PHM Challenge 2023 Methodology description

First the instances were split into segments where each segment includes 5 full rotations of the driven gear. From these segments, common parameters for vibration analysis like variance and root mean squared were calculated. Afterwards a Fourier Transformation was applied to each segment. The resulting frequency spectrum was than transformed using Mini ROCKET (**R**and**o**m **C**onvolution **Ke**rnel **T**ransform). Mini ROCKET had difficulties to extract valuable features from the frequency spectrum, this is likely due to the random nature of ROCKET and the frequency spectrum being very low for most of the frequency domain. Because of this, a logarithmic function was applied to highlight frequencies with lower density. This transformed the frequency spectrum into a richer representation, that Mini ROCKET could take advantage of.

The prediction is divided into 3 parts to handle the unknown degradation states (states, not present in the trainings data). The first part handles the states that are available in the training set. The second part handles the states not available in the training set and the last part handles the operational conditions that are not available in the training set.

In the first part a One vs All approach for each unique operating condition was selected. The features extracted in the preprocessing step from one operating condition were used to train 7 Ridge Classifier (as recommended for the ROCKET Transform), corresponding to the 7 known states. If one model predicted 1 (in contrast to 0) the instance was classified as the respective state. Also, synthetic classes were added that were generated from the known states. The thought behind this is to produce synthetic states close to the known states to increase the certainty a model needs to predict positive. This reduces the probability that an unknown state from the test state gets classified as a known state. This procedure achieved 100% accuracy on the training data, even when some states were handled as unknown states.

In the second part, instances that were not classified as a known degradation state, are handled. This time, one Ridge Regression model was trained for each operating condition. Due to the splitting of the cases in the preprocessing part, there were multiple predictions for each instance. This results here in a probability distribution for each instance. Each of these predictions with a continuous value is rounded to fit into one of the degradation states. Since the first part already identified instance that are part of a known degradation state, predictions of this step, that indicate an association with one of these classes were not adopted. If for example one instance was split into 5 segments and this model classified 4 segments as degradation state 5 (unknown) and 1 as state 4 (known), the probability was set to 0.8 for class 5 and 0 for all other classes.

Finally for the operating conditions with no trainings available, no model was found that produces points over 1, so all probabilities were set to 0. However, the best model that was found was a Regression model trained with instances with similar operating conditions. For this, an ANN is used. The first layer is a custom layer where the features just get passes into the next layer with one connection per feature. Only the added features rotational speed and torque connect to all neurons of the next layer. The final layer is then again, a ridge regression layer, that uses the features that were processed with the operating conditions.

The complete training and inference of training and test/validation set takes less than an hour on an intel i5. This speed is one strength of this algorithm which is thanks to high dimensionality reduction by using a Fourier transformation and the quick feature transform of Mini ROCKET. To my knowledge this combination is a novel approach. Furthermore, for the final model there was not much hyperparameter tuning necessary except for the number of rocket features used and the regularization value alpha for the Ridge models. Also, no deep analysis or domain knowledge was required. In summary the provided model is a fast trained and inference in time suitable model, that can easily be reused for similar problems in vibration analysis.