

Neural Wave Hackathon

Robert Bosch GmbH
Renningen
70465 Stuttgart
Besucher:
Robert-Bosch-Campus 1
71272 Renningen
Telefon +49 711 811-0
www.bosch.com

Jonas Bergdolt, CR/All2
Telefax +49 172 516 2921
Jonas.Bergdolt@de.bosch.com

Project Definition

30. September 2024

1 Motivation

The Bosch Group is a leading global supplier of technology and services with roughly 429,000 associates worldwide. Operating in the mobility, industrial technology, consumer goods, energy and building technology business sectors, manufacturing plays a key role for the company.

One major goal in manufacturing is reducing scrap costs. To achieve this, knowing the causes of scrap is crucial. The goal of this project is to develop an AI solution which determines the root causes of scrap.

2 Problem statement

Imagine the following situation: you are data scientists in a manufacturing plant and suddenly the quality of the produced products decreases leading to higher scrap rates. Factory workers ask you to find the root causes of this change. To accomplish this, they transfer two datasets to you: each contains measurements of 2500 parts, one time before the decrease in quality and one time after it.

3 Explanation of data

Two files:

- Low_scrap.csv (before decrease in quality)
- High_scrap.csv (after decrease in quality)

Each row contains the measurements for one part. Each column contains the measurement values of one specific physical quantity, e.g., *Station2_mp_9*, which is a physical quantity measured at station 2.

The measurements at the stations are taken subsequently. Hence, measurements at station 2 are taken before measurements at station 3. This means that the physical quantities measured at station 3 cannot cause the physical quantities measured at station 2. Vice versa, it might be possible, although does not have to be.

The target parameter we are interested in is called *Station5_mp_85*, whose high values reduce quality and make scrap more likely.

4 Potential solution

The solution should contain:

- A methodology to determine a graph containing the causal relationships between the physical quantities measured at the stations
- A way to detect the root causes of the observed change in *Station5_mp_85* and rank them according to their influence
- A visualization of the graph, the detected root causes and their influence

5 Additional information

No access to company resources required. Some helpful keywords to start are causal discovery, causal inference, causal contributions and explainable AI.

During the hackathon there is the possibility to get your current graph compared to the ground truth. This can only be done 2 times per group. To do this, save the adjacency matrix of your current directed graph as txt file. A “1” indicates an edge from the respective node in the row to the respective node in the column. A “0” denotes the absence of such an edge. The columns need to be ordered in the same way as in the data set.

Send the txt file to Jonas.Bergdolt@de.bosch.com. You will receive a score telling you how close you are to the ground truth and additional information about ground truth relationships in parts of the graph in return.

Introductions to the topics:

- <https://medium.com/causality-in-data-science/introducing-conditional-independence-and-causal-discovery-77919db6159c>

- <https://medium.com/data-science-at-microsoft/beginners-guide-to-causal-discovery-the-what-the-why-and-the-how-65182e97876f>
- https://www.its.caltech.edu/~fehardt/papers/Eberhardt_IJDSA2017.pdf
- <https://arxiv.org/pdf/2407.08602>

30. September 2024

Seite 3 von 3

Libraries:

- <https://github.com/py-why>
- <https://github.com/FenTechSolutions/CausalDiscoveryToolbox>
- <https://github.com/huawei-noah/trustworthyAI/tree/master/gcastle>