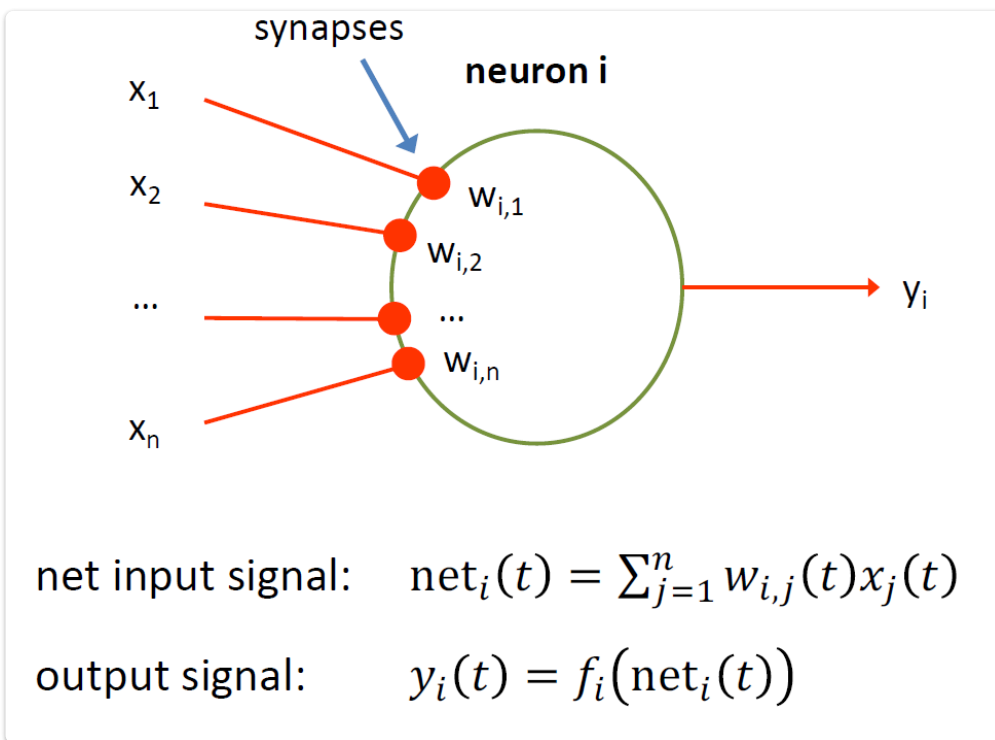


lec12

这个也不重要 没啥东西

人工神经网络（ANN）的基础知识，包括人工神经元的构成及其激活函数，以及如何通过监督学习来训练感知器以调整其权重和阈值，从而实现对输入数据的有效分类和预测

1. 神经网络通过适应其连接模式来学习，以使生物体在实现某些（进化）目标方面改善其行为
2. 连接的强度或是兴奋性还是抑制性取决于接收神经元突触的状态
3. 神经网络通过适当调整其突触的状态来实现学习

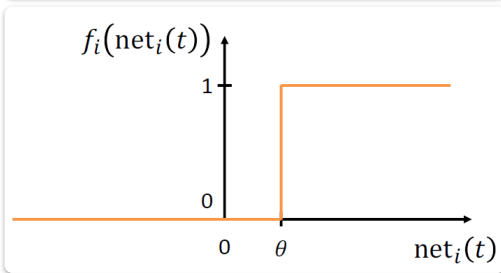


激活函数

- 阈值神经元（Threshold Neurons），如感知器（Perceptron）
- 阈值逻辑单元（Threshold Logic Units），处理二进制输入
- 线性神经元（Linear Neuron），输出是输入的线性组合
- 修正线性单元（ReLU）
- S型神经元（Sigmoidal Neurons），如使用Logistic函数

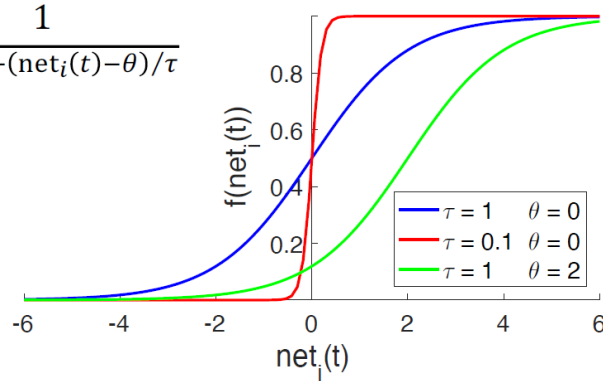
threshold function

$$f_i(\text{net}_i(t)) = \begin{cases} 1 & \text{if } \text{net}_i(t) \geq \theta \\ 0 & \text{otherwise} \end{cases}$$



Sigmoidal

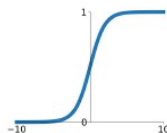
$$f_i(\text{net}_i(t)) = \frac{1}{1 + e^{-(\text{net}_i(t) - \theta)/\tau}}$$



else

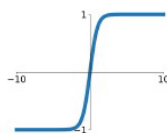
Sigmoid

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$



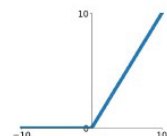
tanh

$$\tanh(x)$$



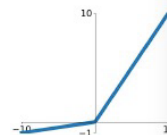
ReLU

$$\max(0, x)$$



Leaky ReLU

$$\max(0.1x, x)$$

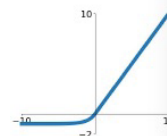


Maxout

$$\max(w_1^T x + b_1, w_2^T x + b_2)$$

ELU

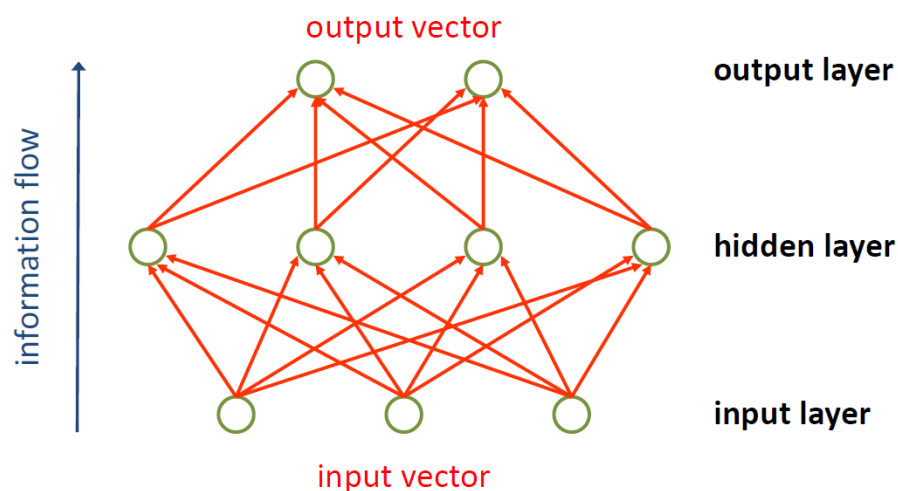
$$\begin{cases} x & x \geq 0 \\ \alpha(e^x - 1) & x < 0 \end{cases}$$



ANN

前馈神经网络

Example: network function $f: \mathbb{R}^3 \rightarrow \{0, 1\}^2$



多层神经网络

1. **输入层 (Input Layer) :**
 - 输入层包含输入向量，不执行任何计算，只是将输入数据传递给下一层
2. **隐藏层 (Hidden Layer) :**
 - 隐藏层从输入层接收输入，并将其输出发送到输出层。隐藏层可以有多个，是神经网络学习和处理复杂模式的关键部分
3. **输出层 (Output Layer) :**
 - 输出层的神经元在应用激活函数后，包含最终的输出向量，即网络的预测结果

设置人工神经网络 (ANN) 的权重和阈值

1. **通常采用：监督学习 (supervised learning) :**
 - 监督学习是一种常见的训练方法
2. **监督学习的过程:**
 - 在监督学习中，我们使用一组向量对来训练ANN，这些向量对称为示例或训练数据
 - 每个向量对包括一个输入向量和一个对应的输出向量
3. **学习特征:**
 - 通过提供足够多的示例，网络可以学习到描述数据的良好特征，从而在一般情况下能够很好地概括数据

监督学习

1. **监督学习的基本概念:**
 - 每个训练对 (x, y) 由一个输入向量 x 和一个对应的输出向量 y 组成
 - 当网络接收到输入 x 时，我们希望它提供输出 y
 - 这些示例描述了我们希望网络学习的函数
2. **泛化能力:**
 - 除了学习这些示例，我们还希望我们的网络能够泛化，即对未经过训练的输入给出合理的输出
3. **权衡 (Tradeoff) :**
 - 在网络准确学习给定示例的能力和其泛化能力之间存在权衡
 - 这个问题类似于为一组数据点拟合一个函数

- 我们可以尝试使用不同次数（复杂度）的多项式来进行拟合

监督学习在人工神经网络（ANNs）中的应用

1. 相同原则适用于人工神经网络：

- **神经元过少**：可能没有足够的自由度来精确逼近所需的函数，导致欠拟合（underfitting）
- **神经元过多**：尽管能完美地学习示例，但额外的自由度可能会导致对未训练输入的表现不合理，即过拟合（overfitting）

2. 缺乏确定的方程：

- 不幸的是，没有已知的方程可以告诉你在特定应用中网络的最佳规模；只有一些启发式方法（heuristics）