

第一章 数据挖掘的概论

**1.1(d)当把数据挖掘看做知识发现过程时，描述数据挖掘所涉及的步骤。

- 数据挖掘是从大量数据中挖掘有趣模式和知识的过程。
- 当其被看作知识发现过程时，其基本步骤主要有：
 - (1).数据清理：清楚噪声和删除不一致数据；
 - (2).数据集成：多种数据源可以组合在一起；
 - (3).数据选择：从数据库中提取与分析任务相关的数据；
 - (4).数据变换：通过汇总或者聚集操作，把数据变换和统一成适合挖掘的形式；
 - (5).数据挖掘：使用智能方法或者数据挖掘算法提取数据模式；
 - (6).模式评估：根据某种兴趣度量，识别代表知识的真正有趣的模式。(7).知识表示：使用可视化和知识表示技术，将已经挖掘到的有用知识给用户呈现出来。

1.2 数据仓库与数据库有什么不同？它们有哪些相似之处？**

- 数据仓库是多个异构数据源在单个站点以统一的模式组织的存储，以支持管理决策。数据仓库技术包括数据清理、数据集成和联机分析处理(OLAP)。数据库系统也称数据库管理系统，由一组内部相关的数据（称作数据库）和一组管理和存取数据的软件程序组成。它们的相似之处：都是通过某个数据库软件，基于某种数据模型来组织、管理数据。

1.5 解释区分和分类、特征化和聚类、分类和回归之间的区别和相似之处。**

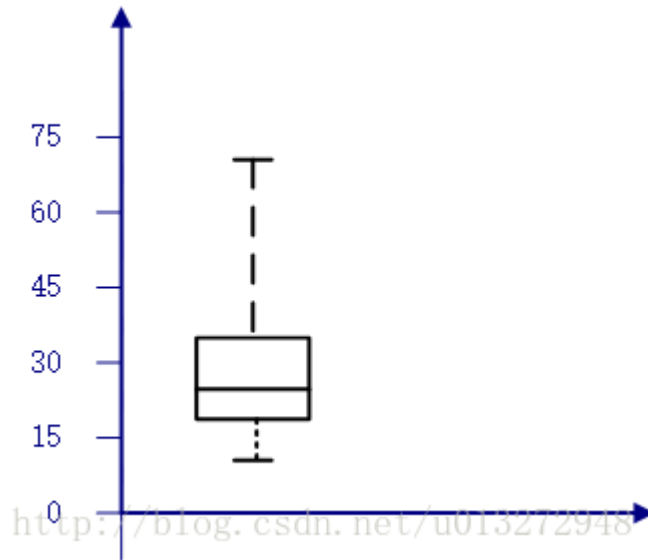
- 区分和分类：数据区分是将**目标类数据对象的一般特性**与一个或多个**对比类对象的一般特性**进行比较；而分类则是找出描述和区分数据类或概念的模式，以便能够使用模型对未知类标号的样例进行**预测**。
- 特征化和聚类：数据特征化是**目标类数据的一般特性或特征的汇总**，即在进行数据特征化时很清楚特征化的这些数据的特点是什么；而**聚类则只是分析数据对象**，按照“最大化类内相似度、最小化类间相似度”的原则**进行聚类或分组**。
- 分类在这一点时已经说过；**回归主要是建立连续值的函数模型**，回归主要用来预测缺失的或难以获得的数值数据值，而不是离散的类标号，同时回归也包含基于可用数据的分布趋势识别。

第二章 认识数据

2.2 假设所分析的的数据包括属性age，它在数据元组中的值（以递增序）为
13,15,16,16,19,20,21,22,22,25,25,25,30,33,33,35,35,35,35,36,40,45,46,52,70.

1. 该数的均值是多少？中位数是什么？该数的均值为29.963，中位数是25。
2. 该数据的众数是什么？讨论数据的模态（即二模、三模等）。该数据的众数为25和35，即该数据是一个双峰的分布，即二模。
3. 该数据的中列数是多少？该数据的中列数为 $(70+13)/2=41.5$ 。
4. 你能粗略的找出该数据的第一个四分位数(Q1Q1)和第三个四分位数(Q3Q3)吗？第一个四分位数为： $[27/4]=7$ 处， $Q_1=20$ ，第三个四分位数为： $7*3=21$ 处， $Q_3=35$ 。

5. 给出该数据的五数概括。根据以上，得到了最小观测值、Q1、Q2、Q3,最大观测值，所以画出其盒图如下：



6. **分位数-分位数图与分位数图有什么区别？** 分位数图(quantile plot)是一种观察单变量数据分布的简单有效方法。首先它显示给定属性的所有数据的分布情况；其次，它绘出了分位数信息（即对于某序数或数值属性X，设 $x_i(i=1,...,N)$ 是按照递增排序的数据，使得 x_1 是最小的观测值， x_N 是最大的观测值）。**统一属性不同取值的分布趋势** 分位数-分位数图(q-q图)则是反映了同一个属性的不同样本的数据分布情况，使得用户可以很方便的比较这两个样本之间的区别或者联系(看距离0)。

散点图:则是说明是**属性间的相关程度****(正负相关**)

如下图所示:

表2.1 AllElectronics的一个部门销售的

单价 (美元)	商品销售量
40	275
43	300
47	250
...	...
74	360
75	515
78	540
...	...
115	320
117	270
120	350

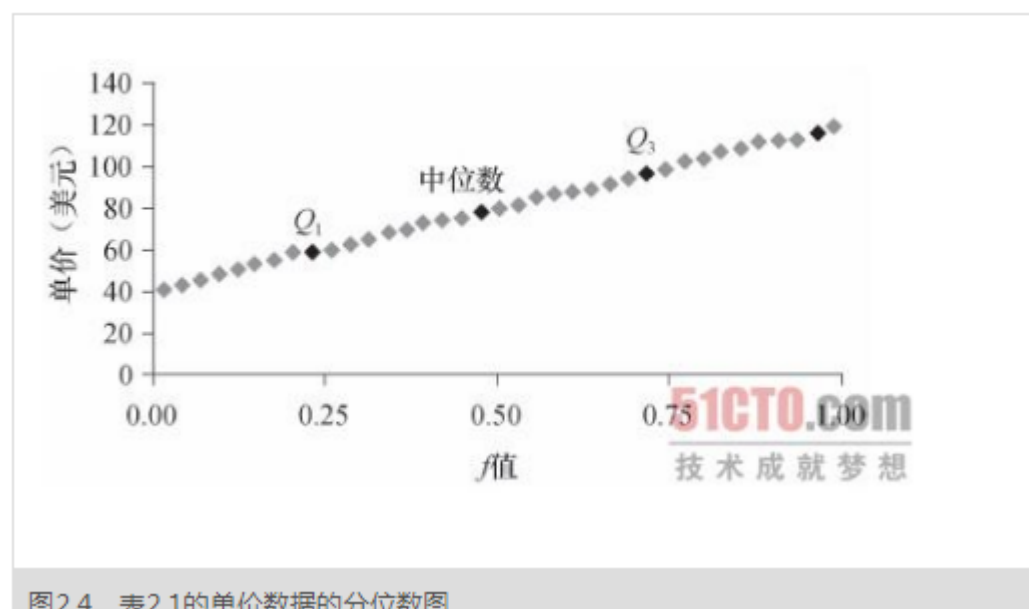


图2.4 表2.1的单价数据的分位数图

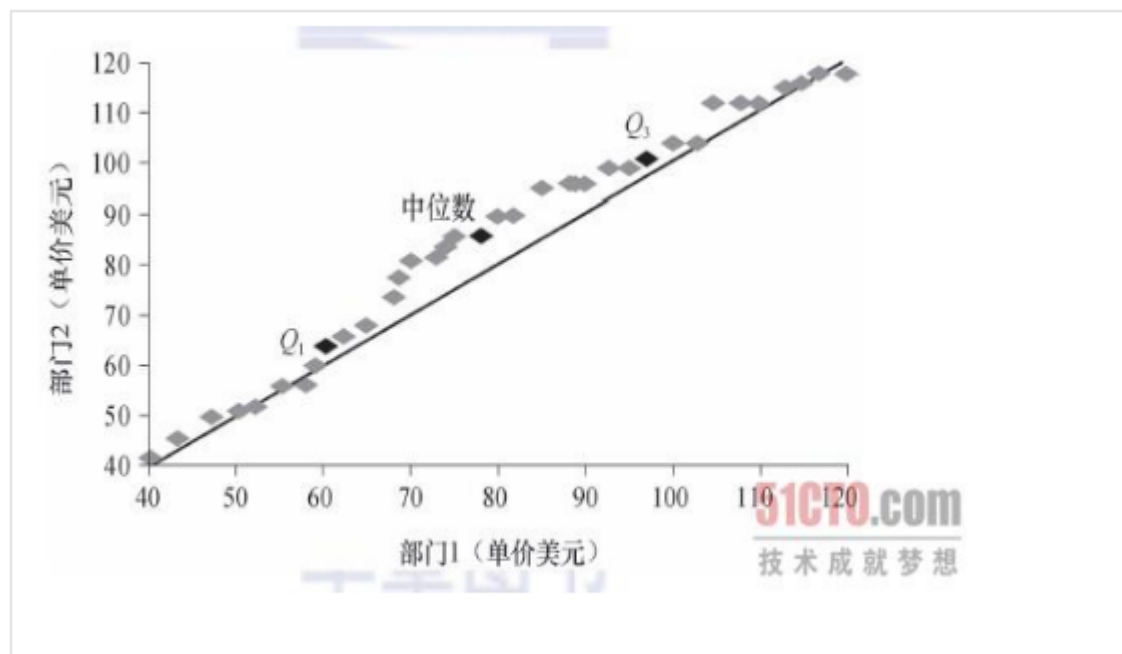


图2.4 表2.1的单价数据的分位数图

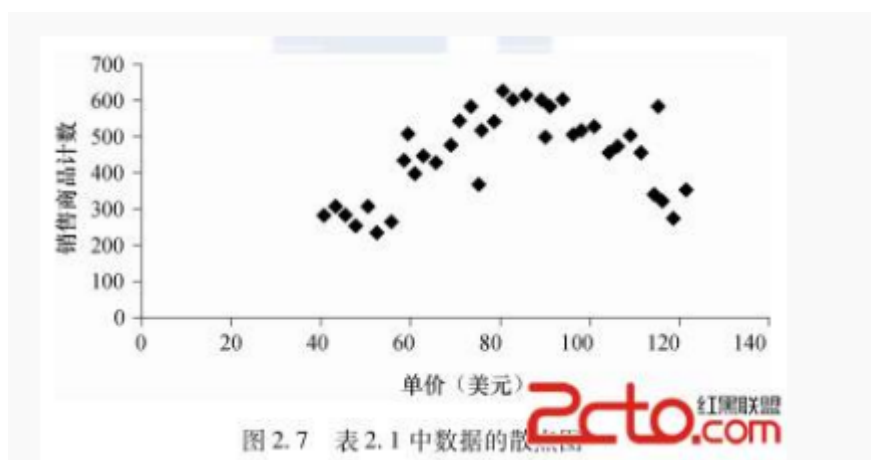


图 2.7 表 2.1 中数据的散点图

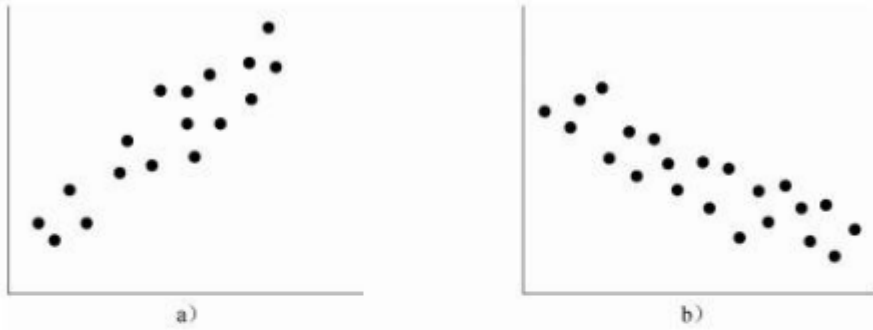


图 2.8 散点图可以用来发现属性之间的相关性：a) 正相关；b) 负相关

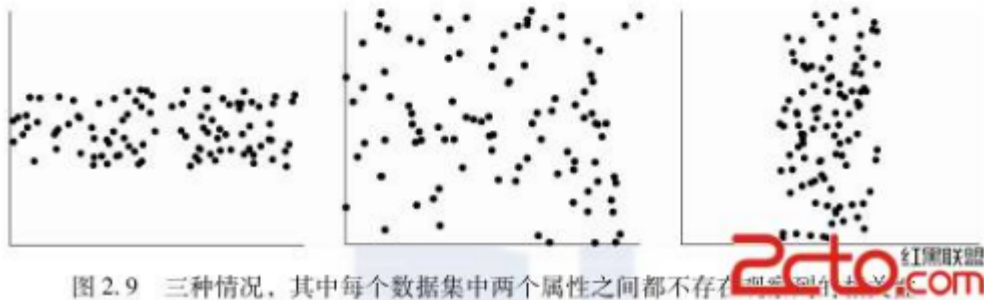


图 2.9 三种情况，其中每个数据集中两个属性之间都不存在相关性

2.3 给定的数据集已经分组到区间，计算该数据的近似中位数。**

1. 解题思路：由于该题目并没有说明某一个年龄对应的人数有多少个，所以一种解题思路就是取每一个年龄区间的中位数乘以其人数，然后再除以总的人数从而计算所有数据的中位数。

$$median = \frac{3 \times 200 + 10 \times 450 + 18 \times 300 + 35 \times 1500 + 65 \times 700 + 95 \times 44}{200 + 450 + 300 + 1500 + 700 + 44} \simeq 35$$

2.5 简要概述如何计算被如下属性描述的对象的相关性：

1. 标称属性

标称属性的相异性可以根据不匹配率去计算：

$$d(i, j) = \frac{p - m}{p}$$

其中， p 为刻画对象的属性总数， m 是匹配的数目（即 i 和 j 取值相同状态的属性数）

2. 非对称的二元属性

非对称的二元相异性可以依据二元属性的列联表去计算，计算公式如下：

$$d(i, j) = \frac{r + s}{q + r + s}$$

具体标号含义详解课本第71页。

3. 数值属性

数值属性可以有闵可夫斯基距离(Minkowski distance)，它是欧几里得距离和曼哈顿距离的推广，定义如下：

$$d(i, j) = \sqrt[h]{|x_{i1} - x_{j1}|^h + |x_{i2} - x_{j2}|^h + \dots + |x_{ip} - x_{jp}|^h}$$

4. 词频向量

词频向量可以用余弦相似度与计算，其计算方式如下：

$$\text{sim}(\vec{x}, \vec{y}) = \cos(\vec{x}, \vec{y}) = \frac{\vec{x} \cdot \vec{y}}{\|\vec{x}\| \|\vec{y}\|}$$

2.6 给定两个被元组(22,1,42,10)和(20,0,36,8)表示的对象。

1. 计算这两个对象之间的欧几里得距离。

$$d = \sqrt{(22 - 20)^2 + (1 - 0)^2 + (42 - 36)^2 + (10 - 8)^2} \simeq 6.7$$

2. 计算这两个对象之间的曼哈顿距离。

$$d = |22 - 20| + |1 - 0| + |42 - 36| + |10 - 8| = 11$$

1. 3. 使用 $q = 3$ ，计算这两个对象之间的闵可夫斯基距离。

这里的 q ，个人认为可能是公式中的 h ，所以计算如下：

$$d = \sqrt[3]{|22 - 20|^3 + |1 - 0|^3 + |42 - 36|^3 + |10 - 8|^3} \simeq 6.15$$

4. 计算这两个对象之间的上确界距离。

$$d = \max_f |x_{if} - x_{jf}| = |42 - 36| = 6$$

第三章 数据预处理

3.3 在习题2.2中，属性age包括如下值13, 15, 16, 16, 19, 20, 20, 21, 22, 22, 25, 25, 25, 25, 30, 33, 33, 35, 35, 35, 35, 36, 40, 45, 46, 52, 70。

(a) 使用深度为3的箱，用箱均值光滑以上数据。说明你的步骤，讨论这种技术对给定数据的效果。首先将以上数据划分为深度为3的箱：(13, 15, 16, 16, 19, 20, 20, 21, 22)，(22, 25, 25, 25, 25, 30, 33, 33, 35)，(35, 35, 35, 36, 40, 45, 46, 52, 70)；其次，用箱均值光滑：(18, 18, 18, 18, 18, 18, 18, 18, 18)，(28, 28, 28, 28, 28, 28, 28, 28, 28)，(43.7, 43.7, 43.7, 43.7, 43.7, 43.7, 43.7, 43.7, 43.7)。对于给定数据，该方法虽然在一定程度上光滑了噪声数据，但是由于箱深度太小，将数据的分布特征也抹去了。

(b) 如何确定该数据中的离群点？可以将以上数据看作到0点的距离，然后通过聚类方法识别以上数据中的离群点。

(c) 还有什么其他方法可以用来光滑数据 可以用回归、聚类等等来检测离群点。

3.4 讨论数据集成需要考虑的问题。在数据集成时，需要考虑诸如实体识别、数据冗余、以及元组的重复等问题。

3.5 如下规范化方法的值域是什么？

- 最小最大规范化。

$$[new_min_A, new_max_A]$$

- z分数规范化。

$$\left[\frac{v_{min} - \bar{A}}{\sigma_A}, \frac{v_{max} - \bar{A}}{\sigma_A} \right]$$

- z分数规范化，使用绝对偏差而不是标准差。

$$\left[\frac{v_{min} - \bar{A}}{s_A}, \frac{v_{max} - \bar{A}}{s_A} \right]$$

- 小数定标规范化

$$[-1, 1]$$

3.6 使用如下方法规范化如下数据组：

200, 300, 400, 600, 1000

- 令 $min = 0, max = 1$, 最小最大规范化；

$$v'_i = \frac{v_i - min_A}{max_A - min_A} (new_max_A - new_min_A) + new_min_A \text{ 规范化后为 } 0, 0.125, 0.5, 0.75, 1。$$

- z分数规范化；

$$v'_i = \frac{v_i - \bar{A}}{\sigma_A}, \bar{A} = 500, \sigma_A = 282.8, \text{ 规范化后为 } -1.06, -0.71, -0.35, 0.35, 1.77。$$

- z分数规范化，使用均值绝对值偏差而不是标准差。

$$v'_i = \frac{v_i - \bar{A}}{s_A}, \bar{A} = 500, s_A = 240, \text{ 规范化后为 } -1.25, -0.83, -0.42, 0.42, 2.08。$$

- 小数定标规范化。

$$v'_i = \frac{v_i}{10^j}, \text{ 规范化后为 } 0.2, 0.3, 0.4, 0.6, 0.1。$$

3.7 使用习题3.3中给出的age数据，回答以下问题：

- 使用最小最大规范化将age值35变换到[0.0,1.0]区间。
0.386
- 使用zz分数规范化，变换age，其中标准差为12.94。
0.386
- 使用小数定标规范化变换35。
0.35
- 指出给定数据，你愿意使用哪种方法。陈述你的理由。
对于大量数据的情况下，可能不容易找到其最大值和最小值，或者对于求其最大值和最小值比较费时，此时我更倾向于z分数规范化。应该会根据具体数据去选择，观察它的分布特性，进而去选择规范化方法。

3.8 使用习题2.4中给出的age和%fat数据数据，回答如下问题：(a) 基于zz分数规范化，规范化这两个属性。(b) 计算相关系数(Pearson积矩系数)。这两个变量是正相关还是负相关？计算他们的协方差。使用zz分数规范化公式计算即可。

3.9 假设12个销售价格记录已经排序，如下所示：

5,10,11,13,15,35,50,55,55,72,92,204,215

使用如下各方法将他们划分成三个箱。

(a) 等频划分。

bin1	5, 10, 11, 13
bin2	15, 35, 50, 55
bin3	72, 92, 204, 215

(b) 等宽划分 等宽分箱：将变量的取值范围分为k个等宽的区间，每个区间当做一个分箱；
 $(215-5)/3=70$ $(215-5)/3=70$, 所以区间的距离为70，又需要将其划分为3个箱。

bin1	5, 10, 11, 13, 15, 35, 50, 55, 72
bin2	92
bin3	204, 215

(c) 聚类 这里以数字间的最大间隔来进行一个简单的聚类；

bin1	5, 10, 11, 13, 15
bin2	35, 50, 55, 72, 92
bin3	204, 215

3.10 使用流程图概述如下属性子集选择过程。

- 逐步向前选择
流程图画起来麻烦，给出定义。自己去画吧，哈哈。
该过程由空属性集作为归约集开始，确定原属性集中最好的属性，并将它添加到归约集中。在其后的每一次迭代，将剩下的原属性集中的最好的属性添加到该集合中。
- 逐步向后删除
该过程由整个属性集开始。在每一步中，删除尚在属性集中最差的属性。
- 逐步向前选择和逐步向后删除的组合
在每一次迭代的过程中，每一步选择一个最好的属性，并在剩余属性中删除一个最差的属性。

第四章:

- 4.3 假定数据仓库包含三个维——*time*、*doctor* 和 *patient*，两个度量——*count* 和 *charge*，其中，*charge* 是医生对一位病人的一次诊治的费用。
- 列举三种流行的数据仓库建模模式。
 - 使用 (a) 中列举的模式之一，画出上面数据仓库的模式图。
 - 由基本方体 [*day*, *doctor*, *patient*] 开始，为列出 2010 年每位医生的收费总数，应当执行哪些 OLAP 操作？
 - 为得到同样的结果，写一个 SQL 查询。假定数据存放在关系数据库中，其模式为 *fee*(*day*, *month*, *year*, *doctor*, *hospital*, *patient*, *count*, *charge*)。

(a) 星形模式。雪花模式。事实星座模式

(b)

(c) time维：由“day”上卷到“year”

patient维：由“patient_name”上卷到“all”

对year=2010进行切片

(d) select doctor,SUM(charge)

from fee

where year=2010

group by doctor

第六章:

支持度3%：意味着3%顾客同时购买牛奶和面包(联合概率)

置信度40%：意味着购买牛奶的顾客40%也购买面包(通过联合概率求条件概率)

频繁项集产生:

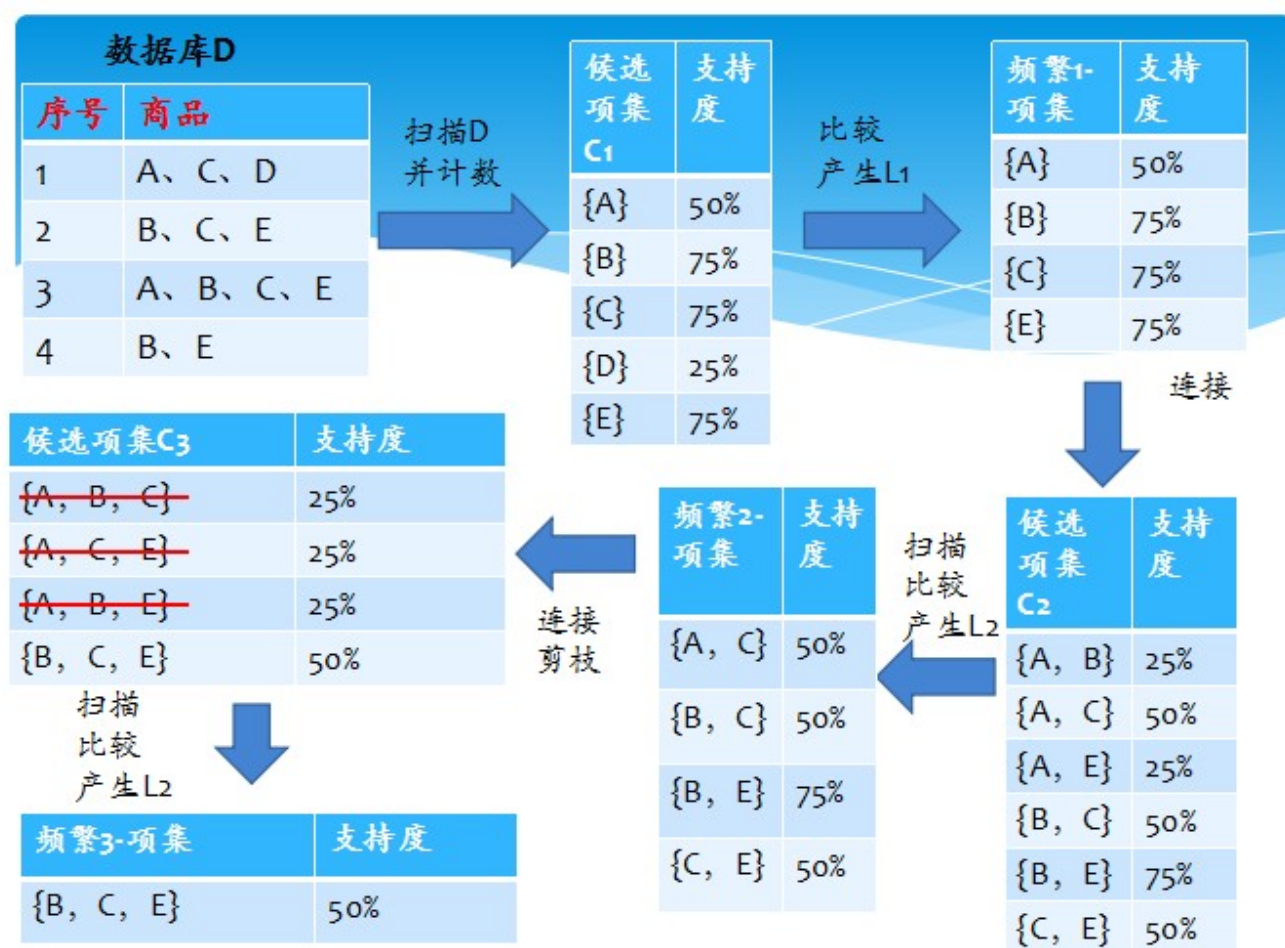


表 5-14 习题 5.3 数据集

Customer ID	Transaction ID	Items Bought
1	0001	{a, d, e}
1	0024	{a, b, c, e}
2	0012	{a, b, d, e}

2	0031	{a, c, d, e}
3	0015	{b, c, e}
3	0022	{b, d, e}
4	0029	{c, d}
4	0040	{a, b, c}
5	0033	{a, d, e}
5	0038	{a, b, e}

- (a) 把每一个事务作为一个购物篮，计算项集{e}、{b, d}和{b, d, e}的支持度。
- (b) 利用(a)中结果计算关联规则{b, d}→{e} 和 {e}→{b, d}的置信度。置信度是一个对称的度量吗？
- (c) 把每一个用户购买的所有商品作为一个购物篮，计算项集{e}、{b, d}和{b, d, e}的支持度。
- (d) 利用(b)中结果计算关联规则{b, d}→{e} 和 {e}→{b, d}的置信度。置信度是一个对称的度量吗？

答：(a) $s(\{e\}) = 8/10 = 0.8$ ；

$$s(\{b, d\}) = 2/10 = 0.2；$$

$$s(\{b, d, e\}) = 2/10 = 0.2.$$

$$(b) c(\{b, d\} \rightarrow \{e\}) = s(\{b, d, e\})/s(\{b, d\}) = 0.2/0.2 = 1;$$

$$c(\{e\} \rightarrow \{b, d\}) = s(\{b, d, e\})/s(\{e\}) = 0.2/0.8 = 0.25.$$

由于 $c(\{b, d\} \rightarrow \{e\}) \neq c(\{e\} \rightarrow \{b, d\})$ ，所以置信度不是一个对称的度量。

- (c) 如果把每一个用户购买所有的所有商品作为一个购物篮，则

$$s(\{e\}) = 4/5 = 0.8；$$

$$s(\{b, d\}) = 5/5 = 1；$$

$$s(\{b, d, e\}) = 4/5 = 0.8.$$

- (d) 利用 c 中结果计算关联规则{b, d}→{e} 和 {e}→{b, d}的置信度,则

$$c(\{b, d\} \rightarrow \{e\}) = 0.8/1 = 0.8$$

$$c(\{e\} \rightarrow \{b, d\}) = 0.8/0.8 = 1$$

置信度不是一个对称的度量

频繁项集计算:

6.6 数据库有 5 个事务。设 $min_sup = 60\%$, $min_conf = 80\%$ 。

TID	购买的商品
T100	{M, O, N, K, E, Y}
T200	{D, O, N, K, E, Y}
T300	{M, A, K, E}
T400	{M, U, C, K, Y}
T500	{C, O, O, K, I, E}

- (a) 分别使用 Apriori 算法和 FP-growth 算法找出频繁项集。比较两种挖掘过程的有效性。
- (b) 列举所有与下面的元规则匹配的强关联规则（给出支持度 s 和置信度 c ），其中， X 是代表顾客的变量， $item_i$ 是表示项的变量（如“A”，“B”等）：

$$\forall x \in transaction, buys(X, item_1) \wedge buys(X, item_2) \Rightarrow buys(X, item_3) \quad [s, c]$$

Apriori:

C1 =	m	3	L1 =	<table> <tr><td>m</td><td>3</td></tr> <tr><td>o</td><td>3</td></tr> <tr><td>k</td><td>5</td></tr> <tr><td>e</td><td>4</td></tr> <tr><td>y</td><td>3</td></tr> </table>	m	3	o	3	k	5	e	4	y	3	C2 =	<table> <tr><td>mo</td><td>1</td></tr> <tr><td>mk</td><td>3</td></tr> <tr><td>me</td><td>2</td></tr> <tr><td>my</td><td>2</td></tr> <tr><td>ok</td><td>3</td></tr> <tr><td>oe</td><td>3</td></tr> <tr><td>oy</td><td>2</td></tr> <tr><td>ke</td><td>4</td></tr> <tr><td>ky</td><td>3</td></tr> <tr><td>ey</td><td>2</td></tr> </table>	mo	1	mk	3	me	2	my	2	ok	3	oe	3	oy	2	ke	4	ky	3	ey	2	L2 =	<table> <tr><td>mk</td><td>3</td></tr> <tr><td>ok</td><td>3</td></tr> <tr><td>oe</td><td>3</td></tr> <tr><td>ke</td><td>4</td></tr> <tr><td>ky</td><td>3</td></tr> </table>	mk	3	ok	3	oe	3	ke	4	ky	3	C3 =	<table> <tr><td>oke</td><td>3</td></tr> <tr><td>key</td><td>2</td></tr> </table>	oke	3	key	2	L3 =	<table> <tr><td>oke</td><td>3</td></tr> </table>	oke	3
	m	3																																																								
	o	3																																																								
	k	5																																																								
	e	4																																																								
	y	3																																																								
	mo	1																																																								
	mk	3																																																								
	me	2																																																								
	my	2																																																								
	ok	3																																																								
oe	3																																																									
oy	2																																																									
ke	4																																																									
ky	3																																																									
ey	2																																																									
mk	3																																																									
ok	3																																																									
oe	3																																																									
ke	4																																																									
ky	3																																																									
oke	3																																																									
key	2																																																									
oke	3																																																									
o	3																																																									
n	2																																																									
k	5																																																									
e	4																																																									
y	3																																																									
d	1																																																									
a	1																																																									
u	1																																																									
c	2																																																									
i	1																																																									

FP-tree的构造:

http://blog.sina.com.cn/s/blog_68ffc7a40100uebg.html

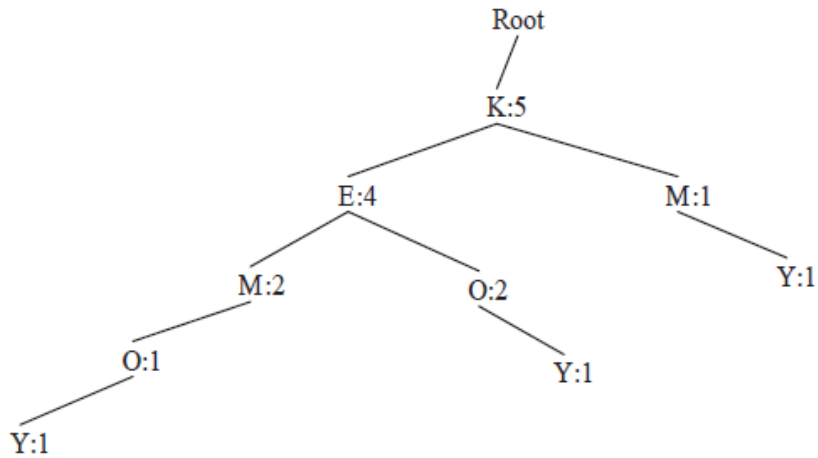


Figure 5.2: FP-tree for Exercise 5.3.

item	conditional pattern base	conditional tree	frequent pattern
y	{ {k,e,m,o:1}, {k,e,o:1}, {k,m:1} }	k:3	{k,y:3}
o	{ {k,e,m:1}, {k,e:2} }	k:3,e:3	{k,o:3}, {e,o:3}, {k,e,o:3}
m	{ {k,e:2}, {k:1} }	k:3	{k,m: 3}
e	{ {k:4} }	k:4	{ k,e:4 }

$$\forall x \in transaction, buys(X, item_1) \wedge buys(X, item_2) \Rightarrow buys(X, item_3) \quad [s, c]$$

$$k, o \rightarrow e \quad [0.6, 1]$$

$$e, o \rightarrow k \quad [0.6, 1]$$

{O,K}→{E}, 支持度0.6, 置信度1 {O,E}→{k}, 支持度0.6, 置信度1

第8章:

(a)修改需要考虑了对每个广义数据元组的计数:从以下两方面考虑:

每个元组的计数必须集成到属性选择的计算中(例如**信息增益**)。

考虑到**确定元组中最常见的类**。

(salary = 26K...30K:

junior

= 31K...35K:

junior

= 36K...40K:

senior

= 41K...45K:

(b) junior

= 46K...50K (department = secretary:

junior

= sales:

senior

= systems:

junior

= marketing:

senior)

= 66K...70K:

senior)

(c)

解一： 设元组的各个属性之间相互独立，所以先求每个属性的类条件概率：

$$P(\text{systems}|\text{junior})=(20+3)/(40+40+20+3+4+6)=23/113;$$

$$P(26-30|\text{junior})=(40+3+6)/113=49/113;$$

$$P(46K-50K|\text{junior})=(20+3)/113=23/113;$$

$$\therefore X=(\text{department}=\text{system}, \text{age}=26\cdots 30, \text{salary}=46K\cdots 50K);$$

$$\begin{aligned}\therefore P(X|\text{junior}) &= P(\text{systems}|\text{junior})P(26-30|\text{junior})P(46K-50K|\text{junior}) \\ &= 23 \times 49 \times 23 / 113^3 = 25921 / 1442897 = 0.01796;\end{aligned}$$

$$P(\text{systems}|\text{senior})=(5+3)/(30+5+3+10+4)=23/52;$$

$$P(26-30|\text{senior})=(0)/53=0;$$

$$P(46K-50K|\text{senior})=(30+10)/52=40/52;$$

$$\therefore X=(\text{department}=\text{system}, \text{age}=26\cdots 30, \text{salary}=46K\cdots 50K);$$

$$\therefore P(X|\text{senior})=P(\text{systems}|\text{senior})P(26-30|\text{senior})P(46K-50K|\text{senior})=0;$$

$$\therefore P(\text{junior})=113/165=0.68;$$

$$\therefore P(\text{senior})=52/165=0.32;$$

$$\therefore P(X|\text{junior})P(\text{junior})=0.01796 \times 0.68=0.0122128 > 0=0=P(X|\text{senior})P(\text{senior});$$

所以：朴素贝叶斯分类器将 X 分到 junior 类。

解二： 设元组的各属性之间不独立，其联合概率不能写成份量相乘的形式。

所以已知： $X=(\text{department}=\text{system}, \text{age}=26\cdots 30, \text{salary}=46K\cdots 50K)$ ，元组总数为： $30+40+40+20+5+3+3+10+4+4+6=165$ 。

先验概率：

当 $\text{status}=\text{senior}$ 时，元组总数为： $30+5+3+10+4=52$ ， $P(\text{senior})=52/165=0.32$ ；

解二： 设元组的各属性之间不独立，其联合概率不能写成份量相乘的形式。

所以已知： $X=(\text{department}=\text{system}, \text{age}=26\cdots 30, \text{salary}=46\text{K}\cdots 50\text{K})$ ，元组总数为： $30+40+40+20+5+3+3+10+4+4+6=165$ 。

先验概率：

当 $\text{status}=\text{senior}$ 时，元组总数为： $30+5+3+10+4=52$ ， $P(\text{senior})=52/165=0.32$ ；

当 $\text{status}=\text{junior}$ 时，元组总数为： $40+40+20+3+4+6=113$ ，
 $P(\text{junior})=113/165=0.68$ ；

因为 $\text{status}=\text{senior}$ 状态没有对应的 $\text{age}=26\cdots 30$ 区间，所以： $P(X|\text{senior})=0$ ；

因为 $\text{status}=\text{junior}$ 状态对应的 $\text{partment}=\text{systems}$ 、 $\text{age}=26\cdots 30$ 区间的总元组数为： 3，所以： $P(X|\text{junior})=3/113$ ；

因为： $P(X|\text{junior})P(\text{junior})=3/113 \times 113/165=0.018>0=P(X|\text{senior})P(\text{senior})$ ；

所以：朴素贝叶斯分类器将 X 分到 junior 类。

- 9.1 下表由取自雇员数据库的训练数据组成。数据已泛化。例如，*age* “31...35”表示年龄在31~35岁之间。对于给定的行，*count*表示*department*、*status*、*age*和*salary*在该行上具有给定值的数据元组数。

<i>department</i>	<i>status</i>	<i>age</i>	<i>salary</i>	<i>count</i>
sales	senior	31...35	46K...50K	30
sales	junior	26...30	26K...30K	40
sales	junior	31...35	31K...35K	40
systems	junior	21...25	46K...50K	20
systems	senior	31...35	66K...70K	5
systems	junior	26...30	46K...50K	3
systems	senior	41...45	66K...70K	3
marketing	senior	36...40	46K...50K	10
marketing	junior	31...35	41K...45K	4
secretary	senior	46...50	36K...40K	4
secretary	junior	26...30	26K...30K	6

设 *status* 是类标号属性。

- (a) 为给定的数据设计一个多层前馈神经网络。标记输入层和输出层节点。
- (b) 给定训练实例 (*sales*, *senior*, 31...35, 46K...50K)，使用 (a) 中得到的多层前馈神经网络，给出后向传播算法一次迭代后的权重。指出你使用的初始权重和偏倚以及学习率。