

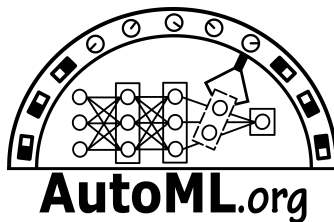
# Zero-Shot AutoML with Pretrained Models

Ekrem Öztürk\*, Fabio Ferreira\*, Hadi Jomaa\*, Lars Schmidt-Thieme, Josif Grabocka, Frank Hutter

ICML 2022

(\*equal contrib.)

<https://github.com/automl/zero-shot-automl-with-pretrained-models>

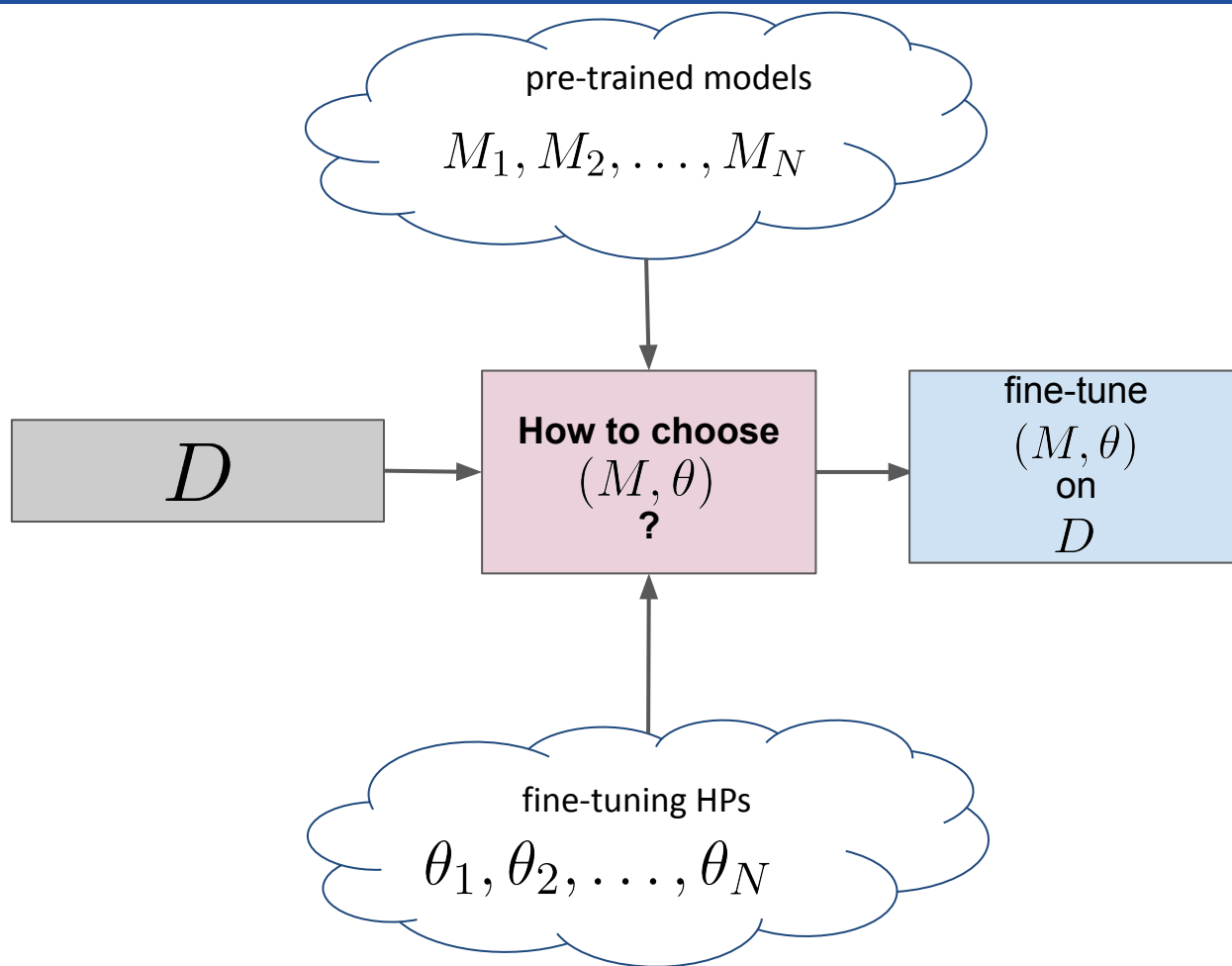


Code





# Motivation



# Motivation

pre-trained models

$M_1, M_2, \dots, M_N$

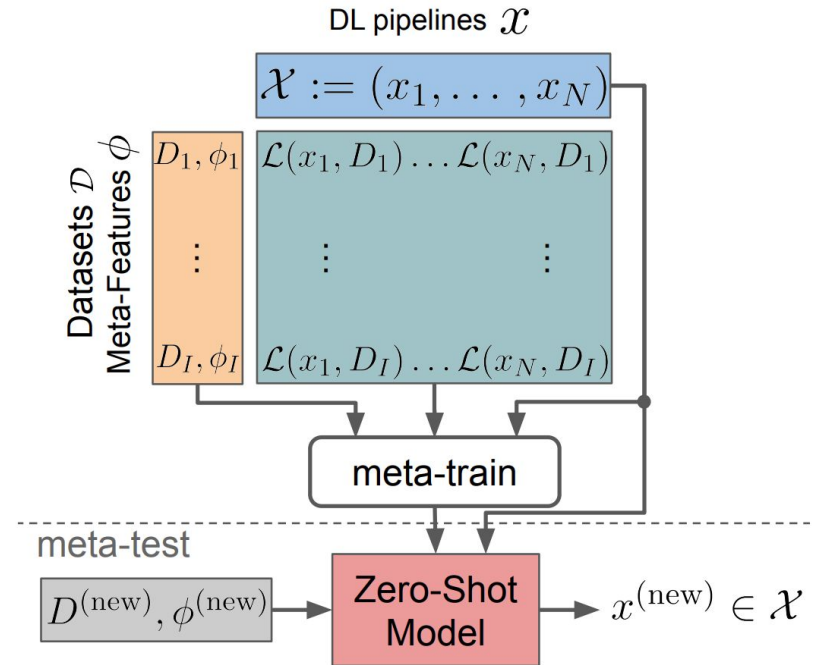
We extend **AutoML** to best make such choices. Our **domain-independent, meta-learning** approach learns a **zero-shot model** that allows to select the **right DL pipeline** given only **trivial meta-features of D**

fine-tuning HPs

$\theta_1, \theta_2, \dots, \theta_N$

# Approach

- We leverage a **cost matrix** to learn a **joint response surface** conditioned on
  - meta-features and
  - DL pipelines
- At meta-test time, the **zero-shot model predicts scores of the pipelines trained on unseen datasets**





# Approach

- **Zero-shot AutoML with Pretrained Models with Algorithm Selection (ZAP-AS)**
  - learn a **selector** between DL pipelines with AutoFolio
  - yields already **very good performance!**
  - **Drawback:** AS abstracts away our DL pipelines as **uncorrelated**
- **Zero-shot AutoML with Pretrained Models with Hyperparameter Optimization (ZAP-HPO)**
  - interprets the space of **DL pipelines as search space for HPO**
  - train an **NN-based selector** with a **pairwise ranking** (of configurations) **objective**

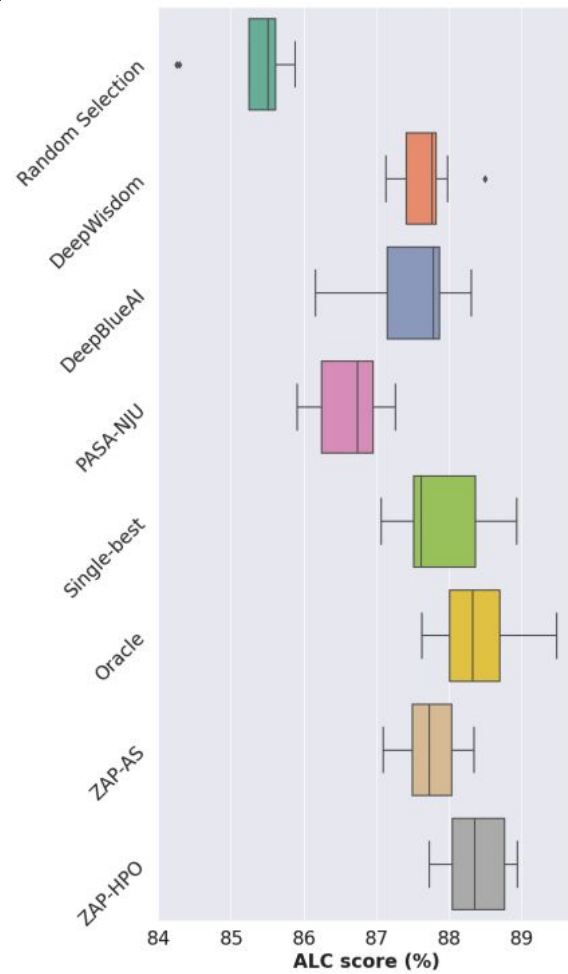
- Meta-Dataset:
  - **set of datasets** with **meta-features**
  - a set of **DL pipelines**
  - **test costs** for the pipelines on the datasets
  
- Datasets:
  - **35 core datasets** from TensorFlow Datasets
  - **Diversity**: small, large, colored, b/w, various image resolutions & domains
  - **augmented** each core dataset **15 times**
  - **525 datasets in total**



# DL Pipeline Design Space & Optimization

- based on the winning approach of the AutoCV competition
- built a highly parameterized space of DL pipelines around this with **26 hyperparameters**
- **Applied HPO to optimize for the Area under the Learning Curve (ALC) score introduced in the AutoDL challenge**
- **Result:** one optimized DL pipeline per dataset
- To arrive at the described **meta-dataset**, we evaluated each **525 pipelines across all 525 datasets**  
→ **Meta-Dataset with 275 625 cost entries**

# Results: Own benchmark







# Results: AutoDL Benchmark

<b>Solution</b>	<b>Rank (ALC)</b>
DeepWisdom	$2.46 \pm 0.13$
DeepBlueAI	$2.76 \pm 0.08$
PASA NJU	$2.62 \pm 0.11$
<b>ZAP-HPO</b>	<b><math>2.16 \pm 0.15</math></b>



# Thank you!

- Please do check out our paper for more details or stop by our (physical) poster at ICML!