AutoML needs to be more interactive, explainable, robust to mistakes, and clear about what it does.

AutoML Adoption: Insights from Interviews
Koen van der Blom¹,4, Alex Serban²,5, Holger Hoos³,1,6, Joost Visser¹
¹Leiden University, ²Radboud University, ³RWTH Aachen University, ⁴Sorbonne University, ⁵Siemens, ⁶University of British Columbia

Summary
https://se-ml.github.io

Survey shows
• 20-30% do not adopt AutoML at all
• Another 50-60% do not completely adopt AutoML

Interviews indicate
• Improved performance and time savings
• High computational costs hold back adoption
• Desire to understand the system and whether it is correct
• Concerns about data misuse and overfitting

Background
• AutoML aims to reduce the workload to apply ML
• Survey: How much is AutoML actually adopted?
• Interviews: What facilitates or inhibits the adoption of AutoML?

What is AutoML?
HPO
NAS
Algorithm selection
Grid search
Supervised classif.
Feature selection
Algo. selection
ML pipeline config.
Feature generation
Algo. conf./design
Data cleaning
Data int./pret./proc.
Grid search

What is used?
HPO
Grid search
Algorithm selection
NAS
Supervised classif.
Feature selection
Model selection
Feature generation
Algo. configuration

Usage
• Use AutoML to fine-tune or as baseline
• Difficulties: many different applications, need business on board, not clear where to use in existing system
• Not always needed (good enough is fine, human-like preferred)
• Use AutoML, but combined with human expert

Benefits
• Already get better results
• Already saves human time, others see the potential
• Potential to get and deploy results faster
• Already reduces complexity (more accessible for non-experts)

Risks
• Unclear what an AutoML system does, and if that is correct
• AutoML may overfit or take advantage of problems in the data
• Less data-expert interaction, miss data and problem details
• Overestimating the abilities of AutoML systems

Challenges
• Computational cost, including for mistakes and debugging
• Need for explainability/visualisation of output model/parameters
• Desire to know what the AutoML system is doing
• Hard to use and/or unclear/outdated documentation

Feature gen. and selection n = 67
Feature gen. and selection n = 51
Model select. and HPO n = 66
Configure algos. or model struct. n = 65

Survey setup
• Target: Teams developing software with ML components
• Question 1: We use automated methods to generate or select features from input data
• Question 2: We perform model selection and hyper-parameter optimisation in an automated way
• Question 3: We use automated methods to configure our algorithms or the structure of our models
• Answers: Not at all; Partially; Mostly; Completely; Implicitly