TUM

Seminar

Explainable / Interpretable AI

In contrast to the continued rise of various deep neural network ML variants, we favor the properties that come with explicit knowledge LCS (Learning Classifier System) -type ML techniques in our IPF project: interpretable and explainable rules which also can be pre-initialized with best-practice prior knowledge content.

This Seminar should provide an overview of existing, explainable ML techniques and of methods to make black-box techniques (NNs) explainable.

Advisors

Florian Maurer