Machine learning and AI for Manufacturing

Abhishek Soni
180100004

1. Abstract

In this term paper, a review is done on the advances and development in Industries and how Artificial Intelligence has changed the functioning of traditional industries. Various challenges and suitability are discussed. Then a review is done on the various machine learning algorithms and data analytics technique and on the potential application areas where these techniques can be used in industries. The term paper focuses on one such areas, called Predictive maintenance, the overall pipeline from machines to the end user which comprises of the data processing and getting useful insight from it. Along with the challenges and difficulties that are presently there is discussed.

2. Introduction

The 4th gen Industrial Revolution, so called "Industry 4.0" was triggered by the rapid development of several fields namely electronics, information and advanced manufacturing technology. The IT influence start penetrating to the core of these manufacturing industries which lead to an increasing collection of data. Even today many industries face the challenges to handle Big Data. This collected big data can be further analyzed and used to create innovative applications, enabling the product and service optimization. The current development has evolved the real physical systems into high-level cyber technologies. The manufacturing environment, is characterized as being stochastic, dynamic and chaotic.

2.1 Challenges in Manufacturing Domain

A very common challenge of ML application in manufacturing is the acquisition of relevant data due to limitation in the availability. Even after obtaining the data, further challenges come are the quality, and composition. it can contain a high degree of irrelevant and redundant information. This requires pre-processing in order to use it to train any model in order to get useful predictions. The preprocessing method generally remains same but could vary with the algorithm of choice. Because pre-processing of data has a critical impact on the results. These so-called missing values present a challenge for the application of ML algorithms.

The other major challenge is to decide which ML technique or algorithm to use. Another challenge is the interpretation of the results because the not only the format or illustration of the output is relevant for the interpretation algorithm to choose (selection of ML algorithm). Especially looking at domains most likely to being optimized, e.g. monitoring and control, scheduling and diagnostics, it becomes apparent that the increasing availability of data is adding another challenge: besides the large amounts of available date (e.g. sensor data), the high dimensionality and variety (e.g. due to different sensors or connected processes) of data as well as the NP complete nature of manufacturing optimization problems present a challenge.

To overcome some of today's major challenges of complex manufacturing systems, valid candidates are machine learning techniques. These data-driven approaches are able to find highly complex and non-linear patterns in data of different types and sources and transform raw data to features spaces, so-called models, which are then applied for prediction, detection, classification, regression, or forecasting.

2.2 Suitability of machine learning in Manufacturing industries

Machine Learning has been found very effective in giving strong argument why its application in manufacturing may be beneficial given the struggle of most first-principle models to cope with the adaptability. An advantage of ML algorithms is the ability to handle high dimensional problems and data. There are many ML algorithms which are capable of handlining large amount of along with very high complexity (high dimensions) like support vector machines. However, that are not also very robust to some issues like over-fitting which can give wrong predictions.
These algorithms in many ways provide opportunities to learn from the dynamic system and adapt itself to the changing environment automatically to a certain extent. With these already existing complexity, combinations of different algorithms, goes by 'hybrid approaches,' are becoming more and more common promising better results than 'individual' single algorithm applications.

3. Structuring of machine leaning techniques and algorithms

Most accepted approach to decide a suitable ML algorithm for a certain problem is as follows:

- First one is to analyze the data and decide to chose which algorithm, to choose supervised or unsupervised as per availability of labels
- Second is to look for the availability of algorithm which can handle the dimensionality given the data
- Third is to study previous application on similar problem, specifically on the structure of the training

Various Machine learning models and algorithms can be classified as per their usability as can be seen in Fig. 1.

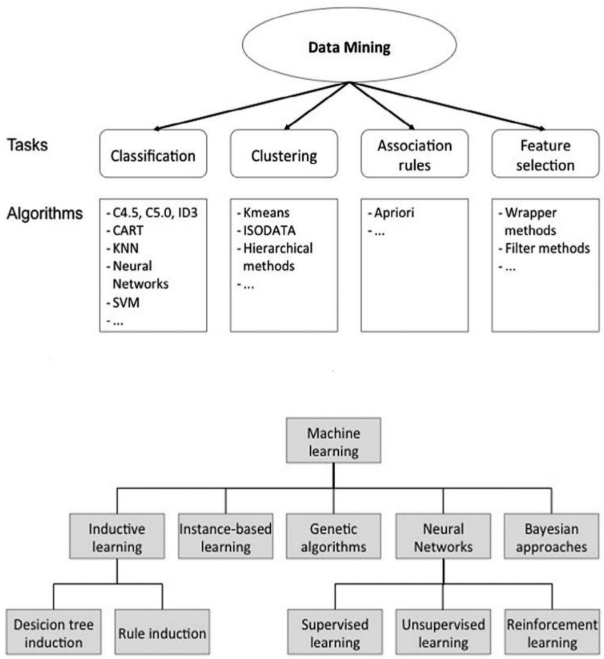

Figure 1 Classification of various ML models and algorithms

3.1 Supervised learning

Supervised ML techniques has its usability due to the data-rich but knowledge-sparse nature of the problems. It benefits from the additional established data provided in manufacturing for statistical process control purposes and as these data are usually labeled. In other words, supervised ML 'is learning from examples provided by a knowledgeable external supervisor'. It generally used for classification.

3.2 Unsupervised Machine Learning

Unsupervised machine learning is another large area of research. In this learning there is no feedback from an external teacher/knowledgeable expert, the output is predicted based on the pattern in the input data.

3.3 Reinforced Learning

It is type of learning in which model is trained to generate the sequence of the decision or we can say the training information based on the environment. The learner has to uncover which actions generate the best results sometimes also termed as 'a special form of supervised learning' which becomes most adequate in situation where there is no knowledgeable supervisor.

4. Application areas of machine learning in Manufacturing Domain

Machine learning has penetrated the core of the industries, and has a large number of application area in manufacturing domain, be it automated assembly, defects detection in the manufactured part, quality management. Maintenance is one of these areas where using ML life expectancy of machine can be improved.

Maintenance of assembly and manufacturing equipment is very crucial to ensure continuous productivity of industry, product quality, on-time delivery, and a safe working environment. Predictive maintenance is an approach that utilizes the condition monitoring data to predict the future machine conditions and makes decisions upon this prediction

Currently the maintenance management in industries can be clustered in three main categories as:

a. Run-to-failure (R2F)— maintenance is only done whenever there a failure occurrence that can hamper the industry production
b. Preventive maintenance (PvM)—where maintenance actions are carried out according to a planned schedule based on time or process iterations. This is better than R2F.
c. PdM — where maintenance is performed based on an estimate of the health status of a piece of equipment. This is advanced management system which requires large amount of data derives from one of Industry 4.0 principles, that has the capability to transforms traditional manufacturing into intelligent sensor-equipped factories where technology is ubiquitous.

This order is in increasing efficiency and complexity.

The 4th industrial revolution, so called Industry 4.0, focuses primarily on creating a advanced digital representation of the physical processes and systems to get better insights on what is going on with the physical processes, what factors are affecting and aiming to create digital factories, i.e., a digital representation of the physical operations that can give insight ranging from technical aspect to business aspects, sometimes called cyber-physical models or digital twins.

5. Predictive Maintenance using Machine Learning

Predictive maintenance is an approach that utilizes the condition monitoring data to predict the future machine conditions and makes decisions upon this prediction. Maintenance approaches in industrial history have evolved over time.

The basic pipeline for the overall from the machine to end user consists of many modules (as in Fig. 2) which are as follows:

5.1 Data collection and Description

The Data Collection module considers different ways to collect data, namely automatic, semi-automated and manual data collection. This module can receive information from multiples sources, i.e. from several machines or assets, or from several departments or services. The collection of the required data is performed considering the IoT technologies.

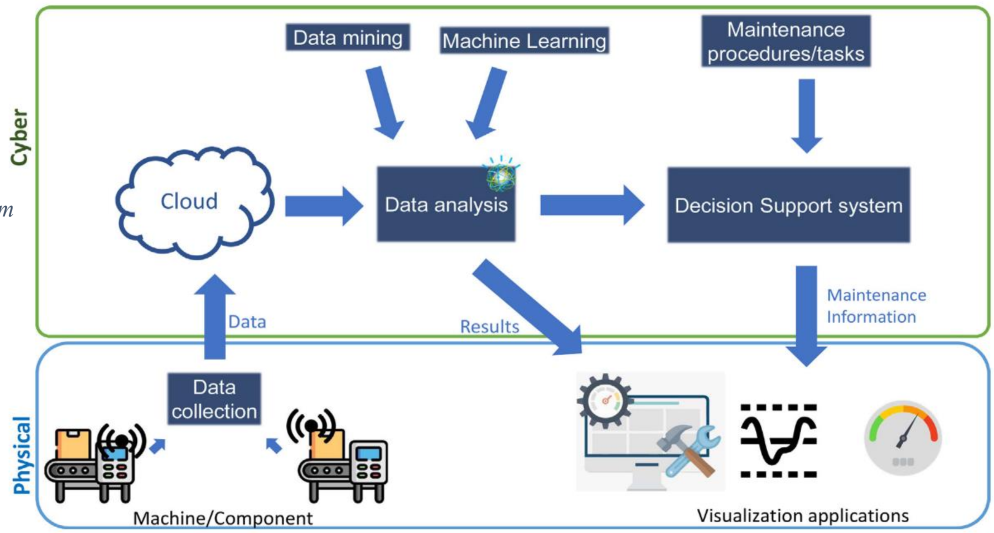

Figure 2 Overall System Design

Internet of things: The connected equipment, regardless of its location in the factory, provides the basis for predictive maintenance. Primarily the sensors send the data whenever machine shows any change in the state which usually is sampled per second or could be any frequency. So, the data received can consists of many Null or missing values. The data collected by these sensors can eventually pass through a preprocessing step before being stored in the Cyber layer of the architecture. So, the very first step towards preprocessing includes replacing these values with any statistical measure like the sliding mean values.

Parameter	Description
Type of defect	Type of defect detected while the visual inspection is performed
Date	Date and time when the defect is detected
Part reference	Identification of the part where the defect is detected
Pressure	Pressure of the hydraulic piston during the operation
Temperature	Temperature in the surroundings of the equipment
Humidity	Humidity in the surroundings of the equipment
Vibration	Vibration of some components of the equipment
Operational noise	Characteristic noise produced during the operation of the equipment

Table 1

5.2 Data Interpretation and Analysis

Interpretation and analysis module use many techniques, namely advanced data analytics, machine learning and cloud technologies to extract out the useful information from the collected data in order to create new monitoring rules and procedures or update the existing ones. A model was developed using supervised learning for the early fault prediction and predictive maintenance was developed. The algorithm was trained from the input data of previous events/faults, labeled accordingly to the type of event.

5.3 Dynamic Monitoring

The Dynamic Monitoring module consists mainly two components, *The Visualization* and *the Early Detection* of Failures. This module receives inputs from the Off-line Data Analysis. which allows a real-time monitoring of specific variables (as in Table 1) which gives useful information about the normal functioning of assets operational limits, and data from the Data Collection module.

5.4 Model development

The output of Data analysis module is to be fed to the choice of Machine learning algorithm. The model can be designed considering many aspects that gives the idea of opportunity cost. A simple model can be developed using the following aspects:

1) Frequency of Unexpected Breaks (ρ_{UB})—percentage of failures not prevented.

2) Amount of Unexploited Lifetime (ρ_{UL})—average number of process iterations that could have been run before failure if the preventative maintenance suggested by the maintenance management module had not been performed the k classifiers work in parallel

Using these aspects given the current costs $c_{UB}(t)$ and $c_{UL}(t)$ at time t operating cost can be defined as:

$$J = \rho_{\mathsf{UB}} c_{\mathsf{UB}} + \rho_{\mathsf{UL}} c_{\mathsf{UL}}.$$

Any maintenance event is triggered by the decision-making logic which is based on the following operating costs minimization philosophy using k different classifiers as:

$$\begin{aligned} j^* &= \arg \min_{j=1,\ldots,k} J^{(j)}(t) \\ &= \arg \min_{j=1,\ldots,k} \rho_{\mathsf{UB}}^{(j)} c_{\mathsf{UB}}(t) + \rho_{\mathsf{UL}}^{(j)} c_{\mathsf{UL}}(t) \end{aligned}$$

5.5 Failure prediction

The prediction can be done in two ways:

i. Classification: This will give the status of the machine in minimal way, whether it is working well or not. The classification problem is not suitable for maintenance managements purposes, because it does not facilitate identification of policies that minimize J which is important in order to improve the life of the machine. This won't give any idea what factors are causing the machine to fail.

$$y_t^{(j)} = \begin{cases} \text{NF}, & \text{if } t \le n_i - m^{(j)} \\ \text{F}, & \text{otherwise} \end{cases}$$

ii. Health estimation: PdM methodology on multiple classifiers basis is used for integral type faults. which describes the failures that happen on a machine due to the accumulative "wear and tear" effects of usage and stress on equipment parts.

The multi classifier gives a freehand to the classification algorithms. There are two well-known and widely used classification techniques, as follows:

a. **Support Vector Machines**: SVMs is probably the most popular approach to classification, it is very efficient in handling high dimension data accurately, even for nonlinear problems, and to the availability of optimized algorithms for their computation

b. **k-Nearest Neighbors**: k-NN is being the simplest classification approach as it requires just computation of distances between samples. In this each point of the input space is labeled according to the labels of its k closest neighboring samples (where distances are computed according to a given metric, often the Euclidean norm).

5.6 Decision Support System

Important piece in this intelligent and predictive maintenance architecture is the decision support system for maintenance technicians during the execution of maintenance interventions. Finally, the predicted output is fed in the decision support system in the Cyber layer, which depending on the results generated by the data analysis, can schedule future maintenance and suggest maintenance routes. It is composed of a database, which can be shared by the database used in the other modules, an engine that processes the maintenance procedures.

5.7 Current Development and challenges

Failure-related data is necessary required to build and test the predictive models. As equipment becomes more reliable, data scientist faces a major challenge of lack of failure related data while developing the model.

A case implementation is shown in the Fig. 3

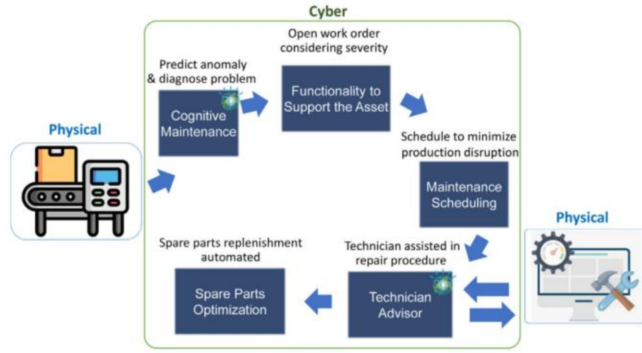

Figure 3 Application Case

6. Conclusions and Discussions

The presented system is robust to variations in the operating costs associated with unexpected breaks and unexploited equipment part lifetime that can occur over time. It also allows the user to dynamically change maintenance policy based on current costs and needs of manufacturing production. Process engineers are provided with a tool that enables them to adjust the performance and the action policy so that the total operating cost is minimized.

References

1. Thorsten Wuest, Daniel Weimer, Christopher Irgens & Klaus-Dieter Thoben (2016) Machine learning in manufacturing: advantages, challenges, and applications, Production & Manufacturing Research, 4:1, 23-45
2. Susto, Gian Antonio; Schirru, Andrea; Pampuri, Simone; McLoone, Seán; Beghi, Alessandro (2015). Machine learning for predictive maintenance: A multiple classifier approach. In IEEE Transactions on Industrial Informatics, 11 (3), pp. 812–820.
3. A. Kanawaday, A. Sane, Machine learning for predictive maintenance of industrial machines using IoT sensor data, International conference on wireless communications, networking and mobile computing (WiCOM), IEEE (2017), pp. 87-90
4. Jovani Dalzochio, Rafael Kunst, Edison Pignaton, Alecio Binotto, Srijnan Sanyal, Jose Favilla, Jorge Barbosa, Machine learning and reasoning for predictive maintenance in Industry 4.0: Current status and challenges, Computers in Industry, Volume 123, 2020, 103298, ISSN 0166-3615
5. A. Cachada et al., "Maintenance 4.0: Intelligent and Predictive Maintenance System Architecture," 2018 IEEE 23rd International Conference on Emerging Technologies and Factory Automation (ETFA), Turin, 2018, pp. 139-146