**随堂练习4 频繁模式挖掘**

1. 当某个项集的支持度高于最小阈值时，它就是频繁项集。（）

A. √

B. ×

1. 频繁项集的所有非空子集（）
2. 必须是频繁的
3. 不一定是频繁的
4. 一定不是频繁的
5. 频繁项集是指（）

A. 支持度不小于预定义的最小支持度阈值的项集

B. 置信度不小于预定义的最小置信度阈值的项集

C. 支持度不小于预定义的最小支持度且置信度不小于预定义的最小置信度的项集

1. （多选）下列不属于FP树结构特点的是（）

A.保存了频繁模式相关的所有信息

B.保留了无关信息，即非频繁项集

C.项在树中按支持度降序排列

D.构造出的FP树大小有可能会超过原始数据库

1. 下列关于Apriori算法，错误的是（）

A. 用频繁的(k–1)-项集可以生成候选的k-项集

B. 候选集生成是Apriori 算法的瓶颈

C. 如果最长的模式是n的话，则需要n次数据库扫描

D. 用数据库扫描和模式匹配计算可以计算候选集的支持度

1. 构造FP-树时，为一个事务增加分支时，沿共同前缀上的每个结点的计数增加1，为前缀之后的项创建结点和链接。 （）

A. √

B. ×

1. FP-Growth算法从FP-树中挖掘频繁项集采用的是自顶向下的方法。（）

A. √

B. ×

1. 下列说法错误的是 （）

A. 闭频繁项集包含了频繁项集的完整信息

B. 极大频繁项集包含了对应频繁项集的完整支持度信息

C. 若频繁项集的数据量非常大，将导致计算机无法计算和存储。为克服这个问题，引入闭频繁项集和极大频繁项集的概念。

D. 使用垂直数据格式挖掘频繁项集,不需要扫描数据库来确定( k+1 ) 项集的支持度

1. 下列关于非频繁模式说法正确的是 ()

A、其支持度小于阈值

B、都不是有趣的

C、包含负模式和负相关模式

1. 具有较高支持度的项集也具有较高的置信度。（）

A. √

B. ×

1. （多选）Apriori算法的计算复杂度受到下列哪些选项的影响。( )

A.支持度阀值

B.项数（维度）

C.事务数