Volume 10, Issue 12, December – 2025 International Journal of Innovative Science and Research Technology
ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25dec776

A Two Stage Artificial Intelligence Based Predictive Maintenance Model for Industry 4.0 Applications

Özkan Bartu Leylek[1]; Ömür Şansal Çenberli[1]; Muhammed Kürşad Uçar[2,3]

[1]Kontrolmatik Technologies, Design Center, 34475, Sarıyer, İstanbul, Türkiye
[2]Sakarya University, Faculty of Engineering, Electrical-Electronics Engineering, 54187, Serdivan, Sakarya, Türkiye
[3]MKU Technology, Sakarya University Technopolis, Serdivan, Sakarya, Türkiye

[1]https://orcid.org/0009-0002-5679-9328
[2]https://orcid.org/0009-0001-6993-5299
[3]https://orcid.org/0000-0002-0636-8645

Corresponding Author: Özkan Bartu Leylek[1]*

Publication Date: 2025/12/20

Abstract:

- ***Objective:***

With the advent of Industry 4.0, the complexity of production systems has increased, making the early prediction of machine failures critical for operational efficiency and continuity. In this study, a two-stage artificial intelligence-based predictive maintenance model is developed to detect machine failures and identify failure types in industrial systems.

- ***Methodology:***

The AI4I 2020 Predictive Maintenance dataset was used in this research. In the first stage, the most significant input features were identified through feature selection based on the Spearman correlation coefficient, followed by the application of a systematic sampling method to address data imbalance. During the fault detection stage, various machine learning algorithms (SVM, Ensemble Trees, Artificial Neural Networks, etc.) were comparatively evaluated. In the second stage, a partial least squares (PLS)-based modeling approach was employed for the classification of failure types.

- ***Results:***

In the first-stage results, the SVM and Ensemble Trees models demonstrated the highest performance, achieving accuracy rates above 92% and an AUC value of 0.98. In the second stage, the PLS based model achieved classification accuracies exceeding 95%, particularly for datasets consisting of six and seven features.

- ***Conclusion:***

The proposed two-stage predictive maintenance model offers a practical artificial intelligence solution that can contribute to the optimization of maintenance planning, enhancement of operational continuity, and reduction of maintenance costs in industrial systems. Owing to its modular structure, the model can be adapted to different production lines and is considered a decision-support tool that can be integrated into Industry 4.0 infrastructures.

Keywords: *Predictive Maintenance; Artificial Intelligence; Industry 4.0; Machine Learning; Fault Prediction.*

How to Cite: Özkan Bartu Leylek; Ömür Şansal Çenberli; Muhammed Kürşad Uçar (2025) A Two Stage Artificial Intelligence Based Predictive Maintenance Model for Industry 4.0 Applications. *International Journal of Innovative Science and Research Technology*,10(12), 1104-1112. https://doi.org/10.38124/ijisrt/25dec776

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25dec776

I. INTRODUCTION

With the advent of Industry 4.0, production systems have evolved into complex structures characterized by a high level of automation and intensive data flow. Within this new manufacturing paradigm, the continuity and performance of machine equipment have become decisive factors in terms of production efficiency and profitability. Machine failures lead to serious consequences such as unplanned downtime, production losses, increased maintenance costs, and reduced product quality [1]. Therefore, the early prediction of failures has become a critical requirement for ensuring continuity in production lines and preventing unexpected interruptions [2], [3]. Predictive maintenance systems minimize downtime by detecting potential failures before they occur and enable uninterrupted production processes [2], [3]. Moreover, these approaches optimize maintenance frequency, thereby reducing both unplanned maintenance costs and unnecessary preventive maintenance activities [4], [5]. Early fault detection also enhances machine reliability and occupational safety, particularly contributing to the prevention of accidents in high-risk industries [2], [6], [7]. In addition, estimating the remaining useful life of equipment enables effective planning of spare parts and maintenance resources, leading to more efficient utilization of operational assets [1], [8]. For these reasons, the early prediction of machine failures in Industry 4.0 environments is considered not merely a technical improvement but a strategic necessity from a production management perspective.

Predictive maintenance approaches developed within the scope of Industry 4.0 extensively utilize machine learning and artificial intelligence techniques to enable early fault detection and optimize maintenance planning. These systems monitor equipment conditions and predict potential failures in advance by analyzing large volumes of data collected from Internet of Things (IoT)-based sensors [9]. In the literature, the most commonly employed approaches for predictive maintenance include supervised learning algorithms (Support Vector Machines-SVM, Random Forest-RF), deep learning models (Convolutional Neural Networks-CNN, Recurrent Neural Networks-RNN), and hybrid combinations of these methods [9], [10], [11]. By leveraging historical operational data, these models identify fault patterns and estimate the remaining useful life of equipment [9], [12]. Furthermore, IoT and big data analytics infrastructures have enabled the real-time implementation of predictive maintenance, while edge computing-based approaches have reduced system response times by allowing data to be processed directly on the production line [13], [14].

Nevertheless, existing approaches in the literature exhibit several limitations. First, the success of predictive maintenance systems is highly dependent on data quality. Noisy, incomplete, or inconsistent data can negatively affect model prediction accuracy and lead to erroneous maintenance decisions [2][7]. Moreover, the integration of data from heterogeneous sources and the achievement of interoperability among different systems remain significant challenges [1]. The high complexity and limited explainability of deep learning-based models make it difficult for operators to trust model decisions [11], [15]. In addition, the sensor infrastructure, data acquisition systems, and computational resources required for deploying such systems involve high costs, which restrict their applicability for small and medium-sized enterprises [11], [16]. Finally, the scalability and adaptability of existing predictive maintenance models to different machine types are limited, necessitating separate model development for each system [15], [16]. Therefore, although current predictive maintenance approaches developed under Industry 4.0 have achieved substantial progress, further research is required to address deficiencies related to data quality, model explainability, implementation cost, and scalability.

To overcome these limitations identified in the literature, this study proposes a two-stage artificial intelligence-based predictive maintenance model. The proposed method consists of a layered structure that aims to determine both the occurrence of machine failures and the classification of failure types under faulty conditions. In the first stage of the model, feature selection based on the Spearman correlation coefficient is applied to identify the most significant variables in the dataset, thereby optimizing the input dimensionality of the system. This approach reduces computational burden while minimizing the impact of noisy features. In the second stage, data imbalance is addressed using a systematic sampling method, ensuring equal contribution of both classes to the model. This strategy provides an effective solution to the data imbalance problem frequently encountered in the literature.

In the first stage, various machine learning algorithms (SVM, Ensemble Trees, Artificial Neural Networks, kNN, etc.) are compared for fault detection, and the model yielding the highest accuracy, sensitivity, and AUC values is identified. In the second stage, using only faulty samples, a partial least squares (PLS)-based model is designed to classify fault types (TWF, HDF, PWF, OSF). The PLS approach is preferred due to its ability to produce stable results in datasets with multicollinearity and to maintain robust performance in data structures with limited sample sizes. Through this two-stage architecture, both general fault detection and specific fault type classification are achieved within a single framework, thereby overcoming the performance and generalization limitations of conventional single-stage models.

In conclusion, the proposed two-stage predictive maintenance model presents a holistic approach that addresses key challenges identified in the literature, including data imbalance, overfitting, explainability, and scalability. The method aims to enhance the reliability of real-time fault prediction and decision-support processes in Industry 4.0 environments and can be considered a predictive maintenance framework adaptable to different production systems in future implementations.

II. MATERIAL AND METHODS

The overall workflow of the proposed two-stage predictive maintenance architecture is summarized through the flow diagram presented in Fig. 1. First, feature selection

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25dec776

based on the Spearman correlation coefficient is applied to the raw data to generate new feature subsets. Subsequently, a data balancing process is performed using the systematic sampling method in order to address the imbalanced distribution among fault classes.

The resulting balanced dataset is then fed into the first stage of the system, referred to as the Machine Fault Artificial Intelligence Model, where the fault status of the machine is determined through binary classification. If a fault is detected by the model, the process proceeds to the second stage, in which the Fault Type Artificial Intelligence Model is activated to identify whether the detected fault belongs to one of the following classes: TWF (Tool Wear Failure), HDF (Heat Dissipation Failure), PWF (Power Failure), or OSF (Overstrain Failure).

- *Data Collection*

The data used in this study were obtained from the publicly available AI4I 2020 Predictive Maintenance dataset, which is widely used in the field of predictive maintenance, without collecting any new measurements or sensor data within the scope of the research [17], [18]. The dataset was generated based on a synthetic yet statistically consistent structure designed to represent operational behaviors observed in real industrial production processes (

Table). The dataset consists of a total of 10,000 observations and includes key features characterizing machine operations, such as product quality type, air and process temperatures, rotational speed, torque, and tool wear. Fault labels were assigned based on five different failure modes representing typical industrial maintenance scenarios. The use of a benchmark dataset standardized the data acquisition phase of the study and provided a reliable and reproducible experimental environment for evaluating the proposed two-stage artificial intelligence models.

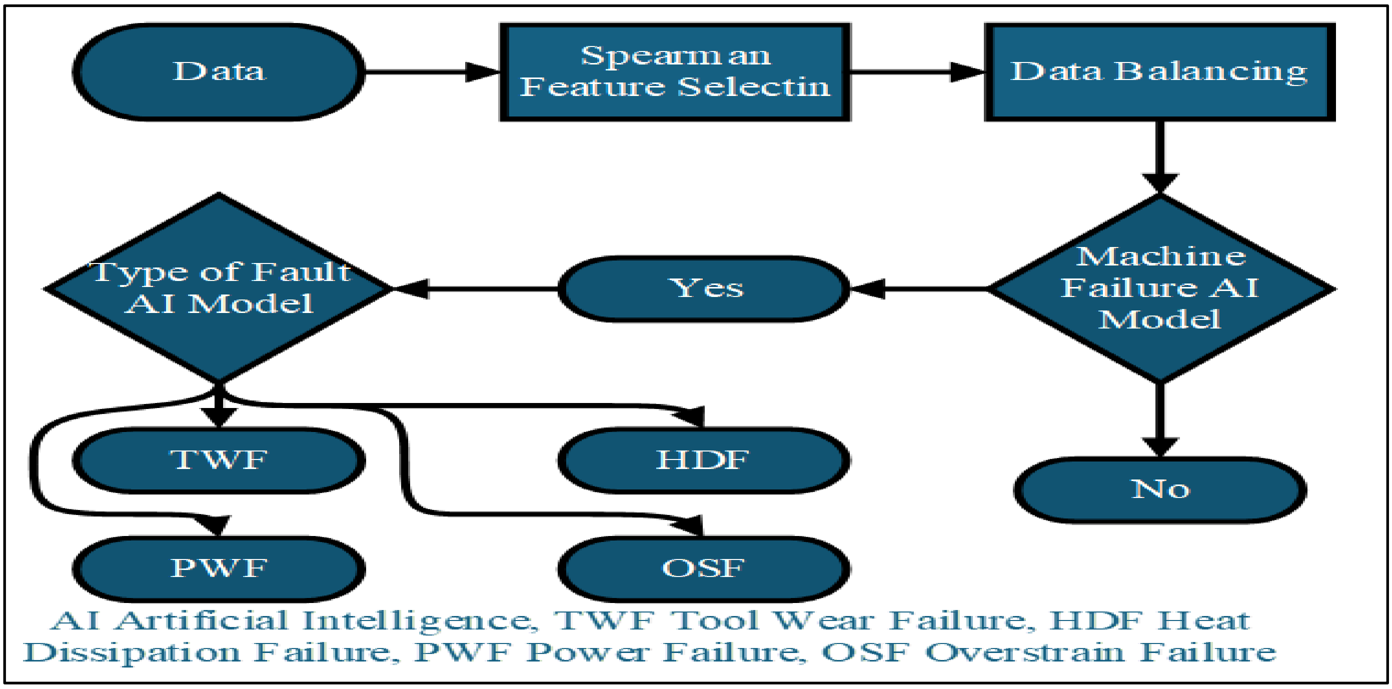

Fig 1 Overall Workflow of the Two-Stage Artificial Intelligence-Based Predictive Maintenance Model

Table 1 Summary of the AI4I 2020 Predictive Maintenance Dataset

Feature	**Description**
Dataset Name	AI4I 2020 Predictive Maintenance Dataset
Total Number of Observations	10,000
Total Number of Variables	14
Data Type	Numerical + Categorical
Dataset Source	UCI Machine Learning Repository
Data Generation Method	Synthetic data representing real industrial processes
Purpose	Predictive maintenance, fault detection, and fault type classification
Input Features	Air temperature, process temperature, rotational speed, torque, tool wear time, product type
Label (Binary)	Machine Failure (Fault: Yes/No)
Label (Multi-Class)	TWF, HDF, PWF, OSF, RNF (As RNF denotes rare, highly imbalanced failures with no consistent patterns, this class was excluded from fault-type specific modeling.)
Class Imbalance	Significant imbalance among fault classes
Application Areas	Fault detection, predictive maintenance, two-stage classification models

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25dec776

Table summarizes the updated structure of the independent variables used in this study and the processing scheme applied to the categorical and numerical features in the dataset. First, since the Type variable representing the product type consists of three categories (L, M, H), it was transformed into three separate binary indicators-H (1/0), L (1/0), and M (1/0)-to enable direct processing by the models. In this way, each category was defined as an independent input feature, thereby preventing potential information loss. The numerical features, namely Air Temperature, Process Temperature, Rotational Speed, Torque, and Tool Wear, contain fundamental physical parameters that represent the operational condition of the machine.

Table 2 Features Used in the AI4I 2020 Dataset

No	Feature Name	Type	Description / Content
1	Type	Categorical	Product type (L, M, H classes)
2	H (1/0)	Numerical	H (Present / Absent)
3	L (1/0)	Numerical	L (Present / Absent)
4	M (1/0)	Numerical	M (Present / Absent)
5	Air Temperature (K)	Numerical	Ambient air temperature (in Kelvin)
6	Process Temperature (K)	Numerical	Process temperature (in Kelvin)
7	Rotational Speed (rpm)	Numerical	Rotational speed of the machine
8	Torque (Nm)	Numerical	Torque generated by the machine
9	Tool Wear (min)	Numerical	Tool wear time (in minutes)

- *Spearman Correlation-Based Feature Selection Algorithm*

In this study, the Spearman correlation coefficient was employed for feature selection. The Spearman coefficient was preferred due to its ability to measure nonlinear monotonic relationships between variables while preserving the statistical interpretability of each feature. In the equation used to compute the correlation coefficient, n denotes the number of data samples, and d_i represents the rank difference for each observation (Equation 1, [19]).

$$r_s = 1 - 6\sum_{i=1}^{n} \frac{d_i^{\,2}}{n \times (n^2 - 1)} \tag{1}$$

In the study, Spearman correlation coefficients were computed for a total of nine features included in the dataset, and the correlation values were ranked in descending order. Based on this ranking, an initial dataset containing only the single feature with the highest correlation was first constructed. Subsequently, a second dataset including the top two features was generated, and the same procedure was followed by progressively adding the top three, top four, and so on. In this manner, nine different datasets, ranging from one to nine features, were obtained by gradually incorporating features selected according to their correlation values.

- *Data Balancing Using Systematic Sampling*

In this study, the data preprocessing pipeline was structured to apply feature selection prior to data balancing in order to minimize information loss. Performing feature selection at the initial stage prevents potential sample loss and distortions in statistical representation that may arise if the balancing procedure is applied before feature extraction or selection [20].

In this study, a data balancing procedure was applied to address the pronounced class imbalance of the Machine Failure label predicted in the first stage. In the original data distribution, the dataset contained 339 failure samples (Machine Failure = 1) and 9,661 non-failure samples (Machine Failure = 0). To enhance the representational power of the failure class during the learning process, a systematic sampling approach was employed. Within this framework, all 339 failure samples were retained, while 339 representative non-failure samples were selected from the 9,661 non-failure instances by applying systematic sampling at intervals of k = 9,661 / 339. As a result, a balanced dataset consisting of two classes with equal sample sizes was obtained. This balancing procedure was applied independently to each of the nine datasets constructed based on Spearman correlation ranking, ranging from one to nine selected features.

- *Machine Learning Algorithms*

In this study, a comprehensive machine learning analysis was conducted in the first layer of the two-stage predictive maintenance approach to determine the machine fault status across nine different feature sets. For each feature set, a total of 32 classical machine learning algorithms were individually trained and their performances were comparatively evaluated. These algorithms included Artificial Neural Networks (ANNs), Support Vector Machines (SVMs), Decision Trees, Quadratic Discriminant Analysis (QDA), Ensemble Tree methods (bagging and boosting), and k-Nearest Neighbors (kNN), among others [20], [21].

In the second layer, a model was developed to classify the type of failure using only samples corresponding to faulty conditions (n = 339). At this stage, a classification approach based on Partial Least Squares (PLS) regression was applied for each feature set [22]. The PLS method was preferred due to its ability to produce stable results in data structures characterized by multicollinearity and limited sample sizes. PLS-based models were trained for the nine different feature sets to predict fault types, namely TWF, HDF, PWF, and OSF. Consequently, in the second layer of the system, dimensionality reduction and classification were integrated within a single methodological framework.

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25dec776

➢ *Performance Evaluation Criteria*

To evaluate the performance of the machine learning models, five primary metrics were employed in this study: Accuracy, Sensitivity, Specificity, F-Measure, and Area Under the Curve (AUC) [20], [23]. In both layers, model evaluation for all nine datasets was conducted using a data splitting strategy of 75% for training and 25% for testing. These metrics were computed based on the elements of the confusion matrix, namely True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN).

Accuracy represents the proportion of correctly predicted samples among all samples and is one of the most commonly used performance metrics in classification problems (Equation **Error! Reference source not found.**). Although accuracy reflects the overall correctness of a model, it can be misleading when applied to imbalanced datasets. Therefore, in this study, accuracy was considered only as a complementary metric, and model performance was interpreted in conjunction with other evaluation metrics.

$$Accuracy=\frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

Sensitivity, also referred to as Recall or the True Positive Rate (TPR), represents the proportion of actual positive samples that are correctly predicted as positive by the model (Equation **Error! Reference source not found.**). This metric is particularly important in critical applications such as fault detection, where minimizing missed positive instances is essential. A sensitivity value close to 1 indicates that the model is highly effective in correctly identifying fault conditions.

$$Sensitivity=TP/TP+FN \quad (3)$$

Specificity, also known as the True Negative Rate (TNR), represents the proportion of actual negative samples that are correctly classified as negative (Equation **Error! Reference source not found.**). In predictive maintenance applications, false alarms in the absence of actual faults may lead to increased maintenance costs; therefore, specificity constitutes an important evaluation metric.

$$Specificity=TN/TN+FP \quad (4)$$

In this study, the F-Measure differs from the conventional F1-score, as it is computed using Specificity instead of Precision (Equation **Error! Reference source not found.**). The metric represents the harmonic mean of Sensitivity and Specificity. Since this measure simultaneously evaluates the ability to correctly detect faults (Sensitivity) and to avoid unnecessary fault alarms (Specificity), it provides a particularly meaningful performance indicator for predictive maintenance applications.

$$FM = 2\times\frac{Sensitivity\times Specificity}{Sensitivity+Specificity} \quad (5)$$

Area Under the Curve (AUC) represents the area under the Receiver Operating Characteristic (ROC) curve and measures the model's ability to discriminate between positive and negative classes. The ROC curve is obtained by plotting the True Positive Rate (TPR, Sensitivity) against the False Positive Rate (FPR) at different threshold levels (Equation **Error! Reference source not found.**).

The AUC value represents the area under the ROC curve and takes values in the range of 0.5 to 1.0. In this context, a value of 1.0 indicates that the model achieves perfect discrimination performance, while values in the range of 0.9-1.0 correspond to very high discriminative power. An AUC range of 0.8-0.9 is considered to reflect good performance, whereas values between 0.7 and 0.8 indicate an acceptable level of discrimination capability. An AUC value of 0.5 denotes that the model performs equivalently to random guessing. Due to these properties, AUC was employed in this study as a robust and complementary metric that provides a general assessment of the model's ability to distinguish between classes.

$$FPR=FP/FP+TN \quad (6)$$

III. RESULTS

In this study, a two-stage artificial intelligence-based model was proposed to enhance the effectiveness of predictive maintenance in industrial production processes. In the first stage of the model, the machine fault status was predicted using machine learning algorithms applied to feature sets selected based on the Spearman correlation coefficient. In the second stage, fault type classification was performed for the faulty conditions. During the implementation process, the AI4I 2020 Predictive Maintenance dataset was utilized, class imbalance was addressed using a systematic sampling method, and the performance of the two-stage model was evaluated across nine different feature sets. The obtained results comparatively demonstrated the impact of different feature sets and algorithm combinations on overall system performance.

When the first-stage model results were examined, it was observed that the overall classification performance generally improved with an increasing number of features (Table 3). For datasets containing a limited number of features (NF = 1-2), the classification accuracy was approximately 80%, whereas a notable performance improvement was observed for Ensemble Trees and SVM-based models, particularly in datasets including three or more features. From the fifth feature set onward, the Ensemble Trees algorithm achieved the highest classification performance, with an accuracy of 91.7%, sensitivity of 96.4%, and an AUC value of 0.97. Similarly, for the dataset containing nine features, both SVM and Ensemble Trees models exhibited balanced prediction performance, achieving 92.9% accuracy and AUC values ranging between 0.96 and 0.98.

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25dec776

Table 3 Performance Results of First-Stage Fault Detection Models

NF	Model	Acc	Sen	Spe	FM	AUC
1	Neural Network	79.90	80.00	79.80	79.90	0.86
1	Quadratic Discriminant	78.70	74.10	83.30	78.43	0.89
2	SVM	81.70	82.10	81.20	81.65	0.86
2	Tree	81.10	79.80	82.40	81.08	0.80
3	Ensemble Trees	87.00	91.80	82.10	86.68	0.92
3	SVM	85.80	85.90	85.70	85.80	0.92
4	SVM	88.80	94.00	83.50	88.44	0.94
4	kNN	88.80	90.50	87.10	88.77	0.93
5	Ensemble Trees	91.70	96.40	87.10	91.51	0.97
5	Neural Network	91.10	90.50	91.80	91.15	0.96
6	Ensemble Trees	91.70	98.80	84.70	91.21	0.96
6	Tree	89.90	94.00	85.90	89.77	0.91
7	Ensemble Trees	91.70	94.00	89.40	91.64	0.97
7	SVM	91.10	95.20	87.10	90.97	0.96
8	Tree	89.90	92.90	87.10	89.91	0.90
8	SVM	89.90	90.50	89.40	89.95	0.94
9	*SVM*	*92.90*	*92.90*	*92.90*	*92.90*	*0.98*
9	Ensemble Trees	92.90	92.90	92.90	92.90	0.96
Acc Accuray, Sen Sensitivity, Spe Specificity, FM F-Measure, AUC Area Under the Curve, NF Number of Feature, SVM Support Vector Machine						

The ROC curve shown in Fig 1 corresponds to the SVM model with a medium Gaussian kernel trained on the dataset containing nine features. The curve being positioned close to the upper-left corner and the high area under the curve (AUC = 0.9756) indicate that the model exhibits a strong ability to discriminate between faulty and non-faulty conditions. The model's operating point achieves a balanced performance by maintaining a low false positive rate (below 7%) while achieving a high true positive rate (approximately 93%).

The second-stage model results were obtained for the classification of the four main fault types (TWF, HDF, PWF, and OSF) using different feature sets (Table 4). Although the models achieved relatively high accuracy values (approximately 91.7%) for datasets containing a single feature (NF = 1), the specificity values being close to zero indicated a tendency toward overfitting. As the number of features increased, a notable improvement in overall model performance was observed. In particular, for the fifth and seventh feature sets, the average classification accuracy ranged between 93.75% and 95.83%, while the corresponding AUC values reached levels between 0.88 and 0.93. For these datasets, classification accuracies exceeding 97% were achieved for the PWF and OSF fault classes. Models constructed with six and seven features produced balanced results in terms of both sensitivity and specificity, thereby demonstrating the highest overall classification performance. In contrast, a decline in accuracy was observed for datasets containing eight and nine features, indicating that an excessive number of features introduced noise effects into the model and consequently reduced its generalization capability. Based on these findings, it was determined that optimal performance in fault type prediction was achieved using PLS-based models with six to seven features.

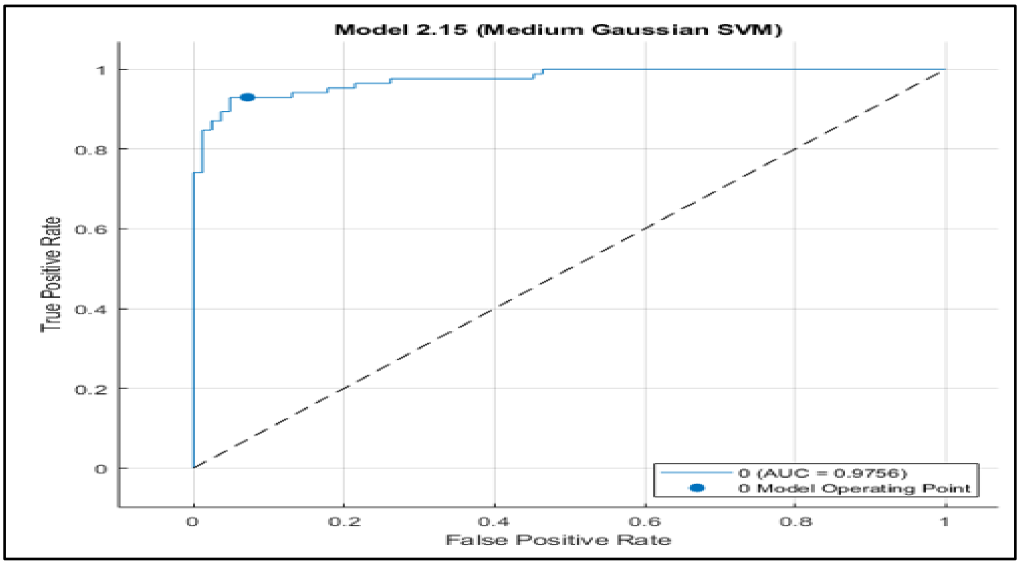

Fig 2 ROC Curve of the Nine-Feature SVM Model

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25dec776

IV. DISCUSSION

The two-stage artificial intelligence-based predictive maintenance model developed in this study demonstrates the significance of data-driven maintenance approaches within the scope of Industry 4.0, as evidenced by the high accuracy rates achieved in both fault detection and fault type classification tasks in industrial systems. The obtained findings are consistent with previous studies emphasizing that predictive maintenance is an effective tool for ensuring production continuity and enhancing operational efficiency [1], [2].

In particular, the results indicating that early fault prediction reduces unplanned downtime, lowers production costs, and improves system reliability are in line with the conclusions of Bahador et al. regarding production reliability and maintenance optimization [1]. Furthermore, owing to the two-stage architecture of the proposed model, both fault detection and fault type classification are performed within a single integrated framework. This represents a notable contribution, as these two processes are typically addressed independently in the existing literature. In this respect, the present findings complement the observations reported by Shukla et al. [2], who highlighted the lack of model integration in predictive maintenance systems [2].

Table 4 Second-Stage Fault Type Classification Results

NF	Label	Acc	Sen	Spe	FM	AUC
1	TWF	50	1.00	NaN	NaN	0.50
1	HDF	50	1.00	NaN	NaN	0.50
1	PWF	50	1.00	NaN	NaN	0.50
1	OSF	66.67	0.98	0.00	0.00	0.49
Mean		*54.17*	*1.00*	*0.00*	*0.00*	*0.50*
2	TWF	86.90	1.00	0.08	0.15	0.54
2	HDF	67.86	0.92	0.26	0.40	0.59
2	PWF	100.00	1.00	1.00	1.00	1.00
2	OSF	63.10	0.93	0.00	0.00	0.46
Mean		*79.46*	*0.96*	*0.34*	*0.39*	*0.65*
3	TWF	89.29	1.00	0.25	0.40	0.63
3	HDF	76.19	0.91	0.52	0.66	0.71
3	PWF	98.81	1.00	0.95	0.98	0.98
3	OSF	95.24	0.93	1.00	0.96	0.96
Mean		*89.88*	*0.96*	*0.68*	*0.75*	*0.82*
4	TWF	89.29	0.99	0.33	0.50	0.66
4	HDF	86.90	0.92	0.77	0.84	0.85
4	PWF	96.43	1.00	0.86	0.93	0.93
4	OSF	95.24	0.95	0.96	0.96	0.96
Mean		*91.96*	*0.96*	*0.73*	*0.81*	*0.85*
5	TWF	90.48	0.99	0.42	0.59	0.70
5	HDF	95.24	0.94	0.97	0.96	0.96
5	PWF	96.43	1.00	0.86	0.93	0.93
5	OSF	92.86	0.93	0.93	0.93	0.93
Mean		*93.75*	*0.96*	*0.79*	*0.85*	*0.88*
6	TWF	92.86	0.99	0.58	0.73	0.78
6	HDF	94.05	0.94	0.94	0.94	0.94
6	PWF	97.62	1.00	0.91	0.95	0.95
6	OSF	98.81	0.98	1.00	0.99	0.99
Mean		*95.83*	*0.98*	*0.86*	*0.90*	*0.92*
7	*TWF*	*94.05*	*0.99*	*0.67*	*0.80*	*0.83*
7	*HDF*	*94.05*	*0.94*	*0.94*	*0.94*	*0.94*
7	*PWF*	*97.62*	*1.00*	*0.91*	*0.95*	*0.95*
7	*OSF*	*97.62*	*0.96*	*1.00*	*0.98*	*0.98*
Mean		*95.83*	*0.97*	*0.88*	*0.92*	*0.93*
8	TWF	15.49	0.03	1.00	0.06	0.57
8	HDF	90.48	0.91	0.90	0.90	0.90
8	PWF	97.62	1.00	0.91	0.95	0.95
8	OSF	37.68	0.09	1.00	0.16	0.53
Mean		*60.32*	*0.51*	*0.95*	*0.52*	*0.74*
9	TWF	86.90	0.92	0.58	0.71	0.75
9	HDF	91.67	0.91	0.94	0.92	0.92
9	PWF	42.50	0.26	1.00	0.41	0.54

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25dec776

9	OSF	89.29	0.88	0.93	0.90	0.90
Mean		*77.59*	*0.74*	*0.86*	*0.74*	*0.78*
Acc Accuray, Sen Sensitivity, Spe Specificity, FM F-Measure, AUC Area Under the Curve, NF Number of Feature, TWF Tool Wear Failure, HDF Heat Dissipation Failure, PWF Power Failure, OSF Overstrain Failure						

Many machine learning-based predictive maintenance models reported in the literature, despite achieving high accuracy, suffer from challenges such as imbalanced data structures, limited generalization capability, and high computational burden [9], [10], [13]. The Spearman correlation-based feature selection and systematic sampling strategy applied in this study address these shortcomings in an effective manner. While Sinha and Das emphasized that data imbalance significantly degrades model performance, the direct integration of data balancing into the training process in this study reduced information loss during learning and resulted in a more balanced predictive performance [13]. Furthermore, Fernández Salazar et al. [10] reported that single-stage machine learning models within Industry 4.0 environments exhibit limited explainability when applied to complex industrial processes [10]. In contrast, the PLS-based modeling approach employed in the second stage of the proposed framework enabled more stable results in data structures characterized by multicollinearity, while simultaneously enhancing model interpretability. This finding provides an alternative solution to the model generalization problem in IoT-based predictive maintenance systems, as also discussed by Madasamy et al. [14].

From an industrial application perspective, the modular structure and two-layer decision mechanism of the proposed model enhance the feasibility of implementing real-time predictive maintenance in production lines. As highlighted by Meriem et al. [15], the scalability of predictive maintenance (PdM) systems and their adaptability to different machine types remain significant challenges in Industry 4.0 environments [15]. The model proposed in this study offers a flexible solution that can be easily adapted to different production lines due to its retrainable structure, which can be configured according to the characteristics of the target dataset. Moreover, in line with architectures suitable for IoT and artificial intelligence integration reported by Caldana et al., the framework developed in this study was designed to operate in conjunction with sensor-based data acquisition systems deployed in industrial environments [16]. Consequently, the proposed model not only presents a theoretical approach at the academic level but also constitutes a practically applicable predictive maintenance solution for real-world industrial settings.

Nevertheless, the present study has certain limitations. The performance of the proposed model was evaluated using a synthetic dataset, namely the AI4I 2020 dataset; therefore, factors such as noise, missing data, and measurement errors encountered in real industrial environments may affect model performance [2]. In addition, as emphasized by Baroud et al., the limited explainability of artificial intelligence-based predictive maintenance systems may restrict the level of trust that human operators place in model decisions [9]. In future work, the integration of Explainable Artificial Intelligence (XAI) techniques is planned to enhance the interpretability of the proposed model, along with validation using diverse real-world industrial datasets. These efforts are expected to strengthen both the generalization capability of the developed system and its level of integration within the Industry 4.0 ecosystem.

V. CONCLUSION

In this study, a two-stage artificial intelligence-based predictive maintenance model was developed for the early prediction of machine failures and the classification of fault types within the scope of Industry 4.0 applications. The proposed approach enhanced the stability and generalizability of the learning process by addressing data imbalance through Spearman correlation-based feature selection and systematic sampling. In the first-stage fault detection analyses, the SVM and Ensemble Trees algorithms achieved high accuracy levels (above 92%) and AUC values of approximately 0.98. In the second stage, the PLS-based classification approach demonstrated classification accuracies exceeding 95%, particularly for datasets containing six and seven features, in identifying fault types.

The obtained results indicate that limitations frequently reported in the existing literature-such as data quality issues, class imbalance, and model complexity-can be substantially mitigated through appropriate data preprocessing and multi-layer modeling strategies. The modular and scalable structure of the proposed model enables straightforward adaptation to different production lines and machine types. In this respect, the study proposes a holistic solution that enhances the reliability of predictive maintenance and improves the accuracy of decision-support processes in industrial systems.

In conclusion, the developed two-stage artificial intelligence-based predictive maintenance model offers significant potential for the optimization of maintenance planning, cost reduction, and the ensuring of operational continuity in production systems. In future studies, validating the model using real industrial data and supporting it with Explainable Artificial Intelligence (XAI) techniques are expected to further enhance both the practical applicability and the reliability of the proposed system.

ACKNOWLEDGMENT

This scientific publication results from the R&D activities of Kontrolmatik Technologies and was prepared within the “Scientific Publication Preparation Training for R&D Centers” provided to the Design Center by Assoc. Prof. Dr. Muhammed Kürşad Uçar. All rights and responsibilities regarding the content rest with Kontrolmatik Technologies.

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25dec776

REFERENCES

- [1]. A. Bahador et al., "Condition Monitoring for Predictive Maintenance of Machines and Processes in ARTC Model Factory," Intelligent Systems Reference Library, vol. 202, pp. 113–141, 2021.
- [2]. K. Shukla, S. Nefti-Meziani, and S. Davis, "A heuristic approach on predictive maintenance techniques: Limitations and scope," Advances in Mechanical Engineering, vol. 14, no. 6, Jun. 2022.
- [3]. S. Maataoui, G. Bencheikh, and G. Bencheikh, "Predictive Maintenance in the Industrial Sector: A CRISP-DM Approach for Developing Accurate Machine Failure Prediction Models," 2023 5th International Conference on Advances in Computational Tools for Engineering Applications, ACTEA 2023, pp. 223–227, 2023.
- [4]. T. Frost, N. Moser, J. Nöcker, and L. Zhukov, "Getting Value from Predictive Maintenance," Mechanisms and Machine Science, vol. 105, pp. 178–185, 2021.
- [5]. N. Es-sakali, M. Cherkaoui, M. O. Mghazli, and Z. Naimi, "Review of predictive maintenance algorithms applied to HVAC systems," Energy Reports, vol. 8, pp. 1003–1012, Nov. 2022.
- [6]. H. M. Hashemian and W. C. Bean, "State-of-the-art predictive maintenance techniques," IEEE Trans Instrum Meas, vol. 60, no. 10, pp. 3480–3492, Oct. 2011.
- [7]. P. Lu, H. Liu, C. Serratella, and X. Wang, "Assessment of Data-Driven, Machine Learning Techniques for Machinery Prognostics of Offshore Assets," Proceedings of the Annual Offshore Technology Conference, vol. 5, pp. 3651–3672, May 2017.
- [8]. F. E. Bezerra et al., "Impacts of Feature Selection on Predicting Machine Failures by Machine Learning Algorithms," Applied Sciences 2024, Vol. 14, Page 3337, vol. 14, no. 8, p. 3337, Apr. 2024.
- [9]. S. Y. Baroud, N. A. Yahaya, and A. M. Elzamly, "Cutting-Edge AI Approaches with MAS for PdM in Industry 4.0: Challenges and Future Directions," Journal of Applied Data Sciences, vol. 5, no. 2, pp. 455–473, May 2024.
- [10]. J. Karina Fernández Salazar et al., "Machine Learning for Predictive Maintenance in Industry 4.0: A Systematic Review of Algorithms and Implementation Cases," in Proceedings of the LACCEI international Multi-conference for Engineering, Education and Technology, Mexico City,: Latin American and Caribbean Consortium of Engineering Institutions, 2025.
- [11]. S. Xie, "Advancing Predictive Maintenance Research Trends: Using Artificial Intelligence for Enhanced Industrial Reliability," Proceedings - 2024 IEEE International Conference on Future Machine Learning and Data Science, FMLDS 2024, pp. 283–288, 2024.
- [12]. A. Belounnasa, F. Brissaud, and P. Faly Ba, "Predictive Maintenance of Natural Gas Regulators by Forecasting Output Pressure with Artificial Intelligence Algorithms," in Proc. of the 31st European Safety and Reliability Conference (ESREL 2021), B. Castanier, M. Cepin, D. Bigaud, and C. Berenguer, Eds., Singapore: Research Publishing, Sep. 2021, pp. 2150–2150.
- [13]. A. Sinha and D. Das, "Data-Driven Techniques for Fault Diagnosis and Predictive Maintenance," in Applied Artificial Intelligence and Machine Learning Techniques for Engineering Applications, CRC Press, 2025, pp. 30–49.
- [14]. S. Madasamy, B. P. Shankar, R. K. Yadav, and K. P. Jayalakshmi, "A Machine Learning Approach in Predictive Maintenance in the IoT Enabled Industry 4.0," Proceedings of the 4th International Conference on Smart Electronics and Communication, ICOSEC 2023, pp. 418–423, 2023.
- [15]. H. Meriem, H. Nora, and O. Samir, "Predictive Maintenance for Smart Industrial Systems: A Roadmap," Procedia Comput Sci, vol. 220, pp. 645–650, Jan. 2023.
- [16]. V. M. Caldana, F. D. G. da Silva, R. A. de Oliveira, and J. F. Borin, "Internet of Things and Artificial Intelligence applied to predictive maintenance in Industry 4.0: A systematic literature review," in Proceedings of the International Conference on Industrial Engineering and Operations Management, Michigan, USA: IEOM Society International, Apr. 2021, pp. 1387–1398.
- [17]. S. Matzka, "Explainable Artificial Intelligence for Predictive Maintenance Applications," Proceedings - 2020 3rd International Conference on Artificial Intelligence for Industries, AI4I 2020, pp. 69–74, Sep. 2020.
- [18]. "AI4I 2020 Predictive Maintenance Dataset - UC Irvine Machine Learning Repository." doi: 10.24432/C5HS5C.
- [19]. R. Alpar, Applied Statistic and Validation - Reliability. Detay Publishing, 2010. [Online]. Available: https://books.google.com.tr/books/about/Uygulamalı_istatistik_ve_gecerlik_guv.html?id=ITk1MwEACAAJ&pgis=1
- [20]. M. K. Uçar, M. Nour, H. Sindi, and K. Polat, "The Effect of Training and Testing Process on Machine Learning in Biomedical Datasets," Math Probl Eng, vol. 2020, pp. 1–17, 2020.
- [21]. M. Fernández-Delgado, E. Cernadas, S. Barro, D. Amorim, and A. Fernández-Delgado, "Do we Need Hundreds of Classifiers to Solve Real World Classification Problems?," Journal of Machine Learning Research, vol. 15, no. 90, pp. 3133–3181, 2014, Accessed: Dec. 11, 2025.
- [22]. H. Herv´ and H. Abdi, "Partial least squares regression and projection on latent structure regression (PLS Regression)," Wiley Interdiscip Rev Comput Stat, vol. 2, no. 1, pp. 97–106, Sep. 2010.
- [23]. T. Fawcett, "An introduction to ROC analysis," Pattern Recognit Lett, vol. 27, no. 8, pp. 861–874, Jun. 2006.