Lifetime prediction of sealing component using machine learning algorithms

Department of Production Management, LTH
Olof Jansson, 2016-09-15

Tetra Pak is a world leading company in the food and packaging industry, with their head quarters situated in Lund, Sweden. They were founded in 1951 and has since provided industrial solutions for packaging food and beverages, such as milk or juice. Traditionally, Tetra Pak has relied heavily on a patent for the packaging material used in almost all packages. With this patent, Tetra Pak's business model has been to sell the packaging machines to its customers at a very low margin, which allows setting up a factory without a huge investment. Once the factory is up and running, Tetra Pak can sell their patented packaging material at a very good margin, and thus have a stable and profitable business running for a long time. Some years ago the patent expired, which brought competitors in the form of large-scale, low cost, Asians manufacturers. These companies could compete very well selling only the packaging material, since they did not have to bear the costs of R&D and such. This has forced Tetra Pak to focus more on providing complete solutions, covering not only the packaging but the whole factory, and even further up and down the value chain. Lately Tetra Pak has put an emphasis on adding services as part of its "complete package"-deal. Existing services are often maintenance related, but there are also some services consultancy services where Tetra Pak helps to benchmark its customers against the customer's competitors.

One step towards expanding the service offering is the introduction of Data-Driven services, i.e. services that in some way are based on the large amount of data that Tetra Pak has collected and stored. This seems like a reasonable step considering that data and data-driven services are very hot topics in many industrial fields. Condition monitoring and predictive maintenance are examples of such data driven services. The idea is not completely new at Tetra Pak; there has been different initiatives and projects in the area since approximately 2005, but it is not until recently that the company is getting closer to delivering such services at a large scale. Lately there has been condition monitoring projects examining knives and servo motors in the machines, and the latest addition is condition monitoring of inductors, which was the object of interest in this master's thesis.

An inductor is an electrical component that Tetra Pak uses sealing packages. It utilizes a technique called induction heating, in which a strong current is run through a copper coil. This will create a magnetic field around the coil, and the magnetic field will induce a current in any material close to it – in this case packaging material. Since the packaging material has a small amount of metal in it, the current will flow in the metal and heat it up, which in turn will melt the plastic layer in the material, and seal the package in a glue-like manner. A problem with examining and replacing sealing components such as inductors is that it is not directly evident that a part is broken. The part might be doing sealings that are 100%, and then slowly drift to 99,8%. The defect is so small that one can't see it at the machine, instead the problem might surface weeks later during transports or in a grocery store. The wear of inductors is mainly caused by corrosion, since the product being packaged (for example fruit juice) is acidic. Different product acidity means different wear, and thus very different lifespans. These two problems make it very interesting to monitor the condition of an inductor.

In this project, machine learning algorithms were used when trying to predict remaining useful life (RUL) of an inductor. The predictive model was built using the analytics tool Microsoft Azure, which is a modern, cloud-based solution. A school-book example for explaing what machine learning is, is to predict housing prices using some variable about the house. In the simplest case, one could try to predict the house price using only the size of the house, measured in square meters, this is displayed Figure 1. There are 100 measurements showing price as a function of house size, and by training the model on more and more data, it will become a better representation of the actual data-set. In the last part of Figure 1 is is evident that the model (representetd by the red line) is a pretty good representation of the data. The example can then be expanded to include more variables, such as year built, number of bathrooms etc. The problem quickly grows and becomes more complex, but the same principles still applies. The same technique can be used in the case of inductors by training the model on Remaining Useful Life (RUL) as a function of some electrical measurements.

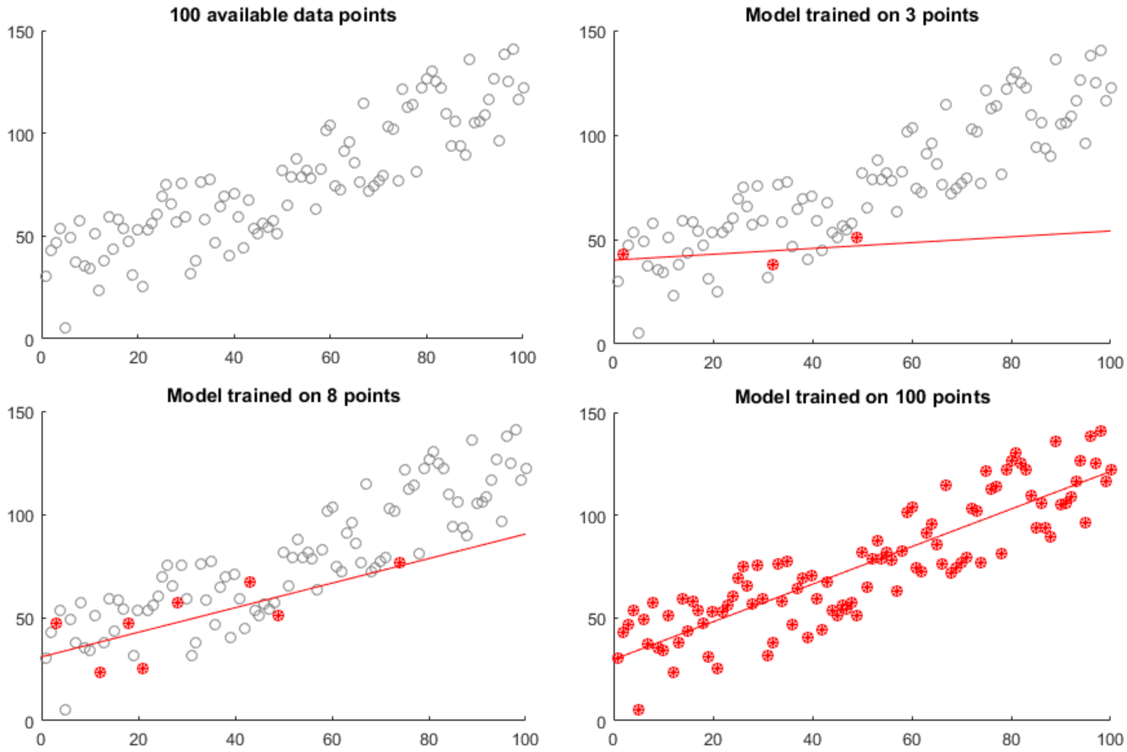

Figure 1: Example of training a predictive model with more and more data.

In the project, historical data from three different machines were used to train a predictive model. There were approximately 8 months of data, 2 inductors per machine, and 13 different measurements recorded every day. After discussing the physical properties of inductors with industry experts, as well as looking at the data from different angles, it was decided to include 3 of the 13 measurements in the predictive model; phase, frequency and impedance. At this point the goal was to find some useful trend that could be used for training a predictive model. Many different approaches were tried, for example calulating mean values each day, cleaning the data for noise etc. The data behaves as pulses, where each pulse represents one sealing, and a breakthrough was reached by looking at all pulses from one day (around 500) and try to find some trend. Figure 2 shows that the daily phase average at a certain time into the pulse will show a significant trend over a few months, and looking closer we can split up the data into cycles, or “run-to-failure”-cycles. This means that an inductor is run from being new and fresh until it’s completely worn out. In total, 6 cycles as the ones shown in Figure 2 were extracted from the data and used to train a predictive model. The same procedure was used with the other variables; impedance and frequency, and apart from daily averages, daily standard deviations were also calculated. This means that each of the 6 cycles contained phase average, phase std. dev., impedance average, impedance std. dev., frequency average and frequency std. dev.

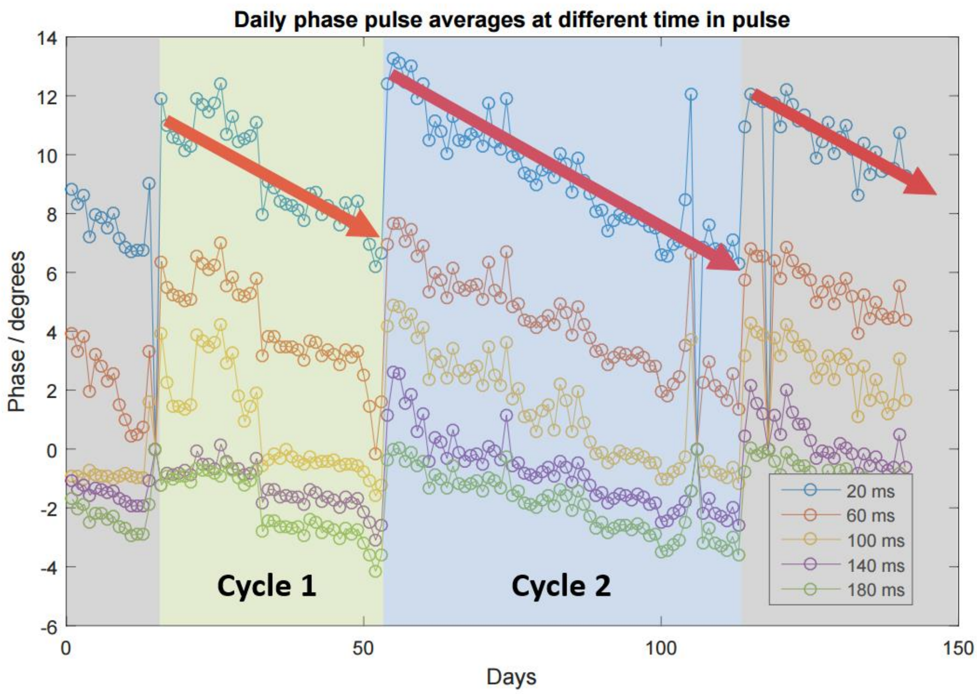

Figure 2: Trends of daily averages of phase pulses at different times into the pulse.

The analytics tool Microsoft Azure was used to build a complete predictive model using 4 different machine learning algorithms and comparing them to each other. As mentioned, there were 6 cycles extracted, and the model used 5 cycles for training and 1 for testing. This procedure was completed 6 times, alternating which cycle that was used for testning. The results of the 6 different runs are displayed in Figure 3. A model with any kind of predictive capability will score between 0 and 1, where 1 is optimal. As seen in Figure 3, no algorithm could consistently produce a high enough score. There is one case where Neural Network Regression scores around 0.95, but since it is only this one time it can only be considered as luck. These results indicate that it is not possible to make meaningful predictions using the current amount of data. However, since there are clear trends in the data as shown in Figure 2, it could very well work if many more cycles are collected.

Figure 3: Scoring the predictive algorithms using coefficient of determination.