DETECTION OF AIR LEAK IN EXHAUST SYSTEM USING ON-THE-ROAD DATA

Acerta Case Study

Acerta
Intelligent Insight From Machine Data

OBJECTIVES

- Prove failure prediction feasibility while lacking domain expertise
- Identify trends indicative of future failures

CHALLENGE

- 250MB of data
- 350 input parameters
- Real-time analysis

RESULTS

- Reduced root-cause analysis time from ~2 *weeks* to 1 *hour*
- Produced algorithm for real-time on-the-road failure prediction
- Precisely identified the 0.029% anomalous data

CONTACT ACERTA

- info@acerta.ca
- +1 (519) 729-5543
- www.acerta.ca
- @AcertaAnalytics
- Acerta Analytics Solutions
- 151 Charles Street, Waterloo, ON, Canada

BACKGROUND

Anomaly detection is currently performed manually, as factory engineers compare collected data series side-by-side and attempt to identify anything out of the ordinary. Thresholds for signals are determined based on complex, and often over-simplified, physical models. When a signal exceeds its threshold, a test failure is triggered. However, malfunctions are not the only cause for changes in signal behavior - it might also change as a result of the driver's actions or various environmental factors. Therefore, it is necessary to analyze not only the signals themselves but also the relationships between them (for example, acceptable engine starting time is dependent on ambient temperature). This makes the process of setting up thresholds very engineering-intensive and time-consuming. When analyzing hundreds of recorded signals, machine learning algorithms potentially have a significant advantage over engineers.

THE PROBLEM

A leading automotive manufacturer requested that Acerta analyze data collected during an on-the-road qualification test for one of its vehicles. As the test driver was operating the vehicle, he noticed unusual behavior which kept worsening with time until ultimately, he had to turn off the vehicle. The client provided 250MB of data recorded from 350 sensors during 80 hours of driving. The objective of the project was to detect any anomaly in vehicle operation by using only the data provided. No information was given regarding the type or model of the vehicle, nor regarding the nature of the anomaly or if one existed in the data at all.

SOLUTION PROCESS

Acerta's team analyzed the data and configured the most suitable machine learning algorithm for the type of data provided. With the assumption that the anomaly must have occurred toward the end of the test drive, the first half of the data was used to train the algorithm on the normal operation of the vehicle's systems. Once trained, the algorithm scanned the entire dataset, and was able to detect an anomaly near the end of the recording, in data generated by 8 out of 350 sensors installed in the car. The algorithm identified an anomalous correlation between the signals representing the oxygen level in the exhaust system, the fuel banks, the engine speed and the cylinder pressure. Looking at the complex relationships between these signals, it is apparent why it was difficult to manually identify the anomaly.

RESULTS

Out of the 281863 rows of data recorded from all of the sensors, the algorithm pointed at 3573 rows as representing an anomaly (the anomaly appears in 0.029% of the data). The results were reported to the client who later confirmed that the algorithm had correctly detected over 98% of the anomalous data. Using the information provided by the algorithm, a combustion engine expert was able to identify the root cause of the problem within 60 minutes. It was determined that the fault originated in an air leak in the exhaust system, which caused the oxygen sensors to spike. This in turn caused the lambda control system to increase the air-fuel ratio in order to compensate for the loss of generated energy.
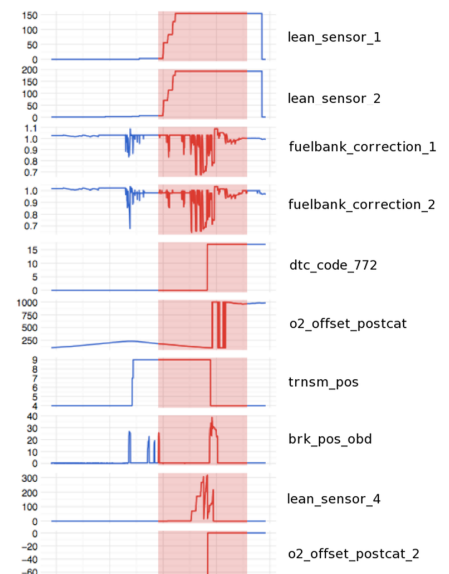
Figure: the portions of the signals which were classified by the algorithm as anomalous (marked in red)