

Cost-Efficiency and Cost-Effectiveness of XAI in Predictive Maintenance

Mrs.G.L.Dayana¹, Mr.Asok kumar², Mr.G.Sabarivasan³, Mr.V.Vijay⁴, Mr.M.Vinithkumar⁵

^{1,2} Assistant professor, V.S.B College of Engineering Technical Campus, Coimbatore, Tamil Nadu, India

^{3,4,5} Undergraduate Student, V.S.B College of Engineering Technical Campus, Coimbatore, Tamil Nadu, India

Emails: dddayanait17@gmail.com¹, ashokcse7192@gmail.com², sabari262004@gmail.com³, vijay1292vijay@gmail.com⁴, sivavinith2005@gmail.com⁵

Abstract

Predictive maintenance has become a fundamental approach in industrial environments to monitor the condition of equipment in order to prevent unexpected failures in machines through the data analysis of the sensors. Conventional artificial intelligence models can predict potential failures based on parameters such as temperature, vibration and pressure. However, these models are often black boxes and offer little interpretability in terms of why they are making the predictions. The lack of explanation can result in decreased trust among the maintenance personnel, as a consequence, unnecessary repairs, longer diagnostic time, and inefficient resources allocation. Explainable Artificial Intelligence overcomes this limitation and gives the ability for understanding the contributing factors of the predicted failures and this leads to more informed and targeted maintenance actions. This research investigates the economic aspects that integrating explainability into prediction maintenance systems poses, both as cost-efficiency and cost-effectiveness. Cost-efficiency is measured in terms of the reduction of unnecessary maintenance interventions and downtime, and cost-effectiveness is measured in terms of the long-term financial benefits in comparison to initial implementation efforts. By facilitating clarity of decisions and improving the transparency of the information, XAI has the potential to improve operational reliability and optimize maintenance strategies. The analysis shows how explainable models can play a role in the sustainable operation of industry by balancing technical performance against economic goals.

Keywords Anemia detection, Deep learning, Image classification, EfficientNetB⁷, Medical diagnostics.

1. Introduction

Industrial systems are increasingly relying on data-driven approaches to ensure the operational reliability and minimize the occurrence of unexpected equipment failures. Predictive maintenance has become a viable solution to use sensor-generated data (e.g. temperature, vibration, pressure, and acoustic signals) to measure the machine's health and anticipate potential failures before breakdowns occur. By moving from reactive and time-based maintenance strategies and moving towards condition-based monitoring, organizations hope to reduce their downtime, increase their equipment life, and optimize their maintenance resources. Artificial intelligence models are at the heart of this change as they are used to analyse vast amounts of real-time and historical data to identify patterns relating to abnormal machine behaviour.

Despite their prediction power, many typical artificial intelligence models are black box systems. They give outputs of whether a machine is likely to fail but do not give any insight into the factors that affect the prediction. This lack of interpretability presents practical problems in industrial environments where maintenance choices inevitably come down to cost, safety and continuation of operation. When the reason for a prediction is not understood, maintenance personnel might conduct additional checks, replace the components unnecessarily, or ignore the prediction because of a lack of faith in the system. Such outcomes can decrease the economic benefits that predictive maintenance is expected to bring and can result in inefficient use of time and financial resources. Explainable Artificial Intelligence brings transparency to predictive models by detecting and

showing the most important variables that are playing a role in making each prediction. Instead of just giving a probability of failure, explainable systems clarify whether certain parameters such as increased trends in vibration or rising temperatures are responsible for the risk assessment. This interpretability is supportive of informed decision-making, builds confidence in automated systems, and facilitates collaboration between human expertise and machine intelligence. In maintenance contexts, better thinking may be used to help technicians zero-in on the actual cause of the problem instead of taking broad or precautionary approaches. From the economic point of view, the integration of explainability can affect cost-efficiency and cost-effectiveness. Cost-efficiency is related to minimizing unnecessary repairs, downtime, and labor and spare part usage. Cost-effectiveness involves evaluating if the long-term financial savings and operational improvements are worth the effort and resources needed to establish explainable systems. This study investigates the role of Explainable Artificial Intelligence in predictive maintenance and assesses the potential of Explainable Artificial Intelligence to bring financial benefits while ensuring technical reliability. By ensuring a relationship between interpretability and operational goals, explainable systems could help to produce more sustainable and economically balanced industrial maintenance approaches.

2. Related Works

Explainable Artificial Intelligence has become an important research problem to help make machine learning models more transparent and interpretable. Mathew et al. [1] talk about some new emerging techniques in explainable artificial intelligence aimed to increase human understanding of AI-driven decisions. Their work emphasizes methods for providing model-level and instance-level explanations, and the need for interpretability in systems where outcomes of the decision process affect operational and strategic actions. The study sets the base for the explainability of artificial intelligence to bolster trust, accountability, and usability of artificial intelligence across domains. The need for explainability in high-stakes applications is also highlighted by Alkhanbouli et al.

[2], who perform a systematic review of explainable artificial intelligence for disease prediction. Although focused on healthcare, the study highlights general challenges associated with black box models, such as lack of transparency, lack of user confidence, and issues associated with adoption. The authors see a need for structured evaluation frameworks to evaluate the interpretability techniques. These findings have relevance for industrial predictive maintenance in which decisions have a direct impact on safety, cost and continuity of operation. Predictive maintenance in manufacturing has been thoroughly reviewed by Benhanifia et al. [3], which analyzes the existing practices and technological advances in the sector. Their review emphasises combining sensor information, machine learning algorithms and condition monitoring systems to predict equipment failures. However, the study also presents challenges, including issues with data quality, implementation complexity, and limited interpretability of advanced predictive models. The lack of explainable mechanisms can make the use of predictive systems less practical despite their high accuracy. Mikołajewska et al. [4] investigate the role of generative AI in digital twins in fault diagnosis for Industry 4.0 and 5.0 environments. Their work shows how digital twin architectures combine AI-based fault detection to simulate the behavior of equipment and assist in predictive maintenance strategies. While generative models are helping to improve diagnostic abilities, the study suggests that the lack of understanding of what models give out is a concern, especially when implementing complex AI elements into industrial ecosystems. This provides a further argument for explainability of advanced maintenance frameworks. Garcia et al. [5] NLP assisted condition monitoring and predictive maintenance: signal processing techniques, hybrid models and implementation challenges. The authors identify major trends such as the use of deep learning approaches and hybrid architectures that incorporate statistical and machine learning approaches. However, they also point to the possible limitation of interpretability due to the growing complexity of models. Implementation challenges include how predictive systems can be integrated into existing industrial workflows, where explainability can play

a vital role in enabling acceptance among practitioners. Condition monitoring and fault diagnosis in industrial robots are studied in Lei et al. paper [6], where methodologies for detecting anomalies and diagnosing failures for robotic systems are reviewed. The study covers the signal analysis, feature extraction and machine learning techniques with robotic maintenance. As diagnostic algorithms become more autonomous in robotic systems, the transparency of these algorithms becomes crucial to ensure reliability and safety. The review proves the importance of having interpretable diagnostic frameworks to support better maintenance planning and reduce the uncertainty in automated systems. Economic considerations are also at the core of the discussion of intelligent systems. Bhatnagar [7] discusses the topic of optimizing costs in FinTech through microservices and serverless architectures. Although it is set in a financial technology topic, the research gives insights into cost-efficiency, resources allocation, and scalable system design. These principles apply to predictive maintenance platforms, where the design of the infrastructure as well as the computational efficiency determine the overall cost-effectiveness. The alignment in technical architecture and financial optimization is very critical in evaluating the AI-based maintenance systems. The use of artificial intelligence in renewable energy systems is discussed by Suci et al. [8], focusing on the predictive maintenance and energy optimization. The role of AI in monitoring renewable energy assets and minimizing operational disruptions is highlighted by the authors. Their findings have shown that predictive models are part of improving system performance, but has also raised issues with the transparency of models and their application in the real world. In energy-critical infrastructures, the ability to give operators an interpretable model that helps them make informed maintenance decisions and manage risk. The discussion on preserving performance versus interpretability has existed for a long time, but Kruschel et al. [9] tests interpretable machine learning models and rebuts the traditional trade-off assumption. However, their analysis implies that interpretable models are capable of competitive performance under certain conditions.

This perspective is important for predictive maintenance applications, where a high level of accuracy has to be balanced by operational transparency. The findings are encouraging the development of explainable models that do not compromise predictive reliability. Zhao et al. [10] show an application of interpretable machine learning for environmental risk assessment, in this case, identifying hazards in European surface waters. By using interpretable models, the study gives clarity about the important environmental factors. Although the domain is different from industrial maintenance, the methodological approach describes how explainable techniques help to improve the understanding of complex systems. The ability to detect important contributing variables is similar in application to the requirements of predictive maintenance systems sensing multiple inputs to a system. Collectively, the reviewed literature says that predictive maintenance has come a long way with machine learning and digital transformation. However, the growing complexity of models has presented challenges in terms of transparency, trust and implementation. Explainable Artificial Intelligence offers a potential way to overcome these concerns through interpretable explanations of model behaviour. While there has been considerable research into explainability in healthcare, environmental science and business systems, the economic aspects of explainability in industrial predictive maintenance continues as an emerging field of research. The integration of XAI into maintenance frameworks holds potential for advancing better decision-making processes, cost-efficiency, and long-term operational sustainability.

3. Methodology

The proposed methodology is to assess the cost-efficiency and cost-effectiveness of Explainable Artificial Intelligence for predictive maintenance systems. The framework combines sensor-based condition monitoring, machine learning-based prediction and explainability mechanisms to understand the technical performance and the economic impact. The methodology is organized under the following subsections shown in Figure 1.

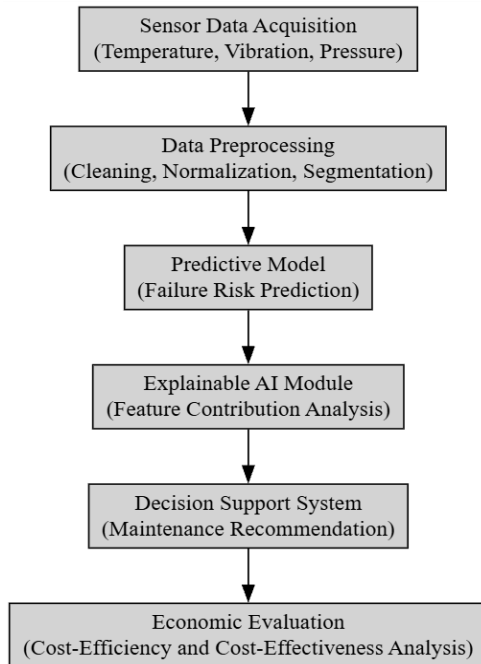


Figure 1 Architecture of the Explainable Predictive Maintenance Framework.

3.1. Data Acquisition And Condition Monitoring.

The first phase involves the constant monitoring of the industrial equipment by use of embedded sensor networks. Operating parameters measured by sensors on machinery include temperature, vibration, pressure, acoustic, rotational speed and load conditions. These parameters suggest very important indicators of machine health and performance stability. Information is relayed to a centralized storage system where it is recorded in the process of time. A preprocessing is to be carried out to guarantee quality and reliability of data. This involves noise reduction of sensor signals, missing or corrupted samples, and time stamping of samples to synchronize with other sensors. The methods used are normalization and scaling that are used to ensure consistency among various units of measurement. Time-Series segmentation is done to transform raw continuous signals into highly structured intervals to be used in extracting features and predicting. Monitoring of condition at this step forms the baseline data of the predictive system. Consistency of data acquisition is imperative as faulty or non-considerate input can significantly decrease the

predictive accuracy as well as the stability of the explanations that are produced in the subsequent phases.

3.2. Feature Engineering And Data Preparation.

Once sensor data has been processed meaningfully, it becomes possible to extract meaningful features in order to represent equipment behavior. Time-series signals are used to get such statistical indicators as mean, variance, trend patterns, and frequency-domain characteristics. The feature selection methods will be used to determine the most applicable indicators that are related to machine degradation or abnormal behavior. The labeling of the data is carried out based on the historical records of maintenance that identify the normal operating conditions and the verified failed conditions. To make the objective model validation, the dataset is split into training and testing subsets. This step will make sure that the predictive model is trained on patterns that represent it and it still has the capacity to extrapolate to unknown conditions of operation. Cautious feature engineering enhances predictive performance as well as interpretability. The explanations produced by explainable techniques are made more understandable by features that can be related to physically meaningful machine parameters.

3.3. Development Of Predictive Model.

Here, a machine learning algorithm is created to predict the possible equipment failures. The extracted features are taken through the model which generates a prediction that would suggest whether a machine is likely to malfunction during a given period of operation. The predictive output can be provided as an outcome of binary classification or a score of risk of failing. Model training Optimizes parameters based on past labeled data. The standard classification measures including accuracy, precision, recall and detection capability are used to do performance evaluation. These measures guarantee that the model is predictable in differentiating between the healthy and unhealthy operating environments. The forecasting element is the fundamental part of the maintenance system. The model at this point of time, however, is a predictable system that is not interpretable. Strong predictive

reliability should also be achieved prior to the incorporation of the mechanisms of explanation where the explainability should supplement and not substitute the predictive performance

3.4.Explainable Artificial Intelligence.

The fourth stage takes into account Explainable Artificial Intelligence methods to find the meaning of predictive results. After the predictive model develops an indication of a failure, explainability techniques are used to find the contribution of each feature to the prediction. These methods are used to determine the strongest operation parameters that determine the failure assessment. The explanations are modeled at the instance level, i.e. every prediction is provided with a description of the factors that contributed to it. As an illustration, high amplitude of vibration or unnatural temperature change can be determined as the major factors. The provided explanation is organized in an understandable format that is not necessitating sophisticated data science knowledge in order to be interpreted by maintenance engineers. Transparency is improved because the explainability aspect builds trust in automated systems. Maintenance staff can know the logic behind prediction and this will minimize uncertainty and lead to knowledgeable action. This step is a direct response to the shortcoming of black-box predictive models.

3.5.Support Framework Maintenance Decision.

After production and explanation generation it incorporates the outputs into a structured decision support mechanism within the system. In case of a possible malfunction, the system displays the risk assessment and the operational factors that have caused it. This information is examined by maintenance engineers to take corrective measures. Technicians will have the opportunity to concentrate on the exact subsystem of abnormal indicators rather than carry out general inspection or even replace a complete assembly. This is a focused maintenance strategy that minimizes unwarranted labor, usage of spares and production downtimes. Maintenance activities and the performance of the machine after used are also captured in the decision support structure. Continuous improvement of both predictive accuracy and explanation relevance is also

aided by this feedback loop. With time, the system will transform to a collaborative platform that involves the coordinated efforts of human knowledge and machine intelligence

3.6.Economic Analysis And Cost Evaluation.

An overall economic analysis is done to determine the economic cost and benefit of introducing explainability to predictive maintenance. Cost-efficiency is studied through the analysis of decreases in the needless maintenance interventions, the usage of spare parts, the amount of time used on diagnostics, and time spent on the downtime. The assessment of cost-effectiveness is based upon the savings of the long-term operational costs in comparison to the effort expended to establish and support the explainable framework. A comparative study is conducted of the traditional predictive maintenance systems and the propose explainable method. The indicators are frequency of maintenance, occurrence of downtimes, and the pattern of resources utilization. This analysis will find out whether the transparency that Explainable Artificial Intelligence has brought about leads to quantifiable economic gains and operational reliability.

3.7.Performance Validation And Impact Assessment.

The last step is to confirm the technical and economic performance. After explainability is incorporated, predictive accuracy is again tested with the aim of making sure that interpretability mechanisms do not lead to a reduction in reliability. The feedback of users is gathered and assessed to determine their clarity, ease and confidence in their decisions. Operational impact is measured against the aspects of better response time, better resource allocation, and better system trust. The proposed methodology offers a unified approach to the evaluation of the role of Explainable Artificial Intelligence in enhancing cost-efficiency and cost-effectiveness within the industry of predictive maintenance systems based on the combination of predictive modeling, explanation habits, structured decision support, and economic assessment.

3.8.Results And Discussion

The application of the proposed explainable predictive maintenance framework has shown that

there are measurable gains in the maintenance decision clarity and operational efficiency as compared to the conventional predictive systems. The predictive element was able to detect possible failure states with the help of sensor-based operation data, whereas the addition of explainability gave a clear understanding of the factors that caused each prediction. Maintenance workers could tell the difference between indicators of critical risks and the minor operational fluctuations, which allowed being more specific in their intervention planning. The existence of instance-level explanations made it less uncertain because of automated outputs and decreased the use of the further manual verification process. This meant that there was less unnecessary inspections and precautionary replacement of components and more focused and planned maintenance was implemented. The framework supported predictive reliability and increased interpretability meaning that the incorporation of explainability did not affect the technical performance shown in the Table 1.

Economic Impact	Moderate cost optimization	Improved cost-efficiency and long-term cost-effectiveness
------------------------	----------------------------	---

Table 1 Comparison of Conventional vs. Explainable Predictive Maintenance:

Parameter	Conventional Predictive Maintenance	Explainable Predictive Maintenance
Model Interpretability	Provides prediction only	Provide prediction with explanation
Decision Transparency	Limited clarity in reasoning	Clear contributing factor
Maintenance Precision	May lead to broad inspections	Supports targeted corrective actions
Resource Utilization	Possible unnecessary repairs	Optimized use of labor and spare parts

Economically, the explicable framework had enhanced cost-efficiency by means of streamlined resource use. Explanatory insight-based targeted repairs minimized the consumption of spare parts and decreased diagnostic time. The machine downtime based on vague predictions was reduced since the maintenance decisions were backed up by clear arguments. The explainable approach also enabled a faster response time and enhanced coordination between the maintenance teams and the outputs of the system as compared to traditional black-box systems. The resulting decrease in both unwarranted interventions and time off over prolonged versions of operation was a part of enhanced cost-efficiency. The need to incorporate the concept of explainability brings about extra computation and implementation factors, but the long-term stability of operations and the increase in the level of confidence in the decision process balance the use of explainability. As mentioned in the discussion, transparency is important to achieve all the potentials of predictive maintenance systems in relation to the economy. Although the traditional models can give good predictive results, they cannot be used in practical applications because of their absence of interpretation, which can decrease the levels of confidence among industrial operators. The findings indicate that explainability enhances human-computer collaboration, which is useful in maintaining sustainable maintenance strategies. The suggested framework will improve financial and operational performance within industrial settings because it balances predictive accuracy with operational transparency.

4. Discussion And Future Work

The future study can elaborate the proposed explainable predictive maintenance framework with the addition of the adaptive and real-time learning processes. The inclusion of continuous learning model would facilitate the system to keep predictive

patterns current over time as new operational information is obtained thus enhancing long-term reliability in changing industrial conditions. It can also be improved by introducing more complex hybrid models that can integrate machine learning and statistical methods to provide a better predictive stability without losing any interpretability. Moreover, in the future, one can work on the enhancement of the explanation outputs clarity and usability by means of user-friendly visualization interfaces designed specifically to serve the maintenance staff. Formatted dashboards depicting feature entries, trend changes, and risk signals in a simplified form might increase the confidence of decisions and lessen the cognitive burden. The other beneficial improvement is the mechanism of incorporating the framework with the digital twins, whereby, predicted failures and corrective measures can be validated through simulations and physical intervention avoided. This kind of integration would help in proactive planning and minimise any disruption of operations. Multi-machine and plant-level monitoring environments would also enhance scalability by expanding the system to accommodate the monitoring environment and allow co-ordinated maintenance strategies of interrelated assets. Future improvement can focus on more profound economic analysis and inter-domain transferability of understandable predictive maintenance systems. Future research may include an in-depth lifecycle cost analysis that includes the cost of infrastructure deployment, the human resource needed to maintain the infrastructure, the overheads incurred in any computation, and the savings in the long-term due to a lower number of breakdowns. Comparative studies on the various industrial sectors including manufacturing, energy systems and automated robotics among others would offer a wider validation of economic benefits. Besides, the discussion of the standardized evaluation metrics of quality of interpretability and decision impact may enhance the methodological rigor and benchmark the traditional and explainable systems. The other possible development would be to incorporate risk assessment modules which would rank the maintenance action according to the safety implications and the criticality of the production. The

aspect of cybersecurity can also be made to guarantee the quality of sensor data and the predictive results in the interconnected industrial settings. Technical refinement, economic modeling, usability improvement, and system scalability can also be combined to make the future further improve the role of Explainable Artificial Intelligence in providing transparent, reliable, and economically sustainable predictive maintenance solutions.

Conclusion

The inclusion of Explainable Artificial Intelligence in predictive maintenance systems allows a more orderly method of enhancing the level of operational visibility and economic output in the industrial setting. Predictive maintenance backed up by sensor-based observation and machine learning allow catching the anomalies of equipment early and preventing its failures. Conventional black-box models, however, tend to restrict the user confidence and real-life processes, as there is no clear rationale behind the making of predictions. The suggestive explainable framework does not have this limitation because it implements interpretability systems that determine the most important operational parameters involved in every failure measurement. The system provides systematic explanation and predictive outputs to the maintenance personnel, making it more effective in clarifying decisions. Such transparency lowers doubts, facilitates targeted corrective measures and cutting-edge wasteful maintenance interventions. This approach has shown that it is possible to achieve explainability without influencing the predictive reliability of the system, thus enhancing the feasibility of data-driven maintenance systems in the practical setting. Economically, the paper draws attention to the significance of explainability in enhancing the cost-efficiency and cost-effectiveness. The efficiency of costs is achieved through the optimization of maintenance resources distribution, decrease in the use of spare parts, less diagnostic time, and minimization of unnecessary downtime. The cost-effectiveness is manifested through the operational advantages in the long run driven by the increased confidence of the decision made, well-organized maintenance planning, and better interaction between human knowledge and automated analytics.

Though the additional system complexity might be offered by the implementation of explainable mechanisms, the improvement of maintenance accuracy and operational stability can be discussed as the way to the sustainable financial results in the long-term. The results indicate that transparency is not a technical addition but a strategic element to the full development of the value of predictive maintenance systems. Explainable Artificial Intelligence allows a more reliable, cost-effective, and sustainable maintenance of industries, by maximizing the interpretability of the operational goals.

Acknowledgment

The authors extend their appreciation to King Saud University for funding this research through Researchers Supporting Project Number

References

- [1]. Mathew, D. E., Ebem, D. U., Ikegwu, A. C., Ukeoma, P. E., & Dibiazue, N. F. (2025). Recent emerging techniques in explainable artificial intelligence to enhance the interpretable and understanding of AI models for human. *Neural Processing Letters*, 57(1), 16.
- [2]. Alkhanbouli, R., Matar Abdulla Almadhaani, H., Alhosani, F., & Simsekler, M. C. E. (2025). The role of explainable artificial intelligence in disease prediction: a systematic literature review and future research directions. *BMC medical informatics and decision making*, 25(1), 110.
- [3]. Benhanifia, A., Cheikh, Z. B., Oliveira, P. M., Valente, A., & Lima, J. (2025). Systematic review of predictive maintenance practices in the manufacturing sector. *Intelligent Systems with Applications*, 26, 200501.
- [4]. Mikołajewska, E., Mikołajewski, D., Mikołajczyk, T., & Paczkowski, T. (2025). Generative AI in AI-based digital twins for fault diagnosis for predictive maintenance in Industry 4.0/5.0. *Applied Sciences*, 15(6), 3166.
- [5]. Garcia, J., Rios-Colque, L., Peña, A., & Rojas, L. (2025). Condition monitoring and predictive maintenance in industrial equipment: An nlp-assisted review of signal processing, hybrid models, and implementation challenges. *Applied Sciences*, 15(10), 5465.
- [6]. Lei, Y., Liu, H., Li, N., Cao, J., Qiao, Y., & Wang, H. (2025). Condition monitoring and fault diagnosis of industrial robots: A review. *Science China Technological Sciences*, 68(1), 1110301.
- [7]. Bhatnagar, S. (2025). Cost optimization strategies in FinTech using [10]microservices and serverless architectures. *Computing*, 19(01).
- [8]. Suci, A. M., Amini, R., Asri, A. K., & Martin, N. (2025). Artificial intelligence in renewable energy: A review of predictive maintenance and energy optimization. *Journal of Clean Technology*, 2(1), 29-44.
- [9]. Kruschel, S., Hambauer, N., Weinzierl, S., Zilker, S., Kraus, M., & Zschech, P. (2025). Challenging the performance-interpretability trade-off: An evaluation of interpretable machine learning models. *Business & Information Systems Engineering*, 1-25.
- [10]. Zhao, L., Chen, J., Wen, J., Li, Y., Zhang, Y., Wu, Q., & Yu, G. (2025). Unveiling PFAS hazard in European surface waters using an interpretable machine-learning model. *Environment International*, 199, 109504.