## Corteva Test Analysis

#### May 16, 2022

[165]: # import libraries

```
import warnings
       warnings.simplefilter(action='ignore', category=FutureWarning)
       import os
       import pandas as pd
       import numpy as np
       from sklearn.preprocessing import MinMaxScaler
       import joblib
       import seaborn as sns
       sns.set(color_codes=True)
       import matplotlib.pyplot as plt
       %matplotlib inline
       from numpy.random import seed
       import tensorflow as tf
       from tensorflow.keras.layers import Input, Dropout, Dense, LSTM,
        →TimeDistributed, RepeatVector
       from tensorflow.keras.models import Model
       from tensorflow.keras import regularizers
[166]: # Test for the set n.2
       dataset = pd.read_csv('/home/developer/Documents/Machine Learning/Corteva/
        ⇔dataset/cleaned/dryer_3_fan_1_2020.csv', sep=';', usecols =_
        →['Motor_Frequency_Cmd_Hz','X_Axis_High_Frequency_RMS_Acceleration','X_Axis_RMS_Velocity',

¬'Z_Axis_High_Frequency_RMS_Acceleration', 'Z_Axis_RMS_Velocity'],)
       # dataset.columns =
        \hookrightarrow ['Motor_Frequency_Cmd_Hz','Timestamp','X_Axis_High_Frequency_RMS_Acceleration','X_Axis_RMS_
        → 'Z_Axis_High_Frequency_RMS_Acceleration', 'Z_Axis_RMS_Velocity']
       dataset.head()
[166]:
        Motor_Frequency_Cmd_Hz X_Axis_High_Frequency_RMS_Acceleration \
                           32,8
                                                                  0,232
                                                                  0,233
       1
                           32,8
       2
                           40,2
                                                                  1,047
                           40,2
       3
                                                                  1,155
```

```
4
                           40,2
                                                                   0,773
         X Axis RMS Velocity Z Axis High Frequency RMS Acceleration \
                       0,326
       1
                        0,38
                                                                0,171
                       0,538
                                                                0,511
       2
       3
                       0,572
                                                                0,516
       4
                       0,605
                                                                 0,47
         Z_Axis_RMS_Velocity
       0
                       3.305
       1
                       1,335
       2
                       0,456
                       0,376
       3
                       0,363
[167]: # Convert values with comma into float
       dataset['X_Axis_High_Frequency_RMS_Acceleration'] =__

¬dataset['X_Axis_High_Frequency_RMS_Acceleration'].astype(str).str.

        →replace(',', '.').astype(float)
       dataset['X Axis RMS Velocity'] = dataset['X Axis RMS Velocity'].astype(str).str.
        →replace(',', '.').astype(float)
       dataset['Z Axis High Frequency RMS Acceleration'] = ...
        →dataset['Z Axis_High Frequency_RMS_Acceleration'].astype(str).str.
        →replace(',', '.').astype(float)
       dataset['Z_Axis_RMS_Velocity'] = dataset['Z_Axis_RMS_Velocity'].astype(str).str.

¬replace(',', '.').astype(float)
       dataset['Motor_Frequency_Cmd_Hz'] = dataset['Motor_Frequency_Cmd_Hz'].
        →astype(str).str.replace(',', '.').astype(float)
       dataset
[167]:
             Motor_Frequency_Cmd_Hz X_Axis_High_Frequency_RMS_Acceleration \
                                32.8
                                                                        0.232
       1
                                32.8
                                                                        0.233
       2
                                40.2
                                                                        1.047
       3
                                40.2
                                                                        1.155
       4
                                40.2
                                                                        0.773
       2690
                                44.0
                                                                        1.340
       2691
                                44.0
                                                                        1.441
       2692
                                44.0
                                                                        1.408
       2693
                                44.0
                                                                        1.363
       2694
                                44.0
                                                                        1.353
             X_Axis_RMS_Velocity Z_Axis_High_Frequency_RMS_Acceleration \
       0
                            0.326
                                                                     0.169
       1
                            0.380
                                                                     0.171
```

2	0.538	0.511
3	0.572	0.516
4	0.605	0.470
•••	•••	•••
2690	1.607	1.053
2691	1.786	1.060
2692	1.395	1.047
2693	2.217	1.085
2694	1.213	1.032
	<pre>Z_Axis_RMS_Velocity</pre>	
0	3.305	
1	1.335	
2	0.456	
3	0.376	
4	0.363	

[2695 rows x 5 columns]

2690

2691

2692 2693

2694

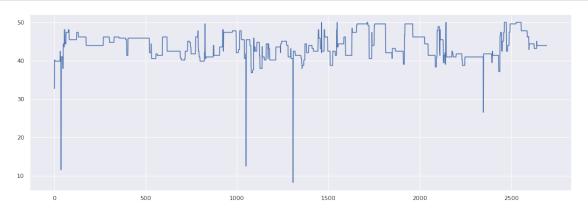
### [168]: dataset['Motor\_Frequency\_Cmd\_Hz'].plot(figsize=(18,6));

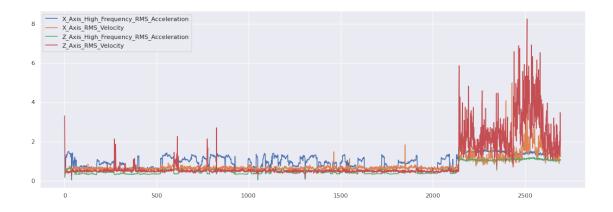
1.264

2.5141.718

3.470

1.954





```
[170]: # Split trainig set and test set

# Training set
training_set = dataset[0:2092]
training_set
```

[170]:	Motor_Frequency_Cmd_Hz	X_Axis_High_Frequency_RMS_Acceleration \
0	32.8	0.232
1	32.8	0.233
2	40.2	1.047
3	40.2	1.155
4	40.2	0.773
•••		•••
2087	38.4	1.252
2088	38.4	1.265
2089	38.4	1.314
2090	38.4	1.301
2091	38.4	1.351
	V A: - DMG U1:+ 7	And High Forence DMG Angelessking
0	•	_Axis_High_Frequency_RMS_Acceleration \
0	0.326	0.169
1	0.380	0.171
2	0.538	0.511
3	0.572	0.516
4	0.605	0.470
•••	•••	•••
2087	0.518	0.688
2088	0.619	0.626
2089	0.566	0.623
2090	0.587	0.609
2091	0.618	0.627

Z\_Axis\_RMS\_Velocity

```
1
                            1.335
       2
                            0.456
       3
                            0.376
       4
                            0.363
       2087
                            0.554
       2088
                            0.505
       2089
                            0.617
       2090
                            0.590
       2091
                            0.582
       [2092 rows x 5 columns]
[171]: # Test set
       test_set = dataset[2093: ]
       test_set
[171]:
             Motor_Frequency_Cmd_Hz X_Axis_High_Frequency_RMS_Acceleration \
       2093
                                41.0
                                                                          0.870
       2094
                                41.0
                                                                          0.895
       2095
                                47.8
                                                                          0.546
       2096
                                47.8
                                                                          0.541
       2097
                                47.8
                                                                          0.546
                                44.0
                                                                          1.340
       2690
                                44.0
       2691
                                                                          1.441
       2692
                                44.0
                                                                          1.408
       2693
                                44.0
                                                                          1.363
       2694
                                44.0
                                                                          1.353
             X_Axis_RMS_Velocity Z_Axis_High_Frequency_RMS_Acceleration \
       2093
                            0.636
                                                                       0.507
       2094
                            0.621
                                                                       0.470
       2095
                            0.746
                                                                       0.369
       2096
                            0.724
                                                                       0.362
       2097
                            0.700
                                                                       0.359
       2690
                                                                       1.053
                            1.607
       2691
                            1.786
                                                                       1.060
       2692
                            1.395
                                                                       1.047
       2693
                                                                       1.085
                            2.217
       2694
                            1.213
                                                                       1.032
             Z_Axis_RMS_Velocity
       2093
                            0.514
       2094
                            0.439
```

0

3.305

```
2095
                           0.537
       2096
                           0.573
       2097
                           0.497
       2690
                           1.264
       2691
                           2.514
       2692
                           1.718
       2693
                           3.470
       2694
                           1.954
       [602 rows x 5 columns]
[174]: # normalize the data
       scaler = MinMaxScaler()
       X_train = scaler.fit_transform(training_set)
       X_test = scaler.transform(test_set)
       scaler_filename = "scaler_data"
       joblib.dump(scaler, scaler_filename)
[174]: ['scaler_data']
[175]: # reshape inputs for LSTM [samples, timesteps, features]
       X_train = X_train.reshape(X_train.shape[0], 1, X_train.shape[1])
       print("Training data shape:", X_train.shape)
       X_test = X_test.reshape(X_test.shape[0], 1, X_test.shape[1])
       print("Test data shape:", X_test.shape)
      Training data shape: (2092, 1, 5)
      Test data shape: (602, 1, 5)
[176]: X_train
[176]: array([[[0.58752998, 0.13655172, 0.08323208, 0.19907407, 1.
                                                                           ]],
              [[0.58752998, 0.13724138, 0.11603888, 0.20216049, 0.37281121]],
              [[0.76498801, 0.69862069, 0.21202916, 0.72685185, 0.09296402]],
              ...,
              [[0.72182254, 0.88275862, 0.2290401, 0.89969136, 0.14422159]],
              [[0.72182254, 0.8737931, 0.2417983, 0.87808642, 0.1356256]],
              [[0.72182254, 0.90827586, 0.26063183, 0.9058642, 0.13307864]]])
```

```
[177]: # define the autoencoder network model
      def autoencoder_model(X):
          inputs = Input(shape=(X.shape[1], X.shape[2]))
          L1 = LSTM(16, activation='relu', return_sequences=True,
                   kernel_regularizer=regularizers.12(0.00))(inputs)
          L2 = LSTM(4, activation='relu', return_sequences=False)(L1)
          L3 = RepeatVector(X.shape[1])(L2)
          L4 = LSTM(4, activation='relu', return_sequences=True)(L3)
          L5 = LSTM(16, activation='relu', return_sequences=True)(L4)
          output = TimeDistributed(Dense(X.shape[2]))(L5)
          model = Model(inputs=inputs, outputs=output)
          return model
[178]: # create the autoencoder model
      model = autoencoder_model(X_train)
      model.compile(optimizer='adam', loss='mae')
      model.summary()
     Model: "model_4"
      Layer (type)
                        Output Shape
                                                        Param #
      ______
                                [(None, 1, 5)]
      input_6 (InputLayer)
      lstm_20 (LSTM)
                               (None, 1, 16)
                                                        1408
      lstm_21 (LSTM)
                                 (None, 4)
                                                         336
      repeat_vector_5 (RepeatVect (None, 1, 4)
      or)
                                (None, 1, 4)
      lstm_22 (LSTM)
                                                         144
      lstm_23 (LSTM)
                                 (None, 1, 16)
                                                         1344
      time_distributed_4 (TimeDis (None, 1, 5)
                                                         85
      tributed)
     Total params: 3,317
     Trainable params: 3,317
     Non-trainable params: 0
[179]: # fit the model to the data
```

```
[179]: # fit the model to the data

nb_epochs = 100

batch_size = 10
```

#### 

```
Epoch 1/100
val loss: 0.0682
Epoch 2/100
val_loss: 0.0603
Epoch 3/100
val_loss: 0.0542
Epoch 4/100
val loss: 0.0438
Epoch 5/100
val_loss: 0.0382
Epoch 6/100
199/199 [============ ] - 1s 7ms/step - loss: 0.0316 -
val loss: 0.0275
Epoch 7/100
val_loss: 0.0250
Epoch 8/100
val_loss: 0.0231
Epoch 9/100
val loss: 0.0218
Epoch 10/100
val_loss: 0.0223
Epoch 11/100
val_loss: 0.0254
Epoch 12/100
val_loss: 0.0229
Epoch 13/100
199/199 [============= ] - 2s 9ms/step - loss: 0.0212 -
val_loss: 0.0218
Epoch 14/100
val loss: 0.0217
Epoch 15/100
199/199 [============ ] - 1s 7ms/step - loss: 0.0214 -
```

```
val_loss: 0.0199
Epoch 16/100
199/199 [=========== ] - 2s 8ms/step - loss: 0.0213 -
val_loss: 0.0222
Epoch 17/100
val loss: 0.0239
Epoch 18/100
val_loss: 0.0225
Epoch 19/100
val_loss: 0.0221
Epoch 20/100
val_loss: 0.0215
Epoch 21/100
199/199 [========== ] - 1s 7ms/step - loss: 0.0213 -
val_loss: 0.0198
Epoch 22/100
val loss: 0.0223
Epoch 23/100
val_loss: 0.0227
Epoch 24/100
val_loss: 0.0208
Epoch 25/100
val_loss: 0.0216
Epoch 26/100
val_loss: 0.0213
Epoch 27/100
val loss: 0.0197
Epoch 28/100
val_loss: 0.0210
Epoch 29/100
199/199 [============ ] - 1s 7ms/step - loss: 0.0211 -
val_loss: 0.0221
Epoch 30/100
val_loss: 0.0210
Epoch 31/100
```

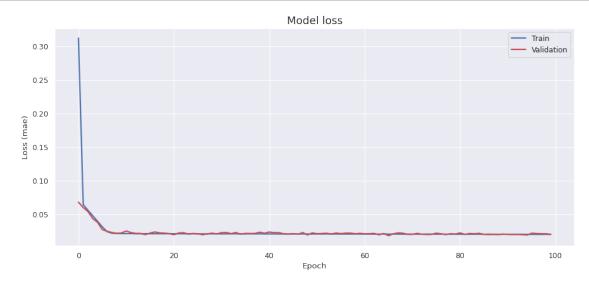
```
val_loss: 0.0228
Epoch 32/100
199/199 [============ ] - 2s 10ms/step - loss: 0.0211 -
val_loss: 0.0230
Epoch 33/100
val loss: 0.0214
Epoch 34/100
val_loss: 0.0231
Epoch 35/100
val_loss: 0.0206
Epoch 36/100
val_loss: 0.0218
Epoch 37/100
val_loss: 0.0217
Epoch 38/100
val loss: 0.0217
Epoch 39/100
val_loss: 0.0235
Epoch 40/100
val_loss: 0.0219
Epoch 41/100
val_loss: 0.0238
Epoch 42/100
val_loss: 0.0228
Epoch 43/100
val loss: 0.0229
Epoch 44/100
val_loss: 0.0213
Epoch 45/100
199/199 [============ ] - 2s 8ms/step - loss: 0.0209 -
val_loss: 0.0206
Epoch 46/100
val_loss: 0.0214
Epoch 47/100
```

```
val_loss: 0.0204
Epoch 48/100
199/199 [========== ] - 2s 8ms/step - loss: 0.0209 -
val_loss: 0.0230
Epoch 49/100
val loss: 0.0194
Epoch 50/100
val_loss: 0.0224
Epoch 51/100
val_loss: 0.0213
Epoch 52/100
val_loss: 0.0216
Epoch 53/100
199/199 [=========== ] - 2s 12ms/step - loss: 0.0210 -
val_loss: 0.0220
Epoch 54/100
val loss: 0.0209
Epoch 55/100
val_loss: 0.0223
Epoch 56/100
val_loss: 0.0215
Epoch 57/100
val_loss: 0.0221
Epoch 58/100
val_loss: 0.0222
Epoch 59/100
val loss: 0.0212
Epoch 60/100
val_loss: 0.0218
Epoch 61/100
199/199 [============ ] - 2s 9ms/step - loss: 0.0207 -
val_loss: 0.0211
Epoch 62/100
val_loss: 0.0214
Epoch 63/100
```

```
val_loss: 0.0217
Epoch 64/100
199/199 [========== ] - 2s 9ms/step - loss: 0.0207 -
val_loss: 0.0196
Epoch 65/100
val loss: 0.0216
Epoch 66/100
val_loss: 0.0183
Epoch 67/100
val_loss: 0.0212
Epoch 68/100
val_loss: 0.0225
Epoch 69/100
199/199 [========== ] - 2s 9ms/step - loss: 0.0206 -
val_loss: 0.0220
Epoch 70/100
val loss: 0.0203
Epoch 71/100
val_loss: 0.0202
Epoch 72/100
val_loss: 0.0218
Epoch 73/100
val_loss: 0.0205
Epoch 74/100
val_loss: 0.0201
Epoch 75/100
val loss: 0.0202
Epoch 76/100
val_loss: 0.0221
Epoch 77/100
199/199 [============ ] - 2s 8ms/step - loss: 0.0206 -
val_loss: 0.0212
Epoch 78/100
val_loss: 0.0199
Epoch 79/100
```

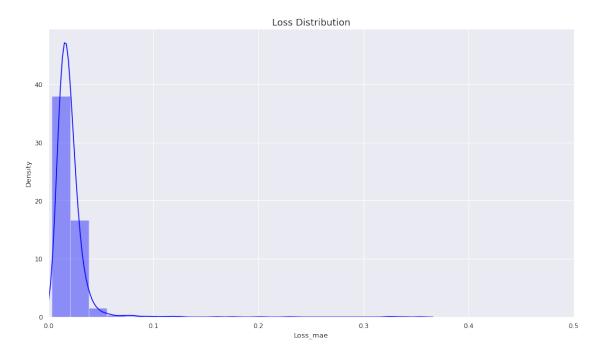
```
val_loss: 0.0213
Epoch 80/100
199/199 [========== ] - 1s 7ms/step - loss: 0.0205 -
val_loss: 0.0208
Epoch 81/100
val loss: 0.0225
Epoch 82/100
val_loss: 0.0203
Epoch 83/100
val_loss: 0.0216
Epoch 84/100
val_loss: 0.0213
Epoch 85/100
199/199 [============ ] - 2s 9ms/step - loss: 0.0205 -
val_loss: 0.0219
Epoch 86/100
val loss: 0.0204
Epoch 87/100
val_loss: 0.0200
Epoch 88/100
val_loss: 0.0203
Epoch 89/100
val_loss: 0.0200
Epoch 90/100
val_loss: 0.0208
Epoch 91/100
val loss: 0.0204
Epoch 92/100
val_loss: 0.0201
Epoch 93/100
199/199 [============= ] - 2s 10ms/step - loss: 0.0204 -
val_loss: 0.0202
Epoch 94/100
val_loss: 0.0199
Epoch 95/100
```

```
val_loss: 0.0193
    Epoch 96/100
    val_loss: 0.0220
    Epoch 97/100
    199/199 [======
                          =======] - 2s 9ms/step - loss: 0.0203 -
    val loss: 0.0216
    Epoch 98/100
    199/199 [=====
                            ======] - 1s 7ms/step - loss: 0.0203 -
    val_loss: 0.0213
    Epoch 99/100
    val_loss: 0.0213
    Epoch 100/100
    val_loss: 0.0205
[180]: # plot the training losses
     fig, ax = plt.subplots(figsize=(14, 6), dpi=80)
     ax.plot(history['loss'], 'b', label='Train', linewidth=2)
     ax.plot(history['val_loss'], 'r', label='Validation', linewidth=2)
     ax.set_title('Model loss', fontsize=16)
     ax.set_ylabel('Loss (mae)')
     ax.set_xlabel('Epoch')
     ax.legend(loc='upper right')
     plt.show()
```



# [211]: # Distribution of Loss Function

#### [211]: (0.0, 0.5)



```
[212]: # calculate the loss on the test set

X_pred = model.predict(X_test)

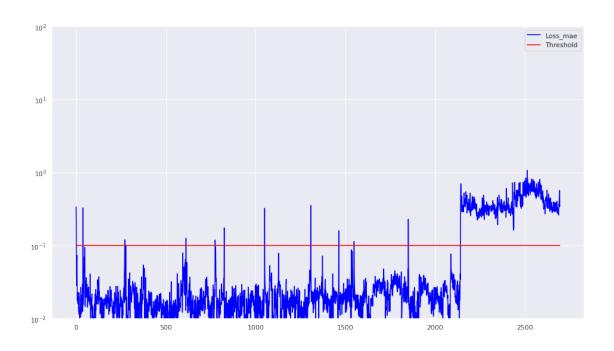
X_pred = X_pred.reshape(X_pred.shape[0], X_pred.shape[2])

X_pred = pd.DataFrame(X_pred, columns=test_set.columns)

X_pred.index = test_set.index
```

```
scored = pd.DataFrame(index=test_set.index)
       Xtest = X_test.reshape(X_test.shape[0], X_test.shape[2])
       scored['Loss_mae'] = np.mean(np.abs(X_pred-Xtest), axis = 1)
       scored['Threshold'] = 0.10
       scored['Anomaly'] = scored['Loss_mae'] > scored['Threshold']
       scored.head()
            Loss_mae Threshold Anomaly
[212]:
       2093 0.042394
                             0.1
                                    False
                                    False
      2094 0.031467
                             0.1
       2095 0.037537
                             0.1
                                   False
       2096 0.034987
                             0.1
                                    False
       2097 0.026520
                             0.1
                                    False
[213]: # calculate the same metrics for the training set
       # and merge all data in a single dataframe for plotting
       X_pred_train = model.predict(X_train)
       X_pred_train = X_pred_train.reshape(X_pred_train.shape[0], X_pred_train.
        \hookrightarrowshape [2])
       X pred_train = pd.DataFrame(X_pred_train, columns=training_set.columns)
       X_pred_train.index = training_set.index
       scored_train = pd.DataFrame(index=training_set.index)
       scored_train['Loss_mae'] = np.mean(np.abs(X_pred_train-Xtrain), axis = 1)
       scored train['Threshold'] = 0.10
       scored_train['Anomaly'] = scored_train['Loss_mae'] > scored_train['Threshold']
       scored = pd.concat([scored_train, scored])
[214]: # plot bearing failure time plot
       scored.plot(logy=True, figsize=(16,9), ylim=[1e-2,1e2], color=['blue','red'])
```

[214]: <AxesSubplot:>



[]: