



# 神经网络与深度学习内容选讲

助教-吴至婧

清华大学 计算机系 智能技术与系统国家重点实验室

2022年04月11日

主要内容取自 复旦大学 邱锡鹏教授

<https://nndl.github.io/>

1

## 神经网络与深度学习：引言



- ◎ 近些年兴起的一种数据挖掘和分析方法
    - ◆ 自然语言处理、音频分析和计算机视觉
  - ◎ 优点
    - ◆ 不依赖于人工构建的手工特征
    - ◆ 借助梯度下降等优化算法，直接从原始数据中学习特征
    - ◆ 泛化能力更强，预测效果更好
  - ◎ 模型结构
    - ◆ 本质上还是由各种基础神经网络构成
    - ◆ 通过“搭积木”的方式构建适用于特定任务的网络模型
- ➡
- ◎ 三大类基础神经网络结构，三者结合可以处理各种常见的数据
    - ◆ 前馈、卷积、循环神经网络
  - ◎ 介绍以理解性为主，不会涉及具体的数学原理

2



## 目录



- 前馈神经网络
- 卷积神经网络
- 循环神经网络
- 如何实现

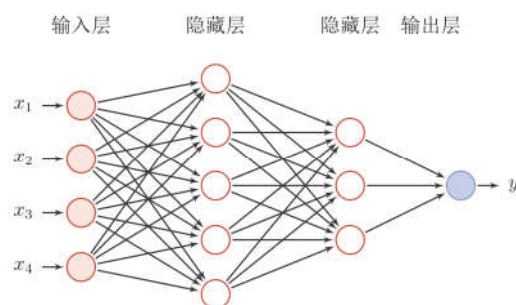
3



## 前馈神经网络 回顾



- 在前馈神经网络中，各神经元分别属于不同的层。
- 整个网络中无反馈（有向无环图）
  - ◆ 整个网络中无反馈（有向无环图）
  - ◆ 参数数量多（ $m \times n$ ）



4



## 前馈神经网络



- 模型

- $y = f^5(f^4(f^3(f^2(f^1(x))))))$

- 学习准则

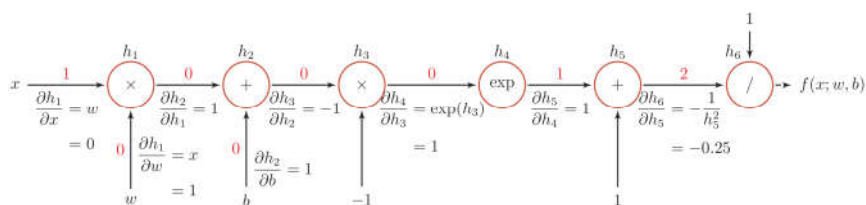
- $L(y, y^*)$

- 优化

- 梯度下降

$$\frac{\partial L(y, y^*)}{\partial f^1} = \frac{\partial f^2}{\partial f^1} \times \frac{\partial f^3}{\partial f^2} \times \frac{\partial f^4}{\partial f^3} \times \frac{\partial f^5}{\partial f^4} \times \frac{\partial L(y, y^*)}{\partial f^5}$$

链式法则，可以自动计算！



5



## 前馈神经网络



- 通用近似定理

- 对于具有线性输出层和至少一个使用“挤压”性质的激活函数的隐藏层组成的前馈神经网络，只要其隐藏层神经元的数量足够，它可以以任意精度来近似任何从一个定义在实数空间中的有界闭集函数。

## 定理 4.1 – 通用近似定理 (Universal Approximation Theorem)

[Cybenko, 1989, Hornik et al., 1989]: 令  $\varphi(\cdot)$  是一个非常数、有界、单调递增的连续函数， $\mathcal{I}_d$  是一个  $d$  维的单位超立方体  $[0, 1]^d$ ， $C(\mathcal{I}_d)$  是定义在  $\mathcal{I}_d$  上的连续函数集合。对于任何一个函数  $f \in C(\mathcal{I}_d)$ ，存在一个整数  $m$ ，和一组实数  $v_i, b_i \in \mathbb{R}$  以及实数向量  $\mathbf{w}_i \in \mathbb{R}^d$ ， $i = 1, \dots, m$ ，以至于我们可以定义函数

$$F(\mathbf{x}) = \sum_{i=1}^m v_i \varphi(\mathbf{w}_i^T \mathbf{x} + b_i), \quad (4.33)$$

作为函数  $f$  的近似实现，即

$$|F(\mathbf{x}) - f(\mathbf{x})| < \epsilon, \forall \mathbf{x} \in \mathcal{I}_d. \quad (4.34)$$

其中  $\epsilon > 0$  是一个很小的正数。

一个两层的神经网络可以模拟任何有界闭集函数。

6

