Introduction to Learning

Alice Gao

Lecture 8

Readings: R & N 18.1 and 18.2

Based on work by K. Leyton-Brown, K. Larson, and P. van Beek

Outline

Learning Goals

By the end of the lecture, you should be able to

- Identify reasons for building an agent that can learn.
- Describe different types of learning.
- Define supervised learning, classification, and regression.
- For supervised learning, describe
  how we choose a hypothesis.
- Describe over-fitting.

Applications

- Medical diagnosis
- Spam filtering
- Facial recognition
- Speech understanding
- Handwriting recognition

Agents that learn

We want to build agents that learn from observations about the world to improve their performance on future tasks.

Why would we want an agent to learn?

What can the agent learn?

- How conditions on the current state map to actions
- What the world is like now
- How the world evolves
- What it will be like if I do action A
- How happy I will be in such a state
- What action I should do now

One class of learning problems

from a collection of example-label pairs,
learn a function that predicts the labels for new examples.

Three Types of learning

- Unsupervised learning: learn patterns in the examples, no labels available
- Semi-supervised learning: learn from a few labeled examples and a large number of unlabeled examples.
- Supervised learning: learn a function that maps from examples to labels

CQ: Supervised or Unsupervised Learning

CQ: We are given information on a user's credit card transactions. We would like to detect whether some of the transactions are fraudulent by finding some transactions that are different from the other transactions. We have no information on whether any particular transaction is fraudulent or not.

Is this a supervised or unsupervised learning problem?

(A) Supervised learning

(B) Unsupervised learning

Supervised Learning

- A training set of N example-label pairs,

$$(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N),$$

  was generated by an unknown function $y = f(x)$.
- x_i represents the i-th example and the y_i represents the label for the i-th example.
- Goal is to discover a function h in a hypothesis space H that approximates f well.

Classification v.s. Regression

- Classification: the label is a discrete value.
- Regression: the label has continuous values.

CQ: Classification or regression

CQ: Is the following problem classification or regression?

You are given historical data on the weather condition (sunny, cloudy, rain, or snow) on a particular day of the year. You want to predict the weather condition of this day next year.

(A) Classification

(B) Regression

(C) This is not supervised learning.

CQ: What type of problem is this?

CQ: Is the following problem classification or regression?

You are given historical data on the price of a house at several points in time. You want to predict the price of this house next month.

(A) Classification

(B) Regression

(C) This is not supervised learning.

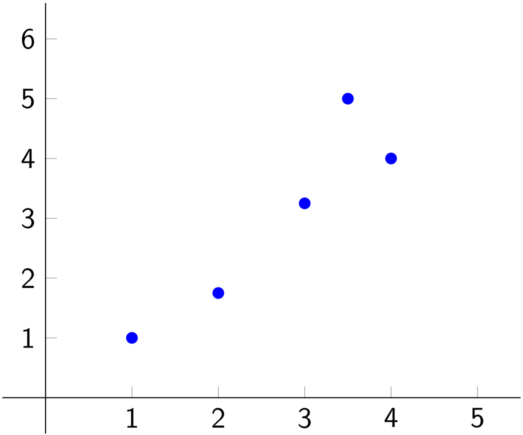
Accuracy of a Hypothesis

We want a hypothesis h that generalizes well
— h correctly predicts the value of y for unseen examples.

Measure the accuracy of the hypothesis
on a **test set** of examples
that are **distinct from the training set**.

An example of regression

Training set: (1, 1), (2, 1.75), (3, 3.25), (3.5, 5), (4, 4)



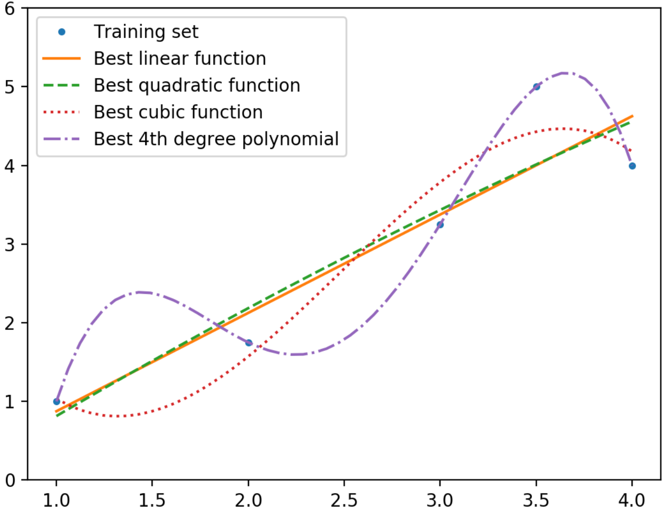
Hypotheses

What are the possible hypothesis spaces H?

- straight line
- piece-wise linear function
- quadratic function
- cubic function

What is the best hypothesis in each space?
A hypothesis is **consistent** if it agrees with all the training examples.

Possible hypotheses



How do you choose among multiple hypotheses?

Trade-off between:

- Complex hypotheses that fit the training data well
- Simpler hypotheses that may generalize better

Ockham's Razor (English philosophy William of Ockham)

The simplest solution tends to be the correct one.

How do we choose among multiple hypotheses?

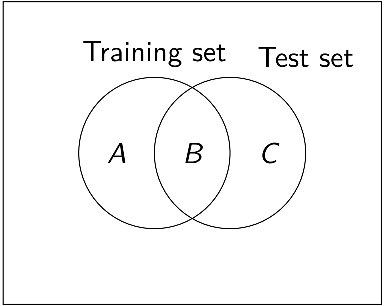
- We prefer the **simplest consistent** hypothesis.
- Sometimes, an even **simpler but inconsistent** hypothesis might generalize even better.

Over-fitting

CQ: Under-fitting as a Venn Diagram

CQ: Consider the Venn diagram below. Suppose that a hypothesis is **under-fitting** the training data.
Which area(s) of the diagram does the hypothesis belong to?

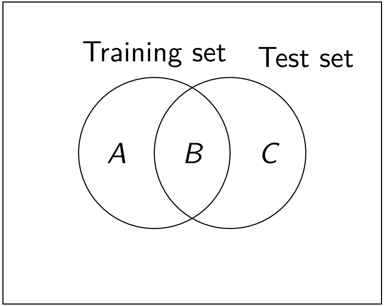
(A) A
(B) B
(C) C
(D) A and B
(E) A, B, and C



CQ: Over-fitting as a Venn Diagram

CQ: Consider the Venn diagram below. Suppose that a hypothesis is **over-fitting** the training data.
Which area(s) of the diagram does the hypothesis belong to?

(A) A
(B) B
(C) C
(D) A and B
(E) A, B, and C



CQ: Which hypothesis is prune to overfitting?

CQ: Suppose that we are considering a simple hypothesis (a straight line) and a complex hypothesis (a 4th degree polynomial). Which of the two is more likely to overfit the training data?

(A) The simple hypothesis

(B) The complex hypothesis

(C) I don't know

Revisiting the Learning Goals

By the end of the lecture, you should be able to

- Identify reasons for building an agent that can learn.
- Describe different types of learning.
- Define supervised learning, classification, and regression.
- For supervised learning, describe
  how we choose a hypothesis.
- Describe over-fitting.