

Machine Learning Introduction

dsai.asia

Asian Data Science and Artificial Intelligence Master's Program



Co-funded by the
Erasmus+ Programme
of the European Union



Readings for these lecture notes:

- Bishop, C. (2006), *Pattern Recognition and Machine Learning*, Springer, Chapters 3, 4, 6, 7.
- Hastie, T., Tibshirani, R., and Friedman, J. (2016), *Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer, Chapters 2, 3, 4, 12.
- Ng, A. (2017), *Supervised Learning*, Lecture note set 1 for CS229, Stanford University.
- Ng, A. (2017), *Generative Learning Algorithms*, Lecture note set 2 for CS229, Stanford University.
- Ng, A. (2017), *Support Vector Machines*, Lecture note set 3 for CS229, Stanford University.

These notes contain material © Bishop (2006), Hastie et al. (2016), and Ng (2017).

1 Introduction

Introduction

What is machine learning?

Machine learning is now near the top of the list of skills U.S. companies want to see in the people they hire.

What's all the fuss, and what is machine learning?

Many tasks we want computers to do are difficult to program directly.

Examples: image recognition, speech recognition, controlling a self-driving car.

Machine learning

A set of tools that let us specify the computer's behavior by giving examples of **how** it should respond in given situations, **without specifying the computation necessary** to formulate that response.

We tell the computer **what** it should decide to do in a situation but not **how** to make the decision.

Introduction

What's a model?

Essential idea: we want to create a **model** from data that can later be **queried** when new situations arise.

Model

A (mathematical) function whose input is a **description of the current situation** and whose output is a **decision, recommendation, or action**.

Introduction

Examples of ML in real life

We are using machine learning every time we

- Use a credit card
- Get a recommendation from Netflix or Amazon
- Ask Google for directions by voice
- Take a ride in our Tesla!

Let's brainstorm about things closer to home that might be using machine learning already or might benefit from it in the near future.

Introduction

The four problems for machine learning

Machine learning comprises perhaps four basic problems:

- **Classification**: place instances into one or more of a set of given discrete **categories**.
- **Regression**: estimate a function from sample inputs/outputs that can later be used for **interpolation** or **extrapolation**.
- **Density estimation**: estimate a probability density function from a sample from the distribution that can later be used, e.g., for **anomaly detection**.
- **Reinforcement**: derive a **policy** that enables an agent to behave optimally in an uncertain environment using **feedback** on the goodness of the outcome over time.

Let's think about the input and output of the model in each of these cases.