

Proceedings of Workshop

AutoML 2017 @ ECML-PKDD:

**Automatic selection, configuration and composition of
machine learning algorithms**

associated with ECML PKDD 2017 conference

Main Aims of the AutoML 2017 Workshop

The aim of this workshop is to provide a platform for discussing recent developments in the areas of meta-learning, algorithm selection and configuration, which arise in many diverse domains and are increasingly relevant today. Researchers and practitioners from all areas of science and technology face a large choice of parameterized machine learning algorithms, with little guidance as to which techniques to use in a given application context. Moreover, data mining challenges frequently remind us that algorithm selection and configuration are crucial in order to achieve cutting-edge performance, and drive industrial applications. Meta-learning leverages knowledge of past algorithm applications to select the best techniques for future applications, and offers effective techniques that are superior to humans both in terms of the end result and especially in the time required to achieve it.

This workshop provides a platform for discussing different ways of exploiting meta-learning techniques to identify the potentially best algorithm(s) for a new task, based on meta-level information, including prior experiments on both past datasets and the current one. Many contemporary problems also require the use of complex workflows that consist of several processes or operations. Constructing such complex workflows requires extensive expertise, and could be greatly facilitated by leveraging planning, meta-learning and intelligent system design. This task is inherently interdisciplinary, as it builds on expertise in various areas of AI.

Main Research Areas

The main research areas of relevance that were included in the *Call for papers* were:

- Algorithm / model selection and configuration
- Meta-learning and exploitation of meta-knowledge
- Hyperparameter optimization
- Automatic generation and evaluation of learning processes / workflows
- Representation learning and automatic feature extraction / construction
- Automatic feature coding / transformation
- Automatic detection and handling skewed data or missing values
- Automatic acquisition of new data (active learning, experimental design)
- Usage of planners in the construction of workflows
- Reinforcement learning for parameter control & algorithm design
- Representation of learning goals and states in learning
- Control and coordination of learning processes
- Meta-reasoning
- Layered learning
- Multi-task and transfer learning
- Learning to learn
- Intelligent experiment design

Organization

Program Co-chairs

Frank Hutter, Univ. of Freiburg, Germany

Holger Hoos, Univ. of Leiden, The Netherlands

Pavel Brazdil, Univ. of Porto / LIAAD InescTec, Portugal

Joaquin Vanschoren, University of Eindhoven, The Netherlands

Program committee

Pavel Brazdil

Marcin Budka

André C. P. Carvalho

Claudia Diamantini

Katharina Eggenberger

Matthias Feurer

Bogdan Gabrys

Roman Garnett

Christophe Giraud-Carrier

Isabelle Guyon

Lisheng Sun

Krzysztof Grabczewski

Holger Hoos, U. Leiden

Frank Hutter

Jörg-Uwe Kietz

Lars Kotthoff

Bernhard Pfahringer

Vid Podpecan

Ricardo Prudêncio

Jan van Rijn

Michael Smith

Carlos Soares

Kate Smith-Miles

Joaquin Vanschoren

Ricardo Vilalta

Program

The event is collocated with ECML-PKDD 2017 and is scheduled for Friday, Sept. 22, 2017. It includes a tutorial in the morning session and a workshop in the afternoon session.

Tutorial

9:00 - 10:30 Session 1 (90 min)

Introduction

Blackbox hyperparameter optimization and automated machine learning

Algorithm configuration

Meta-Learning for Algorithm Selection

10:30 - 10:50 Coffee break (20 min)

10:50 - 12:40 Session 2 (110 min)

Beyond blackbox optimization

Automating workflow design

AutoML systems and applications

The big picture: machine learning for automated algorithm design

Presenters: Pavel Brazdil, Joaquin Vanschoren, Frank Hutter, Holger Hoos

12:40 - 14:00 Lunch break (1h20 min.)

Workshop

14:00 - 15:40 Session 1 (100 min.)

Welcome (5 min.)

Invited Talk by Michele Sebag (30 +10 min.):

Stochastic Gradient Descent: Going as Fast as Possible but not Faster

Poster spotlights 1 (15 min.)

Poster session 1 (40 min.)

15:40 - 16:00 Coffee break (20 min.)

16:00 - 17:40 Session 2 (100 min.)

G Verbruggen and L de Raedt:

Towards Automated Relational Data Wrangling (20+5 min)

G Contardo, L Denoyer and T Artieres:

A Meta-Learning Approach to One-Step Active-Learning (20+5 min)

Poster spotlights 2 (15 min)

Poster session 2 (35 min.)

17:40-18:40 Session 3 (Extra session) (60 min.)

Poster session 3 (25 min.)

Panel & Commentaries on some papers & research lines (30 min.)

Closing (5 min)

The sites of AutoML 2017 @ ECML PKDD 2017

At ECML PKDD 2017 conference site:

<http://ecmlpkdd2017.ijs.si/program.html#Workshops>

At AutoML 2017 workshop site:

<https://sites.google.com/site/automl2017ecmlpkdd/home>

Related Workshops in recent years:

- AutoML@ICML 2017, Sydney, Australia, <http://icml2017.automl.org> (a sister workshop)
- AutoML@ICML 2016, <http://icml2016.automl.org>
- Algorithm Configuration at AAI 2016
- AutoML@ICML 2015, <http://icml2015.automl.org>
- MetaSel-2015, Meta-learning & Algorithm Selection 2015 at ECML/PKDD 2015, metasel2014.inescporto.pt
- MetaSel-2014, Meta-learning & Algorithm Selection 2014, at ECAI 2014, metasel2014.inescporto.pt
- AutoML@ICML 2014, <http://icml2014.automl.org>
- Algorithm Configuration and Selection at AAI 2013

Related resources:

- Hyperparameter optimization library, automl.org/hpolib:
- Analytics, Big Data, Data Mining, & Data Science Resources , www.KDnuggets.com
- OpenML, An open science platform for machine learning, <https://www.openml.org/>

Contents

Invited talk

- Michele Sebag:
Stochastic Gradient Descent: Going as Fast as Possible but not Faster (abstract)

List of papers

- Martin Wistuba:
Bayesian Optimization Combined with Successive Halving for Neural Network Architecture Optimization
- Gust Verbruggen and Luc De Raedt:
Towards Automated Relational Data Wrangling
- Fábio Pinto, Vitor Cerqueira, Carlos Soares and João Mendes-Moreira:
autoBagging: Learning to Rank Bagging Workflows with Metalearning
- Gabriella Contardo, Ludovic Denoyer and Thierry Artieres:
A Meta-Learning Approach to One-Step Active-Learning
- Roxana Istrate, Cristiano Malossi, Costas Bekas and Dimitrios Nikolopoulos:
Incremental Training of Deep Convolutional Neural Networks
- Pieter Gijssbers, Joaquin Vanschoren and Randal Olson:
Layered TPOT: Speeding up Tree-based Pipeline Optimization'
- Miguel Cachada, Salisu Abdulrahman and Pavel Brazdil:
Combining Feature and Algorithm Hyperparameter Selection using some Metalearning Methods
- Silvia Nunes Das Dôres, Carlos Soares and Duncan Ruiz:
Effect of Metalearning on Feature Selection Employment
- Jan N. van Rijn and Frank Hutter:
An Empirical Study of Hyperparameter Importance Across Datasets

Late-breaking papers

- Sergey Muravyov and Andrey Filchenkov:
Meta-learning system for automated clustering
- Alexey Zabashta and Andrey Filchenkov:
NDSE: Instance Generation for Classification by Given Meta-Feature Description