**数据预处理-习题（100分）**

**专业班级：**

**学号：**

**姓名：**

[单选题]**（5分）**1、某超市研究销售记录数据后发现，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？（ ）。

[A]关联规则发现；

[B]聚类;

[C]分类;

[D]自然语言处理;

[单选题]**（5分）**3、当不知道数据所带标签时，可以使用哪种技术使带同类标签与带其他标签的数据相分离？（ ）。

[A]分类；

[B]聚类;

[C]关联分析;

[D]隐马尔可夫链;

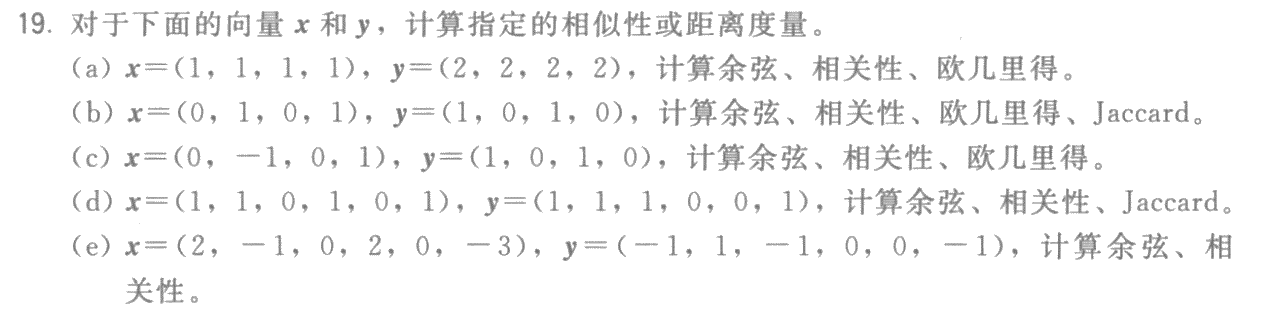
[多选题]**（5分）**2、数据挖掘的预测建模任务主要包括哪几大类问题？（ ）。

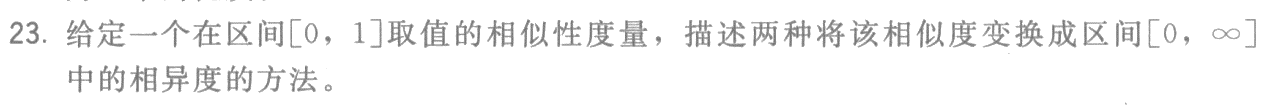
[A]分类 [B]回归 [C]模式发现 [D]模式匹配

[简答题]**（15分）**3、有一组已经排序的数据如下：5，10,11,13，15,35,50,55,72,79,204,215，如果采用等频分箱将这些数据划分为4个箱，15在第几个箱内？

[计算题]**（20分）**7、有数据值2000，3000，4000，6000，10000。求

令min=0,max=1,最小-最大规范化。

[计算题]**（30分）**

[简答题]**（20分）**