AutoML: Algorithm Selection

Overview and Motivation

Bernd Bischl Frank Hutter Lars Kotthoff
Marius Lindauer Joaquin Vanschoren

Algorithm Selection

Given a problem, choose the best algorithm to solve it. [Rice. 1975]

Algorithm Selection

More formally

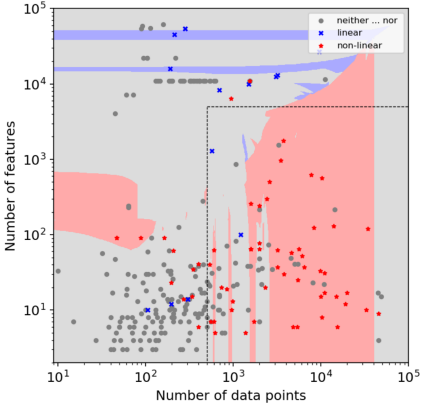
Let

- $p(\mathcal{D})$ be a probability distribution over datasets $\mathcal{D} \in \mathbf{D}$,
- $\mathbf{P}$ a portfolio of algorithms $\mathcal{A} \in \mathbf{P}$, and
- $c : \mathbf{P} \times \mathbf{D} \to \mathbb{R}$ be a cost metric

the *per-instance algorithm selection problem* is to obtain a mapping $s : \mathcal{D} \mapsto \mathcal{A}$ such that

$$\arg\min_{s} \int_{\mathbf{D}} c(s(\mathcal{D}), \mathcal{D}) p(\mathcal{D}) \, \mathrm{d}\mathcal{D}$$

Motivation: Performance Differences [Strang et al. 2018] I



Motivation: Performance Differences [Strang et al. 2018] II