

AISSAI Anomaly Detection Workshop



ID de Contribution: 47

Type: Non spécifié

Multiview Symbolic Regression. How to learn laws from examples

Feature extraction is one of the crucial stages in trying to apply machine learning to real scientific cases. Because of their potentially non homogeneous sampling, feature extraction of time series represents an additional challenge. One of the standard methods to tackle it requires to use a parametric equation which should be versatile enough to describe all the diversity of your dataset. This function is then used to fit each time series and the best parameter values are used as features. The quality of the parametric function determines how much of the data's behavior is encoded inside the features and thus determines how suitable they are for machine learning applications. In particular in the case of anomaly detection it is essential to use the best possible function since small details can make a significant difference between a normal and an anomalous event. In this work we propose a solution to automatically discover optimal parametric functions for a given problem using an adaptation of Symbolic Regression. It is capable of recovering a common parametric equation hidden behind multiple datasets generated using different parameter values. We call this approach Multiview Symbolic Regression (MvSR). I will highlight the potential of MvSR for feature extraction by demonstrating its efficiency on a variety of real scientific datasets. The resulting parametric equations are able to correctly describe the examples from which they were built as well as other unseen similar examples. Applying MvSR on a science case will unlock its optimal parameterization for future anomaly detection pipelines, thus improving chances of future great discoveries

Orateur: RUSSEIL, Etienne

Classification de Session: Contributed