Lecture 1 : Introduction and overview.

AutoML → architecture search, hyperparameter optimization, learning optimizers

→ How can we enable agents to learn skills in the real world?

↓

Robots : agents that actually interact with the world

- can teach us about intelligence.
- must generalize across tasks, object environments, etc.
  (problems that humans actually deal with)
- need some sense of common sense to do well
- supervision can't be taken for granted.

① speeding up RL training ② speeding up training in general
vision based RL

*) The same deep reinforcement algorithm can be applied to a wide array of tasks

*) The problem is with generalization of the learned behaviour over different tasks.

↓

Learn one task in one environment, starting from scratch, relying on detailed supervision and guidance.

(specialists)

(problem with standard RL algorithms
↓
lot of data, lot of supervision)

→ Humans are generalists

↓

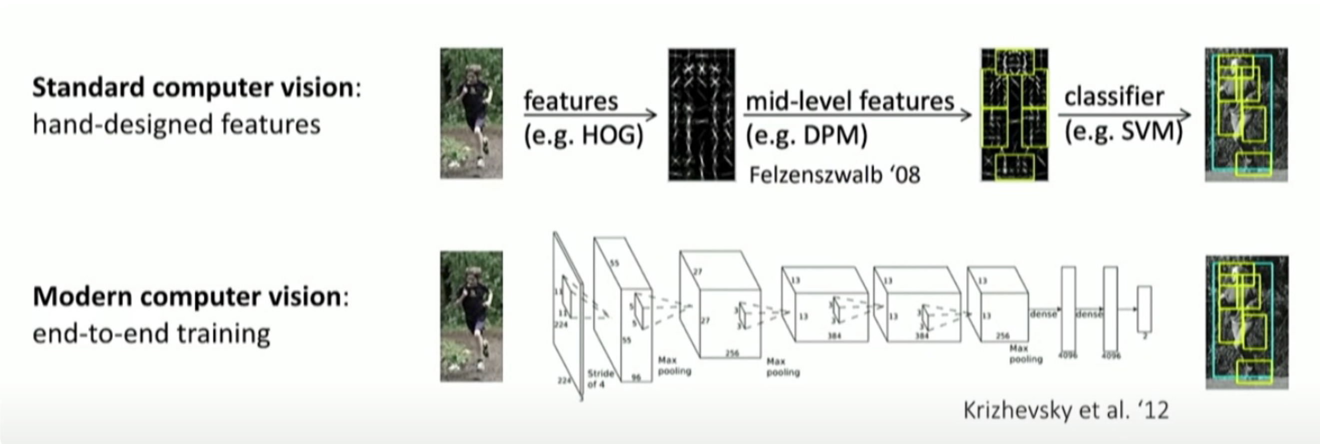
Inspiration for systems that can build upon previous experience to learn more complex tasks, that too quickly.

↓

The above techniques can be thought of teaching a baby to learn to play Go.

→ why should we care about deep multi-task and metalearning?

deep learning allows us to handle unstructured inputs (pixels, languages, sensor readings, etc.) w/o hand engineered features, with less domain knowledge.



They work well on different tasks.

explainable AI

•) one thing that we've picked from Deep Learning is that if we keep networks large, diverse datasets, they can achieve broad generalization.

What if we don't have large datasets?

medical imaging

robotics

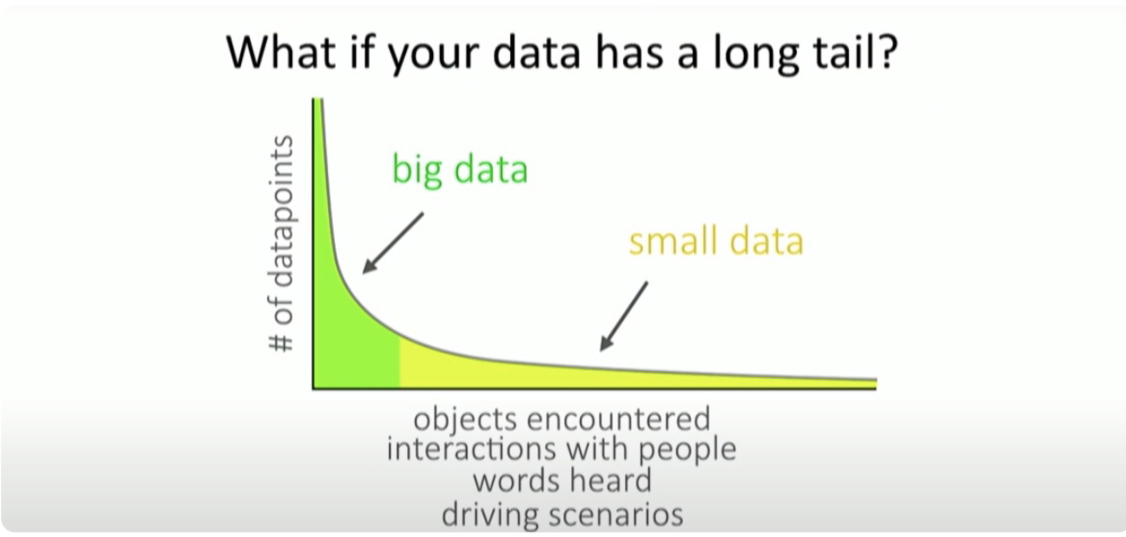
personalized education

translation for rare languages

medicine, recommendations

↓

Impractical to learn from scratch for each disease, each robot, each person, each language, each task.



↓

what if the task distribution has a long tail? i.e. the probability of encountering uncommon data is considerable. This is where supervised learning methods really struggle to perform well.

↓

for example: autonomous driving, where cars can handle common situations, but behave weirdly when given uncommon data?

→ what if we need to learn something new quickly?

few-shot learning : only given a few data points, and the goal is to make predictions about new data points from very small dataset

↓

humans accomplish this, since they aren't learning from scratch; they have previous experience.

→ what is a task?

dataset D

Loss function L

→ model f_θ

machine learning problem, given a dataset and a loss function, we produce a model.

Different tasks can vary based on:

- different objects
- different people
- different objectives
- different lighting objectives
- different words, etc.

Different datasets, Different loss fns.

→ Critical assumption : Different tasks need to share a structure, for these techniques to work.

↓

If tasks don't share a structure, then we're better off using single-task learning.

But there are many tasks with the shared structure.

↓

applies to metalearning as well.

→ Informal problem definitions :

Unsupervised learning & Reinforcement Learning

The multi-task problem: Learn all of the tasks more quickly or more proficiently than learning them independently.

The meta-learning problem: Given data/experience on previous tasks, learn a new-task more quickly and or/more proficiently

multi-task: learning and doing well on a set of tasks whereas in meta-learning, we use experience gained in learning those tasks to do well at new tasks.

→ Doesn't multi-task learning reduce to single-task learning?

domain adaptation

$$\left.\begin{aligned} D &= \cup D_i \\ L &= \Sigma L_i \end{aligned}\right\}$$ becomes a single-task learning

↓

Actually, aggregating the data across tasks & learning a single model is one approach to multi-task learning.

↓

But, we can often to better, by exploiting the fact that the data is coming from different tasks.