

序 列	试题 类型	考点	试题内容	选项A	选项B	选项C	选项D
1	单选题	sklearn相关	<p>下面的代码是使用决策树对鸢尾花数据进行分类的代码，请根据代码选择正确的选项:</p> <pre>```# 第一步: 导入必要的库 from sklearn.datasets import load_iris # 导入鸢尾花数据集 from sklearn.model_selection import train_test_split # 用于划分训练集和测试集 from sklearn.tree import DecisionTreeClassifier # 导入决策树分类器 from sklearn.metrics import accuracy_score, confusion_matrix # 用于评估模型性能 # 第二步: 加载鸢尾花数据集 # 这里我们使用 scikit-learn 自带的鸢尾花数据集 iris = ____ (题目1) ____ # 输出数据集的描述信息, 帮助我们了解数据 print(iris.DESCR) # 特征变量 (花萼长度、花萼宽度、花瓣长度、花瓣宽度) X = iris.data # 目标变量 (鸢尾花的类别: 0 -&gt; Setosa, 1 -&gt; Versicolor, 2 -&gt; Virginica) y = iris.target # 打印特征和目标的基本信息 print("特征数据集的形状:", X.shape) # (150, 4) 表示150个样本, 每个样本有4个特征 print("目标数据集的形状:", y.shape) # (150,) 表示150个样本的类别标签 # 第三步: 划分训练集和测试集 # 我们将数据集按80%训练、20%测试的比例划分 X_train, X_test, y_train, y_test = ____ (题目2) # 打印训练集和测试集的大小 print("训练集的大小:", X_train.shape) print("测试集的大小:", X_test.shape) # 第四步: 训练决策树分类器 # 使用 scikit-learn 中的 DecisionTreeClassifier 来构建模型 clf = DecisionTreeClassifier(____ (题目3)) ____ # 使用训练数据训练模型 clf.fit(X_train, y_train) # 第五步: 模型预测 # 使用测试集进行预测 y_pred = ____ (题目4) # 打印测试集的真实标签和预测结果 print("测试集的真实标签:") print(y_test) print("预测结果:") print(y_pred) # 第六步: 模型评估 # 计算模型的准确率 accuracy = ____ (题目5) print("模型的准确率:") {:.2f}%.format(accuracy * 100)) # 计算并输出混淆矩阵 conf_matrix = confusion_matrix(y_test, y_pred) print("混淆矩阵:\n", conf_matrix);</pre>	pre_train=True	pretrain=True	random_state=42	get_state=42

2	单选题	sklearn相关	<p>下面的代码是使用决策树对鸢尾花数据进行分类的代码，请根据代码选择正确的选项:</p> <pre>```# 第一步: 导入必要的库 from sklearn.datasets import load_iris # 导入鸢尾花数据集 from sklearn.model_selection import train_test_split # 用于划分训练集和测试集 from sklearn.tree import DecisionTreeClassifier # 导入决策树分类器 from sklearn.metrics import accuracy_score, confusion_matrix # 用于评估模型性能 # 第二步: 加载鸢尾花数据集 # 这里我们使用 scikit-learn 自带的鸢尾花数据集 iris = ____ (题目1) ____ # 输出数据集的描述信息, 帮助我们了解数据 print(iris.DESCR) # 特征变量 (花萼长度、花萼宽度、花瓣长度、花瓣宽度) X = iris.data # 目标变量 (鸢尾花的类别: 0 -&gt; Setosa, 1 -&gt; Versicolor, 2 -&gt; Virginica) y = iris.target # 打印特征和目标的基本信息 print("特征数据集的形 状:", X.shape) # (150, 4) 表示150个样本, 每个样本有4个特征 print("目标数据集的形状:", y.shape) # (150,) 表示150个样本的类别标签 # 第三步: 划分训练集和测试集 # 我们将数据集按80%训练、20%测试的比例划分 X_train, X_test, y_train, y_test = ____ (题目2) ____ # 打印训练集和测试集的大小 print("训练集的大小:", X_train.shape) print("测试集的大小:", X_test.shape) # 第四步: 训练决策树分类器 # 使用 scikit-learn 中的 DecisionTreeClassifier 来构建模型 clf = DecisionTreeClassifier(____ (题目3)) ____ # 使用训练数据训练模型 clf.fit(X_train, y_train) # 第五步: 模型预测 # 使用测试集进行预测 y_pred = ____ (题目4) ____ # 打印测试集的真实标签和预测结果 print("测试集的真实标签:", y_test) print("预测结果:", y_pred) # 第六步: 模型评估 # 计算模型的准确率 accuracy = ____ (题目5) ____ print("模型的准确率: {:.2f}%" .format(accuracy * 100)) # 计算并输出混淆矩阵 conf_matrix = confusion_matrix(y_test, y_pred) print("混淆矩阵:\n", conf_matrix)</pre>	load_iris()	get_iris_data()	load_data_iris()	load_iris()
3	单选题	Pytorch相关	如果你想冻结模型中的一部分层的参数，使它们不再更新，应该使用哪个操作？	param.requires_grad = False	torch.no_grad()	optimizer.step()	param.data.zero_()
4	单选题	Pytorch相关	下面哪一个方法是用于保存 PyTorch 模型的参数？	torch.load()	torch.save()	torch.manual_seed()	torch.optim.Adam()
5	单选题	Pytorch相关	如果你想为 GPU 上的张量创建一个与其相同形状但值全为零的张量，应该使用哪个函数？	torch.zeros_like()	torch.ones_like()	torch.full()	torch.randn()
6	单选题	Pytorch相关	下面哪个方法用于将模型从训练模式切换到评估模式？	optimizer.step()	model.train()	model.eval()	torch.save()
7	多选题	机器学习算法相关	关于树结构模型数据归一化问题，下列说法不正确的是：	当样本不同特征之间取值范围相差过大的时候，应当对数据进行归一化处理	需要根据实际情况判断是否对数据归一化	树结构模型是阶跃的，阶跃点是不可导的，因此无需对数据进行归一化	如果算法发生不收敛的情况，即使是树结构模型也需要对数据进行归一化

8	多选题	机器学习算法相关	关于集成学习，下列说法正确的是：	集成学习中的并行学习方法是指算法的个体学习器之间存在强依赖关系 集成学习中的串行学习方法是指算法的个体学习器之间不存在强依赖关系	集成学习中的串行学习方法是指算法的个体学习器之间不存在强依赖关系	Boosting属于串行集成学习方法	随机森林属于并行集成学习方法
9	多选题	机器学习算法相关	关于K-Means算法，下列说法正确的是	K-Means 算法对大数据集有较高的效率并且有可伸缩性	K-Means 算法是一种无监督学习算法	K-Means 算法中的K值需要人工设定，且初始聚类中心随机选择	K-Means 算法中初始聚类中心的选择对聚类结果影响不大
10	多选题	深度学习算法相关	关于BP算法，下列说法正确的是：	BP算法不能用于处理非线性问题	通常情况下，BP算法所需的训练时间较长	BP算法容易陷入局部最优解	使用BP算法进行优化时有可能由于权重调整过大使得激活函数失活
11	多选题	人工智能基础知识	下列场景中，对哪些场景的分析使用了正确的方法？	根据商家最近几年的经营及服务数据，使用聚类算法判断出淘宝商家在各自主营类目下所属上架的层级	根据商家最近几年的成交数据，使用聚类算法预测出用户未来一个月内可能的消费金额	销售笔记本电脑的商家，使用关联规则方法预测可能购买打印机的顾客	根据用户的最近的购买信息，使用决策树算法识别出买家的性别
12	多选题	机器学习算法相关	下列哪些函数可以作为线性不可分支持向量机的核函数？	线性核函数	多项式核函数	径向基核函数	sigmoid核函数
13	多选题	机器学习算法相关	主成分分析获取超平面应具有的性质有：	最大可分性	局部极小性	信息增益最大	最近重构性
14	多选题	机器学习算法相关	下列属于降维算法的有：	主成分分析	奇异值分解	数据离散化	高斯核函数

15	多选题	ai校赛	以下关于层次聚类算法的描述正确的有（ ）	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
16	多选题	Python基础知识	下列哪些python语句是合法的？	$x = y = z = 1$	$x, y = y, x$	$x = (y = z = 1)$	$x += y$
17	多选题	机器学习算法相关	在神经网络中，下面哪种方法可以用来处理过拟合现象？	分批归一化	正则化	Dropout	最小二乘法
18	多选题	机器学习算法相关	可用于数据特征降维的方法有哪些？	SVD	PCA	AutoEncoder	LDA
19	多选题	机器学习算法相关	下列哪些情况可能会导致神经网络训练失败？	梯度消失	梯度下降	鞍点	激活函数死区
20	多选题	机器学习算法相关	关于数据降维，以下说法正确的是：	降维是将训练样本的数据从高维空间转换到低维空间	降维不会对数据产生损伤	通过将为可以更加有效地发觉有意义的数据结构	降维有助于实现数据可视化
21	多选题	机器学习算法相关	在决策树的构建过程中，下列哪些函数可以用于节点划分的衡量指标？	信息增益	信息增益率	基尼系数	核函数
22	多选题	Numpy、Pandas相关	下列关于pandas数据读/写的说法中，正确的是：	read_csv函数可以读取文本格式的结构化数据	read_sql函数可以读取数据库的数据	to_csv函数可以将结构化数据写入csv文件	to_excel函数可以将结构化数据写入Excel文件
23	多选题	深度学习算法相关	卷积神经网络中，关于卷积输入层的通道数、卷积输出层的通道数以及卷积核的通道数，下列说法正确的是：	卷积输入层的通道数与卷积输出层的通道数相等	卷积输入层的通道数与卷积核的通道数相等	卷积输出层的通道数与卷积核的通道数相等	卷积输出层的通道数与卷积核的个数相等

24	多选题	机器学习算法相关	关于逻辑回归（LR）和支持向量机（SVM），下列说法正确的是：	两种算法的损失函数不同，但是目的都是增加对分类影响较大的数据点的权重，减少与分类关系较小的数据点的权重。	两种算法解决的问题不同，逻辑回归是解决回归算法，支持向量机是分类算法	支持向量机可以解决的问题逻辑回归也可以解决，反之不一定	逻辑回归可以解决的问题支持向量机也可以解决，反之不一定
25	多选题	深度学习算法相关	关于seq2seq模型，下列说法正确的是：	seq2seq包含编码器和解码器两部分	seq2seq可用于文本摘要任务	可以将seq2seq理解为一种NxM模型	可以使用attention机制对seq2seq进行优化
26	多选题	人工智能基础知识	使用OneHot（独热码）编码来表示文本，下列说法正确的是：	可以用较低维数来表示高维稀疏数据	会浪费存储空间	不同词的向量正交，无法衡量不同词之间的关系	可以根据两个词在文本中出现的先后顺序进行语义分析
27	多选题	人工智能基础知识	常见的人工智能加速硬件包括：	CPU	GPU	NPU	TPU
28	多选题	人工智能基础知识	关于标准正态分布，正确的是：	均值为1	均值为0	方差为1	方差为0
29	多选题	机器学习算法相关	数据预处理中清除数据集中的“脏”数据，这里的“脏”数据是指：	存储数据的设备受到某些污染物的影响	数据包含错误的记录或者异常点	数据中存在矛盾、有差异的记录	一些数据缺失某些特征或者包含一些缺失的值
30	多选题	Python基础知识	假设 <code>data = (1, 3, 3, 7, 9, 11)</code> ，下列哪些操作是合法的？	<code>data[1 : -1]</code>	<code>data[1 : 7]</code>	<code>list(data)</code>	<code>data * 3</code>
31	多选题	Python基础知识	在python中，哪下列哪个表达式可以创建一个包含10个0的一维数组？	<code>np.zeros(10)</code>	<code>np.zeros(10, )</code>	<code>np.zeros((10))</code>	<code>np.array([0.] * 10)</code>
32	多选题	人工智能基础知识	与机器学习相比，深度学习的主要特点在于：	深度学习对算力的需求小于机器学习。	深度学习可以自动从数据中提取特征。	深度学习可以模拟更加复杂的函数。	通过深度学习算法训练得到的模型更容易被人类所理解。

33	单选题	Numpy、Pandas相关	代码 np.nonzero([1, 0, 2, 4, 3, 0]) 的返回值为：	[1, 2, 4, 3]	[0, 2, 3, 4]	[True, False, True, True, True, False]	[True, 0, True, True, True, True, 0]
34	单选题	Numpy、Pandas相关	代码 df.sort_values(by=['age', 'visits'], ascending=[False, True]) 的运行结果是：	将 DataFrame中的数据先按照 visits降序排列，然后按照 age升序排列	将 DataFrame中的数据先按照 visits升序排列，然后按照 age降序排列	将 DataFrame中的数据先按照 age升序排列，然后按照 visits降序排列	将 DataFrame中的数据先按照 age升序排列，然后按照 visits降序排列
35	单选题	Numpy、Pandas相关	在一个DataFrame中，包含整数格式的age列，代码 df[df.age>3] 运行的结果是：	取出age列中所有大于3的行	取出age列中所有大于3的行的标号	返回一个布尔值的列表，其中age列大于3的行对应的值为True	判断一行age的值是否大于3
36	单选题	Numpy、Pandas相关	关于Pandas中的head函数，下列说法正确的是：	可以默认打印 DataFrame的表头数据	可以默认打印 DataFrame中前5条数据	可以默认打印 DataFrame中前10条数据	可以默认打印 DataFrame的描述信息
37	单选题	Numpy、Pandas相关	下列哪一项可以实现0-9元素的逆序输出？	np.arange(10) [::-1]	np.arange(10) [:-1]	np.arange(10) [1:-1]	np.arange(10) [1:-1]
38	单选题	Numpy、Pandas相关	在Numpy中，哪个函数可以指定创建数组的形状，并且可以用随机数填充这个指定形状的数组？	np.random.random	np.random.random	np.random.rand	np.random.randint
39	单选题	Numpy、Pandas相关	如何获取数组的形状、大小和维度？	arr.shape, arr.size, arr.dim	arr.shape, arr.size, arr.length, arr.ndim	arr.shape, arr.size, arr.length, arr.ndim	arr.size, arr.shape, arr.dim
40	单选题	Python基础知识	下列表达式中，哪个的值一定为True	0.3 == 3 * 0.1	0.3 equals 3 * 0.1	math.equals(0.1, 3 * 0.1)	math.isclose(0.3, 3 * 0.1)

41	单选题	Python基础知识	给定两个变量a、b，要求输出逻辑是：如果a大于b，则返回a-b，否则返回a+b。请问下列哪项可以满足上述逻辑：	(a > b) if (a - b) else (a + b)	(a > b) if (a + b) else (a - b)	&nbsp;(a - b) if (a > b) else (a + b)	&nbsp;
42	单选题	Python基础知识	下列哪个属于可变数据类型：	str	tuple	list	int
43	单选题	深度学习算法相关	以下哪一选项是训练静态词向量的方法？	Fasttext	ELMO	BERT	GPT
44	单选题	深度学习算法相关	1997年，IBM开发的一台超级电脑打败了前世界象棋冠军加里·卡斯帕罗夫，以下哪项是这台计算机的名字？	Deep Mind	Deep Blue	Deep Green	Deep Thinking
45	单选题	深度学习算法相关	在数据较少的情况下，可使用迁移学习完成对图片进行图像识别的任务，下列哪种方法是合适的？	使用现有的数据重新训练整个模型	冻结模型最后几层，只对前面的部分进行调参	冻结模型前面的部分，只对模型最后几层进行调参	对模型的每一层进行评估，只选择其中的少数几层
46	单选题	深度学习算法相关	下列哪一种架构包含有反馈连接？	全连接神经网络	卷积神经网络	循环神经网络	受限玻尔兹曼机
47	单选题	深度学习算法相关	分批归一化的好处都有啥？	是一种非常有效的反向传播算法	是一种非常有效的优化算法	可以将权重分布调整为正态分布	可以大致固定每一层输入的范围
48	单选题	深度学习算法相关	卷积神经网络中Dropout层的作用是：	加快收敛速度	防止过拟合	丰富训练样本	增加正样本强度
49	单选题	深度学习算法相关	关于生成式对抗网络中的两个模型之间的关系，可以用下列哪一对来对比？	组长与组员	警察与小偷	教师与学生	医生与护士
50	单选题	深度学习算法相关	在神经网络中，在特定神经元得到的输出为-0.321，则此神经网络可能使用了哪种激活函数？	ReLU	Sigmoid	tanh	softmax
51	单选题	深度学习算法相关	与传统机器学习方法相比，深度学习的优势在于：	深度学习可以自动学习特征	深度学习完全不需要数据预处理	深度学习不需要调参	以上三条都正确
52	单选题	深度学习算法相关	对于一个图像识别问题，下面那种神经网络可以更好地解决问题？	感知器	多层感知器	循环神经网络	卷积神经网络
53	单选题	深度学习算法相关	当在卷积神经网络中加入池化层时（pooling layer），变换的不变性会被保留吗？	无法确定	需视实际情况而定	会保留	不会保留
54	单选题	ai校赛	( ) 包括数据质量控制和数据治理	数据采集	数据清洗	数据质量检查	数据集成
55	单选题	人工智能基础知识	下列哪一项不属于数据清洗的范畴？	将数据分批次	填补缺失值	删除离群点	解决数据的不一致性
56	单选题	人工智能基础知识	对机器学习性能的评估指的是：	对算法学习能力的评估	对算法学习方法的评估	对算法学习效率的评估	对查准率和查全率的评估

57	单选题	人工智能基础知识	下列方法中，哪个无助于防止过拟合？	L1/L2正则化	Dropout	Batch Normalization	SGD
58	单选题	人工智能基础知识	超市通过对顾客购物小票的分析，得到购买啤酒的人有很大的概率也会购买尿不湿，他们使用了哪种类型的算法？	分类算法	聚类算法	关联规则算法	自然语言处理技术
59	单选题	人工智能基础知识	假设X核Y都服从正态分布，那么 $P(X < 5, Y < 0)$ 是一个：	先验概率	后验概率	联合概率	几何概率
60	单选题	人工智能基础知识	下列哪种方法可能改善模型的欠拟合现象？	从数据集中抽样进行训练	对模型进行裁剪	增加训练过程的迭代次数	使用L1或者L2正则化
61	单选题	人工智能基础知识	“均方根误差”的简写是：	MAE	RMAE	MSE	RMSE
62	单选题	人工智能基础知识	对于多层感知器（MLP），输入层节点数为10，隐藏层节点数为5，从输入层到隐藏层的最大链接数是多少？	小于50	等于50	大于50	无法确定
63	单选题	人工智能基础知识	下列哪个函数不应作为神经网络的激活函数？	$y = \sin(x)$	$y = \tanh(x)$	$y = \max(0, x)$	$y = 2x + 3$
64	单选题	人工智能基础知识	下列哪种算法不属于分类算法：	决策树算法	逻辑回归算法	支持向量机算法	K均值算法
65	单选题	ai校赛	以下关于BERT的说法，正确的是（）。	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
66	单选题	Numpy、Pandas相关	运行下列代码，结果中A的值是？ <code>import pandas as pddf = pd.DataFrame(['A', 'B', 'C', 'D', 'E', 'F'])sum(df.drop_duplicates())</code>	22	15	8	7
67	单选题	Numpy、Pandas相关	关于下列代码运行结果描述中，正确的是： <code>import numpy as npdata_01 = np.arange(6).reshape(3, 2, 1, 1).squeeze()data_02 = np.arange(6).reshape(3, 1, 2, 1).squeeze()</code>	data_01 和 data_02 的形状和数据完全一样。	data_01 和 data_02 的形状不一样但数据不一样。	data_01 和 data_02 的形状不一样和数据一样。	data_01 和 data_02 的形状和数据均不一样。
68	单选题	Python基础知识	以下代码的运行结果是： <code>for i in range(1, 3): print(i)for i in range(2): print(i)</code>	1 3 2 	1 2 0 1 	1 3 0 1 	1 3 0 2 
69	单选题	Python基础知识	已有数据： <code>data = [1, 2, 3, 4, 5, 4, 3, 2, 1]</code> ，下列哪行代码可以去除data中的重复数据？	<code>result = [d for d in list(data)]&lt;br/&gt;</code>	<code>result = [d for d in set(data)]</code>	<code>result = [d for d in dict(data)]</code>	以上都不对
70	单选题	Python基础知识	下列程序的运行结果是： <code>data = [ i for i in range(1, 7) ] data[-1:1:-1]</code>	<code>python [-1, 1, -1]</code>	<code>python [6, 5, 4, 3]</code>	<code>python [1, 2, 3, 4]</code>	程序报错

71	单选题	Python基础知识	下列程序的运行结果是： data = {'name': '张三'} new_data = {'gender': '男'} data.update(new_data) print(data)	{'name': '张三', 'gender': '男'}	{'name': '张三', 'gender': '男'}	{'name': '张三', 'gender': '男'}	程序报错
----	-----	------------	------------------------------------------------------------------------------------------------------------	-------------------------------	-------------------------------	-------------------------------	------