Forecasting Imminent Failures in Electrical Industrial Centrifuge using Machine Learning

Leandro Ferreira Moura Filho
Universidade de Pernambuco
Recife, Brazil
leandro.ferreira.filho@live.com

Rodrigo de Paula Monteiro
Unicap-Icam International School
Universidade Católica de Pernambuco
Recife, Brazil
rodrigo.paula@unicap.br

Diego Pinheiro
Unicap-Icam International School
Universidade Católica de Pernambuco
Recife, Brazil
diego.silva@unicap.br

Patricia Takako Endo
Universidade de Pernambuco
Recife, Brazil
patricia.endo@upe.br

Andrea Maria N C Ribeiro
Unicap-Icam International School
Universidade Católica de Pernambuco
Recife, Brazil
andrea.ribeiro@unicap.br

***Abstract*—Predictive maintenance detects early signs of functional loss in electrical equipment and is a widely used strategy to minimize industrial costs. Machine learning models, the predictive maintenance state-of-the-art, forecast early failures in a myriad of industrial equipment but their application in electrical industrial centrifuge, an equipment whose failure can be critical, was rarely investigated. We hypothesize that machine vibration, temperature readings, lubrication levels, and motor current values help to forecast mechanical failures in an electrical industrial centrifuge. We validate our hypothesis by collecting 23,000 monitoring samples of an electrical industrial centrifuge and its respective history of 757 breakdowns over a 3-yr period, training a Random Forest machine learning model with the first two years, and testing in the third year. Our results show that the trained model can forecast failures with a F1-score of .34, establishing a baseline for forecasting imminent failures in industrial centrifuge. Our proposed approach aids technician inspections, reduces factory downtime, and potentially saves tens of thousands of dollars every year.**

***Index Terms*—Predictive Maintenance. Condition-based Maintenance. Centrifuge. Machine Learning. Random Forest.**

I. Introduction

Machine learning has gained significant popularity recently but its implementation in industrial plants still lags. Intelligent tools are still in early stages in industrial processes, and there is a shortage of skilled professionals who can identify opportunities for its implementation.

While private companies adopts machine learning more widely, it is still relatively rare to find applications of machine learning in the industrial sites and operational lines. One particular area where machine learning can have a significant impact is industrial maintenance [1]. Predicting breakdowns and categorizing anomalies are essential tasks to ensure that operations run smoothly and avoid unexpected downtime.

In an industrial setting, unexpected downtime of production lines or key equipment can lead to significant financial losses [2]. The responsibility falls on the maintenance and reliability department that is in charge of preserving the equipment's integrity and preventing unplanned failures, anomalies or breakdowns of critical plant equipment.

Maintenance strategies have been developed over the years to increase the reliability of operating equipment by detecting and analyzing signals of failures and anomalies in their operation. These signals can be identified by a team with technical expertise, warning that the equipment may require maintenance, as seen in Almeida [3]. Figure 1 is a representative scheme of maintenance strategies, where three fundamental approaches play roles in ensuring equipment efficiency: Predictive, Preventive, and Corrective Maintenance.

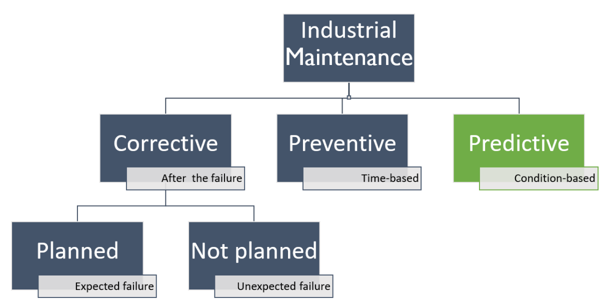

Fig. 1. Tree with the usual maintenance strategies.

One of the most used maintenance strategies is the predictive maintenance (or condition-based maintenance [4]) , which involves intervening in equipment only if signs of functional loss are detected, thereby maximizing the machine's lifespan and avoiding unnecessary maintenance costs [5].

This work aimed to perform a similar analysis as what is typically done in predictive maintenance, which is evaluating the machine's health by monitoring data (i.e, motor current, machine vibration, temperature, and sound), but by utilizing a machine learning model to assess the equipment's health in place of a human technician.

The predictive maintenance strategy is the most cost-effective approach to reducing equipment breakdowns. However, it requires personal analysis and thus is vulnerable to human errors. This work specifically focuses on a piece of equipment in operation using real operational data. We address some of the challenges that implementations of projects in operational industries present and how we overcome them. Furthermore, the equipment used for this project, a critical centrifuge used to perform the separation of grain components, is rarely discussed in related works on the subject, making this work even more valuable.

Early signs of equipment failure may be hard to detect. Additionally, meeting the demand for analysis can be challenging due to the high number of critical equipment, resulting in some equipment being overlooked and vulnerable to breakdowns. This gap can be filled by using a machine learning model as long as the quality of the analysis stays the same (or even gets improved) [6]. In this work, we exclusively focused on failures that remained undetected by the maintenance team. These failures are inherently more challenging to identify, and each instance resulted in industry downtime, causing a substantial impact. Our primary objective is to detect and anticipate these failures actively, enabling on-time interventions, hours or even days before they occur, thereby mitigating potential losses.

This work is organized as follows: Section I, presents the problem, motivation, proposal, and contributions of this work. Section II lists the published works related to the forecasting of failure in industrial equipment. Section III presents the dataset and methodology used in this work. In Section IV, we present and discuss the results achieved, and in Section V, we make the final considerations about the study developed.

II. Related work

A. Survey works

Carvalho et al. [1] systematically analyze works on machine learning applied to predictive maintenance, highlighting the methods explored in this field and the evaluation of the performance of such techniques.

Diez-Olivan et al. [5] conduct a detailed survey of published works on predictive maintenance and the use of machine learning. This study was immensely helpful in mapping possible references with similar applications.

Çinar [6] provides an extensive survey of published works on predictive maintenance with the use of machine learning, highlighting the characteristics of the databases and applications of the studies.

B. Model development works

Liu et al. [7] provide an objective analysis of the implementation of AI (Artificial Intelligence) algorithms in the analysis of faults in rotating machinery from both theoretical and industrial application perspectives.

Ziang, Yang, and Wang [8] perform predictive maintenance analyses, demonstrating the challenges of its implementation. Additionally, they classify industrial applications based on six machine learning and deep learning algorithms, comparing performance metrics for each classification.

Çinar et al. [9] analyze recent advances in machine learning techniques applied to predictive maintenance (focusing on smart manufacturing), using algorithm, equipment and machinery classification, data classification, size, and type.

Ayvaz and Alpay [10] implement machine learning-based approaches in a predictive maintenance system, with data generated from intelligent sensors (using the Internet of Things). They also present details about the best-performing algorithm in an industry.

Kolokas et al. [11] apply algorithms to structure a machine learning model in detecting the need for maintenance for equipment and components in a mass industry.

Amihai et al. [12] scrutinize how one can use vibration readings to categorize failure and to predict the remaining useful life.

These works are related to the application of machine learning in industrial maintenance strategies, providing experiences and ideas that were adapted for the application of this study. However, this study approaches the implementation in a more practical way, analyzing industrial routines and how they were used to collect data and validate the developed model.

III. Materials and methods

With a diverse range of machine learning models, each with its strengths and peculiarities, selecting the appropriate model for the application is crucial for project success. To determine this model, two analyses - one quantitative and one qualitative - were made.

The quantitative analysis involved stratifying published works related to predictive maintenance using machine learning, as detailed by Diez [5]. This work provided valuable insights into the most commonly used models, and these were organized based on their frequency of use.

Diez [5] shows that the Random Forest model is the most commonly used in works with applications similar to this one, and therefore it will be the model used for this work. However, it is worth understanding why the Random Forest is favored in industrial maintenance applications. To achieve this understanding, a qualitative analysis of the use of this model in industrial applications, mainly applications focused on industrial maintenance or predictive maintenance, was elaborated.

Concerning industrial applications, the adaptability of Random Forest to variations is essential, as industrial processes are susceptible to changes in various parameters. Additionally, the model's simplicity during implementation and presentation of results to industry leaders is a distinguishing factor.

The next step was to gather the database to train and test the model. The industrial plant where this work was developed had many equipment reading data generated by instruments in its automation system. Table I demonstrates the data collection used in this work.

TABLE I
SAMPLE OF READING DATA STORED IN THE AUTOMATION SYSTEM.

Equipment	Measurement	Timestamp
Centrifuge	3.238	01/01/2022 17:49:47
Centrifuge	3.072	01/01/2022 17:49:47
Centrifuge	3.186	01/01/2022 17:50:17
Centrifuge	3.178	01/01/2022 17:50:32
Centrifuge	3.055	01/01/2022 17:50:47

Selecting appropriate data to use is crucial to categorize faults. The study conducted by Amihai et al. [12] emphasizes the importance of vibration analysis in predicting and categorizing faults. Temperature readings, lubrication levels, and motor current values are relevant variables for mechanical failures related to vibration and were added to the model's inputs, requiring a correlation analysis between them. In addition to instrument readings, failure data is critical for classification and must also be collected. In the industrial scenario, the management system can extract equipment intervention historical data, usually used to calculate factory performance indicators.

With the input data source defined for the machine learning model, it is necessary to consider the availability of the required minimum input data and the equipment's relevance to the industrial process. Ideally, the selected equipment should be an electrical motor, given this study's focus on related works that analyzed this type of equipment. Therefore, the selection criteria for equipment should include:

- Equipment with motor current, machine vibration and temperature in the bearings measurements, with at least one year of historical data.
- Relevant or critical equipment to the process whose impacts are significant for the final production.
- Electric rotary motor class equipment.

The list of equipment with all prerequisites was raised and analyzed by the maintenance team of the industry.

After considering the available equipment list and the conditions outlined earlier, electrical industrial centrifuge set from the plant was chosen as model equipment, specifically the motor of the separator centrifuge. This equipment is responsible for separating the components of the processed grain using centrifugal force, separating materials with different densities.

Readings of motor current, machine vibration, and temperature in the bearings in the rear and front bearings are respectively named Current 01, Vibration 01, Temperature 01, and Temperature 02. The data history of Lubrication 01, which informs the lubricating oil flow rate to the machine in oil drops per minute, was also collected. This induction motor has a rated power of 300 HP, 02 pairs of poles, and a nominal mechanical speed of 1790 RPM.

The resulting database, composed of all readings taken from the selected parameters from April 2020 to April 2023, was stored in several files, formatted in CSV (Comma Separated Values). With the extracted data, it was also possible to identify some parameters that will be useful for developing the machine learning model.

It was found that the sampling rate of the data collection and storage in the database is 15 seconds, meaning that every 15 seconds, the information collected by the instrumentation in the field is stored in the database. Since it used data from 2020 to 2023, it means that around 6.5 Million inputs were extracted. Table II provides a snapshot of the data from the automation system.

TABLE II
READINGS OF THE SELECTED FEATURES.

Timestamp	Vibration	Temp. 01	Temp. 02	Current	Lubrication
14/03/2021 18:12	1,428	37,596	38,759	x	44,677
14/03/2021 18:13	2,019	x	38,803	158,936	43,677
14/03/2021 18:13	6,737	37,614	39,045	x	x
14/03/2021 18:13	x	x	38,865	158,361	x
14/03/2021 18:13	3,083	37,693	38,939	110,237	x
14/03/2021 18:14	3,155	38,627	40,031	105,787	x
14/03/2021 18:14	3,336	38,771	40,028	x	x
14/03/2021 18:14	3,002	38,624	39,972	x	x
14/03/2021 18:14	3,301	39,675	40,856	x	x
14/03/2021 18:15	3,286	x	41,035	x	x
14/03/2021 18:15	3,114	39,563	x	x	42,675
14/03/2021 18:15	3,257	40,711	40,929	x	x
14/03/2021 18:15	2,913	x	42,001	106,799	x
14/03/2021 18:16	2,843	40,782	41,889	x	x
14/03/2021 18:16	2,942	41,709	41,939	x	x

The history of breakdowns and interventions of the selected equipment can be obtained by querying the management system in the plant. By using transaction queries for work orders mapped to the selected equipment, it is possible to obtain all records of actions.

Since the data is fed by technicians during work, there is no standardization of activity descriptions that are recorded. To ensure that only relevant failures that were ignored by the maintenance team were taken into account, the interventions database was initially filtered so that only corrective maintenance and planned failures were read by the machine learning model. Only failures related to misalignment, rotary head replacement, bearing replacement or failure, and lack of lubrication were kept in the intervention database, aligning the break history with the failures mapped by the selected measurements. Inspection, rearm, cleaning, flow, and other routine interventions, although relevant to equipment maintenance, were not be considered in this work. In the end, around 30 types of failures were used for the development of the model.

A. *Preprocessing*

As usual, it is necessary to perform data preparation before feeding it to the machine learning model for the learning process. This preparation involves eliminating meaningless data, normalization, filling in empty data, as well as some preliminary analysis that helps define the structure of the learning model.

1) Removing outliers and unwanted values: When the data was collected, there was no distinction made regarding the equipment's status at the time of collection. In other words, among the operational data, there were several instances where readings were taken while the equipment was either stationary or undergoing maintenance. These data points are highly detrimental to the analysis since they exhibit distinct characteristics and typically occur after a failure event. We want to prevent these data points from biasing the model and hindering the prediction of failures, thus necessitating their removal from the dataset. To identify the periods of equipment downtime, an additional piece of information called the 'Machine Status' was extracted from the equipment's database. This allowed us to identify the periods when the equipment was not in normal operation and exclude the readings taken during those periods.

After this filtering process, an outlier analysis was performed, aiming to eliminate values that significantly deviate from the usual equipment readings (typically associated with equipment startup and shutdown events). [13] shows how SD (Standard Deviation) can be used as threshold for outliers. Usually, an outlier is defined when its value is outside of the mean ± two to five times SD interval. This approach has issues, as the outliers themselves are used in the SD calculation, but since a very solid data cleansing step was made, this approach with 3 standard deviations is enough to remove any residual outliers in the data.

2) Smoothing the granularity of the dataset: The first step in preparing the dataset is related to the granularity of the data. It is evident that there is a discrepancy between the size of the reading data and the history of breakdowns. Additionally, the raw data obtained from the factory automation system is susceptible to noise, operational transients, and general interference. Therefore, it was decided to smooth the dataset by aggregating the values using the average value for each hourly reading.

Smoothing the data is expected to reduce the amount of noise in the dataset and prevent overfitting. Furthermore, it has made the system easier to interpret and decreased the computational burden. This approach is not recommended if the application was aiming to detect failures with quick evolution. Usually, operational failures or poorly made interventions happens within the one hour interval. But, since those are not the types of failure we are trying to predict, its not expected to have any critical data point lost by utilizing this approach.

This step resulted in a dataset of approximately 25,000 inputs (one input per hour from April 2020 to April 2023).

3) Link between the two data sources: Next step was to define the strategy for linking the two data sources available. One database provides information on current measurement readings, machine vibration, temperature in the bearings, and lubrication; and another database provides information and a history of breakdowns and incidents.

The most plausible decision for the relationship between the two data sources is the date, linking the measurement data to a date with or without a breakdown. Furthermore, this strategy allows us to define a longer time horizon in order to classify the data as Normal Operation or Impending Breakdown. 24 hours was defined as the horizon prior to breakdown and whose data already showed signs of impending breakdown. Thus, it is the goal of the machine learning model to categorize the data between these two conditions.

4) Initial data visualization: A good practice in machine learning applications is to visualize input data through scatter plots. This evaluation provides insights about the graphs, as well as highlights correlation factors between the readings. Figures 2 and 3 show the scatter plots of the databases used in this work. The gray marked points represent data categorized as "Normal Operation", while the orange points represent data categorized as Impending Breakdown".

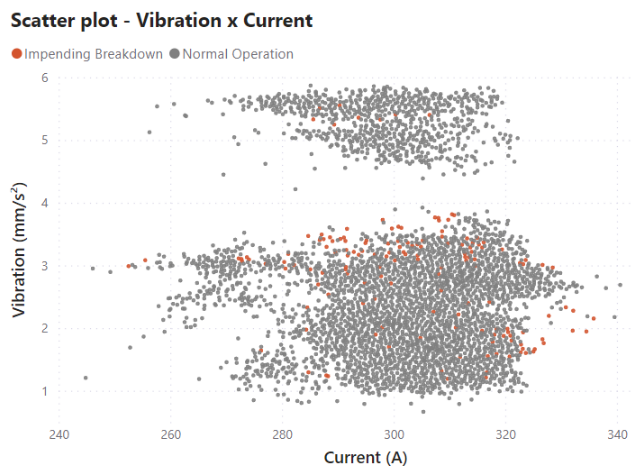

Fig. 2. Scatter Graph between Vibration (mm/s²) and Current(A)

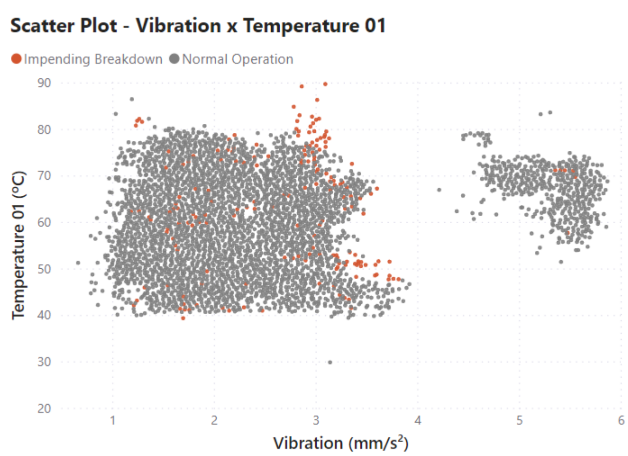

Fig. 3. Scatter Graph between Vibration (mm/s²) and Temperature(ºC)

Visually, it is possible to see that there are areas with a concentration of data categorized as "Imminent failure" in the scatter plots. It will be part of the machine learning model work to simultaneously map these areas through the decision tree forest so that new data can be properly categorized. In addition, some observations can be made:

- The data exhibits imbalanced classes, with the "Normal Operation" class being significantly more prevalent than the "Impending Breakdown" class. This class imbalance can pose challenges for the model, including issues with overfitting and low performance. To address this concern, it is important to implement a balancing strategy.
- The measurement ranges of the input data vary greatly from each other, while some inputs vary from 50 to 70, and others vary from 0 to 500. To improve the analysis, the database will go through a normalization process.

5) Balancing the classes: As mentioned earlier, there is an imbalance in the number of labeled classes in the dataset, as the "Normal Operation" class has over 6,000 registers more than the "Impending Breakdown" class. This imbalance can cause problems for the model and needs to be addressed. For the specific application, three Random Forest Classifiers were trained. The first classifier was trained using the original and imbalanced dataset, while the other two were trained using a balanced dataset.

The first balancing strategy involved randomly removing data inputs from the majority class, aiming to achieve approximate equality between the two classes. The second balancing strategy involved generating synthetic samples in the feature space of the minority class by interpolating between neighboring instances. This technique helps increase the representation of the minority class in the dataset. Both approaches makes the number of occurrences in each class equal, but the last strategy proved to be the most effective, yielding the best results.

6) Data normalization and cleaning: For data cleaning and normalization, we used the preprocessing functions from the *sklearn* library. This library provides commonly used functionalities in data processing steps. This simple step of data normalization ensures that analyses using input data with vastly disparate intervals will be based on an appropriate data model. Also, any input with NaN (Not a Number) values were removed from the dataset.

At the end, there is a database of the selected readings, with approximately 23,000 inputs corresponding to the hourly average values collected from April 2020 to April 2023. Once the input data is ready to be fed into the chosen machine learning model. Machine learning modelling is also done using Python, using the *sklearn.ensemble* library, specifically the set of functions that works with categorization using *Random forest*, called *RandomForestClassifier*.

For the application of this work, the dataset was split in a ratio of 0.8. With the parameters defined and the data set division determined, the model was trained and tested.

B. Implementation

After being trained, the model is ready to predict and categorize new data and be used by the company. For the company's validation, the reading data collected from January 1st, 2023, to March 31st, 2023 was used. During this period, 5 failures with that could have been avoided happened to the equipment. This new dataset went under the same preprocessing steps and was used as input for the trained model. It is the function of the model to classify the new input data as either "Normal operation", written as 0, or "Impending Breakdown", written as 1 by the algorithm.

IV. Results

A. Performance Parameters Achieved

The performance of the machine learning model is usually measured by the percentage of test data that is correctly categorized by the model. For this application, we check the model's F1-score when attempting to predict breakdowns 24 hours in advance. Therefore, for each breakdown, the model must categorize the 24 hourly reading sets as "Impending Breakdown".

The F1-score achieved for the class of interest (Impending Breakdown) was 19.22% for the model trained without the balancing step. The F1-score was increased to 33.58% after applying the Oversampling technique to the test dataset. This Score value is attributed to two factors:

- The inherent difficulty in detecting the selected failures, which were also not detected by the maintenance team of the plant in question.
- The difficulty of the model in meeting the demand imposed during the data labeling to have a 24-hour horizon for fault detection. In other words, although the model can accurately detect the faults, it is unable to extend the detection horizon for the proposed time.

B. A practical approach to the results

To expand on the analysis described in Section IV-A, Figure 4 presents the results of the model in comparison to the breakdowns that occurred from January to March of 2023.

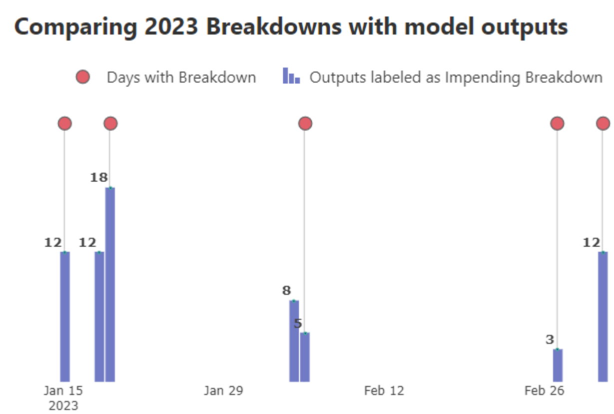

Fig. 4. Comparison between interventions and warnings generated by the model.

The model promptly detected all the failures that occurred this year. However, many of them did not reach the mentioned detection horizon value (labels on the bars can be read as hourly alerts generated before the breakdown), with failures being detected with time of early warning ranging from 3h to 30h, depending on the failure. This indicates that the model is efficient in detecting breakdowns, but we were overly ambitious with the 24-hour horizon for breakdown detection, and a shorter horizon would be more realistic. This is also the reason why the f1-Score is low.

However, alerts hours before the breakdown would be a significant advancement, considering that none of the shown breakdowns was detected by the maintenance team of the industry in question (the equipment in question is critical to the industry and has a robust maintenance strategy).

By detecting the problem early, even just hours before a breakdown occurs, the costs of corrective equipment maintenance can be avoided, and the impacts of equipment downtime can be minimized. This could effectively reduce factory downtime and prevent the loss of tens of thousands of dollars annually. Implementing the developed tool in the plant can lead to significant savings, also ensuring long-term sustainability and competitiveness in the market.

V. Conclusion

Predictive maintenance with automated systems can benefit industrial plants and enhance their processes, but industries slowly implement such innovations. Detecting early signs of functional loss in electrical equipment can aid the forecasting of imminent failures. The prediction and categorization of failures using machine learning models offer numerous advantages to industrial maintenance, spanning from resource conservation to minimizing equipment breakdowns. Machine learning models, however, were rarely employed to forecast imminent failures in electrical industrial centrifuge which are critical equipment commonly in industries whose failure can cause significant losses.

The data used in this work was an innovative valuable resource. It originally consisted of millions of raw monitoring data of an electrical industrial centrifuge extracted from the automation system of a large-scaling operating plant. Subsequently, trained technicians have carefully overseen the data curation over a period of six months. The curated data not only provided valuable insights to improve operational and maintenance processes but also enabled us to train machine learning models to forecast imminent failures in electrical industrial centrifuge.

Training a machine learning model to forecast failures in electrical industrial centrifuge is excessive challenging given the scarcity of historical failures. Such critical equipment, whose failure can cause significant losses, yearly undergoes hundreds of preventive maintenance interventions. In our dataset, only 757 failures, out of $23,000$ total samples, occurred during the 3-yr period. Such extreme imbalance poses a disproportionate challenge to machine learning models to forecast failures. Our trained Random Forest model was capable to detect faults that went unnoticed by the maintenance team, triggering alarms hours before the actual failures occurred.

By integrating machine learning into the predictive maintenance strategy, the industry will reduce factory downtime and mitigate the significant costs associated with corrective maintenance. Such proactive approach enables timely interventions and preventive actions, ensuring smooth operations and optimal equipment performance, ultimately leading to improved productivity and safety.

The Random Forest model trained in this work establishes a baseline F1-score of 0.34 for forecasting failures in electrical industrial centrifuge that can be further improved. In future works, we plan to categorize failure into different classes, enhance missing data imputation, train deep learning models capable of capturing more complex time-dependent patterns, and lastly employ online learning that continuously learns from new data.

References

[1] T. P. Carvalho, F. A. A. M. N. Soares, R. Vita, R. da P. Francisco, J. P. Basto, and S. G. S. Alcalá, "A systematic literature review of machine learning methods applied to predictive maintenance," *Computers Industrial Engineering*, vol. 137, p. 106024, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0360835219304838

[2] M. Bevilacqua and M. Braglia, "The analytic hierarchy process applied to maintenance strategy selection," *Reliability Engineering & System Safety*, vol. 70, no. 1, pp. 71–83, 2000.

[3] P. S. De Almeida, *Manutenção Mecânica Industrial–Conceitos Básicos e Tecnologia Aplicada*. Saraiva Educação SA, 2018.

[4] J. Koochaki, J. A. Bokhorst, H. Wortmann, and W. Klingenberg, "Condition based maintenance in the context of opportunistic maintenance," *International Journal of Production Research*, vol. 50, no. 23, pp. 6918–6929, 2012.

[5] A. Diez-Olivan, J. Del Ser, D. Galar, and B. Sierra, "Data fusion and machine learning for industrial prognosis: Trends and perspectives towards industry 4.0," *Information Fusion*, vol. 50, pp. 92–111, 2019.

[6] Z. M. Çınar, A. Abdussalam Nuhu, Q. Zeeshan, O. Korhan, M. Asmael, and B. Safaei, "Machine learning in predictive maintenance towards sustainable smart manufacturing in industry 4.0," *Sustainability*, vol. 12, no. 19, p. 8211, 2020.

[7] R. Liu, B. Yang, E. Zio, and X. Chen, "Artificial intelligence for fault diagnosis of rotating machinery: A review," *Mechanical Systems and Signal Processing*, vol. 108, pp. 33–47, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0888327018300748

[8] W. Zhang, D. Yang, and H. Wang, "Data-driven methods for predictive maintenance of industrial equipment: A survey," *IEEE Systems Journal*, vol. 13, no. 3, pp. 2213–2227, 2019.

[9] Z. M. Çinar, A. Abdussalam Nuhu, Q. Zeeshan, O. Korhan, M. Asmael, and B. Safaei, "Machine learning in predictive maintenance towards sustainable smart manufacturing in industry 4.0," *Sustainability*, vol. 12, no. 19, 2020. [Online]. Available: https://www.mdpi.com/2071-1050/12/19/8211

[10] S. Ayvaz and K. Alpay, "Predictive maintenance system for production lines in manufacturing: A machine learning approach using iot data in real-time," *Expert Systems with Applications*, vol. 173, p. 114598, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417421000397

[11] N. Kolokas, T. Vafeiadis, D. Ioannidis, and D. Tzovaras, "Forecasting faults of industrial equipment using machine learning classifiers," in *2018 Innovations in Intelligent Systems and Applications (INISTA)*. IEEE, 2018, pp. 1–6.

[12] I. Amihai, R. Gitzel, A. M. Kotriwala, D. Pareschi, S. Subbiah, and G. Sosale, "An industrial case study using vibration data and machine learning to predict asset health," in *2018 IEEE 20th Conference on Business Informatics (CBI)*, vol. 1. IEEE, 2018, pp. 178–185.

[13] J. P. Simmons, L. D. Nelson, and U. Simonsohn, "False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant," *Psychological Science*, vol. 22, no. 11, pp. 1359–1366, 2011, pMID: 22006061. [Online]. Available: https://doi.org/10.1177/0956797611417632