ANOMALY DETECTION: Identification of rare items, events or observations (e.g. malfunctioning of equipment)

Data generated when the equipment approaches failure, or a sub-optimal operation, typically have a different distribution than data from “healthy” equipment.

One of the main techniques for dimensionality reduction is PCA which performs a linear mapping of the data to a lower-dimensional space in such a way that the variance of the data in the low-dimensional representation is maximized. In practice, the covariance matrix of the data is constructed and the eigenvectors of this matrix are computed. The eigenvectors that correspond to the largest eigenvalues (the principal components) can now be used (spanning a new feature space) to reconstruct a large fraction of the variance of the original data.

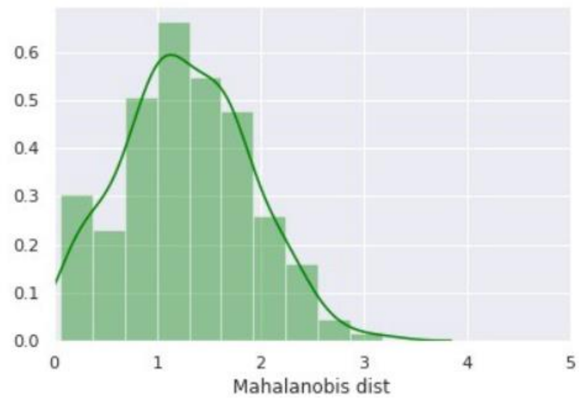
a) Mahalanobis distance (MD): If the distribution is non-spherical (e.g. ellipsoidal), then we would expect the probability of the test point belonging to the set to depend not only on the distance from the center of mass/centroid, but also on the direction. In those directions where the ellipsoid has a short axis the test point must be closer, while in those where the axis is long the test point can be further away from the center. Putting this on a mathematical basis, the ellipsoid that best represents the set’s probability distribution can be estimated by calculating the covariance matrix of the samples. MD is the distance of the test point from the centroid divided by the width of the ellipsoid in the direction of the test point.

   We are only interested in classifying “normal” vs “anomaly”, we use training data that only contains normal operating conditions to calculate the covariance matrix. Given a test sample, we compute the MD to the “normal” class and classify the test point as an “anomaly” if the distance is above a certain threshold. Considering our input variables follow a normal distribution (mean and covariance matrix), MD has been used.

USE CASE

As the equipment was run until failure, data from the first two days of operation was used as training data to represent normal and “healthy” equipment. The remaining part of the datasets for the time leading up to the bearing failure was used as test data, to evaluate whether the different methods could detect the bearing degradation in advance of the failure.

Distribution of MD for training data representing “healthy” equipment:



We can define a threshold value now. From the distribution above, MD > 3 can be defined as an anomaly. The evaluation of this method to detect equipment degradation consists of calculating the MD for data points in the test set and comparing them to the threshold value for flagging an anomaly.

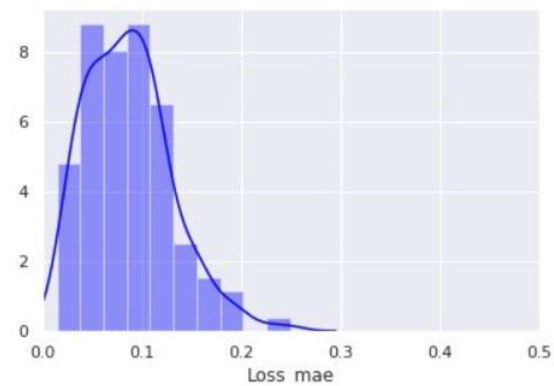
b) NN : An autoencoder is a type of ANN used to learn efficient data codings (representation for a set of data) in unsupervised manner, typically for dimensionality reduction. Along with the reduction side, a reconstructing side is learnt, where the autoencoder tries to generate from the reduced encoding a representation as close as possible to its original input. Architecturally, the simplest form of an autoencoder is a feedforward, non-recurrent NN very similar to the multilayer perceptron (MLP) — having an input layer, an output layer and one or more hidden layers connecting them — but with the output layer having the same number of nodes as the input layer, and with the purpose of *reconstructing* its own inputs.

In the context of anomaly detection and condition monitoring, the basic idea is to use the autoencoder network to "compress" the sensor readings to a lower-dimensional representation, which captures the correlations and interactions (nonlinear as well) between various variables.

The network is trained on data representing the "normal" operating state, with the goal of first compressing and then reconstructing the input variables. During the dimensionality reduction, the network learns the interactions between the various variables and re-constructs them back to the original variables at the output. The main idea is that as the monitored equipment degrades, this should affect the interaction between the variables (e.g. changes in temperatures, pressures, etc.). As this happens, an increased error happens in the network's re-construction of the input variables. By monitoring the re-construction error, one can thus get an indication of the "health" of the monitored equipment. Quantitatively, one uses the probability distribution of the reconstruction error (or loss) to identify whether a data point is normal or anomalous.

USE CASE

Distribution of reconstruction loss (MAE) for the training data representing "healthy" equipment:



From the distribution above, a loss> 0.25 can be defined as an anomaly. The evaluation of the method to detect equipment degradation now consists of calculating the reconstruction loss for data points in the test set and comparing the loss to the defined threshold value for flagging an anomaly.

This would allow predictive measures (maintenance/repair) to be taken in advance of the failure, which means both cost savings as well as the potential importance for HSE (health, safety, environment) aspects of equipment failure.