HUMS2023 Data Challenge Result Submission

Team Name: Accenture-Avanade

Team Members: Behrouz Beheshti, Matthew Thomson

Institutions: Accenture-Avanade

Publishable: Yes

1. Summary of Findings

Table 1 Summary of Analysis Results

#	Detection & Trending	Data file name/number	Comments
1	Consistent detection on at least one signal channel; i.e. the fault indicators remain consistently above the threshold.	Day021_Hunting_SSA_20211208_105651.mat/4	Data reshaped to 4095 X 99 and averaged for 99 columns (the result is going to be an array for each file). Maximum of each array is concatenated into one array. The lower and upper fence on IQR for each sensor is calculated and the outliers detected. The earliest outlier is on the RF2 sensor on the fourth file. Figure 1 & Figure 3
2	Confirmed detection on at least two signal channels; i.e. the fault indicators remain consistently above the threshold.	Day021_Hunting_SSA_20211208_113917.mat/15 Day021_Hunting_SSA_20211208_135820.mat/50 Day022_Hunting_SSA_20211209_124241.mat/125 Day022_Hunting_SSA_20211209_141330.mat/150	Data reshaped to 4095 X 99 and averaged for 99 columns (the result is going to be an array for each file). Maximum of each array is concatenated into one array. The lower and upper fence on IQR for each sensor is calculated and the outliers detected. The earliest outlier on multiple sensors is on IP1,RF2, RL3, and RR4 sensor on the 15th file. Figure 2 & Figure 3 & Figure 4
3	Clear multi-channel indication of the characteristic fault features; i.e. faulty planet gear meshing with both the ring and sun gears.	Day021_Hunting_SSA_20211208_113917.mat/15 Day021_Hunting_SSA_20211208_135820.mat/50 Day022_Hunting_SSA_20211209_141330.mat/150	Data reshaped to 4095 X 99 and averaged for 99 columns (the result is going to be an array for each file). Maximum of each array is concatenated into one array. The lower and upper fence on IQR for each sensor is calculated and the outliers detected. The earliest outlier on multiple sensors is on IP1,RF2, RL3, and RR4 sensor on the 15th , 50th, 125th, and 150th file. Figure 2 & Figure 3 & Figure 4
4	Confirmed trend of fault progression; i.e. a consistent increasing trend started from which file number/name.	Day026_Hunting_SSA_20220114_140718.mat/495	Data reshaped to 4095 X 99 and averaged for 99 columns (the result is going to be an array for each file). Maximum of each array is concatenated into one array. The lower and upper fence on IQR for each sensor is calculated and the outliers detected. The point where the consistent increasing trend start is on RF2 sensor on the 495th file. Figure 5

5	Confirmed trend of accelerated fault progression; i.e. a consistent exponential increasing trend started from which file number/name	Day027_Hunting_SSA_20220118_111018.mat/525	Data reshaped to 4095 X 99 and averaged for 99 columns (the result is going to be an array for each file). Maximum of each array is concatenated into one array. The lower and upper fence on IQR for each sensor is calculated and the outliers detected. The point where the exponential increase starts is o 525th file on all the sensors: IP1, RF2, RL3, RR4. Figure 6

2. Analysis Methods

The analysis is described thoroughly in Section 6.
Below is a brief description of the data preparation and the analysis method:
Step1. 405405 data points are ingested as raw data for each of the 526 files.
Step2. Data is flattened and the values are converted to absolute values.
Step3. Data is reshaped to 4095 X 99-for planet gear and to 11582 X 35-for the ring gear.
Step4. Mean of 99 columns (for a planet gear revolution) and 35 columns (for a ring gear revolution) are calculated and the resulting arrays are kept for next step.
Step5. The maximum of each array (from above) is calculated and the array of 526 values will be shaped to cover all the files.
Step6. The interquartile range (IQR) and the upper and lower quartiles are calculated ana the outliers detected for each channel/sensor.
Step7. Associated filename for each outlier are listed.
Step8. Starting points of consistent increase and the exponential move will be captured from the array in step 5.

3. Illustrating Figures

All the annotated figures are illustrated and explained in Section 4.

4. Characteristic Fault Signatures of Early Detection

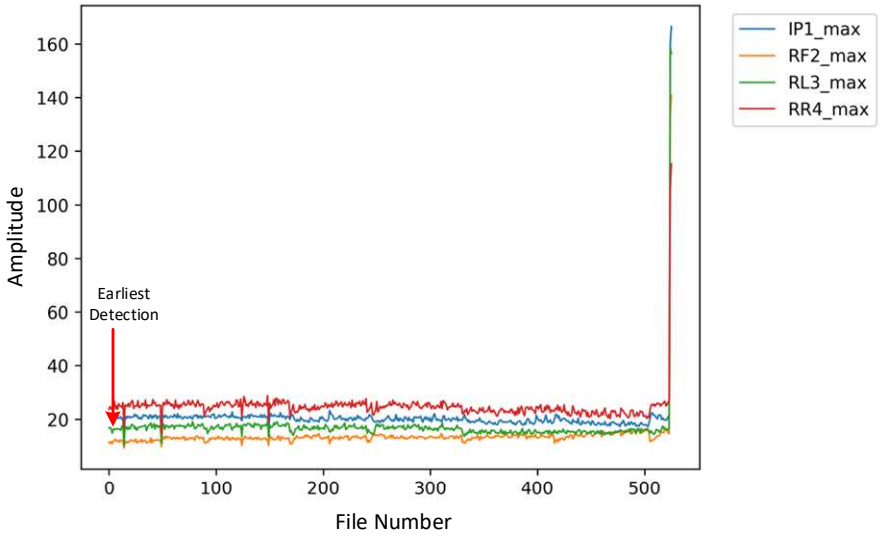

Figure 1. The earliest fault detected; it is on sensor RF2

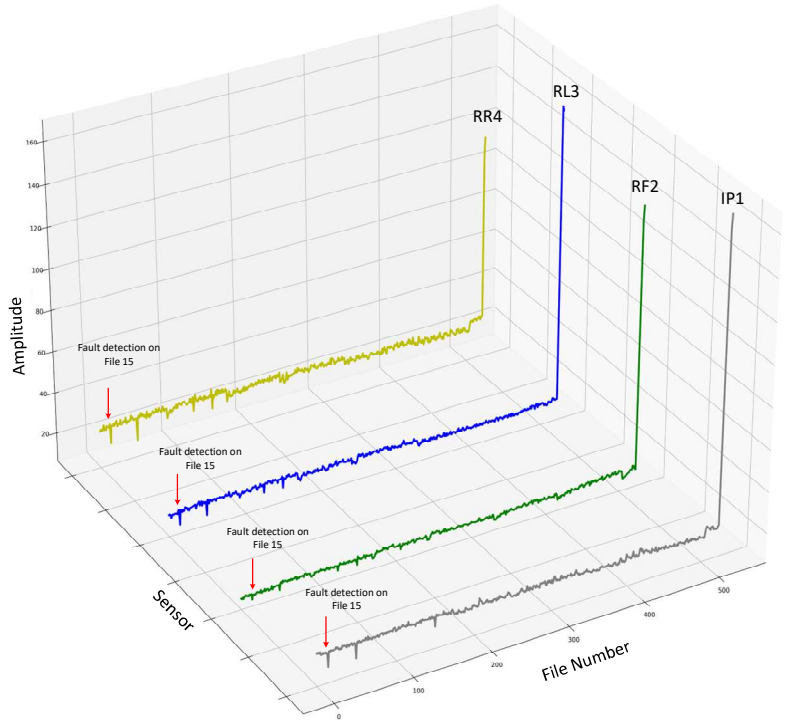

Figure 2. The earliest simultaneous fault detection on multiple sensors IP1, RF2, RL3, RR4

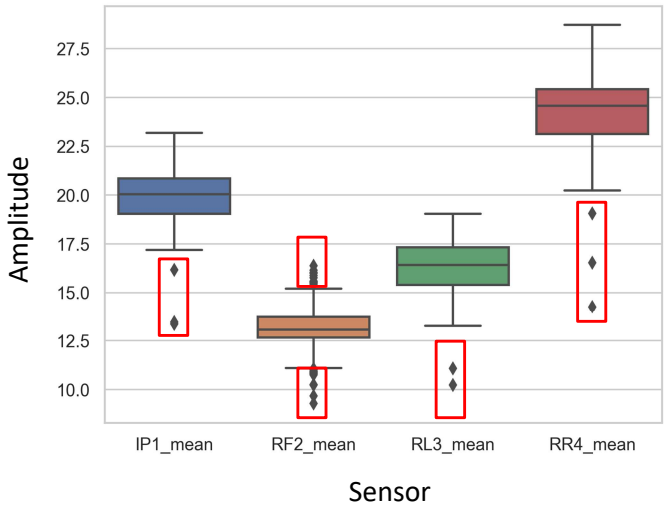

Figure 3. The box plots for the max of mean amplitude values and the outliers for fault detection

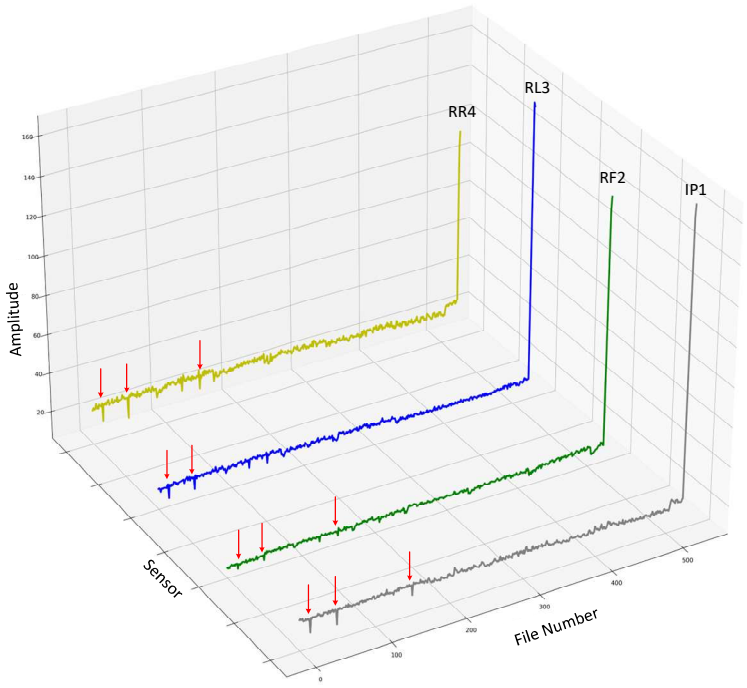

Figure 4. Simultaneous fault detection on multiple sensors

5. Fault Progression Trending Curve

This curve will be compared with the estimated crack growth curve produced by DSTG's fractography analysis.

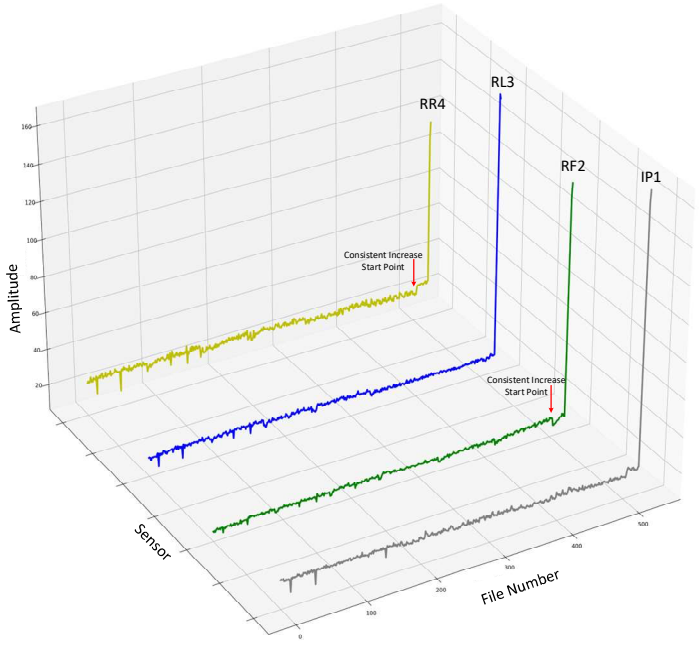

Figure 5. Consistent increase in the fault indicator on RF2 and RR4

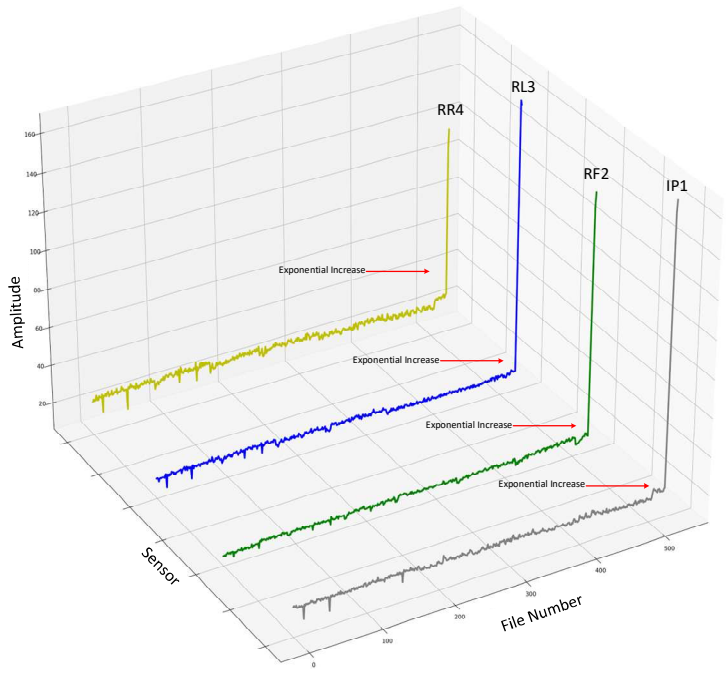

Figure 6. Exponential Increase in all sensors-possible collapse detection

6. Description of Analysis Methods

Input data includes 526 .mat files, each includes [4 channels- one for each sensors X 405405 reading points] from the accelerometers. The data length for each of these files is 6.1 seconds.

Collected data from the lab test on the gearbox and the accelerometer readings were provided as MATLAB files (.mat). The roadmap in Figure 7 shows the data preparation for this analysis:

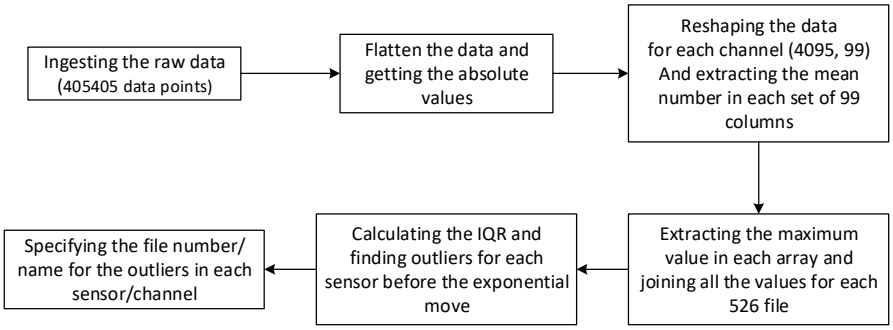

Figure 7. Data pre-processing for the analysis of one channel of one file

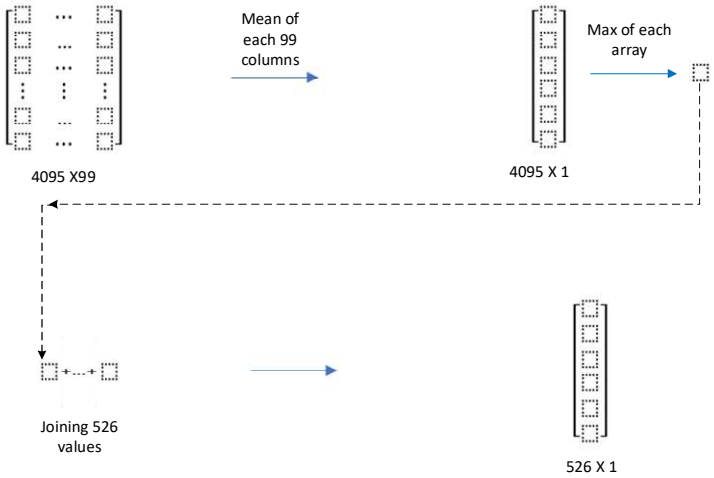

Figure 8. Data preparation for one channel of one file

Figure 8 shows the details of steps 1 to 4 in Figure 7 roadmap.

As shown in this Figure 8, the data is first merged into a 3D matrix (for 526 files and 405405 readings X 4 channels). Then it is divided into 4 sections, for each channel. The absolute value of each 2D matrix of each channel is then kept for the next steps. The data is flattened and then reshaped to (4095 X 99) for the planet gear analysis. Getting the mean of 99 column (one planet gear revolution) results in a 4095 X 1 array. Figure 9 shows the timeseries graph for the whole data and for all channels. The maximum value of each 4095 X 1 array is extracted for each 526 file which results in a final 526 X 1 array for each channel/sensor. Figure 10 shows the max of revolution mean for each file, for all channels.

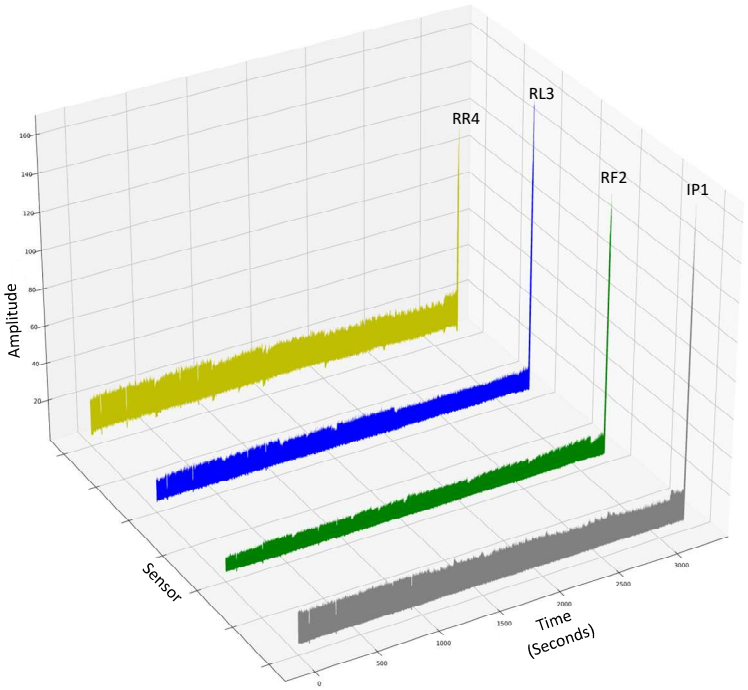

Figure 9. Timeseries graph for the mean of planet gear revolutions

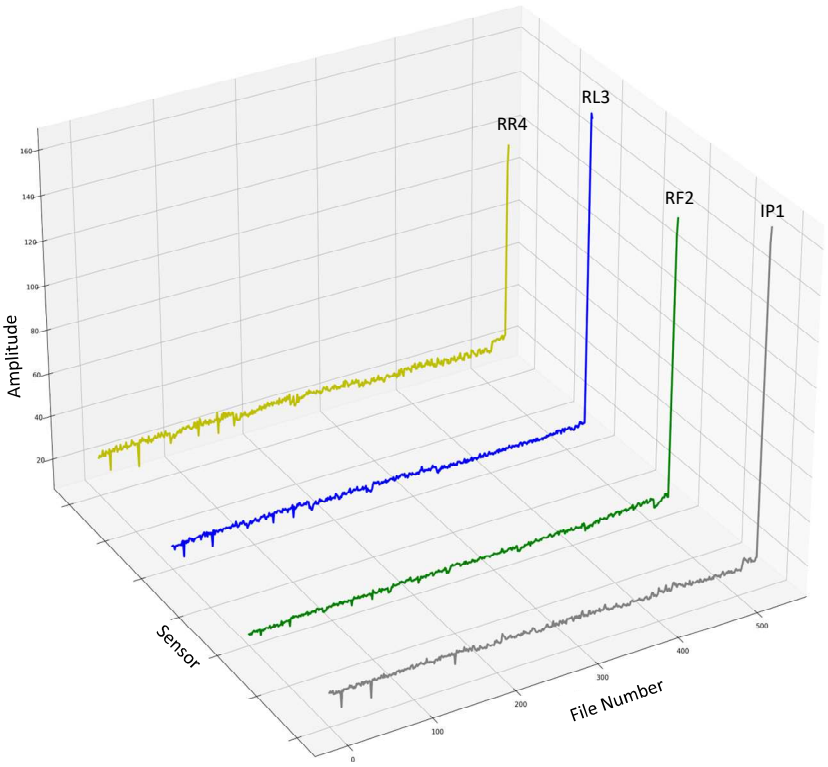

Figure 10. Max of revolution means of each file

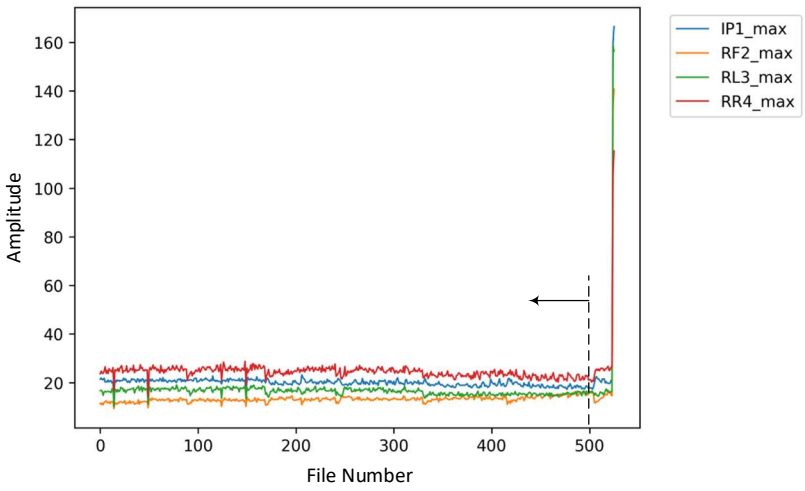

Figure 11. Flat view of the overlayed graphs of the max of revolution means for all channels

Figure 11 shows the flat view of the overlayed graphs of the max of revolution means for all channels. From this graph, the sharp increase (exponential move) in the amplitude happens after file 500. To have a clear picture of the data distribution-without being affected by extreme values after file 500- the final dataset is split between two parts: data before file 500 and after 500. The Interquartile range (IQR), as well as the lower and upper quartiles are calculated for the files before file 500 and these data, along with the

specified outliers for each channel/sensor is used for the analysis in the next steps. Figure 12 illustrates the result for each sensor.

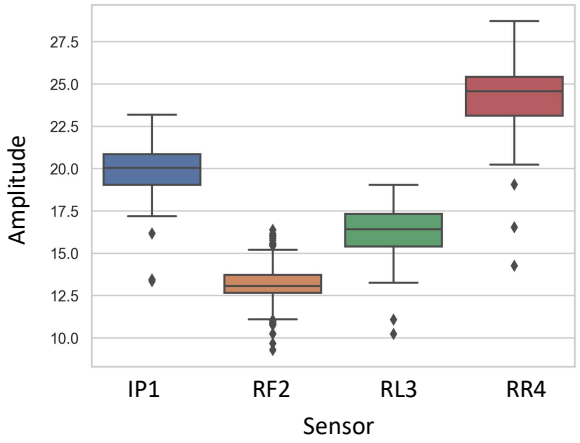

Figure 12. Box plots of the max of revolution means for all channels before file 500

Description of fault detection method

Figure 12 illustrates the density plot view on box plots along with the outliers. The outliers on each channel/sensor reflects the amplitudes on each channel where the fault is located on the specific channel. By cross-referencing these outlier values with the filename-associated with the outlier amplitude-the specific filename can be detected. In other words, the algorithm will go through the list of filenames and their max of revolution mean, and extract the filenames associated with the outlier value.

The first fault is detected solely on the RF2 sensor on the fourth file-as shown on Figure 1 . The next fault on the 15th file is detected simultaneously on all sensors-as shown on Figure 2.

Description of fault trending method

The algorithm will proceed with looking at the total values and point to where the amplitude of the max of the revolution mean reaches to a point where there is a consistent increase in it. This value can be different for each sensor, but the similar fact for all the cases is that this phenomenon happens after 500th file. The exponential change in the max of revolution mean is then visible from Figure 10.

7. Supplement Information

The method described above uses the planet gear by reshaping the data to 4095 X 99 for each channel on each of the 526 files. The same method was used by reshaping the data to 11852 X 35-to cover the ring gear-and the same result (i.e., file number of earliest fault detection, simultaneous fault detection, and consistent increase and exponential change) were achieved.