1st INTERNATIONAL CONFERENCE ON
INTELLIGENT COMMUNICATION AND
COMPUTATIONAL RESEARCH (ICICCR-2020)

A Survey: Predictive Maintenance Modeling using Machine Learning Techniques

Vinayak Tyagi, Uday Chourasia, Prinyanka Dixit, Alpana Pandey and Arundhati Arjaria

University Institute Of Technology RGPV, Computer Science Department, Bhopal
Email: vinayaktyagi77@gmail.com

Abstract:
This is the survey paper which includes abstract of multiple references for better understanding of the topic. This paper includes brief overview about predictive maintenance and some of its application were its used as well as older research is also taken into account which discuss various methodology for Predictive Maintenance using Multiple Classifier Approach by Machine Learning. Predictive Maintenance is introduced into various Industries and sectors. It helps to predict the condition of equipment of machine which are already in use this tells whether the maintenance is required or not. This technique ensures that cost saving has done, as compare to regular maintenance where unnecessary replacements have been done without proper utilization of resources. The main advantage of this Technique is that it can allow timely and convenient scheduling of corrective maintenance and can prevent unexpected failures. If in advance we know which that part needs maintenance we can plan and work accordingly. This is continuous and iterative process the model or algorithm continue to evolve as it used.

1. Introduction

Nomenclature

A PdM (Predictive Maintenance).
B Data Mining.
C Machine Learning.
D Classification Algorithms.
E Remaining Useful Life (RUL) or Health Factor.

As data is continuously increasing there is the change in decision making in organization on many important areas.
Machine learning (ML) approaches have been shown to provide more effective solutions in these fields, these approaches does depend on a multiple parameter in system. Simultaneously, the efficient management of maintenance activities is becoming an important factor such as decreasing the costs associated with downtime and defective products especially in highly competitive semiconductors manufacturing industries.[1]

This topic which elaborates the failures that happen on a machinery due to the additive decomposition and stress on equipment's of machine. Even if no straight evidence of process/machine degradation. PdM tools exploit process and multiple logistic variables collected during production to identify the 'footprints' of this degradation in the data.[1]

The goals of predictive maintenance models can be achieved by three main components: the right data availability, framing the problem appropriately and evaluating the predictions correctly.

Ways for management in maintenance are mainly divided into three main categories, these are as follows:

- **Run-to-Failure (R2F)** – When maintenance is done after the failure. This may also be termed as "Post failure Maintenance". This is the easiest approach because one has to deal after the failure has been occurred. (Hence, it is widely used approach) But its least effective and not at all cost and time worthy as compared to planned and corrective measures taken before the failure. [1]

- **Preventive Maintenance Modeling (PvM)** – When maintenance is done according to the planned and scheduled time. In this approach the failure is prevented but unnecessary corrective measures have been taken which lead to ineffective utilization of resources. This factor leads to increase in cost of maintenance.[1]

- **Predictive Maintenance (PdM)** – When maintenance is performed by probabilistic factor of health status. In this approach the maintenance is performed in pre failure but in effective manner. This technique uses older data for better prediction of failure, commonly defined as *RUL (Remaining Useful Life) or Health factor*. This is mainly done statically and uses of analytical tools with engineering approaches.[1]

Decision making in Predictive maintenance is mainly done by evidence/data. Today the world is revolutionizing with big data and analytics and this what opens the new gateway for advance technologies. These are the few things which can be accomplished by the power of analytics today:

• Predictive Maintenance in textile industries increasing its quality and reducing the production time

• This technology when fused with different hardware like sensors and Algorithms may have the power to completely automate the system which may lead to reduce in human intervention directly or indirectly.
• When analytics is fused with sensors this results in some best medical devices helps the healthcare world.
All of these examples use machine learning in great way to solve out-standing problems.

1.1. Literature Review

The PdM (Predictive Maintenance) by ML (Machine Learning) can be divided into two main categories:
a. **supervised** - where knowledge on the happening of failures is there in the dataset;
b. **unsupervised** - where logistic or process knowledge is there, but maintenance related data does not exist. [1]

In PdM problems, regression technique is generally applied when predicting the Remaining Useful Life (RUL) of an equipment, On the other hand classification technique is applied when discrimination between healthy and unhealthy conditions of the system being evaluated.

The collection of maintenance data is mostly depends on the existing maintenance policies in the system: in the scenario of R2F-Run to Failure policies the data related to a maintenance cycles (the production activities between two successive failures) is available and therefore supervised approaches can be used in another scenario, if PvM-Preventive Maintenance policies are currently been installed or used so maintenance cycle may not be observable (maintenance is generally performed well in advance of any potential failure) so, only unsupervised learning approaches is feasible.[1]

2. Methodology

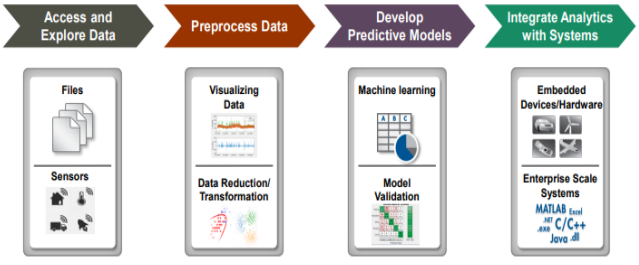

Fig. 1: Working Procedure [4]

All Here is the Working Procedure or an approach towards the implementation.

- **Access and Explore Data** – We need to collect data or download data if needed and then finally Explore it. Usually this type of large data is kept into data centers or databases. This data has been fetched from different sensor installed at different parts of a semiconductor machines and data is stored with respect to time cycles. Exploring data over here means to know the overview of dataset on which an individual had to worked further. This data might be collected on different parameters which are physical like temperature, pressure, humidity and frequency of vibrations the parameters may vary as per the need or output of the machine.

- **Preprocessing Data** – After Accessing and Exploring the data one should have the good idea about the dataset. Now, it's the time for Preprocessing of dataset which basically means the taking out the meaning from the data with the help of its parameters in it and for this the cleaning and visualization is needed to be done on dataset. Cleaning over here means removing the nun value and removing the redundant features from the data like serial numbers, Id numbers. And finally, the visualization has been done of this dataset to have broad picture of what much can be done for example if dimension reduction is needed or not and any sort of transformation has to be done or not this all has been done in preprocessing of data.

- **Developing the Predictive Models** - This the main part of the process where the various classifiers are made using machine leaning and deep learning techniques and these models are further trained into the preprocessed dataset. The data set is divided into two part one for training and other for test or validation purposes. Model validation is also very important part where the trained model has been validated on different types of dataset and its accuracy has been measured.

- **Integrate Analytics with Systems** –Integrate Analytics with systems or deploying the model means that model is available for other business systems. Once, the model has been deployed successfully than others can send the data and get the predictions which may further be used for evolution of the model itself. The model can be further deployed on to various cloud services provider if needed. The trained model be put into the hardware devices or can be deployed on the cloud so that large scale of prediction of new data can be possible at a time.

3. Applications

Predictive Maintenance and Analytics are used is various diverse fields such as automotive, aerospace, healthcare, machinery and much more.

"The value of looking at data analytics is an estimated at 6-8% uplift in production, which is significant. We are talking in the billions." — Amjad Chaudry, Shell Oil [5]

Here is some application which gives the broad idea for what PdM is used for and why it is necessary in involving industries.

- **Mondi Gronau** - Mondi Gronau's plastic production plant delivers about tons of plastic and thin film products annually. Machine failures that result in downtime and wasted raw materials cost Mondi millions of euros each month. To minimize these costs and maximize plant efficiency, Mondi developed a health monitoring and predictive maintenance application. The application uses advanced statistics and machine learning algorithms to identify potential issues with the machines, enabling workers to take corrective action and prevent serious and major problems. This will reduce the costs with predictive maintenance for industrial equipment's and its further process.[5]

- **Safran Snecma** – This is Aircraft engine manufacturing Industry. Which uses Online Engine Health Monitoring System. To monitor engines, Snecma has created an environment for the design and development of health monitoring algorithms. This technology, enables users to integrate algorithmic applications such as input/output, words processing or displays, as well as to continuously and automatically optimize models and algorithms by continuous flow of data from different sensors across the unit or product.
  Some of the application for online engine of health monitoring are: [5]
  - Real time analytics integrated with enterprise service systems helps to monitor and analysis.
  - Predict internal system performance (oil, fuel, lift off, mechanical health, controls etc.)
  - Improve aircraft availability and reduce maintenance costs and time.[5]

- **Respiri** – It's a medical electronic product meant for Asthma patients which collects the data from patients breathing and further used for analyzing. In this technology the defect which is wheeze sound is detected to examine the health of a patients. Respiri is developed for and mainly deployed on mobile app and cloud-based server software. Here are some key points for Embedded-system and cloud-based app for patient diagnosis:
  - Medical device is used to monitor and manage the real time health of asthma patients.
  - Mainly using the analytics in cloud and embedded system.
  - Invokes spectral processing and pattern-detection analytics for defective sound detection on Respiri server in the cloud
  - Gives feedback to patients on their phones.[5]

4. Comparison

Here is comparison of different classifier's which are used for Predictive Maintenance with their brief working. These models have been used in earlier research purpose.

Support Vector Machine (SVM) – SVM is one of the most common approach used for classification techniques. It finds out a line/ hyperplane in multidimensional space that separate out the classes. SVM can handle both regression and classifications tasks and can support multiple continuous and categorical values. For categorical variables a fake value is created mainly in binary format. Although, categorical variables are made of 3 levels, say (A1, B1, C1), is showed by three fake variables:

A1: {1, 0, 0}; B1: {0, 1, 0}; C1: {0, 0, 1}

To have an effective hyperplane, SVMs uses a looped training algorithm, which is further used to reduce an error function.[2]

Kernel Approach:

$$K(\mathbf{X}_i, \mathbf{X}_j) = \begin{cases} \mathbf{X}_i \cdot \mathbf{X}_j & \text{Linear} \\ (\gamma \mathbf{X}_i \cdot \mathbf{X}_j + C)^d & \text{Polynomial} \\ \exp(-\gamma \mid \mathbf{X}_i - \mathbf{X}_j \mid^2) & \text{RBF} \\ \tanh(\gamma \mathbf{X}_i \cdot \mathbf{X}_j + C) & \text{Sigmoid} \end{cases}$$

where $K(\mathbf{X}_i, \mathbf{X}_j) = \phi(\mathbf{X}_i) \cdot \phi(\mathbf{X}_j)$

These kernel function, represented by a dot product of input data points mapped into the higher dimensional features space by transformation. where Gamma is a variable parameter in kernel approach.

The RBF is widely used kernel types in SVMs. Reason for that is their localize and definite responses with in the complete range of the real x-axis.[3]

K-Nearest Neighbors (KNN) –

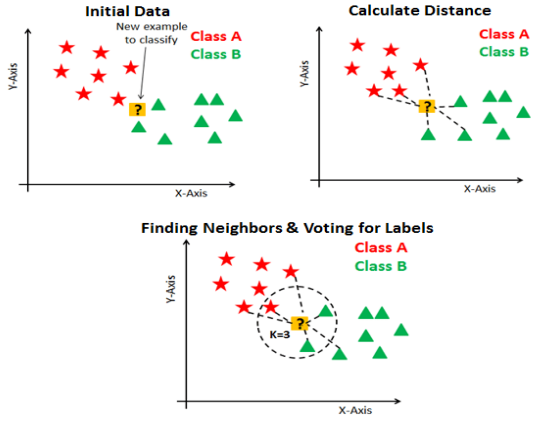

Fig 2: Basic Working diagram of KNN

k-NN (k-Nearest Neighbors) is one of the classification techniques. In this approach distance between the labelled or known data points and the unlabeled or predictive points plays a vital role. In this algorithm the unknown data point closed to the know cluster is considered into its category and this is the simple overview of this algorithm.

The only parameter that is variable in k-NN is k, the value of k is mainly driven by data (mostly elected by cross-validation). As larger as the value of k there will be reduction in noise on the classification, however decision boundaries between classes are less distinct.

	*k*NN			SVM		
m	Accuracy	Precision	Recall	Accuracy	Precision	Recall
1	98.51	69.27	61.69	98.52	69.34	62.96
10	98.47	64.09	67.85	98.47	63.14	70.84
20	98.39	58.00	75.42	98.37	56.61	77.77
29	98.29	53.22	81.72	98.33	54.46	85.79
38	98.15	48.27	87.03	98.16	48.85	88.95
48	97.95	43.31	91.09	97.95	43.36	92.94
57	97.74	39.44	95.30	97.74	39.61	98.15
66	97.52	35.96	97.76	97.57	36.96	99.63
76	97.26	32.74	99.36	97.26	32.86	100.00
85	97.03	30.23	99.94	97.08	30.08	100.00

Fig 3: Older research result of Classification [1]

Acknowledgements

I would like to express my special thanks of gratitude to my teachers and professors, how helped me in doing this research work and I came to know about so many new things. Finally, I would like to thanks my parents and friends for all their support.

REFERENCES

[1] Susto, Gian Antonio, Andrea Schirru, Simone Pampuri, Sean McLoone, and Alessandro Beghi. "Machine Learning for Predictive Maintenance: A Multiple Classifier Approach", IEEE Transactions on Industrial Informatics, 2015.

[2] Aouatif Bencheikh, Abdeljabbar Cherkaoui. "Chapter 25 Transition Model from Articulatory Domain to Acoustic Domain of Phoneme Using SVM for Regression: Towards a Silent Spoken Communication", Springer Science and Business Media LLC, 2019.

[3] Abbas M. Abd, Suhad M. Abd. "Modelling the strength of lightweight foamed concrete using support vector machine (SVM)", Case Studies in Construction Materials, 2017.

[4] Matlab Expo 2017 - Big Data and Machine Learning for Predictive Maintenance.

[5] Predictive Analytics with MATLAB Unlocking the Value in Engineering and Business Data.