Machine Learning to Revolutionize Industrial Maintenance with Advanced Analytics

A Bentley White Paper

Richard Irwin
Senior Product
Marketing Manager

Published:
September 2017

Bentley®
Advancing Infrastructure

www.bentley.com

While machine learning has been researched for decades, its use in applying artificial intelligence in industrial plants and asset operations and maintenance is now advancing exponentially. This is due to the growth in big data and the expansion of the Internet of Things (IoT); the ability to provide the processing power needed to analyze larger data sets; the availability of machine learning methods; and the need for superior predictive and prescriptive capabilities required to manage today's complex assets. While machine learning has typically been linked with industries such as transportation and banking (think self-driving cars and fraud monitoring, respectively), there are many uses for machine learning within the industrial sector. This white paper will focus on some of the principles within machine learning and industries that are primed to take advantage of the application of machine learning to maximize the benefits it brings to improve situational intelligence, performance, and reliability.

It all comes down to knowing what the best fit is for your needs and what type of data you are using. Data comes in many shapes and sizes...

Before starting, it is important to point out that there are many options and techniques available to gain more insights and make better decisions on the performance of your assets and operation. It all comes down to knowing what the best fit is for your needs and what type of data you are using. Data comes in many shapes and sizes, and can consist of time-series, labeled, random, intermittent, unstructured, and many more. All data holds information, it's just a case of using the right approach to unlock it, and this is where the algorithms used within machine learning help decision makers.

The Route to Deeper Understanding

Machine learning makes complex processes and data easier to understand, and it is ideal for industries that are asset and data rich. A great deal of data from various data sources is required in machine learning, and a data scientist or analyst may be needed to help set up and interpret the results. While it is possible to build your own ML platform, this design takes time, specific skills, and investment in a platform such as Microsoft Azure for a secure, private cloud platform for developers and data scientists. Alternatively, purchasing machine learning capabilities off the shelf, as part of an asset performance management software solution, or outsourcing to a third party are options, provided you ensure input from in-house skills.

Supervised Machine Learning

Supervised learning is the most common technique. There are many stages involved to finish with the desired level of accuracy that can then be applied to new data to start the predictive stage, from uploading the data sets, training the data, choosing the algorithms, visualizing the data, and more. Supervised learning falls into two types: classification and regression.

Classification: Classification is typically used for data that can be categorized; for instance, determining whether an email can be classified as genuine or spam. Common algorithms used within classification techniques include logistic regression, decision tree, and neural networks. Predictive maintenance falls into the classification group because it has several possible outcomes that can be categorized as potential equipment problems caused by various parameters, such as levels of risk, health indexes, reasons for failure, etc. This is the same way machine learning is used to predict a medical diagnosis by identifying the symptoms and telling a person his or her diagnosis. Predictive maintenance has a multi-class classification because there are multiple categories or reasons why a piece of equipment will fail, whereas the email example has only two binary states, genuine or spam.

Regression: Regression is used when data has a range, such as sensor or device driven data, and is used to estimate or predict a response from one or more continuous values. The most common algorithm for regression is linear regression. This is one of the most easily applied algorithms because it is easy to interpret and quick to implement. Measuring and predicting a temperature is an example of linear regression because it has a continuous value where the estimate would be easy to train. More complex algorithms, such as non-linear regression, regression tree, and Gaussian process regression, can be applied for complicated models that are more experimental or contain uncertain responses.

> Due to the growth of data from the Industrial Internet of Things (IIoT) and the availability of sensor data across all industries, data that could often arrive unlabeled and unsupervised learning will become more widely used.

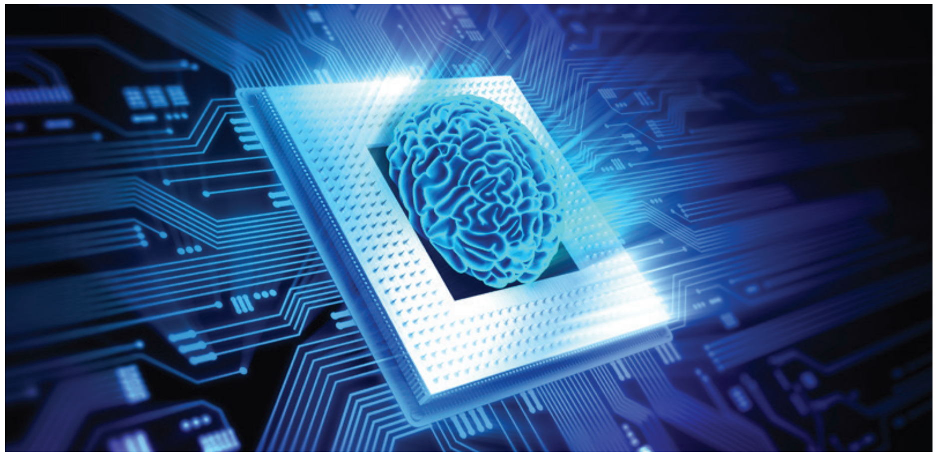
The most common use of regression in the industry is to predict events that have not yet happened, such as demand analysis. The prediction of the number of units a customer or customers will use in the coming days, weeks, or months of any given utility is an example of this predictive regression. Expect to see supervised machine learning to augment or even eventually replace manual reliability-centered maintenance programs to give an "early warning" for asset failures. This approach proves effective if you have a lot of past failure data so that the algorithm can learn what trends occur in a failure and which ones do not. However, past planned maintenance hinders the algorithm's ability to learn; so, a data scientist must be used to interpret the data in the early stages.

Unsupervised Learning

Clustering: While supervised learning usually has an expected outcome to work from and can be trained, unsupervised learning is usually applied when the specific goal is not yet known, or the information of the data is unknown. This means that the data is grouped or clustered together and then meanings are deduced from hidden patterns in the input data by putting the data into similar groups. A basic example of clustering is walking into a room of people you don't know and differentiating the individuals by placing them into different groups based on age, sex, height, etc.

Neural Networks: A neural network, which can be both supervised and unsupervised, is one group of algorithms used for machine learning that models the data using graphs of artificial neurons, or neurons that are a mathematical model that simulate approximately how a neuron in the brain operates. For instance, if a brain was a city, neural networks would be the transportation routes, or networks, within it.

Unsupervised learning can be used to find normal operating modes of your assets and detect trends and anomalies, but requires massive amounts of data and a data scientist to interpret results. Due to the growth of data from the Industrial Internet of Things (IIoT) and the availability of sensor data across all industries, data that could often arrive unlabeled and unsupervised learning will become more widely used. Unsupervised learning can be used as a hybrid learning method as a first step to sort the data into clusters and then be applied to a supervised learning algorithm.



Choosing a machine learning technique to apply depends on the quality of your data, whether the data is labeled, and if there is an expected outcome. Determining the right algorithm for your business purpose will involve a lot of trial and error and will take time, effort, and patience. Once the models have been fine-tuned and the system has been able to learn from past behaviors, only then will predictions be made more accurately. After investing time, the returns from machine learning will be significant and industry changing.

Whatever path is chosen, the benefits that machine learning can bring to big data are only just being brought to fruition. Here are four industries that are at the forefront and are leading the way in this fast-moving digital transformation.

Making the Grid Smarter in the Electric Power Industry

We are all familiar with the term "smart grid" – the electrical supply network that utilizes digital technology and measures to detect and react to usage issues. What areas can machine learning bring to a grid that is already smart? It seems, quite a lot.

In today's turbulent times, electric utility companies are affected by aging assets, increasing energy demand, and higher costs; the ability to recognize equipment failure and avoid unplanned downtime, repair costs, and potential environmental damage is critical to success across all areas of the business.

Predictive maintenance, common throughout all asset intensive industries mentioned here, is again top of the list in any industry, especially utilities, where an enormous number of connected assets are spread across a large network. Consisting of transformers, pylons, cables, turbines, storage units, and more, the potential for equipment failure is high and not without risk, so predicting failures with data and models is one answer to keep the network running smoothly. Another example of how machine learning helps the utilities industry is evidenced through demand forecasting, where predicting usage and consumption from numerous parameters can give a utility an advantage to respond in advance, and balance supply with demand levels. Smart meters can also be leveraged more individually so that customer recommendations regarding efficiency can be made. Machine learning also allows thermal images and video to be analyzed without the human eye to spot differences or anomalies in equipment. Additionally, asset health indexing can be leveraged to automate the analysis of extending asset life with machine learning, which is a low-cost alternative to capital replacement.

Maintaining Safe, Reliable and Predictable Production in the Oil and Gas Industry

In the oil and gas industry, the ability to recognize equipment failure, avoid unplanned downtime, and repair costs and potential environmental damage is critical to success across all areas of the business, from well reservoir identification and drilling strategy, to production and processing. To ensure safe, reliable, and predictable production, identifying equipment failures is one of the main areas where machine learning plays an important role. Predictive maintenance is the failure inspection strategy that uses data and models to predict when an asset or piece of equipment will fail so that maintenance can be planned well ahead of time to minimize disruption. With the combination of machine learning and maintenance applications leveraging IIoT data to deliver more accurate estimates of equipment failure, the range of positive outcomes and reductions in downtime and the associated costs means that it is worth the investment alone.

As well as predicative maintenance, the oil and gas industry has already started using machine learning capabilities in other areas. These include: reservoir modeling, where advanced analytics are used to make improved estimates on the properties of reservoirs based on historical data and models; video analysis that can be employed to detect patterns associated with anomaly detection; and case-based reasoning, which can help by siphoning out numerous parameters that account for well blow outs, leakages, and more, from a large example set of previous cases to come up with solutions to a particular problem. The application of machine learning within the oil and gas industry has the potential to transform the industry, which is even more crucial when production and spending is decreasing.

Save Time and Reduce Costs in Water Utilities

Water utilities also face the challenges of an aging infrastructure, rising costs, tighter regulations, and increasing demand. With that, they can also gain the same benefits offered by machine learning, such as identifying equipment failure before it happens—not just predict a failure, but to determine what type of failure it is. Other benefits of machine learning in the water industry include meeting supply and demand with predictive forecasting and making smart meters “smarter” to help curb waste, such as during water shortages.

Machine learning will help determine the optimal low-cost methods of configuring network transfers, optimizing supply options, enhancing the raw water supply network, and determining the cheapest time and tariffs to transfer water across the network.

Water distribution is an area that can be optimized with the application of artificial intelligence. Machine learning can be used in this scenario to speed up the decision-making process of how demand can be met by analyzing how much water needs to be supplied from the various locations (reservoirs, desalination plants, and rivers), as well as the pumping considerations and water movement, including associated costs and constraints. Machine learning will help determine the optimal low-cost methods of configuring network transfers, optimizing supply options, enhancing the raw water supply network, and determining the cheapest time and tariffs to transfer water across the network.

Flood detection, and distinguishing between what is a flood and what is a blockage, can utilize machine learning by analyzing data from sensors, weather, geospatial location, alarms, and more to give precise predictions and classifications on when and where floods are likely to occur during any given time. These predictions are based on current and historical data from all sources. This would help utilities save time and costs, reduce false alarms, and lessen the impact on the environment.

Smart Manufacturing

Manufacturing generally is one step ahead of most other industries with the application of automation, where robots have replaced repetitive tasks that were once human roles on the production line. Therefore, this stage of early artificial intelligence can already outperform a person through a combination of training and repetition. It can also carry out this task day in, day out, without the need for breaks, sleep, wages, or health insurance.

Manufacturing has always been the main industry when mentioned alongside machine learning, and for good reason, as the benefits are very real.

Manufacturing has always been the main industry when mentioned alongside machine learning, and for good reason, as the benefits are very real. These benefits include reductions in operating costs, improved reliability, and increased productivity — three goals that are the holy trinity of manufacturing. To achieve this, manufacturing also requires a digital platform to capture, store, and analyze data generated by control systems and sensors on equipment connect via the IoT. Preventative maintenance is key in improving uptime and productivity, so greater predictive accuracy of equipment failure is essential with increased demand.

Furthermore, by knowing what is about to fail ahead of time, spare parts and inventory can also use the data to make sure they align with the prediction. Improving production processes through a robust condition monitoring system can give unprecedented insight into overall equipment effectiveness. Other areas of use include quality control optimization to ensure quality is consistent throughout the manufacturing process. For example, adaptive algorithms can be used to inspect and classify defects in products on the production line with pattern recognition to reject defects, from damaged fruit and deformed packaging to car parts.

Early Case Study Examples

The following are two examples that have been applied in different industries, each with different goals and varying machine learning techniques.

Example 1 - Process Manufacturing and Condition Monitoring

The first example is centered around a steel manufacturer who routinely shuts down operations to perform maintenance on its assets, which is very costly. The steel output can sometimes warp or "crimp" during the production process as it travels through different stages. These failures can only be corrected every six months (as well as monthly for smaller fixes) during planned, and very expensive, maintenance that involves long periods of downtime. The main goals of applying machine learning here were to:

- Reduce defects and locate root cause
- Identify key variables that matter the most
- Prioritize assets during shutdown

The first part of the machine learning process was to sort the data into a self-organizing map using neural networks to organize data into 10 distinct classes based on parameters of the steel, such as thickness, weight, and more, as they entered each manufacturing stage. Other techniques included decision trees to learn the pattern of data and to identify which features were important in those patterns; asset health prioritization, to provide ranking and asset health indexing to determine the health of the assets; principle component analysis to reduce the dimensionality of the data; and clustering/anomaly detection, which highlights how each stand deviates from its normal operating mode.

Focusing on assets that are the most at risk optimizes the shutdown, as it is only conducted for a limited time.

What developed was a method for dealing with different types of products, the ability to identify the top variables associated with production defects, and a process for applying anomaly detection to equipment in an industrial plant. It was shown that these processes could reduce the need for extensive analysis of equipment and give operators better tools and more insights to make maintenance decisions. A significant amount of time is spent locating the cause of the issues and performing maintenance. The new algorithm can be run before planning the shutdown, and it can identify which stand to prioritize during shutdowns through analysis of the asset anomaly charts. Focusing on assets that are the most at risk optimizes the shutdown, as it is only conducted for a limited time.

Example 2 - Gas and Chemical Management and Predictive Operational Correction

Within this example, the focus was on preventing failures in cold turbines due to "thawing." The goal was to use machine learning patterns of failure that previously remained undetected until it was too late, which were very costly due to repairs, manpower, and downtime that averaged at nearly 60 lost hours per failure. The data scientist developed a repeatable and reusable method for building predictive models with a combination of R and Azure to predict the failure. With a combination of historical sensor data (over 200) and past failure data, this information was used to train a supervised machine learning model. After data cleansing and some experimentation, algorithms were applied to create the training model that best fit and were suited to the inputs and outputs. Once training was complete after the effectiveness of the model was evaluated, it was applied to the original data set.

By applying an early warning to the cold turbines, so begins the transition from preventative to predictive maintenance. As seen here, this was achievable with decent accuracy, precision, and recall metrics. The early warning help operators prevent a failure by making adjustments and alerting the maintenance team.

Going Digital with Machine Learning

These are just some of the industries that will be early adopters of machine learning. Some are already reaping the benefits in the speed of information delivery, costs, and usefulness. As the technology advances, each industry will learn from each other, further advancing the use and influence of artificial intelligence. Having a machine learning strategy in place will give you unprecedented insight into your operation and will lead to benefits in efficiency, safety, optimization, and decision making. A whole range of problems that once took months to address are now being resolved in a matter of minutes, because of machine learning.