Introduction to Machine Learning

David I. Inouye

What is Machine Learning?

- Machine Learning (ML) is a field of Artificial Intelligence (AI) where computer agents improve their perception, cognition, and action with experience.

> “Machine Learning is about machines improving from data, knowledge, experience, and interaction.” – Manuela Veloso, Carnegie Mellon University

Why Use Machine Learning?

Some problems are too complex to define with explicit, hand-written rules.

The Challenge: Programming by Hand

How would you write a program to identify a cat? You'd have to create rules for every possibility:

- Does it have pointy ears?
- What about whiskers and fur?
- How do you account for different breeds, angles, and lighting?

This approach is brittle and practically impossible to perfect.

The Solution: Learning from Examples

Instead of writing the rules, we let the computer learn them from data.

- It's far easier to collect 10,000 pictures of cats than to write 10,000 rules.
- We show the machine labeled examples, and it learns the underlying patterns itself.

ML finds the rules when examples are easier to find than instructions.

Three Main Types of Machine Learning

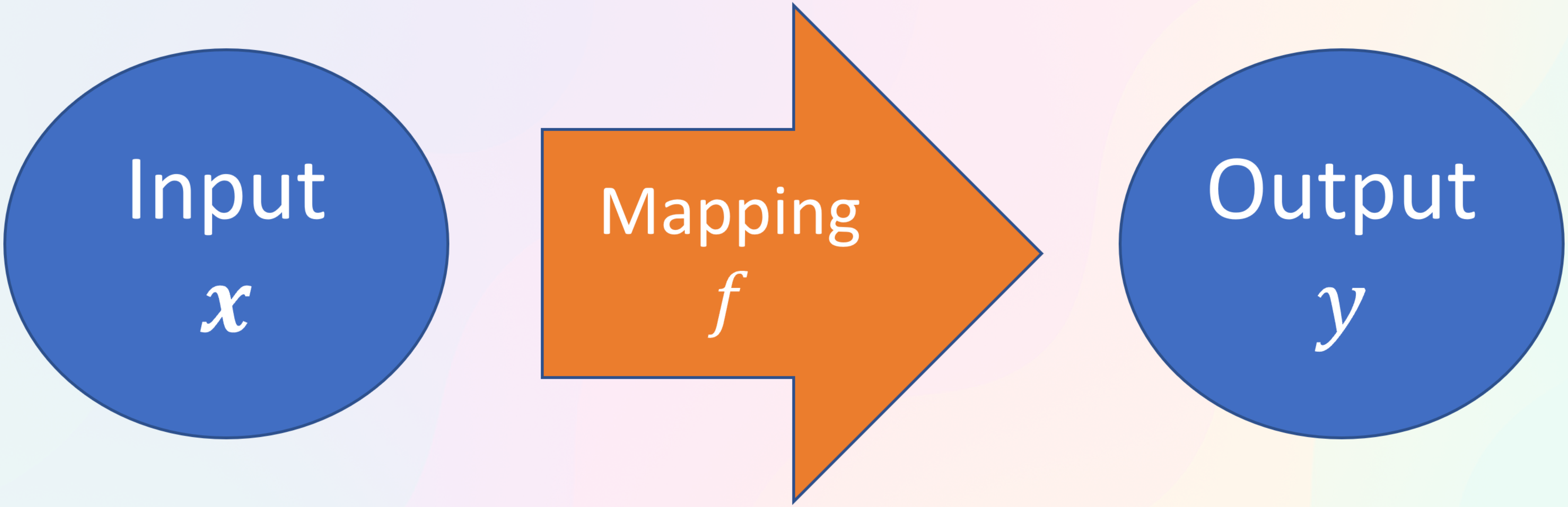
There are three main categories of machine learning:

- **Supervised learning**
  - Regression
  - Classification
- **Unsupervised learning**
  - Dimensionality reduction (PCA)
  - Clustering
  - Generative models
- **Reinforcement learning**
  - Learning from interaction and feedback

Outline

- What is Machine Learning?
- The Three Pillars of Machine Learning
  - Supervised Learning
  - Unsupervised Learning
  - Reinforcement Learning
- The Core Goal: Generalization
- Key Concepts & Limitations

The Goal of **Supervised Learning** is to Estimate a Mapping (or Function) Between Input and Output



The Goal of **Supervised Learning** is to Estimate a Mapping (or Function) Between Input and Output Given Only Input-Output Examples

The Set of Input-Output Pairs Is Called a **Training Set**

Denoted by $\mathcal{D} = \{(x_i, y_i)\}_{i=1}^{n}$

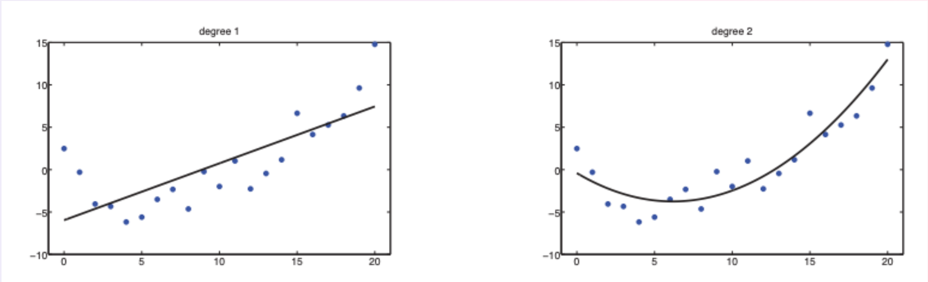
Input x_i

- Called *features* (ML), *attributes*, or *covariates* (Stats). Sometimes just *variables*.
- Can be *numeric*, *categorical,discrete*, or *nominal*.
- Examples
  - [height, weight, age, gender]
  - $[x_1, x_2, \ldots, x_d]$ - A d-dimensional vector of numbers
  - Image
  - Email message

Output y_i

- Called *output*, *response*, or *target* (or *label*)
- *Real-valued/numeric* output: e.g., $y_i \in \mathcal{R}$
- *Categorical*, *discrete*, or *nominal* output: y_i from finite set, i.e., $y_i \in \{1, 2, \ldots, c\}$

If the Output y_i Is Numeric, Then the Problem Is Known as **Regression**



Two plots showing linear (degree 1) and quadratic (degree 2) regression fits to scattered data points.

- Given height x_i, predict age y_i
- Predict GPA given SAT score
- Predict SAT score given GPA
- Predict GRE given SAT and GPA

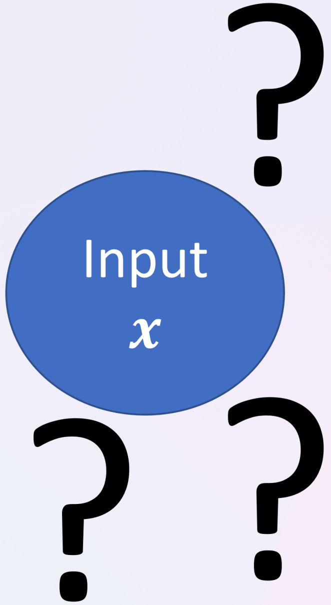
NOTE: Input x does not have to be numeric. Only the output y must be numeric.

If Output Is **Categorical**, Then the Problem Is Known as **Classification**

- Given height x, predict "male" ($y = 0$) or "female" ($y = 1$)
- Given salary x_1 and mortgage payment x_2, predict defaulting on loan ("yes" or "no")

Grid of images showing cats and dogs with predictions.

The Goal of Unsupervised Learning Is to Model or Understand the Input Data Without Labels



- Dimensionality reduction
- Clustering
- Generative models

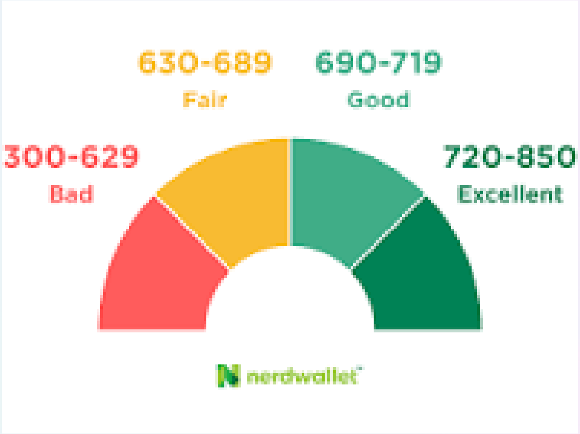
> "What I cannot create I do not understand"
> – Richard Feynman

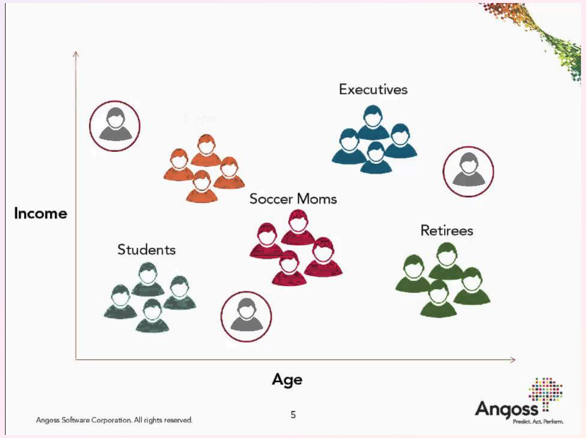
In Unsupervised Learning, the **Training Set** Is Only a Set of Input Values $\mathcal{D} = \{x_i\}_{i=1}^n$

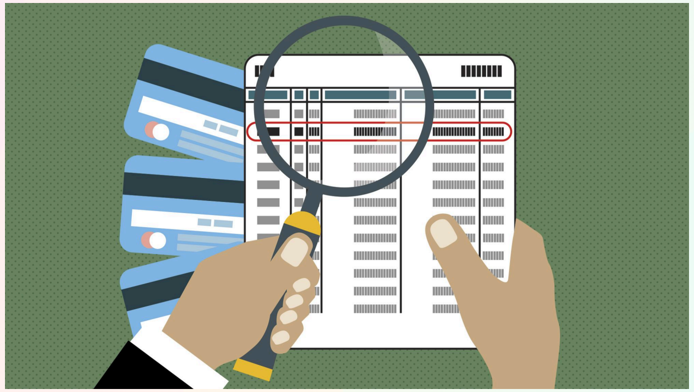
[Dimensionality reduction] Estimate a single number that summarizes all variables of wealth (e.g. credit score)

[Clustering] Estimate natural groups of customers

[Generative Models] Estimate the distribution of normal transactions to detect fraud (anomalies)





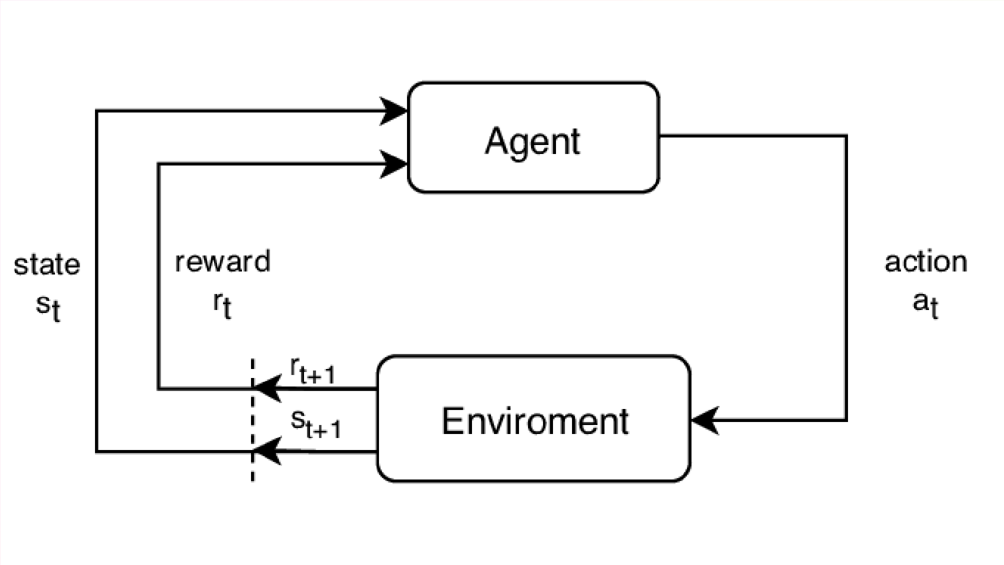


.

Reinforcement Learning: Learning from Interaction

Reinforcement Learning (RL) is about learning to make decisions by trial and error. An "agent" learns to achieve a goal in a complex, uncertain environment by receiving rewards or penalties for its actions.

- No labeled data is required.
- The agent learns from a **reward signal**.
- Focuses on maximizing long-term cumulative reward.



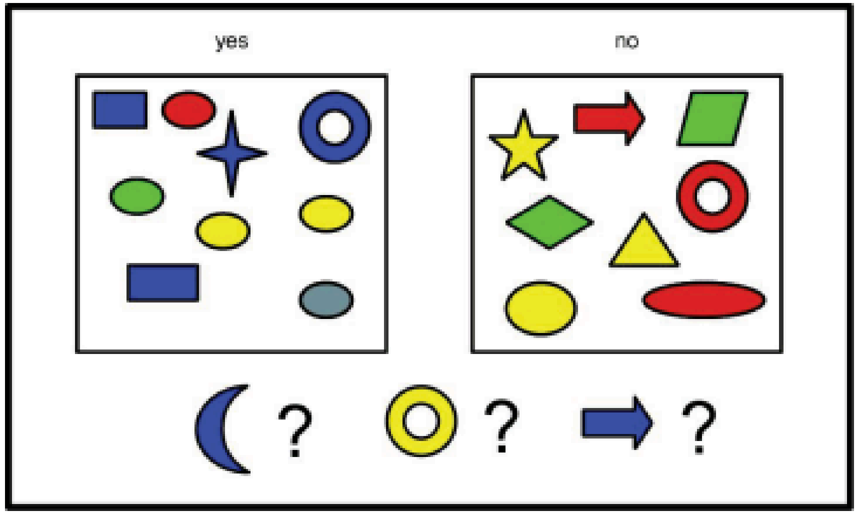
A diagram showing the reinforcement learning loop with an agent, environment, action, state, and reward.

Examples of Reinforcement Learning

RL is used to solve a wide variety of problems where optimal decision-making is key.

- **Robotics:** Training robots to perform complex tasks like grasping objects or navigating a room without bumping into things.
- **Game Playing:** Mastering complex games like Go (AlphaGo) or video games by playing against itself millions of times.
- **Autonomous Driving:** Learning optimal driving strategies based on traffic conditions and road rules.
- **Recommendation Systems:** Personalizing recommendations (e.g., on Netflix or Yahoo) by learning from user interactions to maximize long-term engagement.

Generalization Beyond the Training Set Is the Main Goal of Learning



D features (attributes)

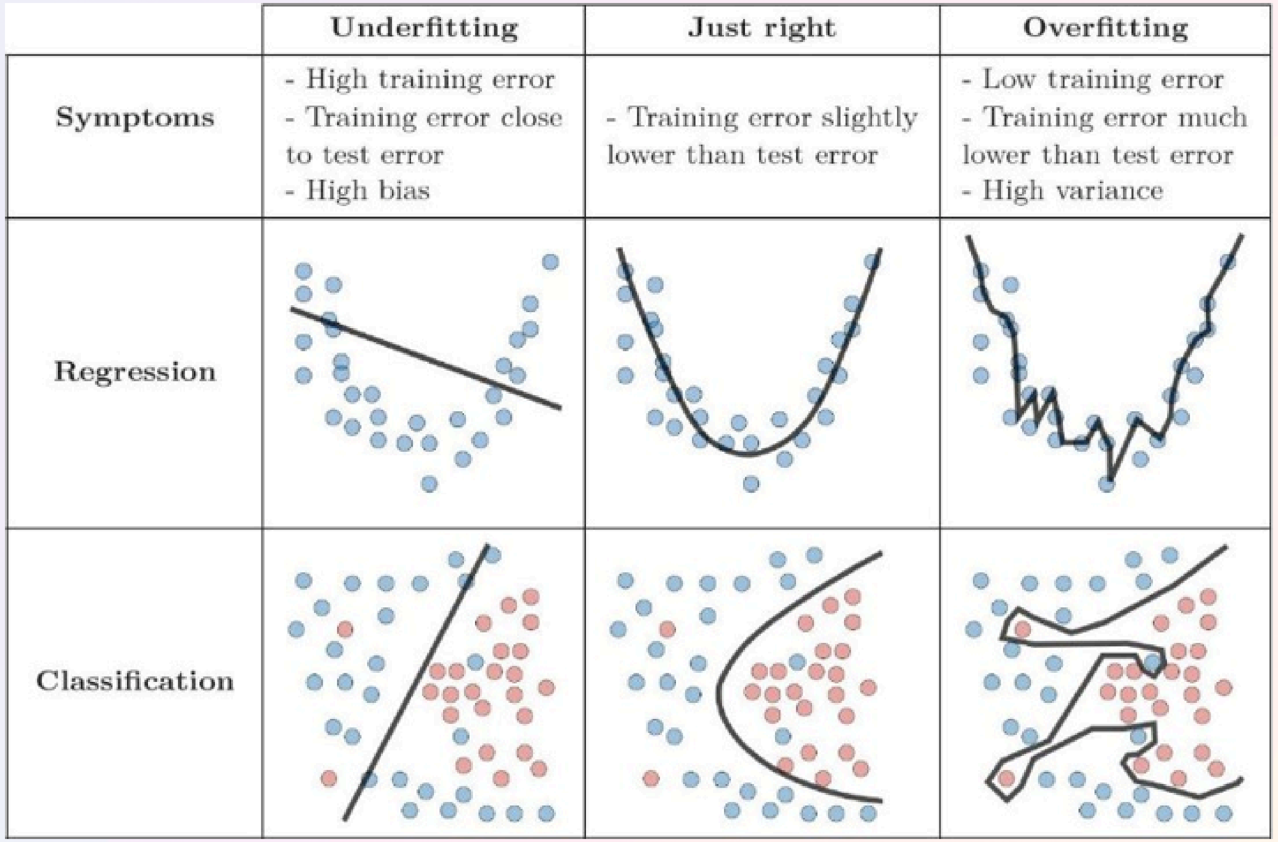
N cases

Color	Shape	Size (cm)	Label
Blue	Square	10	1
Red	Ellipse	2.4	1
Red	Ellipse	20.7	0

- Discuss with partner: Which box (yes or no) do the three new shapes belong to?

Example from Machine Learning: A Probabilistic Perspective, Ch. 1, Kevin P. Murphy, 2012.

Generalization Beyond the Training Set Is the Main Goal of Learning



The balance between overfitting and underfitting was an critical concept for the 2nd (current) wave of AI with statistical learning.

Original source for figure unknown.

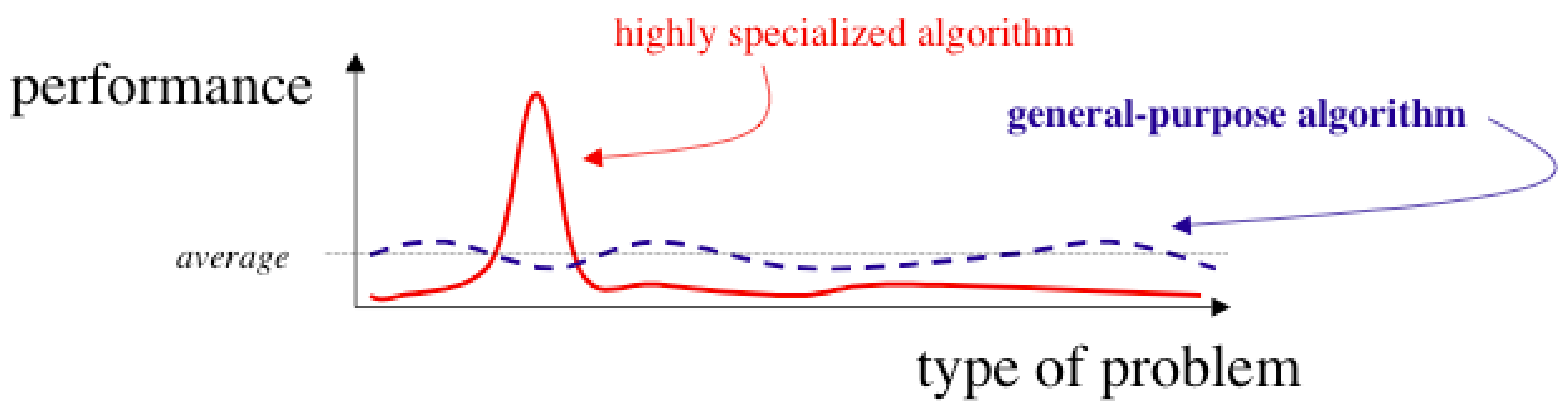
What Does Generalization Look Like for **Unsupervised Learning**?

- **Generalization in dimensionality reduction**
  - Objective on train may be small but test may be large

- **Generalization in generative models** can be understood through the view of log likelihood.

“All models are wrong, but some models are useful.”*

- All models are approximations
- All models make assumptions
- Assumptions are never perfect
- *No Free Lunch Theorem*



*George Box (Box and Draper 1987, page 424).

Questions?