|  |
| --- |
| **------密--封--线--内--不要--答--题--------** |

《人工智能基础检测卷》（个性化）

学校：\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 年级：\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 姓名：\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

一、选择题：每题2分，共20分  
二、填空题：每题3分，共30分  
三、计算题：每题10分，共50分

# 一、选择题

1. 在监督学习中，下列哪个算法最适合用于分类问题？

A. 线性回归

B. K-means

C. 卷积神经网络 (CNN)

D. 支持向量机 (SVM)

2. 下列哪一项是深度学习中常用的激活函数？

A. Sigmoid

B. ReLU (Rectified Linear Unit)

C. Softmax

D. 所有选项都是

3. 在自然语言处理中，用于词向量嵌入的技术是什么？

A. Word2Vec

B. RNN (Recurrent Neural Network)

C. CNN (Convolutional Neural Network)

D. LSTM (Long Short-Term Memory)

4. 哪种算法常被用于无监督学习中的聚类任务？

A. K-means

B. AdaBoost

C. Gradient Boosting

D. Random Forest

5. 在计算机视觉中，用于目标检测的算法是什么？

A. YOLO (You Only Look Once)

B. SVM (Support Vector Machine)

C. PCA (Principal Component Analysis)

D. Naive Bayes

6. 强化学习中的Q-learning属于哪种类型的学习？

A. 监督学习

B. 非监督学习

C. 半监督学习

D. 加强学习

7. 数据挖掘过程中，下列哪一项是预处理步骤？

A. 特征选择

B. 模型训练

C. 模型评估

D. 结果解释

8. 下列哪种技术可用于处理不平衡的数据集？

A. 过采样

B. 降维

C. 主成分分析

D. K-means

9. 在深度学习中，防止过拟合的一种策略是什么？

A. Dropout

B. Batch Normalization

C. L1 正则化

D. 所有选项都是

10. NLP中的序列到序列模型（Seq2Seq）主要用于什么任务？

A. 文本分类

B. 机器翻译

C. 情感分析

D. 命名实体识别

# 二、填空题

1. 在机器学习中，\_\_\_\_\_\_\_\_\_交叉验证是一种评估模型泛化能力的方法.

2. 深度学习的\_反向传播\_\_\_\_\_\_\_\_算法用于更新神经网络中的权重.

3. 在NLP中 ,\_\_\_\_\_\_\_\_是一种衡量词在文档中重要性的指标.

4. 计算机视觉中的\_\_\_\_\_\_\_\_算法用于特征检测和匹配.

5. 强化学习的两个主要组成部分是\_\_\_\_\_\_\_\_\_\_和\_\_\_\_\_\_\_\_\_\_.

6. 数据挖掘的目标之一是发现数据中的\_\_\_\_\_\_\_\_和趋势.

7. 在监督学习中，特征工程的目的是\_\_\_\_\_\_\_\_\_\_.

8. 无监督学习中，主成分分析(PCA)用于降维.

9. 在处理文本数据时，\_\_\_\_\_\_\_模型忽略了单词顺序.

10. 对于强化学习，\_\_\_\_\_\_\_\_\_函数指导智能体采取行动.

# 三、计算题

1. 给定一个二元分类问题，使用逻辑回归模型，假设阈值为0.5，预测概率为0.65，请判断该样本的分类标签.

2. 设计一个简单的深度学习模型，用于手写数字识别，包括网络架构和损失函数的选择.

3. 使用TF-IDF计算给定文档集中关键词的权重.

4. 给定一组图像数据，使用K-means算法进行聚类，确定最佳的聚类数量.

5. 在一个强化学习环境中，计算基于Q-learning策略下智能体获得的最大累积奖励.