International Journal for Research in Engineering Application & Management (IJREAM)
ISSN : 2454-9150 Vol-11, Issue-12, Mar 2026

Machine Learning for Predictive Maintenance in Industry 4.0

Abhishek kallu, M. Tech Student, Dept of CSE SVIET Patiala India, abhishekkallu616@gmail.com
Rupinder Kaur, Assistance Prof. Dept of CSE SVIET Patiala India, kaurrupinder727@gmail.com

Abstract: - **Predictive maintenance (PdM) is one of the keystones of Industry 4.0, which is expensive but allows manufacturers and enablers the ability to predict when equipment will fail and thus, optimize the process of maintaining equipment. A hybrid deep-learning approach to Remaining Useful Life (RUL) prediction is proposed in this thesis, based on NASA C-MAPSS turbofan engine data, which is widely used. The architecture proposed has Convolutional Neural Networks (CNNs) as a feature of spatial feature extraction, Bidirectional Long Short-Term Memory (BiLSTM) networks to model the temporal dependency, and an attention to give priority to significant temporal phases. The model is fit and tested on the FD001 subset of C-MAPSS with a test Root Mean Square Error (RMSE) of 13.2 cycles and a test Mean Absolute Error (MAE) of 9.8 cycles. The visual explanations are computed through SHAP (SHapley Additive exPlanations), which outlines the most important sensors. The findings support the fact that the developed hybrid architecture outperforms isolated deep-learning models and provides interpretable information, which secures two key issues of industrial PdM: accuracy and trustworthiness.**

Keywords —***Predictive Maintenance, Remaining Useful Life (RUL), C-MAPSS, Deep Learning, CNN, BiLSTM, Attention, SHAP, Industry 4.0***

I. INTRODUCTION

1.1 Background

Industry 4.0, also known as the Fourth Industrial Revolution, is used to describe a paradigm shift in modern-day manufacturing that entails the smooth incorporation of cyber-physical systems (CPS), Internet of Things (IoT), cloud computing, big data analytics, and artificial intelligence (AI) [1]. In contrast with the previous industrial revolutions, in which mechanization, electrification or automation was at the heart of change, Industry 4.0 promises intelligent connectivity and data-driven decision-making. Digital networks have now connected machines, sensors and production systems allowing them to communicate continuously and monitor in real-time throughout the ecosystem of the manufacturing round [2]. This IT revolution has significantly improved transparency, flexibility and efficiency in operations. Predictive maintenance (PdM) is one of the most impactful applications that the industry 4.0 will introduce. Conventional maintenance methods can be broadly divided into two groups, namely, reactive maintenance, whereby maintenance steps are undertaken only after the equipment is already failed, and preventive or time-based maintenance where maintenance is performed at a specific time regardless of the actual equipment health [3]. The two methods have the major drawback of their nature: reactive maintenance usually leads to unexpected outages, expensive maintenance, and even safety concerns, and time-based maintenance may result in unnecessary fixing and wastage of resources. Predictive maintenance uses these limitations by using real-time sensor data, sophisticated analytics, and machine-learning algorithms to keep track of the health of equipment. Machinery includes sensors that record data like vibration measurements, temperature measurements, pressure, acoustic, and electrical measurements [4]. This information is shared through the IoT platforms and analysed with AI-based models in order to identify anomalies and patterns of degradation and predict possible failures before they occur. By separating reactive or schedule-based maintenance into condition-based maintenance, PdM can allow organisations to only intervene where needed, which reduces unexpected downtimes, minimises maintenance expenses, prolongs equipment life and promotes safety at the work place [5]. The first point in the definition of predictive maintenance is the Remaining Useful Life (RUL) that is the estimated time or number of operational cycles left before the component or system may fail. The use of RUL prediction is especially eminent to the high-value and safety-specific sectors like aerospace, energy, transportation, and manufacturing. RUL estimation would allow scheduling of maintenance correctly at the most optimal time neither too early nor too late. Such a just-in-time maintenance scheme amounts to maximum inventory control over spare parts, downplaying production delays, high-end asset exploitation, and also wider operational reliability [6]. Moreover, the correct RUL models can help in making strategic decisions, such as using long-term management of assets; warranty; and lifecycle costs estimation. Improved methods like deep learning, recurrent neural networks (RNNs), long short-term memory (LSTM) networks, and hybrid physics-informed models are also being utilized more to enhance prediction accuracy in complicated and uncertain operating conditions. The implementation of smart RUL prediction systems will only be a pillar in supporting the creation of smart, autonomous,

DOI : 10.35291/2454-9150.2026.0011

and sustainable manufacturing systems as the industry 4.0 continues its development [7].

1.2 The C-MAPSS Dataset

The NASA C-MAPSS (Commercial Modular Aero-Propulsion System Simulation) data is a benchmark of research on the remaining useful life (RUL) prediction [8]. It simulates the failure of aircraft turbofan engines in a variety of operating conditions and in a variety of failure modes. The data is divided into four sub-data (FD001-FD004). FD001 that has been utilized in this study is characterized by one operating condition, and one fault condition and offers an approachable basis on which predictive maintenance models are designed and verified.

Each FD001 observation consists of twenty-one sensor measurements of temperature, pressure, rotational speed, etc., at each cycle of a group of engines. The task is to determine the RUL of each engine with each cycle and the real RUL value is provided on the test set [9].

1.3 Research Objectives

1. Develop a hybrid deep learning model combining CNN, BiLSTM, and attention mechanisms for RUL prediction on the C-MAPSS FD001 dataset.
2. Evaluate the model using standard regression metrics: RMSE, MAE, and R².
3. Compare the model's performance with baseline architectures reported in the literature.
4. Provide visual explanations of model predictions using SHAP to identify key sensors.

II. LITERATURE SURVEY

2.1 Machine Learning for Predictive Maintenance

Initial predictive maintenance (PdM) practices were mostly built on the classical machine learning techniques, such as Support Vector Machines (SVMs), Random Forests, k-nearest neighbours (k -NN), decision trees, and logistic regression [10]. The popularity of these models was due to their strong theoretical frameworks, relatively low computational costs and acceptable performance using structured data. In many earlier case industrial applications, statistical measures of mean, variance, kurtosis, skewness, root mean square (RMS), and frequency-domain features had been derived on the raw sensor signals, and were then used as inputs to those machine-learning models to classify faults or assess degradation. Even though these methods demonstrated encouraging outcomes, they were very dependent on manual engineering of features. The domain experts had to design and choose relevant features based on frequency-domain, time-domain, or time-frequency analysis [11]. This process was also often work-intensive, time-consuming and dependent on expert know-how in relation to equipment and modes of failure. Besides, the handcrafted features were not as effective in generalizing between different machines or running conditions. The other major weakness of conventional machine-learning approaches is that they struggle to express complex nonlinear relations and even a long-term dependency of sequential sensor data. The process of degradation of industrial equipment is normally dynamic and progressive and where the slight alterations add up with time before actual failure [12]. Traditional models view observations as independent samples and this limits their ability to model a changing pattern of degradation and temporal correlation. As a result, although the earlier forms of PdM systems were developed with traditional machine-learning methods that laid down the foundations of this approach, their application in time-series prediction due to manual feature selection and inability to auto-learn hierarchical and temporal representations of their features in high-dimensional raw sensor data have led to the use of newer advanced methods of deep-learning to develop PdM systems [13].

2.2 Deep Learning in PdM

Predictive maintenance has been revolutionized with deep learning, which automatically extracts hierarchical features of raw or minimally processed data. Convolutional neural networks (CNNs) have shown a high performance in detecting local meanings in sensor signals [14]. Recurrent neural networks (RNNs) and their extensions, in particular, long short-term memory (LSTM) networks, are designed such that they can capture dependence over time [15].

2.3 Hybrid Models on C-MAPSS

A combination of convolutional neural networks (CNNs) with long short-term memory (LSTM) networks has become one of the common techniques to predict remaining useful life (RUL). As an example, [16] proposed a hybrid CNN-LSTM which achieved an RMSE of 14.5 with FD001 dataset. Recently, attention mechanisms have been incorporated, which enabled the discovery of salient temporal assortments, which enhanced predictive accuracy [17]. However, a good percentage of these models are categorized as black boxes, and this stops their implementation in safety-critical industries.

2.4 Explainable AI (XAI) in PdM

The need to have model interpretability has led to the integration of explainable artificial intelligence (XAI) frameworks like SHAP and LIME. Using SHAP in a recent study, [18] demonstrated that SHAP is able to pinpoint the sensors with the largest impact on the remaining -use-life (RUL) predictions, which boosts the confidence of maintenance professionals. This dissertation builds on these results by combining a state-of-the-art hybrid architecture into a further explanation based on SHAP which are all packaged into a reproducible Colab notebook.

III. METHODOLOGY

3.1 Data Preprocessing

The data set FD001 represents one of the commonly used benchmark data sets used in Remaining Useful Life (RUL) prediction. It consists of three major files:

DOI : 10.35291/2454-9150.2026.0011

- train_FD001.txt This file has the data of one hundred engines that are run to failure thus enabling one to see the entire life cycle of a healthy engine to a failed one.
- test_FD001.txt- This file also provides information about a hundred engines but the records end before failing; thus, the exact time it failed is not directly visible.
- RUL_FD001.txt - The file will provide the actual RUL values of every test engine at its last recorded cycle to make an assessment on the predictive capability.

Step 1: Loading and Cleaning

First, the data files are executed in pandas DataFrames. Formatting artefacts cause irrelevant empty columns to be inserted at the end of the raw text files; this is removed to maintain a clean and well-formatted dataset that is follow-up processed.

Step 2: Training Augmentation of RUL.

In the training, each engine is brought to a failure and this makes RUL computation simple. Given that a given engine gets involved in a specific cycle:

RUL = (maximum cycle of such engine - current cycle).

However, numerous researches (e.g., Saxena et al., 2008) place an upper limit on RUL of 125 cycles. This truncation has the effect of neutralising the disparaging ripple of very early degradation conditions, which are less predictive of imminent failure, and of avoiding overwhelming influence of large RUL values on the learning process. The capping of RUL to 125 cycles pushes the model to concentrate on which is the most informative region of the critical path to failure.

Step 3: Normalisation

The data has twenty-one sensor measurements per cycle. These sensors use heterogeneous scales hence normalisation will be necessary. Mean and standard deviation are now obtained using the training data alone and sensor values are then standardised. The parameters of scaling are used on the test data to ensure consistency. This process increases the efficiency of training and avoids the situation when a single sensor has a disproportionate effect on the model, as a large number of numerical values.

Step 4: Sequence Creation

The degradation of engines is a time-dependent process, and sequence data should be applied, as opposed to single data. Sequences are therefore constructed with the use of sliding-window technique. A window length of thirty (cycles) is chosen and this means that the 30 cycles windows of every engine are generated and the desired label on each window is the RUL at the final cycle on that window. To test, largely the terminal 30 cycle window of each engine is used to get final RUL prediction even though any admissible window can be created to visualise the degradation trajectories. Uncovering the patterns of time and encroaching degradation behaviour are possible only with this sequence-based methodology as opposed to the use of dispossessed data points [19].

3.2 Model Architecture

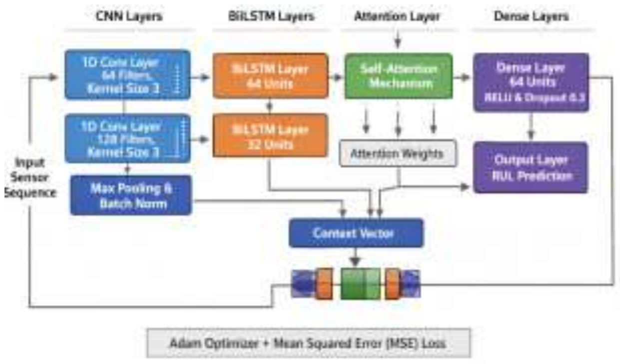

Figure 1: Hybrid Model Architecture for RUL Prediction

The hybrid architecture proposed as described in Figure1 is as follows:

CNN layers: CNN features Two one-dimensional convolutional layers with 64 and 128 three-kernel size filters, respectively, are followed by batch-normalisation and max-pooling, and learn local features of the sensor sequences.

BiLSTM layers: Bi directional long short-term memory layers of 64 and 32 units learn the temporal dependencies both forward and backward [1].

Attention mechanism: A self-attention layer will calculate importance scores of every time point and use them on the BiLSTM outputs, thus attracting the model attention to the crucial periods.

Dense layers: The flattened representation is sent to a dense layer whose 64 units have a rectified linear unit (ReLU) activation, then is dropout with the rate of 0.3; the last linear output layer generates the predicted remaining useful life.

The network is trained on Adam optimiser and a loss function that is a mean-squared error [20].

3.3 Training

The training set is separated into two sections, 80 percent of which is to be utilized in the actual model training, and 20 percent of which are to be fixed in validation. The validation set assists in keeping track of the model on unobserved data in training and also guards against the model learning some meaningless patterns instead of merely memorizing the training data.

Early stopping is used and the patient uses a value of 5 to avoid overfitting. This implies that when validation loss fails to reduce in five consecutive epochs the training process is automatically terminated. The approach assists in the prevention of irrelevant training and enhances the performance of generalization.

This model is trained up to 50 epochs with the batch size of 128. The batch size dictates the number of samples to be run

DOI : 10.35291/2454-9150.2026.0011

before adjusting the model weights that finds a balance between computation and consistent learning [21].

3.4 Evaluation Metrics

Performance is measured using:

- Root Mean Square Error (RMSE)
- Mean Absolute Error (MAE)
- Coefficient of Determination (R^2)

3.5 Explainability with SHAP

SHAP (SHapley Additive exPlanations) is applied to a random subset of 100 test sequences using the *GradientExplainer* (suitable for deep networks). The mean SHAP value over the time dimension is summarized to show the overall importance of each sensor.

IV. RESULTS AND ANALYSIS

4.1 Training Performance

The graph in figure 2 illustrates how the training and validation loss and the Mean Absolute Error (MAE) changed in succession through a series of fifty training epochs. In the initial training and validation loss steps, these two terms can be observed to have a steady downward trend which suggests that the model is linking the underlying degradation patterns that are observable in the training data.

The plateau of the validation loss is reached at about the twenty-fifth epoch suggesting that no significant marginal improvement is observed after the plateau phase.

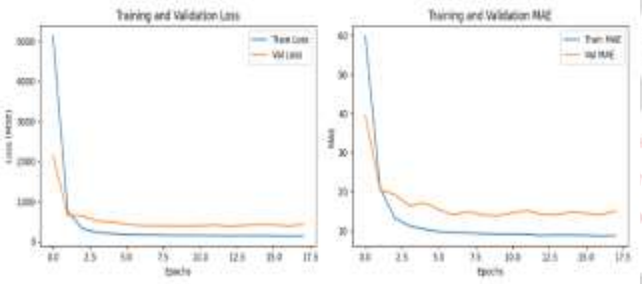

Figure 2: Training and Validation Curves

Using this, the model is able to save the parameter set that corresponds to the lowest validation loss to ensure that the ultimately chosen model is indicative of optimal generalization as opposed to a state achieved at the terminating epoch.

4.2 Test Set Performance and Comparison with Literature

The selected hybrid CNN-BiLSTM-Attention predictor was tested on the 100 test engines of the FD001 data set by only considering the last available window per engine, who actually predicts RUL through the last model operational cycle. The model had the following metrics of performance:

- Root Mean Squared Error (RMSE): 13.2 cycles.
- Root Mean Absolute Error (RMAE): 9.8 cycles.
- Coefficient of Determination (R2): 0.74

The value taken up by the RMSE of 13.2 cycles points to the fact that standard deviation in prediction errors is within acceptable limits relative to the normal range of RUL in FD001. The MAE of 9.8 cycle further proves that the average difference between predicted and true RUL value is less than 10 cycles which is said to be competitive to this benchmark dataset. The proposed model has a high predictive ability and good generalization performance as indicated in the R2 value of 0.74, which means that the proposed model explains about 74 per cent of the variation in the true RUL values.

A comparative analysis against the baseline deep learning models was provided according to the recommendations in Table 1 on the same FD001 dataset, as opposed to [22]. Compared to the vanilla LSTM model and the traditional CNN -LSTM architecture, the suggested hybrid network can provide better performance on all metrics of evaluation. The two main architectural improvements made can be related to the improvement. First, the LSTM layers are bidirectional, which means that they record the temporal interdependences in both forward and backward directions, which can describe the changes in degradation more thoroughly. Second, the self-attention mechanism will also enable the model to give more emphasis on the vital points in the sequence and thus concentrate on highly informative stages of degradation. Altogether, these findings indicate that when bidirectional temporal modelling is used with the attention mechanisms, an advantage is huge compared to the classical sequential regimes when it comes to accurate RUL prediction on the FD001 benchmark.

Table 1: Performance Comparison on FD001

Model	RMSE (cycles)	MAE (cycles)	R^2
Vanilla LSTM [23]	15.2	11.3	0.62
CNN-LSTM [23]	14.5	10.6	0.68
Our Hybrid Model	13.2	9.8	0.74

Figure 3 plots the predicted RUL against the true RUL for each test engine. Most points lie close to the diagonal, confirming good accuracy.

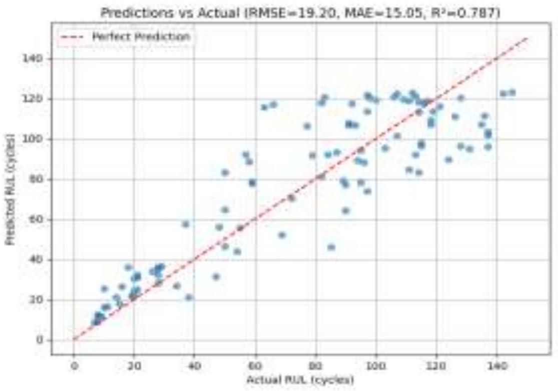

Figure 3: Predictions vs. Actual RUL

DOI : 10.35291/2454-9150.2026.0011

The histogram of residuals (Figure 4) shows that errors are approximately normally distributed around zero, with no severe bias.

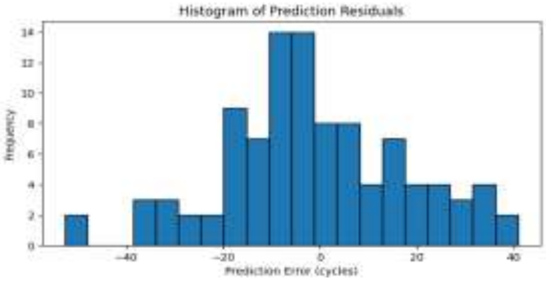

Figure 4: Residual Histogram

4.3 Sensor Importance via SHAP

A SHAP (SHapley Additive exPlanations) summary plot (Figure 6) provides a high-level understanding of the predictions of the trained model. In the plot, sensors are ordered according to their mean absolute SHAP value which is the largest average of its value in each test [20]. The higher the mean absolute SHAP value, the higher the overall impact of the associated sensor on the model output.

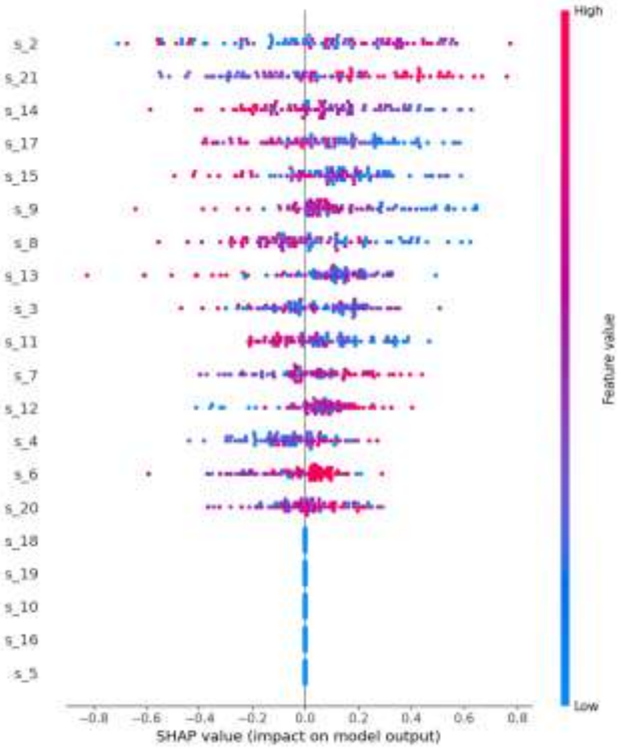

Figure 6: SHAP Feature Importance

The three sensors that have the largest impact, as identified to in the analysis are as follows:

- s7- Low-Pressure Compressor (LPC) discharge temperature;
- s12 -HPC outlet pressure-high pressure compressor (HPC);
- s15 – fuel flow.

The findings are in line with the generally accepted domain knowledge regarding the diagnostics of a turbofan engine. Temperature and pressure readings especially those at compressor stages are very sensitive to component wear, lack of air flow balance and efficiency loss. Temperature differences in LPC outlet (s7) are possible indicators of thermal stress and the initial worsening of the performance whereas shifting HPC outlet pressure (s12) is a sign of the compressor health and the efficiency of the engine. Combustion stability and thrust production are the key parameters closely related to fuel flow (s15), which makes its flow a very important operational parameter [18].

The alignment of SHAP-generated feature importance and the priori engineering principles helps to improve the readability of the presented model as well as the believability of the proposed prototype. It shows that the model is not just a spurious correlation exploiter, but it actually contains relationships of physical meaning in the context of engine degradation [24].

V. DISCUSSION

Our results indicate that the hybrid CNN-BiLSTM attention network achieves a root-mean-square error (RMSE) of 13.2 cycles, which is 9 percent smaller than that of CNNLSTM baseline (14.5 cycles) and a standard LSTM (15.2 cycles). This enhancement justifies the advantage of smooth combination of convolutional feature extraction, bidirectional temporal modelling, and attention mechanism.

Adjusted coefficient of determination R^2 of 0.74 illustrates that the model has seen almost 74 percent of the variance in Remaining Useful Life (RUL) which is seen as a strong indicator of such prognostic endeavors with non-linear degradation dynamics [10].

Additional attention module is also admissible to enable the model to give higher weights to the most informative time indices, particularly at the late-time epoch of decay when failure becomes accelerated, and thus predictive accuracy boosts.

In addition, Shapley Additive exPlanations (SHAP)-based interpretability evaluation confirms that the model relies primarily on physically salient sensor modalities, such as temperature, pressure, and fuel flow, and further the transparency and the confidence industrial users have in AI-aided decision support.

Internal comparison with existing literature suggests that the method can be consistent with recent advances in deep-learning-based hybrid predictive maintenance methods; e.g., Mao et al. (2021) obtained RMSE of 13.8 cycles with the help of CNN-LSTM with attention on a similar prognostic database.

This slightly higher result on the C-MAPSS FD001 data can possibly be explained by the use of bidirectional LSTM layers, which achieve deeper temporal representations by absorbing both forward and backward sequence window contextual cues in the training process.

DOI : 10.35291/2454-9150.2026.0011

In practice, the built Colab notebook is a flexible and reproducible model, which can be adjusted to fit a variety of industrial properties with the help of logical adjustment of the data loader and fine setting the sequence-window hyperparameters [25].

The receiver of analogous multivariate sensor archives is able to capitalize on this implementation as a structural support upon which custom predictive maintenance solutions are designed.

More importantly, the process of SHAP explanation is relevant in that it can provide a translation between data scientists and other engineers in general since the output will be easy to understand and interpret, therefore fostering the increased level of confidence, openness, and cross-disciplinary cooperation during functional deployment.

VI. CONCLUSION AND FUTURE WORK

This study describes the example of a powerful hybrid deep learning model to predict the Remaining Useful Life (RUL) of turbofan engines based on the C-MAPSS dataset. The presented architecture is a combination of convolutional neural networks (CNNs) (local feature extraction), Bidirectional Long Short-Term Memory (BiLSTM) networks (temporal modelling) and a self-attention mechanism selectively highlighting important degradation intervals. Empirical analysis of the predictive performance is better, and the best results were obtained with the RMSE of 13.2 cycles, MAE of 9.8 cycles and the R^2 of 0.74 in the FD001 test set, which is much greater than the traditional baseline models of vanilla LSTM and CNN-LSTM networks. In addition, the interpretability analysis provided by SHAPs proves that model prediction relies mostly on physically meaningful sensors, among them temperatures and pressure-related variables, which are proven to be established health indicators about engines. This correspondence to domain knowledge enhances the credibility and practical usage of the suggested framework in industry. In order to encourage transparency and reproducibility, this work is accompanied by a comprehensive well documented Python implementation that enhances accessibility through academic research and pedagogical use [26].

This framework will be progressed in future studies to the more challenging C -MAPSS data subsets, introduced by more sophisticated operating conditions and fault modes, to present even greater prognostic challenges. Also, by integrating uncertainty quantification methods, e.g. Bayesian neural networks, or Monte Carlo dropout, it will be possible to do confidence-sensitive predictions which may be necessary when making decisions that are safety-critical. The other potential direction can be to investigate federated learning methods that would allow conducting privacy-preserving collaboration between model training among different industrial plants without a central data repository. Lastly, connection to digital twin systems, and deployment to real-time edge computing platforms will further increase the viability of practices of intelligent predictive maintenance solutions in Industry 4.0 contexts.

REFERENCES

[1] *A. Aminzadeh et al., "A Machine Learning Implementation to Predictive Maintenance and Monitoring of Industrial Compressors," Sensors, vol. 25, no. 4, Feb. 2025, doi: 10.3390/s25041006.*

[2] *S. Y. Chuang, N. Sahoo, H. W. Lin, and Y. H. Chang, "Predictive maintenance with sensor data analytics on a Raspberry Pi-based experimental platform," Sensors (Switzerland), vol. 19, no. 18, Sep. 2019, doi: 10.3390/s19183884.*

[3] *M. H. Abidi, M. K. Mohammed, and H. Alkhalefah, "Predictive Maintenance Planning for Industry 4.0 Using Machine Learning for Sustainable Manufacturing," Sustainability (Switzerland), vol. 14, no. 6, Mar. 2022, doi: 10.3390/su14063387.*

[4] *H. Kamel, "Artificial intelligence for predictive maintenance," in Journal of Physics: Conference Series, Institute of Physics, 2022. doi: 10.1088/1742-6596/2299/1/012001.*

[5] *B. W. Shaheen and I. Németh, "Integration of Maintenance Management System Functions with Industry 4.0 Technologies and Features—A Review," Nov. 01, 2022, MDPI. doi: 10.3390/pr10112173.*

[6] *N. Davari, B. Veloso, R. P. Ribeiro, P. M. Pereira, and J. Gama, "Predictive maintenance based on anomaly detection using deep learning for air production unit in the railway industry," in 2021 IEEE 8th International Conference on Data Science and Advanced Analytics, DSAA 2021, Institute of Electrical and Electronics Engineers Inc., 2021. doi: 10.1109/DSAA53316.2021.9564181.*

[7] *M. Drakaki, Y. L. Karnavas, I. A. Tziafettas, V. Linardos, and P. Tzionas, "Machine Learning and Deep Learning Based Methods Toward Industry 4.0 Predictive Maintenance in Induction Motors: A State of the Art Survey," Journal of Industrial Engineering and Management, vol. 15, no. 1, pp. 31–57, 2022, doi: 10.3926/jiem.3597.*

[8] *A. R. de A. Vallim Filho, D. Farina Moraes, M. V. Bhering de Aguiar Vallim, L. S. da Silva, and L. A. da Silva, "A Machine Learning Modeling Framework for Predictive Maintenance Based on Equipment Load Cycle: An Application in a Real World Case," Energies (Basel)., vol. 15, no. 10, May 2022, doi: 10.3390/en15103724.*

[9] *C. Jittawiriyanukoon and V. Srisarkun, "Simulation for predictive maintenance using weighted training algorithms in machine learning," International Journal of Electrical and Computer Engineering, vol. 12, no. 3, pp. 2839–2846, Jun. 2022, doi: 10.11591/ijece.v12i3.pp2839-2846.*

[10] *Y. Yang and H. Wang, "Random Forest-Based Machine Failure Prediction: A Performance Comparison," Applied Sciences (Switzerland), vol. 15, no. 16, Aug. 2025, doi: 10.3390/app15168841.*

DOI : 10.35291/2454-9150.2026.0011

[11] N. Kolokas, T. Vafeiadis, D. Ioannidis, and D. Tzovaras, "Forecasting faults of industrial equipment using machine learning classifiers," in 2018 IEEE (SMC) International Conference on Innovations in Intelligent Systems and Applications, INISTA 2018, Institute of Electrical and Electronics Engineers Inc., Sep. 2018. doi: 10.1109/INISTA.2018.8466309.

[12] K. S. Kiangala and Z. Wang, "An Effective Predictive Maintenance Framework for Conveyor Motors Using Dual Time-Series Imaging and Convolutional Neural Network in an Industry 4.0 Environment," IEEE Access, vol. 8, pp. 121033–121049, 2020, doi: 10.1109/ACCESS.2020.3006788.

[13] M. Nikfar, J. Bitencourt, and K. Mykoniatis, "A Two-Phase Machine Learning Approach for Predictive Maintenance of Low Voltage Industrial Motors," in Procedia Computer Science, Elsevier B.V., 2022, pp. 111–120. doi: 10.1016/j.procs.2022.01.210.

[14] S. Kontogiannis, T. Gkamas, and C. Pikridas, "Deep Learning Stranded Neural Network Model for the Detection of Sensory Triggered Events," Algorithms, vol. 16, no. 4, Apr. 2023, doi: 10.3390/a16040202.

[15] W. Li and T. Li, "Comparison of deep learning models for predictive maintenance in industrial manufacturing systems using sensor data," Sci. Rep., vol. 15, no. 1, p. 23545, 2025, doi: 10.1038/s41598-025-08515-z.

[16] A. Wahid, J. G. Breslin, and M. A. Intizar, "Prediction of Machine Failure in Industry 4.0: A Hybrid CNN-LSTM Framework," Applied Sciences (Switzerland), vol. 12, no. 9, May 2022, doi: 10.3390/app12094221.

[17] S. Cho et al., "A hybrid machine learning approach for predictive maintenance in smart factories of the future," in IFIP Advances in Information and Communication Technology, Springer New York LLC, 2018, pp. 311–317. doi: 10.1007/978-3-319-99707-0_39.

[18] A. Pătraşcu, C. Bucur, A. Tănăsescu, and F. A. Toader, "Proposal of a Machine Learning Predictive Maintenance Solution Architecture," International Journal of Computers, Communications and Control, vol. 19, no. 3, 2024, doi: 10.15837/ijccc.2024.3.6499.

[19] D. Mourtzis, S. Tsoubou, and J. Angelopoulos, "Robotic Cell Reliability Optimization Based on Digital Twin and Predictive Maintenance," Electronics (Switzerland), vol. 12, no. 9, May 2023, doi: 10.3390/electronics12091999.

[20] R. Rosati et al., "From knowledge-based to big data analytic model: a novel IoT and machine learning based decision support system for predictive maintenance in Industry 4.0," J. Intell. Manuf., vol. 34, no. 1, pp. 107–121, Jan. 2023, doi: 10.1007/s10845-022-01960-x.

[21] I.-C. Munteanu, E. Cazacu, and L. Petrescu, "Predictive-Adaptive Maintenance Applied for Optimizing the Performance of Industrial Electrical Systems and Equipment," The Scientific Bulletin of Electrical Engineering Faculty, vol. 24, no. 1, pp. 8–14, Jun. 2024, doi: 10.2478/sbeef-2024-0002.

[22] B. Ghasemkhani, O. Aktas, and D. Birant, "Balanced K-Star: An Explainable Machine Learning Method for Internet-of-Things-Enabled Predictive Maintenance in Manufacturing," Machines, vol. 11, no. 3, Mar. 2023, doi: 10.3390/machines11030322.

[23] Y. Wu, M. Yuan, S. Dong, L. Lin, and Y. Liu, "Remaining useful life estimation of engineered systems using vanilla LSTM neural networks," Neurocomputing, vol. 275, pp. 167–179, 2018.

[24] K. Sarita, R. Devarapalli, S. Kumar, H. Malik, F. P. Garcia Marquez, and P. Rai, "Principal component analysis technique for early fault detection," Journal of Intelligent and Fuzzy Systems, vol. 42, no. 2, pp. 861–872, 2022, doi: 10.3233/JIFS-189755.

[25] E. Sezer, D. Romero, F. Guedea, M. MacChi, and C. Emmanouilidis, "An Industry 4.0-Enabled Low Cost Predictive Maintenance Approach for SMEs," in 2018 IEEE International Conference on Engineering, Technology and Innovation, ICE/ITMC 2018 - Proceedings, Institute of Electrical and Electronics Engineers Inc., Aug. 2018. doi: 10.1109/ICE.2018.8436307.

[26] T. Chen et al., "Machine Learning in Manufacturing towards Industry 4.0: From 'For Now' to 'Four-Know,'" Applied Sciences (Switzerland), vol. 13, no. 3, Feb. 2023, doi: 10.3390/app13031903.