

# A proposed solution architecture for the application of anomaly detection in predictive maintenance

Yapi Donatien Achou

April 8, 2019

## 1 Introduction

Predictive maintenance for machines and industrial equipments can be defined as a maintenance philosophy or more generally a framework with a set of methods used to predict and prevent machine failure in order to avoid unexpected downtime and reduce related human and financial cost.

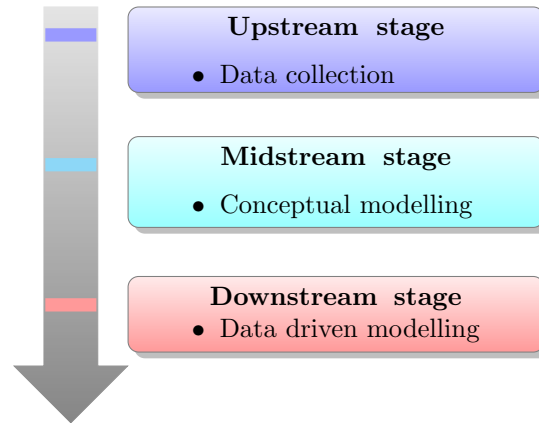
As a framework, predictive maintenance has 4 levels. In level 1, visual inspection of machines or equipments are performed in order to assess any damage. In addition, equipment must fail before they are replaced, which can incur a high production cost. In level 2 and 3, machines characteristics such as vibration, temperature, electrical current, voltage etc are monitored continually or periodically depending on their criticality. This is called condition monitoring. In condition monitoring, the goal is to detect any change in machine normal behaviour, in order to detect failure as early as possible and schedule maintenance accordantly. Maintenance actions are performed periodically regardless of machine health condition. In level 4, big data and machine learning are the main driving forces in detecting failure and planning maintenance. At this level, maintenance action are not performed periodically but are planned according to machine health condition derived from the application of anomaly detection techniques. This reduce unnecessary maintenance action and significantly cuts down maintenance cost as well as increasing machine life time.

According to PriceWaterhouseCoopers (PWC), one of the four largest auditing and consulting company in the world, a survey from 280 companies in Belgium, Germany and the Nederland, revealed that only 11% of companies have reached level 4 [1]. The application of level 4 requires collecting, saving and analysing large amount of data, from which maintenance decision can be made. Anomaly detection methods and more generally supervised and unsupervised learning techniques are used in the analysis phase.

In this report we outline part of the components of the solution architecture for a successful application of level 4.

## 1.1 The solution architecture

Industrial equipments and machines operate under varying conditions, load and requirements. This suggest that the systems are very dynamic in nature and any solution architecture for anomaly detection must be dynamic and adaptive, to say the least. To achieve a dynamic and successful anomaly detection procedure there are three necessary and iterative stages/requirements that must be fulfilled. These stages/requirements are conveniently named as upstream, midstream and downstream stage.



In the upstream stage the data and the metadata are collected. In this stage, often the main focus is on the data, which is collected and stored in a database. Nonetheless more attention should also be placed on the metadata. By metadata we mean all informations about the data as well as any maintenance activity performed on the system. The latter is crucially important for reasons that will be outline shortly.

The midstream stage consists of designing a conceptual model for the whole system. In this stage FMECA or Failure mode, effects and criticality analysis is performed and different data attributes are map to each components of the system. This is done in order to thoroughly understand the system and latter apply appropriate anomaly detection techniques.

lastly in the downstream stage, the anomaly detection techniques are carefully chosen based on the conceptual model designed in the midstream stage. In the next three sections we describe each stage and stress they importance relative to each other and relative to the whole problem of anomaly detection.

## 2 Upstream stage

Most company moving into stage four of the predictive maintenance framework have been collecting data for a period of time. The data is often collected automatically and logged

continuously to a database or data warehouse/data platform. In addition, various information regarding the data are also logged. These information are called metadata. They include detailed description of the data, the name of each time series attribute, the maintenance log, to name a few.

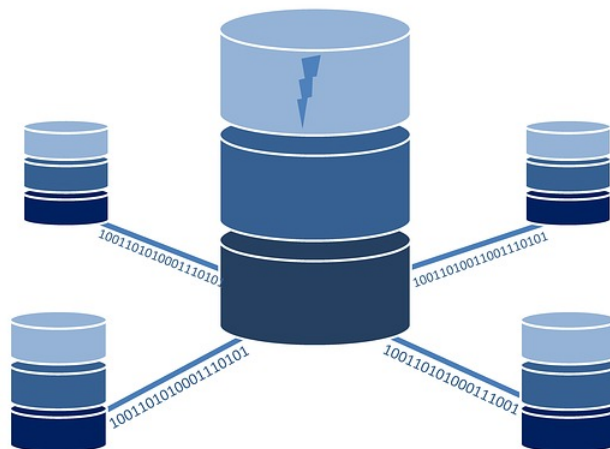


Figure 1: Picture representation of the data collection process

The maintenance log are as important as the time series data. By maintenance log we mean any maintenance activities that have been performed on a machine or system of machines. This includes the exact date, the detailed description of activity performed and the findings if any. The maintenance log will later be used for cross referencing with the time series. By cross referencing we mean the used of the maintenance log to detect a variation in the probability distribution of the data.

Often most companies use excel sheets that are saved locally to log maintenance activities. The downside of this is that the sheets often lack informations as they can be field after the maintenance activities have been performed. In addition, they can not directly be used to auto train an anomaly detection model which has been deployed on a production environment. To mitigate this problem, We suggest that the sheet be replaced by a web application which is accessible on an iPad device like. This way the maintenance activity can be filed in as the maintenance activity are being performed and the maintenance log can be directly push to a database. Let look at a scenario where this can be relevant.

Suppose that a machine learning model has been trained base on the past five years data. This training data probability distribution is generated by a machine. Assume now that few pieces of that machine have been replaced one year later. We have no guaranty that the probability distribution upon which our model has been build is the same. However, with

an automatic maintenance log system such as a web application, an automatic training procedure can be triggered when maintenance activity are logged into the database. Thus we create an automatic triggered base learning or self learning model based on maintenance activities. This will help maintain the integrity of the model.

### 3 Midstream stage

The midstream stage consists of building a conceptual model of the system of machine or machine. The conceptual model is the representation of the system. It allows a better understanding of the system. It is import for a successful development of a data driven model. This stage is a multi disciplinary stage where domain expert and data scientist should work closely to describe and understand the system. Procedure such as Failure Mode Effect and Criticality Analysis or FMECA must be performed to

- Describe the system as a block diagram
- list failure modes and severity of failure
- list cause and effect of each failure mode.

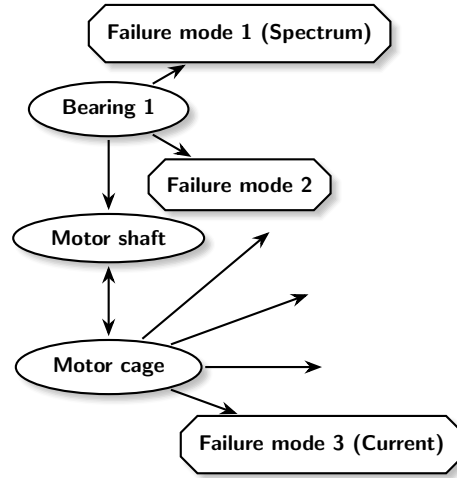


Figure 2: Illustration of an hypothetical conceptual model

Figure 2 shows an hypothetical conceptual mode. In this model the system is reduced to a motor with a shaft, a cage and one bearing. the double arrow between the motor shaft and the motor cage indicates that they is a direct interaction between the two parts. The failure mode of each part is also indicated.

In addition, time series attributes must be map with each part of the system. For example in Figure 2, the envelop spectrum is needed to detect failure mode 1 while current is required to detect failure mode 3. A detail conceptual model can therefore facilitate the data driven modelling where the actual anomaly detection or other statistical/machine learning modelling are performed.

## 4 Downstream stage

After a thorough understanding and analysis of the system, a system of anomaly detection strategies can be derived. This stage is concerned with the machine learning/statistical model training and building. Most of the time we are dealing with time series. Therefore time series analysis as well as supervised and unsupervised methods are applied.

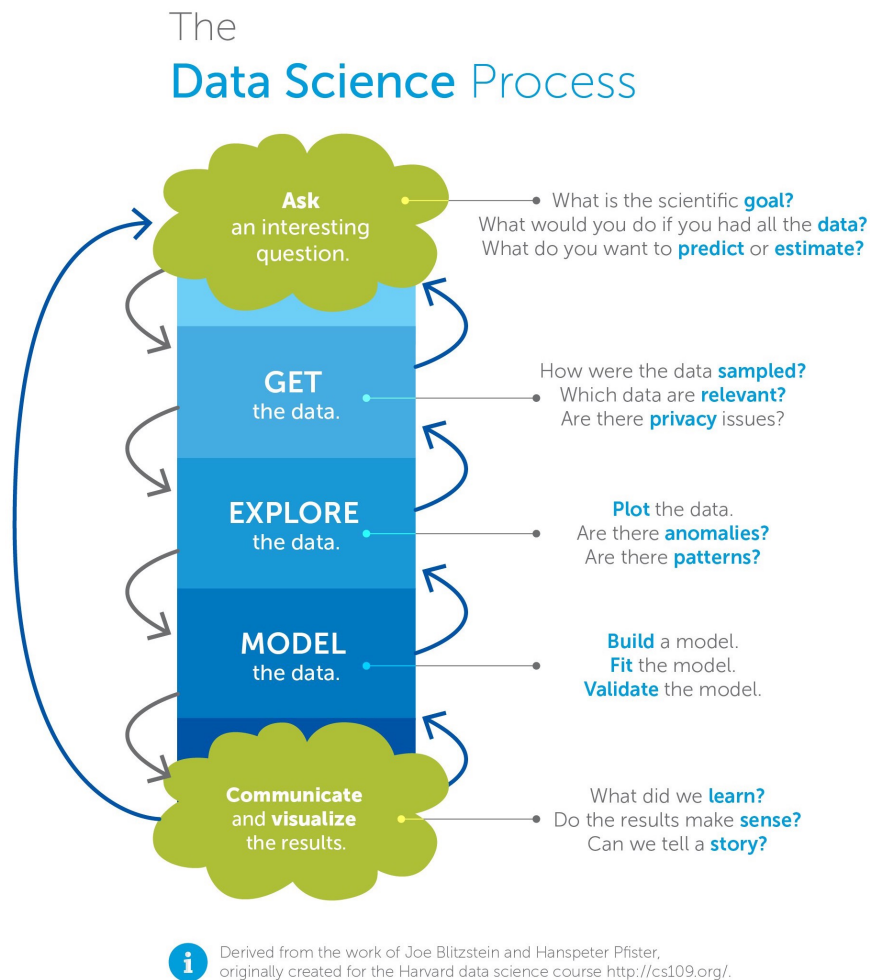


Figure 3: An illustration of the data science process

## 5 Conclusion

## References

- [1] Michel Mulders and Mark Haarman *Predictive maintenance 4.0. Predict the unpredictable*, June 2017