

Mixed signal processing and machine learning methods for bearings fault detection: Application in predictive meaintenance (Project outline)

Yapi Donatien Achou

February 12, 2019

1 Introduction

Predictive maintenance can be defined as a maintenance philosophy with a set of methods used to predict and prevent machine failure in order to avoid unexpected downtime. This maintenance philosophy when correctly implemented, increases machine life time, and reduces maintenance cost [referecne].

In rotating machines, more than 40 % of machine malfunction can be attributed to bearing defect [references].

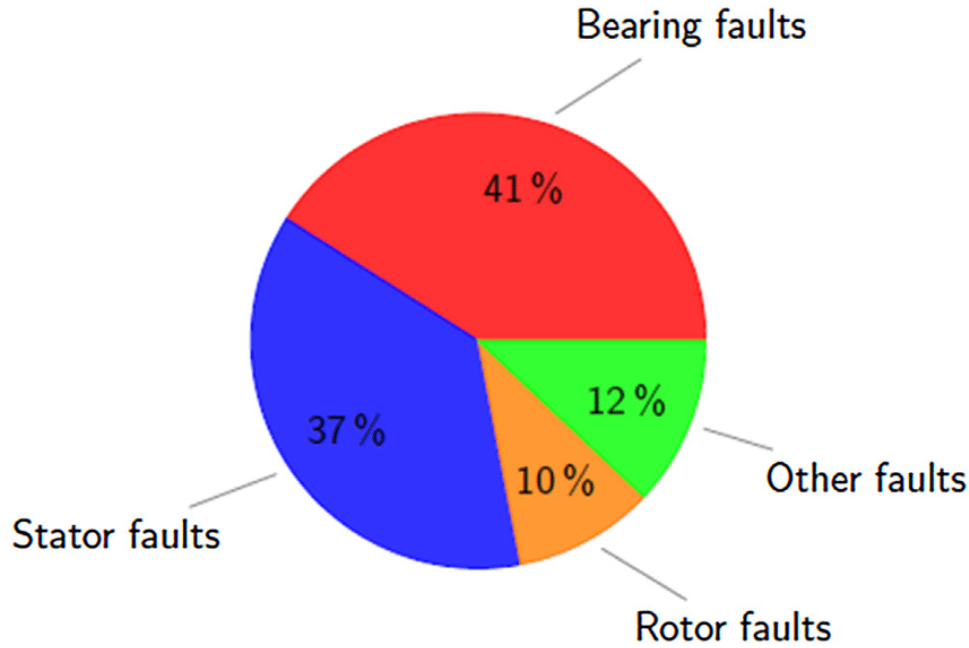


Figure 1: Defect statistic, taken from [reference]

In this project we present a mixed methodology to detect and predict bearing defects. The methodology consists of using signal processing for feature generation and data labelling, and machine learning for defects classification and failure prediction. For a given input dataset, the dataset is decomposed into its subcomponents or basis components. The basis components also called features are then used as input of a supervised learning algorithm for defect classification and failure prediction.

The signal processing methods used are Fourier transform, wavelet transform and Hilbert Huang transform. We focus on ensemble learning and feed forward neural network for classification. Furthermore we show that the back-propagation process in the feed forward neural network can be modelled by an ordinary differential equation, whose solution represents the path of the hidden and output layer weights.

The outline of the project is as followed: In section (chapter) one we give an overview of the signal processing methods. In section (chapter) two we present the machine learning methodology and show that the back-propagation in the feedforward neural network can be modelled as a differential equation. In section (chapter) 3 we present a case study where we apply the methodology defined in section one and two. In section four we present the conclusion of this work.

2 Preview work

3 Signal processing methods

3.1 Overview

We present three signal processing methods for data decomposition or feature generation. The three methods can be described as follow: Given a dataset approximated by a map

$$f : \mathbb{R} \rightarrow \mathbb{R} \quad (1)$$

find a space V spanned by a basis $\{\varphi_j\}_{j=0}^n$ such that

$$f(x) = \sum_{j=0}^{j=n} \alpha_j \varphi_j. \quad (2)$$

In the context of this work, the basis functions $\{\varphi_j\}_{j=0}^n$ are the features derived from the original dataset.

3.2 Fourier transform

3.3 Wavelet transform

3.4 Hilbert Huang transform

4 Statistical approach: dissimilarity measure

In this approach we compute the dissimilarity between each sample. Let X_i and X_j be two samples from the same distribution. Then the Mahalanobis distance between the two random variables with corresponding covariance matrix S and inverse S^{-1} is given by

$$d(X_i, X_j) = \sqrt{(X_i - X_j)^T S^{-1} (X_i - X_j)} \quad (3)$$

The Mahalanobis distance then gives the dissimilarity between the two samples X_i and X_j .

Our dissimilarity measure is inspired by the Mahalanobis distance. Rather than using $X_i - X_j$ we will form a vector $(\text{Var}[X_i], \text{Var}[X_j])$ where $\text{Var}[X_i]$ and $\text{Var}[X_j]$ are the variance of sample X_i and X_j respectively. With this change the new dissimilarity measure is given by

$$d(Z, Y) = (\text{Var}[Z], \text{Var}[Y])^T S^{-1} (\text{Var}[Z], \text{Var}[Y]). \quad (4)$$

Once we have a way to measure the dissimilarity between samples, we set a reference sample X_0 which represents a sample with no defect. Afterwards, we compute the dissimilarity between X_0 and subsequent samples for each bearing. We can implement the above formula in python:

```

def dissimilarity_measure(u,v):
    #inverse of the covariance matrix
    U=np.array([u,v])
    covariance=np.cov(U)
    cov_inv=np.linalg.inv(covariance)
    #dissimilarity between u and v
    V=np.array(list(map(lambda w: np.var(w),[u,v]))).reshape(-1,1).T
    d=np.matmul(np.matmul(V,cov_inv), V.T)
    return d

```

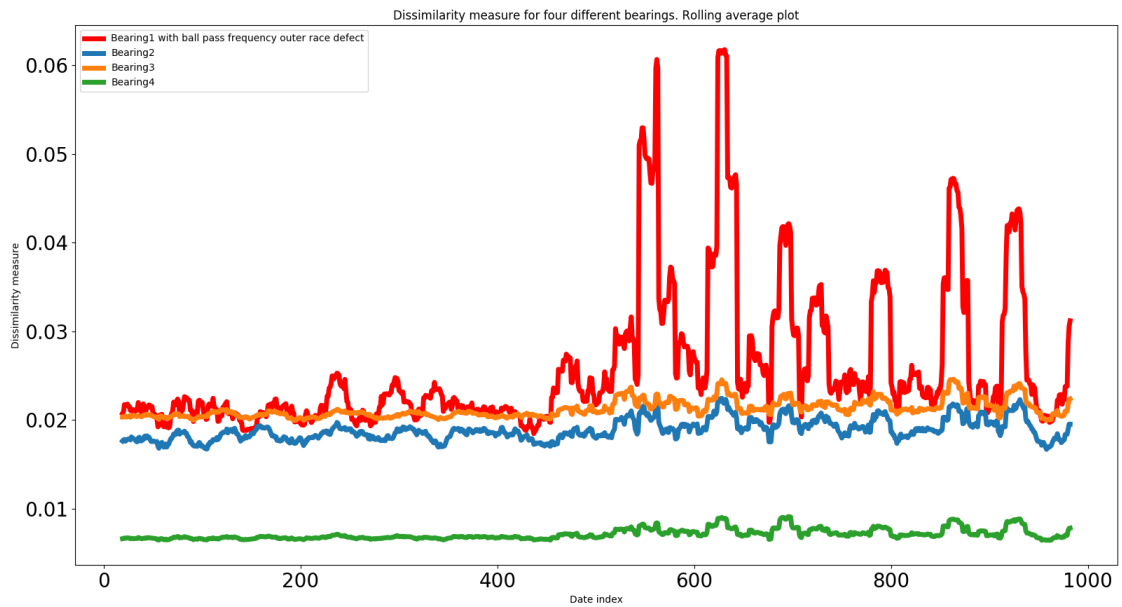


Figure 2: Dissimilarity measure for four bearings. In red a bearing with ball pass frequency outer race defect

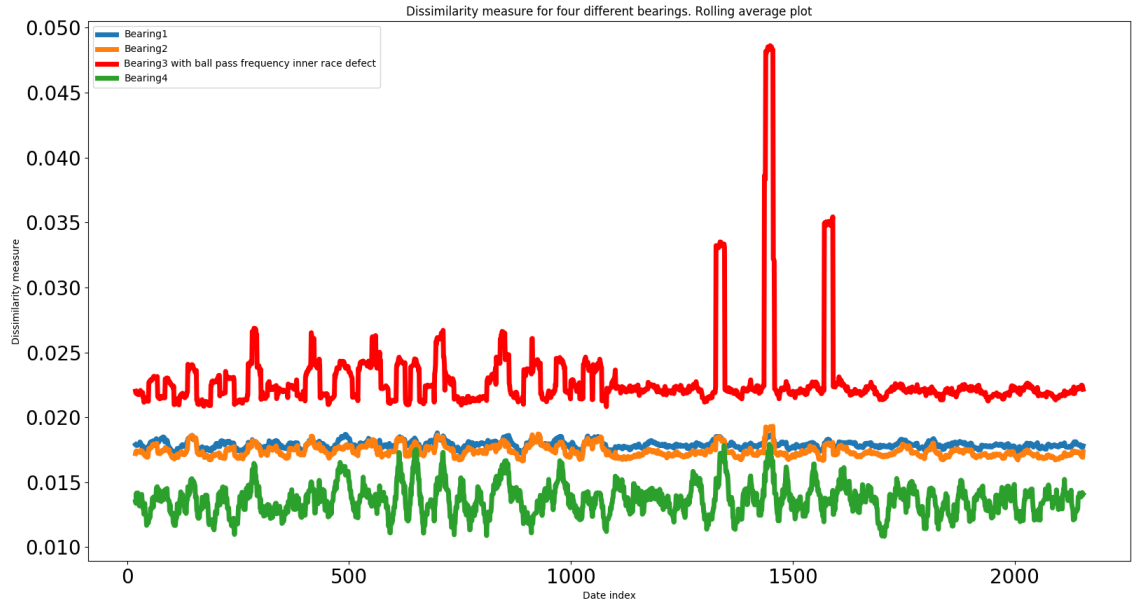


Figure 3: Dissimilarity measure for four bearings. In red a bearing with ball pass frequency inner race defect

As we can see on the above figures, the bearings with outer race and inner race defect have the highest dissimilarity measure. The dissimilarity measures quantifies the health index of a bearing. A higher dissimilarity implies bearing degradation.

5 Mixed method: statistical and wavelet approach

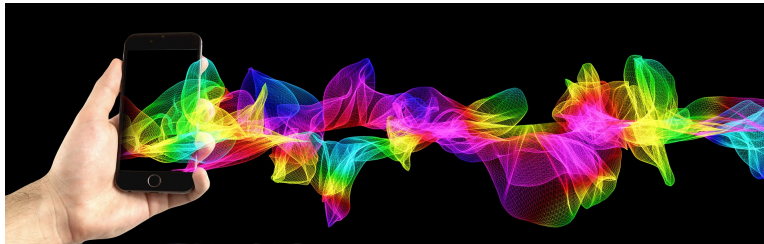


Figure 4:

For the mixed method, we use wavelet transform to generate two extra features from the vibration time signal, namely the discrete detailed coefficient cD and the approximate coefficient cA . The detail coefficient cD represents the high frequency component of the vibration time signal and the approximate coefficient cA represents the low frequency component. For the mother wavelet we use Daubechies 20 or $db20$.

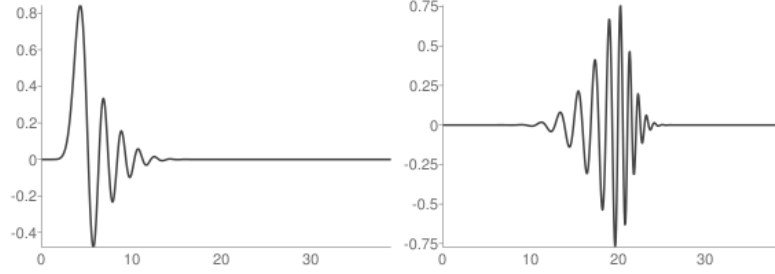


Figure 5: The scaling function φ (left) and the wavelet function Ψ (right)

A pseudo code in python to implement the method:

```

pip install pywavelet

```

```

import pywt
import numpy as np
container = []
mother_wavelet = 'db20'
ref_cA, ref_cD = pywt.dwt(ref_sample, mother_wavelet)
for sample in samples:
    cA, cD = pywt.dwt(ref_sample, mother_wavelet)
    cA_dissimilarity = dissimilarity_measure(ref_cA, cA)
    cD_dissimilarity = dissimilarity_measure(ref_cD, cD)
    container.append((cA_dissimilarity, cD_dissimilarity))

```

Once we have the two extra features, we compute the dissimilarity between a reference sample and subsequent samples for each feature. This process generates a set of points (x, y) that represent the health index of each sample. As can be seen in the figure bellow, a bearing can go through three main stages:

1. A healthy stage characterised by a low health index
2. A warning stage characterised by an increasing health index
3. An alarm stage characterised by a high health index

K.C. Deekshit Kompella a, Venu Gopala Rao Mannam b, Srinivasa Rao Rayapudi

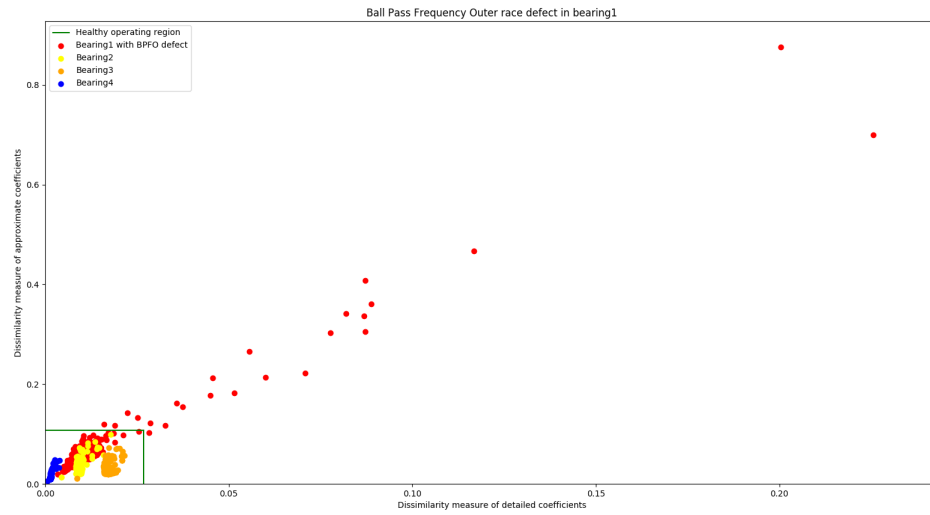


Figure 6:

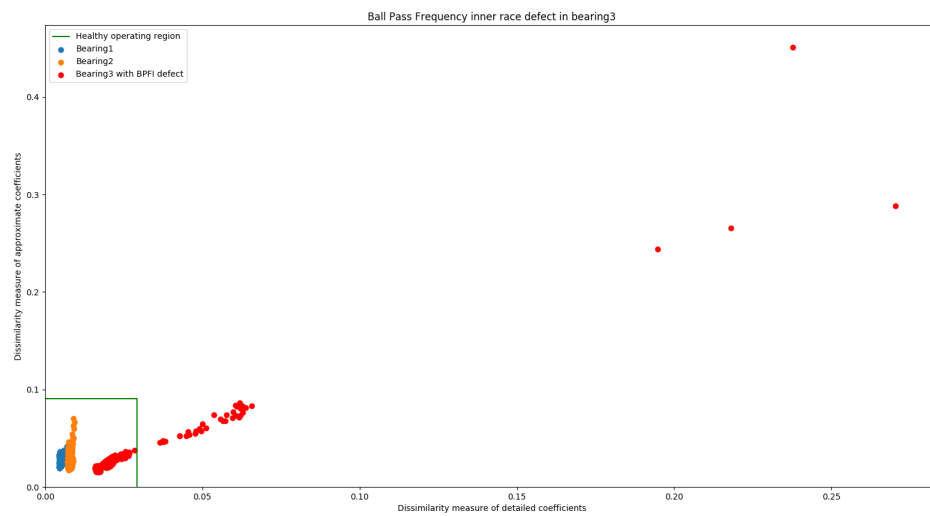


Figure 7:

In the alarm stage, as the degradation becomes more severe, the distance between the points increases.

6 Comparison

The Signal processing method allows finding the nature or type of the fault and monitor the amplitude of the fault frequency. In the mixed method, we can only detect if there is fault but we cant specify for sure which kind of fault. The same goes for the mixed method. However in the mixed method we can estimate the time of failure based on the distance between the points formed by the dissimilarity of the detailed and approximate coefficients of the time signal wavelet transform.

The signal processing method relies on the RPM (the speed of the motor) to compute the fault frequency. If the RPM is inaccurate, it can be challenging to find a fault frequency. The statistical and mixed method only rely on the time signal.

7 Defining the use case

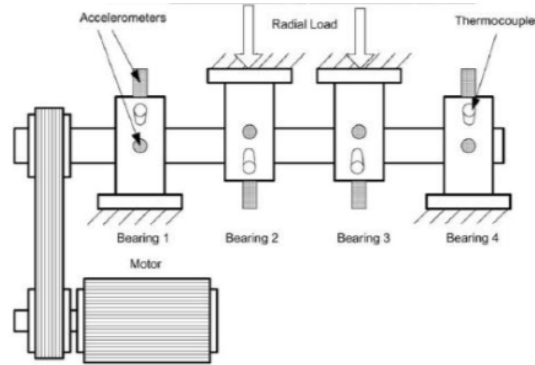


Figure 8: Experimental set up

The data used in this use case was generated by the Intelligence Maintenance system (IMS) [link to the IMS]. Three separate experiments involving four bearings were performed on a motor. In each experiment, a 1-second vibration signal snapshots was recorded every 10 minutes, for a specified time. Each vibrational signal sample consists of 20 480 data points with sampling rate of 20 000 Hz. In this post I will be using the first and the second experiment data [available here (give a link to the data)].