ZDMP Blog

Artificial Intelligence and Machine Learning in Industry 4.0

By Iulian Toma, Software developer, SIMAVI

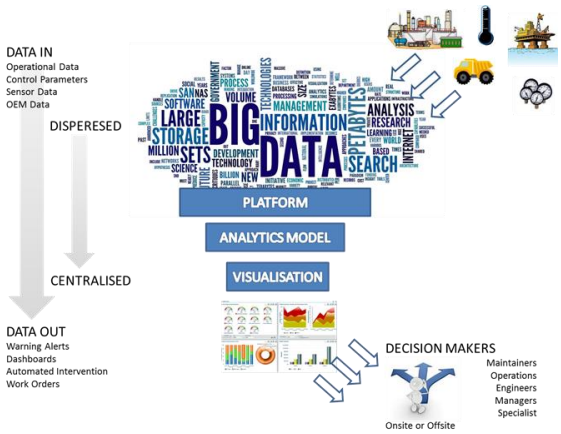
Some questions for you

- How can machine learning optimise quality results?
- What preventive maintenance improvements can be performed with data analytics?
- How to improve data quality for machine learning?

How machine learning can optimise quality results

Artificial Intelligence (AI)/Machine Learning (ML) that learns from situations and condenses the efforts spent towards data quality management saves cost, is scalable to match up with unprecedented data growth, and reduces the processing time by leveraging distributed enhanced computations.

Data quality processes have evolved in their capacity but the demand for pace and efficiency has been proliferating extensively. Data management experts believe that data quality remains a bottleneck due to data volumes and the complexity involved to derive quality insights.



Innovative technologies such as Big Data, AI, ML, have made the application of large-scale advanced data analytics more tangible than before. A transition in data quality process is noticeable from a static rule-based approach to a dynamic, self-adapting, learning-based ML approach in various domains. ML has the potential to assess the quality of data assets, predict missing values, and provide cleansing recommendations, thereby reducing the complexity and efforts spent by data quality experts and scientists. It is always beneficial for an organisation to have a good insight of what data is being procured and consumed, and the business purpose it would solve.

ML assists in deriving a data quality index score to assess data sets' quality and reliability in real-time based on deviation from predicted parameter values. It also has a marked ability to predict trends and identify outliers if trained properly and can make suggestions or take actions on the go.

Businesses look for fast analytics with high-quality insights to deliver real-time benefits based on quick decisions. There is an opportunity for organisations to fine-tune and enhance the current data quality approach using ML techniques. Most leading data quality tools and solution providers have ventured out into ML territory in anticipation of increasing the effectiveness of their solutions. Thus, it has the potential of being a game-changer for businesses in pursuit of improved data quality. ML can complement the contemporary rule-based solutions, and synergistically and gradually lead to process evolution around data quality. The use of ML for data quality assessment and enhancement has promising prospects to churn large data sets and enhance data quality.

What preventive maintenance improvements can be done with data analytics?

There are differing views on the definition of Preventive Maintenance and Predictive Maintenance depending on whether you apply the Reliability Cantered Maintenance definitions or an alternative. Some

people and organisations define Predictive Maintenance as being a particular subset of Preventive Maintenance, with both having the aim of preventing in-service failures of critical equipment.

There are several Predictive Analysis techniques that can be used to determine the variables, combinations of variables, and alarm levels to predict equipment failures:

- **Regression techniques:** These use various forms of statistical analysis, from simple linear regression models to multivariate adaptive regression models in order to predict that equipment failure is about to occur
- **Machine learning techniques:** Put very simply, these emulate the way that humans learn by detecting patterns amongst variables that have, in the past, led to equipment failure, and then using pattern recognition techniques to identify situations similar to those that have occurred in the past

A common issue with all forms of predictive analytics, when it comes to predicting equipment failure, is that there is a need to have failure data to analyse and a shortage of accurate, available data to be able to populate and test these mathematical analytical models which may predict equipment failures is a limitation which needs to be overcome.

Nevertheless, most of the current modelling effort, when it comes to Big Data and predictive analytics in the maintenance space, is currently focused on developing more sophisticated and accurate models to predict the existence of incipient failures.

Increasingly this data is available in real-time, via equipment items fitted with an ever-increasing range of sensors, and which are increasingly connected. The models used to detect the existence of incipient failures are also developing at increasing rates. Machine learning which uses a few inputs can be used to develop more sophisticated models for detecting incipient failures.

Predictive Analytics, combined with real-time condition monitoring may be able to allow estimating the Potential Failure with more accuracy, by collecting information about how measured equipment condition changes over time.

How to improve data quality for machine learning

The ultimate goal of every data scientist is to create a better model with higher predictive accuracy. However, in the pursuit of fine-tuning hyperparameters or improving modelling algorithms, data might actually be the culprit. If the data is of poor quality, regardless of how good a Machine Learning model is, the results will always be sub-optimal at best.

When defects are discovered in products already on the market, it is often too late to remedy and products have to be recalled ensuring the public safety of consumers. In most cases, the defects are results of negligence in quality control of the components or ingredients used in the supply chain.

There is no one-size-fits-all solution for all data quality problems, as the definition outlined above, half of the data quality aspect is highly subjective. However, in the process of data quality assessment, a systemic approach can be used to evaluate and assess data quality. Whilst this approach is largely objective and relatively versatile, some domain knowledge is still required. For example, in the selection of data quality dimensions. Data Accuracy and Completeness might be critical aspects of the data for use case A but for use case B these dimensions might be less important.

ZDMP Blog

ZDMP Links

• **Architecture Component(s)**	Developer Tier
• **Work Package**	WP5
• **Tasks**	T5.6 AI Analytics Designer

References/Acknowledgements

- [1] https://www.assetivity.com.au/articles/maintenance-management/big-data-predictive-analytics-and-maintenance
- [2] https://www.wipro.com/analytics/the-machine-learning-approach-to-data-quality
- [3] https://towardsdatascience.com/how-to-improve-data-preparation-for-machine-learning-dde107b60091