Assessment for the Data Scientist position

Assessment description

Ball bearings are a crucial component in any wind turbine. The condition of the ball bearings is monitored to ensure no unexpected downtime of the turbine.

The ball bearings consists of an outer ring, balls, cage and the inner ring. The ball bearing can be damaged in several ways, where the most common is a dent in either the inner or the outer ring.

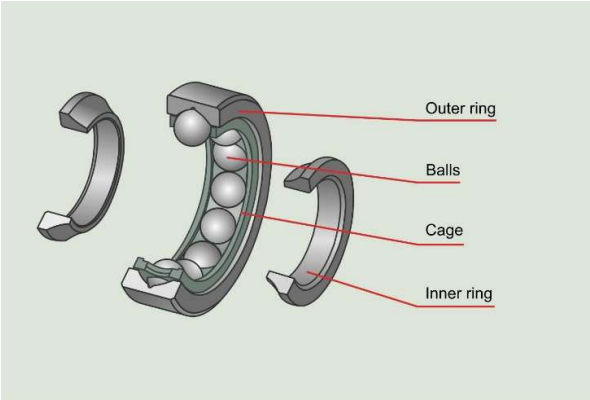

Figure 1 - Ball bearing

Such a dent will cause distinct failure frequencies to appear as a function of the rotation speed of the shaft inside the ball bearing. The "Ball Pass Frequency Outer"(BPFO) is the frequencies which the balls passes over a single dent in the outer ring, this is typically specified as a multiple of rotation speed by the manufacturer.
Every time the ball passes over a dent, it will cause a spike in vibration captured by the data acquisition equipment. This will cause harmonics of the fault frequency(BPFO) to appear in the vibration data as seen in Figure 2. Sometimes these harmonics will also appear at much higher frequencies than seen here, such as $> 25 * BPFO$ and often the low harmonics are not observed.

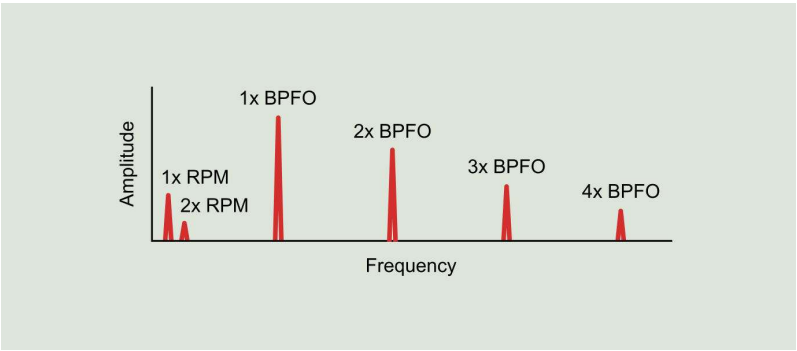

Figure 2 - BPFO fault frequencies

In this assignment data from the Case Western Reserve University Ball bearings dataset is used. Here the rotation speed of the shaft and BPFO is:

$f_{rotation} = 28.68\ Hz$

$BPFO = 3.5848$

$f_{BPFO_0} = f_{rotation} * BPFO = \ 102.81\ Hz$

The data is sampled at $f_s = 12kH$

Provided material

In the Python file "case_western.py" it is shown how to import two HDF5 files containing vibration data from both a good and a faulty bearing. The data from the good bearing is found in "x_baseline.h5" where 40 samples of 1 second each are found. The data from the faulty bearing is found in "x_fault.h5" where 10 samples of 1 second each are found.

Your assignment

1. Can you, by plotting the data in Python, show the difference between a good bearing and a faulty one? Maybe try to identify the BPFO.
2. Now it is time to get creative, how would you use machine learning techniques to distinguish the good bearing from the faulty one?

What is next?

At the next interview we will discuss your findings and methods. You are not expected to do a presentation, but please bring a few plots to show on your pc/tablet.

If in doubt

If you have any questions regarding the assessment, please do not hesitate to contact Asmus at abh@gramjuhl.com.