

A primeira Escola presencial gratuita de Inteligência Artificial do Brasil



Aula Extra 21/03/2020:
Apriori - Cesto de Compras

Apoio:



Professor: Flavia Reis

Objetivo da Aula

A aula sobre Apriori tem como objetivo mostrar um exemplo pratico da aplicação do Modelo Apriori no estudo de Cesto de compras que é uma análise bem comum no Mercado de varejo principalmente.

De forma bem simples, é possível gerar uma análise que agrega muito ao negócio, Trazendo informações que podem maximizar as vendas e ofertas de acordo com as Decisões de negócios aplicadas com as descobertas oferecidas pelo modelo.

Boa aula.

Tópicos

- MBA? Não, eu só vim fazer um curso livre...
- Regras de Associação
- Análise de Cesta de Compras na Prática

MBA - *Market Basket Analysis* (*Análise de Cesta de Compras*)

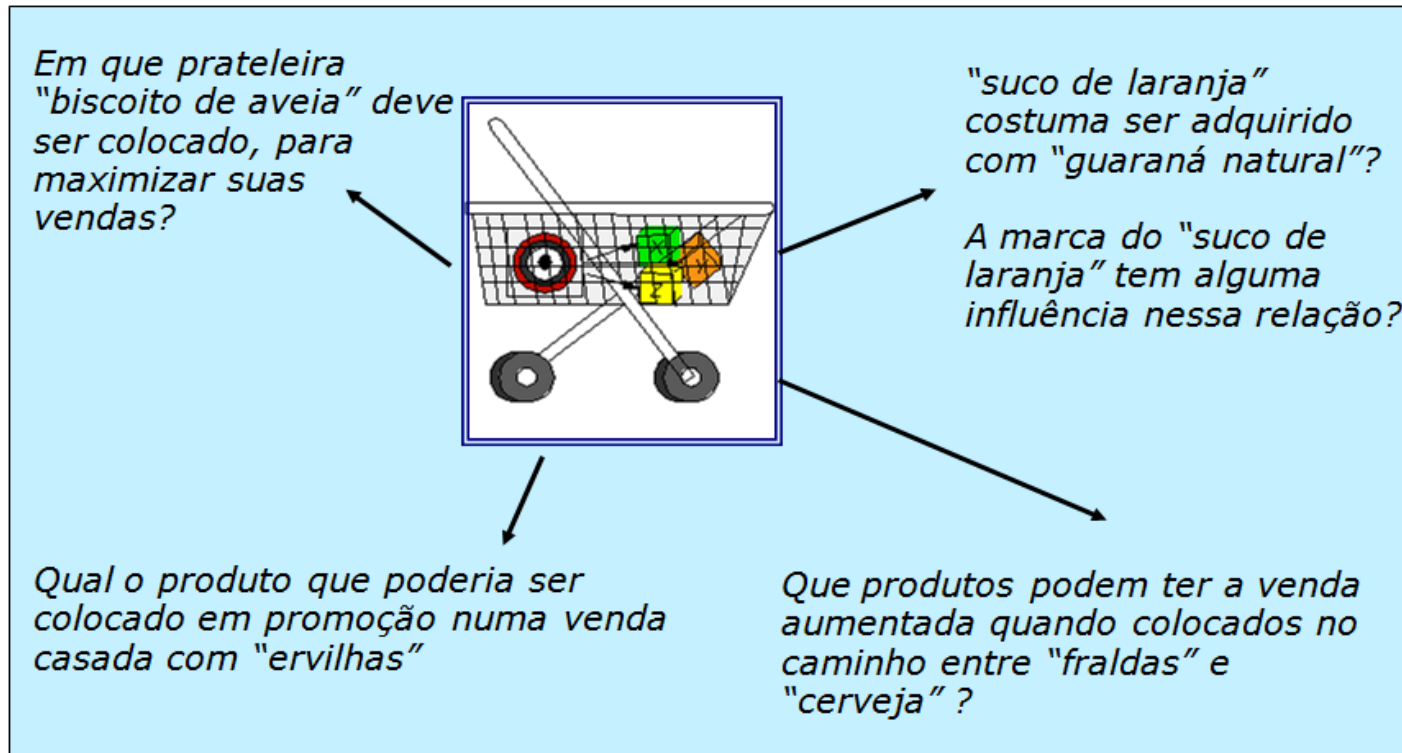
Market Basket Analysis, mais conhecida como análise de cesta de compras, é uma técnica de datamining que utiliza regras de associação para avaliar quais itens são frequentemente comprados juntos e permite entender o comportamento de compras dos clientes.

Saber o que os clientes tendem a comprar em conjunto pode ajudar a ofertar melhor através de *bundles* (), disposição dos itens, entre outras estratégias, alcançando melhores resultados de vendas.



Regras de Associação

Diariamente, surgem muitas dúvidas nos negócios.....



Ras, regras de associação são utilizadas para mostrar o relacionamento entre diversos itens de uma transação.

Regras de Associação

A tarefa de associação tem como objetivo encontrar elementos que implicam na presença de outros elementos em uma mesma transação, ou seja, encontrar relacionamentos ou padrões frequentes entre conjuntos.

Um exemplo de regra de associação poderia ser:
 $\{\text{mussarela, presunto}\} \rightarrow \{\text{pão}\}$



Que nos mostra que quando se compra mussarela e presunto, com um determinado grau de certeza, também compra-se pão

Apriori

Apriori é um dos algoritmos de regras de associação mais conhecido e utilizado em diversas aplicações.

Um conjunto de itens é chamado de *itemset*.

Transação	Itens
1	A, C, D
1	B, C, E
1	A, B, C, E
1	B, E
4	Total

Etapas principais do algoritmo:

- Encontrar os itemsets freqüentes (aqueles com suporte $\geq \text{min_sup}$)
- Gerar as regras a partir dos itemsets freqüentes (com confiança $\geq \text{min_conf}$)

Os valores min_sup e min_conf são parâmetros que devem ser fornecidos ao algoritmo.

Apriori

$$\text{support}(I) = \frac{\text{Number of transactions containing } I}{\text{Total number of transactions}}$$

Ex: Item E
3 / 4 = 75%

Significa que o item E
está presente em 75%
das transações.

Transação	Itens
1	A, C, D
1	B, C, E
1	A, B, C, E
1	B, E
4	Total

$$\text{confidence}(I1 \rightarrow I2) = \frac{\text{Number of transactions containing items } I1 \text{ and } I2}{\text{Total number of transactions containing } I1}$$

Ex: Item B -> E
3 / 3 = 100%

Significa que em 100% das
vezes que consta o item B,
também consta o item E

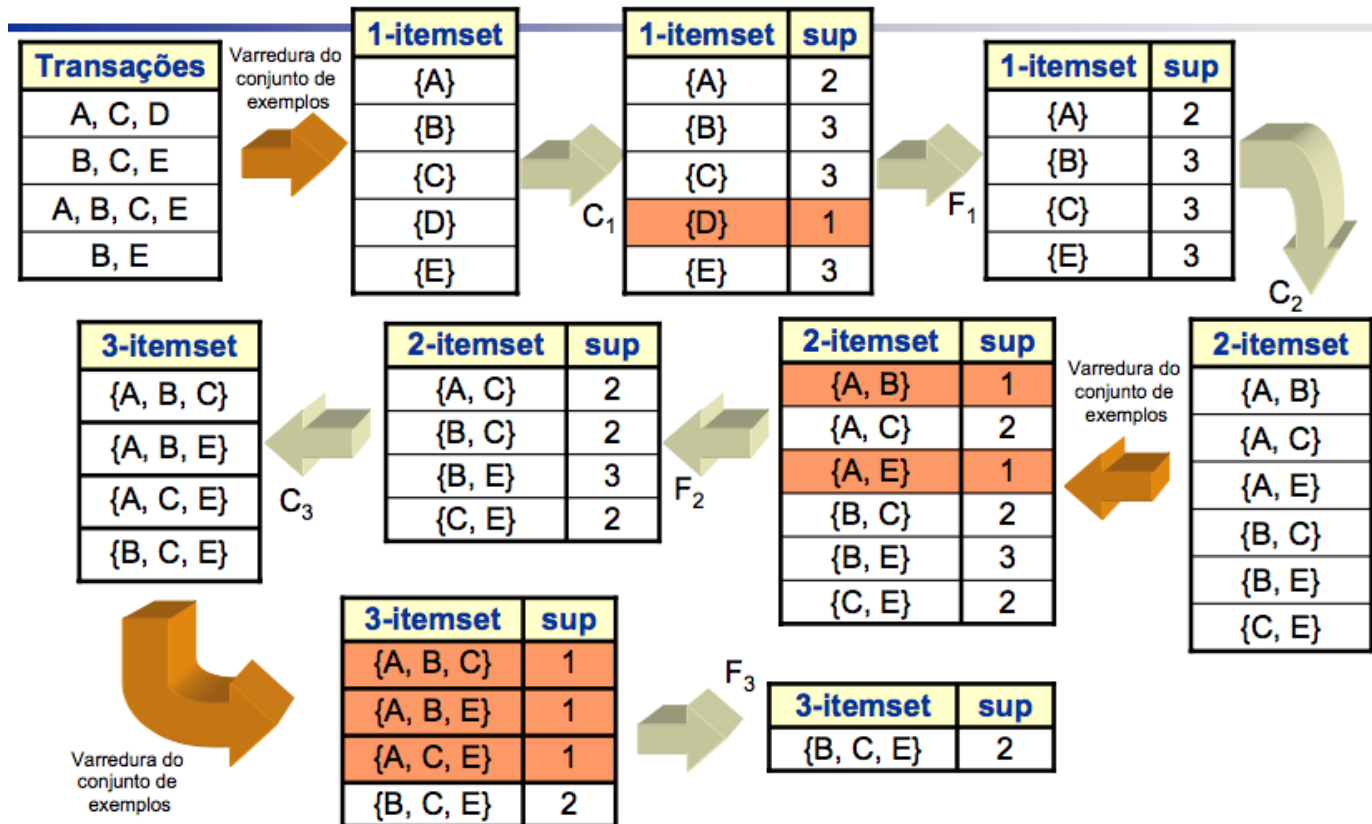
$$\text{lift}(I1 \rightarrow I2) = \frac{\text{confidence}(I1 \rightarrow I2)}{\text{support}(I2)}$$

Ex: Item B -> E
1 / 0.75 = 1,33
Positiva

Correlação de A e B:
> 1 Positiva
< 1 Negativa

Apriori

Exemplo, min_sup=2



Apriori

Redução de regras

Para diminuir a quantidade de regras retornadas, evita-se itens com poucas ocorrências, então define-se um valor mínimo de suporte (min_sup)

Princípio do Itemsets Frequentes

- Qualquer subconjunto de um itemset freqüente também é frequente;
- Se um itemset não é freqüente então não é necessário gerar quaisquer superconjuntos dele como candidatos, porque eles também serão não frequentes

Na geração de itemsets frequentes é utilizada a seguinte idéia:

Usar os 1-itemsets frequentes para gerar 2-itemsets, usar os 2- itemsets frequentes para gerar 3-itemsets e assim por diante.

Cesta de Compras - Python

Mão na Massa??

Bora abrir o Jupyter Notebook e codar....



Códigos

Pasta com notebooks, arquivos usados na análise e ppt

Entre no link abaixo e baixe os arquivos necessários:

<https://github.com/escolalivre-ia>

Exercícios

- 1) Cite outros exemplos de algoritmo usados para Market Basket Analysis, e explique qual a diferença entre eles.
- 2) O que é necessário para uma regra de associação ser considerada forte?
- 3) Através do valor retornado por esse parâmetro, podemos dizer se existe uma relação Direta entre os itens:
 - a) Confidence
 - b) Lift
 - c) Support
- 4) Regras de associação são considerados métodos de mineração de dados:
 - a) aprendizado supervisionado
 - b) não-supervisionado
 - c) Híbridos
 - d) Não são mineração de dados
 - e) NDA

Bibliografia

Apriori Algorithm in Python

<https://www.codespeedy.com/apriori-algorithm-in-python/>

Mineração de dados com Market Basket Analysis - Revista SQL Magazine 111

<https://www.devmedia.com.br/mineracao-de-dados-com-market-basket-analysis-revista-sql-magazine-111/27853>

Algoritmo Apriori - UFSM

<https://www.youtube.com/watch?v=suwzIKi27Ao>

Regras de associação

<http://dcm.ffclrp.usp.br/~augusto/teaching/ami/AM-I-Regras-Associacao.pdf>

Regras de associação

<https://pt.slideshare.net/nortoncg1/regras-de-associao-minerao-de-dados>

Regras de associação

<http://each.uspnet.usp.br/sarajane/wp-content/uploads/2015/11/associacao.pdf>

Regras de associação

https://pt.wikipedia.org/wiki/Regras_de_associa%C3%A7%C3%A3o

Market Basket Analysis

https://www.maxwell.vrac.puc-rio.br/14008/14008_4.PDF

Contatos



www.linkedin.com/in/flaviaxreis

