

1. 单模态表示有哪些？

单模态表示通过线性或非线性映射将单一模态信息（如图像、文本、音频等）转化为高阶语义特征。具体方法包括：

- **图像模态**：CNN特征（如ResNet、VGG提取的高层语义特征）、传统手工特征（如SIFT、HOG）；
- **文本模态**：离散词袋模型（Bag-of-Words）、稠密词向量（Word2Vec、GloVe）、上下文感知嵌入（BERT、Transformer）；
- **音频模态**：时域特征（MFCC、梅尔频谱图）、频域特征（FFT、小波变换）；
- **其他模态**：视频模态的3D-CNN特征、时序模态的RNN/LSTM特征等。

2. 数据集一般由哪些组成？分别的作用是什么？

数据集通常划分为以下三部分：

1. **训练集 (Training Set)**：用于模型参数学习（如神经网络权重更新），占比通常为60%-80%；
 2. **验证集 (Validation Set)**：用于超参数调优和模型选择（如选择最佳epoch、正则化系数），占比10%-20%；
 3. **测试集 (Test Set)**：独立评估模型泛化性能（模拟真实场景），占比10%-20%。
- 此外，**交叉验证（如k折交叉验证）**可减少数据划分偏差，尤其适用于小规模数据集。

3. 简单分类器是什么？简述其过程？

简单分类器包括k近邻（kNN）、线性分类器（如逻辑回归、SVM）等，其核心思想是基于输入特征进行快速决策：

- **kNN分类器**：
 - i. 计算测试样本与训练样本的距离（如欧氏距离、余弦相似度）；
 - ii. 选择距离最近的k个邻居；
 - iii. 投票选择多数类别作为预测结果。
- **线性分类器**（以Softmax为例）：
 - i. 输入特征 x 经线性变换得到得分：

$$s = Wx + b \quad (\text{其中 } W \in \mathbb{R}^{C \times D}, x \in \mathbb{R}^D)$$

- ii. 通过Softmax归一化得分：

$$P(y = c|x) = \frac{e^{s_c}}{\sum_{c'=1}^C e^{s_{c'}}}$$

iii. 选择最大概率类别 $\hat{y} = \arg \max_c P(y = c|x)$ 。

4. 怎么评估方法的性能？

常用评估指标及方法：

- 分类任务：
 - 准确率 (Accuracy)：

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

- 精确率 (Precision) 与召回率 (Recall)：

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad \text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

- **F1分数**：精确率与召回率的调和平均数；
 - **混淆矩阵**：展示TP、TN、FP、FN的矩阵；
 - **ROC曲线与AUC值**：衡量分类器整体性能。
- 回归任务：均方误差 (MSE)、平均绝对误差 (MAE)；
- 验证方法：留出法、交叉验证 (如5折交叉验证)、自助法 (Bootstrap)。

5. 请简述评分函数

评分函数将输入特征映射为类别得分，决定分类决策边界。常见形式：

- 线性评分函数：

$$s_c = w_c^\top x + b_c \quad (\text{对每个类别 } c \in \{1, 2, \dots, C\})$$

- 多层神经网络评分：非线性组合 (如 $s = W_2 \cdot \sigma(W_1 x + b_1) + b_2$)；
- 核方法评分：

$$s(x) = \sum_{i=1}^n \alpha_i y_i K(x, x_i) + b$$

(如SVM中使用核技巧)。

6. 请简述线性分类器

线性分类器通过线性模型 $s = Wx + b$ 计算类别得分，结合决策函数完成分类：

1. 模型形式：

$$f(x) = \text{sign}(w^\top x + b) \quad (\text{二分类}) \quad \text{或} \quad \hat{y} = \arg \max_c (Wx + b)_c \quad (\text{多分类})$$

2. **决策边界**：超平面 $w^\top x + b = 0$ ，将特征空间划分为不同区域；

3. **优化目标**：最小化损失函数（如Hinge Loss或交叉熵），通过梯度下降更新参数。

7. 请简述损失函数，并分别写出常用的几种损失函数

损失函数衡量预测值 \hat{y} 与真实值 y 的差异，指导模型优化方向。常用类型：

- **均方误差 (MSE)** (回归任务)：

$$L = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

- **交叉熵损失 (Cross-Entropy Loss)** (分类任务)：

$$L = - \sum_{i=1}^n y_i \log(\hat{y}_i)$$

- **Hinge Loss** (SVM)：

$$L = \sum_{i=1}^n \max(0, 1 - y_i \cdot \hat{y}_i)$$

- **KL散度** (分布匹配)：

$$L = \sum_{i=1}^n y_i \log \left(\frac{y_i}{\hat{y}_i} \right)$$

8. 请简述人工神经网络的过程

人工神经网络 (ANN) 通过前向传播和反向传播实现端到端学习：

1. **前向传播**：

- 输入层接收特征 $x \in \mathbb{R}^D$ ；
- 隐藏层计算：

$$h^{(l)} = \sigma \left(W^{(l)} h^{(l-1)} + b^{(l)} \right) \quad (\text{其中 } l \text{ 为层数, } \sigma \text{ 为激活函数如ReLU})$$

- 输出层生成预测值 \hat{y} ；

2. **损失计算**：通过损失函数 $L(y, \hat{y})$ 量化误差；

3. **反向传播**：

- 利用链式法则计算梯度：

$$\frac{\partial L}{\partial W^{(l)}} = \frac{\partial L}{\partial h^{(l)}} \cdot \frac{\partial h^{(l)}}{\partial W^{(l)}}$$

- 梯度下降更新参数：

$$W_{t+1}^{(l)} = W_t^{(l)} - \eta \cdot \frac{\partial L}{\partial W^{(l)}}$$

(其中 η 为学习率)。