Artificial Neural Networks
Lecture Notes

Prof. Dr. Sen Cheng

Winter Semester 2019/20

1 Introduction

1.1 Deep neural networks in the news

Artificial intelligence creates perfumes without being able to smell them

How quickly can AI solve a Rubik's Cube? In less time than it took you to read this headline.

Top AI researchers race to detect 'deepfake' videos: 'We are outgunned'

AI is more powerful than ever. How do we hold it accountable?

Warning! Everything Is Going Deep: 'The Age of Surveillance Capitalism'

1.2 What is a neural network?

An artificial neural network is a universal function approximator.

So this means that an ANN can be mathematically described as

$$y = f(x). \tag{1}$$

What is a function? A function is a relation that maps an element in the *domain A* onto a unique element in the *codomain B* (Fig. 1).

$$f : A \longrightarrow B \tag{2}$$

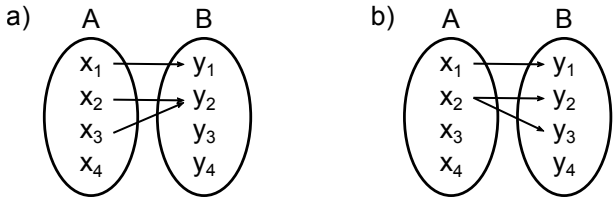


Figure 1: The mapping in a) is a function, the one in b) is not.

In real applications, relationships between two variables are almost never deterministic, i.e., there are unpredictable variations. We can model this with a *random variable* ε that doesn't indicate a fixed value, but a random value drawn from a *statistical distribution*. Distributions are characterized by their *probability density functions* (PDFs) (Fig. 2). With a random variable, we can model the uncertain relation between two variables as a *stochastic function*

$$y = f(x) + \varepsilon. \tag{3}$$

ε is often called a *noise* term. Most commonly we assume that the noise is drawn from a *normal distribution* with mean μ and variance σ^2, a.k.a. *Gaussian distribution*, and denote this by

$$\varepsilon \sim N(\mu, \sigma^2). \tag{4}$$

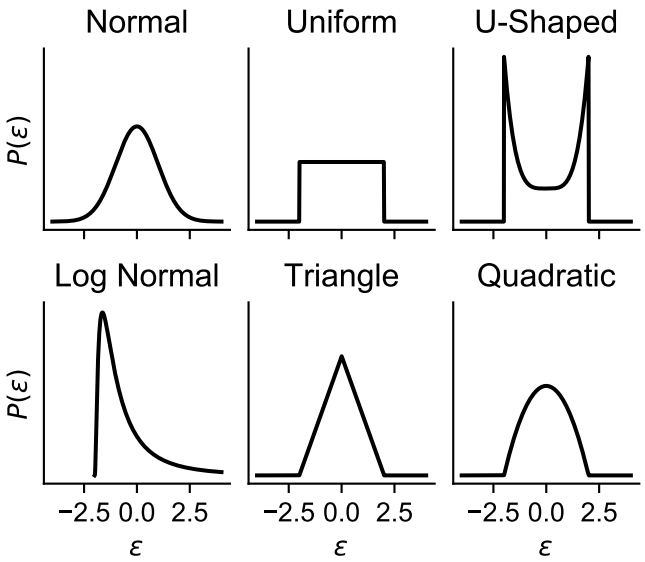


Figure 2: Different classes of distributions.

Many real-world problems can be mapped onto the following mathematical problem: What function best describes the relationship between two variables in the data (X,Y)? Finding such a function is also called *fitting* a function to data. Only recently, it has been shown that ANN are particularly good function approximators. While this class is about ANN, it is important to understand the fundamental principles underlying function fitting. That's why we start with a very general statement of the problem and simple functions that have little to do with ANN.

1.3 Relationship between ANN and AI

artificial neural networks $\subset$ machine learning $\subset$ artificial intelligence (AI).

There are three fundamental paradigms in machine learning: *supervised learning*, *unsupervised learning*, and *reinforcement learning*. This class focuses on supervised learning, however ANN are frequently used in the setting of unsupervised or reinforcement learning.

The literature on ANN is full of tricks and algorithms. We will try to bring out the underlying principles in the lectures. To this end a mathematical description is inevitable. The tutorials will deepen your understanding of the theoretical background by linking it to practical exercises and also enable you to apply the methods in practice.

1.4 Applications

Many real-world applications can be understood as function fitting.

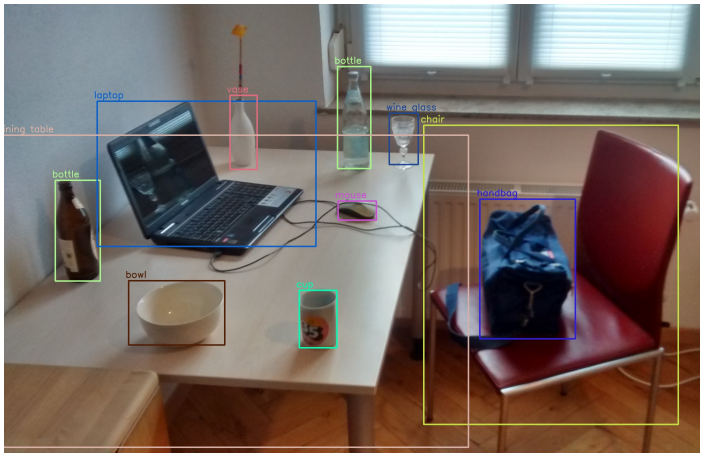
- **Recognition**: Recognizing a visual object, a spoken word or other sensory stimuli is about mapping from a physical stimulus to a semantic label (Fig. 3).
- **Prediction**: predicting a future value (e.g. weather, stock prices, consumer demand, product failure, traffic) is a mapping from past values to future ones

$$(x_{t-a},\ldots,x_t) \longrightarrow x_{t+b} \tag{5}$$

- **Recommender systems** (content-based filtering): mapping from content to like/dislike
- **Fraud/ Spam detection**: mapping from content to fraud/not-fraud or spam/ham (Fig. 3).
- and more

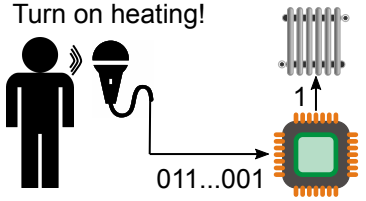
However, there are applications that cannot (easily) be modelled as function fitting.

a)



Source: https://commons.wikimedia.org/wiki/
File:Detected-with-YOLO--Schreibtisch-mit-Objekten.jpg

b)



c)

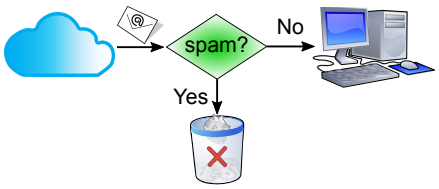


Figure 3: Examples of recognition tasks. a) Object recognition from an image of a scene. b) Speech recognition in the context of smart homes. c) Spam e-mail filtering.

- **Problem solving** (e.g. finding shortest path from A to B, playing games)
- **Recommender systems** (collaborative filtering)
- **Decision making** (avoiding accident in automated driving systems)
- and more

Note, even if a problem cannot be mapped onto function fitting, it turns out that ANN can be an essential component on a subtask, e.g. in deep reinforcement learning.