

MOTTA, Claudia Lage Rebello; GARCIA, Ana Cristina Bicharra; VIVACQUA, Adriana Santarosa; SANTORO, Flávia Maria; SAMPAIO, Jonice de Oliveira. Sistemas de recomendação. In: PIMENTEL, M.; FUKS, H. (Org.). **Sistemas colaborativos**. Rio de Janeiro: SBC/Elsevier, 2011. cap. 15, p. 230-244.

Exercícios comentados

15.1. Considere que a locadora “Alvo” possui um conjunto de filmes para alugar e que tais filmes recebem avaliações de seus clientes. Atualmente consta com o seguinte conjunto de filmes e avaliações apresentados na Tabela 15.A.1. José é um excelente cliente da locadora Alvo. Ele vai ficar em casa sábado e deseja uma sugestão de filme para alugar. Ele vai à locadora e observa que acaba de chegar o filme “Comer, Rezar e Amar”. Utilizando a técnica de filtragem colaborativa, usando a correlação de Pearson, verifique se ele vai gostar do filme. Considere verdadeira a afirmativa de que “ele vai gostar” se a expectativa de avaliação dele for superior a 3. Justifique sua resposta mostrando o passo a passo do método. No seu calculo considere apenas os 3 clientes mais similares a José.

Tabela 15.A.1. Banco de dados de avaliações de filmes da Locadora Alvo.

	Maria	José	João	Pedro	Laura	Manoel	Jaime	Rafael	Lucas	Reinaldo	Felipe	Eva
<i>Gladiator (2000)</i>	4	5	2	3	3	4	4	3	3	4	3	4
<i>Braveheart (1995)</i>	4	4	2	4	3	4	5	2	3	4	3	4
<i>Hannibal (2001)</i>	3	5	2	3	3	5	5	2	3	3	3	3
<i>Black Hawk Down (2001)</i>	4	4	3	3	3	5	4	2	3	4	3	3
<i>The Matrix (1999)</i>	4	4	3	4	3	4	4	2	3	4	3	4
<i>Pulp Fiction (1994)</i>	3	4	3	3	4	5	4	3	3	4	4	4
<i>Training Day (2001)</i>	4	4	3	3	4	4	4	3	3	4	4	3
<i>GangsOf New York (2002)</i>	4	4	3	4	4	5	5	2	3	4	4	3
<i>Pearl Harbor (2001)</i>	4	5	3	4	4	4	5	2	3	3	4	3
<i>Spawn (1997)</i>	3	4	2	3	3	4	4	3	2	3	3	3
<i>TheOriginalKings Comedy</i>	2	3	4	2	4	2	2	5	5	3	3	4
<i>Ice Age (2002)</i>	3	2	5	3	3	3	3	5	5	2	4	3
<i>Finding Nemo (2003)</i>	3	3	4	2	3	2	3	4	5	2	4	4
<i>Zoolander (2001)</i>	3	3	4	2	4	3	2	5	4	3	3	4
<i>Kangaroo Jack (2003)</i>	2	2	4	3	4	3	2	4	5	3	4	4
<i>Bruce Almighty (2003)</i>	2	2	4	3	3	3	2	4	4	3	4	4
<i>Scary Movie (2000)</i>	2	3	4	2	4	3	3	4	5	2	4	3
<i>Bad Boys II (2003)</i>	2	2	5	3	4	3	2	4	4	2	4	4
<i>Men in Black (1997)</i>	3	3	5	2	4	2	2	5	4	3	4	4

<i>Nixon (1995)</i>	5	4	3	3	2	3	4	3	2	3	1	1
<i>The Pianist (2002)</i>	4	3	2	4	2	3	4	2	3	4	2	1
<i>Hamlet (1996)</i>	5	4	3	3	2	3	3	2	3	4	2	1
<i>Seven (1995)</i>	4	3	2	3	1	3	4	2	2	4	1	1
<i>Malena (2000)</i>	5	4	3	3	1	4	3	3	3	3	2	2
<i>Read Review</i>	5	3	2	4	1	4	4	3	3	3	2	1
<i>SurvivingPicasso</i>	5	4	2	3	2	4	3	2	2	4	2	2
<i>A Few Good Men (1992)</i>	4	3	3	3	1	3	4	2	3	4	2	2
<i>TheMothman Prophecies</i>	5	4	3	4	2	4	4	3	2	3	1	2
<i>Goodbye Lenin! (2004)</i>	5	3	2	3	1	4	4	2	3	4	2	1
<i>Philadelphia (1993)</i>	4	4	2	3	1	3	4	3	2	4	1	2
<i>Amelie (2001)</i>		2					2					
<i>Malena (2000)</i>		2					1					
<i>read review</i>		2					2					
<i>Goodbye Lenin! (2004)</i>		2					2					
<i>Vidocq</i>		1					1					
<i>28 Days Later (2003)</i>	3	3	1	4	2	4	3	1	1	5	2	1
<i>Resident Evil (2002)</i>	3	4	1	4	1	4	3	1	2	4	2	1
<i>Sleepy Hollow (1999)</i>	3	4	1	4	1	3	4	2	2	5	1	1
<i>The Ring (2002)</i>	4	3	2	5	1	3	4	1	1	4	1	1
<i>Casablanca</i>	1	3	3	1	4	2	2	4	3	1	4	5
<i>Gone With the Wind</i>	2	2	3	2	4	3	3	4	3	2	5	4
<i>West Side Story</i>	1	3	3	2	4	3	3	4	3	1	5	5
<i>Roman Holiday</i>	2	2	3	1	5	2	3	4	4	1	5	4
<i>An Affair to Remember</i>	1	3	4	2	4	3	2	3	3	1	5	4
<i>The Way We Were</i>	1	2	3	2	5	2	2	3	4	2	4	5
<i>Doctor Zhivago</i>	2	2	3	2	5	2	3	3	3	1	5	5
<i>It's a Wonderful Life</i>	1	2	3	1	4	3	2	4	3	1	5	4
<i>Love Story</i>	2	3	4	2	5	2	3	3	4	1	4	4
<i>City Lights.</i>	1	3	4	1	4	2	2	4	3	1	5	5
MÉDIA DAS AVALIAÇÕES	3,11	3,12	2,95	2,84	3,00	3,27	3,12	3,00	3,11	2,95	3,16	3,02

Usando a correlação de Pearson, chegamos ao conjunto de correlações entre usuários apresentadas na Tabela 15.A.2.

Tabela 15.A.2. Correlações entre as avaliações dos clientes da locadora Alvo.

	Maria	José	João	Pedro	Laura	Manoel	Jaime	Rafael	Lucas	Reinaldo	Felipe	Eva
--	-------	------	------	-------	-------	--------	-------	--------	-------	----------	--------	-----

Maria	1,00	0,59	-0,40	0,68	-0,70	0,55	0,67	-0,52	-0,41	0,72	-0,75	-0,73
José		1,00	-0,46	0,47	-0,35	0,64	0,76	-0,51	-0,44	0,55	-0,45	-0,34
João			1,00	-0,51	0,60	-0,45	-0,56	0,77	0,77	-0,58	0,58	0,59
Pedro				1,00	-0,63	0,62	0,66	-0,70	-0,49	0,78	-0,66	-0,67
Laura					1,00	-0,34	-0,41	0,56	0,57	-0,66	0,88	0,87
Manoel						1,00	0,66	-0,54	-0,40	0,59	-0,29	-0,36
Jaime							1,00	-0,64	-0,45	0,56	-0,45	-0,47
Rafael								1,00	0,72	-0,63	0,56	0,63
Lucas									1,00	-0,46	0,58	0,56
Reinaldo										1,00	-0,73	-0,68
Felipe											1,00	0,85
Eva												1,00

A Tabela a seguir apresenta alguns cálculos intermediários para que entendamos o cálculo da previsão da avaliação do José tendo em vista a avaliação dos seus pares em avaliação de filmes.

Inicialmente devemos calcular a matriz de distância entre os clientes. Usando a correlação de Pearson, chegamos a tabela 15.3:

Tabela 15.A.3. Distância entre os usuários da locadora Alvo.

	Maria	José	João	Pedro	Laura	Manoel	Jaime	Rafael	Lucas	Reinaldo	Felipe	Eva
Maria	0,00	0,41	1,40	0,32	1,70	0,45	0,33	1,52	1,41	0,28	1,75	1,73
José	0,41	0,00	1,46	0,53	1,35	0,36	0,24	1,51	1,44	0,45	1,45	1,34
João	1,40	1,46	0,00	1,51	0,40	1,45	1,56	0,23	0,23	1,58	0,42	0,41
Pedro	0,32	0,53	1,51	0,00	1,63	0,38	0,34	1,70	1,49	0,22	1,66	1,67
Laura	1,70	1,35	0,40	1,63	0,00	1,34	1,41	0,44	0,43	1,66	0,12	0,13
Manoel	0,45	0,36	1,45	0,38	1,34	0,00	0,34	1,54	1,40	0,41	1,29	1,36
Jaime	0,33	0,24	1,56	0,34	1,41	0,34	0,00	1,64	1,45	0,44	1,45	1,47
Rafael	1,52	1,51	0,23	1,70	0,44	1,54	1,64	0,00	0,28	1,63	0,44	0,37
Lucas	1,41	1,44	0,23	1,49	0,43	1,40	1,45	0,28	0,00	1,46	0,42	0,44
Reinaldo	0,28	0,45	1,58	0,22	1,66	0,41	0,44	1,63	1,46	0,00	1,73	1,68
Felipe	1,75	1,45	0,42	1,66	0,12	1,29	1,45	0,44	0,42	1,73	0,00	0,15
Eva	1,73	1,34	0,41	1,67	0,13	1,36	1,47	0,37	0,44	1,68	0,15	0,00

Com isso, chegamos a conclusão que Jaime, Manoel e Maria são os 3 clientes com avaliações mais similares a Jose.

Para calcularmos a previsão da avaliação de José para o filme “Comer, Rezar, Amar”, devemos levar em conta as avaliações dadas ao filme por esses 3 clientes considerados mais similares a Jose.

Logo, podemos esperar que a avaliação do filme “Comer, Rezar, Amar” pelo José seja calculada como: $3,1 + (-0,497/1,99)$, que é igual a 2,85. Logo, podemos considerar que ele NAO vai gostar do filme. Interessante notar que se considerássemos apenas a media das avaliações de José, criaríamos uma expectativa diferente: a de que ele iria gostar do filme.

15.2. Usamos o mesmo exemplo dos filmes da locadora Alvo para explicar o método de filtragem por conteúdo. Para tanto, o conjunto de filmes teve que ser organizado segundo a seguinte categoria: Ação, Drama, Comédia, Romance e Terror, apresentado na Tabela 15.A.4. Qual a expectativa de avaliação de José para o filme “Comer, Rezar, Amar”?

Tabela 15.A.4. Filmes da locadora Alvo, avaliados e organizados por categorias.

Ação	Maria	José	João	Pedro	Laura	Manoel	Jaime	Rafael	Luca	Reinaldo	Felipe	Eva
Gladiator (2000)	4	5	2	3	3	4	4	3	3	4	3	4
Braveheart (1995)	4	4	2	4	3	4	5	2	3	4	3	4
Hannibal (2001)	3	5	2	3	3	5	5	2	3	3	3	3
Black Hawk Down (2001)	4	4	3	3	3	5	4	2	3	4	3	3
The Matrix (1999)	4	4	3	4	3	4	4	2	3	4	3	4
Pulp Fiction (1994)	3	4	3	3	4	5	4	3	3	4	4	4
Training Day (2001)	4	4	3	3	4	4	4	3	3	4	4	3
Gangs of New York (2002)	4	4	3	4	4	5	5	2	3	4	4	3
Pearl Harbor (2001)	4	5	3	4	4	4	5	2	3	3	4	3

Spawn (1997)	3	4	2	3	3	4	4	3	2	3	3	3
Media	3,70	4,30	2,60	3,40	3,40	4,40	4,40	2,40	2,90	3,70	3,40	3,40

Comédia	Maria	José	João	Pedro	Laura	Manoel	Jaimé	Rafael	Luca	Reinaldo	Felipe	Eva
The Original Kings Comedy (2000)	2	3	4	2	4	2	2	5	5	3	3	4
Ice Age (2002)	3	2	5	3	3	3	3	5	5	2	4	3
Finding Nemo (2003)	3	3	4	2	3	2	3	4	5	2	4	4
Zoolander (2001)	3	3	4	2	4	3	2	5	4	3	3	4
Kangaroo Jack (2003)	2	2	4	3	4	3	2	4	5	3	4	4
Bruce Almighty (2003)	2	2	4	3	3	3	2	4	4	3	4	4
Scary Movie (2000)	2	3	4	2	4	3	3	4	5	2	4	3
Bad Boys II (2003)	2	2	5	3	4	3	2	4	4	2	4	4
Men in Black (1997)	3	3	5	2	4	2	2	5	4	3	4	4
Media	2,44	2,56	4,33	2,44	3,67	2,67	2,33	4,44	4,56	2,56	3,78	3,78

Drama	Maria	José	João	Pedro	Laura	Manoel	Jaimé	Rafael	Luca	Reinaldo	Felipe	Eva
Nixon (1995)	5	4	3	3	2	3	4	3	2	3	1	1
The Pianist (2002)	4	3	2	4	2	3	4	2	3	4	2	1
Hamlet (1996)	5	4	3	3	2	3	3	2	3	4	2	1
Seven (1995)	4	3	2	3	1	3	4	2	2	4	1	1

Malena (2000)	5	4	3	3	1	4	3	3	3	3	2	2
read review	5	3	2	4	1	4	4	3	3	3	2	1
Surviving Picasso (1996)	5	4	2	3	2	4	3	2	2	4	2	2
A Few Good Men (1992)	4	3	3	3	1	3	4	2	3	4	2	2
The Mothman Prophecies (2002)	5	4	3	4	2	4	4	3	2	3	1	2
Goodbye Lenin! (2004)	5	3	2	3	1	4	4	2	3	4	2	1
Philadelphia (1993)	4	4	2	3	1	3	4	3	2	4	1	2
	4,64	3,55	2,45	3,27	1,45	3,45	3,73	2,45	2,55	3,64	1,64	1,45

Arte	Maria	José	João	Pedro	Laura	Manoel	Jaimé	Rafael	Luca	Reinaldo	Felipe	Eva
Amelie (2001)		2					2					
Malena (2000)		2					1					
Read review		2					2					
Goodbye Lenin! (2004)		2					2					
Vidocq		1					1					
		1,80					1,60					

Horror	Maria	José	João	Pedro	Laura	Manoel	Jaime	Rafael	Lucas	Reinaldo	Felipe	Eva
28 Days Later (2003)	3	3	1	4	2	4	3	1	1	5	2	1
Resident Evil (2002)	3	4	1	4	1	4	3	1	2	4	2	1
Sleepy Hollow (1999)	3	4	1	4	1	3	4	2	2	5	1	1
The Ring	4	3	2	5	1	3	4	1	1	4	1	1

(2002)												
	3,25	3,50	1,25	4,25	1,25	3,50	3,50	1,25	1,50	4,50	1,50	1,00

Romance	Maria	José	João	Pedro	Laura	Manoel	Jaimé	Rafael	Luca	Reinaldo	Felipe	Eva
Casablanca	1	3	3	1	4	2	2	4	3	1	4	5
Gone With the Wind	2	2	3	2	4	3	3	4	3	2	5	4
West Side Story	1	3	3	2	4	3	3	4	3	1	5	5
Roman Holiday	2	2	3	1	5	2	3	4	4	1	5	4
An Affair to Remember	1	3	4	2	4	3	2	3	3	1	5	4
The Way We Were	1	2	3	2	5	2	2	3	4	2	4	5
Doctor Zhivago	2	2	3	2	5	2	3	3	3	1	5	5
It's a Wonderful Life	1	2	3	1	4	3	2	4	3	1	5	4
Love Story	2	3	4	2	5	2	3	3	4	1	4	4
City Lights.	1	3	4	1	4	2	2	4	3	1	5	5
	1,40	2,50	3,30	1,60	4,40	2,40	2,50	3,60	3,30	1,20	4,70	4,50

Calculando a avaliação média por categoria, por usuário, chegamos aos resultados apresentados na Tabela 15.4.

Tabela 15.4. Média de avaliações por categoria, por usuário, da locadora Alvo.

Usuário	Comédia	Romance	Ação	Arte	Drama	Horror
Maria	2,44	1,4	3,7		4,64	3,25
José	2,56	2,5	4,3	1,8	3,55	3,5
João	4,33	3,3	2,6		2,45	1,25
Pedro	2,44	1,6	3,4		3,27	4,25

Laura	3,67	4,4	3,4		1,45	1,25
Manoel	2,67	2,4	4,4		3,45	3,5
Jaime	2,33	2,5	4,4	1,6	3,73	3,5
Rafael	4,44	3,6	2,4		2,45	1,25
Lucas	4,56	3,3	2,9		2,55	1,5
Reinaldo	2,56	1,2	3,7		3,64	4,5
Felipe	3,78	4,7	3,4		1,64	1,5
Eva	3,78	4,5	3,4		1,45	1

O próximo passo será calcularmos a média das avaliações por categoria, como apresentada na Tabela 15.5.

Tabela 15.5. Média de avaliações por categoria.

Comédia	3,30
Romance	2,95
Ação	3,50
Arte	1,70
Drama	2,86
Horror	2,52

Com isso, podemos calcular a expectativa para a avaliação de Jose aos novos filmes que entraram na locadora, inclusive o “Comer, Rezar e Amar”, como apresentado na Tabela 15.6.

Tabela 15.6. Expectativa de avaliação de filmes novos que entraram na locadora Alvo, pelo José.

Novos Filmes	Classe	Avaliação Média da Categoria	Avaliação média do Filme	Aval média do José por Categoria	Avaliação José para Filme
Comer, Rezar, Amar	comédia	3,30	3,22	2,56	2,50
Love Story	romance	2,95	3,11	2,5	2,64
Hannibal (2001)	ação	3,50	2,78	4,3	3,42
Gangs of New York (2002)	ação	1,70	2,78	1,8	2,94
The Pianist	drama	2,86	2,78	3,55	3,46

(2002)						
Sleepy Hollow (1999)	horror	2,52	2,44	3,5	3,39	

Como mostrado na tabela, também chegaríamos à conclusão que Jose NÃO vai gostar do filme.