PORT SAID ENGINEERING RESEARCH JOURNAL

Faculty of Engineering - Port Said University
Volume (28) No. 3 pp: 103:120

Artificial Neural Network-Based Vibration Data Analysis to Predict Failures of Powder Machines

*Khaled W. Salem [1, *], Ebtisam F. Abdel-Gwad [2] and Hanan K. Kouta [3]*

[1] *Production Engineering and Mechanical Design Department, Faculty of Engineering, Port Said University, Port Said, Egypt, email:* khaled.salem@eng.psu.edu.eg

[2] *Production Engineering and Mechanical Design Department, Faculty of Engineering, Port Said University, Port Said, Egypt, email:* Ebtisam1953@yahoo.com

[3] *Production Engineering and Mechanical Design Department, Faculty of Engineering, Port Said University, Port Said, Egypt, email:* Hanan.Kamel@eng.psu.edu.eg

Corresponding author,* *DOI:* ***10.21608/PSERJ.2024.266793.1317*

Received 31-1-2024
Revised 13-3-2024
Accepted 18-3-2024

ABSTRACT

This study proposes a novel predictive maintenance methodology for a forced blower employed in PVC powder manufacturing at TCI Sanmar, leveraging the power of machine learning. The Multilayer Perceptron (MLP) algorithm is identified as the most effective approach, demonstrating exceptional capability in anticipating machine failures. Following a comprehensive optimization process using GridSearchCV, the optimal MLP configuration achieves impressive performance metrics, including an accuracy of 92.5%, a recall of 95.5%, a precision of 95%, an F-score of 92.7%, and a Matthews Correlation Coefficient (MCC) of 0.85. Furthermore, the chosen MLP configuration consistently delivers high Area Under the ROC Curve (AUC) values of 0.961, surpassing other methods employed within the same industrial setting. This research not only addresses the immediate challenge but also offers valuable insights that can be applied and generalized to broader predictive maintenance strategies. The efficacy of the MLP algorithm underscores its strategic significance in this context, highlighting its potential to revolutionize intelligent and interconnected manufacturing processes. The proposed methodology demonstrably optimizes equipment reliability, minimizes downtime, and fosters proactive maintenance interventions, thereby generating valuable insights that contribute to enhanced industrial efficiency.

Keywords: Maintenance Strategy Optimization, Advanced Manufacturing Integration, Machine Learning Applications, Equipment Health Monitoring, Intelligent Manufacturing Insights

NOMENCLATURE

y_k	The output of the MLP for class k.
$b_j^{(1)}$ and $b_k^{(2)}$	The biases associated with the concealed layer and the result layer, correspondingly.
$w_{ji}^{(1)}$	the weights responsible for connecting the input layer to the hidden layer for a specific neuron j.
$w_{jk}^{(2)}$	The weights that establish the connections between the hidden layer and the output layer for a specific class k.
x_i	The input features.
FN	False Negative
FP	False positive
FPR	False Positive Rate
TF	True Negative
TP	True positive
TPR	True Positive Rate
d	The number of input features.
h	The number of neurons in the hidden layer.
σ	The activation function applied to the summation of weighted inputs for each neuron.

1 INTRODUCTION

The emergence of Industry 4.0 (I4.0) and the subsequent shift towards Industry 5.0 (I5.0) have profoundly transformed industrial maintenance practices. This transformation highlights the growing emphasis on condition monitoring methodologies designed to anticipate faults, minimize operational downtime, and prevent critical failures. Within these technological revolutions, the operational integrity of rolling bearings is paramount, as these crucial components directly influence overall equipment performance and reliability [1], [2].

Decades of research have focused on enhancing the resilience of critical machinery components. The integration of digitization and automation within Industry 4.0, and the subsequent human-centric approach of I5.0, have revitalized maintenance strategies. Intelligent systems leveraging machine learning (ML), artificial intelligence (AI), and evolving deep learning methodologies are now central to strengthening component durability [3], [4].

Maintenance strategies have evolved alongside industrial revolutions, transitioning from reactive paradigms like Run-to-Failure (R2F) to more proactive approaches. The shift from preventive maintenance (PvM), with its preplanned interventions, to condition-based maintenance (CBM), where interventions are triggered based on measured degradation indicators, reflects the changing nature of industrial paradigms [5], [6].

A previous exploration examined the efficacy of established algorithms, such as Support Vector Machines (SVM), Logistic Regression (LR), Random Forest (RF), K-Nearest Neighbors (KNN), and XGBoost, within the realm of predictive maintenance [7]. This study builds upon existing insights and delves deeper into the application of the Multilayer Perceptron (MLP) algorithm within the context of Industry 5.0's evolving industrial landscape.

Within the spectrum of maintenance methodologies, this study identifies the MLP algorithm as a compelling approach that embodies the digital advancements of Industry 4.0 and the human-centric approach of Industry 5.0. Renowned for its multilayered neural networks, the MLP algorithm has the potential to exceed the predictive capabilities of traditional algorithms, offering a powerful alignment with the principles of Industry 5.0 [8], [9].

The present study undertakes an extended exploration of predictive maintenance strategies, with a focus on the MLP algorithm within an established case study framework. This in-depth inquiry aims to evaluate MLP's efficacy in predicting machinery failures, striving to set new benchmarks in predictive maintenance methodologies.

2 LITERATURE REVIEW

This research draws upon a comprehensive literature review to evaluate the accuracy of a predictive maintenance framework tailored for a forced blower within an Egyptian industrial plant. Building on insights from five machine learning (ML) algorithms, this study evaluates the use of ball bearing vibration data to preemptively identify potential machinery failures. Vibration-based maintenance has long been recognized as a highly effective technique for detecting incipient bearing faults, as these faults exhibit unique, identifiable frequency patterns [7]. As demonstrated by the extensive literature review, the emergence of these fault patterns is closely linked to variations in both the machine's weight and operating speed [10].

Márquez et al. [11] addressed asset performance monitoring in capital-intensive industries with long-lasting assets. Their novel approach combined Artificial Neural Networks (ANNs) with Data Mining (DM), specifically using Association Rule (AR) mining software for big data processing. This innovative sequence offers researchers new strategies for developing advanced predictive maintenance models capable of precise asset performance issue detection. The effectiveness of the proposed approach, demonstrated through complex monitoring scenarios, holds promise for improving Operational and Maintenance (O&M) practices. Orrù et al. [12] developed a machine learning framework for early fault detection in centrifugal pumps within the oil and gas sector. Their study utilized archival sensor data (temperature, pressure, and vibration) and employed Support Vector Machine (SVM) and Multilayer Perceptron (MLP) algorithms within the KNIME platform for data processing and training. The model achieved successful fault recognition and classification with high prediction accuracy, providing a reliable decision support system for maintenance personnel to prevent potential equipment failures. Capretz, et al. [13] introduced a Convolutional Neural Network (CNN) framework designed for predictive maintenance (PdM) by transforming 1-dimensional data into a 2-dimensional image-like representation. Their approach, evaluated using datasets from building fans at Western University, outperformed traditional ML techniques. Achieving high accuracy rates of up to 98% for one dataset and 95% for another, this innovative method demonstrates promising potential for asset health prediction, surpassing established techniques in predictive maintenance. Ye, et al. [14] proposed the Adaptive Kernel Sparse Network (AKSNet), a novel deep neural network model for extracting fault features from one-dimensional vibration signals. AKSNet addresses limitations of traditional CNNs by adaptively selecting kernels for multi-scale feature extraction. It employs diverse branches with varying kernels, integrates channel-wise attention for feature fusion, and uses dynamic spatial attention to focus on salient regions. A sparse regularization layer is incorporated to reduce noise and highlight important features. Experiments demonstrate that AKSNet surpasses shallow neural networks and other DNNs in bearing fault diagnosis, showcasing its superior feature extraction capabilities for multi-channel vibration signals. Chen et al. [15] presented an LSTM-based approach for predicting compressor failures in heavy-duty vehicles using aggregated sensor data from 2015 and 2016. Their goal was to optimize uptime through proactive maintenance scheduling. They framed the problem as a classification task and compared LSTM with the currently used Random Forest (RF) algorithm. While RF marginally outperformed LSTM in AUC score, LSTM exhibited greater temporal stability, consistently capturing the transition from healthy to faulty states and vice versa, even after repairs. This stability, crucial for reliable maintenance decisions, suggests that LSTM has the potential for superior practical applicability despite RF's slight performance edge in a single metric. Wu et al. [16] proposed a novel method for accurate failure prognosis in complex equipment maintenance by hybridizing backpropagation neural networks with a quantum-inspired multi-agent algorithm. Drawing inspiration from quantum computing theory and multi-agent systems, this approach optimizes neural network parameters using a multi-agent strategy to address limitations like slow convergence and getting trapped in local minima. Experimental validation, incorporating numerical approximations and real vibrational data analysis from the University of Cincinnati's laboratory, yielded promising results, suggesting the potential of this technique as a valuable tool for industrial machinery failure prediction.

In the study by Rajakumar et al. [17] they investigated methods to improve agricultural machine (AM) efficiency and fault detection using deep learning techniques. They utilized a deep convolutional neural network (DCNN) combined with a Levy flight optimization algorithm (LFOA) on smartphones to enhance fault detection accuracy, successfully identifying six specific faults in agricultural machines. This approach showcases a promising model for real-time health monitoring and fault prediction in AMs, offering potential applications in optimizing machine performance and reducing downtime.

3 METHODOLOGY

3.1 Algorithm Evaluation in the Initial Research - A Comprehensive Overview

The first research paper [18] established a robust foundation by evaluating five widely-used algorithms for maintenance prediction. Support Vector Machines (SVM), known for effective classification in high-dimensional spaces [19], Random Forest (RF), employing an ensemble of decision trees for improved accuracy [20], Logistic Regression (LR), a linear model suitable for binary classification tasks [21], k-Nearest Neighbors (KNN), utilizing proximity-based classification [22] and XGBoost, a gradient boosting algorithm emphasizing model performance enhancement, were meticulously examined [23]. Each algorithm's unique characteristics were scrutinized to provide a thorough understanding of their applicability and performance in the context of maintenance forecasting.

This study focuses solely on the Multilayer Perceptron (MLP), which is a powerful artificial neural network algorithm. We comprehensively compare MLP with previously examined algorithms, evaluating their strengths, weaknesses, and predictive capabilities within the realm of maintenance prediction. Particular emphasis is placed on optimizing MLP parameters to enhance its performance. This focused approach allows for a nuanced understanding of MLP's relative advantages and facilitates the recommendation of an optimal configuration for accurate maintenance forecasting.

3.2 Multilayer Perceptron Algorithm (MLP)

The Multilayer Perceptron (MLP) algorithm has emerged as a prominent tool in predictive maintenance due to its ability to learn complex patterns from historical data and effectively forecast machinery failures. This capability stems from MLP's robust neural network architecture, which allows it to identify intricate relationships within datasets, making it particularly well-suited for prognosticating machinery health and predicting potential failures.

Numerous studies have underscored the effectiveness of MLP in predictive maintenance contexts. In [24] demonstrated the superior predictive power of MLP in forecasting machinery failures, attributing its success to the algorithm's multilayered neural network architecture, which enables it to discern nuanced patterns in data indicative of impending faults. Similarly, in [12] showcased the adaptability of MLP in diverse industrial settings, emphasizing its ability to forecast equipment breakdowns with high accuracy rates compared to other conventional algorithms. Illustrating the MLP algorithm involves elucidating its architecture and functionality. As a type of artificial neural network, the Multilayer Perceptron (MLP) comprises an input layer, one or more hidden layers, and an output layer. These layers consist of interconnected nodes (neurons) that process information using activation functions to generate weighted outputs. During forward propagation, data is fed into the network, triggering computations across layers. This process enables the MLP to learn, adjust its parameters, and refine its predictive capabilities. as suggested by [25] and [26] offer comprehensive insights into the architecture and workings of MLP in predictive maintenance scenarios. These studies delve into the specifics of training the MLP model, adjusting its parameters through backpropagation, and fine-tuning the network to enhance its predictive capabilities for machinery health prognosis.

This amalgamation of real-world studies and theoretical underpinnings serves to highlight MLP's significance in predictive maintenance, both in its practical application

and its theoretical framework, paving the way for enhanced machinery prognostication and minimized downtime in industrial settings. An overview of the key components of a Multilayer Perceptron also shown it in Figure 1: Input Layer: Serving as the initial stratum of the MLP, this layer comprises neurons responsible for receiving input data. Each neuron within this layer corresponds to a distinct feature in the input dataset.
Hidden Layers: Positioned between the input and output layers of a neural network, there may be one or more hidden layers. These concealed layers are composed of neurons that manipulate the input data using weighted connections and activation functions. The number of hidden layers and the quantity of neurons within each layer serve as Hyperparameters that can be fine-tuned to improve the network's performance.
Output Layer: The ultimate layer of the MLP generates the network's output, determined by the specific task it is designed for. In binary classification tasks, it is typical to employ a single neuron with a sigmoid activation function. Conversely, for multi-class classification tasks, the count of output neurons aligns with the number of classes, and a softmax activation function is usually applied.

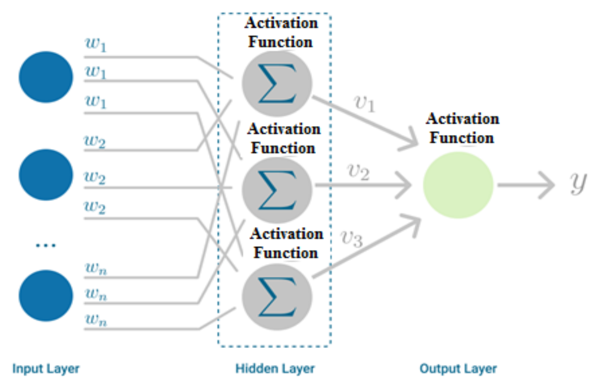


Figure 1: Multilayer Perceptron Network [27].

Every connection linking neurons in adjacent layers is assigned a specific weight, nd each neuron employs an activation function on the weighted sum of its inputs. The network's performance is improved by adapting these weights during training. Multilayer Perceptrons (MLPs) are trained through methods like backpropagation and gradient descent, which aim to minimize a loss function. The selection of the loss function depends on the task at hand, for instance, using mean squared error for regression and cross-entropy for classification tasks [28]. The mathematical Equation (1) of MLP [29],[30] can be written as:

$$y_k = \sigma\left(\sum_{j=1}^{h}\left(w_{jk}^{(2)}.\sigma\left(\sum_{i=1}^{d}(w_{ji}^{(1)}.x_i + b_j^{(1)})\right)\right) + b_k^{(2)}\right) \quad \textbf{(1)}$$

Where:

- y_k: The output of the MLP for class k.
- x_i: The input features.
- $w_{ji}^{(1)}$: the weights responsible for connecting the input layer to the hidden layer for a specific neuron j.
- $w_{jk}^{(2)}$: The weights responsible for establishing connections between the hidden layer and the output layer for a particular class, denoted as k.
- $b_j^{(1)}$ *and* $b_k^{(2)}$: The biases associated with the concealed layer and the result layer, correspondingly.
- σ: The activation function, which is applied to the weighted sum of inputs for each neuron.
- h: The quantity of neurons in the hidden layer.
- d: The count of input features.

3.3 Model Performance Validation

3.3.1 Data Preprocessing for Enhanced Machine Learning

Data preprocessing, a critical first step, transforms raw, unstructured data into a refined format compatible with the Multilayer Perceptron (MLP) architecture. This step ensures optimal model performance and facilitates meaningful analysis. In this project, each machine-generated input is a two-second vibration signal derived from axis-acceleration recordings. These raw time-domain signals are transformed into the frequency domain using Fast Fourier Transform (FFT). This transformation is crucial as features extracted from the frequency domain significantly enhance the MLP model's accuracy when assessing the bearing's condition [31].
Figures 2 and 3 will visually represent the analysis by showcasing samples of the vibration data and the relationship between Root Mean Square (RMS) values (mm/s) and frequency (Hz). These figures, generated from the collected data, will provide practical insights into the preprocessing steps.
Additionally, Table 1 represents a sample of vibration data, will be presented to provide a detailed snapshot of the processed information. This table offers a structured view of the input data, facilitating a clearer understanding of the dataset employed for training and assessing the MLP model. Daily monitoring was conducted over a span of three years, and the readings were identified and categorized into warning and danger zones in accordance with the protocol ISO Standard [32].

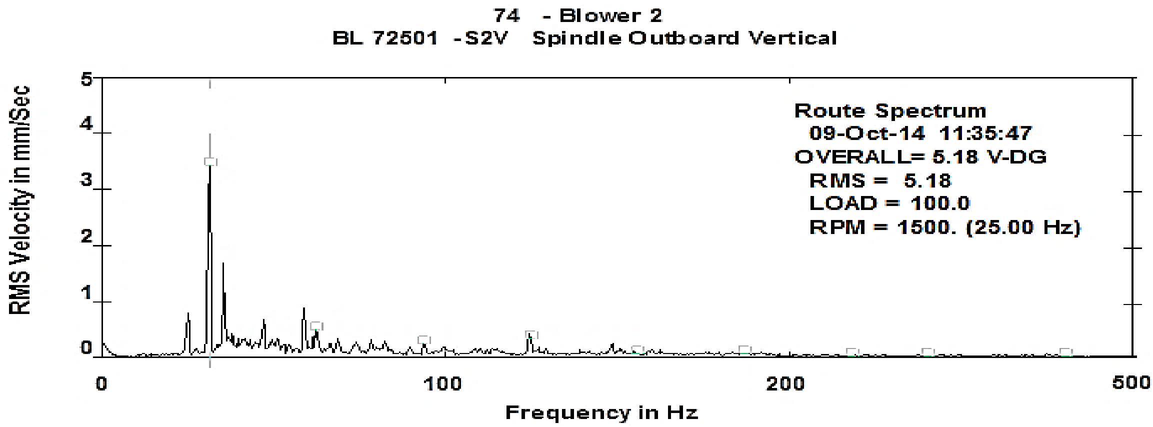


Figure 2: Spindle's outboard bearing diagram meets ISO standards [18].

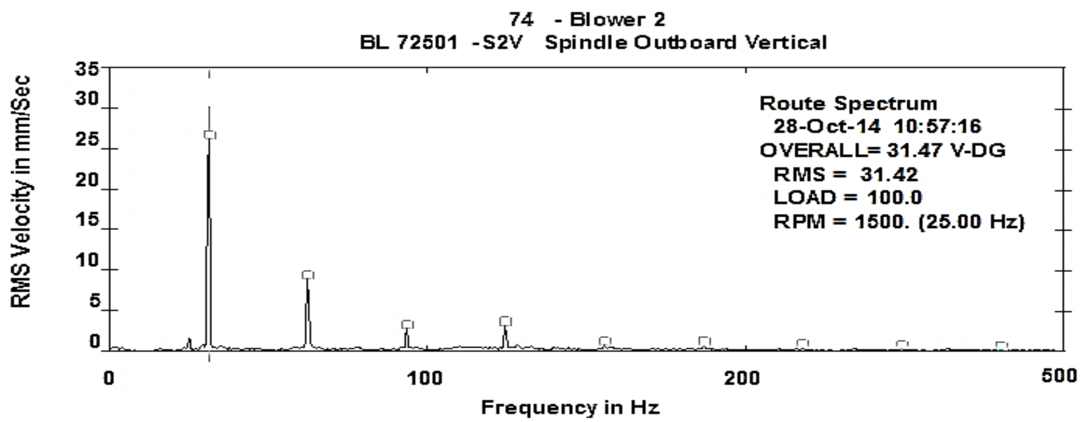


Figure 3: Blower's outboard bearing shows ISO-acceptable faults [18].

Table 1. Sample of Vibration Velocity Measurements

DATE	TIME	RMS mm/Sec	ISO STANDARD	LABEL
01/01/2013	3:54:29	7.061	JUST TOLERABLE	3
02/01/2013	3:54:48	7.629	JUST TOLERABLE	3
03/01/2013	3:55:05	2.702	ALLOWABLE	2
04/01/2013	3:55:23	6.583	JUST TOLERABLE	3
05/01/2013	3:55:45	7.461	JUST TOLERABLE	3
06/01/2013	3:55:58	3.402	ALLOWABLE	2
07/01/2013	3:56:45	6.943	JUST TOLERABLE	3
08/01/2013	3:57:17	4.484	ALLOWABLE	2
09/01/2013	3:57:41	7.578	JUST TOLERABLE	3
10/01/2013	3:58:04	7.322	JUST TOLERABLE	3
11/01/2013	3:58:27	4.707	JUST TOLERABLE	3
12/01/2013	3:58:52	7.318	JUST TOLERABLE	3
13/01/2013	2:10:23	7.48	JUST TOLERABLE	3
14/01/2013	2:10:41	8.704	JUST TOLERABLE	3
15/01/2013	2:10:56	3.797	ALLOWABLE	2
16/01/2013	2:11:11	5.876	JUST TOLERABLE	3
17/01/2013	2:11:29	9.681	JUST TOLERABLE	3
18/01/2013	2:11:43	3.755	ALLOWABLE	2
19/01/2013	2:12:33	8.682	JUST TOLERABLE	3
20/01/2013	2:12:58	4.437	ALLOWABLE	2
21/01/2013	2:13:20	7.196	JUST TOLERABLE	3
22/01/2013	2:13:39	9.919	JUST TOLERABLE	3
23/01/2013	2:14:06	5.759	JUST TOLERABLE	3
24/01/2013	2:14:28	6.687	JUST TOLERABLE	3
25/01/2013	11:10:02	6.159	JUST TOLERABLE	3
26/01/2013	11:10:20	7.124	JUST TOLERABLE	3
27/01/2013	11:10:36	2.655	ALLOWABLE	2
28/01/2013	11:10:54	5.638	JUST TOLERABLE	3

Furthermore, to ensure compatibility with the chosen MLP architecture, data preparation involves meticulous handling of null values within the original dataset, thereby enhancing model performance and analysis accuracy.

3.3.2 Model Evaluation Metrics

After model training, various evaluation metrics are used to assess its performance. These metrics measure the model's ability to accurately predict outcomes when presented with new, unseen data [33]. The confusion matrix serves as a structured table offering insights into the effectiveness of a model's predictions, particularly in classification scenarios. Table 2 below demonstrates how the confusion matrix simplifies the visualization of true positives (TP), false positives (FP), false negatives (FN), and true negatives (TN). A correct prediction of a positive or negative value is termed as a true positive or true negative, respectively. Conversely, a false positive occurs when the model incorrectly predicts a positive value, while a false negative arises when a negative value is inaccurately predicted. An optimal model exhibits high accuracy, with most predictions concentrated on the diagonal of the confusion matrix, indicating successful classifications across the dataset.

Table 2. Confusion Matrix Example

Actual vs Predicted	Positive	Negative
Positive	TP	FP
Negative	FN	TN

To determine the most efficient classification model, each algorithm undergoes testing using a 70/30 training-testing split of the feature-rich dataset. Their performance is evaluated using the following metrics:

- Precision: Measures a classifier's accuracy in identifying true positive instances shown in Equation (2).

$$\textbf{Precision} = \frac{TP}{TP+FP} \quad (2)$$

- Recall: Assesses the model's capability to identify all positive instances. The mathematical equation is expressed as Equation (3).

$$\textbf{Recall} = \frac{TP}{TP+FN} \quad (3)$$

- F-Score: Represents a combined measure of precision and recall. The mathematical Equation shown in Equation (4).

$$\mathbf{F-Score} = 2 \times \frac{\mathbf{Recall} \times \mathbf{Precision}}{(\mathbf{Recall} + \mathbf{Precision})} \quad (4)$$

- Accuracy: Represents the proportion of accurate predictions to the overall predictions generated. The mathematical Equation shown in Equation (5).

$$\mathbf{Accuracy} = \frac{\mathbf{TP} + \mathbf{TN}}{\mathbf{TP} + \mathbf{TN} + \mathbf{FP} + \mathbf{FN}} \quad (5)$$

- Matthews Correlation Coefficient (MCC): Assesses binary classification model accuracy, yielding a score between -1 (disagreement) and +1 (agreement). The mathematical Equation shown in Equation (6).

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP) \times (TP+FN) \times (TN+FP) \times (TN+FN)}} \quad (6)$$

- Receiver Operating Characteristic (ROC) Curve: Illustrates the model's performance across various threshold settings, depicting the balance between the true positive rate (TPR) and false positive rate (FPR). A ROC graph is a two-dimensional Figure that contrasts TPR (sensitivity) on the y-axis and FPR equal one minus specificity on the x-axis to reflect the original two-class problems. TPR and FPR, both of which are defined in Equations (7), (8) and (9).

$$\mathbf{sensitivity} = \mathbf{TPR} = \frac{\mathbf{TP}}{\mathbf{TP} + \mathbf{FN}} \quad (7)$$

$$\mathbf{specificity} = \mathbf{TNR} = \frac{\mathbf{TN}}{\mathbf{FP} + \mathbf{TN}} \quad (8)$$

$$\mathbf{FPR} = 1 - \mathbf{specificity} = \frac{\mathbf{FP}}{\mathbf{FP} + \mathbf{TN}} \quad (9)$$

3.4 Hyperparameter Tuning with GridSearchCV for Multilayer Perceptron: Optimizing Neural Network Performance

Hyperparameter tuning is essential for optimizing ML models, especially Multilayer Perceptrons (MLPs). Parameters such as learning rate, batch size, and the number of hidden layers significantly impact MLP performance. GridSearchCV (Grid Search Cross-Validation) enables a systematic exploration of the hyperparameter space, evaluating the MLP across various combinations. This process helps identify the optimal configuration for enhanced accuracy and generalization. For example, optimizing the number of neurons in each layer and selecting appropriate activation functions can significantly improve the MLP's ability to learn complex patterns.

As exemplified in [34], the application of GridSearchCV provides a structured and efficient method for navigating through Hyperparameter configurations, ensuring the fine-tuning of MLPs for optimal performance across diverse datasets and. By employing a systematic hyperparameter tuning approach, this study unlocks the full potential of Multilayer Perceptrons, paving the way for the development of highly effective Artificial Neural Network (ANN) models applicable to diverse machine learning domains.

3.5 Case Study and Data Collection

3.5.1 Overview of TCI Sanmar Chemicals facility

This section investigates the functionality of an industrial forced air blower employed at the TCI Sanmar Chemicals facility in Port Said, Egypt. Notably, this facility is a prominent producer of Caustic Soda, PVC, and Calcium Chloride Granules within the MENA region. The specific focus of this investigation lies on the role of the blower in the final stage of PVC production.

3.5.2 Configuration of the Forced Air Blower

Figure 4 illustrates the forced air blower's configuration, which includes three elevated stages and a ground stage. In the production process, a slurry of PVC powder and water is treated. Water expulsion, powder drying, cyclone separation, and granule size control occur across consecutive stages. Rotary valves facilitate material transportation.

3.5.3 Technical Specifications Overview

The design and technical operating specifications of the blower are comprehensively outlined, detailing key parameters for analysis. Operating conditions encompass atmospheric air and gas handling, with an operating temperature range between Normal (114000 °C) and Design (135500 °C). The flow capacity is 1.12 m^3/h, with a density at suction of 50 kg/m^3. The pressure at suction is 750 mmwc, and the discharge pressure varies from Normal (29 mmwc) to Maximum (35 mmwc). Ambient conditions include temperatures of 80°C in Summer and 76°C in Spring, with relative humidity at the suction being a significant factor.

In terms of mechanical design, the blower operates at a fan speed of 1490 rpm, featuring an impeller diameter of 1726 mm. The critical speed is set at 2285 rpm, and the fan efficiency stands at 84.5%. The blower's power specifications include an installed power of 400 KW and a shaft power of 360 KW.

3.5.4 Measuring Points and System Components

Strategically positioned measuring points within the system, highlighted in Figure 5, play a vital role in data collection and performance analysis.

3.5.5 Vibration Data Monitoring and Piezoelectric Transducer Specifications

To monitor and analyze vibration data, the CSI accelerometer by Emerson is utilized. The analysis process is supported by the AMS Suite Machinery Health Analyzer Ver 5.51 software. Furthermore, specific details about the piezoelectric transducer employed for vibration data collection are provided. The transducer exhibits a sensitivity of 100 mV/g with a ± 5% margin of error, and its operational frequency range spans from 0.5 Hz to 10000 Hz with a ± 0.5 dB tolerance [35].

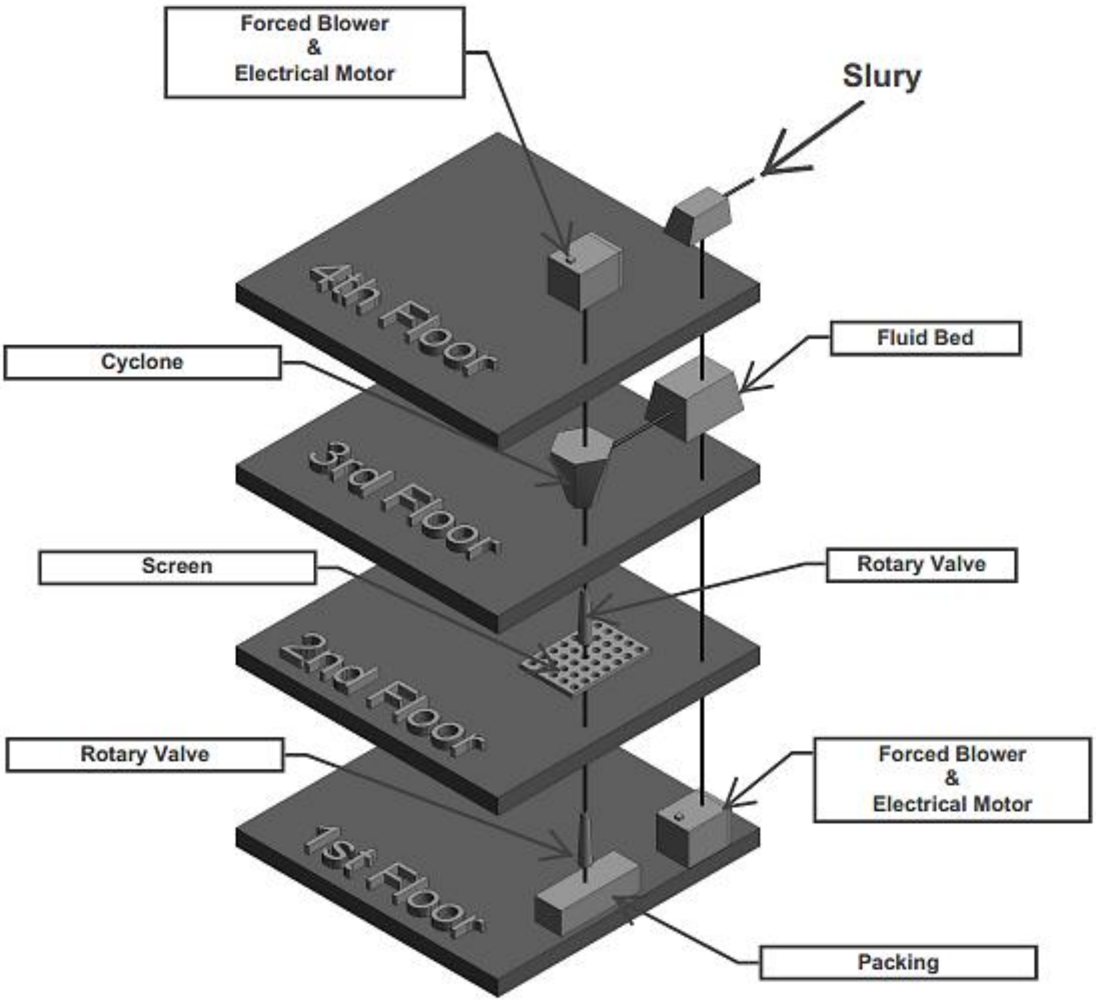
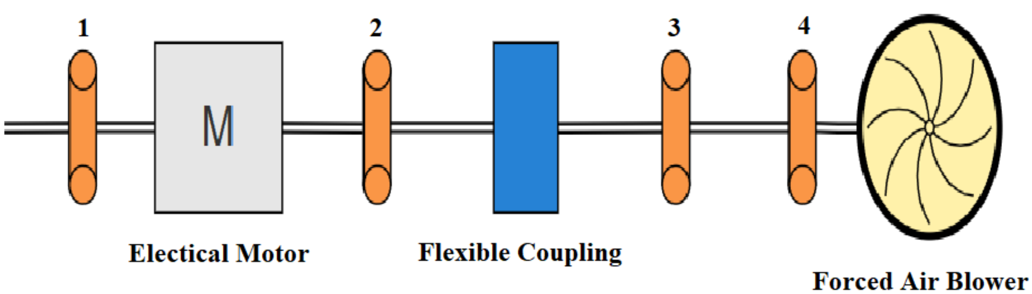


Figure 4: Schematic diagram of the structural composition of the production.



1: Inboard Bearing in Electrical Motor.
2: Outboard Bearing in Electrical Motor.
3 : Inboard Bearing in Forced Air Blower.
4 : Outboard Bearing in Forced Air Blower.

Figure 5: Schematic diagram of the blower showing the vibration measuring points.

4 RESULTS AND DISCUSSION

Building upon the initial research that explored various machine learning algorithms (SVM, RF, LR, KNN, and XGBoost), this study delves into the Multilayer Perceptron (MLP) algorithm for enhanced predictive modeling. The shift to a more complex neural network approach yielded promising results. Rigorous experimentation using Python, with a 70/30 data split for training and evaluation, demonstrated MLP's superiority over the previously utilized models. This highlights the effectiveness of neural network paradigms for capturing complex data patterns and ultimately improving predictive modeling outcomes.

4.1 Comparison of Performance Metrics for Classification Algorithms

Each metric offers valuable insights into the algorithm's performance and guides the model selection process. Below, we explain each metric and its impact on decision-making:

- Accuracy reflects the overall correctness of predictions, indicating the percentage of instances correctly classified. A higher accuracy implies a more reliable model, making it a crucial metric for performance evaluation. Results for each algorithm (SVM, RF, LR, KNN, XGB, MLP) are displayed in Figure 6, allowing for a visual comparison. However, in class-imbalanced cases, accuracy may be insufficient as it neglects false positives or false negatives. Thus, other metrics should be considered.

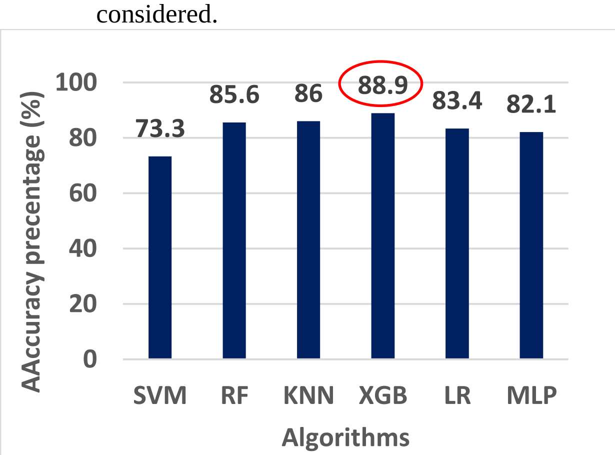


Figure 6: Comparison of Accuracy Metric for Six Algorithms.

The accuracy graph reveals XGBoost as the leading model, achieving an accuracy score of 0.889. This metric measures the overall correctness of the model's predictions.

- Precision signifies the model's accuracy in identifying positive instances among predicted positives, minimizing false positives. Figure 7 presents results for each algorithm, allowing precision comparison. Higher precision indicates fewer false positives, crucial in high-cost scenarios like machine failure.

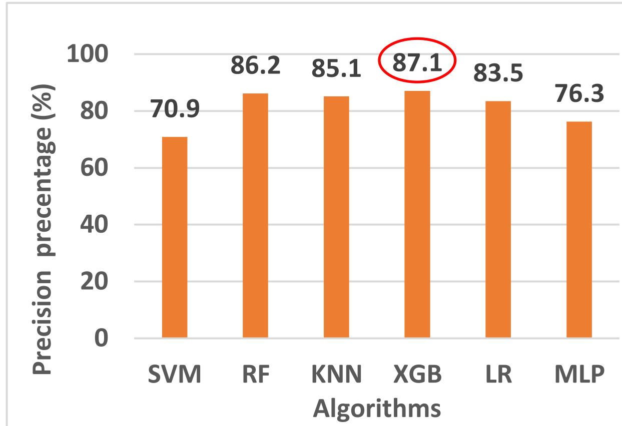


Figure 7: Comparison of Precision Metric for Six Algorithms.

The precision graph reveals a significant distinction for XGBoost, which achieves a score of 0.871. This indicates its superior ability to minimize false positive predictions compared to other models.

- Recall, also recognized as sensitivity, centers on reducing false negatives. In Figure 8, the results for each algorithm will be depicted, facilitating a comparison of their recall performance. Elevated recall indicates a decreased rate of false negatives, a factor of importance in situations where overlooking positive instances incurs high costs, as seen in failure prediction.

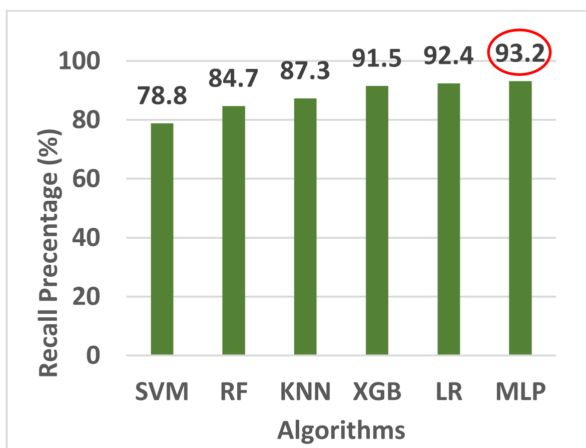


Figure 8: Comparison of Recall Metric for Six Algorithms.

The Recall graph highlights MLP as the leading model with a score of 0.932. This metric emphasizes MLP's effectiveness in capturing a higher proportion of actual positive instances within the dataset, making it a favorable choice when capturing all positives is a priority.

- In the F-score, the results for each algorithm will be displayed in Figure 9, enabling a comparison of their F-score performance. The F-score is valuable in situations where there is an uneven distribution between classes or when both precision and recall carry equal importance.

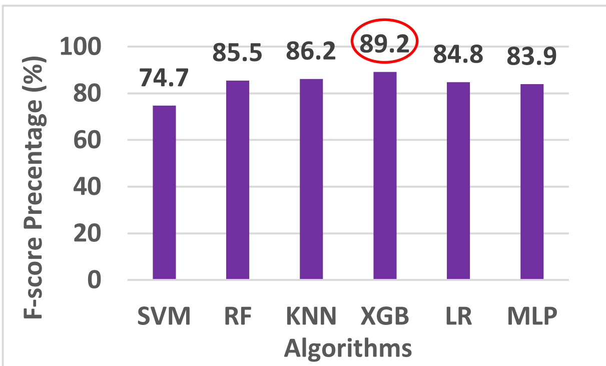


Figure 9: Comparison of F-score Metric for Six Algorithms.

XGB is retained its superiority in the F-score graph with a score of 0.892, signifying a balanced precision and recall trade-off. Showcases XGB's the aptitude to sustain equilibrium between making precise positive predictions and capturing authentic positive instances.

- In the MCC Figure 10, the results for each algorithm will be presented, allowing for a comparison of their MCC performance. MCC ranges from -1 to 1, where 1 indicates a perfect classification, 0 represents a random classification, and -1 implies a totally incorrect classification.

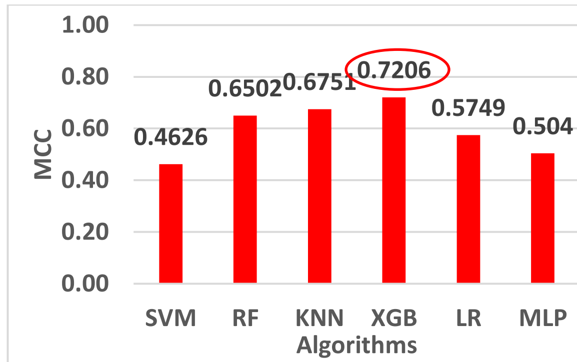


Figure 10: Comparison of MCC Metric for Six Algorithms.

In the MCC graph, a robust performance is demonstrated by XGB, scoring 0.720. This metric assesses the quality of binary classifications, with XGB's high MCC indicating a substantial agreement between predicted and actual classes.

In Table 3, the metrics' results, as illustrated in [18], are presented alongside those for the MLP algorithm with default parameters compiled from all comparison figures.

Table 3. Results of the Metrics before Using GridSearchCV

Model	Default Setting/default parameters	Accuracy	Precision	Recall	F-score	MCC
SVM	C=1.0, kernel='rbf'	0.733	0.709	0.788	0.747	0.4626
RF	Number of estimators=100, minimum samples split=2	0.856	0.862	0.847	0.855	0.6502
KNN	Number of neighbors=3, weights='uniform', algorithm='auto	0.860	0.851	0.873	0.862	0.6751
XGB	Number of estimators=20000, maximum depth=10	0.889	0.871	0.915	0.892	0.7206
LR	C=1.0	0.784	0.835	0.924	0.848	0.5749
MLP	activation = ' ReLU ', learning rate = 0.01 hidden layer sizes = (5,)	0.821	0.763	0.932	0.839	0.504

4.2 Impact of Hyperparamter Adjustment and GridSearchCV on Results

The purpose of this evaluation is to assess the effectiveness of Hyperparameter tuning using GridsearchCV in optimizing the performance of the algorithms. By visualizing the changes in the metrics, it becomes possible to analyze the impact of these adjustments on the overall effectiveness of the models. Selecting whose change will affect the results of metrics and, consequently, the choice of the most suitable model.

Following the implementation of Hyperparameter tuning with GridSearchCV employing a 5-fold cross-validation technique across various classification algorithms, a remarkable transformation in performance metrics has emerged. These metrics serve as vital indicators in assessing model effectiveness and suitability. The provided Figures for each metric vividly depict the substantial improvements observed:

- Post-Hyperparameter tuning, accuracy metrics exhibit a significant uplift across all algorithms (SVM, RF, LR, KNN, XGB, MLP). This enhancement underscores a marked increase in overall correctness in predictions, ensuring a more reliable model performance. Figure 11 accompanying this metric depict the evident rise in accuracy values.

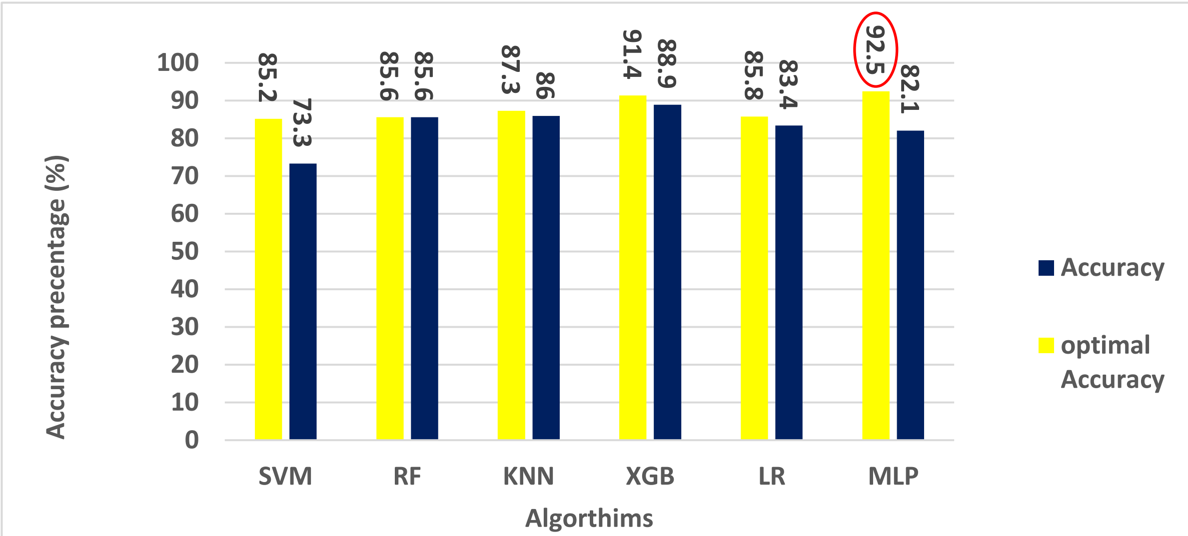


Figure 11: Comparison of Accuracy Metric before and After Adjustments.

Following grid search hyperparameter tuning, the accuracy graph reveals a notable shift, with MLP emerging as the frontrunner (0.925) among the six algorithms. This signifies a significant improvement in MLP's predictive capability compared to both its previous performance and other models.

- Precision metrics reveal noteworthy advancements, highlighting the models' refined ability to correctly identify positive instances while markedly reducing false positives. A lowered rate of false positives, vividly illustrated in the associated Figure 12, signifies the increased precision achieved post-tuning. This enhancement is particularly crucial in fields like medicine or fraud detection where minimizing false positives is critical.

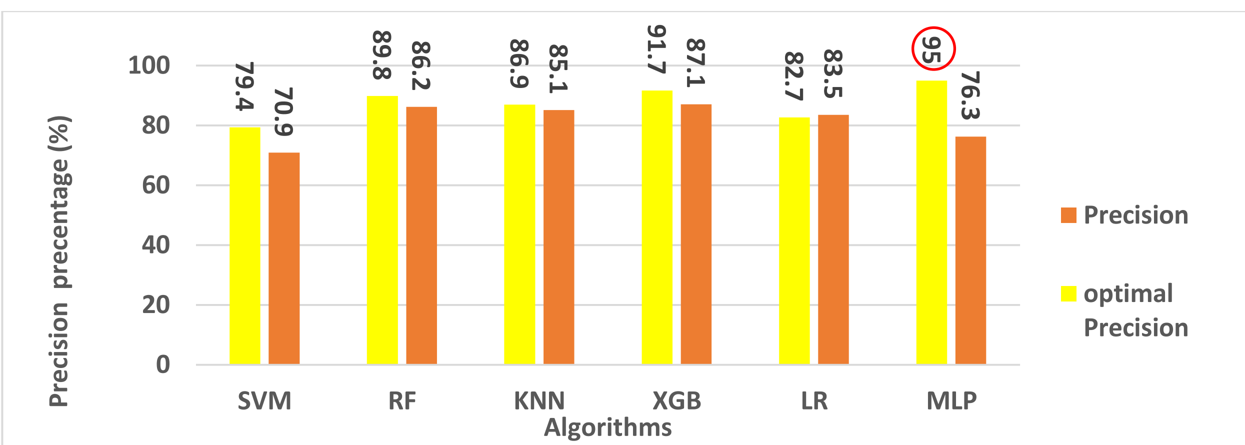


Figure 12: Comparison of Precision Metric before and after Adjustments.

In the precision comparison graph, MLP excels further with a precision score of 0.950, signifying its ability to limit false positive predictions more effectively than other algorithms. The substantial increase in precision underscores the successful optimization of MLP's parameters, making it a standout choice for precise positive classifications.

- Following hyperparameter tuning, recalculated metrics reveal positive changes, particularly in recall. This indicates an enhanced ability for models to correctly identify true positive instances and minimize false negatives, as evidenced by the significant improvements in Figure 13. These enhancements translate to a reduced rate of missing positive instances, leading to more accurate captures.

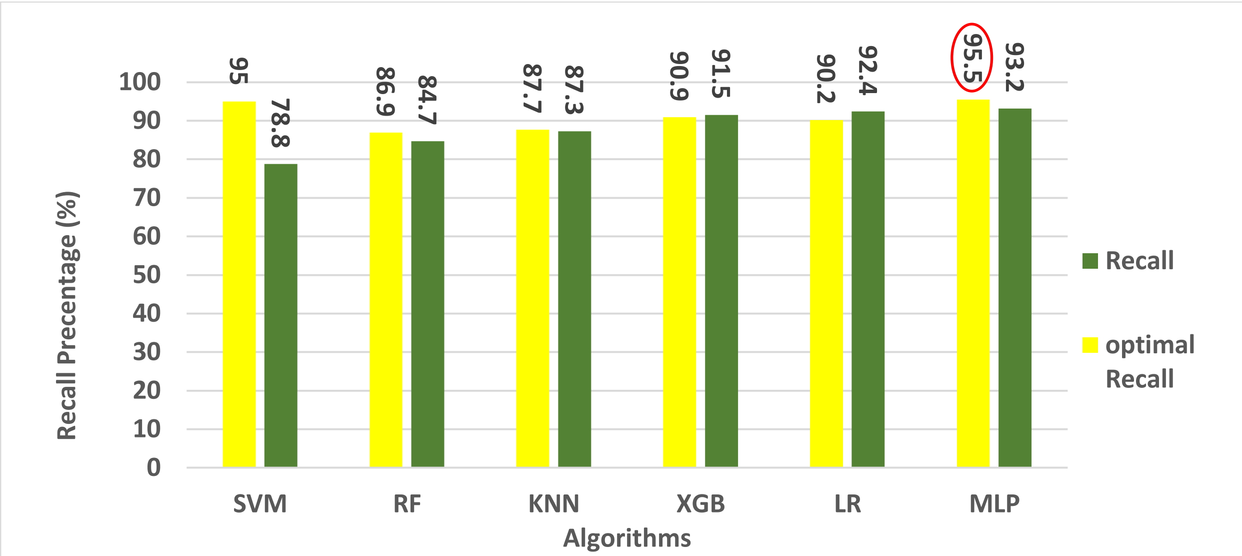


Figure 13: Comparison of Recall Metric before and after Adjustments.

The Recall graph shows MLP as the top-performing algorithm with a score of 0.935 post-Hyperparameter Tuning Grid Search. This highlights MLP's efficiency in capturing more actual positive instances compared to its previous performance and other models, showcasing its robustness in positive detection.

- The F-scores, acting as a harmonious blend of precision and recall, exhibit remarkable progress across all algorithms post-tuning. The Figure 14 associated with this metric portray a noticeable rise in F-score values, indicating a more balanced performance between precision and recall. This balanced improvement ensures that both false positives and false negatives are equally addressed, crucial in scenarios with uneven class distributions.

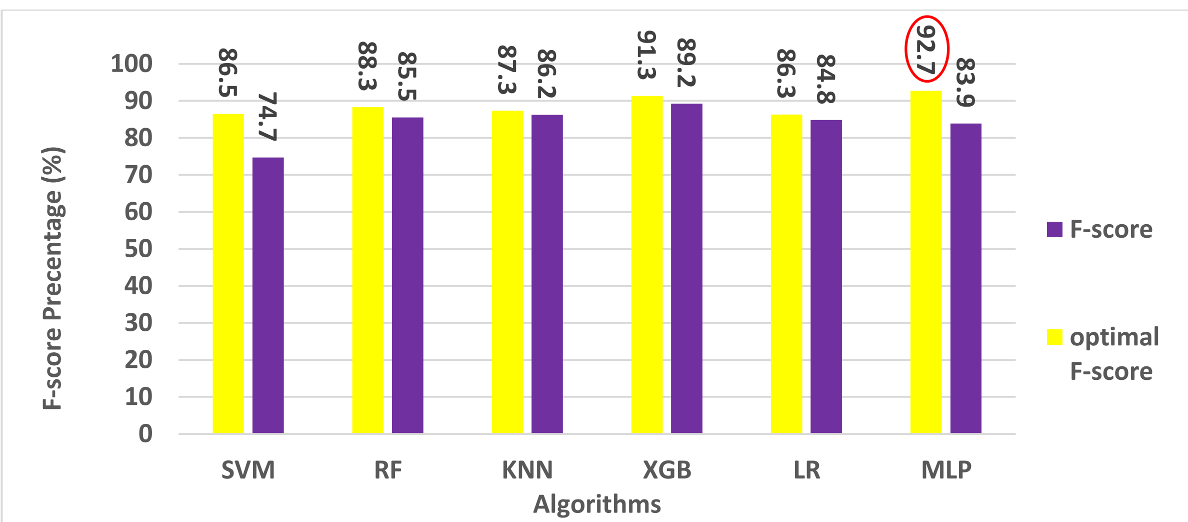


Figure 14: Comparison of F-score Metric before and after Adjustments.

After grid search hyperparameter tuning, MLP remains dominant in the F-score graph, achieving a remarkable score of 0.927. This metric emphasizes MLP's ability to balance precision and recall, highlighting its consistent performance in making accurate positive predictions and effectively capturing true positives.

- MCC values, after the Hyperparameter tuning, showcase a significant boost across all algorithms. This reflects an improved correlation between observed and predicted binary classifications. The Figure 15 provided for MCC highlight the substantial increase in these coefficients, underlining better classification accuracy. MCC proves particularly beneficial in handling imbalanced datasets. These metrics collectively offer a comprehensive evaluation of model performance. Accuracy measures overall correctness, precision focuses on minimizing false positives, recall targets false negatives, F-score balances precision and recall, and MCC considers the correlation in binary classifications.

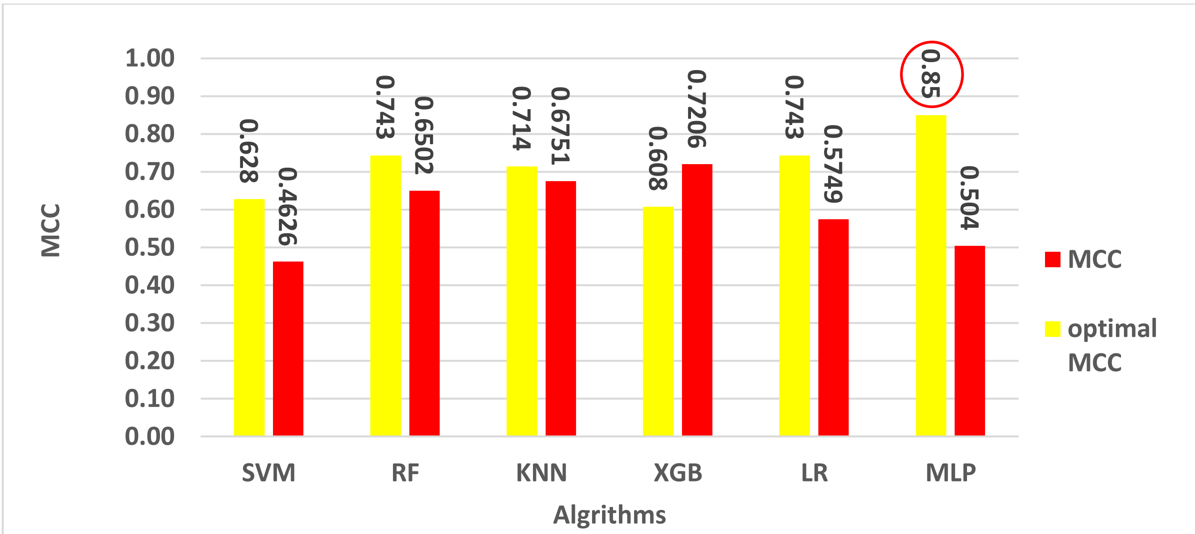


Figure 15: Comparison of MCC Metrics before and after Adjustments.

The MCC comparison graph showcases MLP's exceptional performance following hyperparameter tuning, achieving a significantly improved score of 0.850. This substantial increase signifies MLP's enhanced capability for reliable binary classification, demonstrating a strong agreement between predicted and actual classes. This positions MLP as a highly dependable model choice.

The results post-GridSearchCV are represented in Table 4, compiled from all comparison figures.

Table 4. Obtained results after adjusting the Hyperparameter values for each algorithm

Model	Best parameters	Accuracy	Precision	Recall	F-score	MCC
SVM	C=10, kernel=linear	0.852	0.794	0.951	0.865	0.628
RF	minimum samples split=3, number of estimators=6000	0.885	0.898	0.869	0.883	0.743
KNN	algorithm=auto, number of neighbors=5, weights=distance	0.873	0.869	0.877	0.873	0.714
XGB	maximum depth=3, number of estimators=10000	0.914	0.917	0.909	0.913	0.800
LR	C=10	0.856	0.827	0.902	0.863	0.608
MLP	Activation = ' relu ', learning rate = 0.001, hidden layer sizes = (16,16)	0.925	0.950	0.955	0.927	0.850

4.3 Comparative Analysis

The algorithms' predictive performance was evaluated using Area Under the Curve (AUC) graphs. Under optimal hyperparameters, the Multilayer Perceptron (MLP) achieved the highest accuracy (92%), followed by Extreme Gradient Boosting (XGB) (91%), Random Forest (88%), K-Nearest Neighbors (KNN) (87%), Logistic Regression (LR) (85%), and Support Vector Machine (SVM) (85%).

With default hyperparameters, the algorithms' initial prediction accuracies were: XGBoost (88%), KNN (86%), Random Forest (85%), Logistic Regression (83%), MLP (82%), and SVM (73%). The ROC curve plots the true positive rate (y-axis) against the false

positive rate (x-axis) as the classification threshold varies. Both the ROC curve and the Area Under the Curve (AUC) metric are used to evaluate the performance of classification models.
The code displayed in the Figure 16 (a) and (b) encapsulates the process of generating ROC (Receiver Operating Characteristic) curves utilizing the best parameters obtained through GridSearchCV shown in Figure 17.

```
from sklearn import metrics
import matplotlib.pyplot as plt
from sklearn.metrics import roc_curve, roc_auc_score
import seaborn as sns
from xgboost import XGBClassifier
sns.set()
```

(a)

```
from pandas.core.common import random_state

classifiers=[LogisticRegression('C'=10,'kernel'=linear),SVC(probability=True),

KNeighborsClassifier('algorithm'=auto,'number_of_neighbors'=5,    'weights'=distance),

XGBClassifier('maximum depth'=3, 'number_of_estimators'=10000),

RandomForestClassifier('minimum_samples_split'=3, 'number_of_estimators'=6000),

MLPClassifier('Activation' = relu , 'learning_rate' = 0.001,'hidden_layer_sizes' = (16,16))]

# color-marker-linestyle
CML = ['go-', 'r*--', 'b^-', 'cs--', 'mx-','k+-']
# Define a result table as a DataFrame
result_table = pd.DataFrame(columns=['classifiers', 'fpr','tpr','auc'])
# Set name of the classifiers as index labels
result_table.set_index('classifiers', inplace=True)
result_table
```

(b)

Figure 16: Key Steps in ROC Analysis and Classifier setup for predictive Modeling in Python. (a) Necessary Library imports for ROC Analysis. (b) Classifier setup and Result Compilation with Pandas Data Frame.

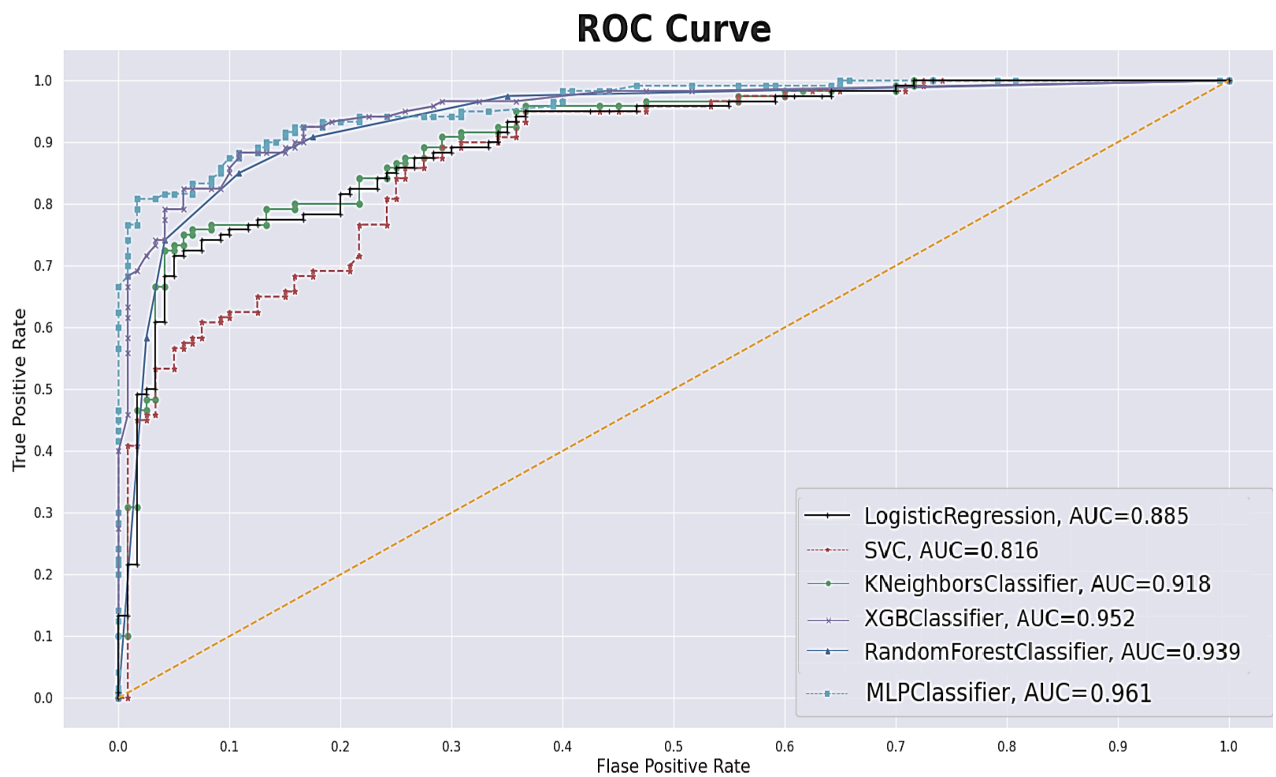


Figure 17: ROC Curve Analysis.

A diagonal line on the ROC curve indicates random classification, where true positives and false positives occur at equal rates. The AUC (Area Under the Curve) measures the classifier's performance by calculating the area between the ROC curve and the diagonal line. A higher AUC signifies better performance, allowing for comparison between different models [36]. The diagonal line on the ROC plot, connecting (0, 0) and (1, 1), represents random chance performance. This scenario implies that the model's ability to distinguish between classes is no better than random guessing, with equal probabilities of correct and incorrect predictions. Points above the diagonal indicate better-than-chance performance, meaning the model's true positive rate (correctly identifying positive cases) exceeds its false positive rate (incorrectly identifying negative cases). Conversely, points below the diagonal suggest worse-than-chance performance, where the model has a higher rate of false positives than true positives.

5 CONCLUSION AND FUTURE WORK

5.1 Conclusion

In conclusion, this research offers valuable insights that extend beyond the immediate scope of the study, concentrating on the formulation of an efficient predictive maintenance methodology for a forced blower in PVC powder manufacturing at TCI Sanmar. The significant outcomes of investigation are as follows:

- The utilization of machine learning algorithms demonstrated remarkable adaptability in analyzing maintenance-related data from TCI Sanmar Chemicals' industrial air blower, representing a substantial advancement in predictive maintenance strategies and promising more accurate and proactive equipment upkeep.

- In a comprehensive evaluation that considered various algorithms used in previous research related to the same case—namely, SVM, LR, KNN, XGBoost, RF and MLP algorithms stood out as the top performance. With an MCC of 0.850 and consistently high AUC values, the MLP algorithm showcased unparalleled efficacy in accurately foreseeing machine failures compared to its counterparts.

- The robust performance of the MLP algorithm reaffirms its superiority over other algorithms used in the same industrial setting, underscoring its reliability as a predictive maintenance solution. Its ability to outperform previously employed algorithms in predicting potential equipment malfunctions establishes it as a dependable tool for proactive maintenance interventions.

- This exploration underscores the strategic importance of employing the Multilayer Perceptron algorithm in predictive maintenance frameworks. Its precision in predicting faults not only outshines its counterparts from previous research but also presents a valuable opportunity for industrial sectors to enhance equipment reliability, minimize downtime, and optimize maintenance schedules for improved operational efficiency.

5.2 Future work

The proposed proof-of-concept, which shows how machine learning aids in the analysis and provision of early data from industrial facilities to anticipate the status of industrial machinery, was the subject of the study. Based on a thorough case study on vibration monitoring the following considerations are recommended in future work:

- Increasing data volume as a larger training dataset can potentially improve model performance by providing the model with more information to learn from.
- Merging machinery malfunction causes with sensor data enhances predictive capabilities, enabling the model to anticipate malfunctions and identify their specific types or causes, offering actionable insights for tailored maintenance strategies.
- Creating an integrated online platform that seamlessly collects data from machines, utilizing sensors as a prime example, and directly feeds it into the algorithm, constitutes an efficient system for predictive maintenance processes.
- Utilizing deep learning algorithms can significantly enhance the efficiency of predictive maintenance strategies.

Acknowledgments

The authors would like to express their sincere gratitude to TCI Sanmar Chemicals, a pivotal PVC powder manufacturing facility in Port Said, Egypt. The invaluable assistance, particularly the provision of historical data for the forced blower from TCI Sanmar Chemicals and the knowledge sharing were important and essential to the successful completion of this study.

6 REFERENCES

[1] C. O. Klingenberg, M. A. V. Borges, and J. A. do V. Antunes, “Industry 4.0: What makes it a revolution? A historical framework to understand the phenomenon,” *Technol. Soc.*, vol. 70, p. 102009, 2022, doi: https://doi.org/10.1016/j.techsoc.2022.102009.

[2] F. Tao, H. Zhang, A. Liu, and A. Y. C. Nee, “Digital Twin in Industry: State-of-the-Art,” *IEEE Trans. Ind. Informatics*, vol. 15, no. 4, pp. 2405–2415, 2019, doi: 10.1109/TII.2018.2873186.

[3] P. K. Reddy, V. Q. Pham, and N. Deepa, “Industry 5 . 0 : A Survey on Enabling Technologies and Potential Applications Industry 5 . 0 : A Survey on Enabling Technologies and Potential Applications,” no. July, 2021, doi: 10.1016/j.jii.2021.100257.

[4] S. Nahavandi, "Industry 5.0—A Human-Centric Solution," *Sustainability*, vol. 11, no. 16, p. 4371, Aug. 2019. doi: 10.3390/su11164371.

[5] L. Bokonda, O. T. Khadija, and N. Souissi, “Predictive analysis using machine learning : Review of trends and methods,” no. November, 2020, doi: 10.1109/ISAECT50560.2020.9523703.

[6] A. Keleko, B. Kamsu-Foguem, R. Houe Ngouna, and A. Tongne, “Artificial intelligence and real-time predictive maintenance in industry 4.0: a bibliometric analysis,” *AI Ethics*, vol. 2, Mar. 2022, doi: 10.1007/s43681-021-00132-6.

[7] F. Raza, “AI for Predictive Maintenance in Industrial Systems AI for Predictive Maintenance in Industrial Systems Abstract,” no. November, 2023, doi: 10.13140/RG.2.2.27313.35688.

[8] Werbi´nska-Wojciechowska, S.; Winiarska, K. Maintenance Performance in the Age of Industry 4.0: A Bibliometric Performance Analysis and a Systematic Literature Review. Sensors 2023, 23, 1409. https://doi.org/10.3390/s23031409

[9] A. Orailoglu and G. Goos, *Systems : Architectures , Modeling , and Simulation*. 2021. doi: 10.1007/978-3-031-04580-6.

[10] C. Wang, J. Tian, F. Zhang, Y. Ai, and Z. Wang, “Dynamic modeling and simulation analysis of inter-shaft bearing fault of a dual-rotor system,” *Mech. Syst. Signal Process.*, vol. 193, p. 110260, 2023, doi: https://doi.org/10.1016/j.ymssp.2023.110260.

[11] A. C. Márquez, A. de la Fuente Carmona, and S. Antomarioni, “A process to implement an artificial neural network and association rules techniques to improve asset performance and energy efficiency,” *Energies*, vol. 12, no. 18, 2019, doi: 10.3390/en12183454.

[12] P. F. Orrù, A. Zoccheddu, L. Sassu, C. Mattia, R. Cozza, and S. Arena, “Machine learning approach using MLP and SVM algorithms for the fault prediction of a centrifugal pump in the oil and gas industry,” *Sustain.*, vol. 12, no. 11, 2020, doi: 10.3390/su12114776.

[13] W. Silva and M. Capretz, “Assets Predictive Maintenance Using Convolutional Neural Networks,” in *2019 20th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD)*, 2019, pp. 59–66. doi: 10.1109/SNPD.2019.8935752.

[14] Z. Ye and J. Yu, “AKSNet: A novel convolutional neural network with adaptive kernel width and sparse regularization for machinery fault diagnosis,” *J. Manuf. Syst.*, vol. 59, pp. 467–480, 2021, doi: https://doi.org/10.1016/j.jmsy.2021.03.022.

[15] K. Chen, S. Pashami, Y. Fan, and S. Nowaczyk, “Predicting Air Compressor Failures Using Long Short Term Memory Networks BT - Progress in Artificial Intelligence,” 2019, pp. 596–609.

[16] W. Wu, M. Liu, Q. Liu, and W. Shen, “A quantum multi-agent based neural network model for failure prediction,” *J. Syst. Sci. Syst. Eng.*, vol. 25, no. 2, pp. 210–228, 2016, doi: 10.1007/s11518-016-5308-2.

[17] M. P. Rajakumar, J. Ramya, and B. U. Maheswari, “Health monitoring and fault prediction using a lightweight deep convolutional neural network optimized by Levy flight optimization algorithm,” *Neural Comput. Appl.*, vol. 33, no. 19, pp. 12513–12534, 2021, doi: 10.1007/s00521-021-05892-0.

[18] K. Salem, E. AbdelGwad, and H. Kouta, “Predicting Forced Blower Failures Using Machine Learning Algorithms and Vibration Data for Effective Maintenance Strategies,” *J. Fail. Anal. Prev.*, vol. 23, no. 5, pp. 2191–2203, 2023, doi: 10.1007/s11668-023-01765-x.

[19] T. Sexton, M. P. Brundage, M. Hoffman, and K. C. Morris, “Hybrid datafication of maintenance logs from AI-assisted human tags,” *Proc. - 2017 IEEE Int. Conf. Big Data, Big Data 2017*, vol. 2018-Janua, pp. 1769–1777, 2017, doi: 10.1109/BigData.2017.8258120.

[20] K. Miner-Romanoff, “Student Perceptions of Juvenile Offender Accounts in Criminal Justice Education,” *Am. J. Crim. Justice*, vol. 39, no. 3, pp. 611–629, 2014, doi: 10.1007/s12103-013-9223-5.

[21] V. Nasteski, “An overview of the supervised machine learning methods,” *Horizons.B*, vol. 4, no. December 2017, pp. 51–62, 2017, doi: 10.20544/horizons.b.04.1.17.p05.

[22] T. Kebabsa, N. Ouelaa, and A. Djebala, “Experimental vibratory analysis of a fan motor in industrial environment,” *Int. J. Adv. Manuf. Technol.*, vol. 98, no. 9–12, pp. 2439–2447, 2018, doi: 10.1007/s00170-018-2391-1.

[23] T. Chen and C. Guestrin, “XGBoost: A scalable tree boosting system,” *Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, vol. 13-17-Augu, pp. 785–794, 2016, doi: 10.1145/2939672.2939785.

[24] Z. M. Çinar, A. A. Nuhu, Q. Zeeshan, O. Korhan, M. Asmael, and B. Safaei, “Machine learning in predictive maintenance towards sustainable smart manufacturing in industry 4.0,” *Sustain.*, vol. 12, no. 19, 2020, doi: 10.3390/su12198211.

[25] J. Eriksson, "Machine Learning for Predictive Maintenance on Wind Turbines – Using SCADA Data and the Apache Hadoop Ecosystem," Master's thesis, Linköping University, Department of Computer and Information Science, 2020, LIU-IDA/LITH-EX-A--2020/008--SE.

[26] Y. Lei, N. Li, L. Guo, N. Li, T. Yan, and J. Lin, “Machinery health prognostics: A systematic review from data acquisition to RUL prediction,” *Mech. Syst. Signal Process.*, vol. 104, pp. 799–834, 2018, doi: https://doi.org/10.1016/j.ymssp.2017.11.016.

[27] H. Mohamed, M. Zahran, and O. C. Saavedra, “Assessment of Artificial Neural Network for Bathymetry Estimation Using High Resolution Satellite Imagery in Shallow Lakes: Case Study El Burullus Lake,” *EighteenthInternational Water Technol. Conf.*, vol. 18, no. December, pp. 12–14, 2015.

[28] Y. Lecun *et al.*, “Gradient-based learning applied to document recognition To cite this version : HAL Id : hal-03926082 Gradient-Based Learning Applied to Document Recognition,” vol. 86, no. 11, pp. 2278–2324, 2023.

[29] K. G. Kim, "Deep Learning," Healthc Inform Res, vol. 22, no. 4, pp. 351-354, Oct. 2016, doi: 10.4258/hir.2016.22.4.351.

[30] C. M. Bishop, "Pattern Recognition and Machine Learning," 2006. [Online]. Available: https://www.microsoft.com/en-us/research/uploads/prod/2006/01/Bishop-Pattern-Recognition-and-Machine-Learning-2006.pdf

[31] W. J. Lee, H. Wu, H. Yun, H. Kim, M. B. G. Jun, and J. W. Sutherland, “Predictive maintenance of machine tool systems using artificial intelligence techniques applied to machine condition data,” *Procedia CIRP*, vol. 80, pp. 506–511, 2019, doi: 10.1016/j.procir.2018.12.019.

[32] I. Standard, “INTERNATIONAL,” vol. 995, 1995.

[33] “Brownlee, Jason. Tour of Evaluation Metrics for Imbalanced Classification.” https://machinelearningmastery.com/tour-of-evaluation-metrics-for-imbalanced-classification/

[34] J. Bergstra and Y. Bengio, “Random Search for Hyper-Parameter Optimization,” *J. Mach. Learn. Res.*, vol. 13, pp. 281–305, Mar. 2012. doi: 10.1561/2200000006.

[35] Micromega, “Industrial Accelerometer (4..20mA),” no. June, 2017, [Online]. Available: https://micromega-dynamics.com/wp-content/uploads/2016/08/IAC-I-03-EN-Rev2p3.pdf

[36] “Understanding AUC ROC Curve Towards Data Science - Understanding AUC - ROC Curve Sarang Narkhede - StuDocu.” https://www.studocu.com/row/document/कठमणड-वशववदयाय/machine-learning/understanding-auc-roc-curve-towards-data-science/8902167