

1. 下面（ ）方法提出来，用来描述对计算机智能水平进行测试。
A. 香农定理 B. 费马定理 C. 摩尔定律 D. **图灵测试**
2. 下列哪种情况是图灵测试的内容？
A. 两机对话，其中一机的智力超过另一机时，说明智者机器通过了图灵测试
B. 当机器与人对话，两者互相询问，人分不清机器是人还是机器，说明它通过了图灵测试
C. 当人与人对话，其中一人的智力超过另一人时，说明智者通过了图灵测试
D. **当机器骗过测试者，使得询问者分不清是人还是机器时，说明它通过了图灵测试**
3. 人工智能诞生于（ ）年
A. **1956** B. 1957 C. 1965 D. 1955
4. 图灵测试用来测试计算机系统的哪方面能力？
A. 存储能力 B. 网络能力 C. **智能能力** D. 计算能力
5. 以下关于人工智能概念的描述中，错误的是
A. 人工智能是对人的意识、思维过程进行模拟
B. 人工智能英文翻译是 Artificial Intelligence
C. 人工智能使计算机能实现更高层次应用
D. **人工智能就是机器学习**
6. 我们日常生活中的上下班，指纹扫描打卡考勤、虹膜扫描，应用到了人工智能中的（ ）。
A. 机器翻译 B. 虚拟现实 C. 机器翻译技术 D. **模式识别**
7. 不属于人工智能的学派是
A. 连接主义 B. **机会主义** C. 符号主义 D. 行为主义
8. AI 的英文全拼是
A. Automatic Information B. **Artificial Intelligence**
C. Artificial Information D. Automatic Intelligence
9. 图灵实验是为了判断一台机器是否具备智能的实验。✓
10. 下列不是知识表示法的是（ ）
A. 谓词表示法 B. 产生式表示法 C. 状态空间表示法 D. **计算机表示法**
11. 确定性知识是指（ ）知识。
A. **可以精确表示的** B. 正确的 C. 在大学中学到的 D. 能够解决问题的
12. 设 P 是谓词公式，对于 P 的任何论域，存在 P 为真的情况，则称 P 为（ ）。
A. **永真式** B. 推理式 C. 据取 D. 永假式
13. 谓词公式 G 是不可满足的，当且仅当对所有的解释（ ）
A. **G 都为假** B. G 可为真也可为假
C. G 有时为假有时为真 D. G 都为真
14. 常见的语义网络中的关系有（ ）
A. have 关系 B. Composed-of 关系 C. A-Member-of 关系 D. **以上全对**
15. 归纳推理是（ ）的推理。
A. **从个别到一般** B. 从一般到一般 C. 从一般到个别 D. 从个别到个别
16. 以下推理不正确的是（ ）
A. 如果下雨，则地下湿；地下不湿，所以没有下雨
B. 如果 x 是金属，则 x 能导电；铜是金属，所以铜能导电
C. 小贝喜欢可爱的东西；哈士奇可爱；所以小贝喜欢哈士奇。
D. **如果下雨，则地上是湿的；没有下雨，所以地上不湿**

17. 关于下列的推理你的观点是：（ ）
- (1) 如果行星系统是以太阳为中心的，则金星会显示出位相的变化；
 - (2) 金星显出位相变化；
 - (3) 所以行星系统是以太阳为中心的。
- A. **不对，因为使用了肯定后件的推理。**
 - B. 对，因为符合自然演绎推理的规则。
 - C. 对，因为经典逻辑推理是从一组已知为真的事实出发，直接运用经典逻辑的推理规则推出结论的过程，而上题符合这个特征所以是对的。
 - D. 不对，因为使用了否定前件的推理。
18. 经典命题逻辑和谓词逻辑的语义解释只有两个：真和假，使用 0 和 1 表示。✓
19. 命题是可以判断真假的语句。✗
20. 谓词公式 $\forall x (P(x) \vee \exists y R(y)) \rightarrow Q(x)$ 中量词的辖域是（ ）。
- A. $\exists y R(y)$
 - B. **$\forall x (P(x) \vee \exists y R(y))$**
 - C. $Q(x)$
 - D. $P(x)$
21. 设 $L(x)$: x 是演员, $J(x)$: x 是老师, $A(x, y)$: x 佩服 y 。那么命题“所有演员都佩服某些老师”符号化为（ ）
- A. $\forall x L(x) \rightarrow A(x, y)$
 - B. **$\forall x (L(x) \rightarrow \exists y (J(y) \wedge A(x, y)))$**
 - C. $\forall x \exists y (L(x) \wedge J(y) \rightarrow A(x, y))$
 - D. $\forall x \exists y (L(x) \wedge (J(y) \wedge A(x, y)))$
22. 谓词公式 $\forall y P(y)$ 取真值为 1 的充分必要条件是（ ）
- A. 存在某些 y , 使 $P(y)$ 都取真值 1
 - B. **对任意 y , 使 $P(y)$ 都取真值 1**
 - C. 存在 y_0 , 使 $P(y_0)$ 取真值 0
 - D. 存在一个 y_0 , 使 $P(y_0)$ 取真值 1
23. 与谓词演算公式 $(\exists x)(P(x)) \vee (\exists y)(Q(y))$ 等价的公式是（ ）
- A. $(\forall x)(P(x) \wedge Q(x))$
 - B. $(\exists x)(P(x) \wedge Q(x))$
 - C. $(\forall x)(P(x) \wedge (\forall y)Q(y))$
 - D. **$(\exists x)(P(x) \vee Q(x))$**
24. MGU 是指
- A. 基替换
 - B. 最一般替换
 - C. **最一般合一**
 - D. 最一般谓词
25. 在公式中 $(\forall y)(\exists x)P(x, y)$ 存在量词是在全称量词的辖域内, 我们允许所存在的 x 可能依赖于 y 值。令这种依赖关系明显地由函数所定义, 它把每个 y 值映射到存在的那个 x 。这种函数叫做（ ）
- A. 依赖函数
 - B. **Skolem 函数**
 - C. 多元函数
 - D. 决定函数
26. $A \rightarrow B) \wedge A \Rightarrow B$ 是（ ）
- A. **假言推理**
 - B. 附加律
 - C. 拒收
 - D. 其他三个选项都对
27. 与谓词演算公式 $\neg (\forall x)(\forall y)(P(x, y) \vee Q(x, y))$ 等价的公式是（ ）
- A. **$(\exists x)(\forall y)(\neg P(x, y) \vee \neg Q(x, y))$**
 - B. $(\exists x)(\exists y)(\neg P(x, y) \wedge \neg Q(x, y))$
 - C. $(\forall x)(\exists y)(\neg P(x, y) \vee \neg Q(x, y))$
 - D. $\neg (\exists x)(\exists y)(\neg P(x, y) \vee \neg Q(x, y))$
28. 下列那个不是子句的特点?
- A. **子句中可以有析取词 (\vee)**
 - B. 子句间是没有合取词的 (\wedge)
 - C. 子句通过合取词连接句子 (\wedge)
 - D. 子句间是没有析取词的 (\vee)
29. 设有代换: $q = \{f(y)/x, z/y\}$, $l = \{a/x, b/y, y/z\}$ 则最一般合一为（ ）
- A. $\{f(y)/x, y/z\}$
 - B. **$\{f(b)/x, y/z\}$**
 - C. $\{f(b)/x, z/y\}$
 - D. $\{f(b)/x, z/b\}$
30. 用归结反演证明定理时, 若当前归结式为（ ），则定理得证。
- A. 原子谓词
 - B. 包孕式
 - C. **空子句**
 - D. 永真式
31. 从全称判断推导出特称判断或单称判断的过程, 即由一般性知识推出适合于某一具体情况的结论的推理是（ ）
- A. **演绎推理**
 - B. 默认推理
 - C. 归结推理
 - D. 单调推理

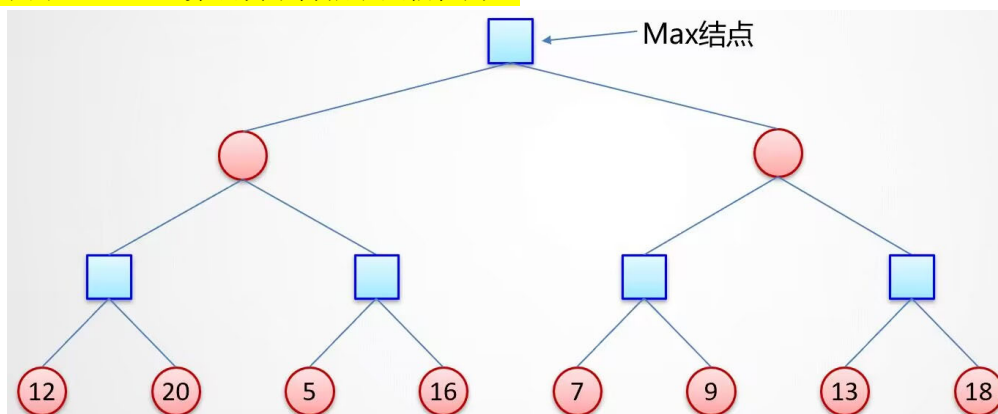
32. 反演归结（消解）证明定理时，若当前归结式是（ ）时，则定理得证。
A. 子句集 B. 包孕式 C. **空子句** D. 永真式
33. 下列选项哪个不是完备的规则策略（ ）
A. 线性归结 B. 删除策略 C. **线性输入策略** D. 支撑集策略
34. 下列说法不正确的是（ ）
A. 永真性：如果谓词公式 P 对个体域 D 上的任何一个解释都取得真值 T ，则称 P 在 D 上是永真的
B. 不可满足性：如果谓词公式 P 对于个体域 D 上的任何一个解释都取得真值 F ，则称 P 在 D 上是永久假的，如果 P 每个非空个体域上均永假，则称 P 永假
C. **永真性：如果谓词公式 P 在个体域 D 上，存在一个解释都取得真值 T ，则称 P 在 D 上是永真的**
D. 可满足性：对于谓词公式 P ，如果至少存在一个解释使得公式 P 在此解释下的真值为 T ，则称公式 P 是可满足的
35. 演绎推理的三段论式不包括（ ）
A. 结论 B. 小前提 C. **推理** D. 大前提
36. 归结策略大致可分为两大类：
A. 删除策略 递归策略 B. 限制策略 循环策略
C. **删除策略 限制策略** D. 递归策略 循环策略
37. 设甲、乙、丙三人中有人从不说真话，也有人从不说假话，某人向着三人分别提出一个问题：谁是说谎者
甲答：“ B 和 C 都是说谎者”；
乙答：“ A 和 C 都是说谎者”；
丙答：“ A 和 B 至少有一个是说谎者”。谁是老实人，谁是说谎者？
A. 丙是老实人，乙是说谎者 B. **丙是老实人，甲是说谎者**
C. 甲是老实人，丙是说谎者 D. 甲是老实人，乙是说谎者
38. 在深度优先搜索策略中，open 表的数据结构是（ ）
A. 根据估价函数值重排 B. 先进先出 C. 后进后出 D. **先进后出**
39. 盲目搜索策略不包括下列哪个？
A. 深度优先搜索 B. **全局最优搜索**
C. 有界深度优先搜索 D. 广度优先搜索
40. 关于“与或图”表示知识的叙述，正确的有（ ）
A. “与或图”就是用“and”和“or”连续各个部分的图形，用来描述各部分的因果关系。
B. **“与或图”就是用“and”节点和“or”节点组合起来的树形图，用来描述某类问题的求解关系。**
C. “与或图”就是用“and”和“or”连续各个部分的图形，用来描述各部分之间的不确定关系。
D. “与或图”就是用“and”节点和“or”节点组合起来的树形图，用来描述某类问题的层次关系。
41. 关于“与或图”表示知识的叙述，错误的有（ ）
A. “与或图”能方便地表示陈述性知识和过程性知识。
B. 用“与或图”表示知识方便使用程序设计语言表达，也便于计算机存储处理。
C. “与或图”表示知识时不一定同时有“与节点”和“或节点”。
D. **能用“与或图”表示的知识不适宜用其他方法表示。**

42. 已知初始问题的描述，通过一系列变换把此问题最终变为一个子问题集合；这些子问题的解可以直接得到，从而解决了初始问题。这种知识表示法叫（ ）
- A. 状态空间法 B. 谓词逻辑法 C. 语义网络法 D. **问题归约法**
43. 如果问题存在最优解，则下面几种搜索算法中，（ ）必然可以得到该最优解。
- A. 启发式搜索 B. 深度优先搜索
C. 有界深度优先搜索 D. **广度优先搜索**
44. 广度优先搜索算法中，OPEN 表的数据结构实际是一个（ ）
- A. 平衡树 B. **二叉树** C. 队列 D. 单链表
45. 深度优先搜索算法中，OPEN 表的数据结构实际是一个（ ）
- A. 队列 B. 平衡树 C. 二叉树 D. **单链表**
46. 启发式搜索一定比盲目式搜索好 **×**
47. 囚徒困境说明（ ）
- A. 如果没有某种约束，局中人也可在（抵赖，抵赖）的基础上达到均衡；
B. **双方都独立依照自己的利益行事，则双方不能得到最好的结果；**
C. 每个局中人在做决策时，不需考虑对手的反应
D. 双方都依照自己的利益行事，结果一方赢，一方输；
48. 囚徒困境说明个人的理性选择不一定是集体的理性选择。 **✓**
49. 博弈中知道越多的一方越有利。 **×**
50. 在博弈中如果某博弈方改变策略后收益增加则另一博弈方收益减少。 **×**
51. “二人零和”博弈中只有“敌、我”二方，并且双方的利益完全对立，其赢得函数之和为零。 **✓**
52. 博弈的最本质特征是：博弈双方的行动既相互影响又相互依赖。 **✓**
53. 将原始数据进行集成、变换、维度规约、数值规约是在以下哪个步骤的任务？
- A. **数据预处理** B. 分类和预测 C. 频繁模式挖掘 D. 数据流挖掘
54. 层感知器，其最大的缺点是只能解决线性可分的分类模式问题，要增强网络的分类能力唯一的方法是采用多层网络结构。与单层感知器相比，下面（ ）不是多层网络所特有的特点。
- A. 含有一层或多层隐含单元 B. 激活函数采用可微的函数
C. **神经元的数目可以达到很大** D. 具有独特的学习算法
55. 假定你现在解决一个有着非常不平衡类别的分类问题，即主要类别占据了训练数据的 99%。现在你的模型在测试集上表现为 99% 的准确度。那么下面哪一项表述是正确的？
- 1 准确度并不适合于衡量不平衡类别问题
2 准确度适合于衡量不平衡类别问题
3 精确率和召回率适合于衡量不平衡类别问题
4 精确率和召回率不适合于衡量不平衡类别问题
- A. 2 and 3 B. 1 and 4 C. 2 and 4 D. **1 and 3**
56. 神经网络模型训练过程的主要目的是让损失函数取得如下哪种结果？
- A. **损失函数值尽可能小** B. 损失函数方差值尽可能大
C. 损失函数值尽可能大 D. 损失函数方差值尽可能小
57. 想要了解上海市小学生的身高，需要抽取 500 个样本，这项调查中的样本是？
- A. 上海市全部小学生 B. **从中抽取的 500 名学生的身高**
C. 从中抽取的 500 名小学生 D. 上海市全部小学生的身高

58. 下面梯度下降说法错误的是？
- 随机梯度下降是深度学习算法当中常用的优化算法之一
 - 随机梯度下降是梯度下降中常用的一种
 - 梯度下降算法速度快且可靠**
 - 梯度下降包括随机梯度下降和批量梯度下降
59. 如果使用数据集的全部特征并且能够达到 100% 的准确率，但在测试集上仅能达到 70% 左右，这说明
- 欠拟合
 - 过拟合**
 - 模型很棒
 - 其它都不对
60. 下面哪种不属于数据预处理的方法？
- 离散化
 - 聚集
 - 变量代换
 - 估计遗漏值**
61. 当不知道数据所带标签时，可以使用哪种技术促使带同类标签的数据与带其他标签的数据相分离？
- 分类
 - 聚类**
 - 隐马尔可夫链
 - 关联分析
62. 以下对 k-means 聚类算法解释正确的是
- 不能自动识别类的个数,随机挑选初始点为中心点计算**
 - 能自动识别类的个数,随机挑选初始点为中心点计算
 - 能自动识别类的个数,不是随机挑选初始点为中心点计算
 - 不能自动识别类的个数,不是随机挑选初始点为中心点计算
63. 在以下不同的场景中，使用的分析方法不正确的有
- 根据用户最近购买的商品信息，用决策树算法识别出淘宝买家可能是男还是女
 - 用关联规则算法分析出购买了汽车坐垫的买家，是否适合推荐汽车脚垫
 - 根据商家最近一年的经营及服务数据，用聚类算法判断出天猫商家在各自主营类目下所属的商家层级
 - 根据商家近几年的成交数据，用聚类算法拟合出用户未来一个月可能的消费金额公式**
64. 下列哪个模型属于无监督学习？
- 逻辑回归
 - 决策树
 - KNN 分类
 - DBSCAN**
65. 如果在大型数据集上训练决策树，为了花费更少的时间来训练这个模型，下列哪种做法是正确的（ ）
- 减少树的深度**
 - 增加树的深度
 - 减少树的数量
 - 增加学习率
66. 使用决策树分类时，如果输入的某个特征的值是连续的，通常使用二分法对连续属性离散化，即根据是否大于 / 小于某个阈值进行划分。如果采用多路划分，每个出现的值都划分为一个分支，这种方式的最大问题是：
- 计算量太大
 - 验证集表现良好，测试集表现很差**
 - 验证集表现很差，测试集表现很好
 - 验证集和测试集表现很差
67. 我们想在大数据集上训练决策树，为了减少训练时间，我们可以
- 对决策树模型进行预剪枝**
 - 减少树的数量
 - 增大学习率（Learning Rate）
 - 增加树的深度
68. 某超市研究销售纪录数据后发现，买啤酒的人很大概率也会购买尿布，这种属于数据挖掘的哪类问题？
- 分类
 - 自然语言处理
 - 聚类
 - 关联规则发现**

83. 下列哪个选项是神经网络的代表
A. **残差网络** B. ID3 C. 梯度提升树 D. A*算法
84. 在图像智能处理中，图像分类的目标主要是要确定什么内容？
A. 对象背景 B. 对象边框 C. **对象类别** D. 对象位置
85. 神经网络的“损失函数”(Loss function)衡量的是
A. dropout 损失的信息量 B. 训练集与测试集之间的差距
C. **预测值与真实值之间的差距** D. pooling 损失的信息量
86. 下列哪项关于模型能力(model capacity)的描述是正确的？（模型能力指神经网络模型能拟合复杂函数的能力）
A. Dropout 的比例增加，模型能力增加 B. 学习率增加，模型能力增加
C. **隐藏层层数增加，模型能力增加** D. 其它都不正确
87. 假设我们已经在 ImageNet 数据集(物体识别)上训练好了一个卷积神经网络。然后给这张卷积神经网络输入一张全白的图片。对于这个输入的输出结果为任何种类的物体的可能性都是一样的，对吗？
A. 看情况 B. 对的 C. **不对** D. 不知道
88. 当在卷积神经网络中加入池化层(pooling layer)时，变换的不变性会被保留，是吗？
A. **是** B. 否 C. 不知道 D. 看情况
89. 对于一个图像识别问题(在一张照片里找出一只猫)，下面哪种神经网络可以更好地解决这个问题？
A. 循环神经网络 B. 感知机 C. 多层感知机 D. **卷积神经网络**
90. 考虑某个具体问题，你可能只有少量数据来解决这个问题。不过幸运的是你有一个类似问题已经预先训练好的神经网络。可以用下面哪种方法来利用这个预先训练好的网络？
A. 对新数据重新训练整个模型
B. 对每一层模型进行评估，选择其中的少数来用
C. 把除了最后一层外所有的层都冻住，重新训练最后一层
D. **只对最后几层进行调参(fine tune)**
91. 可以在自然语言处理中用于语义关系提取的神经网络技术是以下哪种？
A. 递归神经网络 B. **循环神经网络** C. 卷积神经网络 D. 残差神经网络
92. 用户希望对以下这句话进行预测“我的祖国是中国，我的母语是”，则应该采用以下哪种神经网络结构进行处理？
A. **循环神经网络** B. 卷积神经网络 C. 生成对抗网络 D. 多层感知机
93. 深度学习是当前很热门的机器学习算法，在深度学习中，涉及到大量的矩阵相乘，现在需要计算三个稠密矩阵 A,B,C 的乘积 ABC,假设三个矩阵的尺寸分别为 $m \times n$, $n \times p$, $p \times q$, 且 $m < n < p < q$, 以下计算顺序效率最高的是
A. A(BC) B. AC(B) C. **(AB)C** D. 所有其它几项计算效率都相同
94. 在神经网络中，“梯度消失”问题的主要来源是
A. **梯度趋近于零** B. 被 Dropout 丢弃 C. 被 Pooling 丢弃 D. 梯度为负数
95. 相对于普通的神经网络，循环神经网络(RNN)的“循环”主要体现在
A. 训练过程中的反向传播次数更多 B. 训练经过一定轮次之后将参数归零
C. 每个节点自循环 D. **深层节点的输出会反过来影响浅层节点**
96. (多选)人工智能中(深度学习)网络使用的常用激活函数有
A. **relu** B. **tanh** C. **sigmond** D. sinh

97. （多选）对于神经网络，以下哪些是超参数？
- A. **神经网络的学习率（learning rate）**
 - B. 神经网络的偏置（bias）
 - C. 神经网络的权重（weights）
 - D. **神经网络的层数（number of layers）**
98. 下列哪个神经网络结构会发生权重共享？
- A. **卷积神经网络**
 - B. 全连接神经网络
 - C. 感知机
 - D. **循环神经网络**
99. 下面哪些属于循环神经网络？
- A. **LSTM**
 - B. CNN
 - C. AlexNet
 - D. **RNN**
100. 可以通过将所有权重初始化为 0 来训练网络。✗
101. 下面哪种卷积神经网络结构引入了残差网络结构？
- A. LeNet
 - B. GoogleNet
 - C. AlexNet
 - D. **ResNets**
102. 在反向传播过程中，首先被计算的变量（ ）的梯度，之后将其反向传播。
- A. **激活函数**
 - B. 特征映射
 - C. 损失函数
 - D. 连接权重
103. 卷积神经网络 VGG16 名称中 16 指的是（ ）
- A. VGG 发表的第 16 代网络
 - B. 网络中有 16 层的参数需要训练
 - C. 论文发表于 2016 年
 - D. **网络共有 16 层**
104. 下列选项中，不是人工智能算法中的学习方法是？
- A. 深度学习
 - B. 强化学习
 - C. **重复学习**
 - D. 归纳学习
105. 下列哪个选项是神经网络的代表？
- A. ID3 算法
 - B. AO*算法
 - C. **卷积神经网络**
 - D. 梯度提升树
106. 下列哪项算法是深度学习的重要基础：
- A. A*算法
 - B. 最小生成树算法
 - C. **SGD 反向传播**
 - D. 最大流—最小割算法
107. 下列哪一项在神经网络中引入了非线性？
- A. 随机梯度下降
 - B. 池化函数
 - C. 卷积函数
 - D. **修正线性单元（ReLU）**
108. 在一个神经网络里，知道每一个神经元的权值和偏差值是最重要的一步。如果以某种方法知道了神经元准确的权重和偏差，就可以近似任何函数。实现这个最佳的办法是：
- A. 随机赋值，祈祷他们是正确的
 - B. **赋予一个初始值，通过检查跟最佳值的差值，然后迭代更新权重**
 - C. 搜索所有权重和偏差的组合，直到得到最佳值
 - D. 其他说法都不正确
109. 图中 MiniMax 算法决策树根节点估值为：



- A. 20
 - B. 18
 - C. **16**
 - D. 5
110. 从已知事实出发，通过规则库求得结论的产生式系统的推理方式是
- A. **正向推理**
 - B. 反向推理
 - C. 双向推理
 - D. 一般推理

111. 以下哪种方法不属于特征选择的标准方法：

- A. 抽样 B. 嵌入 C. 过滤 D. 包装

112. 以下哪些算法是分类算法？

- A. DBSCAN B. K-Mean C. A*算法 D. ID3

113. () 是一个观测值，它与其他观测值的差别如此之大，以至于怀疑它是由不同的机制产生的。

- A. 边界点 B. 离群点 C. 质心 D. 核心点

114. 简单地将数据对象集划分成不重叠的子集，使得每个数据对象恰在一个子集中，这种聚类类型称作 ()

- A. 非互斥聚类 B. 模糊聚类 C. 层次聚类 D. 划分聚类

115. 关于 K 均值和 DBSCAN 的比较，以下说法不正确的是 ()

- A. K 均值很难处理非球形的簇和不同大小的簇，DBSCAN 可以处理不同大小和不同形状的簇。
B. K 均值丢弃被它识别为噪声的对象，而 DBSCAN 一般聚类所有对象。
C. K 均值可以发现不是明显分离的簇，即便簇有重叠也可以发现，但是 DBSCAN 会合并有重叠的簇。
D. K 均值使用簇的基于原型的概念，而 DBSCAN 使用基于密度的概念。

116. 如果“线性回归”模型完美地拟合了训练样本（训练样本误差为零），则下面哪个说法是正确的？

- A. 测试样本误差始终为零 B. 出现了欠拟合
C. 出现了过拟合 D. 测试样本误差不可能为零

117. 下列关于线性回归分析中的残差 (Residuals) 说法正确的是？

- A. 残差均值总是大于零 B. 残差均值总是小于零
C. 残差均值可以随意 D. 残差均值总是为零

118. 构建一个最简单的线性回归模型需要几个系数（只有一个特征）？

- A. 3 个 B. 2 个 C. 4 个 D. 1 个

119. (多选) 常用的池化层有哪些？

- A. AveragePooling B. MedianPooling C. MinPooling D. MaxPooling

120. (多选) 常见的知识表示方法有：

- A. 状态空间表示法 B. 计算机表示法 C. 谓词表示法 D. 产生式表示法

121. (多选) 关于梯度下降法叙述正确的有？

- A. 精度下降方法迭代时将沿着梯度方向进行更新
B. 梯度下降方法迭代时将沿着负梯度方向进行更新
C. 梯度方向是使得函数值下降最快的方向
D. 梯度方向是使得函数值上升最快的方向

122. (多选) 针对 α - β 过程首先使搜索树的某一部分达到最大深度，这时计算出某些 MAX 节点的 α 值，或者是某些 MINI 节点的 β 值。其中，如下说法中正确的有：

- A. MINI 节点的 β 值永不增加 B. MINI 节点的 β 值永不下降
C. MAX 节点的 α 值永不下降 D. MAX 节点的 α 值永不增加

123. (多选) Relu 激活函数的优点有：

- A. 解决了梯度消失、爆炸的问题 B. 输出不是以 0 为中心
C. 加速网络训练 D. 计算方便，计算速度快，求导方便

124. （多选）如下那个表述是正确的？选出所有正确的选项。

- A. 初始化 K-means 的标准化方法是设置每个簇等于零向量。
- B. 既然 K-means 算法是无监督学习算法，所以不会过拟合数据，因此聚类的类别越多越好。
- C. 如果担心 K-means 陷入局部最优，一个解决这个问题方法就是尝试使用多个随机的初始化值。
- D. 对于一些数据集来说，“正确的”K 值（簇的数量）是不确定的，因此人类专家需要根据数据的特性去确定 K 值。

125. （多选）对于原始数据，存在的问题可能有：（ ）

- A. 不完整
- B. 含噪声
- C. 不一致
- D. 重复

126. （多选）通过监督学习进行二分类模型训练过程中，可能会遇到正负样本数量不平衡的情况（比如正样本有 50 万但是负样本有 100 万），以下哪些方法可以对此进行恰当的处理？

- A. 复制两份正样本参与到训练中
- B. 将所有数据加入训练集，充分利用全部数据
- C. 从 100 万负样本中随机抽取 50 万
- D. 正样本的权重设置为 2，负样本权重设置为 1

127. （多选）观察样本次数如何影响过拟合？注意：所有情况的参数都保持一致。

- A. 观察次数少，不容易发生过拟合
- B. 观察次数多，不容易发生过拟合
- C. 观察次数少，容易发生过拟合
- D. 观察次数多，容易发生过拟合

128. 利用归结原理证明定理时，若得到的归结式为空集，则结论成立。✓

129. Relu 激活函数在零点是不可导的。✓

130. 如果两个变量相关，那么它们一定是线性关系。✗