrusion detection approach for securing industrial Internet of Things. Alexandria Engineering Journal, 81, pp.371-383.

[12] Friha, O., Ferrag, M.A., Benbouzid, M., Berghout, T., Kantarci, B. and Choo, K.K.R., 2023. 2DF-IDS: Decentralized and differentially private federated learning-based intrusion detection system for industrial IoT. Computers & Security, 127, p.103097.

[13] Misbha, D.S., 2022, December. Detection of Attacks using Attention-based Conv-LSTM and Bi-LSTM in Industrial Internet of Things. In 2022 International Conference on Automation, Computing and Renewable Systems (ICACRS) (pp. 402-407). IEEE.

[14] Schmieg, T. and Lanquillon, C., 2024. Time Series Representation Learning: A Survey on Deep Learning Techniques for Time Series Forecasting. In International Conference on Human-Computer Interaction (pp. 422-435). Springer, Cham.

[15] Yang, H. and Deng, F., Residual Life Prediction Method of Industrial Equipment Based on Deep Learning and Attention Mechanism.

[16] Chander, N. and Upendra Kumar, M., 2024. Enhanced pelican optimization algorithm with ensemble-based anomaly detection in industrial internet of things environment. Cluster Computing, pp.1-19.

[17] Isah, A., Shin, H., Oh, S., Oh, S., Aliyu, I., Um, T.W. and Kim, J., 2023. Digital Twins Temporal Dependencies-Based on Time Series Using Multivariate Long Short-Term Memory. Electronics, 12(19), p.4187.

[18] Smmarwar, S.K., Gupta, G.P. and Kumar, S., 2023. AI-empowered malware detection system for industrial internet of things. Computers and Electrical Engineering, 108, p.108731.

[19] Assagaf, I., Sukandi, A., Abdillah, A.A., Arifin, S. and Ga, J.L., 2023. Machine predictive maintenance by using support vector machines. Recent in Engineering Science and Technology, 1(01), pp.31-35.

[20] Seon, J., Lee, S., Sun, Y.G., Kim, S.H., Kim, D.I. and Kim, J.Y., 2023. GraphSAGE with contrastive encoder for efficient fault diagnosis in industrial IoT systems. ICT Express, 9(6), pp.1226-1232.

[21] Venkatasubramanian, S., Raja, S., Sumanth, V., Dwivedi, J.N., Sathiaparkavi, J., Modak, S. and Kejela, M.L., 2022. Fault diagnosis using data fusion with ensemble deep learning technique in IIoT. Mathematical Problems in Engineering, 2022(1), p.1682874.