

## Predictive maintenance of Green hydrogen infrastructure: A literature review

*Shanti Swamy<sup>1</sup>, Sashikant Gupta<sup>2</sup>*

<sup>1</sup> Post Doctoral Researcher, Lincoln University, Malaysia; <sup>2</sup> Adjunct Professor, Lincoln

University, Malaysia;

pdf.shantiswamy@lincoln.edu.my

shashigupta@lincoln.edu.my

---

**Abstract:** Green hydrogen, generated through various ways, is considered as a clean energy carrier that can lead to zero carbon emissions in several industrial sectors. The reliability of green hydrogen infrastructure is critical as far as the efficiency is concerned. The very important issue is sustainability involved in hydrogen based system to avoid costly downtime and equipment failure. Maintenance has normally been a crucial and vital process with the machinery being repaired or replaced only when they break down. This kind of approach results in long unexpected wastage of time, and costly affair and decreased operational efficiency. It basically depends on data analysis, sensors, and machine learning algorithms for equipment failures before they occur, enabling the organizations to go for maintenance before hand. Also, the layer-wise relevance propagation is adapted for the better understanding of the model parameters. Pros and cons of predictive maintenance are discussed in this paper through the current literatures.

**Keywords:** Predictive maintenance; natural language processing; algorithms ; Layerwise relevance propagation; Green hydrogen

---

### Introduction

Predictive maintenance has been developed as an important area of interest for Green hydrogen manufacturers in recent years, as it permits for the proactive identification and knowledge of equipment issues before they are of serious nature. Through the implementation of various algorithms such as random forests, gradient boosting, and deep learning, machine learning can be adopted for better prediction of failures.

This research work proposes a predictive maintenance system for green hydrogen system, designed to detect potential equipment malfunctions before they result in significant operational failures. Natural language models and Explainable AI models based on Bidirectional Encoder Representations from Transformers (BERT) and layer-wise relevance

propagation are proposed here for the text classification and proper interpretation and understanding of the model. One block diagram is proposed for justifying the idea of interpretability and trust-worthy model by using layerwise relevance propagation.

## Related work

The paper suggests any damage before it occurs. It also informs about the working of PdM, the significance of sensors, and the tools that is considered as the base of PdM.[1]It highlights the issues of PdM, like usage of quality data, and the importance of skilled workers. The process will encompass its influence to various sectors related to Power, medicine, conveyance, and, the dependability of crucial systems in these area can be guaranteed.

Grey, blue, and green hydrogen are the primary hydrogen production methods which are compared with respect to the hydrogen electrolysis process, and different electrolyzer technologies are included.[2].Hydrogen can be the ingredient in fuel cells to generate power using a chemical reaction instead of combustion, producing only water and heat as byproducts. There are many applications for example in day today life based on green hydrogen, because Hydrogen has low energy density and It can be available with the help of compression and liquefaction, no matter it is costly. Hydrogen in liquid form needs almost one-third of its heating value. Transporting fluid hydrocarbon fuels is more cheaper than hydrogen distribution [3].Because of hydrogen dimensions, it is more prone to leakage, which should be avoided as far as storage and transportation are concerned. The proposed PdM algorithm permits a very interesting decision rule which is taken care for maintenance management, and relevant for big and critical data issues.[4].Multiple classification modules with different prediction horizons are provided to achieve different performance tradeoffs with respect to occurrences of sudden breaks and unexploited lifetime, and forwarding the same to a maintenance decision system to reduce cost to a certain extent.

Knowing the text of log messages in the dashboard, upcoming failure can be understood before hand. The unstructured log messages are cleaned and grouped into structured form which can be extracted significantly[5].The recommendation from data insight can be fed to text preparation for recognizing with maintenance details. The maintenance normally utilize the output from sensors which are temperature or vibration and these can be linked to the degradation process of the system. Further, sensor data are not so easily utilized for the next step, so event logs generated by the machine are used instead. To recognize patterns which are preceded a failure incident, an event-based method with registered events are developed. The paper suggests a remarkable algorithm with text classification on the cyber-physical infrastructure of an Industry. The aim is to enhance the accuracy of prediction with the given industrial figures.[6]The author deliberates to manage the production processes of manufacturing industries efficiently. The algorithms display execution time which is not

affected by the dimension, and seamless manipulation without facing any other computational costs.

The detection of anomalies can be experienced by using the supervised machine learning technique like feed-forward neural network.[7]It is advisable to use a supervised learning algorithm when there are samples with common range.

Isolation Forest also works well in high dimensional problems which have a large number of irrelevant attributes, and in situations where training set does not contain any anomalies.[8].

Methods for machine learning interpretability can be classified according to various criteria.[9]Without machine learning there can be no interpretable machine learning. Therefore we have to guess where machine learning is heading before we can talk about interpretability.

Decisions now will decide which countries and companies lead and profit from the new hydrogen age[10]The author summarizes the development in the energy sector.

Comparison table is introduced here for the proper understanding of different algorithms.The table suggests the methods with the challenges

The following table [1] shows the different techniques with respect to advantages and challenges:-

Methods	Elaboration	Pluses	Challenges
ML	Uses techniques to study data for Pattern recognition and predictive modelling	Accurate predictions, adaptable to various components,understands from previous data.	Works on labeled data for training, complex methods computationally costly
Internet of things	Collection of . real-time temperature ,vibration	remote access to data,easy detection	Sensors are costly Analysis complex Analysis
Big Data	anomalies for vast data	Can work on large patterns,good for real time datasets.	Data security is the main concern
Cloud computing	data analysis provides storage	Good collaboration.	internet connectivity is a must
Edge Computing	Latency is reduced.	transmission costs is reduced,	additional hardware is required.
Cognitive Computing	Decision-making with AI	Adaptive to	Basically for biased

		changes.	models.
Digital Twins	virtual environment of physical structure and simulations thereby	Accurate simulations,	Integration is complex,
Software-Predictive maintenance	Analysis of patterns	Integration with available systems	Implementation is the crux of the method
NLP +EAI(proposed models)	Text classification ,textual information,statistical features of log messages	Maintenance records data is used .Explainable methods to interpret	Works well with any change but needs huge volume of data

Table1.Comparison between differents tools and technologies

### Key Contribution

A natural processing based text classification framework integrated with Layerwise Relevance Propagation (LRP) is applied to identify and interpret critical failure patterns in green hydrogen systems. Transformer-based models, such as BERT (Bidirectional Encoder Representations from Transformers), for contextual text embeddings that excels in NLP tasks for better understanding is to be implemented. Layerwise Relevance Propagation (LRP) is the technique after classifier to note words/sentences needed for model's prediction. By backtracking through the layers and mapping the relevance of each input feature to the model's output is studied. In this research work ,with the textual datasets either in the form of Operator logs ,Fault descriptions, Sensor anomaly reports in natural language, and Historical maintenance notes, the model is pre-trained and understood with EAI models. The block diagram for the implementation is the idea behind this paper. Figure 1 shows the block diagram of proposed method which is more transparent and efficient. The real database output of the registers ,and maintenance logs from the green hydrogen sector for validation can be accessed from the local generation unit.

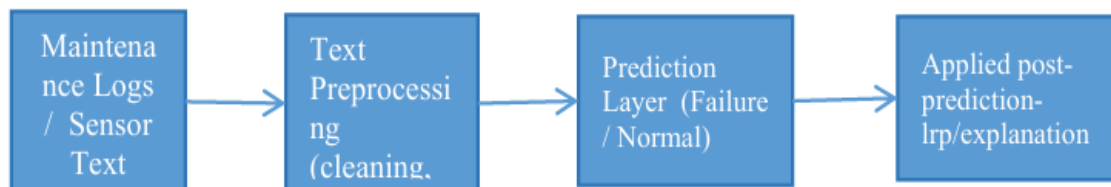


Figure 1. Block Diagram of proposed method

## Method, Experiments and Results:

The predictive maintenance solution presented in this paper is a novel combination of two models in the optimization of green hydrogen scenario. By focussing on the Isolation Forest algorithm for anomaly detection, the system is treated as an effective and scalable approach to improving the efficiency and reliability of hydrogen production machinery. Achievements to date confirm the system's impact, while future developments promise even greater functionality and performance. A dummy dataset with volume, temperature, and flow rate on real-time anomaly detection using the Isolation Forest algorithm is tested which has significantly reduced unexpected machine failures, improving both reliability and efficiency.

## Discussions for further work

- 1)Transformer-based models: The model is fit for handling huge datasets.
- 2)To apply machine learning algorithm, which will achieve more than 95% accuracy in predicting equipment failures
- 3)SHAP tool will assign a numerical value to make a prediction. It provides transparency and trustworthy concepts to understand the box properly. This technique can also be considered with the actual datasets.

## Conclusion

Anomaly detection mechanism is addressed with the Isolation forest algorithm here for the dummy dataset to understand the process. This unsupervised machine learning technique separates anomalies with decision trees, separating odd data points from normal ones.

Random selection is the way to select a feature and a split value from the data to build trees. Anomalies are quickly isolated based on its isolation score. For the better accuracy and score, the proposed method is to be implemented on real datasets.

The proposed method for the increase in efficiency in prediction of failures with NLP and EAI models is studied.

## References

- 1.Manan Patel, Jalpesh Vasa, Bimal patel, "Predictive Maintenance: A Comprehensive Analysis and Future Outlook,"IEEE Explore, 22 February 2024, <https://doi.org/10.1109/INCOFT60753.2023.10425122>

2.Terry Tadlock, P.E,Omar Rubio, PMP , Dragan Ristanovic, P.E"THE GREEN HYDROGEN REVOLUTION – INTEGRATING HYDROGEN INTO INDUSTRIAL APPLICATIONS, *IEEE*, 978-1-6654-9716-9/22/\$31.00 ©2022

3.Mamta Motiramani, Alok kumar singh,"AI-ML techniques for green hydrogen: A comprehensive review,"<https://doi.org/10.1016/j.nxener.2025.100252>

4.Gian Antonio Susto, Andrea Schirru, Simone Pampuri,"Machine Learning for Predictive Maintenance: A Multiple Classifier Approach, **Page(s)**: 812 - 820,18 August 2014, [10.1109/TII.2014.2349359](https://doi.org/10.1109/TII.2014.2349359)

5.Duc-Minh Pham<sup>1</sup>,"Text Mining Framework for Predictive Maintenance in Manufacturing," Nov.,2022[http://dx.doi.org/10.3850/978-981-18-6021-8\\_OR-08-0264.html](http://dx.doi.org/10.3850/978-981-18-6021-8_OR-08-0264.html)

6. Alberto Postiglione,"Predictive Maintenance with Linguistic Text Mining" , *Mathematics* **2024**, 12(7), 1089; <https://doi.org/10.3390/math12071089>

7. chang, K., Hutter, M., & Jin, H. (2009). A New Support Vector Method for Anomaly Detection. *Proceedings of the 2019 IEEE International Conference on Machine Learning*, 8(2), 567-574.

8. Liu, F. T., Ting, K. M., & Zhou, Z.-H. (2008). Isolation Forest. *Proceedings of the 8th IEEE International Conference on Data Mining*, 413–422.

9. Molnar, C. (2022). *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*. Independently published.

10.Dunn, S., Ruyck, M. D., & Kilpatrick, P. (2021). Advances in Green Hydrogen Production and Storage Technologies. *International Journal of Hydrogen Energy*, 46(45), 23456-23471

11.Zhang, Y., Zhang, J., Wang, Y., & Liu, B. (2020). Predictive Maintenance Based on Anomaly Detection Using Hybrid Models. *Journal of Mechanical Engineering Research*, 14(5), 1021-1030.

12.Machairas, V., Boutalis, Y., & Nikolakopoulos, G. (2020). Predictive Maintenance Using Sensor Fusion and Machine Learning for Hydrogen Production Plants. *Renewable Energy Research and Applications*, 16(7), 35- 50.

13. López-Ruiz, A., Rueda-Bayona, J. G., & Arroyo-Castro, J. (2019). Overview of Predictive Maintenance Approaches for Hydrogen Systems. *Hydrogen Energy Journal*, 44(3), 1781-1791.
14. Tang, Z., Zhang, L., & Liu, L. (2020). A Real-Time Predictive Maintenance Strategy for Industrial Equipment Based on Machine Learning. *Procedia Manufacturing*, 52, 357-364.
15. Aziz, M., & Oda, T. (2021). Green Hydrogen Production: Methods and Technologies. *International Journal of Hydrogen Production*, 46(1), 578–591.
16. Shah, S., Ghosh, P., & Bhatt, S. (2022). Deep Learning Techniques for Predictive Maintenance in Renewable Energy Systems. *Journal of Computational Intelligence and Applications*, 11(3), 198-215.