Process Mapping

BigDataStack Software Component developed by UPRC

Description

The Process Mapping component targets the problem of selecting the best algorithm along with a set of values for the algorithm's input parameters, from a set of candidate algorithms, given a specific data analysis task, in an automatic way. Its role is to automatically map a step of a process to a specific algorithmic instance from a given pool of algorithms, thereby achieving so-called "process mapping".

Features

• Data pre-processing.
• Feature extraction from datasets.
• Algorithm selection.
• Hyperparameter tuning.

Areas of Application

Automated Machine Learning.

Process Mapping

Market trends & opportunities

Existing frameworks for AutoML (research prototypes or systems: AutoML, AutoWeka, TPOT, Google's Vizier) mainly provide solutions in the context of supervised learning, such as classification and regression but none of those frameworks provides a solution for the problem of Automated Machine Learning in an unsupervised context.

Customer benefits

The Process Mapping delivers benefits for Data analysts, Data scientists and Business analysts. In all the above cases, the analyst/scientist can gain significant time when applying machine learning techniques to new datasets. The reason is that the Process Mapping component automatically selects the best algorithm and its parameterization, thus eliminating the need for testing various algorithms and configurations manually, until acceptable performance is observed.

Technological novelty

Solving the problem of AutoML for unsupervised learning (clustering) is considered as the main technological novelty addressed by the Process Mapping component.

TRL level: 4

Find the Open Source code here:

- http://bigdatastack-tasks.ds.unipi.gr/gpoul/process_mapping

bigdatastack.eu