**ANOMALY DETECTION**

**with**

**ARTIFICIAL NEURAL NETWORK MODELLING**

**EC 48W.01**

**GROUP ESPOR**

Alperen Yıldız 2015402174

Aşkın Caner Akkuş 2014300300

Feyzi Can Bağbozan 2013300177

Muhammed Ali Metin 2013300312

Barış Koç

**ABSTRACT**

*In this research, we tried to determine times when the machines perform in anomalistic situations. After our analysis, models and Machine Learning algorithms that we use predict anomaly times of the machines. This project aims to determine those anomaly times and constitute preventive maintenance processes. We used two different algorithms of Machine Learning which are Unsupervised Artificial Neural Network and Clustering to find the best fitting model.*

**I) INTRODUCTION**

Today, we have very competent production systems, thanks to rapidly developing technology. Modern cyber-physical systems make us enable to control and monitor the production processes better with their software and physical components. They have computer-based algorithms. They are enable to work with many parameters and enable to monitor many sensors. So in the supply side, producers are more capable now. On the other hand, the demand for more various and more complex products is increasing. And a lot of new producers are entering the sectors. In such a competitive market, producers need to invest on their production technologies for can competing with their competitors and meeting the demands of the consumers. Therefore the producers have much more complex production systems today. But that complexity of the systems comes with another problems. Monitoring that production systems which work with many factors is quite difficult for human operators. Monitoring the production process, maintenance of the efficiency of the production components are very important in such a competitive environment. A production component has a lifespan. To optimize the production process, to predict the lifespans of the components is necessary. If operators can detect and locate anomalous behaviors as earlier as possible, they may maintain better the production order. If it is late, it is clear that to reach again the necessary efficiency condition would be more costly.

In this project we used some data analyzing techniques and machine learning for using data that comes from production components in monitoring the production process. We tried to observe the significant changes in a data from a production component that has many sensors and we tried to figure that changes in the data in an easier way for human operators.

**II) METHODS REVIEW AND BACKGROUND**

**A) Methods**

In this study, we have implemented two different methods of Machine Learning.

1. Unsupervised Artificial Neural Network
2. Clustering

Because of the misfit of Supervised Artificial Neural Network, we implemented Unsupervised one. Our data and modelling is more suitable for Unsupervised one because we tried to find thing we don’t have in variables, data.

**B) Background**

As can be seen from the introduction part, “modern cyber-physical systems make us enable to control and monitor the production processes better with their software and physical components. They have computer-based algorithms. They are enable to work with many parameters and enable to monitor many sensors.” Therefore, by using those facilities, we tried to determine predictive maintenance times of the machines.

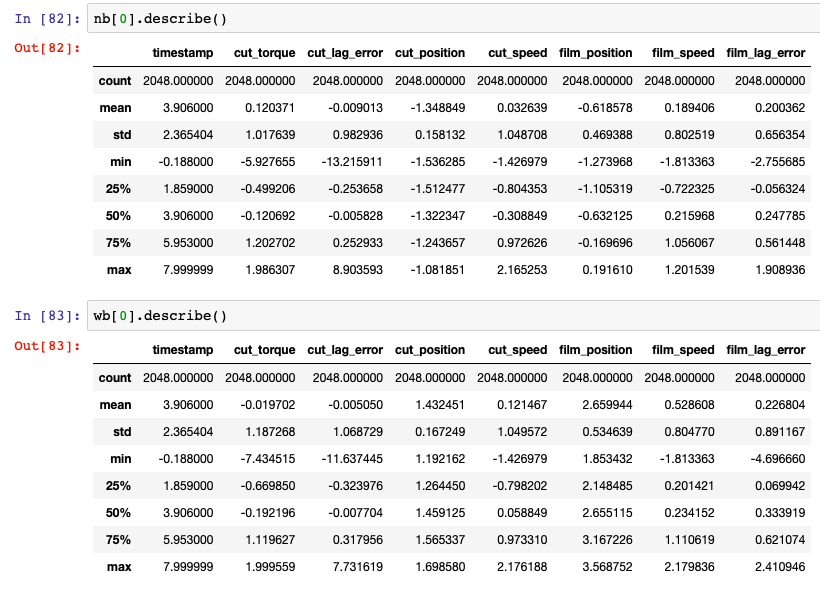
Automatization and digitalization have been spreading through all production systems. Therefore, both maintenance systems and quality-control processes should be supervised over data sets taken from those machines. Moreover, all those processes produce complicated and huge data sets, this makes use of Machine Learning algorithms necessary.

**III) DATA CLEANING AND RECONSTRUCTION**

By using the StandardScaler library, data has been normalized by using mean and standard deviation of the normal position of the machine.

**IV) STATISTICAL ANALYSIS AND DATA EXPLORATION**

**A) DESCRIPTION OF DATA**

****

Data consists of six data sets containing 1 second signal snapshots recorded at specific intervals. Each data set consists of 2048 points with the sampling rate set at 2 kHz.

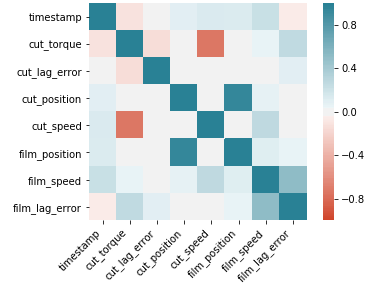
3 of the data sets are recorded in the normal state of the machine where the other 3 of the data sets are recorded after the blade is worn. The data sets are named as: nb[0], nb[1], nb[2], wb[0], wb[1], wb[2] where “nb” indicates “new blade”; and as the same way “wb”, “worn blade”. The data sets are recorded in the same order they are named with long time intervals in between. The blade wears more as the data sets are recorded.

Figure above shows statistical values of relevant variables. (Ex:timestamp has mean 3.9060.)

Variables in the columns: timestamp, cut\_torque, cut\_lag\_error, cut\_position, cut\_speed, film\_position, film\_speed, film\_lag\_error

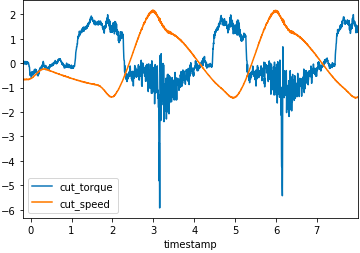
**B) CORRELATION BETWEEN VARIABLES**

As can be seen from the first below figure, there is high negative correlation between cut\_torque and cut\_speed and also high positive correlation between cut\_position and film\_position

****.

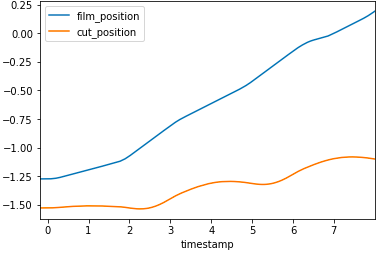
**Correlation Figures:**

**1.Correlation Figure**

****

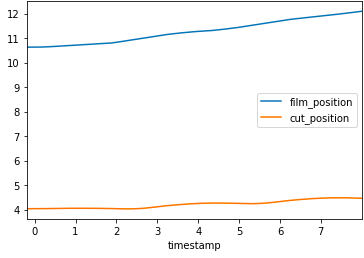
The negative correlation between the cut torque and cut speed is shown in the figure.

**2.Correlation Figure**

****

The correlation between film position and cut position can be seen in the correlation graph where the condition of the machine is normal.

**3.Correlation Figure**

****

Third correlation figure indicates the anomaly situation plot, meaning the machine is not in normal condition. The correlation seems to fade out compared to the 2. correlation figure.

**C) SELF-ORGANIZING MAPS**

After the correlation analysis, we searched for new techniques. And we decided to use Self-Organizing Maps by using machine learning. In this part we try to answer the question that “What is a self-organizing map” and then explain that how we got meaningful graphs from our data by using Self-Organizing Maps.

In the previous parts we told about the difficulties of monitoring the well-developed cyber-physical production systems of today. Production components, machines, motors have some predicted lifespans but actually too many environmental factors (like temperature and workloads) may affect the working of these components. Operators may face many problems such as unexpected faults, degradation of production quality, or decrease in efficiency in a production process due to these unpredicted changes in working of the machines. All of these, and repair or change the components may be very costly.

Producers need to have a dynamic detection system to avoid the possible costs due to such problems in a working production system. At this point engineers build some models that make the operators enable to observe anomalous behaviors of production components. Those models usually contain some predicted behaviors of a given system and compare them with the real, working system. These models detect anomalous behaviors when there is a significant deviation between in working of the predicted system and the real system. Those models may be built manually by experts’ observations but it would be costly. Working with well knowledgeable experts would be expensive and it also costs time, when that work is done manually for modern, complex cyber-physical production systems. Learned models from data would be a better option here. Data-driven monitoring approach is new and popular for production systems. When we work with learned models from data we can use self-organizing maps to detect and locate anomalous behavior for predictive maintenance of a system.

Self-organizing map is a type of artificial neural network. It is also called as Kohonen map or network, by the name of the Finnish professor Teuvo Kohonen who has introduced the approach. A self-organizing map projects a data comes from several sensors, on a two dimensional grid. Each neuron has a weight vector. And the distance between two vectors on the grid may tell us something significant.

We worked with random batch training approach to learn a self-organizing map from data. We firstly gave random variables to the weight vectors at the beginning. And we acquired a diverse starting point for the training stage. Self-organizing maps give us a best matching unit (BMU) by finding the neurons that have the shortest distance to the sample from the all input samples.

The self-organizing maps detect anomalous behaviors by calculating the “quantization error”. The quantization error is the calculated distance of a sample to the best matching unit’s weight vector.

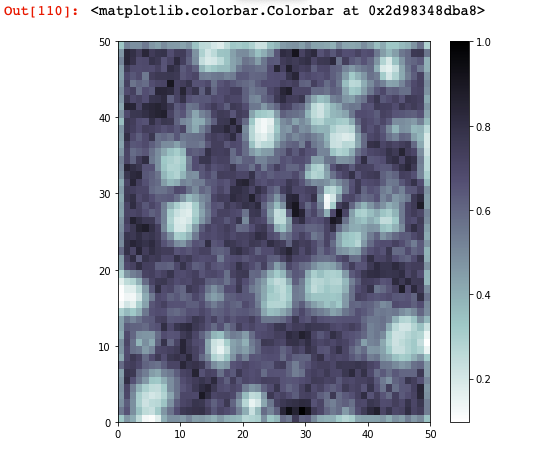
Features of Self-Organizing Map:

- Dimensionality Reduction: Self-organizing maps are used to compress data that has many dimensions. Self-organizing map represents data in many dimensions as two-dimensional output.

- Competitive Learning: Self-organizing map activates only one neuron for a input vector by choosing best one. The neuron that is activated by Self-organizing map is called as “best matching unit”.

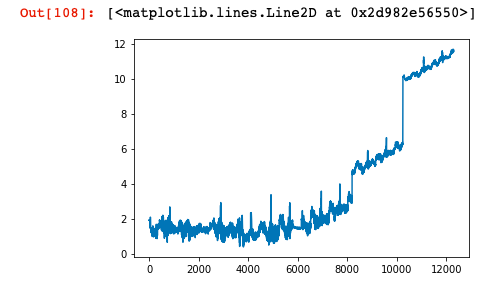
- Topological Preservation: In self-organizing map the data points which near each other in a multi-dimensional data are preserved also nearby each other in the output in the two-dimensional figure. Self-organizing map clumps the vectors that have similar features.

We used first part of our data that comes from beginning of the component’s working process, and trained our full data. Trained data gave us the map below.



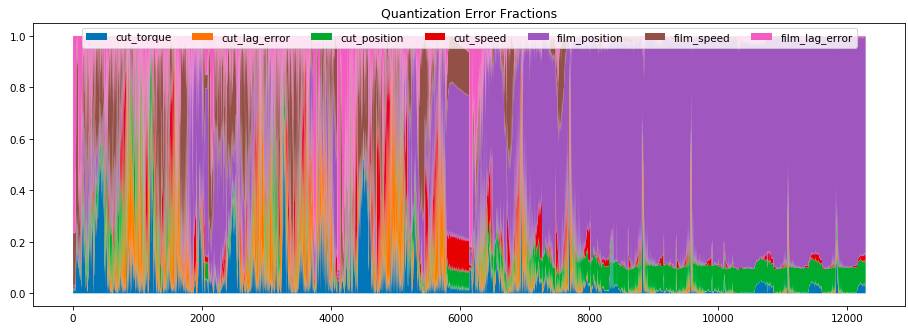
This self organizing map represents the density of the neurons, the darker the colour is the more dense the neurons are. The neurons also represent the normal situation of the machine. The quantization error is calculated by the distance between input and the best matching unit. The inputs from a normal behaving machine are expected to locate on the dense areas on the map. When the inputs begin to locate away from the dense areas, this indicates the machine is shifting away from its normal condition.

The total quantization error calculated representing the health of the system.



Since the data sets are 2048 signals long, the first 6144 timestamps represent the normal behaviour of the machine. The timestamps between 6144 and 8192 represent the data set Worn Blade 1. Since the data sets are recorded with some time interval in between, the error rate jumps at timestamps 8192 and 10240, the transitions to next data sets. Since the anomaly is detected in the system, it comes to the localization of the anomaly. The anomaly is simply located by calculating the error fractions of components. The error fractions of the components are expected to vary with time and show no dominancy.

The error fractions are calculated:



As it can be seen on the quantization error fractions, after the timestamp near 6000, there seem to be increase of the error fraction of film position. The increase in the film position levels off after some point and stay dominant, indicating that the anomaly is caused by film position. The anomaly is detected just at the beginning of degradation.

**V) METHODS AND TOOLS OF USE**

Using Unsupervised Artificial Neural Network, this study aims to detect anomalies when they begin to occur. It is called “Unsupervised” because the thing we want to find is not depended or contained in variables in columns.

The normal condition is learned by the model to represent normal condition of the machine. With the time the recorded signals shifts away from its normal condition, it is expected that the error increases.

**VI) RESULTS**

Since it is important to avoid unexpected failures in industrial systems, it is critical to detect the anomaly or degradation at the very first beginning. The model implied provides experts to detect the anomaly and operate even before the anomaly can be detected by human observation. This prevents both system breakdowns and high costs due to new components.

An unsupervised artificial neural network is trained using the new machine, the map created represents the normal condition. Using this map, the degradation of the system is calculated and the anomaly is located.

**VII) RESOURCES AND REFERENCES**

We took our data from Kaggle. By using the StandardScaler library, data has been normalized by using mean and standard deviation of new position of the machine:

[*https://www.kaggle.com/inIT-OWL/vega-shrinkwrapper-runtofailure-data*](https://www.kaggle.com/inIT-OWL/vega-shrinkwrapper-runtofailure-data)

We used Self Organizing Maps to analyze the data:

# *Self-Organizing Maps for Anomaly Localization and Predictive Maintenance in Cyber-Physical Production Systems Von Birgelen, Alexander / Buratti, Davide / Mager, Jens / Niggemann, Oliver*

[*https://www.sciencedirect.com/science/article/pii/S221282711830307X*](https://www.sciencedirect.com/science/article/pii/S221282711830307X)