



Mineração de Dados

Ativ. 10 - Regras de Associação

Prof.: Alexandre M. A. Maciel

Estudantes:

Nicolau Calado Jofilsan

Renan Costa Alencar

Sthéfano Henrique Mendes Tavares Silva

João Victor Oliveira de Albuquerque



Atividade 10 – Regras de Associação

1. Escolha algum algoritmo de REGRAS DE ASSOCIAÇÃO apresentados na aula, aplique ao seu projeto.

- **Apriori**
- FP-Growth

Defina a metodologia experimental.

- Dados contínuos (não sequenciais);
- Gerados a partir de reações involuntárias (não faz sentido predizer);
- Dados gerados com base em micro expressões, valores próximos e pequenos.

Solução

- Matriz inversa generalizada de Moore-Penrose – Polarização dos valores.

Resultado da polarização

Cluster 1 – Tédio
5.432 linhas

	A	B	C	D	E	F	G	H	I	J
1		angry	disgusted	fear	sad	surprised	happy			
2	0	0	1	0	1	1	0			
3	1	0	1	0	1	1	1			
4	2	0	1	0	0	1	1			
5	3	0	1	0	1	1	0			
6	4	0	1	0	1	1	0			
7	5	0	1	0	1	1	0			
8	6	0	1	0	1	1	0			
9	7	0	1	0	1	1	0			
10	8	0	1	0	1	1	1			
11	9	0	1	0	1	1	0			
12	10	0	1	0	1	1	0			
13	11	0	1	0	1	1	0			
14	12	0	1	0	1	1	0			
15	13	0	1	0	1	1	0			
16	14	0	1	0	1	1	1			
17	15	0	1	0	1	1	0			
18	16	0	1	0	1	1	0			
19	17	0	1	0	1	1	0			
20	18	0	1	0	1	1	0			
21	19	0	1	0	1	1	0			
22	20	0	1	0	1	1	0			
23	21	0	1	0	1	1	0			
24	22	0	1	0	1	1	0			

Atividade 10 – Regras de Associação

Definindo um valor limite para o valor de suporte e calculando o valor de suporte.



```
df_new = apriori(data, min_support = 0.2, use_colnames = True)  
df_new
```



	support	itemsets
0	0.607327	(angry)
1	0.493741	(disgusted)
2	0.545655	(fear)
3	0.540317	(sad)
4	0.444772	(surprised)
5	0.497054	(happy)
6	0.347938	(angry, fear)
7	0.288660	(angry, sad)
8	0.207658	(angry, surprised)
9	0.385677	(happy, angry)
10	0.242636	(fear, disgusted)
11	0.272091	(disgusted, sad)
12	0.292526	(surprised, disgusted)
13	0.231959	(fear, sad)
14	0.399300	(happy, fear)
15	0.249080	(surprised, sad)
16	0.301546	(happy, angry, fear)

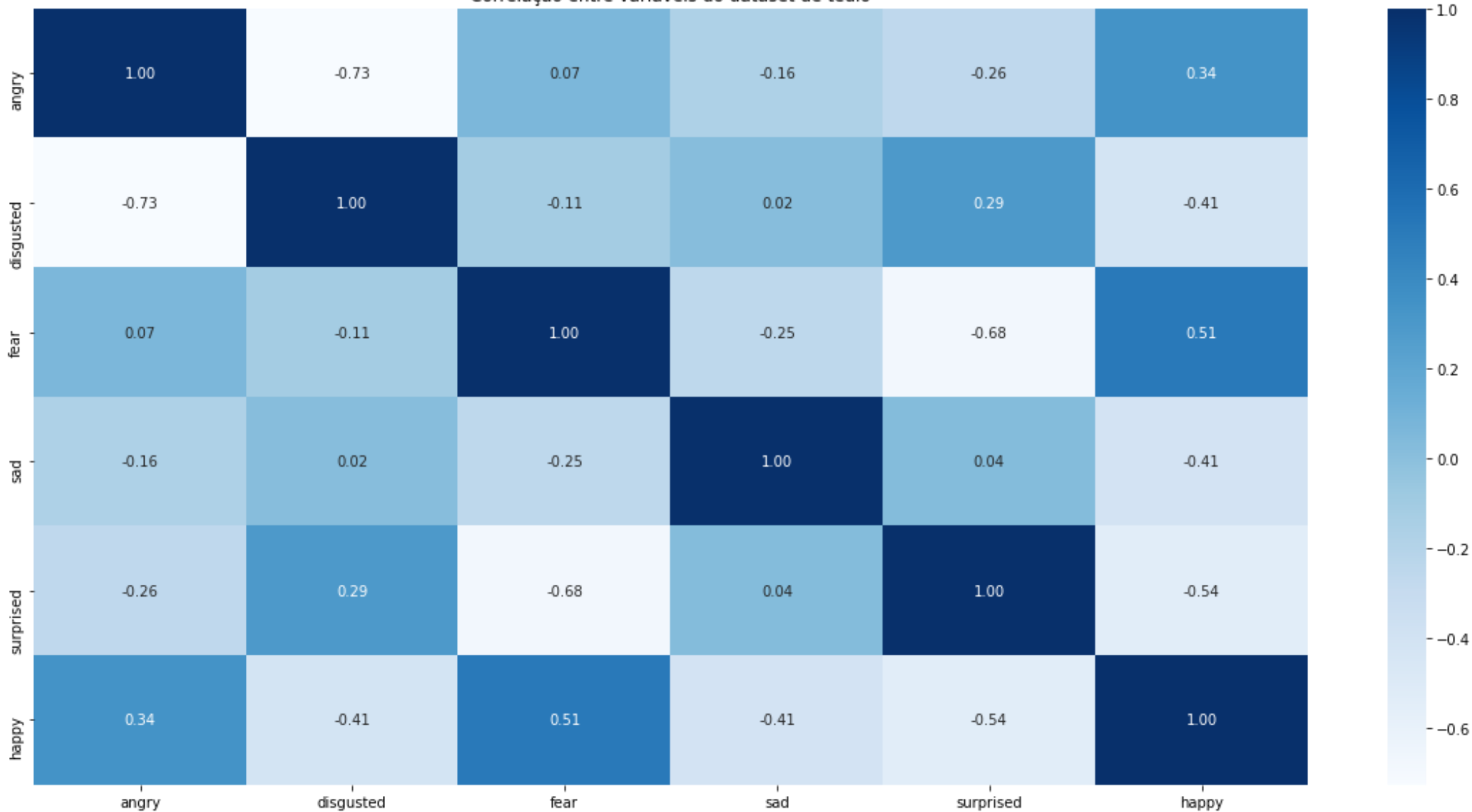
Atividade 10 – Regras de Associação



```
df_ar = association_rules(df_new, metric = "confidence", min_threshold = 0.6)
df_ar
```

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction
0	(fear)	(angry)	0.545655	0.607327	0.347938	0.637652	1.049932	0.016547	1.083690
1	(happy)	(angry)	0.497054	0.607327	0.385677	0.775926	1.277608	0.083803	1.752425
2	(angry)	(happy)	0.607327	0.497054	0.385677	0.635041	1.277608	0.083803	1.378087
3	(surprised)	(disgusted)	0.444772	0.493741	0.292526	0.657699	1.332073	0.072924	1.478987
4	(happy)	(fear)	0.497054	0.545655	0.399300	0.803333	1.472236	0.128080	2.310227
5	(fear)	(happy)	0.545655	0.497054	0.399300	0.731781	1.472236	0.128080	1.875133
6	(angry, happy)	(fear)	0.385677	0.545655	0.301546	0.781862	1.432885	0.091099	2.082827
7	(happy, fear)	(angry)	0.399300	0.607327	0.301546	0.755187	1.243460	0.059040	1.603970
8	(angry, fear)	(happy)	0.347938	0.497054	0.301546	0.866667	1.743605	0.128602	3.772091
9	(happy)	(angry, fear)	0.497054	0.347938	0.301546	0.606667	1.743605	0.128602	1.657784

Correlação entre variáveis do dataset de tédio



Obrigado

Fim do Documento