

INRAE



université
PARIS-SACLAY

➤ Hyperparameter optimization and AutoML

Alberto TONDA

*UMR 518 MIA-PS, INRAE, AgroParisTech, Université Paris-Saclay
UAR 3611, Institut des Systèmes Complexes de Paris Île-de-France*

➤ Outline

- Hyperparameter optimization
- Common strategies
- Advanced strategies
- AutoML
- Auto...optimization?



➤ Hyperparameter optimization

- ML algorithms have **hyperparameters**
 - Number of trees in an ensemble
 - Number of layers in a neural network
 - Type of kernel in a support vector machine
 - ...
- Values of hyperparameters have influence on performance
- How do we select the best values?

➤ Hyperparameter optimization

- Let's frame it as an optimization problem
 - **Mixed** integer/floating point/categorical
 - **Candidate solution**: set of hyperparameter values
 - **Objective function**: performance (in a cross-validation)
 - Objective function is **expensive** to evaluate
- What types of optimization algorithms could work here?

➤ Common strategies

- Grid (exhaustive) search
 - Select a few meaningful values of parameters
 - Try all possible combinations
 - *sklearn.model_selection.GridSearchCV*
- Random search
 - Lists for discrete parameters and distributions for continuous
 - Makes a user-specified number of attempts and returns best
 - *sklearn.model_selection.RandomizedSearchCV*

➤ Common strategies

- Surrogate models
 - Evaluate candidate solutions on a few samples
 - Remove the ones that are clearly among the worst
 - *sklearn.model_selection.HalvingGridSearchCV*
 - *sklearn.model_selection.HalvingRandomSearchCV*
- Luckily, **some other people** know optimization!

➤ Advanced strategies

- Bayesian Optimization!
- Surrogate models!

scikit-optimize

Sequential model-based optimization in Python

[Getting Started](#)

[What's New in 0.8.1](#)

[GitHub](#)

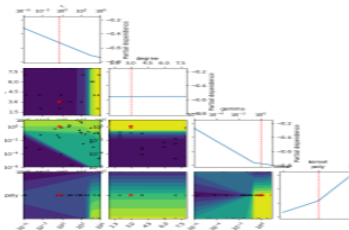
- Sequential model-based optimization
- Built on NumPy, SciPy, and Scikit-Learn
- Open source, commercially usable - BSD license

BayesSearchCV

Scikit-learn hyperparameter search wrapper.

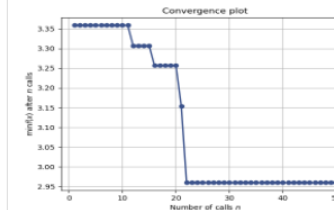
Search for parameters of machine learning models that result in best cross-validation performance

Algorithms: BayesSearchCV



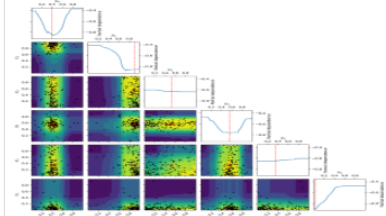
Tuning

Tuning a scikit-learn estimator with skopt



Visualizing

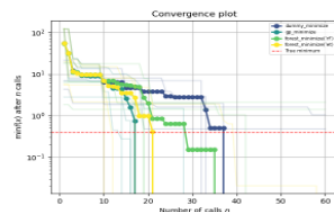
Visualizing optimization results



Comparing surrogate models

Comparing surrogate models

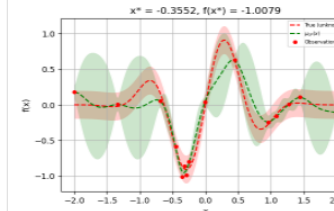
Algorithms: dummy_minimize, gp_minimize, forest_minimize



Bayesian optimization

Bayesian optimization with skopt

Algorithms: gp_minimize



➤ Advanced strategies

- Can we do better?
- Here we are selecting best hyperparameters for a model...
- ...but we already chose a specific model/algorithm
- Not to mention, preprocessing!
 - Dimensionality reduction
 - Feature construction (e.g. polynomial features)
 - Normalization (several different possibilities)
- Can we optimize **all this**? If so, what kind of representation?



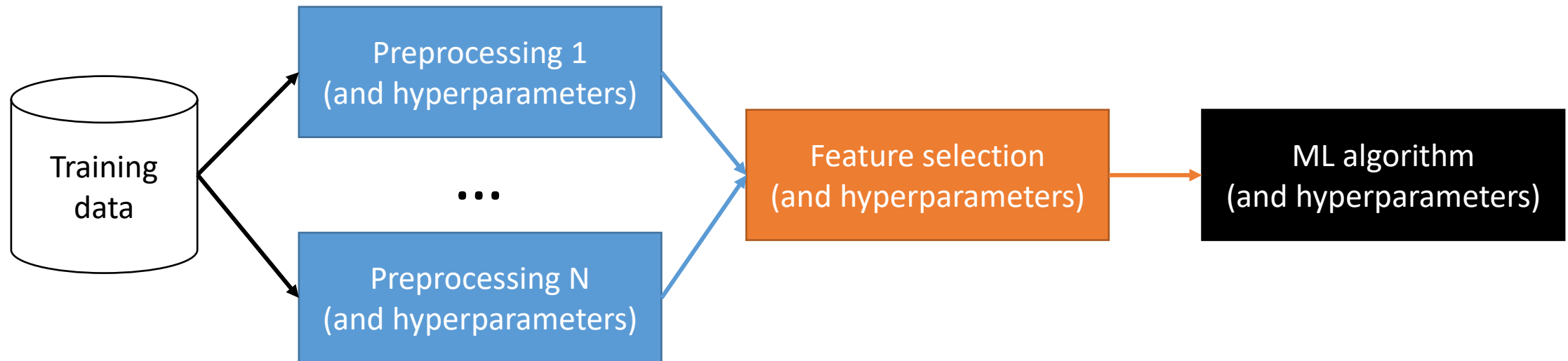
➤ AutoML

- The idea of optimizing *all choices*
 - Candidate solution: preprocessing, algorithm, hyperparameters...
 - Objective function: performance in a cross-validation
- Several different approaches, budding field
- Example: TPOT (and TPOT-2)
 - Tree-based Pipeline Optimization Tool
 - <https://epistasislab.github.io/tpot/>



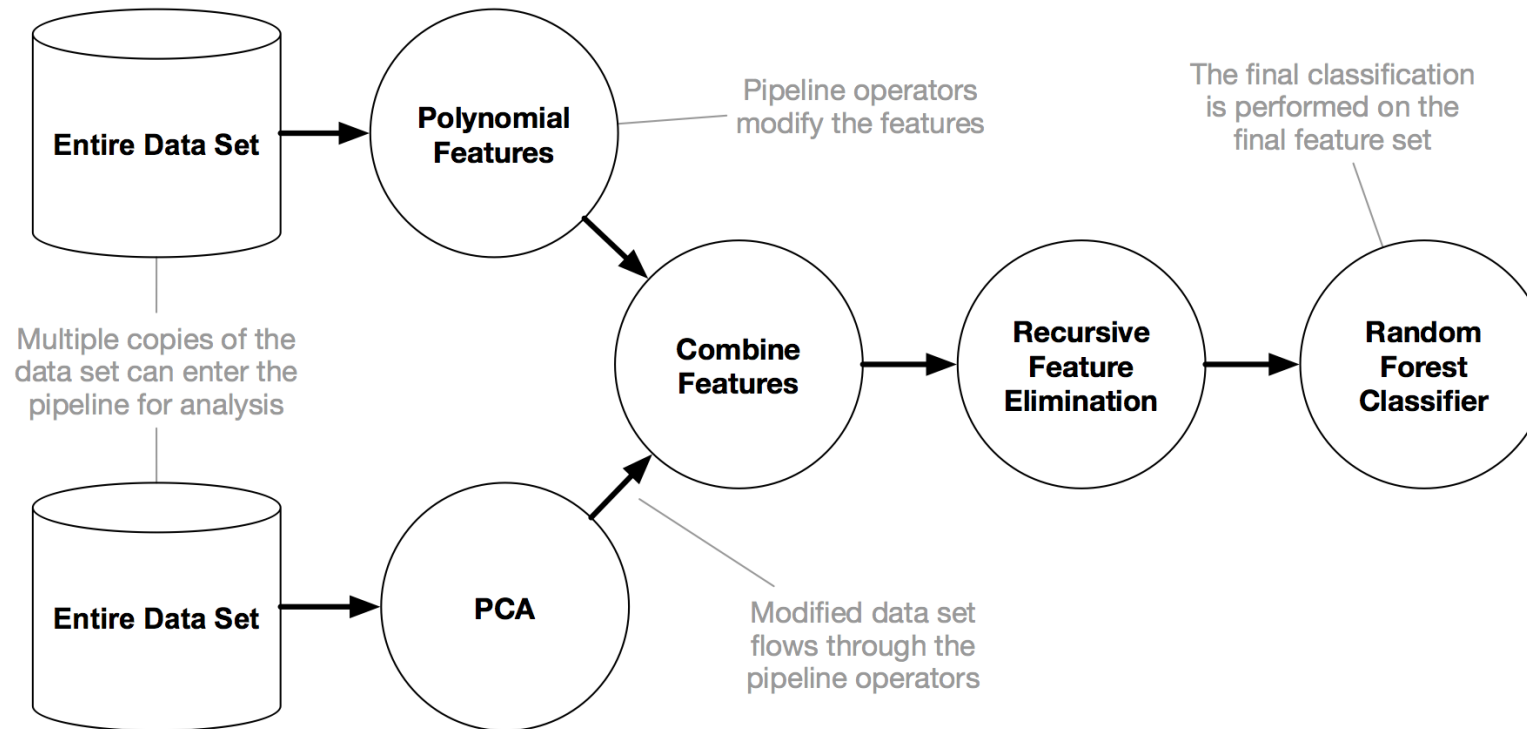
➤ AutoML

- Structure of a candidate solution



➤ AutoML

- It's Genetic Programming!



➤ AutoML

- TPOT/TPOT-2 is just one example of AutoML
 - Auto-WEKA (Java)
 - Auto-sklearn (variants of Bayesian optimization)
 - DataRobot (company, offers an online service)
 - AutoGluon
 - H2O AutoML
 - ...
 - **Auto-Pytorch: Neural Architecture Search**
- Take a look at the bibliography to know more

➤ Auto...optimization?

- Meta-optimization
- “Algorithm selection” / “Automated algorithm design”
 - Finding the best algorithm for the target optimization problem
 - Carola Doerr (CNRS) is one of the leading experts out there
 - Olivier Teytaud (Facebook) works on implementation
- Dagstuhl seminars discuss the state of the art

Jun 18 – Jun 23, 2023 • Dagstuhl Seminar 23251

Challenges in Benchmarking Optimization Heuristics

[Anne Auger](#) (INRIA Saclay - Palaiseau, FR) • [Peter A. N. Bosman](#) (CWI - Amsterdam, NL) • [Pascal Kerschke](#) (TU Dresden, DE) • [Darrell Whitley](#) (Colorado State University - Fort Collins, US)

Aug 13 – Aug 18, 2023 • Dagstuhl Seminar 23332

Synergizing Theory and Practice of Automated Algorithm Design for Optimization

[Martin S. Krejca](#) (Ecole Polytechnique - Palaiseau, FR) • [Marius Lindauer](#) (Leibniz Universität Hannover, DE) • [Manuel López-Ibáñez](#) (University of Manchester, GB) • [Katherine M. Malan](#) (UNISA - Pretoria, ZA)



The logo for INRAE, consisting of the word "INRAE" in a bold, teal, sans-serif font.The logo for Université Paris-Saclay, featuring the text "université PARIS-SACLAY" in a purple, sans-serif font, with a small purple dot above the "é" in "université".

➤ Questions?

Bibliography

- James et al., *An Introduction to Statistical Learning with Applications in Python*, 2023
- Feurer et al., *Efficient and Robust Automated Machine Learning*, 2015
- Hutter et al., *Automated Machine Learning*, 2019

Images and videos: unless otherwise stated, I stole them from the Internet. I hope they are not copyrighted, or that their use falls under the Fair Use clause, and if not, I am sorry. Please don't sue me.