

2023-2024 学年秋季学期《模式识别与机器学习》课程

第 14 章作业题

一、选择题

1. (多选题) 下列关于半监督学习的说法正确的是 ()
 - A. 在半监督学习中, 未标记样本和有标记样本需要是独立同分布的
 - B. 主动学习只是依赖人工标注, 没有对未标记样本的数据特点加以利用
 - C. 直推学习可能无法对新样本进行预测
 - D. 求解生成式半监督学习的 EM 算法的结果不依赖于初始值
2. (单选题) 以下关于半监督学习的说法错误的是 ()
 - A. 生成式半监督学习方法可基于高斯混合模型, 并使用 EM 算法求解
 - B. 生成式半监督学习方法对先验知识的要求较弱
 - C. 聚类假设和流形假设是利用未标记样本的两种常见要求
 - D. 协同训练假设数据拥有两个及以上充分且条件独立的视图

二、计算题

1. (半监督图学习) 现有 3 个有标记样本 $\{(\mathbf{x}, y)\}$ 如下:

$$\{((-1, -1), -1), ((-1, 1), -1), ((1, 0), 1)\}$$

其中 \mathbf{x} 为二维向量, $y \in \{-1, 1\}$ 。试利用半监督图学习判断未标记样本 $\{(-1, 0), (2, 0)\}$ 的标签, 其中亲和矩阵基于 k 近邻 (k 取 2) 构建, 即:

$$(W)_{ij} = \begin{cases} 1, & \text{如果 } \mathbf{x}_i \text{ 是 } \mathbf{x}_j \text{ 的 } k \text{ 近邻, 或 } \mathbf{x}_j \text{ 是 } \mathbf{x}_i \text{ 的 } k \text{ 近邻} \\ 0, & \text{否则} \end{cases}$$