An AI Framework for Predictive Maintenance with a Foundation Rooted in Physics-Based Principals

K. Berrada[1], B. Herrou[2]

[1]*Industrial techniques laboratory, Faculty of sciences and techniques, Sidi Mohamed Ben Abdellah University, Morocco, Kenza.berrada@usmba.ac.ma*
[2]*Higher school of technology, Morocco*

Abstract: In the field of production in general, the physical models that govern degradation are only effective in particular cases, but they are highly explanatory. On the other hand, machine learning models, although working under all conditions with high accuracy, remain unexplained due to their complexity, posing a challenge for engineers in the field. This article presents a symbolic approach to modeling failure modes from data. This approach integrates causal analysis, performance level estimation and risk analysis based on operational safety parameters evaluated according to the criteria of severity, frequency and probability. Using techniques such as deep learning (artificial intelligence), parsimony-driven model selection and symbolic regression, the aim is to minimize a variance function using defined operators. The end result is a simple, accurate function that works under all conditions, while remaining explainable to domain engineers and preserving the predictive capabilities of the neural network.

Keywords: Causal analysis, Performance level, Deep learning, Symbolic regression, Neural network.

1. Introduction

Failures in engineered systems can have serious consequences. Effective maintenance policies are needed to prevent failures and reduce the associated costs. However, limitations in modeling and data availability can make failure prediction difficult and lead to costly and inefficient preventive maintenance. The availability of large quantities of data and advanced technologies has ushered in the era of the digital twin, which can predict asset performance and degradation. Following a risk assessment-oriented decision-making process, recommendations are provided for practical corrective and mitigating actions to be undertaken by maintenance technicians. This methodology underscores predictive maintenance, a strategy extensively investigated in the last decade for its effectiveness as a proactive maintenance approach. The fundamental concept involves anticipating the future trajectory of system health to inform decision-making. Predictive maintenance requires predictive analysis, modeling of failure interactions and optimization of maintenance work. It combines forecasting, knowledge of the asset's functional model and a maintenance policy grounded in risk assessment.

Establishing a predictive maintenance framework involves implementing a decision-making process that connects online asset monitoring with maintenance execution. This process starts with promptly identifying anomalies in real-time operational data.. Next, it involves identifying and modeling failure modes. Finally, it involves estimating the risk of asset failure and optimizing the scope and timing of maintenance.

In this paper, we propose a new symbolic model to model failure modes from maintainability, reliability, availability and safety data of each system and its nth subsystem, using causal analysis, deep learning (artificial intelligence), performance and degradation level estimation by risk assessment through training an artificial neural network on causal data, to generate well-calibrated probabilities. The article is a breakthrough both in terms of:

- Definition of a general framework defining a significant breakthrough in predictive maintenance where each brick can be optimized using AI.
- Modeling failure modes and estimating performance and degradation levels through risk assessment using an AI breakthrough where we create an artificial neural network informed by physics, then distill the model into an interpretable physical equation inspired by symbolic regression with causality data known by the engineer in order to estimate the precise probability of failure risk and therefore plan dedicated maintenance operations in advance.

1.1 Definition of general Framework (The predictive maintenance Framework)

Physics-based machine learning combines data-driven analysis and physics-based modeling to balance the two approaches. It uses real-time sensor data, historical information and maintenance records for anomaly detection and prediction. The data-driven approach is commonly used for rapid event prediction due to time constraints, but long-term predictions are limited by the inability to extrapolate equations and the uncertainty of sensor noise. Monitoring critical failure modes is difficult in harsh environments without nearby sensors, and the complexity of engine assembly requires modeling of interactions and combined effects. To address the limitations inherent in exclusively data-driven methodologies, it is essential to incorporate domain knowledge into the analytical framework of predictive maintenance - When it comes to data availability and quality issues in harsh environments or with limited sensor coverage, domain engineers can use data pre-processing techniques to compensate for gaps or inconsistencies in the available data. In addition, they can explore methods of merging data from different sources to improve data quality and reliability. In harsh environments, they may also consider using specific sensors and technologies that are better suited to these conditions. The aim is to maximize data availability and quality for more reliable and accurate results- .This hybrid approach aims to enhance accuracy by modeling the governing equations of functional behavior and asset health. The integration of domain knowledge actively guides and supports the decision-making process within predictive maintenance. Real-time data is employed for anomaly detection, while the identification of failure modes, risk assessment, and maintenance optimization rely on domain knowledge and experience. The decision-making process begins by identifying the mode of failure, then assesses risks and estimates performance and degradation levels for each system. The distinction between functional anomalies linked to the malfunctioning of engine subsystems and component ageing/degradation phenomena can be established when estimating performance levels. These two analysis paths can be consolidated when component degradation results in system malfunction, or they can operate independently when degradation has no impact on the functional state of the asset and remains undetected until final failure. The decision-making process culminates in the maintenance optimization stage. During this step, the risk trend and estimated time to failure are analyzed to formulate an action plan that effectively mitigates the risk and restores the engine to proper operation. This can be illustrated by the following model (Figure 1):

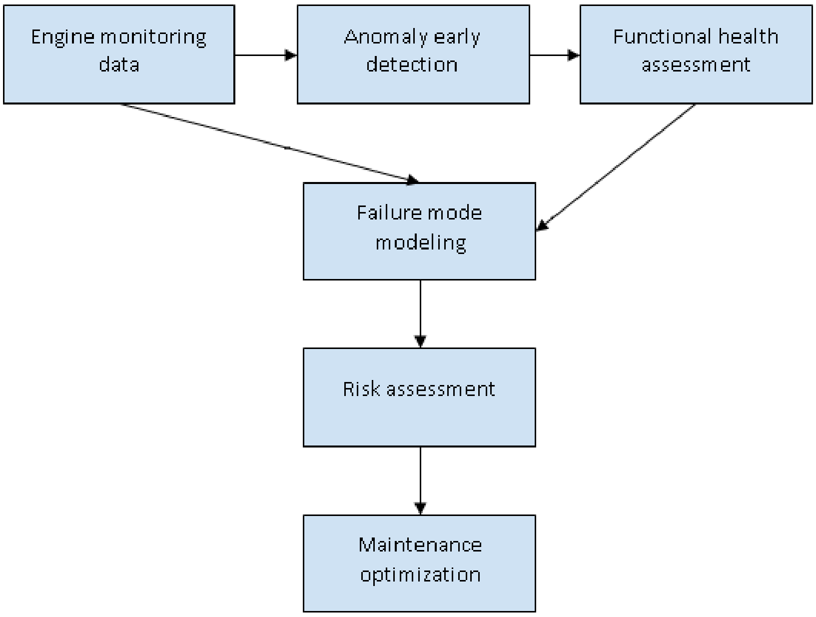

Figure 1: *Schematic diagram of a predictive maintenance analytical framework*

1.2 Anomaly detection with conventional AI

Functional health assessment serves as the foundational layer of predictive maintenance, as the primary goal of maintenance is to reinstate the functional health of assets and their anticipated performance [1]. The pathway for functional assessment involves the identification, classification, and differentiation of anomalies in acquired and computed operational parameters, with a specific emphasis on discerning functional anomalies from sensor failure signatures. While our framework simplifies this aspect for clarity, we will provide an intuition for the reader nonetheless.

Anomaly detection is the technique of identifying rare events or observations that may arouse suspicion by being statistically different from the rest of the observations. Mathematical complexity is the anomaly induced by statistical noise due to sensor failure. Usually, a Kolmogorov-Smirnov statistical test is employed to compare an anomalous signal with signals deemed normal based on specific signal-dependent rules (e.g., during steady-state or transient operation relative to speed) [2].

1.3 Critical failure modes

Critical failure modes are typically characterized by utilizing both engine operational data and historical maintenance data. The choice between a data-driven or knowledge-driven modeling approach depends on factors such as problem complexity, data availability, and understanding of failure modes. In situations with abundant data, a data-driven approach suffices for forecasting and predictive maintenance, particularly when the cause of failure is unknown. However, this approach has limitations, as the results lack easy interpretability, and long-term forecasts may lack precision. Operational data is readily available, but maintenance data is constrained as it is collected at end-of-life or inspection intervals, leading to an imbalance between input data (operational data) and output data (damage measurements collected during maintenance). The model, in this scenario, lacks proper constraints, rendering the solution underdetermined. Another drawback of the data-driven approach is its inability to capture the interaction of various failure modes contributing to the overall degradation observed during inspections. Nevertheless, data-driven risk assessment helps bridge this gap by estimating the performance and degradation levels of each system.

1.4 Risk assessment

On the basis of the failure modes entered and the operational data for each system making up the machine, i.e. the operational safety parameters (reliability, maintainability, availability and safety), we estimate the risk of each parameter for the three criteria of severity, frequency and probability. As a result, each system has an estimate of its level of performance based on a rating grid that encompasses the estimate of the 3 criteria in relation to the four operational safety parameters. High performance levels will be eliminated and only the weakest will be trained in the neural network (threshold to be defined in the algorithm).

These sets of predictors will be used as input to train our artificial neural network, in order to study the interactions between all the data and provide us with a probability in the form of an equation that we need to distill through symbolic regression to obtain a calibrated, interpretable probability. The result obtained must target a high-risk, low-performance or high-degradation system.

1.5 Proposed model

Our model will follow the following methodology, schematized in Figure 2:

a. Analysis of the causes which directly influence the parameters of operational safety (FMDS).

b. Estimation of the performance level of each system based on the causal analysis, and definition of a performance level acceptance threshold.

c. Training of a set of regularized neural network predictors.

d. Simplifying the neural network into a brief mathematical representation through symbolic regression.

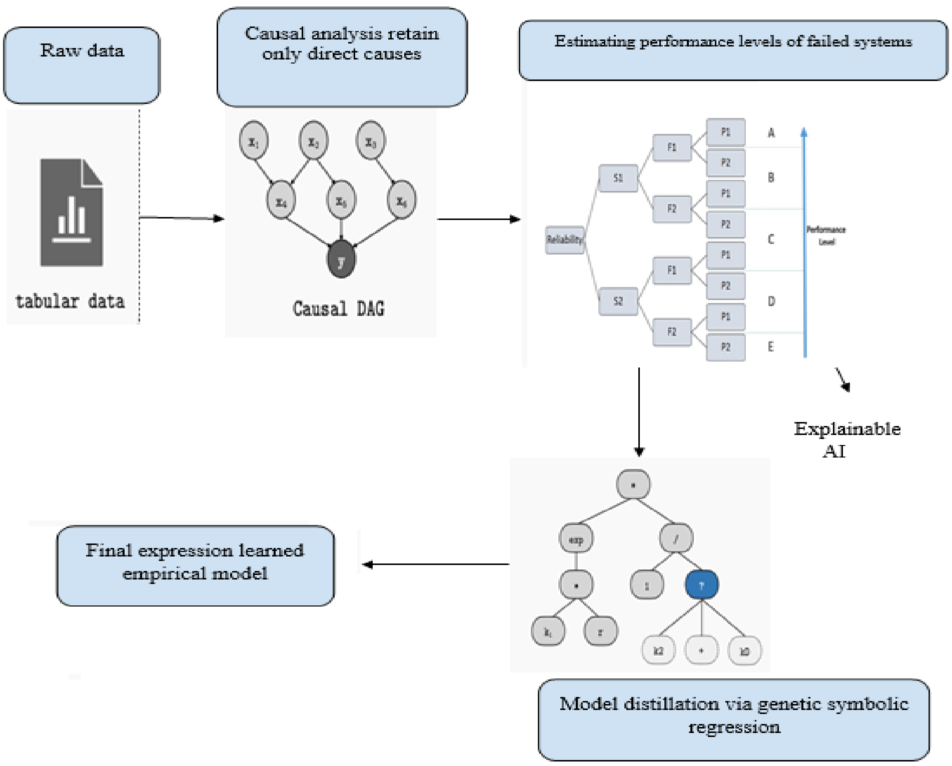

Figure 2: *visualization of the proposed model*

a. Analysis of the causal pathways of degradation :

Creating a causal machine learning model necessitates exposing it solely to parameters with causal relevance. Given that degradations stem from a complex dynamic system, pinpointing their true cause involves a multitude of extrinsic operating conditions that are multidimensional and challenging to capture in all their intricacies. Despite this complexity, many physical effects are influenced by material or fluid state parameters, the characteristic aggregate parameters featured in theoretical models and integral to operational degradation predictions. This study endeavors to establish a model linking state parameters to degradation observations, with the goal of deducing the relative importance of physical effects. The Directed Acyclic Graph (DAG) serves as the most suitable tool for causal analysis. In a causal DAG (figure 3), nodes represent variables, and edges (A → B) signify that A is the cause of B, typically in the probabilistic sense where the probability distribution P(B) depends on A. Our focus is on identifying parameters that encompass the majority of causal pathways leading to targeted degradation. It's important to note that there is no guarantee that a given learning procedure will converge to the true causal model. Consequently, we cannot rely on the causal consistency of a model by design. Instead, we conduct a verification of learned models to strike the optimal balance between causal consistency and predictive performance.

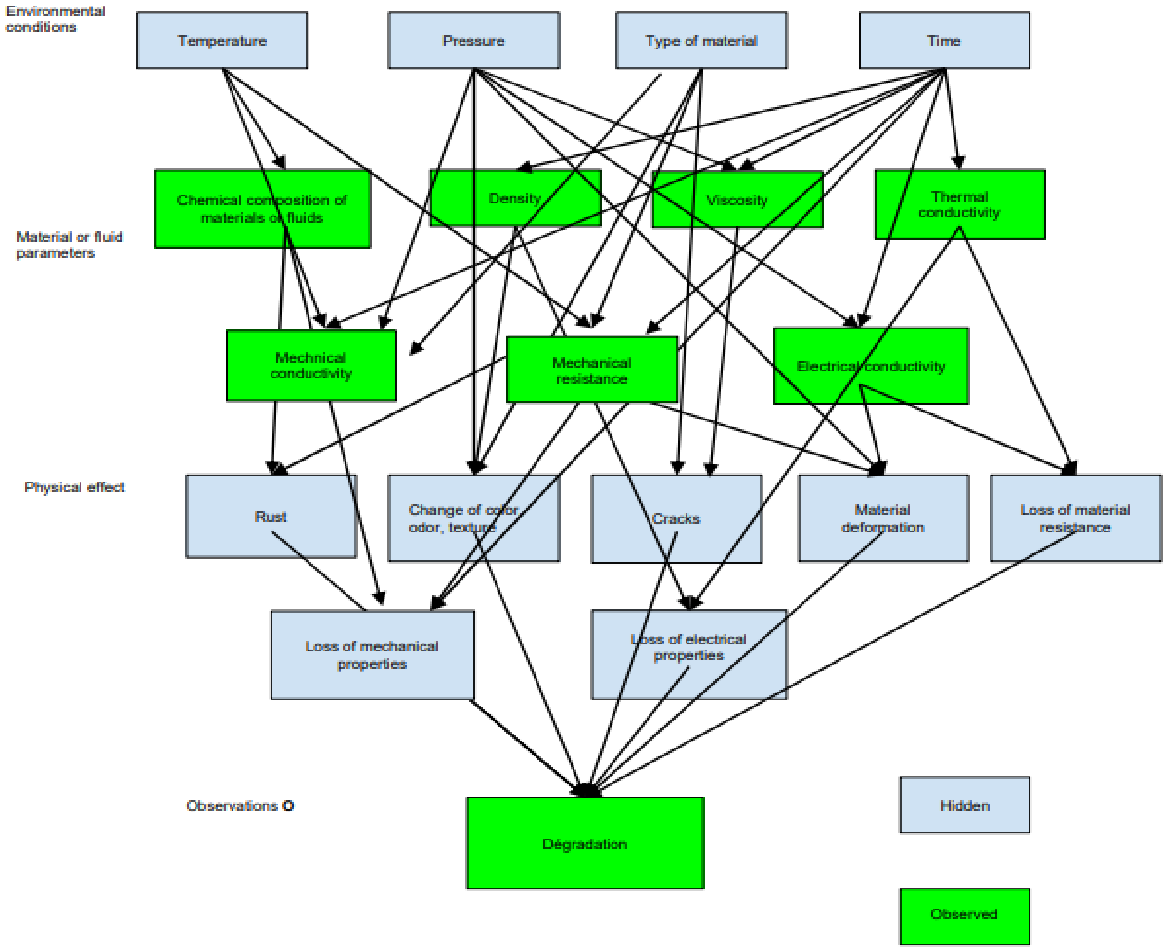

Figure 3: *Cause analysis of degradation in DAG format.*

To validate the predictive maintenance framework and guarantee the accuracy and reliability of the results, we carry out rigorous tests using validation and test data sets to evaluate model performance. This allows us to measure the accuracy of predictions and compare results with actual data. Next, we use cross-validation techniques to assess the robustness of the model by testing it on data different from those used for training. This enables us to detect any problems of over- or under-fitting of the model. In addition, we perform sensitivity analyses to assess the impact of variations in input variables on model predictions. This helps us to understand the factors that most influence results, and to identify potential sources of uncertainty. Finally, we also compare the model's predictions with the knowledge and expertise of engineers in the field to validate the results and guarantee their reliability. These combined methods enable us to validate the predictive maintenance framework and provide accurate, reliable results for making informed maintenance decisions.

To verify the learned models and strike a balance between causal consistency and predictive performance, we use an iterative approach. We start by training the model using historical data and optimizing predictive performance. We then analyze the results to assess causal consistency and adjust the model accordingly. We use techniques such as residual analysis and validation plots to assess the model's causal consistency. This enables us to detect systematic errors or biases in the predictions and correct them. We also seek to understand the relationship between the model's input and output variables, to ensure that it is consistent with the causal logic of the phenomenon under study. If we identify inconsistencies, we adjust model parameters or add constraints to improve causal consistency. In short, we iterate between optimizing predictive performance and assessing causal consistency to find the best balance between the two. This enables us to develop models that are both accurate in their predictions and consistent with the underlying causal relationships.

b. Estimate the performance level of each system based on causal analysis, and define a performance level acceptance threshold :

In order to pinpoint the root cause of the degradation, the idea is to estimate the performance level of the degraded system or sub-system by analyzing the four operational safety parameters (FMDS) according to three criteria: severity, frequency and probability [3]. Each parameter is assigned a performance level rating ranging from A to E (high performance to low performance), corresponding to an analysis of risk in terms of severity, frequency and probability. The overall performance of the system or subsystem is given according to a rating of three levels: high performance, moderate performance or low performance. In our study, we set a moderate performance level for maintenance. In other words, our model will only provide causes for moderate- or low-performance systems. Real-time data analysis is performed through online asset monitoring, utilizing a multi-head neural network. Each input head within the network is assigned a distinct subset of the complete parameter set (x). This allocation aims to restrict non-causal interactions between parameters, allowing for more focused and meaningful results.

c. Training a set of regularized neural network predictors :

The likelihood of detecting degradation based on the material or fluid state can be expressed as the sum of nonlinear functions, each uniquely relying on a subset of state parameters representing distinct causal paths. To capture these arbitrary functions, we employ a neural network with fully connected layers (FCN), allowing for efficient training on extensive datasets. These functions are amalgamated into a single multi-headed FCN with a linear output layer. The neural network outputs a scalar, $p \in (-\infty, \infty)$, representing the logarithm of degradation occurrence concerning the material or fluid state. During training, we utilize the Adam optimizer and back propagation to minimize cross-entropy loss. The resultant equation is intricate and may pose challenges for engineers in terms of interpretation.

d. Simplifying the neural network into a brief mathematical representation through symbolic regression.

A neural network is a machine-learning model inspired by the workings of the human brain. It is composed of several layers of interconnected neurons. Each neuron receives inputs, performs calculations and transmits an output to other neurons. The weights and biases of the connections between neurons are adjusted during learning to enable the network to make decisions and predictions. The neural network learns from the data, adjusting these weights and biases to optimize its performance. Symbolic regression plays an important role in the distillation process of the neural network model. It transforms the complex model into a simple, understandable mathematical equation. This makes it easier to interpret the relationships between the model's input and output variables. In short, symbolic regression helps to make the model more explicit and interpretable, which is essential for understanding its operation and predictions. The accuracy of the simplified model is a crucial aspect when using symbolic regression. Domain engineers are careful to maintain a balance between simplicity and accuracy. This can be achieved by using variable selection techniques, adjusting the symbolic regression parameters and regularly evaluating the performance of the simplified model against the original model. In this way, a concise and interpretable mathematical representation can be obtained, while preserving as much of the accuracy of the original model as possible. At this stage, when distilling the neural network model, uncertainties can be handled in a number of ways. Domain engineers can take confidence intervals into account to assess the reliability of the simplified model's predictions. They can also perform sensitivity analyses to assess the impact of variations in input variables on model predictions. In this way, they can better understand the uncertainties associated with the simplified model and make informed decisions accordingly.

In this phase, we employ PysSR, an open-source tool specifically designed for Symbolic Regression – a machine learning task focused on discovering an interpretable symbolic expression that optimizes a given objective. The primary metric used by PySR to evaluate equation quality is based on parsimony, expressed through the derivative of predictive performance concerning model complexity. When the correct model is identified, increasing complexity typically results in only marginal improvements in performance, often attributable to noise in the data. Our goal in this context is to identify an expression (f) within the space of potential expression graphs, utilizing specific operators to closely replicate the log-probability of degradations predicted by the neural network on dataset x. The use of a restricted set of operators is pivotal to ensure the interpretability of the final expression, and we opt for simple symbols to allow expressions akin to theoretical models such as $P \sim A\exp(B)$. All input features are normalized. PySR generates a set of candidate expressions and presents the user with the best Pareto solutions based on increasing complexity. Ultimately, we manually select the best solution by choosing the expression with the highest parsimony score while encompassing all input features.

2. Conclusion

In the field of production in general, physical models governing degradation only work in special cases, but are highly explainable, or machine learning models which work under all conditions with good precision, but which are not explainable by engineers in the field, given their complexity. The aim is to propose a machine learning model that works under all conditions and that can be explained, through the study of the causality diagram (what I can measure and what governs what is measured), then we estimate the performance level of the systems or subsystems potentially involved following a risk analysis of the operational safety parameters according to the severity criteria, frequency and probability, then we inject these inputs into the neural network, which gives us a very complicated (non-explainable) equation, which we distill using symbolic regression to minimize a variance function with the defined operators, to finally have a simple, precise function that works under all conditions, explainable by engineers in the field and which retains the predictive capabilities of the neural network. The scheme starts with early detection of anomalies based on live operational data, continues with identification and modeling of failure modes, and ends with estimation of asset failure risk and optimization of maintenance scope and timing. Among the potential future directions for research and improvement in the field of predictive maintenance, we recommend the development of predictive maintenance models for specific industries: Each industry has its own maintenance challenges and needs. By focusing on specific domains such as aerospace, energy or automotive, it is possible to develop predictive maintenance models more tailored to the needs of each industry. There are some specific challenges and limitations that need to be addressed in future predictive maintenance work. These include:

1. Data collection and integration; 2. Data confidentiality and security; 3. Model complexity; 4. Interpretability of results; 5. Cost and return on investment;

References

1. F. Calabrese, A. Regattieri, M. Bortolini, M. Gamberi, and F. Pilati, “applied sciences Predictive Maintenance : A Novel Framework for a Data-Driven , Semi-Supervised , and Partially Online Prognostic Health Management Application in Industries,” pp. 1–28, 2021.
2. H. W. Lilliefors, “On the Kolmogorov-Smirnov Test for Normality with Mean and Variance Unknown,” J. Am. Stat. Assoc., no. September 2013, pp. 37–41, 1967.
3. K. Berrada and B. Herrou, “Method for Assessing the Overall Performance Level of a Production System Through A Reverse Probabilistic Study Considering Operational Reliability,” vol. 5, no. 3, pp. 114–118, 2023.