

序号	试题内容	答案	选项A	选项B	选项C	选项D
1	在目标检测任务中，有很多的算法可以使用，以下哪个选项不是目标检测算法（）。	B	SSD	KNN	FasterR-CNN	YOLO
2	图像平滑处理的目的是（）。	A	去除噪声	模糊图像	提高图像对比度	改变图像颜色
3	你正在开发一个基于深度学习的口罩识别系统，任务是从图像中检测人脸并判断是否佩戴口罩。在数据收集与标注阶段，以下哪种做法是最合适的。	C	仅收集少量高质量标注数据，因为标注成本高	收集大量未标注数据，并使用无监督学习方法进行训练	收集尽可能多的标注数据，并确保标注质量，即使成本较高	忽略数据标注，直接使用预训练模型进行迁移学习
4	以下关于词向量的说法，正确的是（）。	B	词向量只表示单词本身，和单词的语义、语法没关联	词向量能体现语义相近的词在空间上距离较近	词向量的维度都是一样的，不会有变化	词向量没法用于中文文本处理
5	在进行模型部署时，以下哪种方式常用于将训练好的模型部署为可以通过网络接口对外提供服务的应用	B	直接把模型文件拷贝到服务器指定目录	使用 Flask 框架进行封装	将模型保存到本地磁盘就完成部署	仅用 Python 解释器运行模型代码
6	以下关于目标检测算法中模型评估指标的描述，错误的是（）。	C	平均精度均值（mAP）综合考虑了不同类别目标的检测精度，能较为全面地衡量目标检测模型在多类别目标检测任务中的整体性能表现	召回率（Recall）反映了模型正确检测出的目标数量占实际所有目标数量的比例，召回率越高说明模型遗漏的目标越少，但仅看召回率不能完全评判模型好坏	交并比（IoU）主要用于衡量预测框与真实框的重叠程度，在实际应用中只要预测框与真实框的IoU 大于 0，就可认定该目标被准确检测到了	准确率（Precision）体现的是模型预测为正例的样本中真正为正例的比例，在目标检测任务中，它与召回率相互制约，通常需要综合二者来分析模型性能
7	以下关于目标检测方法的说法，错误的是（）。	C	Faster R - CNN 是两阶段目标检测算法，先通过区域建议网络（RPN）生成候选区域，再对这些候选区域进行分类和定位，检测精度较高，但速度相对较慢	SSD（Single Shot MultiBox Detector）是单阶段目标检测算法，直接在不同尺度的特征图上进行目标检测，速度较快，但对小目标的检测效果往往不如两阶段算法	对于所有目标检测任务，只要使用更复杂、参数更多的模型，就一定能得到更好的检测结果，无需考虑数据集特点和硬件资源	目标检测中的非极大值抑制（NMS）算法用于去除重叠度较高的冗余检测框，只保留得分最高的检测框，从而提高检测结果的准确性

8	以下关于深度学习基础的数据标注相关内容，错误的是（）。	B	对于序列标注任务（如词性标注），转移概率矩阵可辅助提升标注准确性，它记录了标签间转移的统计规律	主动学习用于数据标注时，模型主动挑选最有价值样本让标注员标注，可减少标注工作量，但初始标注集需随机选大量样本	半监督学习标注数据，利用少量标注数据和大量未标注数据训练模型，未标注数据经伪标签法等转为有标签数据参与训练	元学习用于数据标注场景，可快速适应新标注任务，通过学习不同任务的共性，模型能在新任务上用少量标注样本高效学习
9	以下关于使用 LabelImg 和 X-ALabeling 等软件进行数据标注的说法，错误的是（）。	C	LabelImg 支持多种标注格式的导出，如 PASCAL VOC XML 格式和 YOLO TXT 格式，方便适配不同的深度学习框架	在使用 X-ALabeling 标注复杂场景图像时，若目标之间存在遮挡，可利用其提供的分层标注功能，将被遮挡和未被遮挡部分分别标注，提高标注的精准度	LabelImg 主要用于目标检测任务的矩形框标注，不支持对标注框添加属性信息，而 X-ALabeling 可以为标注对象添加丰富的属性	无论是 LabelImg 还是 X-ALabeling，在标注过程中如果发现之前的标注有误，都可以通过撤销操作或者直接修改标注信息来进行更正
10	在使用 LabelImg 工具进行图像标注时，以下哪种标注格式是它默认可以直接输出的	B	COCO 格式	Pascal VOC 格式	YOLO 格式	KITTI 格式
11	以下关于目标检测算法选型的说法，错误的是（）。	B	若应用场景对实时性要求极高，即使模型准确性稍低，也应优先选择运行速度快的算法	仅依据精度指标来选择算法即可，无需考虑运行速度、资源消耗和实际场景需求	当资源有限时，应选择资源消耗少、效率高的轻量级算法以适配应用场景	利用精度、召回率、F1 分数等指标评估算法，能辅助我们在选型时做出更明智的决策
12	在 Transformer 的训练过程中，以下哪种方法用于计算损失。	A	交叉熵损失	均方误差损失	对比损失	铰链损失
13	以下关于BERT的说法，正确的是（）。	C	BERT 是一种单向的语言模型，只能处理从左到右的文本序列	BERT 在预训练阶段只采用了掩码语言模型（Masked Language Model）这一种任务进行训练	BERT 能够同时捕捉文本的上下文语义信息，是双向的语言模型	BERT 在进行下游任务微调时不需要再进行任何参数调整，可直接使用预训练模型
14	<pre>import cv2 # 读取彩色图像 image = cv2.imread('test_image.jpg') # 转换为灰度图像 gray_image = cv2._____(image, cv2.COLOR_BGR2GRAY)</pre>	A	cvtColor	convertToGray	grayScale	toGray

15	在深度学习目标检测算法中，以下关于 RetinaNet 和 SSD 的说法，哪一个错误的	C	RetinaNet 和 SSD 都属于单阶段目标检测算法	RetinaNet 引入了 Focal Loss 来解决目标检测中正负样本不平衡的问题，而 SSD 没有使用 Focal Loss	SSD 通过在不同特征图上设置多个默认框（锚框）来检测不同大小的目标，RetinaNet 没有使用锚框机制	一般来说，RetinaNet 在检测小目标时的性能可能优于 SSD
16	在深度学习目标检测算法中，以下关于各类算法的描述，正确的是？	C	两阶段目标检测算法（如 Faster R - CNN）通常比单阶段目标检测算法（如 YOLOv5）检测速度快，因为两阶段算法可以更高效地筛选候选区域	基于无锚框（Anchor - free）的目标检测算法（如 CenterNet）由于不需要预设锚框，所以在检测密集小目标时通常不如基于锚框（Anchor - based）的算法（如 SSD）	基于关键点（Key - point based）的目标检测算法（如 CornerNet）主要通过检测目标的关键点（如角点）来确定目标的位置和边界框，这种方法对于目标的姿态变化和遮挡具有较好的鲁棒性	单阶段目标检测算法（如 RetinaNet）的精度普遍高于两阶段目标检测算法（如 Mask R - CNN），因为单阶段算法直接对整个图像进行预测，能够获取更全面的信息
17	在强化学习中，“探索”的主要目的是（）。	B	只选择当前已知的最优动作，以获取最大奖励	尝试不同的动作，去发现可能带来更高奖励的新行为	按照固定顺序依次执行所有动作	避免执行任何可能带来负奖励的动作
18	在经典物体识别卷积神经网络中，ResNet（残差网络）提出“残差连接”的主要目的是（）。	B	降低网络的计算复杂度，加快运算速度	使网络更容易收敛，解决深度网络训练时的梯度消失或梯度爆炸问题	增加网络对不同尺寸输入图像的适应性	直接提升网络对物体的分类准确率
19	以下哪一项是模型优化过程中常用的方法	D	增加训练数据量	减少模型复杂度	调整学习率	以上都是
20	以下关于模型评估的叙述中，哪一项是错误的？	A	在分类任务中，准确率（Accuracy）是评估模型性能的唯一可靠指标	交叉验证（Cross - Validation）可以有效减少模型评估中的方差，提高评估结果的稳定性	在类别不平衡的数据集中，F1 分数比准确率更能反映模型的性能	混淆矩阵（Confusion Matrix）可以帮助分析模型在不同类别上的表现
21	以下关于模型优化与评估的叙述中，哪一项是正确的？	C	在神经网络中，批量归一化（Batch Normalization）的主要作用是减少模型的训练时间，但对模型的泛化能力没有影响	在超参数调优中，网格搜索（Grid Search）比随机搜索（Random Search）更高效，尤其是在超参数空间较大时	在模型评估中，AUC - ROC 曲线下的面积越大，说明模型在不同阈值下的分类性能越好	使用更多的隐藏层和神经元总是可以提高模型的性能，因为这会增加模型的表达能力
22	梯度下降算法中，损失函数曲面上轨迹最混乱的算法是以下哪种算法？	A	SGD	BGD	MGD	MBGD

23	在图像数据清洗中，以下哪一项操作通常用于解决图像尺寸不一致的问题？	B	对图像进行直方图均衡化	使用图像插值方法（如双线性插值）调整尺寸	对图像进行边缘检测	对图像进行颜色空间转换
24	关于循环神经网络中“一对多”结构，以下正确的是？	B	常用于图像生成，输入图像特征，输出多个不同风格图像	适用于文本生成，输入主题，输出多个语句，输出有先后顺序依赖	多用于语音识别，输入语音片段，输出多个独立语音特征	应用在视频分类，输入视频帧，输出多个不同类别标签且无关联
25	在人脸识别的机器学习项目中，关于特征工程与数据预处理，以下说法正确的是？	B	数据预处理只需将图像裁剪成统一大小，特征工程只提取人脸肤色特征	数据预处理要去除损坏图像，特征工程提取人脸五官位置特征	数据预处理先划分训练集和测试集再去除噪声，特征工程只考虑人脸脸型特征	数据预处理不用处理图像噪声，特征工程只提取人脸轮廓特征
26	以下哪种数据标注任务可能需要更多的专业知识？	B	简单图像分类	医疗影像标注	普通文本情感标注	道路标识标注
27	( ) 包括数据质量控制和数据治理	D	数据采集	数据清洗	数据质量检查	数据集成
28	以下关于常见回归算法的描述中，错误的是？	B	线性回归用最小二乘法拟合数据，适合自变量和因变量呈线性关系的场景	岭回归是通过加 L1 正则化来避免过拟合的	决策树回归可以处理非线性的数据	支持向量回归 (SVR) 允许部分数据点偏离预测超平面
29	以下关于深度学习中常见目标检测算法的描述，错误的是？	C	R-CNN系列算法先通过选择性搜索等方法生成候选区域，再对这些区域进行分类和位置回归	Fast R-CNN 在 R-CNN 基础上进行了改进，它将整张图像送入卷积神经网络提取特征，然后针对每个候选区域在特征图上进行操作，提升了检测速度	Faster R-CNN 进一步优化，通过引入区域生成网络 (RPN) 来自动生成候选区域，取代了之前的选择性搜索等传统方法，但它的检测速度依然很慢，很难应用于实时场景	YOLO (You Only Look Once) 系列算法摒弃了分区域生成候选框再检测的思路，而是将目标检测任务当作回归问题，直接从图像中预测出目标的类别和位置，检测速度较快，可用于实时检测
30	以下关于个性化分析方法中的聚类分析，说法错误的是？	C	聚类分析可以将用户或物品划分成不同的组，同一组内的成员具有较高的相似性	K - 均值 (K - Means) 聚类是一种常用的聚类算法，它需要预先指定聚类的数量 K	聚类分析的结果可以直接用于个性化推荐，无需考虑其他因素	层次聚类可以分为凝聚式层次聚类和分裂式层次聚类两种类型
31	在自然语言处理中，以下关于注意力机制的说法，错误的是？	C	注意力机制可以让模型在处理序列数据时，自动地关注到序列中不同部分的重要性	自注意力机制是注意力机制的一种特殊形式，它计算序列中每个元素与其他所有元素之间的关联	在多头注意力机制中，多个注意力头的计算结果是完全相同的，只是为了增加模型的计算量	注意力机制可以与卷积神经网络 (CNN) 或循环神经网络 (RNN) 结合使用，提升模型的性能
32	以下关于注意力机制在图像领域应用的描述，正确的是？	C	注意力机制在图像领域只能应用于卷积神经网络，无法与其他类型的网络结合	空间注意力机制只关注图像的通道信息，不考虑像素的空间位置	通道注意力机制通过计算不同通道之间的相关性，为每个通道分配不同的权重	在图像分类任务中使用注意力机制，会增加模型的计算复杂度，但不会提升模型的准确率

33	以下关于图像分割算法的描述，正确的是？	C	K-Means 聚类算法能精准分割复杂形状物体	阈值分割法适用于所有图像分割	基于CNN的方法（如 U-Net）对医学影像分割效果好	区域生长法无需考虑条件就能快速分割
34	以下关于聚类方法中距离度量方式选择的描述，正确的是？	B	K-Means 聚类只能用欧氏距离，其他距离都不适用	文本数据聚类用余弦相似度，比欧氏距离更合适	层次聚类计算簇间距离，统一用曼哈顿距离效果最好	DBSCAN 算法只能用切比雪夫距离，才能识别不同密度区域
35	在图像生成任务中运用注意力机制，下列表述正确的是？	D	注意力机制在图像生成里仅能关注图像整体特征，无法聚焦局部细节	引入注意力机制后，图像生成模型训练时间会大幅缩短，且不会影响生成图像质量	生成对抗网络（GAN）结合注意力机制时，通常只在生成器中使用，判别器用不到	自注意力机制应用于图像生成时，能让模型学习到图像中不同区域之间的长距离依赖关系
36	在图像分类中，以下关于模型评估指标的说法正确的有？	ABCD	准确率（Accuracy）是指分类正确的样本数占总样本数的比例，适用于各类别样本分布均匀的情况	精确率（Precision）关注的是预测为正类的样本中实际为正类的比例，在处理类别不平衡问题时比准确率更有参考价值	召回率（Recall）是指实际为正类的样本中被正确预测为正类的比例，在需要尽可能找出所有正类样本的场景中很重要	F1 - score 是精确率和召回率的调和平均数，它综合考虑了精确率和召回率，能更全面地评估模型性能
37	以下关于聚类方法的说法，正确的有？	ABCD	K - 均值（K - Means）聚类算法对初始聚类中心的选择比较敏感，不同的初始中心可能导致不同的聚类结果	层次聚类可以生成聚类的层次结构，分为凝聚式和分裂式两种，凝聚式是从单个样本开始逐步合并成大的簇	DBSCAN 聚类算法是基于密度的聚类方法，能够发现任意形状的簇，并且可以识别出噪声点	高斯混合模型（GMM）聚类假设数据是由多个高斯分布混合而成，通过估计每个高斯分布的参数来进行聚类
38	以下关于卷积神经网络（CNN）的说法，正确的有？	ABC	卷积层通过卷积核在输入数据上滑动进行卷积操作，能够自动提取数据的局部特征	池化层的主要作用是对特征图进行下采样，减少数据量和计算量，同时增强特征的平移不变性	全连接层将前面卷积层和池化层提取的特征进行整合，实现从特征到具体类别的映射	批归一化层在卷积神经网络中可以加速模型收敛，减少内部协变量偏移问题，但会增加过拟合的风险
39	以下关于模型优化与评估的表述，正确的有？	AC	对于线性回归模型，若发现残差呈现出明显的规律性，说明模型可能存在设定误差，需要进一步调整模型	在使用 Dropout 进行模型优化时，它只能应用于全连接层，对卷积层无效	信息增益率是决策树算法中用于特征选择的指标，相较于信息增益，它能更好地处理具有较多取值的特征	在评估模型时，若使用分层抽样来划分训练集和测试集，对于类别分布不均衡的数据集，可能会导致测试集无法准确反映真实的类别分布情况
40	以下关于模型训练的说法，正确的有？	ABD	在训练深度学习模型时，使用批量归一化层可以加快收敛速度，并且在一定程度上缓解过拟合问题	训练循环神经网络（RNN）时，梯度消失或梯度爆炸问题可以通过使用长短时记忆网络（LSTM）或门控循环单元（GRU）来缓解	对于大规模数据集，为了提高训练效率，通常采用全量数据进行梯度计算，即每次更新模型参数时使用整个数据集	学习率调度策略（如学习率衰减）可以根据训练的轮数动态调整学习率，有助于模型更好地收敛到最优解

41	关于模型评估中的交叉验证，以下哪些说法是正确的？	ABD	交叉验证可以减少模型评估的方差	K折交叉验证将数据集分为K个子集，依次使用其中一个子集作为验证集，其余作为训练集	交叉验证只适用于小数据集，对大数据集没有意义	留一法交叉验证是K折交叉验证的一种特例，其中K等于数据集的大小
42	关于模型评估中的混淆矩阵，以下哪些说法是正确的？	ABD	混淆矩阵用于描述分类模型的预测结果与真实标签之间的关系	混淆矩阵可以用于计算精确率、召回率和F1 Score	混淆矩阵只适用于二分类问题	混淆矩阵中的对角线元素表示模型预测正确的样本数。
43	以下关于回归算法的说法，正确的有？	ABC	线性回归假设自变量和因变量之间存在线性关系，其目标是找到一组最优的系数，使得预测值与真实值之间的误差平方和最小	岭回归是在线性回归的基础上加入了 L2 正则化项，它可以有效解决特征之间存在多重共线性的问题	多项式回归通过增加自变量的高次项来拟合数据，能够处理自变量和因变量之间的非线性关系，但容易出现过拟合	支持向量回归（SVR）只能处理线性数据，无法拟合非线性的回归任务
44	以下关于特征工程与数据预处理的说法，正确的有？	ABD	对于偏态分布的数值型特征，进行对数变换可以使其分布更接近正态分布，有利于一些对数据分布有要求的模型	当处理高维稀疏数据时，主成分分析（PCA）是一种有效的降维方法，但它会损失部分信息	数据清洗时，对于异常值的处理，只能采用删除的方式，以保证数据的质量	对于时间序列数据，提取诸如移动平均值、季节性等特征，有助于模型更好地捕捉数据的变化规律
45	以下哪些是语义分割的典型应用？	AB	自动驾驶中的道路场景理解	医学图像中的器官分割	图像风格迁移	视频目标跟踪
46	以下哪些是 Transformer 模型的主要组成部分？	AB	编码器	解码器	卷积层	循环层
47	以下关于层次聚类算法的描述正确的有（ ）	ABCD	分为凝聚式层次聚类和分裂式层次聚类	凝聚式层次聚类是从每个数据点作为单独的类开始，不断合并相似的类	分裂式层次聚类是从所有数据点属于一个类开始，不断分裂出不同的类	最终的聚类结果可以通过绘制树状图（Dendrogram）来直观呈现并辅助选择合适的聚类数量
48	在模型选择与训练阶段，以下需要考虑的因素有（ ）	ABCD	任务类型	数据规模	计算资源	模型复杂度
49	在划分训练集和测试集时，以下可用于确保数据分布一致性的方法有（ ）	AB	随机打乱后按固定比例划分	分层抽样	对数据排序后按顺序划分	基于数据哈希值划分

50	以下关于人脸识别模型训练相关要点的描述，正确的有？	ABCD	训练数据集的质量至关重要，需要包含大量不同场景（如不同光照、角度、人员类型等）下的人脸图像，以保证模型能学习到全面的特征	为了提升模型的泛化能力，在训练过程中可以采用交叉验证的方法，将数据集合理划分成多个子集，轮流作为训练集、验证集和测试集，优化模型参数	选择合适的损失函数很关键，比如对于人脸识别的分类任务（区分不同身份的人脸），可以选用交叉熵损失函数来衡量模型预测结果与真实标签之间的差异	训练时设置合适的学习率很有必要，学习率过大可能导致模型无法收敛，学习率过小则会使训练过程过于缓慢，难以在合理时间内达到较好的训练效果
----	---------------------------	------	--	--	--	--