

3 ANOMALY DETECTION TECHNIQUES TO KICKSTART YOUR ANALYTICS JOURNEY IN THE CHEMICAL INDUSTRY

EXECUTIVE SUMMARY

In recent years, there has been increasing pressure within process-enabling industries to innovate, especially in a world of big data. With the existence of the Internet of Things (IoT) and advancements within cloud-computing technologies, it is evident that the implementation of advanced analytics is imperative in order to obtain a competitive advantage. These methods combined with domain knowledge allow transformation of the data into meaningful business insights. This paper specifically focuses on anomaly detection as it has **widespread implications across the chemical space**; specifically, fault detection and diagnosis, process safety hazards and asset monitoring.

Within the Chemical Industry, the implementation of AI (Artificial Intelligence) / ML (Machine Learning) based solutions is currently in its incipient stages; however, we see 3 clear areas where anomaly detection can be used within the large volumes of data to identify quality / performance issues:

1. **Predictive Maintenance Using ML:** We can use AI / ML methods to **help determine the conditions of equipment and predict when maintenance should be performed**. From a strategic perspective, this leads to significant cost savings, higher predictability and standardization throughout the process.
2. **Image Analysis for Asset Monitoring:** We can use image processing techniques to **identify and study different types of faults, generating characteristically different patterns used for fault detection**. These processing techniques leverage deep learning's capabilities to automatically learn fault-related features from images.
3. **Dynamic Risk Assessment for Operating Facilities:** We can build a **dynamic approach to risk-based safety management** that has the capability of **learning from past lessons and dealing with unexpected events and providing the right support**. This employs deep learning methods that utilize the large volume of data collected.



1. PREDICTIVE MAINTENANCE USING MACHINE LEARNING

Within the realm of predictive maintenance, some common problems are predicting remaining useful life (RUL) so that you can schedule maintenance; identify irregular behavior; and diagnose failures to recommend mitigation actions.

Traditionally, predictive maintenance is performed using Supervisory Control and Data Acquisition (SCADA) systems. This is a computer system that gathers and analyzes real-time data and is set up with human-configured rules but does not account for dynamic behavioral patterns of machinery or contextual data relating to the process.

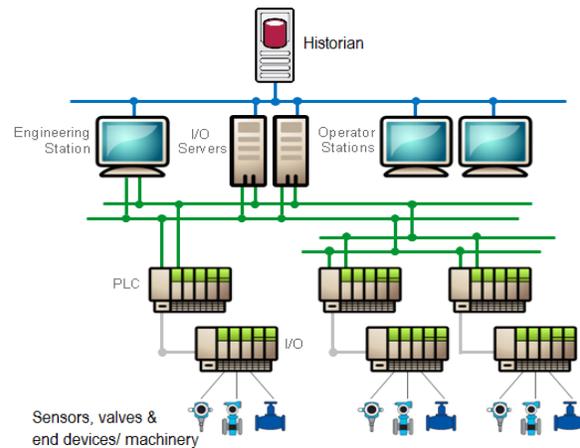


Figure 1: Traditional SCADA System ⁷

Using ML techniques, we can utilize data from different sources – specifically:

- **Operational Technology (OT) Data** – From the production floor, sensors, programmable logic controllers (PLC), SCADA systems, etc. ¹
- **Information Technology (IT) Data** – Enterprise resource planning (ERP) data, customer relationship management (CRM) data, quality data, manufacturing execution system (MES) data, etc. ¹
- **Contextual Data** – Synchronicity between machines, production related information, usage history, operational conditions, machine features, etc. ¹

The 2 most commonly-asked questions in predictive maintenance use different types of ML models, specifically:

1. RUL Prediction employs regression techniques. This requires every event to be labeled; furthermore, every path to failure would correspond to a unique regression model. If there are multiple types of failure and multiple ways that the system can reach failure, then each of these ‘paths to failure’ has its own model ¹.
2. Failure Prediction within a given time frame employs classification techniques. Generally speaking, the ‘given time frame’ in this instance is in the context of the next few days / cycles. Again, this requires every event (data point) to be labeled. Different types of failure correspond to different classes, and so it is likely that you face a multi-class problem. These classification techniques can



also be used to predict the likelihood of failure due to one of the root causes / the most likely root cause of a given failure ¹.

Implementing an ML-based solution to this problem removes a lot of guesswork made by a facilities' staff and allows them to focus on more relevant issues. Moreover, it reduces plant downtimes which keeps clients happy and saves costs.

2. IMAGE ANALYSIS FOR ASSET MONITORING

Recent advancements in image technology have resulted in increasing resolutions on the order of tens of megapixels. This allows one to observe more intricate details in the images. Sequentially, the steps in image processing are to first collect images; then pre-process images e.g. segment them; then use the relevant image processing technique on the sample (discussed further), then detect any faults on the result of the processed image; and lastly extract all features / estimate necessary parameters.

In this example, we are using convolutional neural networks (CNN) as our image processing technique (refer to the diagram on the right). The traditional NN architecture consists of input layers, hidden layers and output layers, and these hidden layers are used to understand the complex structures of data ². CNNs are unique in terms of the NN architecture, in that the convolution operation allows the network to be deeper with fewer features ⁵. This is important when looking for patterns in an image; for example, the first few layers can identify corners / edges, etc., and these patterns are then passed down to deeper layers to recognize more intricate features ³. This property of CNNs makes it successful for pattern recognition (and hence fault detection).

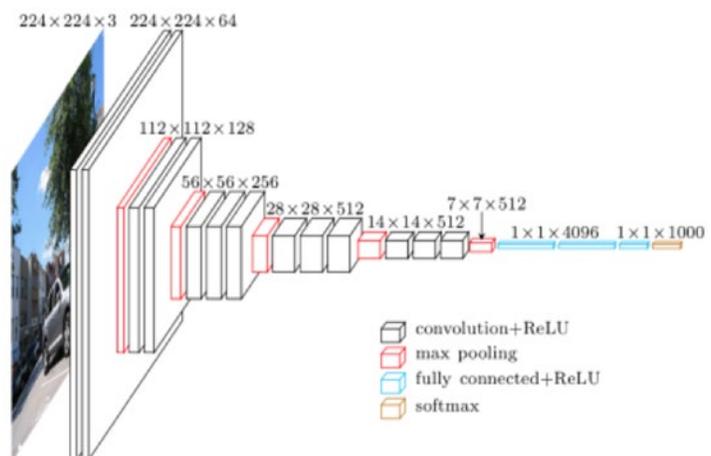


Figure 2: CNN Layering Structure ⁶

For successful image processing, the data needs to be completely / correctly labeled (i.e. we need relatively large and equal numbers of images labeled as defective and non-defective). This would be used to train the CNN, which can be used against new images to detect faults. Lastly, the benefit of using the CNNs for this application is that it can interpret that fault / defect area in the image with heat maps. This allows employees to be able to proactively pinpoint and thus inspect the fault on an asset.



Image processing techniques eliminate the need for Quality Assurance (QA) technicians to inspect assets / equipment manually, which allows for excess labor to be utilized elsewhere. Furthermore, there are limitations to the little faults that the human eye can easily detect; whereas using an automated method will be able to flag these.

3. DYNAMIC RISK ASSESSMENT FOR OPERATING FACILITIES

Traditionally, risk is denoted as a function of what can go wrong (i.e. scenario), its likelihood (i.e. probability), the severity of the scenario (i.e. consequence), and the level of knowledge that we have of the risk (i.e. knowledge) ⁴. Currently, most operating facilities use static risk assessments, which explains risk at fixed reference time points. These studies involve huge numbers of scenarios that cover operating conditions, manning levels, maintenance activities, etc. More importantly, the static nature of these assessments means that these need to be repeated every few years.

The key to the dynamic risk management is the concept of initial conditions – which represents the existing state of a system at risk prior to a hazardous event. Given these initial conditions, the system should undergo a learning process that reflects its risk, its interaction with other systems, etc., to determine the next temporary state of operations ⁴. These responses, when repeated over time will define the dynamic approach to adapting to risk. The goal is to be able to achieve dynamicity, cognition, data processing, emergence and implementation.

We can use deep learning methods to assess risk across these processes; specifically feed-forward neural networks (FFNN). This processes a large quantity of information via different indicators such as physical conditions of a plan, number of failures of equipment components, maintenance logs, number of emergency preparedness exercises, etc. ⁴ These indicators would span from normal operations as well as past ‘risk-oriented’ events which are used to train these models. Once the model has learned to categorize risk, it uses the knowledge of these indicators to determine risk index in real time. Considering the large amount of data available, we can assume that different indicator readings are collected at every time

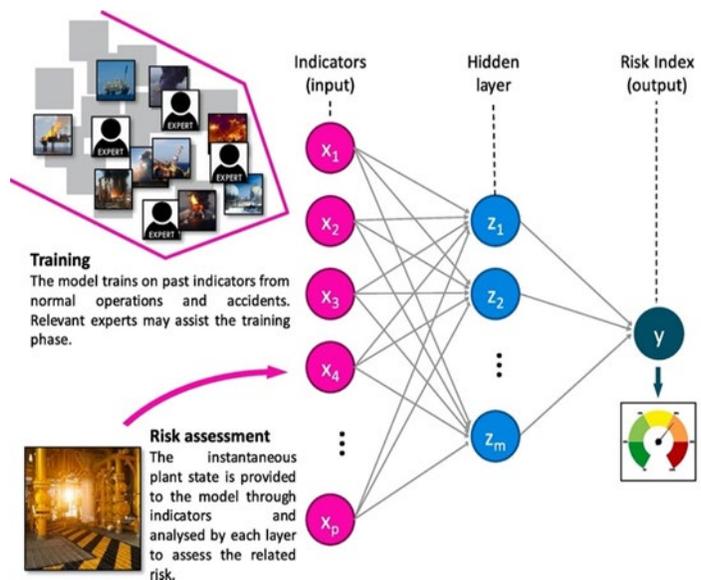


Figure 3: FFNN for Risk Mapping ⁴



3 ANOMALY DETECTION TECHNIQUES TO KICKSTART YOUR ANALYTICS JOURNEY IN THE CHEMICAL INDUSTRY

interval for the last decade. If necessary, these indicators need to be transformed into its derivative with respect to time, which are then used to define inputs into our FFNN.

It is important to recognize the high costs and consequences of any accidents at these facilities; hence the need to ensure high reliability in performance in a risk environment. Any failure has the potential to cost lives, equipment and damage surrounding communities. A dynamic approach to risk mapping allows us to ensure reliable performance in a risk environment, better design training to manage equipment and tasks, as well as have a better understanding of interactions between actors and components of an operating environment in state of risk.

CONCLUSION

These 3 methods described above are a subset of techniques that fall within the larger umbrella of predictive maintenance. Despite the qualitative benefits of implementing these techniques, it is important to understand the quantitative value of this larger effort. In fact, a **McKinsey Global Institute report** entitled *The Internet of Things: Mapping the Value Beyond the Hype* suggested that manufacturer's savings from predictive maintenance methods could total **\$240-630 billion worldwide by 2025**⁸. Considering the massive scale of this initiative; combined with the nascence of using learning-based approaches, it paves as an exciting future for the chemical industry – especially for the first movers and early adopters.



Kaizen Thought Leadership:
**3 ANOMALY DETECTION TECHNIQUES TO
KICKSTART YOUR ANALYTICS JOURNEY IN
THE CHEMICAL INDUSTRY**

KaizenAnalytix.
June 2020
Adi Banerjee

CITATIONS

Gonfalonieri, Alexandre. "How to Implement Machine Learning For Predictive Maintenance." Towards Data Science, 7 Nov. 2019, towardsdatascience.com/how-to-implement-machine-learning-for-predictive-maintenance-4633cdbe4860. [1]

Deka, Partha. "Quality Inspection in Manufacturing Using Deep Learning Based Computer Vision." Towards Data Science, 18 Dec. 2018, towardsdatascience.com/quality-inspection-in-manufacturing-using-deep-learning-based-computer-vision-daa3f8f74f45. [2]

Mohan, Arun, and Sumathi Poobal. "Crack Detection Using Image Processing: A Critical Review and Analysis." ScienceDirect, June 2018, www.sciencedirect.com/science/article/pii/S1110016817300236. [3]

Paltrinieri, Nicola, et al. "Learning about Risk: Machine Learning for Risk Assessment." ScienceDirect, Oct. 2019, www.sciencedirect.com/science/article/pii/S0925753518311184. [4]

Anonymous. "An Intuitive Explanation of Convolutional Neural Networks." The Data Science Blog, 11 Aug. 2016, ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets. [5]

Migdal, Piotr. "Simple Diagrams of Convolved Neural Networks." Medium, 15 Sept. 2018, medium.com/inbrowserai/simple-diagrams-of-convoluted-neural-networks-39c097d2925b. [6]

Kamal El Hawary, AbdelRahman. "Basic Control System: Get Familiar with PLC, DCS, SCADA." LinkedIn, 18 Aug. 2019, www.linkedin.com/pulse/basic-control-system-get-familiar-plc-dcs-scada-abdelrahman-kamal. [7]

---. "The Benefits of Predictive Maintenance." Shell, www.shell.com/business-customers/lubricants-for-business/industry-insights/the-benefits-of-predictive-maintenance.html. Accessed 1 June 2020. [8]

About the Authors: Adi Banerjee is a Data Scientist at [Kaizen Analytix](https://www.kaizenanalytix.com).

