

Regras de Associação

SCC0230 - Inteligência Artificial

Solange Oliveira Rezende

Conteúdo

1. Introdução
2. Conceitos Básicos
3. Algoritmo APRIORI para Geração de Itemsets Frequentes e Regras de Associação
4. Gerando Regras de Associação
5. Medidas de Interesse
 - a. Lift
 - b. Rule Interest
 - c. Convicção
 - d. Outras medidas de interesse objetivas

Regras de Associação

Objetivo:

Encontrar tendências que possam ser usadas para entender e explorar padrões de comportamento dos dados.

Exemplo:

Uma Base de Dados de um supermercado teria como regra o fato de que 80% dos clientes que compram um produto Q , também adquirem, na mesma ocasião o produto W . No qual 80% é o fator de confiança da regra.

Problema:

Analisar um grande volume de conhecimento extraído no formato de regras.

Regras de Associação

- Identifica itens que tem grande probabilidade de ocorrerem juntos em uma mesma transação da base de dados.
- Comumente utilizada em análise de cestas de mercado.



- Exemplo de uma regra de associação:
fralda → *cerveja*, indicando que o cliente que compra *fralda*, tende a comprar *cerveja*.

Introdução

A descoberta de relações entre itens em grandes quantidades de transações de negócios podem ajudar em muitos processos de tomada de decisões

- Análise de comportamento do consumidor
- Recomendação de itens
- Definição de layouts
- Desenvolvimento de estratégia de marketing

As transações podem estar armazenadas em diferentes formatos

Transações

```
transações.txt
1 leite pão manteiga café
2 coca-cola fandangos
3 vinho queijo_frescal
4 arroz feijão milho ervilha
5 banana maçã pêra
```

Figura: Exemplo de um arquivo de transações no qual os itens de uma **transação estão dispostos de maneira sequencial**

transaction ID	milk	bread	butter	beer
1	1	1	0	0
2	0	0	1	0
3	0	0	0	1
4	1	1	1	0
5	0	1	0	0

Figura: Exemplo de um arquivo de transações dispostos em uma tabela na qual **cada coluna corresponde a um item**

case ID	attribute
TRANS_ID	ITEM_ID
11	B
11	D
11	E
12	A
12	B
12	C
12	E
13	B
13	C
13	D
13	E

Figura: Exemplo de um arquivo de transações no qual cada **item de uma transação corresponde a** uma entrada na base de dados

Conceitos Básicos: Itemset

Um ***itemset*** é um conjunto de atributos ou itens.

Ou seja, o itemset $\mathcal{I} = \{l_1, l_2, \dots, l_m\}$

é o conjunto formado pelos itens l_1, l_2, \dots, l_m .

Exemplos:

{pão, manteiga}

{fralda, cerveja, arroz}

{banana, arroz, feijão, pão}

Conceitos Básicos: Suporte

O **suporte** indica a frequência com que um *itemset* ou com que A e B ocorrem juntos no conjunto de dados.

Suporte de $A \rightarrow B$: Porcentagem das transações de D que contém $A \cup B$

Exemplos:

transaction ID	milk	bread	butter	beer
1	1	1	0	0
2	0	0	1	0
3	0	0	0	1
4	1	1	1	0
5				0

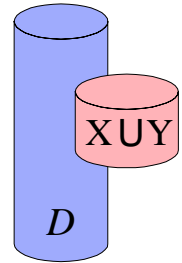
X1	X2	X3
a	b	c
d	e	f
a	b	c

$$P(A \cup B) = \frac{\text{cont}(A \& B)}{|T|}$$

$$\text{sup}(\{\text{milk}, \text{bread}\}) = 2 / 5$$

$$\text{sup}(\{\text{milk}, \text{butter}\}) = 1 / 5$$

$$\text{sup}(\{\text{milk}, \text{butter}, \text{beer}\}) = 0$$



suporte

O *Itemset*: **de** tem $\text{sup} = 1$

A Regra: **ab** \rightarrow **c** = **abc** tem $\text{sup} = 2$

Conceitos Básicos: Regra de Associação

Uma **regra de associação** é uma implicação da forma

$$A \Rightarrow B \quad \text{qual} \quad A \subset \mathcal{I}, B \subset \mathcal{I}, A \neq \emptyset, B \neq \emptyset, \text{ e } A \cap B = \emptyset$$

Exemplo:

$$\{\text{arroz}\} \Rightarrow \{\text{feijão}\}$$

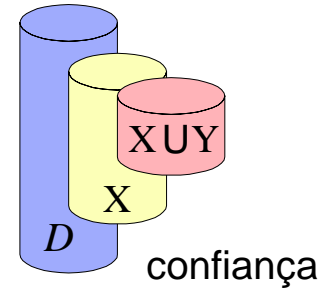
$$\{\text{pão, manteiga}\} \Rightarrow \{\text{leite}\}$$

Conceitos Básicos: Confiança

Sejam A e B dois conjuntos de itens. A **confiança** de uma regra $A \Rightarrow B$ é a probabilidade condicional da transação conter o conjunto de itens B , dado que contém o conjunto A .

Ou seja,

$$conf(A \Rightarrow B) = P(B|A) = \frac{sup(A \Rightarrow B)}{sup(A)}$$



Exemplo:

transaction ID	milk	bread	butter	beer
1	1	1	0	0
2	0	0	1	0
3	0	0	0	1
4	1	1	1	0
5	0	1	0	0

$$sup(\{\text{milk}, \text{bread}\}) = 2 / 5$$

$$sup(\{\text{milk}\}) = 2 / 5$$

$$conf(\text{milk} \Rightarrow \text{bread}) = \frac{sup(\text{milk} \cup \text{bread})}{sup(\text{milk})} = 1$$

Definição

Seja **D** uma Base de Dados composta por:

- um conjunto de itens $I = \{I_1, I_2, \dots, I_m\}$ ordenados lexicograficamente (*itemset*);
- e por um conjunto de transações **T**, em que cada transação t_i é composta por um conjunto de itens tal que t_i está contido em **I**.
- Uma Regra de Associação é uma implicação na forma **A** \Rightarrow **B**, em que **A** e **B** estão contidos em **I** e a intersecção de **A** e **B** é vazia.
- A regra **A** \Rightarrow **B** ocorre no conjunto de transações **T** com confiança **c** se **c%** das transações em **T** em que ocorre **A** também ocorre **B**.
- A regra **A** \Rightarrow **B** tem suporte **s** se em **s%** das transações em **D** ocorre **A** \cup **B**.

Conceitos Básicos

- Em geral, a mineração de regras de associação pode ser dividida em dois passos:
 - **Encontrar todos os itemsets frequentes:** cada um dos itemsets irá ocorrer com frequência igual ou superior ao limiar determinado pelo usuário (suporte mínimo)
 - **Gerar regras de associação fortes a partir dos itemsets frequentes:** gerar regras usando os itemsets frequentes extraídos no passo anterior de forma que a confiança das regras seja maior que um limiar de confiança definido pelo usuário (confiança mínima)
- O segundo passo é muito menos custoso computacionalmente, portanto o custo computacional para a extração de regras é determinado pelo primeiro passo.

Conceitos Básicos

- As regras de associação são geradas se estas satisfazem ambos o suporte e confiança mínimos
- Os valores de suporte e confiança mínimos são definidos pelo usuário ou especialistas de domínio
- Uma análise adicional pode ser realizada para descobrir correlações estatísticas interessantes entre os itens associados.

Algoritmo Apriori

- O algoritmo **Apriori** [Agrawal and Srikant, 1994] é utilizado para a extração dos itemsets frequentes que são utilizados para gerar as regras de associação
- É executado um processo iterativo no qual k-itemsets são utilizados para gerar (k+1)-itemsets
 - Primeiramente, o conjunto de 1-itemset frequente é gerado (L1)
 - A seguir, L1 é utilizado para gerar o conjunto de 2-itemsets frequentes (L2)
 - L2 é utilizado para gerar L3 e assim por diante, até que nenhum k-itemset seja encontrado
- Em geral, um conjunto de dados que contém k itens pode gerar potencialmente $2^k - 1$ itemsets frequentes

Algoritmo Apriori

Um *lattice* pode ser utilizado para listar todos os possíveis itemsets

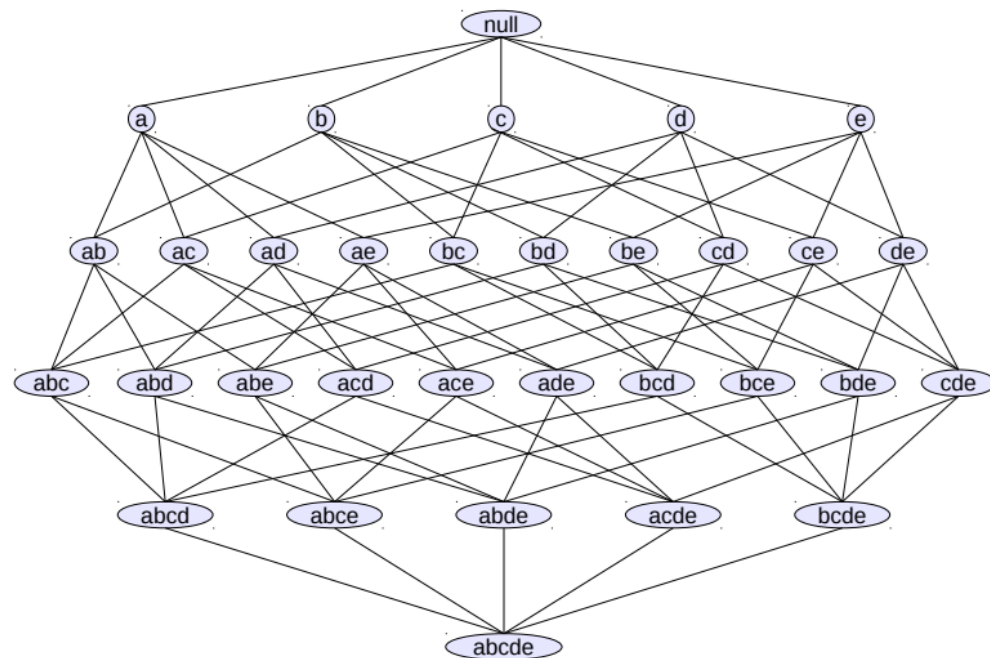


Figura: Lattice dos conjuntos de itemsets gerados por $\mathcal{I} = \{a, b, c, d, e\}$

Algoritmo Apriori

- Para melhorar a eficiência da geração de itemsets frequentes em cada nível, uma propriedade importante é utilizada para reduzir o espaço de busca
- **Propriedade do Apriori**
 - **“Todo subconjunto não vazio de um itemset frequente também é frequente”**
 - Se um itemset I não satisfaz o suporte mínimo, a adição de um item A em I não fará com que $(A \cup I)$ seja frequente.

Algoritmo Apriori

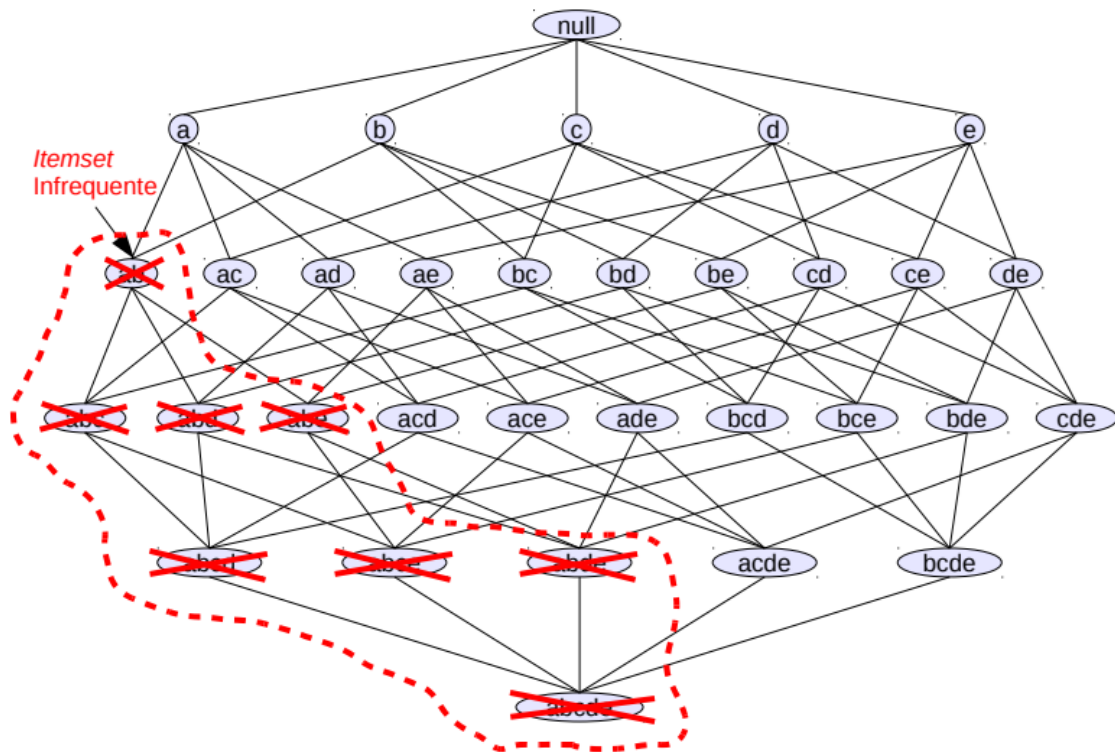


Figura: Redução do espaço de busca pela utilização da propriedade do Apriori

Algoritmo Apriori

- A princípio, existem várias formas de gerar os candidatos a itemsets
- Os requisitos para a geração dos itemsets frequentes são:
 - Não se deve gerar muitos candidatos desnecessários (um candidato é desnecessário se ao menos um de seus subconjuntos é infrequente)
 - Deve-se assegurar que o conjunto de candidatos é completo (nenhum itemset frequente é excluído da geração de candidatos)
 - Não se deve gerar o mesmo itemset candidato mais de uma vez

Algoritmo Apriori

- **Método $F_{k-1} \times F_1$**
 - **Estender cada (k-1)-itemset** com os 1-itemsets frequentes
 $\{\text{Cerveja, Fralda}\} \cup \{\text{Pão}\}$ formam o itemset candidato
 $\{\text{Cerveja, Fralda, Pão}\}$
 - **O método é completo** pois cada k-itemset é composto de um (k-1)-itemset e um 1-itemset frequente
 - Não prevê que o mesmo itemset candidato não seja gerado mais de uma vez. Exemplo:
 $\{\text{Cerveja, Fralda, Leite}\}$ pode ser gerado pelas junções de:
 - $\{\text{Fralda, Pão}\} \cup \{\text{Cerveja}\}$
 - $\{\text{Cerveja, Leite}\} \cup \{\text{Fralda}\}$
 - $\{\text{Fralda, Leite}\} \cup \{\text{Pão}\}$
 - ...

Algoritmo Apriori

- **Método $F_{k-1} \times F_1$**
 - Uma forma de **evitar a geração de candidatos duplicados** é mantendo uma **ordem lexicográfica dos itemsets frequentes**
 - Cada $(k-1)$ -itemset X é então estendido com itens frequentes que são lexicograficamente maiores que X
 - {Cerveja, Fralda} pode ser aumentado com {Leite}
 - {Fralda, Pão} não pode ser aumentado com {Cerveja}
 - Ainda pode haver geração de candidatos desnecessários
 - Supondo que temos os candidatos {Cerveja, Fralda} e {Leite}
 - Será gerado o itemset {Cerveja, Fralda, Leite}
 - Porém, pode ser que {Cerveja, Leite} não seja frequente, e portanto {Cerveja, Fralda, Leite} também não será

Algoritmo Apriori

- **Método $F_{k-1} \times F_1$**
 - A geração de um **itemset candidato** é feita pela **junção de dois** (k-1)-itemsets **somente se** os primeiros (k-2)-itemsets são idênticos
 - Por exemplo, os itemsets **{Cerveja, Fralda}** e **{Cerveja, Leite}** são unidos para formar o itemset candidato **{Cerveja, Fralda, Leite}**
 - Por exemplo, não é necessário unir os itemsets {Cerveja, Fralda} com {Fralda, Leite} para formar o itemset {Cerveja, Fralda, Leite} pois esse já foi obtido pela junção dos itemsets {Cerveja, Fralda} e {Cerveja, Leite}

Algoritmo Apriori

- **Redução do número de transações verificadas em iterações posteriores**
 - Uma transação que não contém nenhum *k-itemset* frequente não irá conter um $(k+1)$ -*itemset* frequente
 - Marcar ou remover esta transação
- **Particionar os dados para encontrar itemsets candidatos**
 - **Fase I:** dividir o conjunto de transações D em n partições não sobrepostas
 - Um itemset que é frequente em D deve ser frequente em ao menos uma das partições
 - Se um itemset não é frequente em uma das partições, também não deve ser frequente em D
 - **Fase II:** fazer a contagem do suporte considerando somente os itemsets extraídos na Fase I

Algoritmo Apriori

- Amostragem
 - Escolher um conjunto aleatório de exemplos S em D e então executar a busca por itemsets frequentes em S
 - *Trade-off* entre acurácia e eficiência
 - Normalmente é utilizado um suporte menor em S para evitar que se percam itemsets frequentes
- ...

Além do algoritmo Apriori, outras técnicas para a geração de itemsets frequentes podem ser encontradas na literatura:

- FP-growth [Han et al., 2000]
- Apriori TID [Agrawal et al., 1996]
- *Opus* [Webb, 1995];
- *Direct Hasing and Pruning* (DHP) [Adamo, 2001];
- *Dynamic Set Counting (Dic)* [Adamo, 2001];
- *Charm* [Zeng, Kohavi e Manson, 2001];
- *FP-growth* [Zeng, Kohavi e Manson 2001];
- *Closet* [Zeng, Kohavi e Manson 2001];

Principais Algoritmos

❑ **Apriori**

Rakesh Agrawal and Ramakrishnan Srikant. Fast Algorithms for Mining Association Rules. In Proc. of the 20th Int'l Conference on Very Large Databases (VLDB '94), Santiago, Chile, September 1994. <http://citeseer.comp.nus.edu.sg/agrawal94fast.html>

❑ **FP-Growth** - que usa uma árvore de padrões frequentes (FP-Tree)

Han, J., Pei, J., and Yin, Y. 2000. Mining frequent patterns without candidate generation. In Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data (Dallas, Texas, United States, May 15 - 18, 2000). SIGMOD '00. ACM Press, New York, NY, 1-12.

DOI= <http://doi.acm.org/10.1145/342009.335372>

Considerações sobre os Algoritmos

- Os diversos algoritmos devem gerar sempre o mesmo conjunto de conhecimento.
- Mas qual é a diferença entre eles?
 - Forma como os dados são carregados na memória;
 - Tempo de processamento;
 - Tipos de atributos (Numéricos, Categóricos);
 - A maneira como os algoritmos geram os *itemsets*;

Gerando Regras de Associação

- Uma regra de associação pode ser extraída por particionar o itemset Y em dois subconjuntos não vazios, X e $Y - X$, tal que $X \Rightarrow Y - X$ satisfaça a confiança mínima
- Seja $X = \{a, b, c\}$, pode-se gerar as seguintes regras
 - $\{a, b\} \Rightarrow \{c\}$
 - $\{a, c\} \Rightarrow \{b\}$
 - $\{b, c\} \Rightarrow \{a\}$
 - $\{a\} \Rightarrow \{b, c\}$
 - $\{c\} \Rightarrow \{a, b\}$
- Cada k -itemset frequente Y pode produzir $2^k - 2$ regras de associação (ignorando regras com $\emptyset \Rightarrow Y$ e $Y \Rightarrow \emptyset$)
- Computar a confiança de uma regra de associação não requer verificações adicionais no conjunto de transações

Gerando Regras de Associação

Exemplo: Extrair regras com sup-min = 30% e conf-min = 70%

Tabela: Conjunto de Transações

T	Leite	Café	Cerveja	Pão	Manteiga	Arroz	Feijão
1	N	S	N	S	S	N	N
2	S	N	S	S	S	N	N
3	N	S	N	S	S	N	N
4	S	S	N	S	S	N	N
5	N	N	S	N	N	N	N
6	N	N	N	N	S	N	N
7	N	N	N	S	N	N	N
8	N	N	N	N	N	N	S
9	N	N	N	N	N	S	S
10	N	N	N	N	N	S	N

Gerando Regras de Associação

Exemplo

- **Calculando o suporte dos 1-itemsets**

- $\text{sup}(\{\text{Leite}\}) = 20\%$
- $\text{sup}(\{\text{Café}\}) = 30\%$
- $\text{sup}(\{\text{Cerveja}\}) = 20\%$
- $\text{sup}(\{\text{Pão}\}) = 50\%$
- $\text{sup}(\{\text{Manteiga}\}) = 50\%$
- $\text{sup}(\{\text{Arroz}\}) = 20\%$
- $\text{sup}(\{\text{Feijão}\}) = 20\%$

T	Leite	Café	Cerveja	Pão	Manteiga	Arroz	Feijão
1	N	S	N	S	S	N	N
2	S	N	S	S	S	N	N
3	N	S	N	S	S	N	N
4	S	S	N	S	S	N	N
5	N	N	S	N	N	N	N
6	N	N	N	N	S	N	N
7	N	N	N	S	N	N	N
8	N	N	N	N	N	N	S
9	N	N	N	N	N	S	S
10	N	N	N	N	N	S	N

- Somente {Café}, {Pão} e {Manteiga} são frequentes
- **Combinando os 1-itemsets frequentes e calculando o suporte**
 - $\text{sup}(\{\text{Café}, \text{Manteiga}\}) = 30\%$
 - $\text{sup}(\{\text{Café}, \text{Pão}\}) = 30\%$
 - $\text{sup}(\{\text{Manteiga}, \text{Pão}\}) = 40\%$



Os três 2-itemsets
são frequentes

Gerando Regras de Associação

Exemplo

- Calculando a confiança para **regras geradas de {Café, Pão}**
 - $\text{conf}(\{\text{Café}\} \Rightarrow \{\text{Pão}\}) = 30\% / 30\% = 100\%$
 - $\text{conf}(\{\text{Pão}\} \Rightarrow \{\text{Café}\}) = 30\% / 50\% = 60\%$
- Calculando a confiança para **regras geradas de {Café, Manteiga}**
 - $\text{conf}(\{\text{Café}\} \Rightarrow \{\text{Manteiga}\}) = 30\% / 30\% = 100\%$
 - $\text{conf}(\{\text{Manteiga}\} \Rightarrow \{\text{Café}\}) = 30\% / 50\% = 60\%$
- Calculando a confiança para **regras geradas de {Manteiga, Pão}**
 - $\text{conf}(\{\text{Manteiga}\} \Rightarrow \{\text{Pão}\}) = 40\% / 50\% = 80\%$
 - $\text{conf}(\{\text{Pão}\} \Rightarrow \{\text{Manteiga}\}) = 40\% / 50\% = 80\%$

Gerando Regras de Associação

Exemplo

- $\text{sup}(\{\text{Café, Manteiga, Pão}\}) = 30\%$
- Calculando a confiança para regras geradas de $\{\text{Café, Manteiga, Pão}\}$
 - $\text{conf}(\{\text{Café, Pão}\} \Rightarrow \{\text{Manteiga}\}) = 30\% / 30\% = 100\%$
 - $\text{conf}(\{\text{Café, Manteiga}\} \Rightarrow \{\text{Pão}\}) = 30\% / 30\% = 100\%$
 - $\text{conf}(\{\text{Manteiga, Pão}\} \Rightarrow \{\text{Café}\}) = 30\% / 40\% = 75\%$
 - $\text{conf}(\{\text{Café}\} \Rightarrow \{\text{Pão, Manteiga}\}) = 30\% / 30\% = 100\%$
 - ...

Medidas de Interesse

- **O número de regras geradas utilizando somente suporte e confiança geralmente é muito grande**
 - As vezes muitas Regras não interessantes
 - Dificuldade de análise por parte do usuário
- Itens com suporte muito alto acabam por compor a maioria das regras
- **Mesmo com valores de suporte e confiança altos, as regras de associações podem expressar relações não válidas**
- Medidas de interesse são comumente utilizadas para avaliar a qualidade de uma regra de associação

Medidas de Interesse

Tabela: Conjunto de regras de exemplo

ID	Regra	Sup _A	Sup _B	Sup	Conf
1	{cenoura} \Rightarrow {batata_inglesa}	0,7701	0,8175	0,7038	0,9138
2	{acerola} \Rightarrow {ovos}	0,0427	0,8886	0,0379	0,8889
3	{filé_de_viola} \Rightarrow {açúcar_refinado}	0,0877	0,8649	0,0758	0,8649
4	{milho_verde} \Rightarrow {ervilhas}	0,3294	0,3791	0,2701	0,8201
5	{fruta_do_conde} \Rightarrow {melancia}	0,0450	0,1422	0,0308	0,6842
6	{banana_nanica} \Rightarrow {banana_prata}	0,1209	0,7607	0,0735	0,6078

Medidas de Interesse - Independência entre itens

- **A confiança da regra 3 ($\{\text{filé_de_viola}\} \Rightarrow \{\text{açúcar_refinado}\}$) é 0,8649**
 - Se observarmos, o suporte de $\{\text{açúcar_refinado}\}$ é 0,8649, portanto a compra de $\{\text{filé_de_viola}\}$ não aumenta nem diminui a chance de comprar $\{\text{açúcar_refinado}\}$
 - A compra de $\{\text{açúcar_refinado}\}$ independe da compra de $\{\text{filé_de_viola}\}$
- O mesmo acontece para a regra 2 ($\{\text{acerola}\} \Rightarrow \{\text{ovos}\}$)

Medidas de Interesse - Independência entre itens

- Os itemsets são independentes se

$$\text{conf}(A \Rightarrow B) = \text{sup}(B)$$

- Lembrando que $\text{conf}(A \Rightarrow B) = \text{sup}(A \cup B) / \text{sup}(A)$

- Portanto, os itemsets são independentes se

$$\text{sup}(A \cup B) = \text{sup}(A) \times \text{sup}(B)$$

$\text{sup}(A) \times \text{sup}(B)$ representa o **suporte esperado** (sub_esp) de $A \cup B$

Medidas de Interesse - Dependência negativa entre itens

- A confiança da regra 6 {banana_nanica} => {banana_prata} é 0,6078
 - A probabilidade de qualquer cliente comprar {banana_prata} é 0,7607
 - A compra de {banana_nanica} diminui a chance de comprar {banana_prata}
- Os itemsets de uma regra possuem dependência negativa se $sup(A \cup B) < sup_esp(A \cup B)$

Medidas de Interesse - Dependência positiva entre itens

- **A confiança da regra 4 {milho_verde} => {ervilhas} é 0,8201**

 - A probabilidade de qualquer cliente comprar {ervilhas} é 0,3791
 - A compra de {milho_verde} aumenta a chance de comprar {ervilhas}
 - Os clientes que compraram {milho_verde} tem maior probabilidade de comprar {ervilhas}
- **Os itemsets de uma regra possuem dependência negativa se $sup(A \cup B) > sup_{esp}(A \cup B)$**

Medidas de Interesse

- A medida confiança, por não considerar a dependência entre os itemsets, pode gerar um número muito grande de regras que apresentam relacionamentos falsos
- Utilização de outras análises por meios de medidas de interesse objetivas são importantes para selecionar regras interessantes

Medidas de Interesse: Lift

A medida **Lift**, também conhecida por *Interest*, é uma das mais utilizadas para avaliar dependências entre itemsets

$$\text{lift}(A \Rightarrow B) = \frac{\text{conf}(A \Rightarrow B)}{\text{sup}(B)} = \frac{\text{sup}(A \cup B)}{\text{sup}(A) \times \text{sup}(B)}$$

- $\text{lift}(A \Rightarrow B) = 1$: A e B são independentes
- $\text{lift}(A \Rightarrow B) > 1$: A e B são positivamente dependentes
- $\text{lift}(A \Rightarrow B) < 1$: A e B são negativamente dependentes
- Intervalo: $[0, \infty[$
- Simétrica

Quanto maior o valor de Lift, mais interessante a regra

Medidas de Interesse: Rule Interest

A medida ***Rule Interest***, também conhecida por Piatetsky-Shapiro, novidade ou *leverage*, também é utilizada para avaliar dependência entre os itemsets

$$ri(A \Rightarrow B) = sup(A \cup B) - sup_esp(A \cup B)$$

- $ri(A \Rightarrow B) = 0$: A e B são independentes
- $ri(A \Rightarrow B) > 0$: A e B são positivamente dependentes
- $ri(A \Rightarrow B) < 0$: A e B são negativamente dependentes
- Intervalo: $[-0,25 ; 0,25]$
- Simétrica

Quanto **maior o valor da medida**, mais interessante é a regra

Medidas de Interesse: Convicção

Objetivo é avaliar a “implicação” da regra

$$\text{conv}(A \Rightarrow B) = \frac{\text{sup}(A) \times \text{sup}(\neg B)}{\text{sup}(A \cup \neg B)}$$

- Avalia o quanto A e $\neg B$ se afastam da independência
- Caso exista independência completa entre o antecedente e o conseqüente da regra, o valor da convicção será igual a 1
- Regras onde o antecedente nunca aparece sem o conseqüente (confiança de 100%) terão valor de convicção igual a ∞
- Intervalo: $[0, \infty[$
- Assimétrica

Mineração de Regras de Associação

Ao se tratar de mineração de regras de associação, pode-se dividir este processo em três etapas (AGGARWAL; PROCOPIUC; YU, 2002; AGRAWAL; IMIELINSKI; SWAMI, 1994; NGUYEN et al., 2015):

- **Pré-processamento:** Preparação da base para a etapa de extração e preparação do domínio dos dados, podendo ocorrer a remoção de itens não interessantes.
- **Extração de padrões:** São efetuados os cálculos de medidas, construção dos *itemsets* frequentes e construção das regras de associação propriamente ditas.
- **Pós-processamento:** efetua-se a remoção de regras não interessantes, divisão do domínio e redução do número de regras a serem exploradas pelo usuário.

Outras medidas de interesse

- Outras medidas de interesse objetivas podem ser encontradas em:
 - [Geng and Hamilton, 2006]
 - [Guillet and Hamilton, 2007]
 - [Tan et al., 2002]