

CAUSALITY DISCOVERING

Mykola Rabchevskiy

There is reason to believe that most **Artificial Intelligence** developers are confident that statistical data analysis (more specifically, analysis of the correlation between observed values) is an adequate tool for discovering causal relationships.

In this regard, the idea arose to **invite readers to find causal relationships** in the two cases represented by the respective datasets and then discuss the results.

AGI engineering is a reader-supported publication. To receive new posts and support my work, consider becoming a free or paid subscriber.

The datasets are small enough to be imported into **Excel** or its equivalent: a thousand samples of four observed factors. The goal of the analysis is **to discover causal relationships between these factors**.

Correlation coefficients between the factors of the first data set:

1 2 3 4

1 1.0000 0.5889 -0.0386 0.0396

2 0.5889 1.0000 0.3504 0.4965

3 -0.0386 0.3504 1.0000 -0.0021

4 0.0396 0.4965 -0.0021 1.0000

Correlation coefficients between the factors of the second data set:

1 2 3 4

1 1.0000 0.9449 0.0299 0.0226

2 0.9449 1.0000 0.0480 0.1369

3 0.0299 0.0480 1.0000 -0.0523

4 0.0226 0.1369 -0.0523 1.0000

Original datasets (comma-separated format): ***data1.csv*** and ***data2.csv*** at <https://github.com/mrabchevskiy/dataset>

After two or three weeks, the results will be the subject of discussion.

We are waiting for your thoughts, decisions, and comments!

AGI engineering is a reader-supported publication. To receive new posts and support my work, consider becoming a free or paid subscriber.