

浙江大学 2024—2025 学年 春夏 学期

《人工智能基础（A）》期末考试

课程号：CS1241G，开课学院： 计算机科学与技术学院

考试试卷：√ A 卷、B 卷（请在选定项上打√）

考试形式：√ 闭、开卷（请在选定项上打√），允许带 普通计算器 入场

考试日期：2025 年 6 月 20 日，考试时间： 120 分钟

诚信考试，沉着应考，杜绝违纪。

题序	一	二	三	四	五	总 分
得分						
评卷人						

一、判断题（共 20 小题，每小题 1 分，共 20 分）

- 1.1956 年达特茅斯会议是现代人工智能基础学科确立的开始。
- 2.贪心算法每步都是求解最优解，因此能够解决所有最优解问题。
- 3.弱人工智能能模拟和执行特定任务，人类智能有泛化能力。
- 4.行为主义强调通过在与环境的交互中试错来不断优化。
- 5.使用 Adam 优化器无需对学习率做指数衰减也能获得良好收敛。
- 6.决策树是无监督学习模型。
- 7.分类模型输出离散值，回归模型输出连续值。
- 8.微调大模型需要解冻并更新全部参数。
- 9.确定了卷积核的尺寸和步长就能确实卷积完的图像大小。
- 10.Softmax 函数所有类别的概率值的和为 1。
- 11.PCA 计算得到的特征向量的最大值，能解释主成分最多的方差。
- 12.CBOW 由中心词生成周边词。
- 13.语音转文字不是自然语言处理。
- 14.DeepSeek-R1 能输出思考过程，完全解决了以前大模型不可解释的问题。

- 15.当一种系统在复杂性增加临界点，会出现子系统或小规模版本中未存在的行为或特性，这种能力称为“涌现”。
- 16.BERT 是 Transformer 架构，但是只有编码器。
- 17.Transformer 架构最出色，因此其他模型可以全部淘汰。
- 18.提高大模型在专业领域的可靠性，最切实可行的方法是扩大参数规模。
- 19.用 AI 结合 CT、B 超、磁共振诊断就是多模态大语言模型。
- 20.Transformer 结构中，每个编码器之间必须保证词向量的形状相同。

二、单选题（共 20 小题，每题 1 分，共 20 分）

- 1.下列什么学派是运用神经网络的（）

A.符号主义B.联结主义C.行为主义D.进化学派
- 2.下列关于 PyTorch 的说法正确的是（）

A.PyTorch 不能自动求导B.PyTorch 只能应用于小型机器学习系统C.PyTorch 是一种开源的机器学习框架D.PyTorch 只能使用 CPU 训练
- 3.配置环境使用的工具是

A.Jupyter notebookB.GitHubC.AnacondaD.Pytorch
- 4.下列什么方法是监督学习

A.k-meanB.PCAC.垃圾邮件识别D.特征可视化
- 5.在回归分析中，常用什么损失函数

A.信息熵B.均方误差C.交叉熵D.F1 分数
- 6.下列说法错误的是

A.验证集作用是调整参数B.测试集用来检验最终选择最优的模型的性能如何C.训练集采用所有数据最好D.数据集划分避免重合
- 7.过拟合表现为

A.训练集良好，测试集不佳B.训练集不佳，测试集良好C.泛化能力强D.参数少
- 8.卷积核的数量决定了卷积神经网络的哪一个特征

A.速度B.特征图大小C.特征图数量D.感受野
- 9.“层次特征学习”描述的是下面哪个模型

- A.K—means B.决策树 C.CNN D.SVM

10. 文本分割成最小处理单元的方法被称为

- A. 词向量化 B. 分词 C. 回忆缺失 D. 命名实体识别

11. 下面程序的输出大意是选择 A(2,2)、B(4,4)为两类中心，到 A 距离小于到 B 距离归为 class1，否则归为 class2。请问(1,3)(5,1)(3,5)(6,5)哪些点在 class1

```
1. point1= [2,2]
2. point2= [4,4]
3.
4. def distance(point1, point2):
5.     return ((point2[0] - point1[0]) ** 2 + (point2[1] - point1[1]) ** 2) **
0.5
6.
7. class1 = []
8. class2 = []
9.
10. points = [[1, 3], [5, 1], [3,5], [6,5]]
11.
12. for point in points:
13.     if distance(point1, point) < distance(point2, point):
14.         class1.append(point)
15.     else:
16.         class2.append(point)
17.
```

- A. [(1,3)(5,1)] B. 回忆缺失 C. (1,3) D. [3,5]

12. RNN 不能处理长序列的主要原因是:

- A. 参数量太大 B. 维数诅咒
C. 梯度消失和梯度爆炸 D. 回忆缺失

13. “洞帘水挂水帘洞，山果花开花果山” 的词汇表和句子的 token 分别为

- A. 8, 14 B. 14, 10 C. 10, 10 D. 8, 8

14. “怒发冲冠凭栏处，_____” 正确填充了“潇潇雨歇”。GPT 自回归预测几次得到“潇潇雨歇”

- A. 1 B. 4 C. 10 D. 11

15. 下面的余弦相似度，哪一个相似度最高

- A. $\frac{\sqrt{2}}{2}$ B. $\frac{\sqrt{3}}{2}$ C. $\frac{1}{2}$ D. 0

16. “小朋友们在花园里快乐地追赶着五彩缤纷的蝴蝶”与“孩子们在公园中欢快地追逐着色彩斑斓的蝴蝶”的 Jaccard 相似度最接近以下哪个

- A. 0.32 B. 0.50 C. 0.67 D. 回忆缺失

17. 词袋模型的缺点

- A. 未登录词块 B. 不考虑顺序和没有上下文的语义信息
C. 无法区分多义词和同义词 D. 回忆缺失

18. 以下哪项是 clip 的作用和目标

- A. 把文本和图像编码成计算机可识别的 B. 加强图像输出
C. 训练图像编码器和文本编码器的参数 D. 实现图像与文本的匹配

19. 像素为 128×128 的图片，三通道，即每个像素点占三个字节大小（8 位）。现在用计算机生成 1000 张这样的图片，图片数据用 numpy 存储，np.array(uint8)总共占用计算机的内存大小是

- A. 16MB B. 24MB C. 48MB D. 72MB

20. 现在有 A[0.98,0.32,0.25]、B[0.72,0.65,0.43]和 C[0.65,0.55,0.13]，用 Jaccard 相似度判断哪两个更相似

- A. A 和 B B. B 和 C C. A 和 C D. 都一样

三、多选题（共 15 小题，每小 2 分，共 30 分）

1. 中国开发的深度学习框架有

- A. PaddlePaddle B. MindSpore C. TensorFlow D. PyTorch

2. 有监督学习的特点

- A. 一定要有数据和标签 B. 常见模式有分类和回归
C. 回忆缺失 D. 聚类和降维是典型案例

3. 强化学习的要素有

- A. 环境 B. 状态 C. 行为 D. 奖励

4. 扩散模型需要

- A. 前向加入噪声 B. 判别器 C. 后向去除噪声 D. U—NET

5. 下列什么方法是监督学习

- A. k—mean B. PCA C. 决策树 D. 回归

6. 关于大模型微调，正确的是

- A. 必须修改所有参数 B. 预训练和微调的学习率必须一致

- C. 回忆缺失 D. 专业适配
7. BP 的局限性有
- A. 局部最小值 B. 收敛速度慢
- C. 隐节点选取缺乏理论指导 D. 无法进行非线性运算
8. Transformer 的编码器和解码器都有的是
- A. MLP B. 多头注意力层 C. 跨层和归一化 D. 前馈神经网络
9. LSTM 的优势有
- A. 并行处理全部计算步 B. 缓解梯度消失和爆炸
- C. 引入门控机制 D. 参数较少
10. 下面哪些是实体命名识别任务
- A. 情感分类 B. 词性分析 C. 机器翻译 D. 地名识别
11. 下面属于多模态任务的是
- A. 视觉问答 B. 图文跨模态检索 C. 文生图 D. 图像描述及指称表达
12. 计算词相似度的方法这个是有
- A. 余弦相似度 B. Jaccard 相似度 C. Softmax D. 欧式距离
13. 下面两处可以选择的函数是
- ```
1. model = Sequential([
2. Dense(??, activation='relu', input_shape=(784,)),
3. Dense(??, activation='relu'),
4. Dense(10, activation='_____')
5.])
6.
7. model.compile(optimizer='adam',
8. loss='_____',
9. metrics=['accuracy'])
10.
```
- A. relu; binary\_crossentropy B. sigmoid; binary\_crossentropy
- C. relu; categorical\_crossentropy D. softmax; categorical\_crossentropy
14. 有一幅尺寸为  $10 \times 256$  的图像，对该图像按位与位置操作，那么得到的图像尺寸可能是
- A.  $10 \times 256$  B.  $5 \times 256$  C.  $10 \times 512$  D.  $20 \times 256$
15. 回忆缺失
- A. 回忆缺失 B. 回忆缺失 C. 回忆缺失 D. 回忆缺失

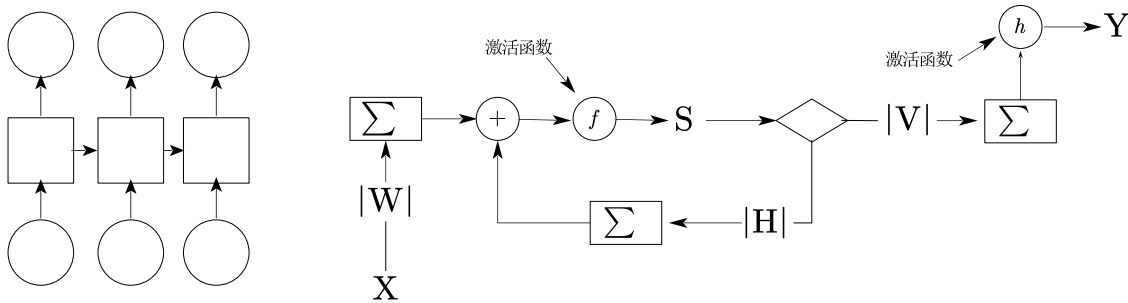
四、计算题（共 3 小题，共 15 分）

1. 尺寸为  $64 \times 64$  的三通道彩色图片，卷积核  $k=3$ ，步长  $s=1$ ，填充  $p=1$ ，卷积核个数  $N=16$ 。

- (1) 求输出图像的  $W_{out}$ 、 $H_{out}$ 。
- (2) 若有 Bias 存在，求卷积核要学习的参数数量（含有 Bias，每个卷积核包含 1 个偏置值）

2. 用梯度下降法， $f(x)=2x^2+3x+1$ ，起始点  $P(4,45)$ ，迭代 5 轮，学习率 0.1，求之后五个点，保留两位小数，按  $P_n(X_n, Y_n)$  的格式。

3. 多对多同步 RNN，输入层、隐含层（一层）、输出层的神经元均为一个， $W=[0.3, 0.1, 0.5]$ ，激活函数都是 Sigmoid， $H=[-2]$ ， $V=[2]$ ， $S_0=[0]$ ， $X=[[2, 1, 3], [1, 2, 3], [3, 1, 2]]$  无偏置，求输出序列 Y，保留两位小数（自然对数  $e=2.72$ ）



五、解答题（共 3 小题，共 15 分）

1. (4 分) 根据以下代码回答问题

```
1. optimizer = optim.Adam(model.parameters(),lr=1e-3,)
2.
3. for epoch in range(10):
4. model.train()
5. for images, labels in train_loader:
6. optimizer.zero_grad()
7. outputs = model(images.to(device))
8. loss = criterion(outputs, labels.to(device))
9. loss.backward()
10. optimizer.step()
11.
```

- (1) optimizer.zero\_grad(), loss.backward() 两行代码的作用是什么，为什么不能颠倒顺序。
- (2) 训练了几轮。
- (3) 讲讲 adam 相比于 SGD 做了哪两个更新规则上的优化。
- (4) 讲讲这份代码做了什么。

2. (4 分) 详细阐述隐空间的概念以及在人工智能方面的应用。

3. (7 分) 根据图灵测试的局限性和改进方法，人工智能伦理安全，幸存者偏差，写一篇关于人工智能面对的挑战以及你如何实现可信人工智能的方案的综合，不少于 600 字小作文。

工作留名区

试鸢、もち mochi、落拓、ajuan223、ouean、好客的山雀、冰糖葫芦慢、咻光咻