## 第一讲习题（概述）

1. 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. 频繁模式挖掘

B. 分类和预测

C. 数据预处理

D. 数据流挖掘

1. （多选）通过数据挖掘过程所推倒出的关系和摘要经常被称为：()

A. 模型

B. 模式

C. 模范

D. 模具

1. （多选）数据挖掘的预测建模任务主要包括哪几大类问题？ ()

A. 分类

B. 回归

C. 模式发现

D. 模式匹配

1. （多选）数据挖掘算法的组件包括：()

A. 模型或模型结构 B. 评分函数 C. 优化和搜索方法 D. 数据管理策略

1. （多选）以下哪些学科和数据挖掘有密切联系？()

A. 统计 B. 计算机组成原理 C. 矿产挖掘 D. 人工智能

1. 数据挖掘的主要任务是从数据中发现潜在的规则，从而能更好的完成描述数据、预测数据等任务。 ()
2. 对
3. 错
4. 数据挖掘的目标不在于数据采集策略，而在于对于已经存在的数据进行模式的发掘。（）
5. 对
6. 错
7. 寻找模式和规则主要是对数据进行干扰，使其符合某种规则以及模式。（）
8. 对
9. 错
10. Web数据挖掘是通过数据库中的一些属性来预测另一个属性,它在验证用户提出的假设过程中提取信息. （）
11. 对
12. 错
13. 聚类（clustering）是这样的过程：它找出描述并区分数据类或概念的模型(或函数)，以便能够使用模型预测类标记未知的对象类。 （）
14. 对
15. 错
16. 分类和回归都可用于预测，分类的输出是离散的类别值，而回归的输出是连续数值。()
17. 对
18. 错
19. 分类模型的误差大致分为两种：训练误差（training error）和泛化误差（generalization error）. ()
20. 对
21. 错