AutoML: Meta-Learning

Introduction

Bernd Bischl Frank Hutter Lars Kotthoff
Marius Lindauer Joaquin Vanschoren

Intro: humans can easily learn from a single example

thanks to years of learning (and eons of evolution)

Canna Indica ‘Picasso’

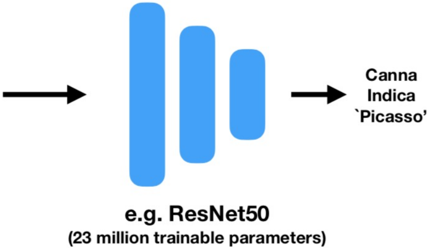
?

train

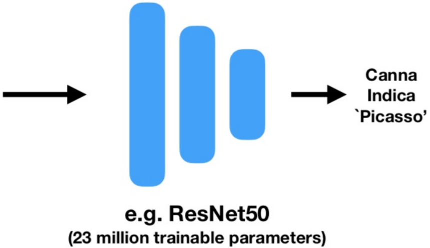
later →

test

Can a computer learn from a single example?



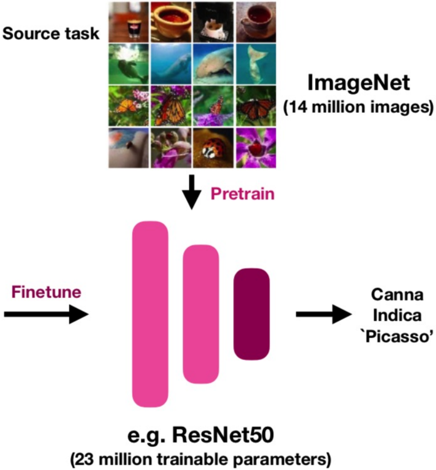
Can a computer learn from a single example?



That won't work :) Humans also don't start from scratch.

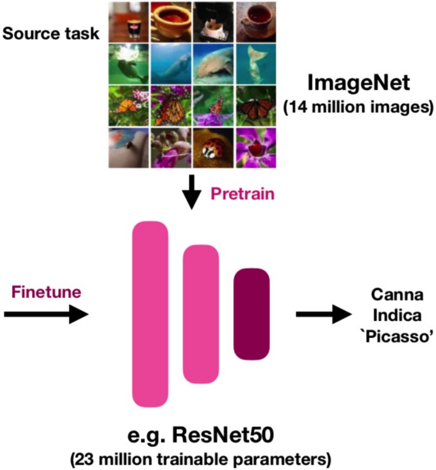
Transfer learning?

Target task



Transfer learning?

Target task



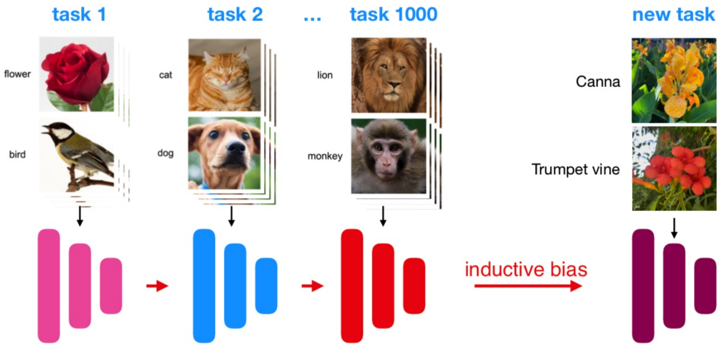
A single source task (e.g. ImageNet) may not generalize well to the test task.

Meta-learning

Learn over a series (or distribution) of many different tasks/episodes

Inductive bias (or prior): learn assumptions that you can to transfer to new tasks

Prepare yourself to learn new things faster

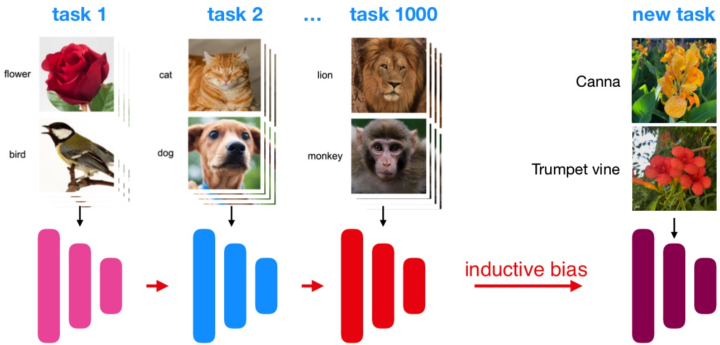


Meta-learning

Learn over a series (or distribution) of many different tasks/episodes

Inductive bias (or prior): learn assumptions that you can to transfer to new tasks

Prepare yourself to learn new things faster

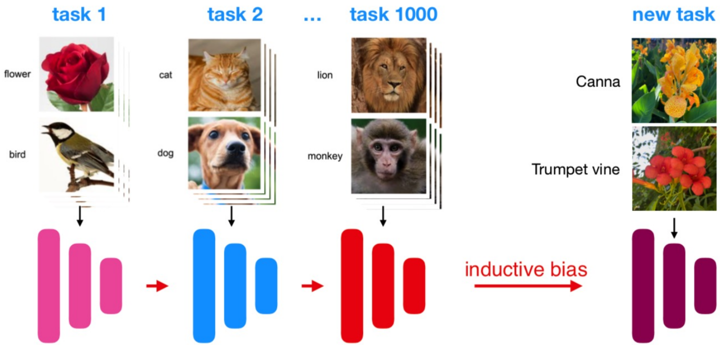


Closely related to continual learning, online learning, multi-task learning

Meta-learning

Learn over a series (or distribution) of many different tasks/episodes

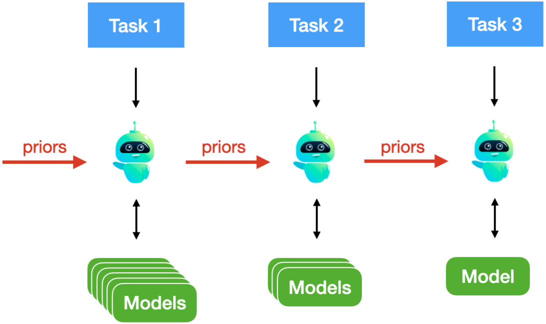
Inductive bias (or prior): learn assumptions that you can to transfer to new tasks
Prepare yourself to learn new things faster



Useful in many real-life situations: rare events, test-time constraints, data collection costs, privacy issues,...

Meta-learning for AutoML?

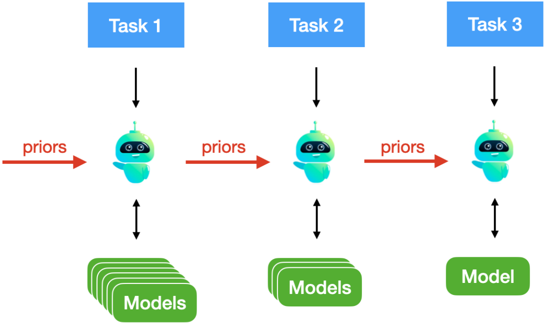
AutoML systems should not start from scratch, they should learn across tasks
Human ML experts also get better over time



- Current AutoML systems require humans to hard-code a set of assumptions
- What if AutoML systems could learn these by themselves?

Meta-learning for AutoML?

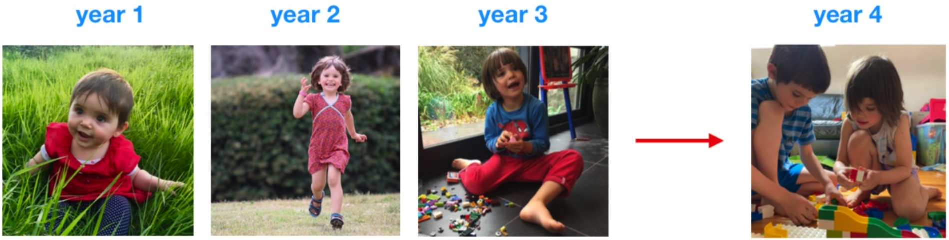
AutoML systems should not start from scratch, they should learn across tasks
Human ML experts also get better over time



- Meta-learning can make AutoML much more efficient
- Vice-versa: AutoML can make meta-learning more robust

Inspired by human learning

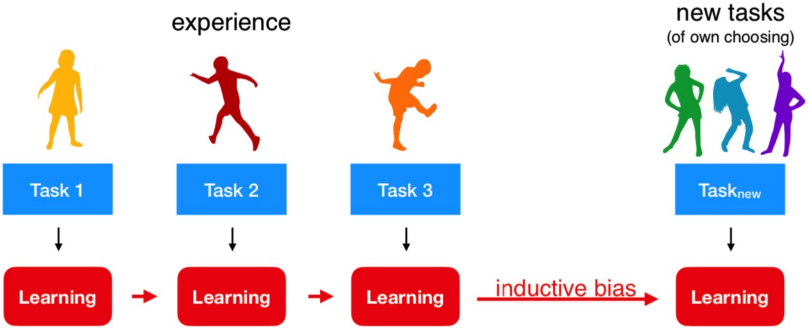
We don't transfer from a single source task, we learn across many, many tasks
We have a 'drive' to explore new, challenging, but doable, fun tasks



Human-like Learning***

humans learn across tasks: less trial-and-error, less data, less compute

new tasks should be related to experience (doable, fun, interesting?)

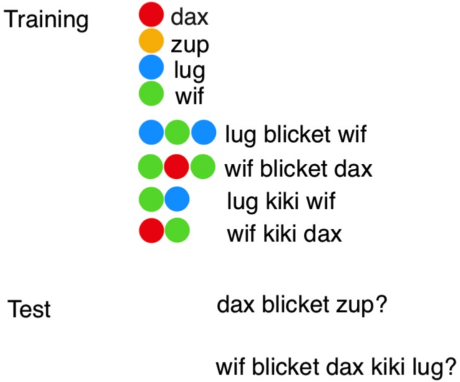


key aspects of fast learning: compositionality, causality, <u>learning to learn</u>

Lake et al. (2017) Building machines that learn and think like people.

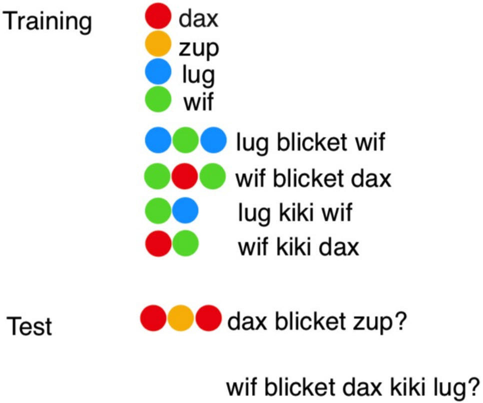
Inductive bias (in language)

which assumptions do we make?



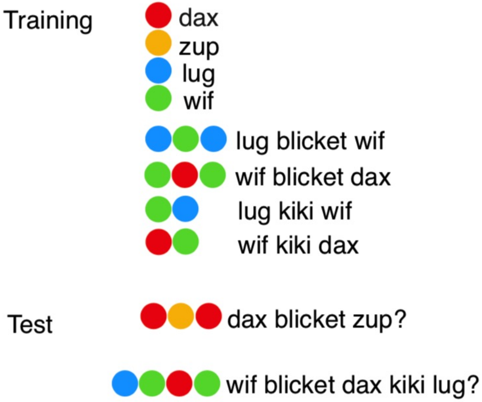
Inductive bias (in language)

which assumptions do we make?



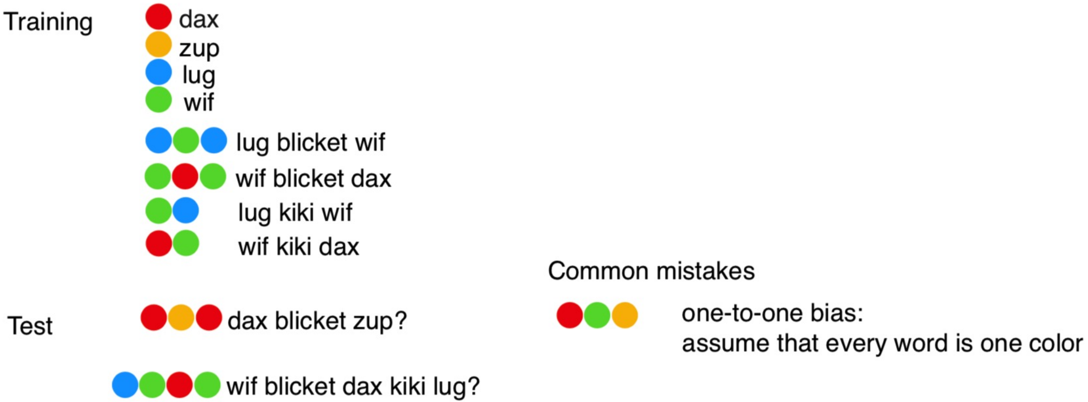
Inductive bias (in language)

which assumptions do we make?



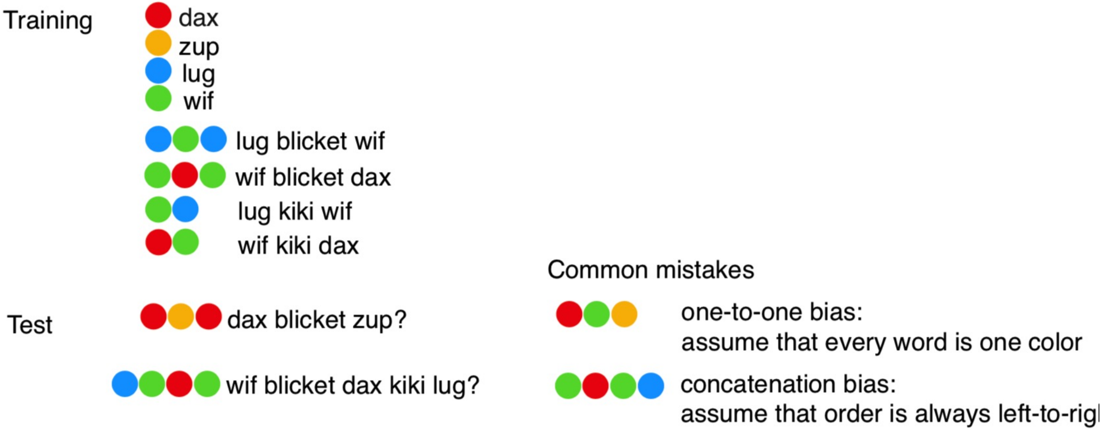
Inductive bias (in language)

which assumptions do we make?



Inductive bias (in language)

which assumptions do we make?



What if there is no training data?

we can still solve problems by making assumptions

Item pool 

Test

zup?

zup zup?

dax zup?

zup tufa?

zup wif zup?

zup wif blicket?

blicket wif zup?

What if there is no training data?

we can still solve problems by making assumptions

Item pool 

Test  zup?

zup zup?

dax zup?

zup tufa?

zup wif zup?

zup wif blicket?

blicket wif zup?

What if there is no training data?

we can still solve problems by making assumptions

What if there is no training data?

we can still solve problems by making assumptions

What if there is no training data?

we can still solve problems by making assumptions

What if there is no training data?

we can still solve problems by making assumptions

What if there is no training data?

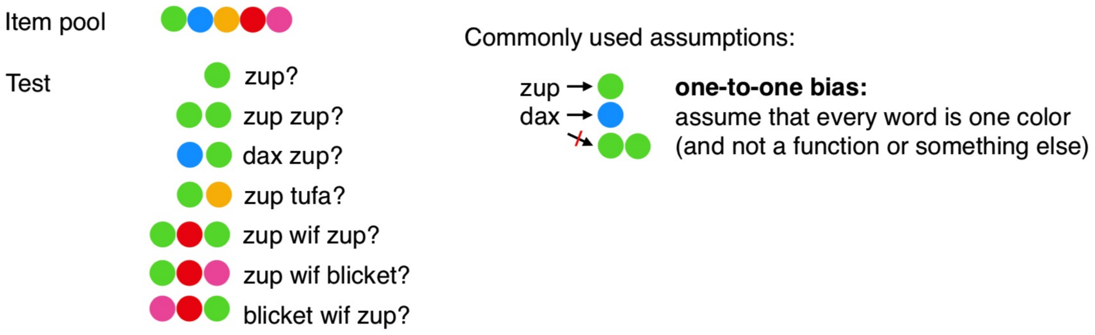
we can still solve problems by making assumptions

What if there is no training data?

we can still solve problems by making assumptions

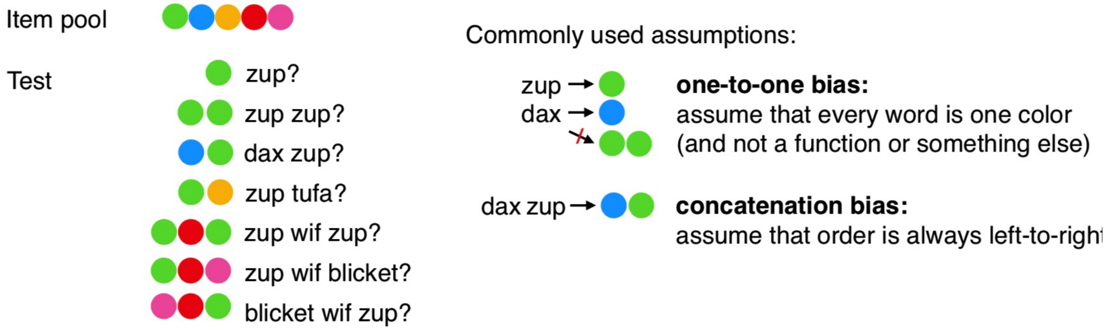
What if there is no training data?

we can still solve problems by making assumptions



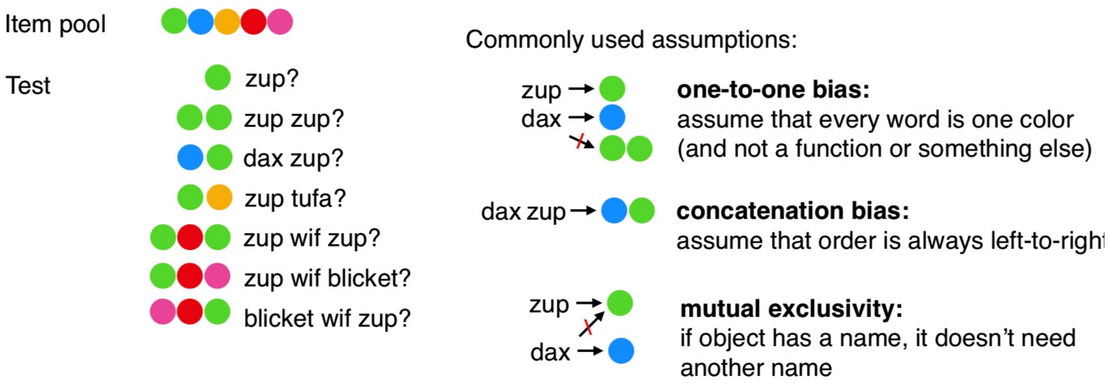
What if there is no training data?

we can still solve problems by making assumptions



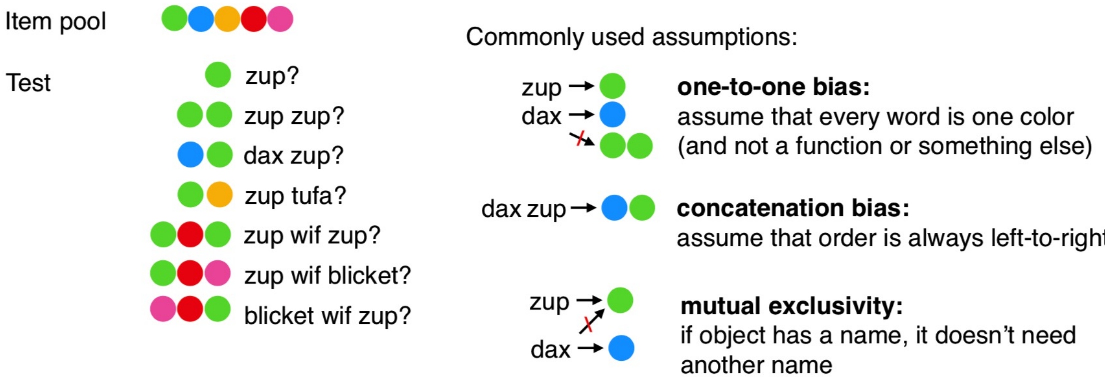
What if there is no training data?

we can still solve problems by making assumptions



What if there is no training data?

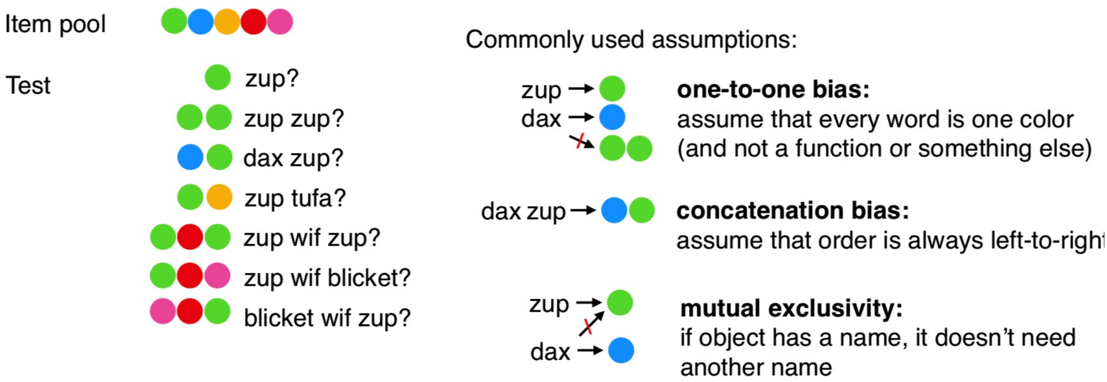
we can still solve problems by making assumptions



Humans *assume* that words have consistent meanings and follow input/output constraints

What if there is no training data?

we can still solve problems by making assumptions

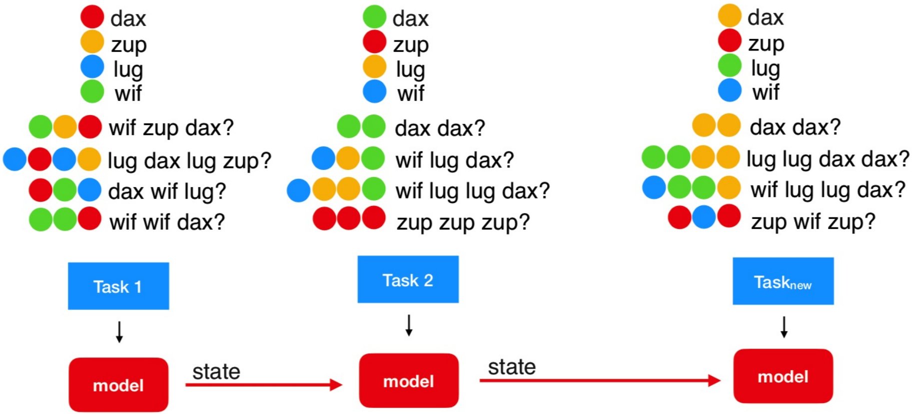


Humans *assume* that words have consistent meanings and follow input/output constraints

These assumptions (inductive biases) are necessary for learning quickly

Meta-learning inductive biases

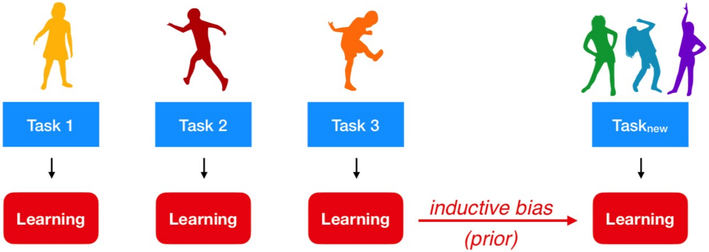
Capture useful assumptions from the data - that can often not be easily expressed



Lake (2019) Compositional generalization through meta sequence-to-sequence learning.

Meta-learning goal

learn *minimal* inductive biases from prior tasks instead of constructing manual ones
should still generalize well (otherwise you meta-overfit)

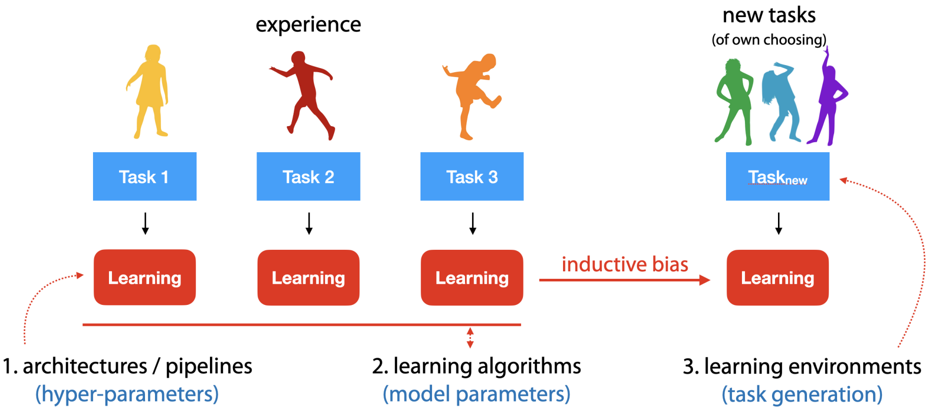


Inductive bias: any assumptions added to training data to learn more effectively. E.g:

- Instead of general model architectures, learn better architectures (and hyperparameters)
- Instead of starting from random weights, learn good initial weights
- Instead of standard loss/reward function, learn a better loss/reward function

What can we learn to learn?

3 pillars



Clune (2019) AI-GAs: AI-generating algorithms.