

一、选择题

1. () 的主要议题是如何实现和优化机器的自我学习。
A. 机器学习 B. 人工智能 C. 深度学习 D. 人机共生
2. 目前, 机器学习领域所面临的主要挑战包括 ()。
A. 过拟合 B. 维度灾难
C. 特征工程 D. 算法的可扩展性
3. 以下描述中正确的是 ()。
A. 有监督学习常用于从数据中发现未知的模式信息
B. 当训练集中的部分样本缺少标签信息时, 通常采用半监督学习
C. 训练集中的每个示例 (Example) 均带有自己的输出值——标签 (Label) 时, 采用有监督机器学习
D. 有监督学习和无监督学习的区别在于学习过程中是否引入人的监督作用
4. 以下算法中属于有监督学习的是 ()。
A. KNN B. k-means C. 逻辑回归 D. 主成分分析
5. 以下算法中属于无监督学习的是 ()。
A. KNN B. k-means C. 逻辑回归 D. 主成分分析
6. () 主要研究的是如何协助自治 Agent 的学习活动, 进而达到选择最优动作的目的。
A. 强化学习 B. 有监督学习 C. 无监督学习 D. 半监督学习
7. 通常, 在机器学习中造成“过拟合”的主要原因有 ()。
A. 所训练出的模型过于复杂 B. 所训练出的模型过于简单
C. 特征属性太多, 但训练样本太少 D. 数据集的规模太大
8. 在机器学习中, () 用于算法选择和参数调整。
A. 训练集 B. 测试集 C. 验证集 D. 候选集
9. () 率指模型“错误地”预测了样本的类别, 样本的预测类别为“正例”。
A. TP B. FP C. TN D. FN
10. () 指一些在低维度空间上表现较好的算法很可能在高维度输入数据上效果差或效率低, 甚至不可行。
A. 算法可扩展性 B. 算法鲁棒性
C. 过拟合 D. 维度灾难



二、调研与分析题

1. 调查并对比分析机器学习领域的国际顶级会议及学术期刊。
2. 调查分析基于 Python 进行机器学习的常用第三方包及其特征。
3. 调查并对比分析机器学习开发包 TensorFlow、PyTorch 与 Weka。
4. 结合自己的专业领域，调研自己所属领域常用的机器学习方法、技术与工具。
5. 调研常用机器学习工具软件（包括开源系统），并进行对比分析。

