

# 2023-2024 学年秋季学期《模式识别与机器学习》课程

## 第 4 章作业题

### 一、选择题

1. (多选题) 下列针对激活函数 (如 ReLU, Sigmoid, Tanh 等) 的说法, 正确的是 ( )
  - A. 它们必须是可导的
  - B. 它们可以在某些点不可导, 但这些不可导的点必须很少
  - C. 它们可以是任意的连续函数
  - D. 它们必须是非线性的
2. (单选题) 下列哪类谓词逻辑问题不能用简单感知器解决? ( )
  - A. 与
  - B. 或
  - C. 非
  - D. 异或

### 二、简答题

请分析采用反向传播算法训练神经网络时, 学习率过大或过小时会产生什么负面影响。

### 三、证明题

试证明: 多层感知机中节点的激活函数如果采用线性函数, 网络无法实现非线性映射。

### 四、计算题

给定前向传播神经网络如下图所示,  $\{\mathbf{x}^{(i)}, y^{(i)}\}, i = 1, \dots, N$  为样本,  $\mathbf{x}^{(i)} \in \mathbb{R}^m$  为输入。

$$v_j^{(i)} = \sigma \left( \sum_{k=1}^3 w_{k,j} x_k^{(i)} \right), \quad \sigma(z) = \frac{1}{1 + \exp(-z)}$$

输出  $\hat{y}^{(i)} = \sigma(\boldsymbol{\theta}^T \mathbf{v}^{(i)})$ , 损失函数为

$$\mathcal{L} = -\frac{1}{N} \sum_{i=1}^N (y^{(i)} \log \hat{y}^{(i)} + (1 - y^{(i)}) \log(1 - \hat{y}^{(i)}))$$

1. 求  $\frac{\partial \mathcal{L}}{\partial \theta_2}$ 。
2. 设学习率为  $\eta$ , 求用随机梯度下降法更新  $w_{1,2}$  的表达式。

