



CRM介绍

华东理工大学计算机系
霍吉

作业

- 1、霓裳公司在线销售流行女装。该公司希望通过关联算法分析哪些流行元素结合在一起会令服饰热卖。该公司对最近网店的2000件服饰作了分析，其中：
- 200件仅含有蕾丝元素，100件仅含亮片元素，300件仅含花朵元素；
- 400件是蕾丝加亮片元素，500件是蕾丝加花朵元素，200件是亮片加花朵元素；
- 100件是蕾丝、亮片加花朵元素；
- 200件没有蕾丝、亮片及花朵元素。
- 假设A代表蕾丝元素，B代表亮片元素，C代表花朵元素。
- (1) 求最小支持度为35%的频繁项目集；
- (2) 求频繁项目集中兴趣度大于等于1的所有关联组合。

2

(1) 频繁项目集 (评分20分)

蕾丝A: $200+400+500+100=1200$ $1200/2000=60\%$

亮片B: $100+400+200+100=800$ $800/2000=40\%$

花朵C: $300+500+200+100=1100$ $1100/2000=55\%$

AB: $400+100=500$ $500/2000=25\%$

BC: $200+100=300$ $300/2000=15\%$

AC: $500+100=600$ $600/2000=30\%$

所以频繁项目集: {A, B, C}

(2) 兴趣度大于1的关联组合 (评分30分)

AB的兴趣度=A和B的支持度/ (A的支持度*B的支持度) = $25\% / (60\% * 40\%) = 1.04 > 1$

BC的兴趣度=B和C的支持度/ (B的支持度*C的支持度) = $15\% / (40\% * 55\%) = 0.68 < 1$

AC的兴趣度=A和C的支持度/ (A的支持度*C的支持度) = $30\% / (60\% * 55\%) = 0.91 < 1$

所以AB是符合要求的组合

作业

- 2、云裳电子商务公司为男士或女士定制专属礼服，客户在线提交自己的尺寸以及喜欢的流行元素。公司与客户在线确定服装的设计图和面料，最终将礼服快递给客户。请分析以下问题。
- 营销数据库中客户的样例如下。用归一化绝对值分别为定制服装的次数、消费总额建立距离函数，然后用欧几里得几何距离函数建立组合函数。请问若有一位客户定制服装次数为4，消费总额为38890元，则该客户与哪个样例相似可以采用相似的优惠策略。

客户样例号	定制服装的次数	消费总额
1	10	18000
2	6	60000
3	2	10000

3

4

(评分20分)

步骤1:

$|10-4|/(10-2)=0.75$ $|38890-18000|/(60000-10000)=0.4178$

$|6-4|/(10-2)=0.25$ $|38890-60000|/(60000-10000)=0.4222$

$|2-4|/(10-2)=0.25$ $|38890-10000|/(60000-10000)=0.5778$

(评分30分)

步骤2:

$d1-4 = \text{Sqrt}(0.75^2 + 0.4178^2) = 0.8585$

$d2-4 = \text{Sqrt}(0.25^2 + 0.4222^2) = 0.4907$

$d3-4 = \text{Sqrt}(0.25^2 + 0.5778^2) = 0.6286$

因为 $d2-4 < d3-4 < d1-4$

所以和样例2相似

(注: 步骤1和步骤2合在一起也可以😊)



谢谢大家!

华东理工大学计算机系
霍吉

5