

# ANOMALY DETECTION IN BIOSENSOR WAVEFORMS

## EXECUTIVE SUMMARY

SARA HALL, JUSTINE FILION, SAISREE GR, NEETHU GOPALAKRISHNA

Siemens Healthineers specializes in manufacturing medical diagnostic equipment, an example of which is the epoc® blood system, which is used to perform comprehensive blood analyses. One of the main components of this system is a test card which contains the biosensors necessary to perform the testing for a variety of analytes contained in blood samples. To measure these analytes, the test card is inserted into the reader portion of the system, the sample is injected into the test card and the results are reported by the host.

Before being shipped to customers for use, these test cards must undergo comprehensive quality control testing. If unsuccessful results are yielded during this process, an operator may manually inspect the raw data, which includes time-series signals, and the test cards in order to determine the root cause of failure. However, this is time consuming, and misclassification may occur due to subjectivity in manual review. As such, this project aimed to create an unsupervised machine learning pipeline that characterized these unsuccessful readings into different groups so that their occurrences over time can be tracked and tied back to manufacturing. This tool will help identify possible root causes of failure which can then be addressed, improving the manufacturing process.

One of the root causes that has been investigated occurs when an electrical circuit disruption (ECD) occurs during the testing of the card. Having the labels for these anomalies allowed us to evaluate whether or not our unsupervised pipelines were able to form clusters that could potentially be tied back to root causes. Thus, we used these labels to diagnose whether or not our pipelines were successful.

Before we could apply any machine learning models to cluster the signals, we had to pre-process them. Pre-processing included aligning the signals with respect to a semantically consistent time point, making sure all the signals were of the same length, and performing noise reduction. Various machine learning models, including K-Means, K-Shape and Gaussian Mixture Modelling were then used to cluster the unsuccessful readings. We obtained four promising pipelines that were successful to varying degrees in clustering the ECD errors from other unsuccessful readings. Furthermore, we observed distinguishable shapes across the different clusters which could be used in the future to identify and track potential root causes of failure.

To conclude, we successfully developed four pipelines that led to clusters with a high concentration of ECD errors and other clusters with characteristic signal shapes. This will be useful to Siemens Healthineers for helping identify root causes of failure and improve test card manufacturing.