

Lista 4

Inteligência Artificial

Iyan Lucas Duarte Marques¹, Samir do Morim Cambraia¹, Wesley Filemon¹

¹Instituto de Ciências Exatas e Informática - Pontifícia Universidade Católica Minas Gerais (PUC-MG)

1. Questão 01

A partir da base de dados **PãoeManteiga Sim.csv** disponível no CANVAS, pede-se:

Produto	Suporte
Leite	$2/10 = 0.2$
Café	$3/10 = 0.3$
Cerveja	$2/10 = 0.2$
Pão	$5/10 = 0.5$
Manteiga	$5/10 = 0.5$
Arroz	$2/10 = 0.2$
Feijão	$2/10 = 0.2$

Tabela 1. Itemset = 1

Produto	Suporte
Café e Pão	$3/10 = 0.3$
Café e Manteiga	$3/10 = 0.3$
Pão e Manteiga	$4/10 = 0.4$

Tabela 2. Itemset = 2

Produto	Suporte
Café, Pão e Manteiga	$3/10 = 0.3$

Tabela 3. Itemset = 3

1.1. Encontre os itensSetse as regras de associação do problema, utilizando-se o algoritmo APRIORI

- *[Café, Pão]*
 - Se Café \rightarrow Pão Confiança= $3/3 = 1$
 - Se Pão \rightarrow Café Confiança= $3/5 = 0.6$
- *[Café, Manteiga]*
 - Se Café \rightarrow Manteiga Confiança= $3/3 = 1$
 - Se Manteiga \rightarrow Café Confiança= $3/5 = 0.6$
- *[Pão, Manteiga]*
 - Se Pão \rightarrow Manteiga Confiança= $4/5 = 0.8$
 - Se Manteiga \rightarrow Pão Confiança= $4/5 = 0.8$

- **[Café, Pão e Manteiga]**

- Se Café e Pão → Manteiga Confiança = $3/3 = 1$
- Se café e manteiga → Pão Confiança = $3/3 = 1$
- Se Pão e Manteiga → Pão Confiança = $3/4 = 0,75$
- Se Café → Pão e Manteiga Confiança = $3/3 = 1$
- Se Pão → Café e Manteiga Confiança = $3/5 = 0.6$
- Se Manteiga → Café e Pão Confiança = $3/5 = 0.6$

1.2. Rode a base de dados no WEKA (ou qualquer ferramenta que você esteja trabalhando), habilite o parâmetro para exibir os ItemsSets e verifique se o que você encontrou no item 1 está correto. Utilize os mesmos valores de suporte = 0.3 e confiança de 0.8

Ao todo foram encontrados 12 tipos de regras em relação a quem levou café e manteiga temos a regra se café → manteiga vai mostrar que todas as pessoas que levaram café também levou manteiga mas apenas 0.6 dos que levaram manteiga levaram café. Já em pão e manteiga ambos tiveram o mesmo resultado aqueles que levaram pão e manteiga e manteiga e pão tiveram o resultado de 0.8. Em relação aos três, nos temos que quem levou café e pão também levou manteiga, quem levou café e manteiga também levou pão, e se levou café também levou pão e manteiga, agora quem levou pão e manteiga também levou pão tiveram o resultado de 0,75. E quem levou pão também levou café e manteiga teve o resultado de 0.6, e quem levou manteiga e também levou café e pão também teve o resultado de 0,6.

2. Questão 02

A partir da base de dados PãoeManteigaSimNao.csv disponível no CANVAS, pede-se:

2.1. Realize os mesmos procedimentos da questão anterior e discuta sobre a quantidade de regras encontradas e sobre o tipo de regras

Produto	Suporte
Leite	$2/10 = 0.2$
Café	$3/10 = 0.3$
Cerveja	$2/10 = 0.2$
Pão	$5/10 = 0.5$
Manteiga	$5/10 = 0.5$
Arroz	$2/10 = 0.2$
Feijão	$2/10 = 0.2$

Tabela 4. Itemset = 1

- **[Café, Pão]**

- Se Café → Pão Confiança= $3/3 = 1$
- Se Pão → Café Confiança= $3/5 = 0.6$

- **[Café, Manteiga]**

- Se Café → Manteiga Confiança= $3/3 = 1$
- Se Manteiga → Café Confiança= $3/5 = 0.6$

Produto	Suporte
Café e Pão	$3/10 = 0.3$
Café e Manteiga	$3/10 = 0.3$
Pão e Manteiga	$4/10 = 0.4$

Tabela 5. Itemset = 2

Produto	Suporte
Café, Pão e Manteiga	$3/10 = 0.3$

Tabela 6. Itemset = 3

- *[Pão, Manteiga]*
 - Se Pão \rightarrow Manteiga Confiança = $4/5 = 0.8$
 - Se Manteiga \rightarrow Pão Confiança = $4/5 = 0.8$
- *[Café, Pão e Manteiga]*
 - Se Café e Pão \rightarrow Manteiga Confiança = $3/3 = 1$
 - Se café e manteiga \rightarrow Pão Confiança = $3/3 = 1$
 - Se Pão e Manteiga \rightarrow Pão Confiança = $3/4 = 0,75$
 - Se Café \rightarrow Pão e Manteiga Confiança = $3/3 = 1$
 - Se Pão \rightarrow Café e Manteiga Confiança = $3/5 = 0.6$
 - Se Manteiga \rightarrow Café e Pão Confiança = $3/5 = 0.6$

Ao todo foram encontrados 12 tipos de regras em relação a quem levou café e manteiga temos a regra se café \rightarrow manteiga vai mostrar que todas as pessoas que levaram café também levou manteiga mas apenas 0.6 dos que levaram manteiga levaram café. Já em pão e manteiga ambos tiveram o mesmo resultado aqueles que levaram pão e manteiga e manteiga e pão tiveram o resultado de 0.8. Em relação aos três, nos temos que quem levou café e pão também levou manteiga, quem levou café e manteiga também levou pão, e se levou café também levou pão e manteiga, agora quem levou pão e manteiga também levou pão tiveram o resultado de 0,75. E quem levou pão também levou café e manteiga teve o resultado de 0.6, e quem levou manteiga e também levou café e pão também teve o resultado de 0,6.

3. Questão 03

Considere a seguinte tabela:

3.1. Crie um arquivo arff ou CSV com esta tabela

Vide anexo: questao3.csv.

3.2. Com o método APRIORI, com os demais parâmetros default, na ferramenta Weka, descubra 15 regras

Apriori =====

Minimum support: 0.55 (6 instances) Minimum metric {confidence}; 0.9 Number of cycles performed: 9

Generated sets of large itemsets:

Size of set of large itemsets L(1): 8

Size of set of large itemsets L(2): 13

Size of set of large itemsets L(3): 5

Best rules found:

1. Banana=Sim 7 ==> Shampoo=Sim 7 ;conf:(1)< lift:(1.22) lev:(0.12) [1] conv:(1.27)
2. Banana=Sim 7 ==> Protetor solar=Não 7 ;conf:(1)< lift:(1.22) lev:(0.12) [1] conv:(1.27)
3. Pasta de dente=Não 7 ==> Protetor solar=Não 7 ;conf:(1)< lift:(1.22) lev:(0.12) [1] conv:(1.27)
4. Creme para mãos=Não 7 ==> Protetor solar=Não 7 ;conf:(1)< lift:(1.22) lev:(0.12) [1] conv:(1.27)
5. Banana=Sim Protetor solar=Não 7 ==> Shampoo=Sim 7 ;conf:(1)< lift:(1.22) lev:(0.12) [1] conv:(1.27)
6. Banana=Sim Shampoo=Sim 7 ==> Protetor solar=Não 7 ;conf:(1)< lift:(1.22) lev:(0.12) [1] conv:(1.27)
7. Banana=Sim 7 ==> Shampoo=Sim Protetor solar=Não 7 ;conf:(1)< lift:(1.38) lev:(0.17) [1] conv:(1.91)
8. Ervilha=Sim 6 ==> Pasta de dente=Não 6 ;conf:(1)< lift:(1.57) lev:(0.2) [2] conv:(2.18)
9. Ervilha=Sim 6 ==> Protetor solar=Não 6 ;conf:(1)< lift:(1.22) lev:(0.1) [1] conv:(1.09)
10. Pasta de dente=Não Shampoo=Sim 6 ==> Protetor solar=Não 6 ;conf:(1)< lift:(1.22) lev:(0.1) [1] conv:(1.09)

3.3. Altere a confiança para 0.6 e veja o que ocorre. Registre os resultados dos experimentos.

Apriori =====

Minimum support: 0.65 (7 instances) Minimum metric <confidence>: 0.6 Number of cycles performed: 7

Generated sets of large itemsets:

Size of set of large itemsets L(1): 7

Size of set of large itemsets L(2): 5

Size of set of large itemsets L(3): 1

Best rules found:

1. Banana=Sim 7 ==> Shampoo=Sim 7 ;conf:(1)< lift:(1.22) lev:(0.12) [1] conv:(1.27)
2. Banana=Sim 7 ==> Protetor solar=Não 7 ;conf:(1)< lift:(1.22) lev:(0.12) [1] conv:(1.27)
3. Pasta de dente=Não 7 ==> Protetor solar=Não 7 ;conf:(1)< lift:(1.22) lev:(0.12) [1] conv:(1.27)
4. Creme para mãos=Não 7 ==> Protetor solar=Não 7 ;conf:(1)< lift:(1.22) lev:(0.12) [1] conv:(1.27)
5. Banana=Sim Protetor solar=Não 7 ==> Shampoo=Sim 7 ;conf:(1)< lift:(1.22) lev:(0.12) [1] conv:(1.27)
6. Banana=Sim Shampoo=Sim 7 ==> Protetor solar=Não 7 ;conf:(1)< lift:(1.22) lev:(0.12) [1] conv:(1.27)

7. Banana=Sim 7 ==ç Shampoo=Sim Protetor solar=Não 7 ;conf:(1)ç lift:(1.38) lev:(0.17) [1] conv:(1.91)
 8. Protetor solar=Não 9 ==ç Shampoo=Sim 8 ;conf:(0.89)ç lift:(1.09) lev:(0.06) [0] conv:(0.82)
 9. Shampoo=Sim 9 ==ç Protetor solar=Não 8 ;conf:(0.89)ç lift:(1.09) lev:(0.06) [0] conv:(0.82)
 10. Shampoo=Sim Protetor solar=Não 8 ==ç Banana=Sim 7 ;conf:(0.88)ç lift:(1.38) lev:(0.17) [1] conv:(1.45)

4. Questão 04

Abra o arquivo iris.arff e execute o algoritmo Apriori.

4.1. Porque ele não está habilitado?

Porque o algoritmo só funciona com valores nominais.

4.2. Discretize então os dados numéricos em 5 grupos (bins) (PreProcess/filters/unsupervised/attribute/discretize). Quantas regras foram encontradas? Analise algumas regras

=== Run information ===

Scheme: weka.associations.Apriori -N 10 -T 0 -C 0.6 -D 0.05 -U 1.0 -M 0.1 -S -1.0 -c -1 Relation: iris-weka.filters.unsupervised.attribute.Discretize-B10-M-1.0-Rfirst-last-precision6 Instances: 150 Attributes: 5 sepalwidth sepalwidth petalwidth petalwidth class === Associator model (full training set) ===

Apriori =====

Minimum support: 0.2 (30 instances) Minimum metric ;confidenceç: 0.6 Number of cycles performed: 16

Generated sets of large itemsets:

Size of set of large itemsets L(1): 6

Size of set of large itemsets L(2): 3

Size of set of large itemsets L(3): 1

Best rules found:

1. petalwidth='(-inf-0.34]' 41 ==ç class=Iris-setosa 41 ;conf:(1)ç lift:(3) lev:(0.18) [27] conv:(27.33)

2. petalwidth='(-inf-1.59]' 37 ==ç class=Iris-setosa 37 ;conf:(1)ç lift:(3) lev:(0.16) [24] conv:(24.67)

3. petalwidth='(-inf-1.59]' petalwidth='(-inf-0.34]' 33 ==ç class=Iris-setosa 33 ;conf:(1)ç lift:(3) lev:(0.15) [22] conv:(22)

4. petalwidth='(-inf-1.59]' 37 ==ç petalwidth='(-inf-0.34]' 33 ;conf:(0.89)ç lift:(3.26) lev:(0.15) [22] conv:(5.38)

5. petalwidth='(-inf-1.59]' class=Iris-setosa 37 ==ç petalwidth='(-inf-0.34]' 33 ;conf:(0.89)ç lift:(3.26) lev:(0.15) [22] conv:(5.38)

) 6. petalwidth='(-inf-1.59]' 37 ==ç petalwidth='(-inf-0.34]' class=Iris-setosa 33 ;conf:(0.89)ç

lift:(3.26) lev:(0.15) [22] conv:(5.38)
) 7. class=Iris-setosa 50 ==> petalwidth='(-inf-0.34]' 41 ;conf:(0.82); lift:(3) lev:(0.18)
 [27] conv:(3.63)
 8. petalwidth='(-inf-0.34]' 41 ==> petallength='(-inf-1.59]' 33 ;conf:(0.8); lift:(3.26)
 lev:(0.15) [22] conv:(3.43)
 9. petalwidth='(-inf-0.34]' class=Iris-setosa 41 ==> petallength='(-inf-1.59]' 33 ;conf:(0.8);
 lift:(3.26) lev:(0.15) [22] conv:(3.43)
 10. petalwidth='(-inf-0.34]' 41 ==> petallength='(-inf-1.59]' class=Iris-setosa 33 ;conf:(0.8);
 lift:(3.26) lev:(0.15) [22] conv:(3.43)

4.3. Refaça a discretização, agora considerando 3 grupos . Qual o resultado?

=== Run information ===

Scheme: weka.associations.Apriori -N 10 -T 0 -C 0.6 -D 0.05 -U 1.0 -M 0.1 -S
 -1.0 -c -1 Relation: iris-weka.filters.unsupervised.attribute.Discretize-B10-M-1.0-Rfirst-
 last-precision6 Instances: 150 Attributes: 5 sepalwidth sepalwidth petallength petalwidth
 class === Associator model (full training set) ===

Apriori =====

Minimum support: 0.2 (30 instances) Minimum metric ;confidence;: 0.6 Number
 of cycles performed: 16

Generated sets of large itemsets:

Size of set of large itemsets L(1): 6

Size of set of large itemsets L(2): 3

Size of set of large itemsets L(3): 1

Best rules found:

1. petalwidth='(-inf-0.34]' 41 ==> class=Iris-setosa 41 ;conf:(1); lift:(3) lev:(0.18)
 [27] conv:(27.33)
 2. petallength='(-inf-1.59]' 37 ==> class=Iris-setosa 37 ;conf:(1); lift:(3) lev:(0.16) [24]
 conv:(24.67)
 3. petallength='(-inf-1.59]' petalwidth='(-inf-0.34]' 33 ==> class=Iris-setosa 33 ;conf:(1);
 lift:(3) lev:(0.15) [22] conv:(22)
 4. petallength='(-inf-1.59]' 37 ==> petalwidth='(-inf-0.34]' 33 ;conf:(0.89); lift:(3.26)
 lev:(0.15) [22] conv:(5.38)
 5. petallength='(-inf-1.59]' class=Iris-setosa 37 ==> petalwidth='(-inf-0.34]' 33 ;conf:(0.89);
 lift:(3.26) lev:(0.15) [22] conv:(5.38)
) 6. petallength='(-inf-1.59]' 37 ==> petalwidth='(-inf-0.34]' class=Iris-setosa 33 ;conf:(0.89);
 lift:(3.26) lev:(0.15) [22] conv:(5.38)
) 7. class=Iris-setosa 50 ==> petalwidth='(-inf-0.34]' 41 ;conf:(0.82); lift:(3) lev:(0.18)
 [27] conv:(3.63)
 8. petalwidth='(-inf-0.34]' 41 ==> petallength='(-inf-1.59]' 33 ;conf:(0.8); lift:(3.26)
 lev:(0.15) [22] conv:(3.43)
 9. petalwidth='(-inf-0.34]' class=Iris-setosa 41 ==> petallength='(-inf-1.59]' 33 ;conf:(0.8);
 lift:(3.26) lev:(0.15) [22] conv:(3.43)

10. petalwidth='(-inf-0.34]' 41 ==> petallength='(-inf-1.59]' class=Iris-setosa 33 ;conf:(0.8);
lift:(3.26) lev:(0.15) [22] conv:(3.43)

5. Questão 05

Considerando a base de dados do WEKA: Supermaket.arf, pede-se:

5.1. Anote os resultados usando o algoritmo APRIORI

=== Run information ===

Scheme: weka.associations.Apriori -N 10 -T 0 -C 0.9 -D 0.05 -U 1.0 -M 0.1 -S -
1.0 -c -1 Relation: supermarket Instances: 4627 Attributes: 217 [list of attributes omitted]
=== Associator model (full training set) ===

Apriori =====

Minimum support: 0.15 (694 instances) Minimum metric ;confidence;: 0.9 Num-
ber of cycles performed: 17

Generated sets of large itemsets:

Size of set of large itemsets L(1): 44

Size of set of large itemsets L(2): 380

Size of set of large itemsets L(3): 910

Size of set of large itemsets L(4): 633

Size of set of large itemsets L(5): 105

Size of set of large itemsets L(6): 1

5.2. Analise as regras. Veja o que você acha sobre elas e o que você pode fazer para melhorá-las

Best rules found:

1. biscuits=t frozen foods=t fruit=t total=high 788 ==> bread and cake=t 723 ;conf:(0.92); lift:(1.27) lev:(0.03) [155] conv:(3.35)
2. baking needs=t biscuits=t fruit=t total=high 760 ==> bread and cake=t 696 ;conf:(0.92); lift:(1.27) lev:(0.03) [149] conv:(3.28)
3. baking needs=t frozen foods=t fruit=t total=high 770 ==> bread and cake=t 705 ;conf:(0.92); lift:(1.27) lev:(0.03) [150] conv:(3.27)
4. biscuits=t fruit=t vegetables=t total=high 815 ==> bread and cake=t 746 ;conf:(0.92); lift:(1.27) lev:(0.03) [159] conv:(3.26)
5. party snack foods=t fruit=t total=high 854 ==> bread and cake=t 779 ;conf:(0.91); lift:(1.27) lev:(0.04) [164] conv:(3.15)
6. biscuits=t frozen foods=t vegetables=t total=high 797 ==> bread and cake=t 725 ;conf:(0.91); lift:(1.26) lev:(0.03) [151] conv:(3.06)
7. baking needs=t biscuits=t vegetables=t total=high 772 ==> bread and cake=t 701 ;conf:(0.91); lift:(1.26) lev:(0.03) [145] conv:(3.01)
8. biscuits=t fruit=t total=high 954 ==> bread and cake=t 866 ;conf:(0.91); lift:(1.26) lev:(0.04) [179] conv:(3)

9. frozen foods=t fruit=t vegetables=t total=high 834 ==, bread and cake=t 757 ;conf:(0.91);
lift:(1.26) lev:(0.03) [156] conv:(3)
10. frozen foods=t fruit=t total=high 969 ==, bread and cake=t 877 ;conf:(0.91); lift:(1.26)
lev:(0.04) [179] conv:(2.92)