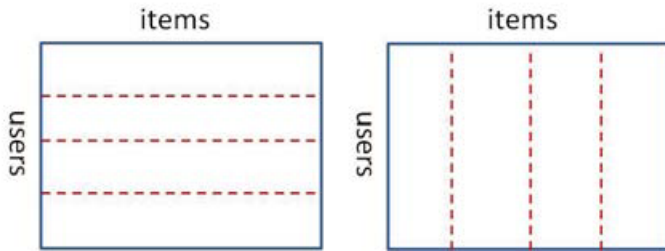


## Recommendation



Recommendation Item-based



## Item based recommendation

Weighted One Slope

Gegevens

*De gebruikte dataset is gelijk aan die van User-Item*

Gegevens

Data set

Artikelen : 101 t/m 106

Gebruikers : 1 t/m 7

Waardering : 1.0 t/m 5.0

	101	102	103	104	105	106
1	2.5	3.5	3.0	3.5	2.5	3.0
2	3.0	3.5	1.5	5.0	3.5	3.0
3	2.5	3.0	-	3.5	-	4.0
4	-	3.5	3.0	4.0	2.5	4.5
5	3.0	4.0	2.0	3.0	2.0	3.0
6	3.0	4.0	-	5.0	3.5	3.0
7	-	4.5	-	4.0	1.0	-

[dataset](#) [java vul code](#)

Een stukje van de data verzameling :

*userid, itemid, rating*

```

1,101,2.5
1,102,3.5
1,103,3.0
1,104,3.5
1,105,2.5
1,106,3.0
2,101,3.0
2,102,3.5
2,103,1.5
2,104,5.0
2,105,3.5
2,106,3.0
3,101,2.5
3,102,3.0

```

Inlezen

```

0:: entrykey=1  userid=1
    (101,2.5) (102,3.5) (103,3.0) (104,3.5) (105,2.5) (106,3.0)
1:: entrykey=2  userid=2
    (101,3.0) (102,3.5) (103,1.5) (104,5.0) (105,3.5) (106,3.0)
2:: entrykey=3  userid=3
    (101,2.5) (102,3.0) (104,3.5) (106,4.0)
3:: entrykey=4  userid=4
    (102,3.5) (103,3.0) (104,4.0) (105,2.5) (106,4.5)
4:: entrykey=5  userid=5
    (101,3.0) (102,4.0) (103,2.0) (104,3.0) (105,2.0) (106,3.0)
5:: entrykey=6  userid=6
    (101,3.0) (102,4.0) (104,5.0) (105,3.5) (106,3.0)
6:: entrykey=7  userid=7
    (102,4.5) (104,4.0) (105,1.0)

```

Item - Item : rating tabel

Algoritme : [slope-one](#)

trek de rating (voorkeur) van elkaar af (waarde= ItemRatingI - ItemRatingJ) en  
aan het eind reken het gemiddelde uit

```

for every item i
  for every other item j
    for every user u expressing preference for both i and j
      add the difference in u's preference for i and j to an average

```

[Euclidean distance](#) [Bron] Oreilly Programming Collective Intelligence 2007

blz 24  $d(\text{oreilly}) = \text{som}(\text{rating1} - \text{rating2})^2 : \text{recommendation\_value} = 1/(1+d)$

[Bron] [wiki euclidean](#)  $d(\text{wiki}) = \text{Wortel}(\text{som}(\text{rating1} - \text{rating2})^2)$

tbv  $\text{recommendation\_value} = 1/(1+d)$

[Bron] [wiki One slope](#)

[Bron] [Mahout Item](#)

[Bron] [Mahout algoritme](#)

For every item i the user u expresses no preference for  
 for every item j that user u expresses a preference for  
   find the average preference difference between j and i  
   add this diff to u's preference value for j  
   add this to a running average  
 return the top items, ranked by these averages

## One Slope

OPBOUW:  $\text{userid} \text{ itemid1 itemid2} : \text{verschil\_rating} = \text{rating2} - \text{rating1}$

101 102	1) 1 = 3.5 - 2.5	101 103	1) 0.5 = 3.0 - 2.5
	2) 0.5 = 3.5 - 3.0		2) -1.5 = 1.5 - 3.0
	3) 0.5 = 3.0 - 2.5		4) -0.5 = 3.0 - 3.5
	5) 1 = 4.0 - 3.0		5) -1.0 = 2.0 - 3.0
	6) 1 = 4.0 - 3.0		
-----	gemid_verschil = 4/5 = 0.8		gemid_verschil = -2.5 / 4 = -0.625
	Eucli_wiki = 1.87029		Eucli_wiki =

[excel file met berekeningen](#)

One slope uitleg

[Uitleg One Slope](#)

Euclidean & Difference Similarity

[Euclidean Similarity tabel](#)

	101	102	103	104	105	106
101	x	0,348331477	0,348331477	0,240253073	0,449489743	0,387425887
102		x	0,255396793	0,309016994	0,188637865	0,320377241
103			x	0,207991597	0,320377241	0,298935084
104				x	0,191825366	0,252650309
105					x	0,294298055

[Gemiddelde verschil \(One slope\) tabel](#)

	101	102	103	104	105	106
101	-	4.0/5=0.8	-2.0/3=-0.6666666666	6.0/5=1.2	0.0/4=0.0	2.0/5=0.4
102	-	-	-5.0/4=-1.25	2.0/7=0.285714285714285	-8.0/6=-1.3333333333333333	-1.0/6=-0.1666666666666666
103	-	-	-	6.0/4=1.5	1.0/4=0.25	4.0/4=1.0
104	-	-	-	-	-9.5/6=-1.5833333333333333	3.5/6=-0.5833333333333334

105	-	-	-	-	-	2.5/5=0.5
-----	---	---	---	---	---	-----------

Aan te bevelen artikelen

Bijvoorbeeld user=7

Alle items waar hij/zij **geen** voorkeur voor heeft gegevens is: 101, 103, 106

Alle items waar hij/zij **WEL** voorkeur voor heeft gegevens is: 102, 104, 105

Euclidean (uitwerking)

**Methodiek:** [Euclidean \(wiki\)](#)

bekend bij user 7		ON bekend bij user 7					
		factor	product	factor	product	factor	product
artikelen	rating	101	101	103	103	106	106
102	4,5	0,348331477	1,567491647	0,255396793	1,149285569	0,320377241	1,441697585
104	4	0,240253073	0,961012292	0,207991597	0,831966388	0,25260309	1,01041236
105	1	0,449489743	0,449489743	0,30377241	0,30377241	0,294298055	0,294298055
som factor		1,038074293		0,7671608		0,867278386	
som product			2,977993682		2,285024367		2,746408
som_product/som_factor			2,868767391		2,978546827		3,166697157

Aanbevolen (van hoog naar laag) : 106, 103, 102

Difference (uitwerking)

**Methodiek:** [Difference](#)

bekend bij user 7		ON bekend bij user 7					
		factor	som	factor	som	factor	som
artikelen	rating	101	101	103	103	106	106
102	4.5	-0.8	3.7	-1.25	3.25	-0.16666666666666666	4.333333333333333
104	4.0	-1.2	2.8	-1.5	2.5	-0.5833333333333334	3.4166666666666665
105	1.0	-0.0	1.0	-0.25	0.75	+0.5	1.5
gemiddelde			7,5/3=2,5		6,5/3=2,167		9,25/3=3,083

Aanbevolen (van hoog naar laag) : 106, 103 , 102

### One Slope Difference TIP

Voor One Slope heb je naast het gemiddelde verschil ook het aantal nodig om een weging te kunnen berekenen

bijv : 101 - 102 : verschil = -0,8 en aantal = 5

Weighted Difference (uitwerking)

**Methodiek:** [Weighted Difference](#)

bekend bij user 7		ON bekend bij user 7					
		101		103		106	
		(rating+verschil)* aantal	totaal	(rating+verschil)* aantal	totaal	(rating+verschil)* aantal	totaal
102	4.5	(4.5+-0.8)*5	18.5	(4.5+-1.25)*4	13.0	(4.5+-0.16666666666666666)*6	26.0
104	4.0	(4.0+-1.2)*5	14.0	(4.0+-1.5)*4	10.0	(4.0+-0.5833333333333334)*6	20.5
105	1.0	(1.0+-0.0)*4	4.0	(1.0+-0.25)*4	3.0	(1.0+0.5)*5	7.5
som_totaal/som_aantal			36,5/14=2,607		26/12=2,167		54/17=3,176

Aanbevolen (van hoog naar laag) : 106, 101 , 103