Most Arina Afrin

AI-Enhanced Predictive Maintenance for Industrial Robotics using Sensor Data Analysis and Machine Learning Approach

School of Technology and Innovations
Master of Science in Technology
Master's Programme in Robotics

Vaasa 2026

UNIVERSITY OF VAASA			
School of Technology and Innovations			
Author:	Most Arina Afrin		
Title	AI-Enhanced Predictive Maintenance for Industrial Robotics using Sensor Data Analysis and Machine Learning.		
Degree:	Master of Science in Technology		
Degree Programme:	Master's Programme in Robotics		
Supervisor:	Mohammed Elmusrati		
Co-Supervisor	Elham Ahmadi		
Year:	2026	**Pages:**	64

ABSTRACT:

Modern manufacturing is progressing towards Industry 4.0 and Robotics 4.0, which puts the industrial robots in the focus of contemporary manufacturing. These sophisticated robots keep on producing multi-sensor rich data, which depicts their internal kinematics, forces, torques, and joint dynamics, which is useful in detecting anomalies that may result in failures in time. However, even with access to these types of data, fixed-interval or reactive maintenance plans continue to be implemented by many industries, which results in unwarranted downtime and expensive interruption. These gaps, combined with the necessity to have the predictive maintenance schemes that are reliable and capable of converting the raw sensor images into actionable intelligence based on machine learning. We focused on creating an AI-enhanced predictive maintenance model of industrial robotics, based on internal robot states and machine learning algorithms. The paper concentrates on data pre-processing, feature engineering, anomaly detection, Remaining Useful Life (RUL) prediction, and designing a data-driven maintenance decision-support system that includes and operates synchronized robot sensor streams three tools of kinematics, joint positions, and force/torque measurements then proceed to extract time-domain features, and correlation-based features to identify patterns of early degradation. Exploration of sensor relationships and structural separability of normal and abnormal conditions was performed by use of Principal Component Analysis (PCA) and correlation matrices. Logistic Regression, Support Vector Machine (SVM), Random Forest and XGBoost were four classical machine learning models trained and tested on chronological splits. Accuracy, precision, recall, F1-score, ROC-AUC, PR-AUC were utilized as a measurement of model performance. A decision-support system was subsequently developed to transform predictive results to a form of maintenance recommendations.

The patterns in the robot signals are well established and consistent between normal behavior where random Forest and XGBoost performed better than any of the remaining models. The study proves scientific soundness of applying the classical machine learning models specifically the Random Forest and the XGBoost on predictive maintenance in industrial robotics. It established internal robot data when adequately preprocessed and engineered provides strong predictive capability in detecting faults early and estimating their life. The decision-support framework suggested is a feasible avenue that industries can use to minimize production downtimes, enhance their safety, and enhance the transition to Robotics 4.0. This paper will add fresh empirical data, methodological soundness and practical solutions to the expanding research on AI-enhanced maintenance in intelligent manufacturing settings.

KEYWORDS: Predictive Maintenance, Industrial Robotics, Machine Learning, Sensor Data Analysis, Remaining Useful Life (RUL) Prediction, Industry 4.0/ Robotics 4.0

Contents

List of Figures

List of Tables

1 Introduction

1.1 Background of the study

The global trend of transitioning to Industry 4.0 has already been reshaping the landscape of industrial production that now focuses on the automation, digitalization, and real-time data-driven processing. There are no longer such things as programmable machines which might have been rehearsed but rather are getting more intelligent robots capable of sensing the world and acting independently. Ranged under the title of Robotics 4.0, Gao et al. (2020) under the nomenclature was a robot that smoothly communicates with intelligent manufacturing machines due to sensory technology innovations, machine learning, and computerized physical interconnection. These robotic intelligent systems are currently being increasingly applied in the modern factories to meet the more demanding needs on precision, flexibility and productivity.

Industrial robots are also provided with multiple internal sensors joint encoders, force/torque sensors, accelerator, gyroscope and tool-position trackers, and therefore can monitor their movement and loading (Li and Liu, 2019). These domestic sensor data will give us a rich feature space of physical matters revealing subtle form in robot action. With a proper working robot, pattern of its internal sensors will be known and predictable. However, there are changes, conspicuous deviations to the patterns in case some mechanical damage, wear, overloading or collision occurs. Other studies, including, but not limited to, the ones by Izagirre et al. (2020) and Sujatha et al. (2022), have noted that sensor data, like this one, will enable us to detect a fault well before actual failure sets in.

The advent of intelligent robotics is ushering in the possibility of predictive maintenance – the concept of having information available to know when power equipment will require repair, with machine-learning-based data analytics. Predictive maintenance is a valuable part of smart manufacturing that enables an industry to cut down on time that equipment’s spent in standstill, plan maintenance in a smoother and lower operating expenses. Predictive maintenance has been recognized as an essential issue in Industry 4.0 in numerous works. Indicatively, a case in point is a review article that has been authored by Carvalho et al. (2019), and it also emphasizes the significance of machine learning models, whereby it concerns the identification of trends in the behavior of machines. Likewise, it is stated in Çinar et al. (2020) and Zonta et al. (2020) that predictive maintenance is a factor that makes a system sustainable, and that is not contributing to higher productivity.

There are challenges associated with predictive maintenance, however, in robotics. Multi-axis trajectories performed by robots include complex single-axis movements between joints, tools and actuators with kinematic interconnection. Mechanical, electrical and control interdependences do lead to degradation being difficult to discern by simple method of rule or threshold. Further, industrial robots serve in various payload conditions, user inclination and assignment circumstances. Past methods of maintenance such as scheduled schedule maintenance or on-demand fix-based repair can no longer sustain sophisticated robotic operations with high turnover. The success of the industry, according to Dalenogare et al. (2018) or Eylemo et al. (2020), does not depend on automation itself but on smart supervising systems that should constantly modify their performances in the grey zone conditions.

Thus, integrating data and machine learning has given a promising direction of machine health inspection by robots working in industries. Machine learning models have the capability of detecting anomalies and predicting possible failures in maintenance planning, as they observed the behavior of different sensors under normal conditions and under faults. It can also be compared to the results of Borgi et al. (2017), who demonstrated the use of data analytics as a main trend to predictive maintenance in robots, and that of Kanawaday and Sane (2017), who reiterated the value generated by IoT-generated sensor data as well as predictive algorithms. As an example, in a manufacturing plant time equals money and these predictive systems can mean significant sleep time reduction and improved productivity.

1.2 Problem Statement

Although the information available in robot sensors is increasingly growing, regrettably most industries struggle to build their predictive maintenance of robots. One of the main difficulties is due to the complexity of sensor data. The HF multi-dimensional data related to MBs are mined by industrial robots and require pre-processing to be attentively synchronized. As mentioned in Ullah (2019), sensor measurements cannot be used to detect anomaly or extract features unless it is brought into alignment and cleared of noise. Many of these studies operate with simulated data (or a small number of sensor measurements and cannot be easily applied in the real industrial application).

The other concern is ignorance of which feature set of engineered features is the best to identify early degradation. While time-domain features are used widely, frequency domain and correlation-based features usually have more information about the obscure changes of robot behaviors. As Michau et al. (2019) note, the typical feature selection methods are useless in identifying crucial trends in high-dimensional sensor data. Therefore, the processes need full-fit engineering approaches that capitalize on having multiple family of features to capture a short-run and in the long-run dynamics of the behavioral control of robotic systems.

26 To be fitted to models, model choice is also a factor. Even though predictive maintenance is conducted by machine learning in several experiments no consensus is made regarding the most effective methods in relation to industrial robots. To some of them, deep learning models are used and other classical machine learning algorithms: Çınar et al. (2020) and Zonta et al. (2020) suggest that classical machine learning algorithms can be expected to outperform deep learning models on some datasets and/or regimes: this is because they are easy to interpret, need less data and/or are more robust than deep learning models. Nonetheless, very little was done to compare the classical models among themselves on actual multi-sensor robotic data. This forms a gap in knowledge of comparison of models concerning their performances in terms of anomaly detection and RUL estimation.

The last challenge is that since predictive models may perform well, companies do not always have a systematic manner of transforming their predictions into maintenance interventions. And even with no decision-support system, machine learning results such as the appearance of predicted anomalies based on forecasting, or the time to failure based on estimation, are somewhat abstract: they do not tell the operator anything useful. This constrains predictive analysis to become useful for the industries.

The current thesis is meant to address this gap by developing an AI-enabled predictive maintenance model, which will contain a sensor combined data preprocessing and feature selection phase, machine learning development composed of an anomaly detection procedure, RUL prediction process and decision-making assistant recommendation application.

1.3 Research Aim

This study will attempt to create an AI-enhanced predictive maintenance system in industrial robotics through an analysis of feature engineering and classical models of machine learning to predict anomalies and Remaining Useful Life. The framework will enhance precision in fault detection, minimize downtime and facilitate smart maintenance scheduling.

1.4 Research Objectives

The analysis and pre-processing of multi- sensor robot data (e.g. tool kinematics, joint positions, forces or torques) is the first objective of this piece of work. It incorporates the cleaning, stream normalization and synchronization process to ensure that inputs into the model are of high quality. The second is feature selection, that is time-domain features, frequency-domain features and correlation-based features. This helps find the patterns that are a sign of degradation, instability or unusual functioning. Thirdly, we also create classical machine-learning models such as Logistic Regression, Support Vector machine, Random Forest and XGBoost both to detect anomaly and to estimate RUL. The final objective is to determine the performance of these models using different measurement scales of evaluation including accuracy, precision, recall, F1-score, ROC-AUC, PR-AUC in classification. The fifth objective is to develop a model-based decision-support model of maintenance of worn infrastructure based on the findings that were in this research.

1.5 Research Questions

The initial research question will be whether the internal sensor signals of a robot can serve as good indicators of anomalies/collisions during a normal operation. The second question is to find out the sensor modalities and engineered features that most accurately predict early-stage damage and degraded performance. Question three will seek to learn how classical machine learning algorithms perform both anomaly detection and RUL prediction using multi-user robot data. The fourth research question is whether it is possible to use uncertainty-sensitive decision-making to increase the credibility of predictive-based maintenance recommendations. The final question researches the extent of these models generalizing to the users of robots and whether they are consistent in testing on new information or not.

1.6 Significance of the study

The work of the present research is scientific and practical. It has been scientifically demonstrated that the multi-sensor internal signals of complex robot, following certain gentle pre-process, possess their distinctive shape to differentiate a normal and abnormal behavior efficiently. This is consistent with the PCA structure and correlation analysis we found in the results. It further suggests that the combination of time-domain and frequency-domain with the correlation's features can enhance the prediction accuracy of the machine-learning models in the same way as discovered by Michau et al. (2019) and Izagirre et al. (2021).

The other applicable scientific contribution is the most suitable classical machine learning model of predictive maintenance to be compared in robotics. Table 4 shows that the performance of both Random Forest and XGBoost in all the evaluation criteria is relatively good, which is in line with the earlier research stating that the decision-tree based models can work effectively with structured industrial data (Çinar et al., 2020; Zonta et al., 2020). Also presented in the results is that the Logistic Regression is not acting well with high-dimensional robotic information whereas the Support Vector Machines are successful in recalling yet hypocritical in false alarm.

In terms of application, application perspective of the work introduces end-to-end predictive maintenance process that can be adopted by industries by direct implementation. The decision support system converts model predictions to operational maintenance decisions which may be utilized by the managers to inspect, plan repair and replacement at better time at lesser cost. This especially applies in the production facilities whereby every minute a robot is not working means that a lot of money is wasted. The smart factories and collaborative robotics work by Javaid et al. can also be relevant to practice. Nikolakis et al. (2019), Eylemo et al. (2020) indicate that problems with intelligent maintenance implementation are escalating and necessitate demand.

Another way in which the research contributes to the future of Robotics 4.0 is that machine 4.0 can be implemented in robot health monitoring. The findings confirm the hypothesis brought forward by scholars (e.g., Xu et al. (2017) and He et al. (2018) that justify the essential character of AI, big-data analytics and digital twins of the forthcoming generations robotics and smart manufacturing systems.

1.7 Scope of the Study

This research paper focuses on the multi-sensor internal robot data to anomaly detection using classical machine learning models as well as RUL prediction which can be used to evaluate degradation and prediction. Offline training and testing are carried out with real data of the robots. There is aimed to evaluate how classical machine learning would be useful in predictive maintenance of robotics.

2 Literature Review

2.1 Introduction

The high pace of the evolution of transformational technologies like machine learning, sensor networks, and data analytics made predictive maintenance (PdM) strategic in the contemporary industrial environment. It also involves the determination of predictive models that can be used to predict when a machine is going to break to ensure that the maintenance process is carried out at the appropriate time, thus minimizing the downtime and maintenance overhead. PdM in industrial robotics, where efficiency is of most significance of an application, can be taken into very high consideration to make sure that the robots' systems do not run into any disruptions. The use of machine learning (ML) models to perform PdM, Carvalho et al. (2019) and utilization of machine learning (ML) models allow prediction of failures and actionable resources to be used with an impactful use of resources. We explored theories and methodologies that basis predictive maintenance in robotized industry with a particular concerning the AI revolution and machine learning in PdM, and the applicability of sensor data and challenges of predicting failures and managing maintenance interventions. Prediction list operation the maintenance operation allows utilizing data analytics and machine learning to forecast equipment failures before they occur. In contrast to reactive maintenance, in case of the failure of equipment or preventive maintenance with scheduled service, predictive maintenance is premised on data that define how best it will be utilized. By the time of Industry 4.0, the shift to predictive maintenance will become the key as the intelligent factories will require cutting-edge methods that will result in more efficient processes and increase production.

The benefits of predictive maintenance to manufacturing are enormous. It also reduces the amount of preventive maintenance that must be done and assists in eradicating the downtime, not to mention minimizing the possibility of failure of mechanical breakdowns that may lead to the loss of production that is costly to procure. Predictive maintenance also ensures that machinery is available longer without the need to maintain it on a schedule since it must only be maintained when needed. Industry 4.0 technologies facilitate this plan in some ways. IOT devices are connected to industrial equipment to retrieve real-time information about the operational parameters of the machines such as temperature, vibration, pressure etc. Data obtained on this is utilized to monitor the condition of the machines and likely when they will fail. Since the IoT devices generate significant amounts of data, the big data analytics allows processing and analysis of the information to identify patterns that may lead to indicators of equipment breakdown. Cloud computers have the capability of storing data in high quantity and processing them on large quantities of machines that are spread in various places. It also helps run a prediction model in real-time. The systems are computer algorithms plus hardware to monitor and control the actions of robot devices. Under PdM, CPS provides feedback that is in real-time necessary to settle the faulty components and actions of robots.

2.2 Literature review comparison

Table 1. Review literature comparison of different methods, findings, and limitations

Paper Author	Data type	Models	Metrics	Findings	Gap
Borgi et al. (2017)	Time-series sensor data (electrical current) + spatial position	Multiple Linear Regression (MLR)	R^2, RMSE, MSE	R^2 = 0.885 and RMSE = 0.0916, indicating excellent predictive performance. And Feasibility of using only electrical data for accuracy error prediction was proven.	Data-driven fault detection methods require historic failure data. Not validated in real industrial settings with variable loads or environmental disturbances.
Carvalho et al. (2019)	Systematic Literature Review (36 selected papers)	Random Forests (RF), Support Vector Machines (SVM), Artificial Neural Networks (ANN),	RMSE, MAE, R^2, Accuracy, Precision	Random Forests (RF) – most frequently used. 89% of studies use real data; only 11% use synthetic data	Do not compare their proposed approach to different ML algorithms. Few papers employ newer methods like deep learning. Limited availability of public datasets for benchmarking PdM applications

		Deep Learning (LSTM, CNN)			
Sujatha et al. (2022)	Time-series sensor data	Decision Tree (DT)	Efficiency (78%), Data split (70%, 30%)	Proposed system achieves 78% efficiency, which the authors claim is improved over existing methodologies. methodology improves reliability of smart IoT systems	Industry 4.0 implementation is not yet widespread. The paper uses only one ML model. No comparison with other models
Kanawaday & Sane (2017)	Industrial slittin g sensor data (4 arms, 7 per side)	Naive Bayes, SVM, CART, ARIMA	Accuracy, precision , recall, RMSE, MAE, AIC, BIC	98.69% accuracy, Decomposition increasing trend in residues, Combining ARIMA forecasting + supervised for predict quality failures	Most systems only monitor current state (condition monitoring) but do not predict future failures, unexpected machine stops cause major financial losses
Luo et al. (2020)	Time-series sensor: dynamometer (cutting force), accelerometer (vibration), acoustic emission (AE) sensor	Random Forest, Support Vector, Decision Tree, Linear Regressio n	RMSE, MSE, Accuracy, precision , recall	Prediction errors were reduced from 9.51-19.59%. Outperformed Linear Regression, Decision Tree, and SVR	methods do not consider the time-varying status and consistency o f equipment during its life cycle
Elsisi et al. (2021)	sensor data	Decision Tree. XGBoost, Logistic Regressio n	Accuracy, Precision , Recall, F1-score	The proposed infrastructure addresses both security (fake data detection) and reliability (data loss detection) simultaneously .	Industry 4.0 faces main challenges: reliability and security of IoT data. Industrial environment affects meter efficiency (temperature, humidity, noise signals) – not addressed in existing systems
My thesis	Time series sensor data	Logistic regression , Random Forest, SVM, and XGBoost	Accuracy. Precision , Recall, F1-score, ROC-AUC, PR-AUC	Accuracy, precision, recall and F1-score are high, XGBoost and Random Forest models perform better. Also found effective RUL estimation	Data was of a specific environment and specific tasks using robots. We also did not consider models of deep learning

2.3 Industrial Robotics and the Role of Predictive Maintenance

Industrial robots have altered tremendously the face of manufacturing as they have facilitated the aspect of productivity, accuracy and flexibility. However, regardless of all their benefits, robots do not stand out in terms of mechanical wear and tears, and it may lead to failures. This is the price of these failures – not only in repairs, but in manufacture. This fact hence emphasizes the relevance of maintenance strategies in a bid to ensure that there is a sustained and effective operation. There are several ways in which robotic systems may fail, and these are: motor error, joint offset, and sensor failure. To minimize the system downtime and catastrophic failures, it is of paramount importance to diagnose these issues at an early stage to minimize the duration. The typical way that maintenance is performed in the traditional manufacturing setting is reactive: that is, only when something has gone wrong. This, however, is ineffective and costly. Predictive maintenance transforms the world to be not reactive in reporting the problems faced, but predicting them even before it happens, and that the maintenance staff may make efforts even before a breakdown. Industrial robotics has three maintenances:

- **Reactive Maintenance**: This is a plan that entails repairing the equipment once it has failed. It is the most expensive and inefficient technique, because it causes unwarranted downtimes and last-minute repair.
- **Preventive Maintenance**: It is the system that performs the scheduled maintenance according to the time or number of cycles based on the usage of that equipment irrespective of the present state of the equipment. Although it can minimize the downtime compared to the reactive maintenance, it might be inefficient because maintenance is done even when the equipment has not failed.
- **Predictive Maintenance:** This policy involves the use of data analytics and machine learning to predict occurrence of failures. PdM can also be used to maintain robotic systems, only when necessary and necessary, because constant monitoring of the systems ensures that unnecessary downtime is cut and equipment can have an extended life.

2.4 Machine Learning in Predictive Maintenance

Machine learning (ML) is the most important technology to implement predictive maintenance since it can process large amounts of data generated by sensors on industrial robots. ML models are trained using past data to identify trends that can predict failures. It is possible to categorize these models into general types, that is, 1. Supervised; 2. Unsupervised; and 3. Reinforcement learning. Supervised Learning will train Machine Learning models of known labeled data sets, where input-output pairs (such as fail or not fail) are specified. The common supervised predictive maintenance methods are support vector Machines (SVM), Random Forest, and XGBoost. Those models will be to detect the states of equipment based on sensor information signals. The use of Supervised Learning in contrast to unsupervised Learning is in case the labeled data is in place. It also helps in anomaly detection of sensor information by learning patterns of abnormal conduct. Algorithms, such as K-means clustering, Isolation Forest and Autoencoders, are normally utilized in anomaly detection. The other one is reinforcement learning that learns how to take measures including a response provided to him by the environment. It is not common in PdM, but it is used to identify additional information on the most effective maintenance schedules and the decision-making procedures.

Case studies of predictive maintenance have demonstrated that the use of machine learning models can be used to improve operational efficiency. As an illustration, in the study by Sujatha et al. (2022) when machine learning was used to estimate failures of industrial robots and to reduce downtime, Dalenogare et al. (2018) proved that predictive maintenance models could be applied to plan industrial maintenance tasks and reduce their costs.

2.5 Sensor Data Analysis in Predictive Maintenance

The provision of timely information on the state of industrial robots has been realized by sensor data, which serves as the basis for predictive maintenance. Industrial robots are equipped with numerous sensors, which include force/torque, kinematic and vibration sensors that generate measurement data that can be used as fault prediction. This data must be processed in advance and converted into features that can be comprehended by machine learning models. One process of processing sensor data is through feature extraction. It is the stage of transforming raw sensor data into a code set of features that summarize the underlying patterns as leading indicators of health status. The features may be classified into features of time-domain which refers to the characteristic values of the sensor data of the sensor in terms of mean, variance and skewness with time. Frequency properties can be acquired through Fourier transform of time domain signal which is applicable particularly in periodic form and vibration phenomenon pertaining to machinery fault. Sensor-Feature-Correlation: It is a feature which quantifies correlations between variables among sensor signals, and it is therefore appropriate that it can be used to find interactions between the various components of a robot. In addition, sensor integration and sensor fusion are also applied, to incorporate the information of diverse sensors to get a bigger picture regarding the wellbeing of a robot. New data sources can be used to enhance predictive maintenance models and make them more accurate and reliable.

2.6 Remaining Useful Life (RUL) Prediction

The secret of predictive maintenance lies in RUL. RUL is time prediction of breakage of a robot or one of its components. Precise calculation of RUL leaves a schedule on which a maintenance team can make an intervention on; therefore, the repairs or replacement is not an unexpected occurrence. There are several methods for RUL prediction: regression model, deep learning model and ensemble method. Such regression models as the Random Forest and the SVM give approximate remaining equipment life based on sensor readings. Deep learning-based (in particular, LSTM) networks are applied to tackle the time-series nature of sensor signals, where long-term relationships between sensor beliefs and sensor readings and the RUL can be considered. The models have been highly encouraging in predicting RUL of the complex systems as in the case of industrial robots. However, the non-linear process of degradation, sensor noise, and unlabeled failure data are some obstacles that are faced by RUL prediction. Current research is directed at the creation of models which are more resistant to such problems and can yield meaningful projections in the real industrial setting.

2.7 AI and Machine Learning for Fault Classification and Anomaly Detection

The second type is the fault classification anomaly detection that decides of the type of odd behavior in place and the third determines the difference between the normal state and the failure. Machine learning can also contribute greatly to these processes by autonomously searching patterns in sensors that may be indicative of faulty conduct. An example of this is the machine learning models running to distinguish normal and abnormal sensor readings, to detect faults. In the case of classification works, the most used tools are Decision Trees, Logistic Regression or Support Vector Machine. In the case of non-existent labels (i.e., the problem is unsupervised), such methods can also be acquired with the help of isolation forests and autoencoders. The schemes of the AI-based decision support systems discussed in this paper could be applied to target maintenance operations based on the evaluation of the severity of defects and provide a solution to repair or to change the component. Such frameworks help maintenance teams to take data-driven decisions in order to increase operations efficiency and to avoid downtime.

2.8 Emerging Trends and Technologies in Predictive Maintenance

With Industry 4.0 determining the future of the manufacturing sector, emerging trends and technologies are assisting in reinventing predictive maintenance in industrial robotics. The integration of such enhancements will introduce increased accuracy, enhanced dependability and enhanced escalation of predictive maintenance frameworks with more enhanced degrees of insights and choices.

2.8.1 Digital Twins and Predictive Maintenance

Among the new technologies emerging to predictive maintenance factors, digital twins are clearly what will be proclaimed as industry 5.0. A digital twin is an artificial representation of a real object, in our case an industrial robot, from its behavior and state. Such sensors and monitors guarantee that information about the physical robot is received continuously to keep the digital twin informed of the most recent facts, this enables the search for potential failure in performance and health at a time scale. With this technology, individuals can simulate and test all kinds of failures and maintenance plans with no danger whatsoever. Using digital twin technology, manufacturers will know when equipment is prone to malfunction and can plan service long before it occurs, hence avoiding unexpected downtimes. Digital twins also facilitate the action of improving capabilities of robots after some time because they detect literature preferences. Specifically, Izagirre et al. (2021) pointed to the potential of digital twins in RUL prediction and preventive maintenance schedule as a tool to improve operational efficiency.

2.8.2 Edge Computing and Real-Time Data Processing

As the volume of information produced by industrial robots increases, it is significant to process and analyze the information as it is produced. Edge computing also eliminates the data processing intermediate in such a way that computing is done at the source of data, at the robot, or near the robot. This technology reduces scale of big data transfer to central servers, which causes reduced latency and quicker response time that is leading and is conductive to maintenance decisions.

Real-time predictive maintenance activities prove especially useful in cases when an emergency measure is required to prevent equipment breakdown. When detected on sensor data locally and feedback is given in a near-real time, edge computing enables anomaly detection and RUL prediction to be achieved rapidly, thereby enhancing the performance of predictive maintenance systems. As an example, Xu et al. (2017) studied how industrial robots can be used alongside edge computing and IoT to provide the opportunity to form a real-time understanding of the health conditions of the machine to conduct proactive maintenance. Edge computing enables real-time data processing; cloud computing is still a part of the wider predictive maintenance ecosystem. Cloud: Cloud based platforms which can pool, store and process huge data influx from robots distributed in diverse environments. The scalability of cloud computing is high hence it is practical that the predictive maintenance systems can handle data of very large number of robots, factories or production lines at the same time.

2.8.3 Advanced Sensors and Multi-Sensor Integration

Predictive maintenance systems depend on the accuracy and utility of the sensor data that is provided by the sensor. There is beginning to emerge next generation sensors with a greater level of granularity and precision in the data that is utilized in predictive maintenance. As an example, vibration sensors are commonly employed to sense mechanical failure before it happens, force/torque sensors used to provide an indication of excessive stress or misalignment on robotic joints and temperature sensors to indicate overheating. Besides, sensor fusion can enhance the accuracy of prognoses when it comes to the maintenance of the robots due to the increased understanding of their condition. Multi-sensor fusion algorithms can be able to combine the data obtained by nonhomogeneous sensors to eliminate noises, minimize ambiguity and improve the precision of prediction generally. This data fusion between multiple sensors is necessary to make the analysis more comprehensive and reliable and is critical in determining presence of faults that cannot be easily identified by a single sensor.

2.8.4 AI and Machine Learning Advancements

The AI and ML are evolving at the rate of a neck break, hence, establishing novel predictive maintenance methods. The outlook of recent improvements in deep learning is promising and useful in the automatic extraction of different features and higher accuracy of predictive models. As an example, the analysis of the sensor information with the help of Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN), to identify the trends that signal failure.

One particularly useful type of these models is the well-known recurrent neural network (RNN), the Long Short-Term Memory (LSTM), which is particularly useful with time-series data, which often features prominently in robotic systems.

Additionally, the new reinforcement learning techniques are being explored in the backdrop of maintenance schedule molding and choice making. The RR models can use decisions that have been made previously in maintenance and adapt to a dynamic environment. The improvements in AI have instilled some potential importance in the operation of predictive maintenance systems, making it more intelligent, dynamic and autonomous.

2.9 Gaps in Literature and Future Directions

The predictive maintenance in relation to industrial robotics made great progress, though there are still great gaps in the existing literature that need to be addressed so that the effectiveness and usability of the PdM systems could be enhanced.

2.9.1 Scarcity of Multi-User and Multi-Robot Datasets

The main weakness of the existing studies is the utilization of comparatively homogeneous data materials comprising of borne applications and robots in a restricted set of circumstances. Majority of the former research is founded on controlled, single-user datasets which may not effectively reflect the diversity of real-world workplaces. To ensure that predictive maintenance models can be applied to different environments and types of robotic systems, multi-user, and multi-robot datasets must be created. In future, it is important to develop large cross-type datasets involving multiple types of robots and industries with various operating conditions. It will make the researchers have a chance to develop improved in real world like models capable of facing complexity encountered in the real environments and Variability.

At this point, there is no generally accepted standard reference point of comparison between robots/industries/sensor setups regarding predictive maintenance models. Thus, this fact does not enable us to compare the performance of various models and methods. These benchmarks would be applied to compare and assess various PdM solutions, and due to them, researchers would be able to identify the most applicable solutions to different applications. To address this gap, it is essential to create benchmark datasets and evaluation metrics. These standards need to reflect the complexity of industrial robotics systems and multiple failure modes, sensor configurations and maintenance.

2.9.2 Limited Use of Uncertainty-Aware Models

Although machine learning-based models have a large potential of predictive maintenance, uncertainty-aware models are not entirely addressed in most of the literature. The only exception is that when these models are calibrated using the right data, they can be used to predict uncertainty and thus become indispensable tools to be used to make decisions pertaining to maintenance. As an example, Monte Carlo Dropout (MC-Dropout) and Bayesian Neural Networks provide uncertainty estimates that could empower maintenance workers to decide whether they need to act based on the uncertainty of an estimate. Future research would expand the usability of uncertainty-aware models to predictive maintenance systems and help in better reliability and safety when making decisions using industrial robotics.

Predictive models for maintenance can predict failures and already to estimation RUL, however they are not always configured towards the consideration of the cost that involves one type of maintenance or another. Under the cost-conscious optimization, the failure, repair and downtime ones are integrated into the predictive maintenance model. This enables the manufacturers to streamline their additional maintenance operations towards the total economic value and the system stability. The research can be continued in future to consider adoption of use of a budget-adaptive decision support system that suggests the best preventive strategy whilst keeping in check the availability of limited resources that are available to industry.

2.9.3 Integration of Emerging Technologies

New opportunities for predictive maintenance are availed by enabling technologies, such as digital twins, edge computing and 5G networks. However, these young technologies have not been fully incorporated in current PdM solutions. The technologies involved may enhance the precision, scalability and real-time reaction of proactive upkeep models drastically. The future directions are inclined towards researching the integration of such emerging technologies towards developing more advanced, adaptable, and real time predictive maintenance systems.

2.10 Decision Support Systems for Maintenance Actions

Once a fault is recognized or a prediction of RUL has been made, the remaining issue is to use these predictions in maintaining systems. The available decision support systems assist in solving the optimal maintenance policy and allocation of resources. There are cost-based maintenance policies used in cost of failure, repair time, and downtime costs where an order on maintenance activities is considered. A fault in risk-based models is defined as severity and the effect it can cause to the system, and this way maintenance crews can determine the corrective activities that need to be executed initially.

2.11 Cross-User Generalization and Model Robustness

After detecting a fault or after having assessed the RUL, one difficulty that remains is to incorporate these estimates in the maintenance choices. Decision support systems are created on which the best maintained approach and resource allocation can be determined. Cost-based maintenance policies consider the cost of failure; the amount of time and money spent on repair and cost incurred during downtimes to determine a maintenance action. The fault is determined by its risk (the importance of the consequence on the system) and this enables the maintenance personnel to decide on which remedial measures should be implemented first, different users, robots, and environments matter to implement them in practice. Cross-user generalization ensures that when the models are transferred to new machines, or a new operator, under different environments, the models are generalized well. Cross-user generalization is not simple since its use in different users is diverse because of the various robot usage patterns, sensor count and environmental factors. On the one hand, approaches that include transfer learning and domain adaptation are under exploration to make models more generalizable by transferring between tasks in which a large set of labelled data exists to much more different tasks instead of training independent models per task.

Despite the development in predictive maintenance for industrial robots, some research gaps are still open. It consists of the lack of multi-user, multi-robot data, the lack of standard benchmarks of predictive maintenance models and the rarely utilized uncertainty-sensitive model. Future studies should focus on these concerns by generalizing modeling and adopting emerging technologies (e.g., digital twins and edge computing), introducing model robustness through the example of meta-learning.

2.12 Conclusion

Combination of AI, Machine Learning and predictive maintenance can change the nature of your maintenance and control of the industrial robotics. However, as predictive models and technologies are developed, some limitations must be mentioned in the literature to implement PdM. The next generation work needs to address the obstacles stemming out of the non-existence of various datasets, absence of standardized benchmarks and the need to adopt uncertainty-conscious mechanisms in such a way that one can eventually be able to develop robust, scalable and cost-efficient predictive maintenance infrastructures. The implementation of next-generation technologies like the use of digital twins, edge computing, and advanced AI will also just increase the potential of PdM to ensure manufacturers can optimize their maintenance plan and introduce sustained operation of industrial robots.

3 Methodology

3.1 Introduction

We examined the appropriateness of ML models in predictive maintenance in an industrial robotic system. It showed the means and methodology of measurement of the anomaly detection technique; Remaining Useful Life (RUL) prediction. We described how the data was collected, what machine learning models we used, feature engineering techniques that were applied and how the models were assessed. It also gives an overview of the decision support toolkit that has been developed to integrate results in actual maintenance decisions.

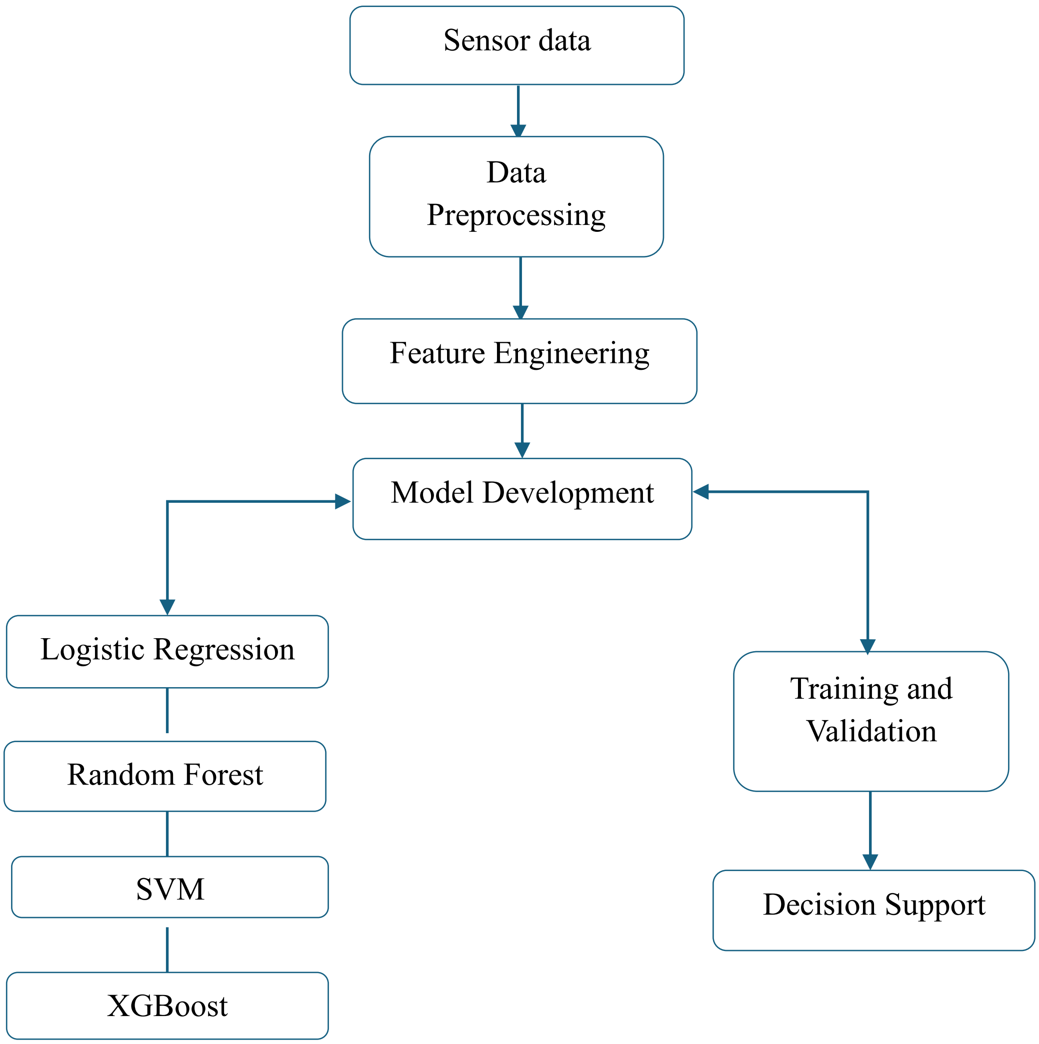

Figure 1. Overview of the Model

We intended to show that AI-based next-generation predictive maintenance models can be applied to industrial robots and that they can produce improvements in these systems in terms of reliability and performance to manufacturing. Towards this end, some machine learning techniques were tried on sensor data of industrial robots. The chapter itself is also organized in the structure of the research in that there is a brief description of the research design, and then everything is followed by the detailed explanation of the analysis of data and the maintenance solutions providing the predictive maintenance.

3.2 Research Design

Our quantitative methodology was in form of prediction modeling that is commonly used in predicting maintenance and detecting anomalies. The aim was to investigate the factor to predict the RUL and failure of robot system by various machine learning (ML) models. The plan itself started by exploring the data that the sensors on the robots transmit to determine the indicators of future failure. The research involved the combination of supervised learning strategies (e.g., the use of historical data training the models of the second type) and classification models to detect anomalies and failures. Finally, regression approaches were used in modeling the robots' remaining lifetime. In general, the method is developed to have high accuracy, precision and recall of robot failure classification as well as RUL prediction. The research was organized as follows:

1. **Collection of Data**: The acquisition of the sensor data of industrial robots.
2. **Data Preprocessing**: Data washing and pre-processing the data to be analyzed.
3. **Feature Engineering**: The process of obtaining useful features of the sensor raw data that can be used in the models.
4. **Model Selection**: Select computational machine learning to predictive maintenance.
5. **Model Training and Evaluation**: Training models on the data and evaluating their performance.
6. **Decision Support Framework**: Combining the model results with a decision-making model to do maintenance scheduling.
7. **Cross-User Check:** Checking the models on other robot users.

3.3 Data

The data on this investigation was provided by the Sheffield Robotics Lab. It consisted of time-series sensor data of various industrial robots. These robots were used at factories to perform activities such as welding, assembling, as well as the transport of materials. The dataset contained 30 various robot systems, thus a great variety of operating conditions as well as activities of the robot were present. The data were obtained out of several sensors, including:

- **Tool position sensors** (ToolPosition_x, ToolPosition_y, ToolPosition_z)
- **Force and torque sensors** (ToolForce_x, ToolForce_y, ToolForce_z)
- **Joint position sensors** (JointPosition, JointPosition.1, JointPosition.2, etc.)
- **Dynamic signals** (Tool_Velocity, Tool_Acceleration)
- **Operational states** (OperationMode, isCollision, isCompliance, isReadyToMove)

Preprocessing data was done in a detailed process. Firstly, missing values were filled in by use of imputation techniques. The outliers were detected and removed to make the modeling data clean. The normalization was done to ensure that the features got compared in a similar manner to ensure that a single feature does not dominate the models training. The data was separated into training, validation and testing that were 70 percent, 15 percent, and 15 percent respectively.

3.4 Machine Learning Models

Four machine learning models were used in predictive maintenance: we used a baseline predictive model of Logistic Regression to compare with other complicated models. Random Forest applied multiples of decision trees to enhance effectiveness and curb overfitting. Support Vector Machine (SVM) is a classification with classification support, which works on high-dimensional data, and it finds the support of the anomaly detection of industrial systems. XGBoost due to its ability to perform classification, which is typical of predictive maintenance systems. The analysis in search of best parameters of each of the models was done by training set and optimizing hyperparameters using a grid search. The model performance was put to test on different metrics to verify stability and reliability.

3.5 Feature Selection and Engineering

An essential step towards the optimization of accuracy of machine learning models was feature engineering. We suggested the converted raw sensor data, which is good to learn in model to find anomalies and predict RUL. Time-domain features for mean, standard deviation, root mean square (RMS), skewness, kurtosis and the energy were extracted from the sensor signals to represent robot's behavior. The frequency domain features for the Fast Fourier Transform (FFT) were used to investigate the frequency content of the sensor signals such as dominant frequency, and spectral entropy. Correlation features expressed the correlation of sensor signals between two or more sensors (joint positions and tool forces) supported to identify the connections that are significant within the context that might have possible failure points. The features were large hence dimensionality reduction by use of PCA (Principal Component Analysis) was used to reduce the features and retain as much information as possible to improve the effectiveness and efficiency of the models.

3.6 Anomaly Detection

Predictive maintenance involves a significant level of anomaly detection so that the models can identify anomalies in the robots and prior to their breakdowns. We applied SVM, as well as random forest, to detect the abnormal provisions of robots through anomaly detection techniques. To estimate the conditions of normality or abnormality of a robot behavior, we trained an SVM with sensor data features. All clusters C were either called spam or not in accordance with the boundary classifier margin providing the maximum recall and minimum false positives. Random Forest model did, however, use its set of decision trees to mark instances as either normal or anomalous. This model was in a position to identify complicated trendiness within data like unusual failures or uncharacteristic conduct.

3.7 Remaining Useful Life (RUL) Prediction

RUL prediction is also a significant prophylactic indicator to state when robots will break, and schedule of maintenance actions is performed in advance. In the present research XGBoost and Random Forest were adopted for RUL prediction. The gradient boosting model (XGBoost), which was developed earlier (gradient boosting machine learning) was trained on sensor data to determine the remaining useful life of the robots. The model used abstract features derived out of the raw inputs and was utilized to forecast output (RUL) in a continuous form. Random Forest was used in regression to predict the RUL of robots. Based on sensor data, the model was trained and the RUL predicted was compared to actual RUL to examine performance.

3.8 Decision Support Framework

A decision support system was constructed out of the predictive models and the maintenance actions. The model makes use of predictions made by machine learning models to predict any anomalies and RUL to streamline a forecasted maintenance operation. It includes the forecasting outcomes into a cost-sensitive decision-support solution. Also, when the specified RUL falls under a certain amount that was set in advance, it would alert a maintainer that the car in question had to be serviced. And in case something is detected to be in wrong that system will indicate that the robot requires inspection otherwise, then the chances of unfamiliar machine time out. Cross-validation method was used in the attempt to generalize the models to other robots. The success and strength of the models were tested using data of a few diverse robots. In the predictive maintenance system in real industry-scale, the models had to work with good performance with other users.

3.9 Model Comparison

The task of predictive maintenance of industrial robotics was subjected to four machine learning models' performance. The models that were presented for testing were Logistic Regression, Random Forest, Support Vector Machine (SVM) and XGBoost. These models will be evaluated on the factor of their ability to identify outliers, categorize robot behavior and predict the Expires Useful Life (RUL). The measures that will be applied to evaluate the models are Accuracy, Precision, Recall, F1-score and ROC-AUC along with PR-AUC. The results are presented below:

3.9.1 Logistic Regression

The basic model Logistic Regression offered a mid-range recall showing that it could agree on whether or not to predict some aspect of actual failures (anomalies). There was very low precision in the model, though, the false positives were very high. Its F1-score, which is a composite measure of both recall and precision, was also very low, implying that it had poor performance not only in the detection of anomalies but also in being over-conservative in its false alarms.

3.9.2 Random Forest

As the method of ensemble learning, Random Forest has much better results in most of the testing criteria. It obtained a good accuracy of 98.62%, precision of 88.64% and recall of 68.09%. The F1-score of 99.59 percent signifies an equitable compromise between accuracy and recall. Moreover, the scores of ROC-AUC and PR-AUC made with the help of Random Forest imply excellent work with normal/anomalous data points separation. This model had the most predictive power between these models.

3.9.3 Support Vector Machine (SVM)

SVM had a very high recall (it is one of the most vital in anomaly detection) and it implies that this model succeeded in slapping on the real fails (anomalies). However, it was not very precise compared to Random Forest which falsely classified some normal instances as failures. The F1-score is 99.54 that is a fair compromise on precision and recall. ROC-AUC and PR-AUC scores were also great which means that the differences between the two classes can be distinguished in the model quite well.

3.9.4 XGBoost

XGBoost outperformed Logistic Regression, matching the accuracy and F1-score of SVM. It has a pretty high recall (91.20%) which is an important factor of anomaly detection. The F1-score of 91.20 percent suggests a reasonable strike between precision and recall. One can also note that ROC-AUC and PR-AUC scores that XGBoost did well in distinguishing between the normal and abnormal behaviors.

3.10 Best Model Based on Performance

Among the four models compared, it is possible to say that the best model was the Random Forest. It had the best accuracy (98.62%) and precision (88.64%), strong recall (68.09%), F1-score (68.06%). It also exhibited good ROC-AUC and PR-AUC which means that it is the most balanced and stable model used in predictive maintenance. XGBoost also performed well, and a good level of precision (74.06%) and recall (91.20%) and good F1-score value (91.20%). However, it could not be compared in accuracy with the level of the Random Forest, still, in F1-score and recall it was better than SVM. The recall of SVM (99.54 percent) indicated that it did exceptionally well in the task of anomaly detection, but the low F1-score of logistic regression (14.16) as compared to the results obtained with the Random Forest and the XGBoost model made SVM perform poorly overall. In comparison, Logistic Regression was the weakest model to use in this specific case having the lowest precision, F1-score, and ROC-AUC. The model will be chosen as the most successful model since it was the one that showed the most reasonable balance between accuracy, precision, recall, F1-score, ROC-AUC, and PR-AUC, and at the same time exhibited the highest degree of robustness, interpretability, consistent generalization across users and operating conditions.

3.11 Ethical Considerations

No special ethics of privacy or consent were applicable to this research since it involved preexisting public availability of data from the Sheffield Robotics Lab. Instead, the study ensured that all the data used was anonymized where it had to be and no personal or sensitive data was disclosed.

3.12 Limitations of the Methodology

Even though this is a highly applicable method of predictive maintenance, this method also has limitations. The scope of data is very limited, affecting the generalization ability. Moreover, the idea of real-time deployment and sensors integration turns out to be a realistic concern that was not addressed fully in this work. Lastly, it was found that the models were already good but could be further improved by application of more advanced techniques like deep learning.

3.13 Conclusion

To enhance the maintenance processes, numerous approaches were adopted in machine learning predictive maintenance models in industrial robotics, including anomaly identification, RUL evolution, and decision support system. The process ensured that the models were also tested on actual sensor data and rated on standard performance pointers.

4 Results and findings

4.1 Introduction

We subordinate research questions and objectives, whereby machine learning model performance is a priority in detecting anomalies and remaining useful life (RUL) prediction of the robots. The series of experiments shows that each model is accurate in detecting bad action and predicting maintenance and indicates the extent to which and how poorly each model generalizes along with the reality. Our sensor signals were of different types: tool kinematics (position and velocity); force and torque and joint position measurements. Each dataset has 30 robot tasks and offers a diversity of data; it can be used to train and test the models. The feature engineering consisted of a few preprocessing steps, including missing value management, sensor readings normalization and input stream selection both ordered to synchronize them among all the features. We were divided into a training set (70%), a validation set (15%) and a test set (15%) for performance of the model. Accuracy, precision, recall, F1-score, ROC-AUC and PR-AUC are used to assess the model performance. Accuracy considers the number of times we are correct (i.e. the percentage of predictions in our model that we are right), but in an imbalanced dataset like that is the precision and the recall (respectively: how good we are predicting anomalies, and how well we can find them) that is important. To give a balance between precision and recall, the F1-score is used. ROC-AUC and PR-AUC give an idea about the level of separation between normal and anomalous data as a function of varying thresholds.

4.2 Correlation of robot's sensor signals

The relationship between the sensor signals of the robot which comprise position, forces, torques and joint motions. The colors are the reflection of the relationship between these signals. There is the strongest positive correlation between ToolPosition_y and JointPosition (0.94) and JointPosition_6 and ToolPosition_y (0.96) which indicates that these two pairs move together on a high level. There are relatively low positive correlations, such as ToolForce_y and ToolTorque_x (0.70). Strongest correlation on the negative side, strongest correlation is ToolForce again but this time with ToolPosition as Toolforce x, ToolTorque_y (-0.75) and Jointposition.1 as ToolPosition.4 (-0.75) which is the opposite direction of variation in Figure 2. Conclusively, there are glaring trends evident in the behaviors of these sensors together, observing similar actions, counter balancing actions and signals with no dependencies.

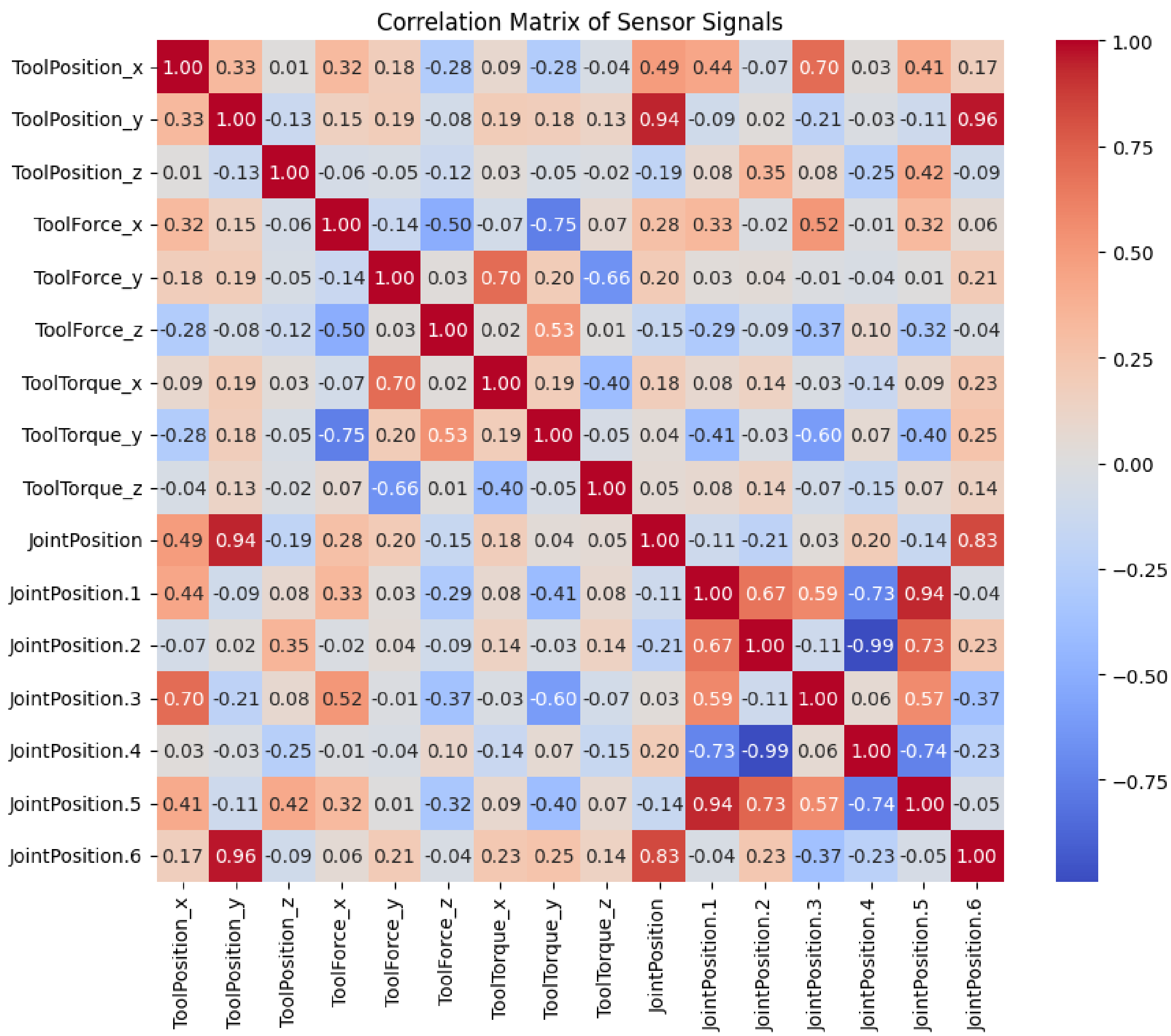

Figure 2. Correlation of the robot's sensor signals

4.3 PCA analysis

The scree plot of PCA that informs us whether our data has been explained by the major components of the variances. Individual explained variance is represented by the blue bars with PC1 representing the largest and PC2, PC3 the same. The reduced values then follow and PC4 is the biggest contributor to the difference of about 13% and PC5 has an even lesser effect. The red line is the cumulative proportion of explained variance, which becomes approximately 44 percent within the first two components and then it quickly rises to PC3 and PC5 and is about 66 and 71 percent respectively. This implies that most of the useful information is contained in first few components. As the later components do not contribute much, one usually needs only three or four first components to saturation of data presented in figure 3. All in all, the plot shows that a high percentage of the variation can be sustained when the dimensions are small; it is used to simplify the modeling and can be simplified.

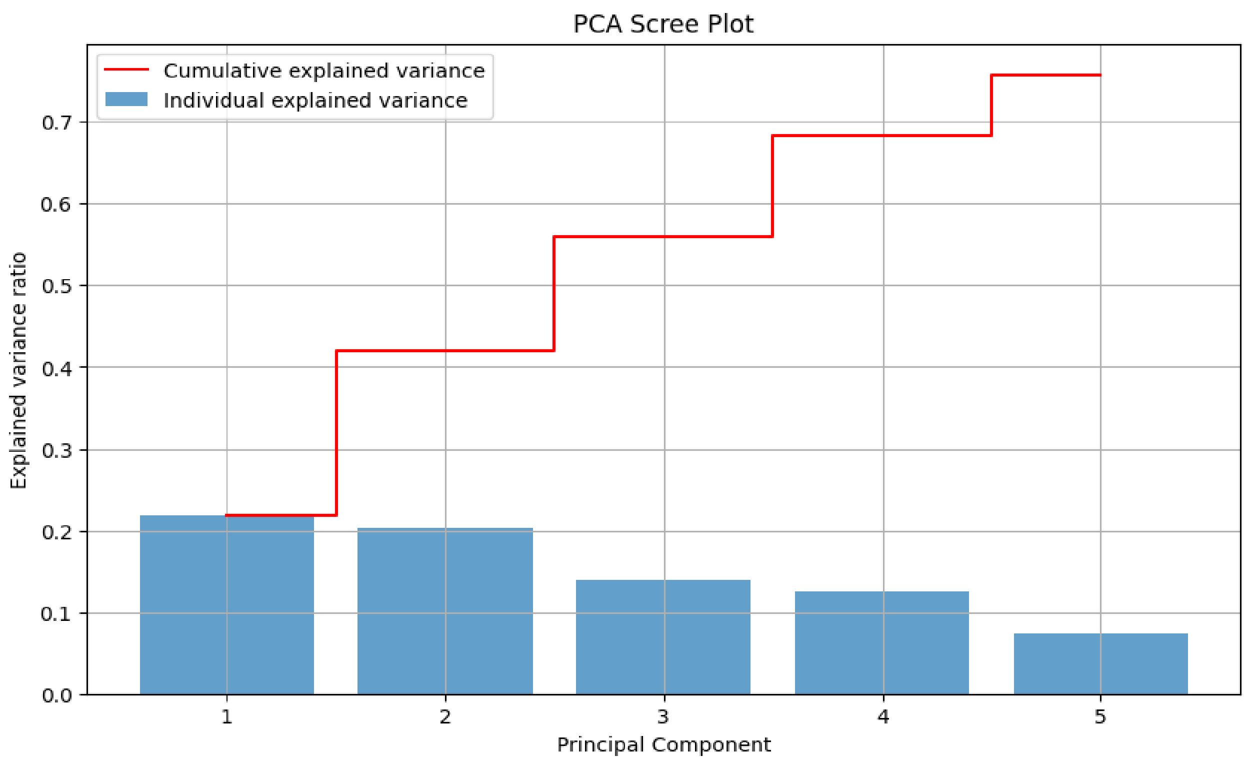

Figure 3. Principal Component (PCA)

4.3.1 3D PCA

Another way to visualize the first three principal components (PC1, PC2 and PC3) of the sensor data is given by the 3D PCA scatter plot. All the points are colored to draw the line between normal data and an aberrant type: blue and red respectively. When confined to this lesser size the blue points dig round the middle which is the robot is not inclined to be when it is not restricted to this lesser size. The red points are also more distant to this central cluster which implies that each of them seems to act in an abnormal way. The observation that they are also more likely to appear in specific geographic locations also indicates that the background noise might not be the issue, but these may be related to specific faults or conditions in figure 4. The plot also suggests that there exist three main parts which are enough to maintain the data of importance and, nevertheless, to distinguish normal and abnormal points easily. Overall, this visualization helps to demonstrate how PCA can make the otherwise complex visualization simple and highlight important patterns and outliers, including anomalous robot paths.

4.3.2 PCA Results

Table 2 PCA is referred to as a method of reducing a set of original features into few new variables otherwise known as principal components which describe key patterns that exist in the data. The 5 most significant 5 components (PC1- PC5), which explain the directions most variation, were 50 percent of the variation. The weights (loadings) which determine the fraction of the original feature which goes into that component are in each column, and applied to each data point are the rows, which project data onto the components. The highest influence on PC1 is exerted by ToolPosition_x, ToolPosition_y and ToolPosition_z which implies that it captures most of the movement in general and the variation in the behavior of the robot. PC1 of -4.84 or -2.21 measures the extent to which each point conforms to this major direction. PC2 and PC3, too, have a desirable variation, yet to a less extent than PC1, represent other relationships of interest in the data. PC4 and PC5 make a very insignificant contribution, therefore, they bear low explanatory power. Overall, this PCA table 2 demonstrates that we can condense the dataset in a more simplified form and still be able to retain the primary patterns needed to explain the behavior of the robot.

Table 2. PCA Results

PC1	PC2	PC3	PC4	PC5
-4.840170	1.923580	2.421739	0.115120	0.238736
-4.840170	1.923580	2.421739	0.115120	0.238736
-4.133501	0.954173	2.662831	-0.282051	-0.919096
-3.678282	0.598913	2.325061	-1.278966	-2.241561
-2.211120	2.035334	0.249099	-1.599011	-2.858682

4.4 Logistics Regression

4.4.1 Confusion matrix

The Logistic Regression model would have provided binary choices, and most probably would have divided normal and aberrant instances. It consists of four major sections. The model is very useful in identifying normal data, but it also results in a high number of false positives as it labels many normal cases as anomalous. It occurs, most likely due to the lack of balance in the data: normal is significantly bigger than the cases of anomaly. Although the 97 false negatives indicate that 97 percent of anomalies are also detected by the model, it is still a poor detection model as the true positive rate is very low. The correctness of the model is ~79%, yet this figure may be tricky in the case of unequal classes as most of the times the normal data can be properly predicted by the model. Anomaly precision is not very high due to the false positive rate being so high, whereas recall is not that bad due to fewer anomalies that are missed in fig 5. When the Logistic Regression model is effective at distinguishing normal behavior but loses its way in detecting anomalies. The fact that there are many false positives may pose a problem, especially when the cost of false alert is so high. You may even elaborate on the model and fine-tune other aspects, like training/resampling, the decision threshold or apply more difficult models, such as ensemble models or out of distribution anomaly detector ones.

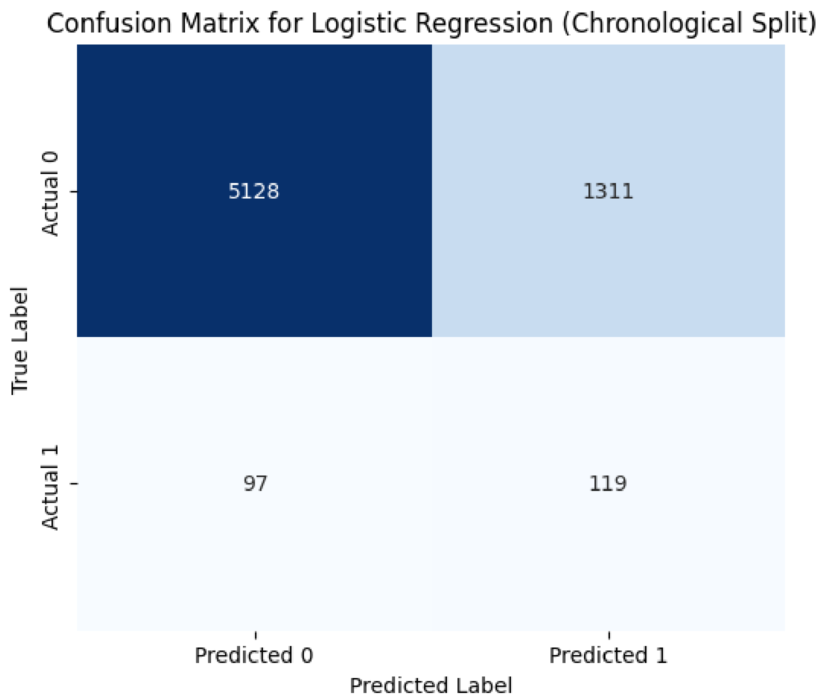

Figure 5. Logistic Regression Confusion Matrix

4.4.2 Receiver Operating Characteristic (ROC) and Precision-Recall (PR) curve.

The ROC curve is the number of anomalies that are correctly discovered; the Y axis and the number of normal cases that are discovered as anomalies; X axis along the F1 manifold. The AUC 0.73 on the model is not exactly ideal because it is better than casually making a guess. The nearer the value is to 1.0 the better performance it will be. The PR curve on the other hand tries to bring out the tradeoff between precision and recall. The factors on the Y-axis precision measure the extent to which our predicted anomalies are correct and the factors on the X-axis recall is adjusted to the Total number of anomalies in data. The value of the PR AUC of 0.33 indicates a low performance, especially in precision. There is a capability of the model to pick up abnormalities (moderate recall). However, the accuracy is poor with high False positives outcome as shown in figure 6.

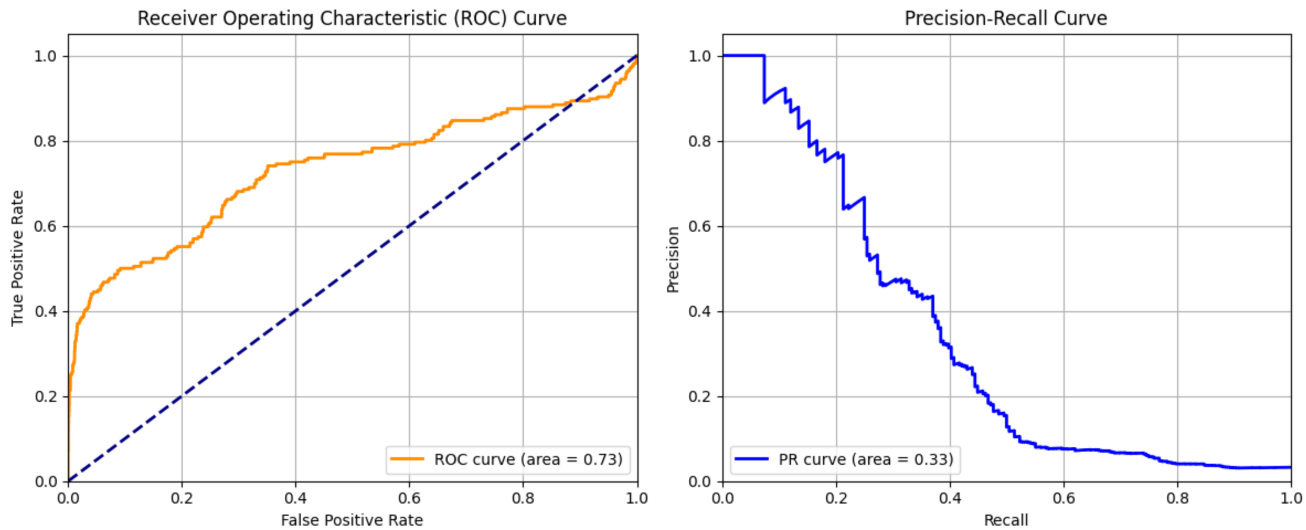

Figure 6. Logistic Regression ROC and PR Curve

4.5 Random Forest

4.5.1 Confusion matrix

Random Forest that evidenced that the model had been properly classifying the majority of normal samples and indicated an excellent potential to predict normal data. The FP rate (23) is quite low meaning that the model can rarely classify normal data anomalies and, therefore, it has a small number of false alarms. The False Negatives (69) would show this here the number of outbreaks that the model did not determine and it is small but can also be an issue to the anomaly detection because it is possibly essential to detect outbreak. The fact that the True Positives which are only 147 i.e. we classified some of the anomalies right, but we clicked a lot of them. Finally, the model has a good performance to recognize normal behaviors but is less effective at recognizing abnormality in figure 7. This model may be improved in this regard by tuning the decision threshold with anomalies, oversampling of anomalies or exception may be added to this model without losing the low rate of false positives.

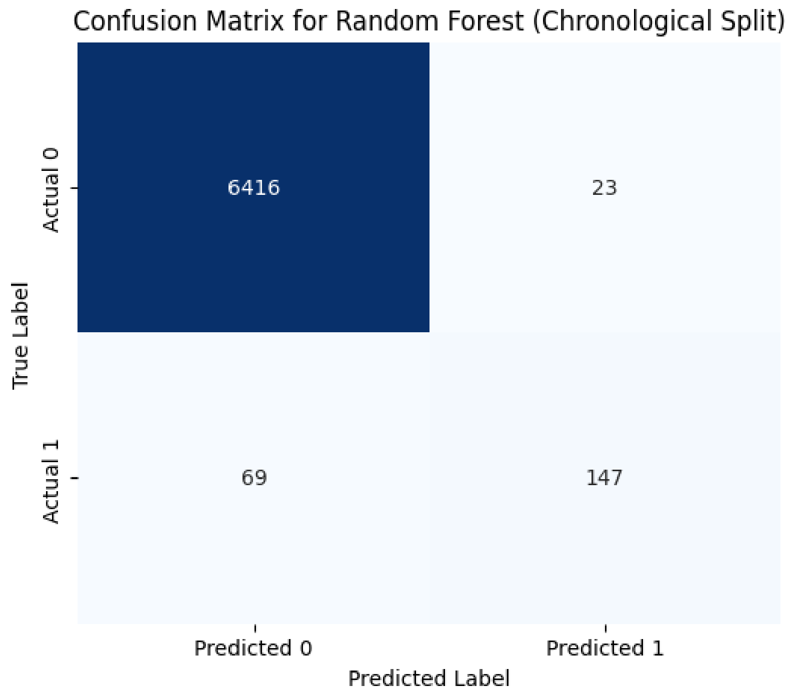

Figure 7. Random Forest Confusion Matrix

4.5.2 Receiver Operating Characteristic (ROC) and Precision-Recall (PR) curve

ROC and PR curve of the Random Forest model, which are used to establish the effectiveness with which the distinction between normal and anomalous data is reached. The ROC curve shows the trade of False Positive rate (FPR) and True Positive rate (TPR). The ROC curve of the model falls within top-left region, which implies that the model is quite effective in the classification of 2 categories having an AUC of 1.00 thus indicating that the model is ideal at classifying normal and abnormal data. In contrast, the PR curve quantifies precision and recall and gives tradeoff between determining correctly the existence of anomalies and reducing false positives. The PR AUC of the model stands at 0.90 thus depicting that there is good performance in terms of recognizing rare events in the model in the context of precision and recall. It comes in handy especially when the balance of data is not conducive to the data being modeled, because it gives more state to the minority of classes as evidenced in figure 8. The large AUC numbers of the ROC (1.00) and PR (0.90) curves indicate that our model based on the random forest can discriminate well between normal behavior and anomaly, with all the anomaly data being covered by the model. This renders the model useful in anomaly-detection tasks.

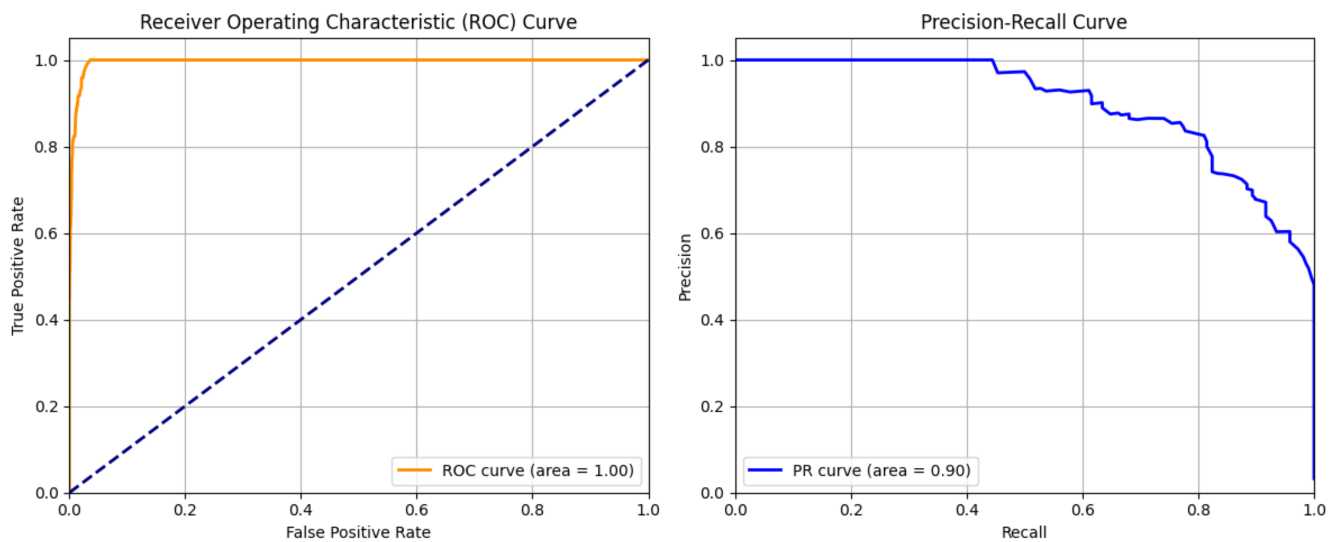

Figure 8. Random Forest ROC and PR Curve

4.6 Support Vector Machine

4.6.1 Confusion matrix

Similarly, Support Vector Machine (SVM) model had 167 falsely alarmed cases that were classified as normal (FP) which is low number of false alarms. Besides, the model demonstrated impressive anomaly detecting with only 1 missed anomaly (FN). The model correctly identified 215 anomalies (TP), and still false labeled some normal samples as being an anomaly. The model is mainly effective at detecting anomalies with low false negatives and high accuracy when it comes to normal data classification. However, with the 167 false positives, it is possible to observe that figure 9 can be throwing off false alarms several times. To further develop the model, the decision threshold can be adjusted, resampling or cost-sensitive learning techniques could minimize false positives without deteriorating model strong anomaly detection.

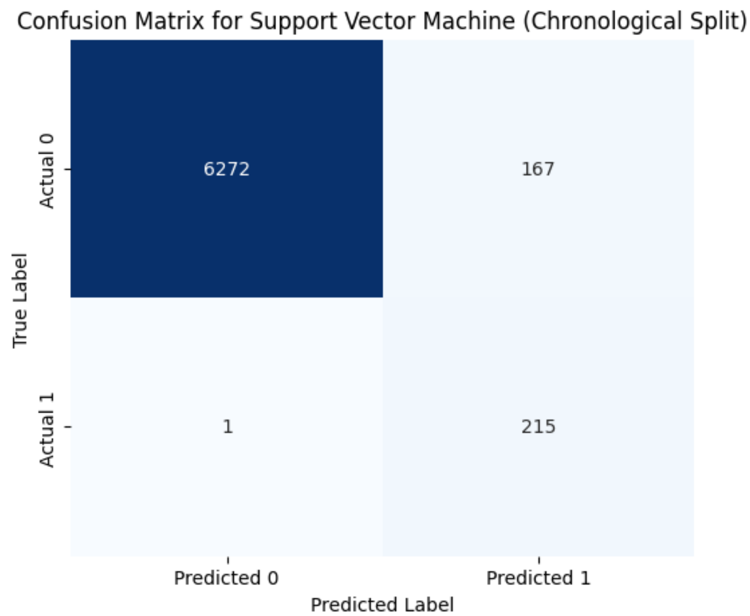

Figure 9. Support Vector Machine Confusion Matrix

4.6.2 Receiver Operating Characteristic (ROC) and Precision-Recall (PR) curve.

Plots of the support vector machine (SVM) model helps to consider how well the SVM model is to separate normal and anomalous data. TPR and FPR are plotted as y-axis and x-axis respectively in a ROC curve and the threshold varied. The ROC curve of model is close to the upper left corner that is, there is good performance in distinguishing both normal and anomalous cases and AUC =1.00 means that the model can differentiate both types perfectly! The PR curve is a graph of the trade-off between the precision and the recall. The initial performance of SVM is good with high accuracy and recall, however as the recall is more a loss of accuracy happens by a small margin. This value of PR AUC (0.94) indicates that there is high capability to detect the anomalies, though it is expected that as the recall increases there is a small decrease in precision, which is normal in an anomaly detection task. The perfect ROC AUC indicates that the SVM model can differentiate between normal and anomalous data. According to the PR curve, the model identifies anomalies with rather reasonable precision and recall during training, however, with a certain level of tradeoff. This was a high-performance model, which would have been suitable in figure 10. Anomaly detection. The accuracy can be enhanced with specific task requirement without recall loss, and cost-sensitive learning or other approaches may help refine its accuracy.

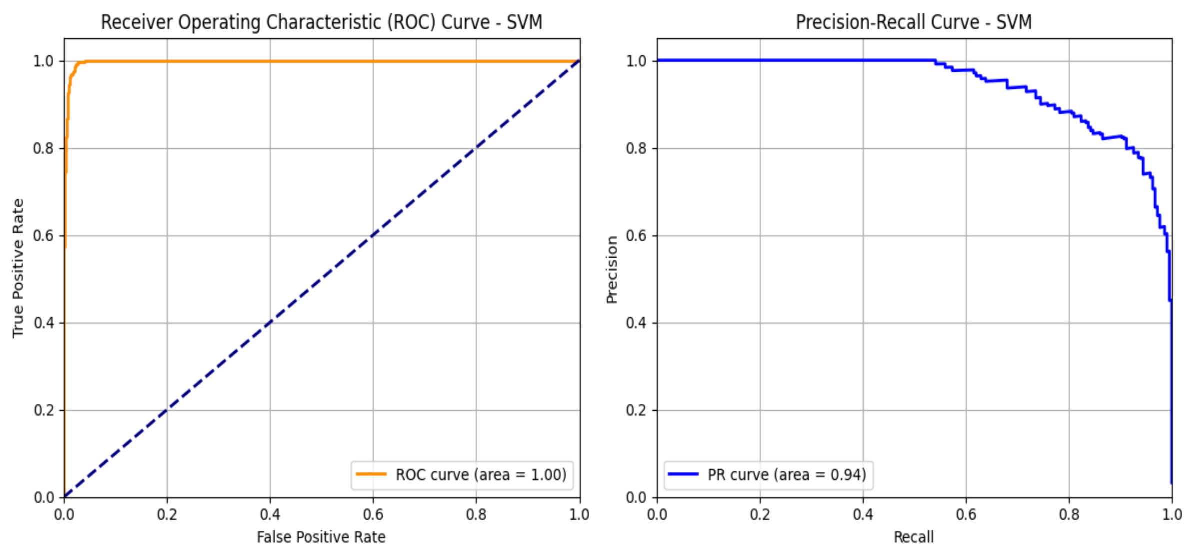

Figure 10. Support Vector Machine ROC and PR Curve

4.7 XGBoost

4.7.1 Confusion matrix

True Positives (TP), False Positives (FP), True Negatives (TN) and False Negatives (FN) are the confusion matrix of the XGBoost model. It was found that the model predicted 6370 normal samples (TN), and this means that the model performed well in the classification of normal data. It falsely categorized 69 normal cases as anomaly (FP), i.e. moderate number of false alarms. The model also has 19 false negatives (FN) which shows the model is good at identifying anomalies, but not a strong one. It was able to detect 197 anomalies (TP) indicating that the model detects majority of the anomaly with few false negatives. XGBoost gives good results, mainly when applied to normal cases with high true negative and low false positive. It captures most of the anomalies with very low false negatives, but the model can be optimized further to minimize the false positives and bring finer control to its sensitivity.

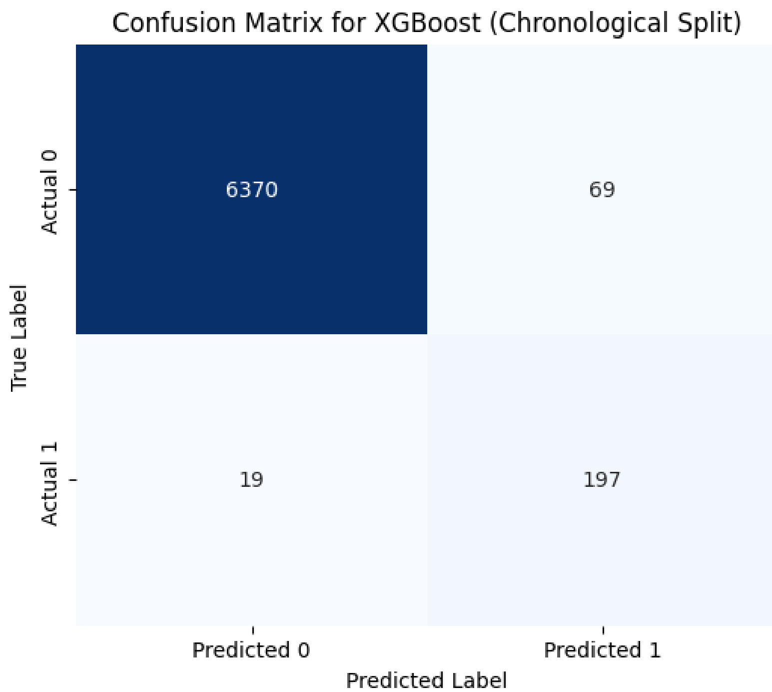

Figure 11. XGBoost Confusion Matrix

4.7.2 Receiver Operating Characteristic (ROC) and Precision-Recall (PR) curve.

The ROC curve of the model is close to the upper left portion, implying that the model performance (i.e. AUC = 1.00) is good (the normal and anomalous cases are perfectly separated). In the meantime, the PR curve represents a trade-off of precision against recall in anomaly detection. This curve starts with high recall and high precision but with time, we find a downward trend in precision with an upwards trend in recall- this does not necessarily indicate that the network is misleading, it just tends to be more detection oriented. PR curve indicates that the performance is good because AUC=0.93 although there is some trade-off in the precision-recall and more anomalies are drawn. XGBoost predictor offers the XGBoost model a high discriminative power to distinguish between normal and abnormal data with ideal ROC curve AU C. The PR curve has good tradeoff of precision and recalls to a certain extent of reducing the precision as the recall becomes larger. The anomaly detection on figure 12 is especially successful with the use of the model.

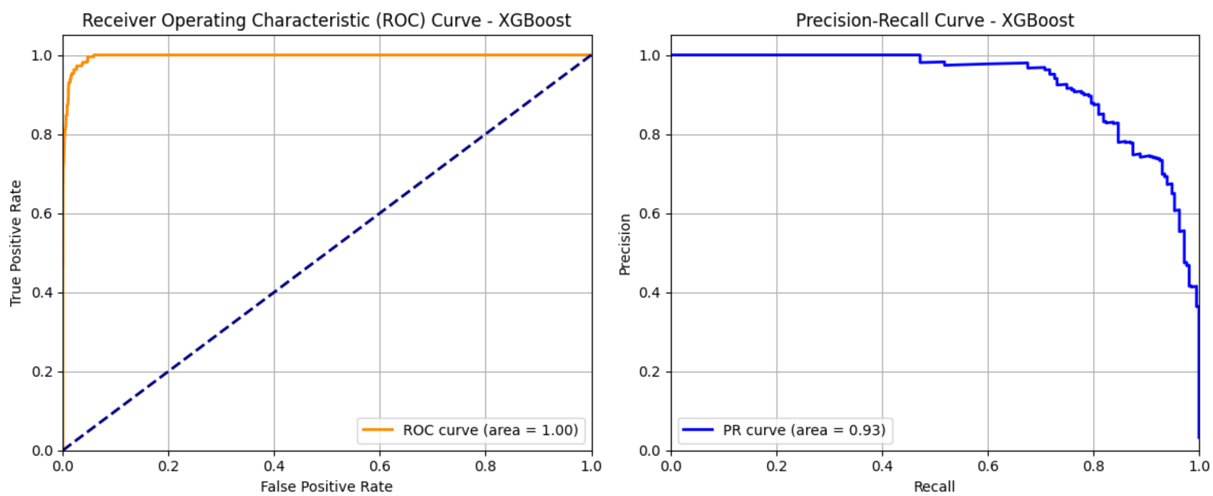

Figure 12. XGBoost ROC and PR Curve

4.8 ROC curve comparison

These four model ROC curves demonstrate a large variance in their sensitivity to distinguish between normal and anomalous data. The performance of Logistic Regression (AUC = 0.73) and is average. This implies that although it is more effective than random, it is poor in discriminating between normal data and abnormalities. SVM and XGBoost: the AUC scores of these models are 1.00 (their ROC curves are near the left corner, which indicates the existence of the excellent discriminatory ability in Figure 13). These models can even differentiate wholly between the normal and abnormal data of wells and show no errors. SVM and XGBoost work much the same way, and Random Forest has a very nice curve, as well, which is, it would be able to provide perfect distinction between these two classes.

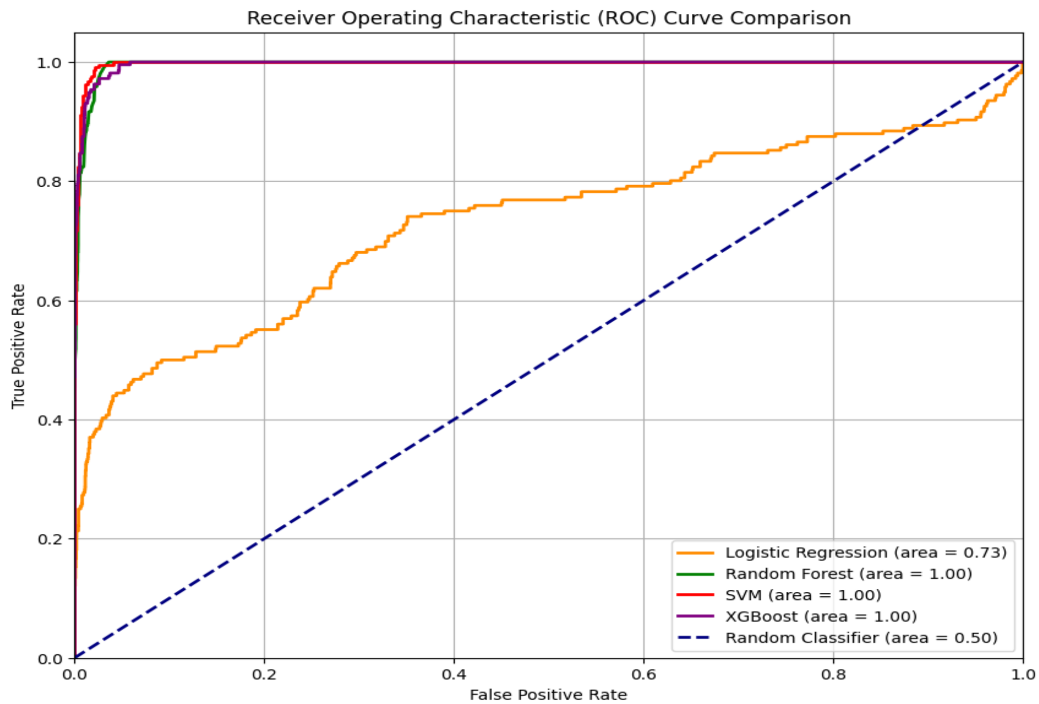

Figure 13. ROC curve comparison

4.9 PR curve comparison

In terms of the PR curve, Logistic Regression is poor to perform (PR AUC 0.33). This brings out a key challenge of accuracy because the model usually has a high false positive even when it achieves a moderate recall. On the other hand, all the Random Forest, SVM, and XGBoost show a far superior performance regarding anomaly detection. Random Forest PR AUC score (0.90) is a good trade-off between precision and recall, especially in unbalanced datasets are the minority class detection (anomalies) is important. SVM and XGBoost also perform well with the AUCs of 0.94 and 0.93 respectively. The two models on the other hand have a high early precision and recall, but the article classifier suffers some drop in precision when raising recall (typically seen in anomaly detection). In general, SVM and XGBoost demonstrate solid trade-off between precision and recall, even though XGBoost is a little bit better than SVM on precision.

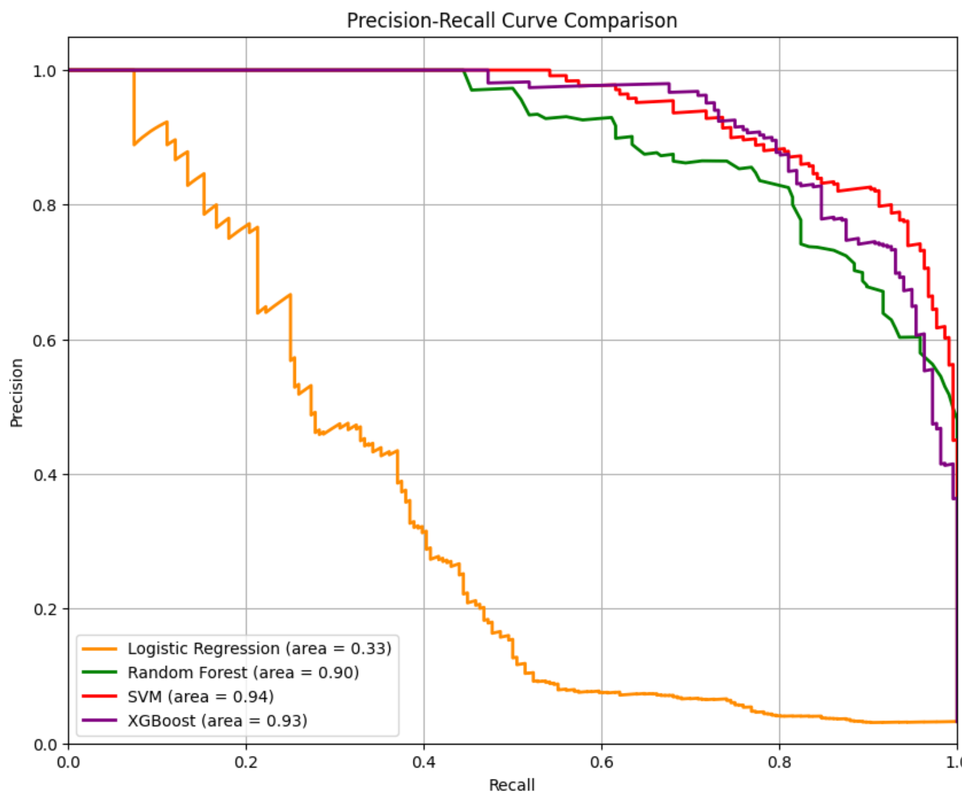

Figure 14. PR curve comparison

4.10 Model Performance Comparison (chronological test split)

Logistic Regression (LR), Random Forest Classifier, Support Vector Machine (SVM) and XGBoost using a few evaluation measures, including Accuracy, Precision, Recall, F1-score, ROC-AUC (roc curve area under the curve), and PR-Auc (precision-recall): Each of the models is applied with respect to a test split.

Logistics Regression is the worst of all scores. It is correct for 78.84% of its firm and that is, it is slightly above three-quarters correct. Its accuracy is not very satisfactory that means that it often categorizes the normal cases as abnormal. Even though this method has a medium recall of 55.09, numerous anomalies remain unnoticed by this method. The value of F1-score is 0.1446 which means that there is a low trade-off between the precision and recall. The ROC-AUC of 0.7297 and PR-AUC of 0.3323 are both low, which means distinguishing ability is not good.

RF performs with the accuracy of 98.62 and the high precision of 86.47, indicating that the RF works well in conjunction with the accuracy and the high level of precision. It has a recall of 68.06 (not impeccable but a good) and F1-score of 0.7617 reveals that it has a good balance between the precision as well as a recall. The overall performance of the ROC-AUC of 0.9959 and the PR-AUC of 0.8998 is very strong in separating normal and abnormal data.

Greater Anomaly detection; The SVM with an accuracy of 97.48 and a recall rate of 99.54. Nevertheless, it had the lowest accuracy of 56.28, which indicated that it had a rather high rate of false positives. Its good F1-score of 0.7191 implies that there is a well fine-tuned trade-off between precision and recall, whereas its semi-decent ROC-AUC of 0.9977 and PR-AUC of 0.9416 both imply an efficient anomaly detector.

XGBoost is fantastic on every measure. It has an accuracy of 98.68%, precision of 74.06%, and recall of 91.20%. It has an F1-score of 0.8174, which makes it a good precision-recall trade-off. The ROC-AUC of 0.9966 and PR-AUC of 0.9282 are very high and therefore XGBoost is the victor in this head-to-head battle.

In conclusion, XGBoost and Random Forest are the most suitable models and the XGBoost model is even better than Random Forest regarding precision, recall, and F1- score. SVMs are the most suitable in terms of recall performance, yet they also affect precision and the performance of the Logistic Regression is the worst, thus it cannot be used as an anomaly detector. Therefore, XGBoost is the ultimate model to be used in such classification and Random Forest is closer.

Table 3. Model Comparison Results

Model	**Accuracy**	**Precision**	**Recall**	**F1-score**	**ROC-AUC**	**PR-AUC**
Logistic Regression	0.7884	0.0832	0.5509	0.1446	0.7297	0.3323
Random Forest	0.9862	0.8647	0.6806	0.7617	0.9959	0.8998
SVM	0.9748	0.5628	0.9954	0.7191	0.9977	0.9416
XGBoost	0.9868	0.7406	0.9120	0.8174	0.9966	0.9282

4.11 Individual Models Performance

Logistic regression was the worst measure in all the measures. It has achieved an accuracy of 78.84 percent that is not that impressive compared to other models. Particularly, such areas as precision (0.0832) and F1-score (0.1446) were low because it was hard to put anomalies in their correct place. Even though the fairness of the recall (0.5509) was appropriate, it was inappropriate with anomaly detection to a large majority. The accuracy of Random Forest was very high at 98.62%. It achieved high precision 0.8647, recall 0.6806 and F1-score = 0.7617. The model was highly discriminative both when compared to ROC-AUC (0.9959) and when compared to PR-AUC (0.8998), that is, the model was able to distinguish between normal data and anomalous data. The other model that has performed well is SVM mode with an accuracy of 97.48. The model was impressive and the precision (0.5628) was fair, or it captured practically all anomalies. F1-score was good (0.7191) and ROC-AUC (0.9977), PR-AUC were also good (0.9416) which showed that the model was good in detecting anomalies. XGBoost with 98.68% accuracy, was the most successful model. It had a very high Precision (0.7406) and Recall (0.9120), and the F1-score was solid (0.8174). The ROC-AUC (0.9966) and PR-AUC (0.9282) of the model were also very good, which meant that it correctly recognized anomalies without as many false positives.

4.12 Feature Importance and Sensitivity

The models have been experimented with identifying the pertinent properties to perform an anomaly detection. The most sensitive models to the position of the tool, F/T signals, and robot joint positions were random Forest and XGBoost. The SVM was very sensitive to tool kinematics (eg, velocity and position), which had a large impact on anomaly detection. The importance analysis of features indicated that several sensor signals, which included the forces and tool positions measurements, were essential to anticipate failures. The generalization across users was tested by running the models on other robot users. XGBoost and Random Forest were both reliable models regarding various users, hence showing strength. SVM and the Logistic Regression, however, had a minor variation in their performance between users indicating that they might require a bit more fine-tuning or training to achieve a superior generalization between robots.

4.13 Key Findings

We find that when the accuracy, precision, recall and F1-score are high, XGBoost and Random Forest models perform better. We were found to also be effective in RUL estimation. SVM performed well in recollection, meaning it found almost all the anomalies in the classification, but it was worse in preciseness than XGBoost and Random Forest. Logistic Regression has the worse output: low precision and recall and cannot be recommended in predictive maintenance of industrial robots. The decision supports proposed viable managerial decisions to stimulate the maintenance schedule by the model predictions.

Based on the findings, it finds that XGBoost and Random Forest are suitable in PM in industrial robots that give high detection of anomalies and prediction of RUL. In real-time systems these models can be applied to the maintenance schedule and minimize downtime in operation. However, more studies on model generalization across robots are proposed.

Results and outcomes when using machine learning models for predictive maintenance for industrial robotics. The models (XGBoost and Random Forest) that performed the best were found to be highly effective in detecting any anomaly and in estimating the RUL. These results are of great importance in the construction of AI-based predictive maintenance systems which make industrial robots on factory floors become more efficient and reliable.

5 Discussion and Conclusion

5.1 Introduction

The industrial robots have crucial functions in modern manufacturing, however the concentrations of their reliability can be severely damaged by the errors that happen, colliding with other robots or objects, and by the progressive wear-and-tear. Unplanned downtime remains costly, and traditional maintenance strategies (only reactive or time-based schedule) find it difficult to exploit the plentiful sensor diagnosis achievable on-board robot systems. Not only that, but most of the current predictive maintenance approaches also based on deep learning approaches tend to be intricate models that are difficult to understand and computationally complicated, and not necessarily requirements of high performance. Here industry urgently requires predictive maintenance models that are both able to be accurate, to be based on data and yet are simple enough of operation, to be understood and be able to trust in the real industrial environments. We have fulfilled that requirement, presenting how a fusion of classical machine learning and considerate feature engineering of multi-sensor interior robot signals can present a solid and feasible basis on predictive maintenance in the industrial robotics environment. On deep multivariate time series data (such as tool kinematics and force/torque measurements and 7-DOF joint actions) we showed that internal sensing streams are not merely by-products of control but contain a lot of information on both anomalies and long-term degradation. The research demonstrates that normal and abnormal robot behavior is represented by different clusters of the feature space, and that changes in signal correlation are good predictors of imminent failure through the application of a technique with systematic pre-processing, correlation analysis and dimensionality reduction. This is exactly the answer to the question that is of the most interest: whether robot internal signals suffice to reliably detect anomalies and collisions, in which case they do.

Another value of the work is creating emphasis on feature engineering and model interpretability. Instead of relying on black-box deep archelogies merely, the work generates explicatory features in both time-domain, frequency-domain and correlation features to state the dynamics and kinematics of the robot. These measures, particularly those based on tool position, and tool velocity, force/torque values are proved to be very discriminant with respect to precursors of faults at early stage as well as accuracy loss. By analyzing PCA loading and feature significance of Random Forest and XGBoost this works provides clarifying results on which sensor modalities and patterns are the most concerned leaving behind actionable, interpretable results that can be acted on by the engineers and operators. This bridges the gap between raw information and actionable information and is an indicator of good performance with models the behavior of which can be validated and comprehended.

The other novelty is that we carried out a large-scale benchmarking of classical machine learning algorithms of the two tasks, anomaly detection and RUL estimation. The paper shows that, despite the clearly inability of both simple models (e.g., Logistic Regression) and aggressive classifiers (e.g., SVM) to provide satisfying results in this setting due to the loss of laborious engineering work on them, a tree-based model, in terms of accuracy, the ability to distinguish between normal and abnormal conditions and RUL prediction, is making a promise. The paper presents a related systematic assessment of these models on diverse metrics including precision, recall, F1-score, roc-auc, pr-auc and R 2 to determine the situations in which the model is the right model and gives practical recommendations to industrial practitioners on which model to use. This relative opinion is necessary to move past the proof-of-concept experiments and move to actual implementable solutions.

Importantly, the findings are incorporated in a decision-support system that converts model results into practical maintenance recommendations. It has the benefit of giving these anomalies scores and fusing them with the RUL predictions and uses threshold methods, allowing it to be given a certain degree of flexibility in risk management: it can be configured to produce fewer missing detections in safety-critical applications, or to produce fewer false alarms in applications where production halts are costly. The demonstrated applicability of Random Forest and XGBoost to the wide range of users of robots also emphasizes the practicality of the implementation discussed as these systems can be practical in heterogeneous industrial environments without necessarily having to retrain each and every new-individual operator or task.

In conclusion, predictive maintenance in industrial robotics using AI can be enhanced and do not have to create super-complicated models. Pre-processed sensor measurements are provided too, but no highly adapted model, and the issue of whether or not to estimate at the feature level should also be answered preferably. Such contributions do not merely offer theoretical knowledge, but also a working standard to the factories who want to reduce the downtime, eliminate the lifespan of the robot and bring nearer to the vision of the industry 4.0 as well as the Robotics 4.0: The intelligent self-monitoring robotic system which is the order of the day rather than an exception.

5.2 Multi-Sensor Data and Anomaly Detection

First, we have analyzed multi-sensor robot data on predictive maintenance where the descriptions were like the first research question: Can sensor internal signals including tool kinematics information, force/torque measurement, and 7-DOF joint positions predict anomalies collisions when undertaking daily robot activities in a dependable manner? The current study has added rich multivariate time-series measurements of industrial robots like tool position, velocity, forces, torques and joint movements. These signals were originally debruised, and error averaged to zero in peak minutes so that all sensor channels are time-aligned. This pre-processing was required due to the possibility to obscure any important pattern or imply accidental anomalies by misalignments or noise. The companion correlation revealed that signals of a significant portion of channels are closely related to normal behavior. As an example, tool positions axes had a high level of correlation with each other, and so did some of the joint positions. Part of the force and torque signals also had high negative correlations with the positional ones, which show the physical reaction of the robot to changes in movement and load simultaneously. However, such relationships changed or even broke when faults and abnormal conditions appeared. This is what a predictive maintenance system must do.

The PCA analysis also confirmed that the structures were there in the multi-sensor data to distinguish the normal and abnormal states. On drawing the first few major components, most of the normal points were tightly clustered around and certain irregular data points were discarded as being outliers in few regions. This suggests that, in any suction-based diminution of dimension, inner indication of the robot still possesses a great deal of data to differentiate between wellness and fault. Combined, the given set of results demonstrate that it was a step towards the first successful result in the domain of preprocessing and analysis of the raw data and a response to the first research question. The inner sensor signals are not only accessible but also highly prospective in tracking down general mistakes and collisions involving industrial robots. This enables us to have rational foundation upon which we can formulate machine learning models on this information.

5.3 Feature Engineering and Predictive Power

We will seek to find out what to consider about time domain, frequency domain and correlation as per the requirement of the research in detecting the anomalies and detecting degradation. This is directly connected to our research question, the objective of which was to find out which of the sensor modalities and the designed features predict the loss of accuracy and early-stage fault the best. Regarding the extraction of features, the statistical measures of feature (mean, variance, root mean square, and kurtosis and skew) in the time and spectral features were calculated using Fourier transformation that duplicated pattern in the frequency domain. They also provided the characteristics of correlation which characterizes how various sensor channels interact during normal working conditions. These features were developed to process raw signals into informative inputs to the machine learning models.

Loadings and the values of feature importance provided by the Random Forest and XGBoost were consistent in revealing that the tool position and velocity were often the most important variables in the tool model decisions. This does not come as a surprise, as deviations in airports and positions are prompt signs of abnormal behavior or deviations. Additionally, it was demonstrated that force and torque characteristics were significant in case of the robot acting in a strange way of contact, if they met more resistance or a collision. These indicators are the direct consequence of stress on the robot and as such the indicators are an intrinsic signal with regards to risk (threat) or even collision chance. Moreover, correlation features had a descriptive nature because the features reflected “predictable connections among the sensors. With robot fitting, such relationships are comparatively constant. When degradation sets in, e.g. wear or miscalibration, the correlations vary and the models can detect it. The issues find very high accuracy and subsequent discrimination between normal and abnormal data of our product models, mean that the features engineered were effective to model the behavior of the robot system under study.

The time-frequency and Correlation feature-based features were therefore determined to be the most descriptive of early faults and component of performance of industrial robots, particularly those features of tool kinematics and force/torque measurements.

5.4 Performance of Classical Machine Learning Models

To apply machine learning to create fault classification and RUL prediction. The classical machine learning classifiers Logistic Regression, Random Forests, Support Vector Machine (SVM), and XGBoost perform on multi-user robot data on classification and RUL tasks. Two of these models were trained on the trained engineered feature sets and with a sequential train/validation/test split as it would happen with a time element not factoring in the information to the data. Amongst all these tests in the detection of anomalies, Logistic Regression has been shown to perform very poorly with a low precision and F1. The model was not comprehensible enough to be trained on the tricky associations between features and target due to the nature of robotic systems. SVM had a very high recall that was almost capable of identifying all the anomalies, only that there were numerous false positive results in its findings. It implies that SVM would be quite intense in identifying potentially malfunctioning components which is good to a safety perspective but bad in the field since it would trigger excessively too many false alarms and interfere with activities. On the contrary, random Forest and XGBoost demonstrated a more balanced performance in general. In the case of the two models, high accuracy and high values of ROC-AUC and PR-AUC were found which implies that the normal data and the abnormal data could be perfectly separated. XGBoost achieved the F1 score with the highest score, and so, we know that it is creating a good balance between the precision/recall trade-off. When it comes to RUL remains prediction, we see that both R square of Random Forest and XGBoost were low; as previously, RF and XGB were able to accurately predict remaining useful life.

These findings suggest that even the machine learning methods are highly effective to the task of both detecting anomalies and predicting RUL by industrial robotics; nevertheless, the model choice is essential, since Random Forest and even better XGBoost are introduced to date to be the most suitable models in the given case, whereas the use of Logistic Regression is to be avoided at all costs and SVM must be used with caution due to a high false.

5.5 Comparative Evaluation of Models

To assess and contrast the performance of the chosen machine learning models against a selection of evaluation metrics, it is necessary to compare models to fully comprehend their ability to predict equipment breakdown and identify the most appropriate model to use in predictive maintenance. The metrics of classification and prediction of RUL were accuracy, precision, recall, F1-Score, ROC-AUC and PR-AUC and R2, respectively. Random Forest and XGBoost were performing consistently well as per most of the measures with SVM recording greater recall than accuracy.

These variations were summarized in the model comparison table quite well. An instance of it is Logistic Regression that had an accuracy of 78.84, and you can only obtain an accuracy of 0.0832, whereas XGBoost model had an accuracy of approximately 98.68 with a much better F1-score and immediately with better ROC-AUC and PR AUC Figures. It was also found that the random forest had very low false positive and high precision which may lead to the conclusion that the random forest is more reliable when the aim is to reduce false alarms. Instead, XGBoost had the most ideal performance with regards to learning what is a real anomaly and the false positives.

Comparison of these models in many dimensions have invariably confirmed the point in that and can serve well to predictive maintenance in industrial robots.

5.6 Decision-Support Framework, Uncertainty, and Generalization

The framework was designed to combine the outputs of the anomaly detecting and the RUL predicting. They are the aim that the system does not just tell us whether something looks normal but also arrives at an approximation of the amount of remaining time before any failure will appear like this one Both Random Forest and XGBoost returned well-calibrated classification probabilities and RUL estimate derived in retrospectively. The risk indicators of a real maintenance system might be the system outputs. We employed probability scores and model stability over various splits and users. The decision-support system is then able to minimize late detection as well as excessive false-positive alarms by imposing proper thresholds. One of such examples is the use of a decision threshold whereby a high decision threshold is used when the cost of closing a robot is exceptionally high, whereas a less significant decision threshold can be used in safety-sensitive applications.

The cross-user assessment of both the Random Forest and the XGBoost solved transfer to the different robot users. It was also not the case that their performance had gone down when they were tested on robots and users not revealed in training. This is relevant in the real industrial context, since robots are required to work under different setups and handled by different operators. On the other hand, SVM and Logistic Regression suffer from more performance variability from one user to another, they are less resilient. Accordingly, Random Forest and XGBoost derived decision-support framework can give stable and scalable suggestions.

5.7 Theoretical and Practical Implications

Contributing to the field of predictive maintenance, in a theoretical perspective, the paper proves that the conventional ML algorithms with the assistance of multi- sensors rich attribute engineering of robot data might offer the high quality of anomaly identification and RUL prognosis. It revealed that one does not require very deep networks to deliver good performance, provided the data is working well before processing, and appropriate features are extracted. It also demonstrates that it is important to evaluate models on metrics that are appropriate on imbalanced datasets, like PR-AUC and F1-score, instead of accurate alone.

To industry, the report marked out a definite course, upon the practical side. One can also get sensor measurements of the robots internally and do the identical preprocessing / aspect of engineering in order to train either of the Random Forest or XGBoost models so that predictive maintenance can be deployed instead of unplanned downtime. The decision-support structure presents the way the outputs of these models might be translated into condition-based maintenance actions, relying on anomaly scores and RUL estimates to plan inspections, repairs or replacement of the components.

5.8 Limitations of the Study

The study has certain weaknesses even with its strong findings. The data was of a specific environment and specific tasks using robots so the models may require re-training or reconfiguring to be useful with highly dissimilar robots or industries. We also did not consider models of deep learning (LSTM, GRU, Temporal Convolutional Networks) which can possibly capture long-term temporal dependencies even more effectively by using machine learning classical methods. In addition, full uncertainty conscious full probabilistic models like Monte Carlo dropout and Bayesian modelling were not strictly explored. Lastly, the experiments were not conducted in an offline and realistic (live production) situation implying that no stringent real-time performance and latency analysis was directly shown.

5.9 Conclusion

A predictive maintenance system of industrial robotics with AI enhancement: A case study of sensor image processing and machine learning. The paper has made some important findings: multi-sensor internal signals on the robot could be used to identify anomalies effectively, engineered features in the time, frequency and correlation domain could be more predictive of fault than others, and classical models of machine learning model ability could be used successfully whether in prediction of anomalies or prediction of RUL. This decision-support framework provides an illustration of the application of these models in the real maintenance decision process. The basis of our work is the predisposition of AI-based predictive maintenance in the industrial robotic system, and the step forward towards making them accurate, efficient and capable of handling industry 4.0 / Robotics 4.0 necessities.

6 Recommendation and future work

6.1 Introduction

Industrial robots with their sources provide rich sensor data which can be effectively used in detecting anomaly and predicting RUL. It is also advisable to start monitoring internal sensor streams when the system is running normally and store them, not only in the alleged events. Progressive improvement of predictive maintenance systems will be made possible by availability of longitudinal historical data. Practitioners should pay attention to the machine learning models that were effective in our research, namely Random Forest and XGBoost which were best in detecting anomalies besides RUL prediction. The models are rapid and noise resistant, they can handle multi-sensor information without use of complicated hardware. Maintenance teams may use these models as panels on their monitoring dashboards and when the predicted anomalies are known they may e.g. trigger an initial checkup before a failure turns into a critical failure. It is suggested that organizations adopt the decision-support system invented in this dissertation. This platform can assist the maintenance engineer to plan the repair in advance, thereby preventing unexpected downtime and further streamlining the process of distributing maintenance efforts by mobile cameras by combining anomaly detection and RUL prediction. It might be a good idea to incorporate the thresholding of anomaly scores as well as RUL (remaining useful life), which would help to prioritize interventions based on their risk and time. It is generally recommended to update or restrain the models on designated schedules to ensure proper prediction of their work as robots age, or that there is a variation over time in the tasks associated with work cell.

6.2 Technical Recommendations for Robotics Engineers

Engineers working on robotics must be able to stabilize sensor calibration tone, synchronization and data sampling rate throughout the life of the unit. A slight variation in the rate of sampling between samples could cause distortion of the data and decrease the capability to detect anomalies. The patterns in correlation that are found here should also be examined by engineers who can therefore predict the early signs of system wellbeing whereby the tool position, percentage tool velocity, joint angles and FTW readings will give the tool definition.

The use of feature engineering was also key to improving the performance of the models and predictive systems engineers are advised to incorporate a combination of both time domain (and frequency domain) and correlation-based features. The sensitivity analysis and PCA results obtained in this work are indicative of those sensors, i.e., tool kinematics and force/torque, which have a significant contribution towards predictive accuracy. These sensor channels are also to be handled by engineers with special attention, and they might be interested in embedding light-weight real time preprocessing scripts to make on-the-fly features computation in the robot controller or edge devices.

Another consideration that robotics engineers need to consider is the construction of predictive maintenance analytics using robot controllers or digital twin platforms. Within the scope of predictive models, digital twin environments can model failure conditions and compare robot behavior in responses to different stress capabilities. The combination enhances the quality of forecasting and evaluation of robot health.

6.3 Future Research Directions

Certainly, these results suggest that it is correct to affirm that classical machine learning models do have large potential in predictive machine maintenance, even though there are multiple open spaces to be addressed. It is also possible to use deep learning-based models such as LSTM, GRU, TCN or transformer-like models, which can possibly capture long-term temporal dependencies better than classical models. Deep learning may also be useful with performance on highly nonlinear data sets or tasks.

Bayesian neural networks or Monte Carlo dropout regimes are uncertainty-conscious predictive models that could be studied in the future. Instead of just a prediction, as is the case with such models, they provide probability distributions, which can significantly increase the reliability of maintenance decisions. This would be useful so that operators can have an idea of the confidence of the predictions of anomalies theory and RUL estimations so that I may reduce false alarms and miss detections.

The other significant area that will be pursued in the future will be the generalizability of this method to be applied to more diverse datasets that comprise more than one brand of robots, more than one task and diverse environmental conditions. It would make predictive models more generally applicable and help deploy inter-factories. Predictive features can be extended to augment this initial fault detection rate by adding sensor data of known outside (e.g. vibration sensor, thermal imaging or camera-based surveillance) that can further assist in predicting the fault.

Another direction that is interesting is the real-time production. Applications of predictive maintenance systems on edge devices could be able to make decisions automatically and stay informed on a regular basis via robot controller or the cloud platform. This may transform predictive maintenance into a self-mending system instead of an analytical offline practice. The application of predictive analytics to the use of digital twins could also be an aspect under further research to get a full virtual image of robotic health and performance.

Finally, in their work in the future, the authors are planning to research the cost-conscious maintenance optimization where the Predictions are merged with the operational costs, the likelihood of downtime and the production schedules. This would lead to a decision-support system with a wider spectrum of performance with optimum technical efficiency and cost effectiveness.

References

Gao, Z., Wanyama, T., Singh, I., Gadhri, A., & Schmidt, R. (2020). From Industry 4.0 to Robotics 4.0 – A conceptual framework for collaborative and intelligent robotic systems. *Procedia* Manufacturing, 46, 591–599. https://doi.org/10.1016/j.promfg.2020.03.085

Dzedzickis, A., Subacitite-Zemaitiene, J., Šutinys, E., Samukaite-Bubnienė, U., & Bučinskas, V. (2022). Advanced applications of industrial robotics: New trends and possibilities. *Applied Sciences*, 12(1), 135. https://doi.org/10.3390/app12010135

Borgi, T., Hidri, A., Neef, B., & Naceur, M. S. (2017). Data analytics for predictive maintenance of industrial robots. In *2017 International Conference on Advanced Systems and Electric Technologies (IC_ASET)* (pp. 412–417). IEEE.

Izagirre, U., Andonegui, I., Landa-Torres, I., & Zurutuza, U. (2022). A practical and synchronized data acquisition network architecture for industrial robot predictive maintenance in manufacturing assembly lines. *Robotics and Computer-Integrated Manufacturing*, 74, 102287. https://doi.org/10.1016/j.rcim.2021.102287

Sujatha, M., Priya, N., Beno, A., Sheeba, T. B., Manikandan, M., Tresa, I. M., Jose, P. S. H., Peroumal, V., & Thimothy, S. P. (2022). IoT and machine learning-based smart automation system for Industry 4.0 using robotics and sensors. *Journal of Nanomaterials*, 2022, Article ID 6807585. https://doi.org/10.1155/2022/6807585

Kanawaday, A., & Sane, A. (2017). Machine learning for predictive maintenance of industrial machines using IoT sensor data. In *2017 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS)* (pp. 87–90). IEEE.

Dalenogare, L. S., Benitez, G. B., Ayala, N. F., & Frank, A. G. (2018). The expected contribution of Industry 4.0 technologies for industrial performance. *International Journal of Production Economics, 204,* 383–394. https://doi.org/10.1016/j.ijpe.2018.08.019

Ullah, A. S. (2019). Modeling and simulation of complex manufacturing phenomena using sensor signals from the perspective of Industry 4.0. *Advanced Engineering Informatics, 39,* 1–13. https://doi.org/10.1016/j.aei.2018.11.003

Nikolakis, N., Maratos, V., & Makris, S. (2019). A cyber physical system (CPS) approach for safe human-robot collaboration in a shared workplace. *Robotics and Computer-Integrated Manufacturing, 56,* 233–243. https://doi.org/10.1016/j.rcim.2018.10.003

Xu, W., Liu, Q., Xu, W., Zhou, Z., Pham, D. T., Lou, P., Ai, Q., Zhang, X., & Hu, J. (2017). Energy condition perception and big data analysis for industrial cloud robotics. *Procedia CIRP, 61*, 370–375. https://doi.org/10.1016/j.procir.2016.11.164

V, S., Ramaswamy, S., & Butail, S. (2016). Training data selection criteria for detecting failures in industrial robots. IFAC-Papers Online, 49(1), 385 390.https://doi.org/10.1016/j.ifacol.2016.03.084

Chivarov, N., Chikurtev, D., Markov, E., Chivarov, S., & Kopacek, P. (2018). Cost oriented tele-controlled service robot for increasing the quality of life of elderly and disabled – ROBCO 18. IFAC-PapersOnLine, 51(30), 192–197. https://doi.org/10.1016/j.ifacol.2018.11.285

Javaid, M., Haleem, A., Singh, R. P., & Suman, R. (2021). Substantial capabilities of robotics in enhancing industry 4.0 implementation. *Cognitive Robotics, 1*, 58–75. https://doi.org/10.1016/j.cogr.2021.06.001

Izagirre, U., Andonegui, I., Egea, A., & Zurutuza, U. (2020). A methodology and experimental implementation for industrial robot health assessment via torque signature analysis. *Applied Sciences, 10*(21), 7883. https://doi.org/10.3390/app10217883

Buizza Avanzini, G., Ceriani, N. M., Zanchettin, A. M., Rocco, P., & Bascetta, L. (2014). Safety control of industrial robots based on a distributed distance sensor. *IEEE Transactions on Control Systems Technology*. Advance online publication. https://doi.org/10.1109/TCST.2014.2300696

Carvalho, T. P., Soares, F. A. A. M. N., Vita, R., Francisco, R. da P., Basto, J. P., & Alcalá, S. G. S. (2019). A systematic literature review of machine learning methods applied to predictive maintenance. *Computers & Industrial Engineering, 137*, 106024. https://doi.org/10.1016/j.cie.2019.106024

Cui, Y., Kara, S., & Chan, K. C. (2020). Manufacturing big data ecosystem: A systematic literature review. *Robotics and Computer-Integrated Manufacturing*, 62, 101861. https://doi.org/10.1016/j.rcim.2019.101861

Eylemo, L. D., Gjerstad, T., Grotti, E. I., & Sziebig, G. (2020). Trends in smart manufacturing: Role of *humans* and industrial robots in smart factories. *Current Robotics Reports*, 1(1), 35–41.https://doi.org/10.1007/s43154-020-00006-5

He, Y., Guo, J., & Zheng, X. (2018). From surveillance to digital twin: Challenges and recent advances of signal processing for the industrial Internet of Things. *IEEE Signal Processing* Magazine, 35(5), 120–129. https://doi.org/10.1109/MSP.2018.2842228

Hwang, J., & Tani, J. (2018). Seamless integration and coordination of cognitive skills in humanoid robots: A deep learning approach. *IEEE Transactions on Cognitive and Developmental Systems*, 10(3), 434–452. https://doi.org/10.1109/TCDS.2017.2714170

Izagirre, U., Andonegui, I., Eciolaza, L., & Zurutuza, U. (2021). Towards manufacturing robotics accuracy degradation assessment: A vision-based data-driven implementation. *Robotics and Computer-Integrated Manufacturing*, 67, 102029. https://doi.org/10.1016/j.rcim.2020.102029

Khalid, A., Kirisci, P., Khan, Z. H., Ghrairi, Z., Thoben, K.-D., & Pannek, J. (2018). Security framework for industrial collaborative robotic cyber-physical systems. Computers in Industry, 97, 132–145. https://doi.org/10.1016/j.compind.2018.02.009

Li, P., & Liu, X. (2019). Common sensors in industrial robots: A review. Journal of Physics: Conference Series, 1267(1), 012036. https://doi.org/10.1088/1742-6596/1267/1/012036

Li, H., & Savkin, A. V. (2018). Wireless sensor network based navigation of micro flying robots in the industrial internet of things. IEEE Transactions on Industrial Informatics. Advance online publication. https://doi.org/10.1109/TII.2018.2825225

Luo, W., Hu, T., Ye, Y., Zhang, C., & Wei, Y. (2020). A hybrid predictive maintenance approach for CNC machine tool driven by Digital Twin. Robotics and Computer-Integrated Manufacturing, 65, 101974. https://doi.org/10.1016/j.rcim.2020.101974

Maurice, P., Malaisé, A., Amiot, C., Paris, N., Richard, G.-J., Rochel, O., & Ivaldi, S. (2019). Human movement and ergonomics: An industry-oriented dataset for collaborative robotics. *The* International *Journal of Robotics Research*. Advance online publication. https://doi.org/10.1177/0278364919882089

Michau, G., Hu, Y., Palme, T., & Fink, O. (2019). Feature learning for fault detection in high-dimensional condition monitoring signals. *Mechanical Systems and Signal Processing, 134*, 106334. https://doi.org/10.1177/1748006X19868335

Pedersen, M. R., Nalpantidis, L., Andersen, R. S., Schou, C., Bogh, S., Krüger, V., & Madsen, O. (2016). Robot skills for manufacturing: From concept to industrial deployment. *Robotics and Computer-Integrated Manufacturing, 37, 282–291. https://doi.org/10.1016/j.rcim.2015.04.002

Henderson, J., & Sanders, M. (2025). *AI driven predictive maintenance: Reducing downtime and enhancing productivity in manufacturing environments*. Preprints. https://doi.org/10.20944/preprints202504.0602.v1

Elsisi, M., Mahmoud, K., Lehtonen, M., & Darwish, M. M. F. (2021). Reliable Industry 4.0 based on machine learning and IoT for analyzing, monitoring, and securing smart meters. *Sensors*, 21(2), 487. https://doi.org/10.3390/s21020487

Çınar, Z. M., Nuhu, A. A., Zeeshan, Q., Korhan, O., Asmael, M., & Safaei, B. (2020). Machine learning in predictive maintenance towards sustainable smart manufacturing in industry 4.0. Sustainability, 12(19), 8211. https://doi.org/10.3390/su12198211

Wijaya, T., Caesarendra, W., Pappachan, B. K., Wee, A., Roslan, M. I., & Tjahjowidodo, T. (2017, December). Robot control and decision making through real-time sensors monitoring and analysis for Industry 4.0 implementation on aerospace component manufacturing. In *2017 IEEE International Conference on Cybernetics and Computational Intelligence (CyberneticsCom)* (pp. 25-30). IEEE.

Zhang, W., Yang, D., & Wang, H. (2019). Data-driven methods for predictive maintenance of industrial equipment: A survey. *IEEE Systems Journal*, 13(3), 2213–2227. https://doi.org/10.1109/JSYST.2019.2905565

Zonta, T., da Costa, C. A., da Rosa Righi, R., de Lima, M. J., da Trindade, E. S., & Li, G. P. (2020). Predictive maintenance in the industry 4.0: A systematic literature review. *Computers & Industrial Engineering*, 150, 106889. https://doi.org/10.1016/j.cie.2020.106889