Replacing the Ex-Def Baseline in AutoML by Naive AutoML
Felix Mohr (Universidad de La Sabana), Marcel Wever (Paderborn University)

The Ex-Def Baseline
- Is just to cross-validate all learners and pick the best one (no pre-processors considered).
- Proposed (and used only) in the 2013 Auto-WEKA paper.
- Claimed to be outperformed but in fact (much) stronger than one would expect.
- We extend this baseline by also including pre-processors, meta (ensembles) learners, and some random parameter tuning.

Naivety Assumption

Formulation
Let $\phi(D, c_1 \circ .. \circ c_{k+1})$ be the prediction performance of the pipeline consisting of transformers $c_1, .., c_k$ and estimator $c_{k+1}$ on dataset $D$. Naive assumption:

$$c_i^* \in \arg\min_{c_i} \phi(D, c_1 \circ .. \circ c_{k+1})$$

is invariant to the choices of $c_1, .., c_{i-1}, c_{i+1}, .., c_{k+1}$. If we consider the search space an urn and denote as $Y$ the event to observe an optimal pipeline in the urn, then

$$P(Y | c_1, .., c_{k+1}) \propto P(c_1, .., c_{k+1} | Y)P(Y)$$

$$= P(c_i | Y) \prod_{j=1, j \neq i}^{k+1} P(c_j | Y)P(Y)$$

Implication
The naivety assumption allows to optimize all parts in isolation. Or at least to have only a loose coupling.

Naive AutoML

Idea
Optimize different aspects of the ML pipeline in sequentially isolated stages.

Results (ranks after 1h runtime, summarizing 67 datasets from AutoML Benchmark)