Regular Project Bearing Failure Report

Michał Filipiak
Patryk Grochowicz
Natalia Nazarczuk
Dominik Borkowicz
Piotr Pilis

Faculty of Mathematics and Information Science

Abstract

Bearing failures have a significant impact on industry machines and the economy. Many various reasons can cause it. It is essential to understand why it appears to quickly spot a defective device as the loss of the productivity of such a machine can be pretty expensive. In our project, we prepare two forecasting models, one based on Recurrent Neural Network second which is predicting time-to-failure variable. We also prepared Exploratory Data Analysis which will be explained in the further chapters of the report. Our results are based on data generated by NSF I/UCR Center for Intelligent Maintenance Systems.

1 Introduction

In our project, we were provided with data collected during three separate measurements. The device used for collecting data was four bearings installed on the shaft. In these measurements, the speed was constant at 2000 RPM provided by a motor coupled to the shaft via rub belts. What is more, a radial load of 6000 pounds was applied onto the shaft and bearing by a spring mechanism. All bearings were force lubricated. For measurements, High Sensitivity Quartz ICP accelerometers were installed in data set 1, two of them for each bearing y and x-axis and only one for data sets 2 and 3. Sensor placement is shown in Fig 1. In the results provided by NSF I/UCR Center for Intelligent Maintenance Systems, all failures occurred after exceeding the designed lifetime of the bearing, which is more than 100 million revolutions.

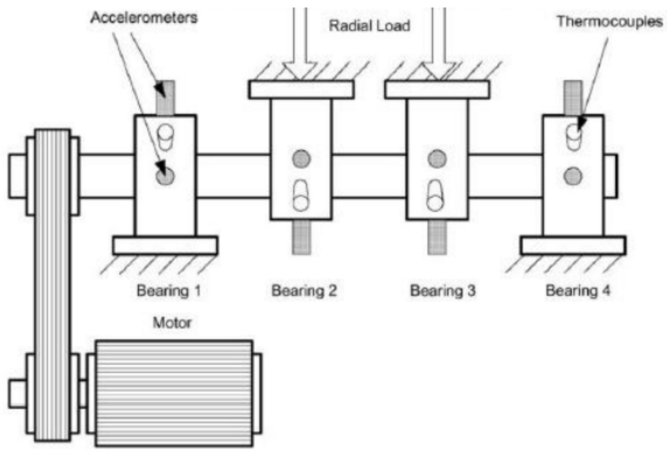

Fig. 1: Bearing test rig and sensor placement

The 3 data sets we were provided with describes a test-to-failure experiments. All the observations recorder in data sets were taken in very specific way every 10 minutes, a vibration signal snapshot was collected, with a sampling rate of 20kHz. This means every 10 minutes, during a 1 second period, 20480 data points were added to the set

2 Exploratory Data Analysis

2.1 Data

For Exploratory Data Analysis, we read data from three data sets containing the measurements made during the experiment. For performance reasons, we are sampling the data from each file in regular intervals to obtain around 1000000 rows in each set. What is more, we also manually added a column to each of the data set containing the time after the start of the experiment given in minutes.

	Bearing 1 x	Bearing 1 y	Bearing 2 x	Bearing 2 y	Bearing 3 x	Bearing 3 y	Bearing 4 x	Bearing 4 y	Time
0	-0.115	-0.027	0.044	-0.073	-0.137	-0.032	-0.129	-0.181	0
1	-0.132	-0.024	-0.115	-0.063	-0.007	-0.149	-0.095	-0.161	0
2	-0.125	-0.178	-0.291	-0.186	0.024	-0.398	-0.037	-0.120	0
3	-0.095	-0.146	-0.229	-0.068	-0.085	-0.168	-0.042	-0.103	0
4	-0.073	-0.002	0.005	-0.117	-0.176	-0.171	-0.144	-0.083	0
...	...	...	...	...	...	...	...	...	...
963523	-0.220	-0.164	-0.171	-0.427	0.347	0.193	-0.330	-0.081	21120
963524	-0.159	0.000	-0.066	-0.181	0.518	0.176	-0.017	-0.049	21120
963525	0.085	-0.188	-0.210	-0.012	0.457	0.081	-0.349	-0.176	21120
963526	-0.032	-0.107	-0.100	0.181	0.303	-0.222	0.105	-0.137	21120
963527	-0.168	-0.068	-0.449	0.190	0.117	1.196	-0.037	-0.122	21120

Fig. 2: First data set

The first data set contain measurements made by accelerometers installed on x and y of the bearings and an additional time column.

	Bearing 1	Bearing 2	Bearing 3	Bearing 4	Time
0	-0.049	-0.071	-0.132	-0.010	0
1	-0.037	0.017	-0.044	0.032	0
2	-0.002	0.002	-0.159	-0.002	0
3	0.010	-0.054	0.166	-0.066	0
4	-0.056	-0.081	-0.088	-0.002	0
...	...	...	...	...	...
1007611	-0.002	0.000	0.000	-0.002	9830
1007612	0.000	0.000	0.000	-0.002	9830
1007613	-0.002	0.000	0.000	-0.005	9830
1007614	-0.002	-0.002	-0.002	-0.005	9830
1007615	-0.002	-0.002	-0.002	-0.002	9830

Fig. 3: Second data set

The second data set contains measurements made by an accelerometer installed on the bearings and an additional time column.

	Bearing 1	Bearing 2	Bearing 3	Bearing 4	Time
0	0.034	0.264	0.039	-0.046	0
1	0.044	-0.066	-0.054	0.122	0
2	0.046	-0.046	-0.115	0.029	0
3	-0.059	0.281	0.010	-0.139	0
4	0.022	0.049	0.015	0.015	0
...	...	...	...	...	...
999187	0.000	0.002	0.005	0.000	63230
999188	0.002	0.005	0.005	0.000	63230
999189	0.000	0.002	0.005	0.000	63230
999190	0.000	0.002	0.005	0.000	63230
999191	0.002	0.005	0.005	0.002	63230

Fig. 4: Third data set

The third data set contains measurements made by an accelerometer installed on the bearings and an additional time column.

2.2 Distributions

After reading the data we checked their distribution for further analysis. We have done this step for all data set.

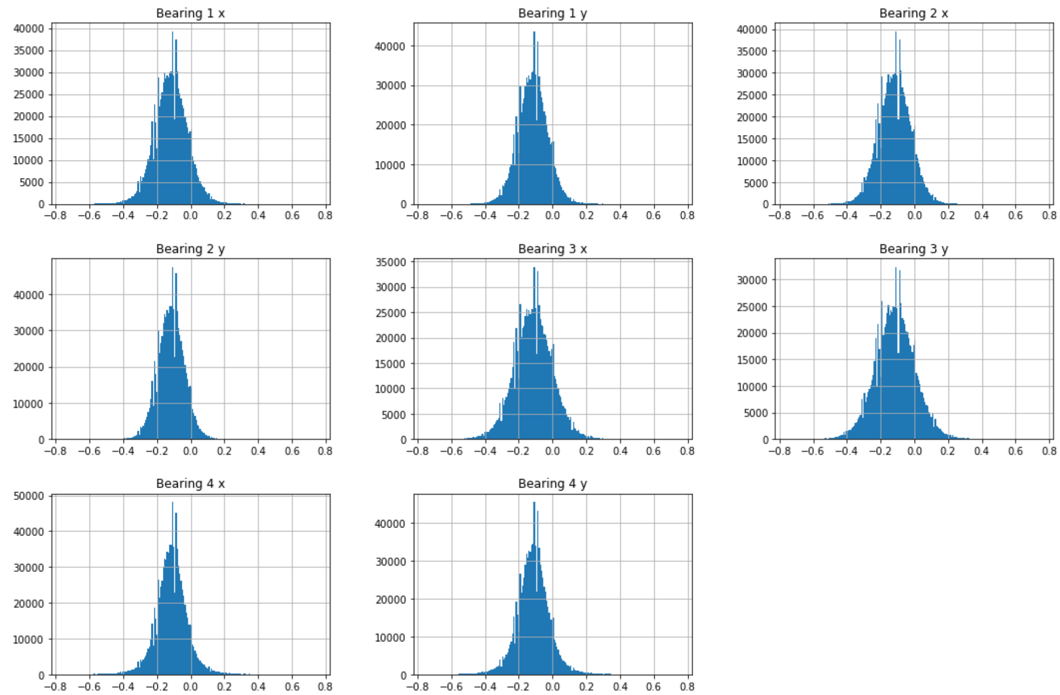

Fig. 5: First data set distribution

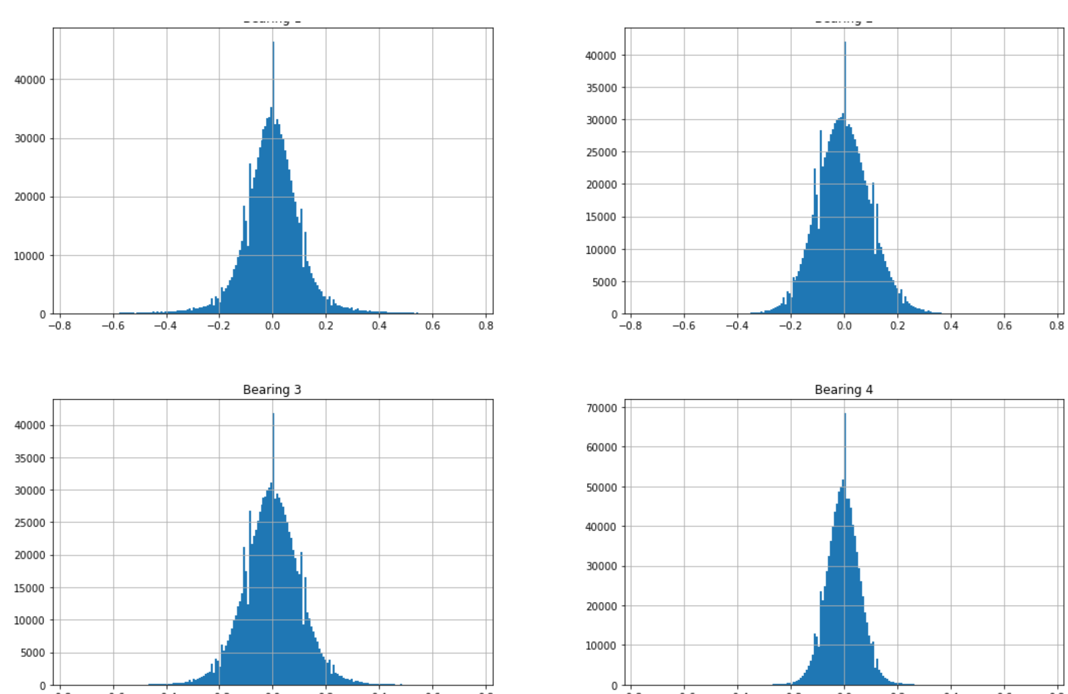

Fig. 6: Second data set distribution

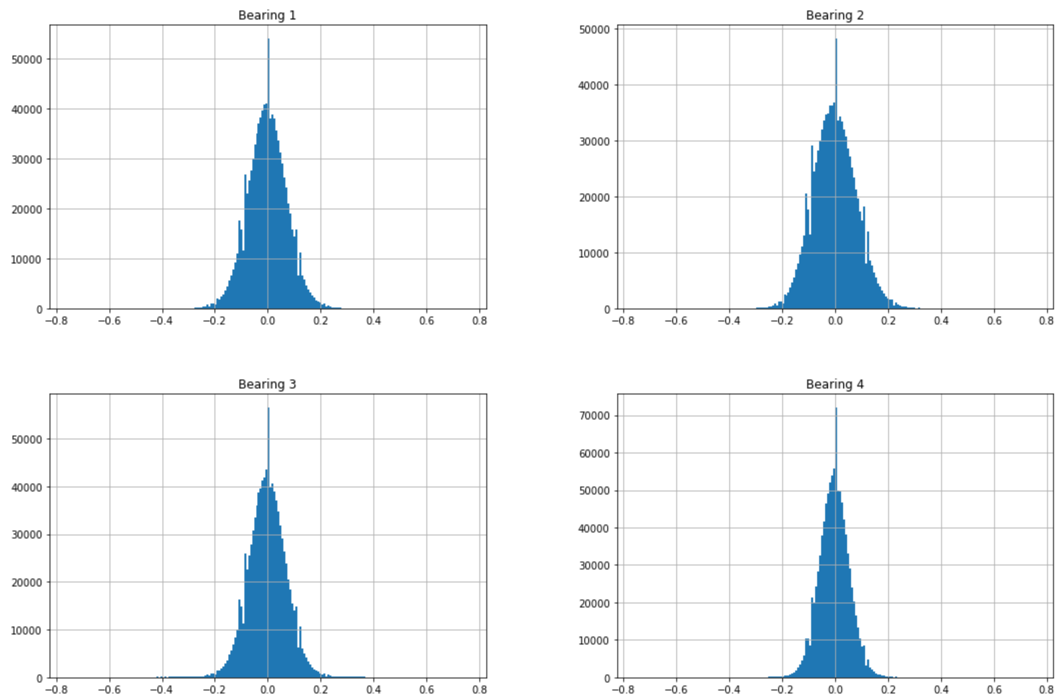

Fig. 7: Third data set distribution

One can notice that this is a normal distribution. To be precise, it is normal leptokurtic distribution. It is visible because of the spikes in the midpoint that mean the positive kurtosis is more significant than in a bell-shaped distribution. What is more, there is no visible skewness.

2.3 Outliers

After checking the distribution we analysed the outliers for all data sets. We have done this to check for the data that escape normality.

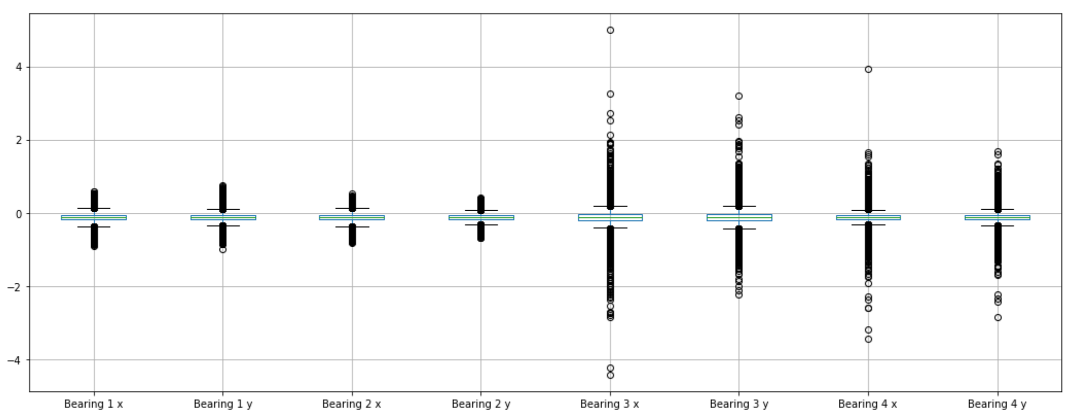

Fig. 8: First data set outliers

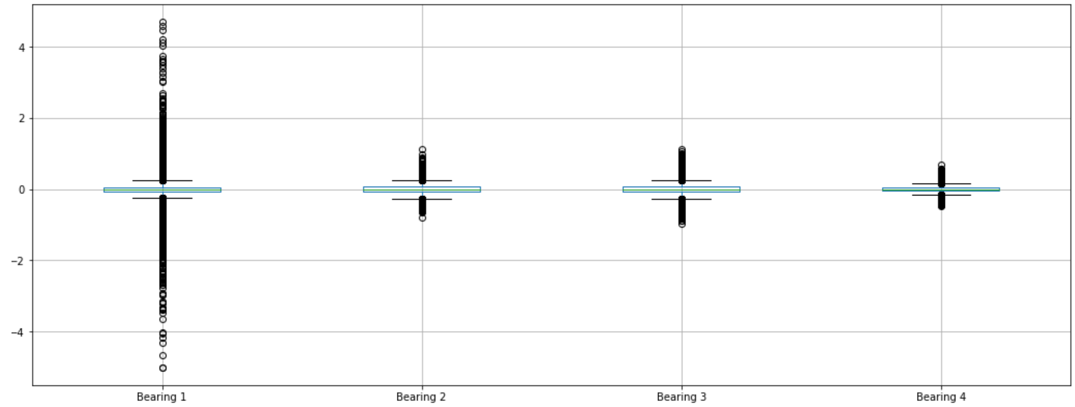

Fig. 9: Second data set outliers

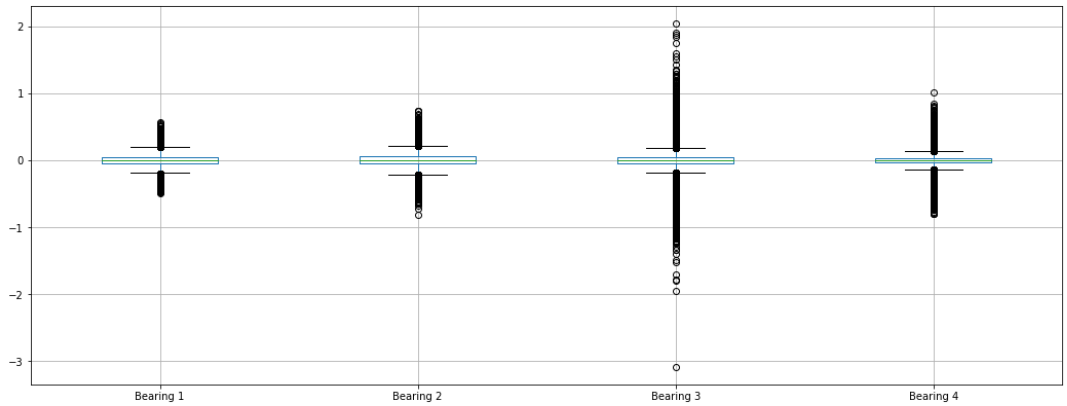

Fig. 10: Third data set outliers

From the results obtained through graphs one can notice that in the first data set bearing 3 and bearing 4 has the most data points that significantly differ from other observations. What is more in the second data set bearing 1 was the one that has the most points that escape normality and in the third data set one can notice that bearing 3 has the most abnormalities. This concludes the results obtained by the NSF I/UCR Center for Intelligent Maintenance Systems as in their experiment in first measurement bearing 3 and bearing 4 were the ones that the race defect occurred, in second measurement bearing 1 was the one that race failure occurred and in third measurement bearing 3 was the one that race defect occurred.

2.4 Correlation Diagrams

After checking outliers we checked the correlation between time and collected data in data sets. We done this by ploting the correlation diagrams.

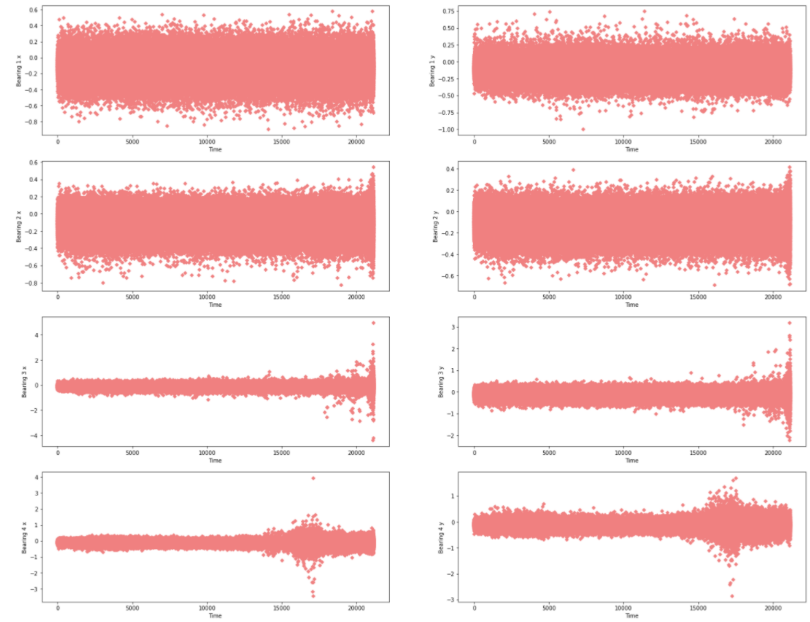

Fig. 11: First data set correlation diagrams

From the result of the second diagram, one can notice that in bearing three and bearing four at the end, there is the broadest range of measurements values.

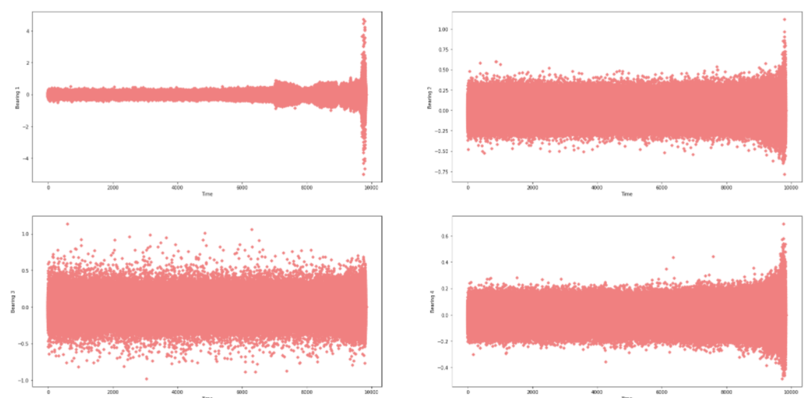

Fig. 12: Second data set correlation diagrams

From the result of the first diagram, one can notice that in bearing one at the end, there is the broadest range of measurements values.

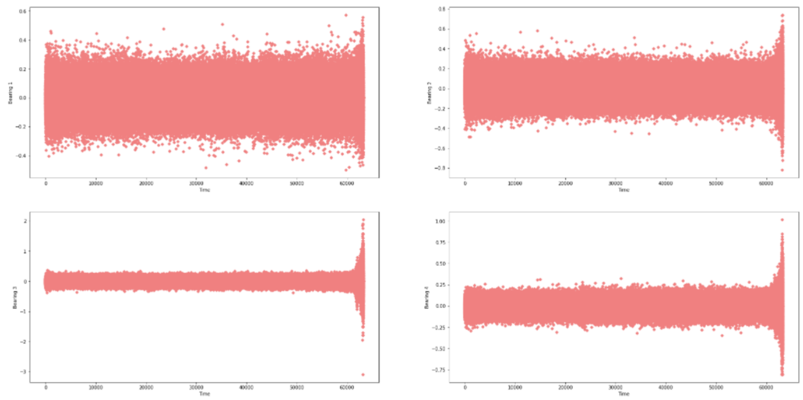

Fig. 13: Third data set correlation diagrams

From the result of the third diagram, one can notice that in bearing three at the end, there is the broadest range of measurements values.
All of these diagrams can be connected with the result obtained by NSF I/UCR Center for Intelligent Maintenance Systems as those bearings that had the broadest range of measurement values were the ones that the defect occured.

3 Recurrent Neural Network forecasting model

3.1 Introduction

Recurrent neural networks are a broad class of networks where the same parameters are applied again and again even as the inputs change as time passes. In our work we divided it into single step model that predicts one future result and multi step model which predicts multiple future results. For the purpose of comparing results we calculated mean absolute error which will show how far our predictions are from data points.

3.2 Single step model

3.2.1 Baseline

Before building a trainable model we decided to create a performance baseline as it will be used to compare with later made more complicated models.

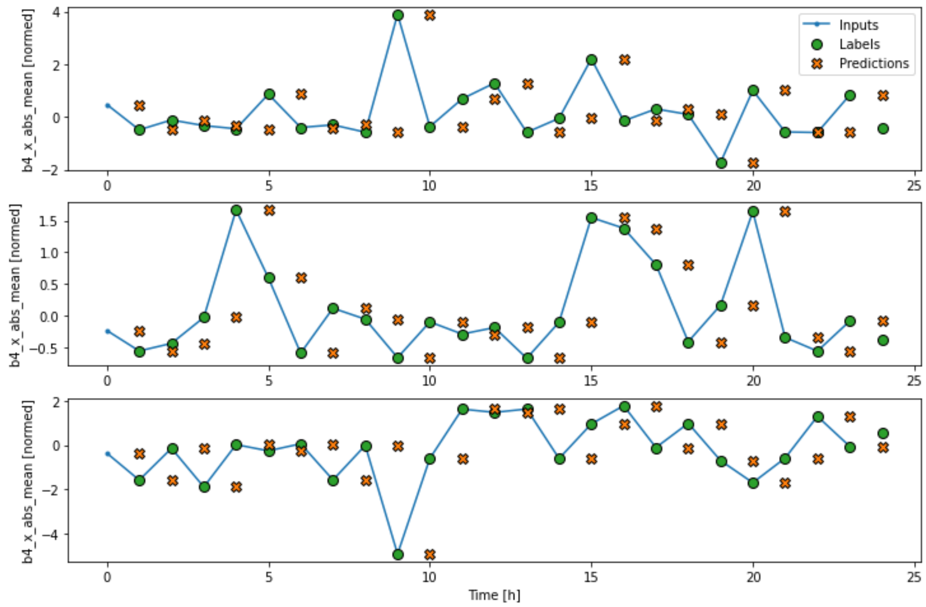

Fig. 14: Baseline result

The result of our baseline model was the value of absolute mean error equal to 1.0381. This means that if we take a random point from data set our prediction will be 1.0381 away from real value.

3.2.2 Linear transformation

We used this model as it is the simplest trainable model you can apply to this task is to just insert linear transformation between the input and output.

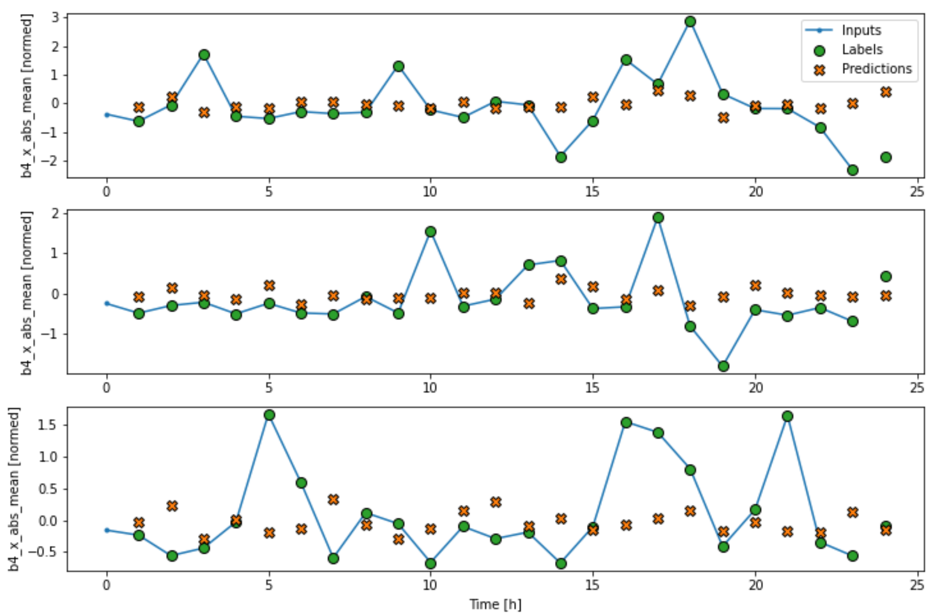

Fig. 15: Linear result

The result of our model was the value of absolute mean error equal at the start to 0.9350 and at the end 0.8201. This shows that while training, our model reduced the mean error by 0.1149. This means that we obtained predictions that are much closer to actual values in comparison to baseline.

3.2.3 Long short-term memory

For this model, we used a Long short-term memory network, which is the type of Recurrent Neural Network, which allow to remember past data in memory. Typically LSTM is composed of a cell, an input gate, an output gate, and a forget gate. The cell is responsible for remembering data over an arbitrary time interval and the gates are responsible for regulating data flow into and out of the cell.

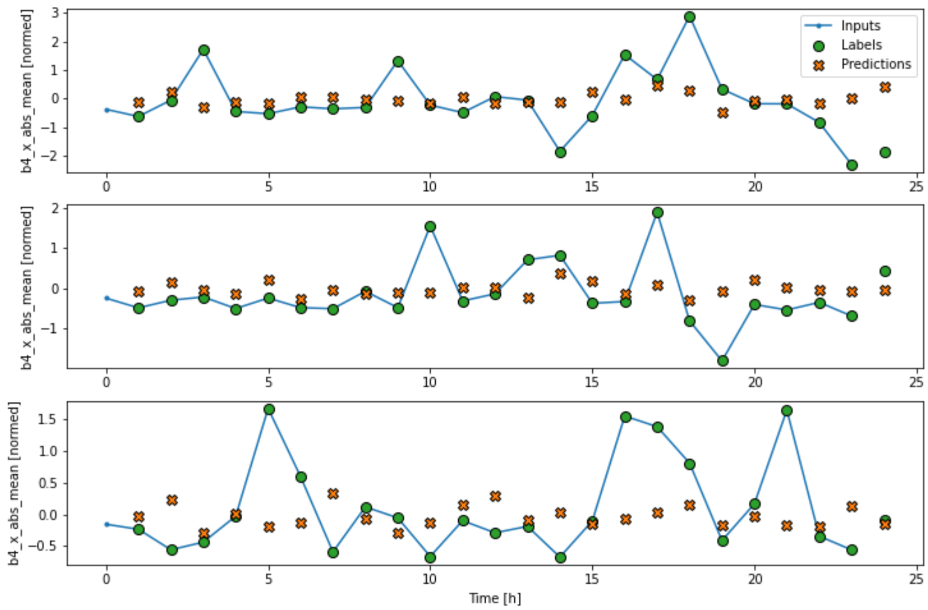

Fig. 16: LSTM result

The result of our model was the value of absolute mean error equal at the start to 0.7043 and at the end 0.6966. This shows that while training, our model reduced the mean error by 0.0077. The difference we obtained while training model was smaller than in linear transformation between the input and output; however, an absolute mean error in this model is much smaller. This shows that our predictions are closer to actual values.

3.3 Multiple step model

3.3.1 Baseline

In multiple step model we as well created performance baseline to use it for comparison with later made more complicated models.

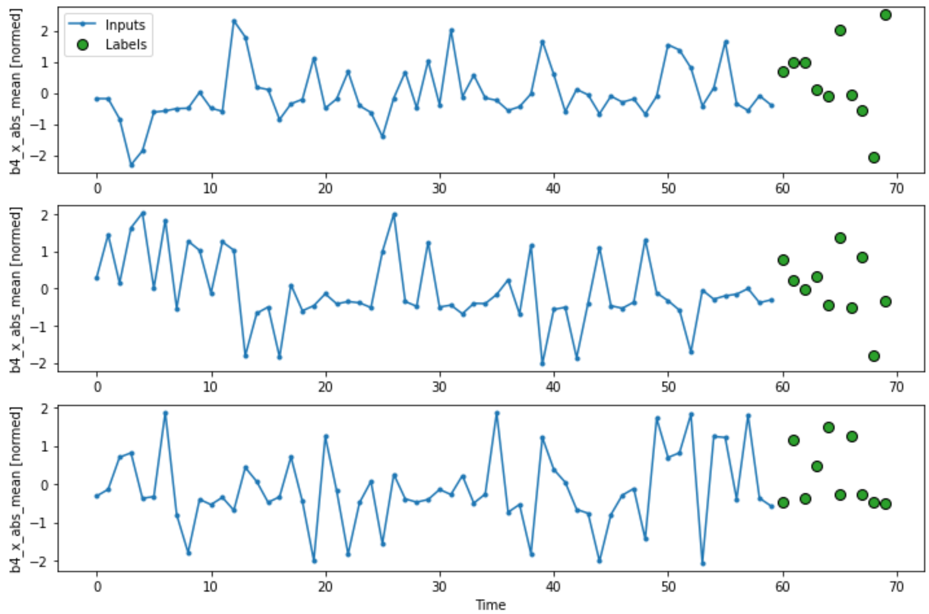

Fig. 17: Baseline result

The result of our baseline model was the value of absolute mean error equal to 0.8391. This means that if we take a random point from data set our prediction will be 1.0381 away from real value.

3.3.2 Linear transformation

For this model, we used like in the single step model a linear transformation between the input and output.

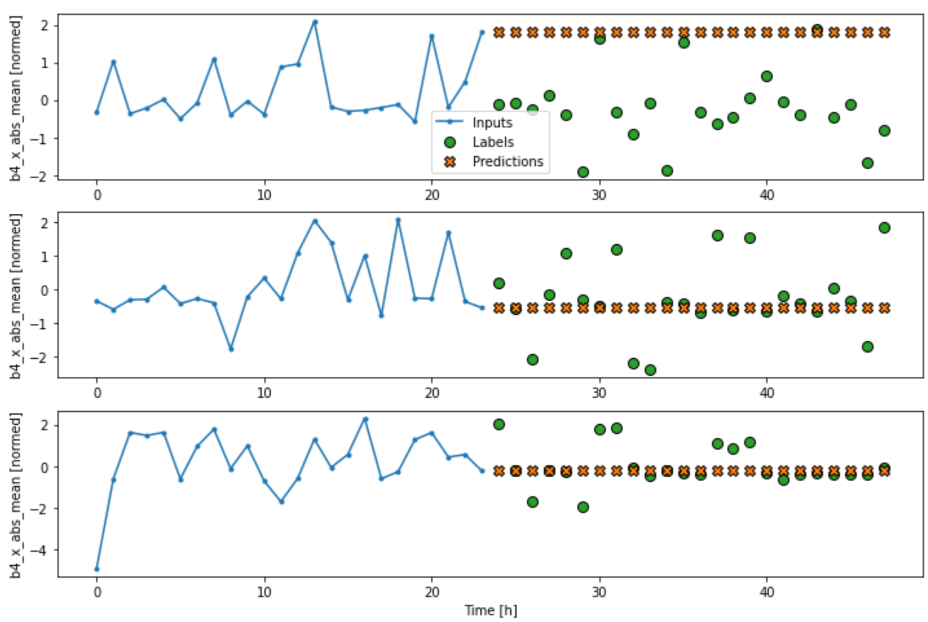

Fig. 18: Linear result

The result of our model was the value of absolute mean error equal at the start to 0.6223 and at the end 0.6220. This shows that while training, our model reduced the mean error by 0.0003. This means that we obtained predictions that are much closer to actual values in comparison to baseline.

3.3.3 Long short-term memory

For this model, we used like in the single step model a Long short-term memory network.

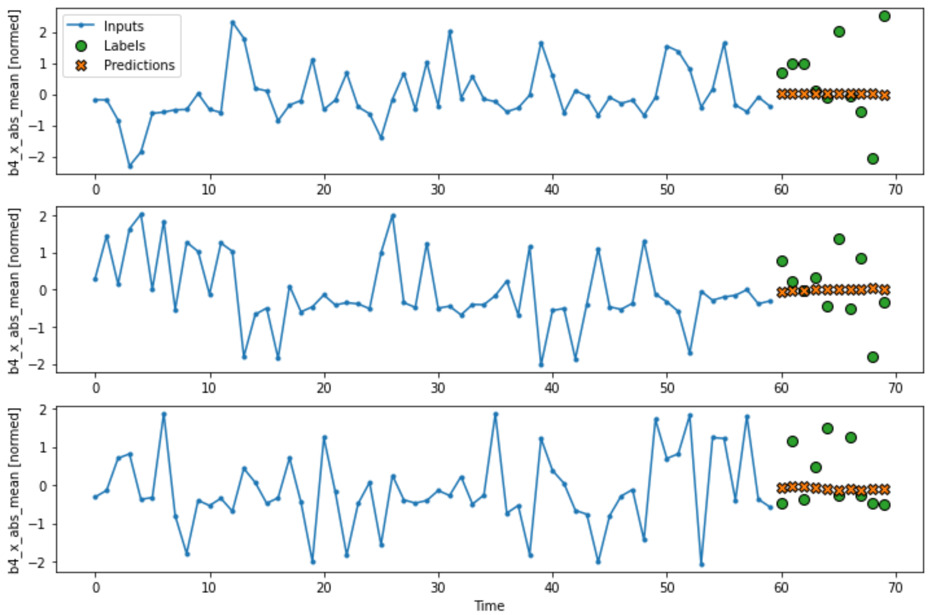

Fig. 19: LSTM result

The result of our model was the value of absolute mean error equal at the start to 0.6492 and at the end 0.6466. This shows that while training, our model reduced the mean error by 0.0026. The difference we obtained while training model was bigger than in linear transformation between the input and output; however, and an absolute mean error in this model is a little bit bigger. This shows that our predictions are further to actual values.

3.3.4 Autoregresive RNN

In some cases it may be helpful for the model to decompose this prediction into individual time steps. Then, each model's output can be fed back into itself at each step and predictions can be made conditioned on the previous one, like in the classic Generating Sequences With Recurrent Neural Networks. One clear advantage to this style of model is that it can be set up to produce output with a varying length.

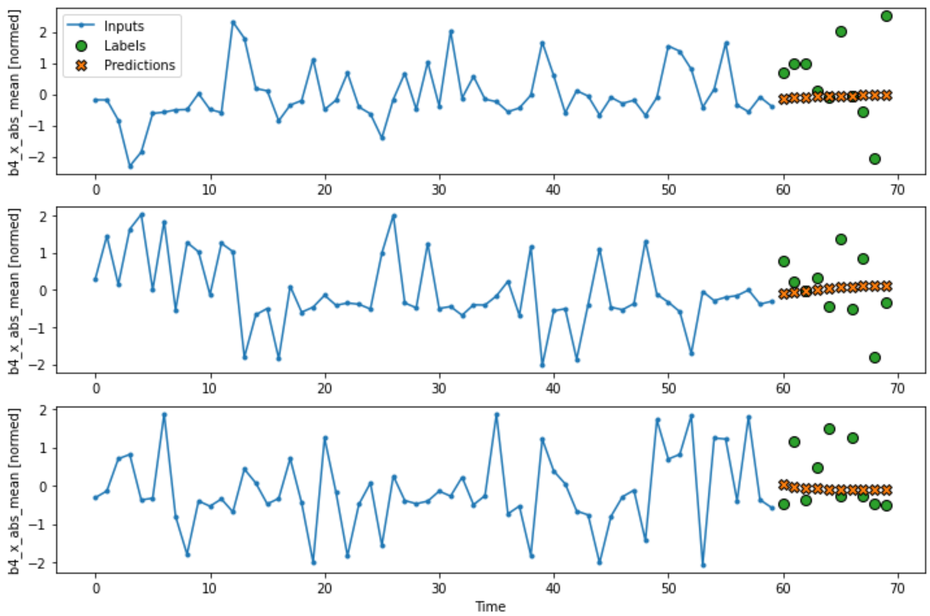

Fig. 20: Autoregresive RNN result

The result of our model was the value of absolute mean error equal at the start to 0.6494 and at the end 0.6471. This shows that while training, our model reduced the mean error by 0.0023. Compared to the LSTM model, we obtained the value of absolute mean error is bigger, so our predictions are slightly further from actual values.

4 Time-to-failure prediction

4.1 K-means algorithm

K-means is an clustering algorithm based on distance between future points.Due to data discontinuity and aging factors showing regardless of any feature, use of clustering methods yielded in unsatisfactory results. There was not also a single cluster which could predict bearing failure.

4.2 Support Vector Machine

Support Vector Machines are based on hyperplane's points optimization. It can perform both supervised and unsupervised machine learning.
We used SVM regression algorithm to predict time-to-left variable. Our results Yielded -7 R^2 value and 430 RMSE error.

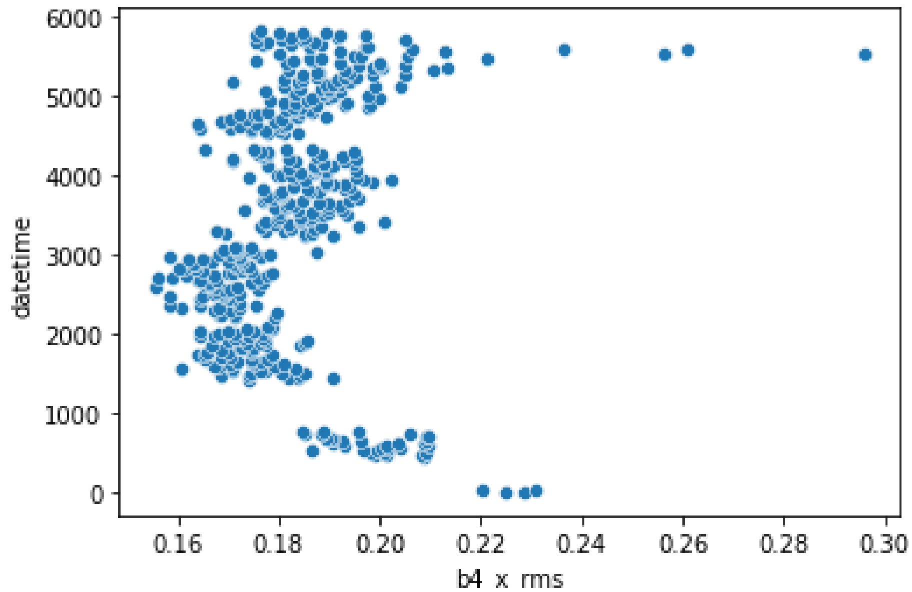

Fig. 21: Dependence of actual time-left value on RMS of signal

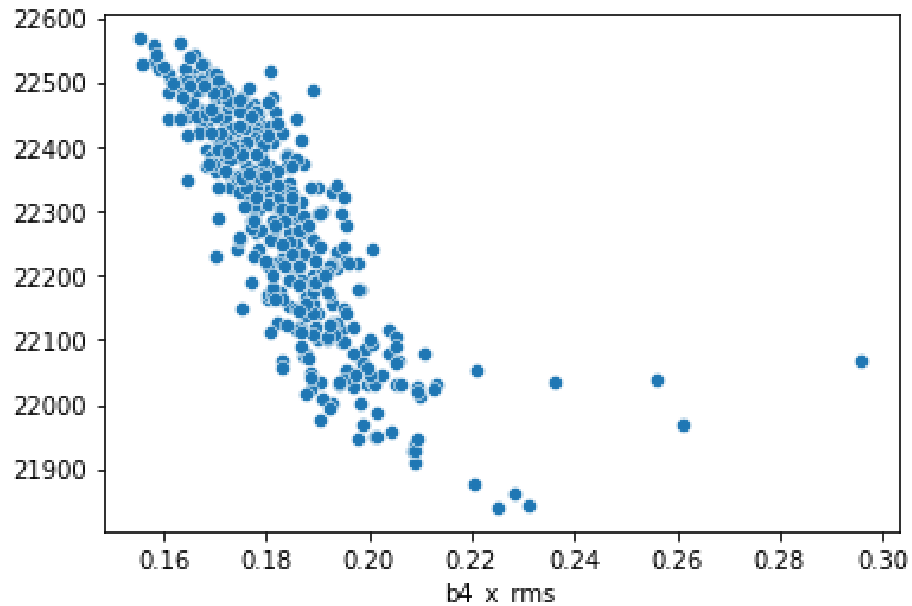

Fig. 22: Dependence of predicted time-left value on RMS of signal

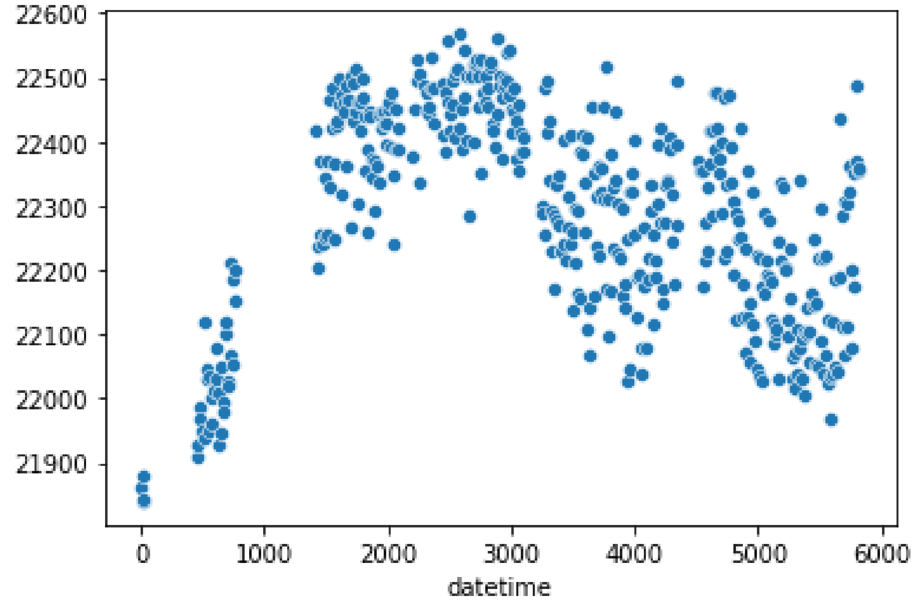

Fig. 23: Dependence of actual value on predicted value

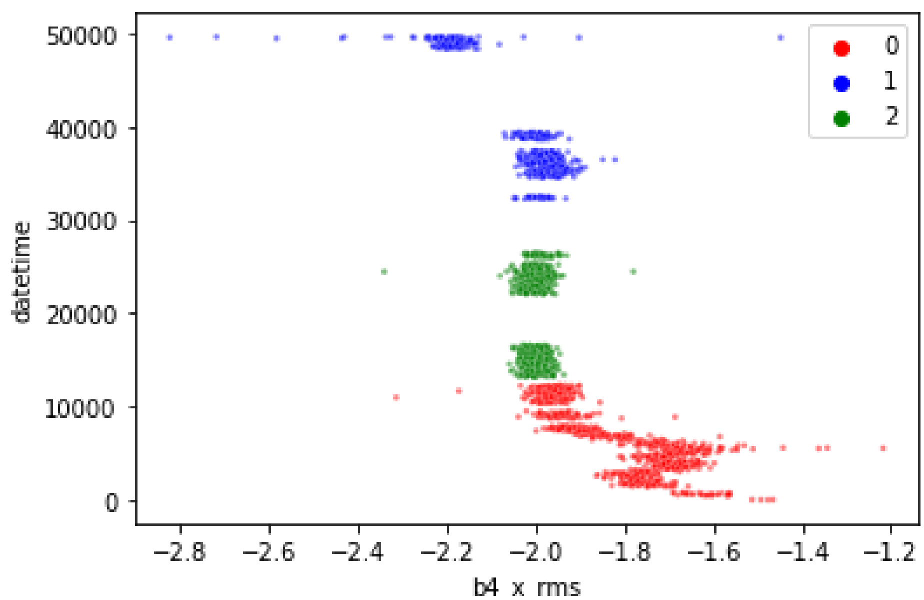

Fig. 24: Distribution of RMS to Time

According to the literature, RMS is the indicator of Bearings inefficiency and correlates with the health of the bearing. So from the obtained data, we can predict bearing failure 100 hours before it occurs in real life.

5 References

1. https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository/bearing
2. Practical Time Series Analysis written by Aileen Nielsen
3. https://www.machinecurve.com/index.php/2019/10/04/about-loss-and-loss-functions/loss-functions-for-regression
4. Hidden Markov Models for failure diagnostic and prognostic
5. A data-driven failure prognostics method based on mixture of Gaussians hidden Markov models, Tobon-Mejia, Diego Alejandro and Medjaher, Kamal and Zerhouni, Noureddine and Tripot, Gerard, Reliability, IEEE Transactions on, Vol. 61 No. 2, 491–503, 2012
6. https://en.wikipedia.org/wiki/Long_short-term_memory