Recommender Systems

Daro VAN
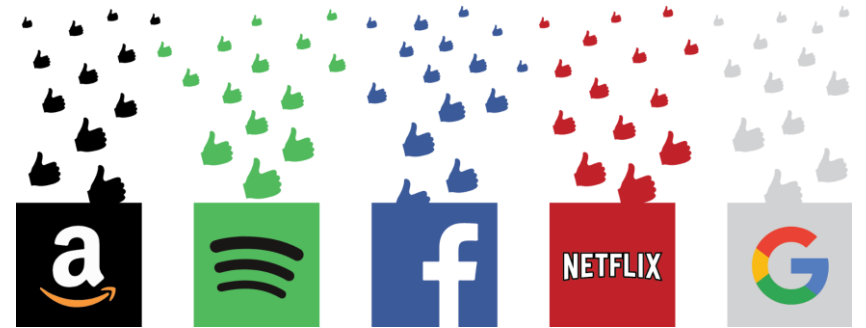
What are recommender systems?

Recommender systems captures the pattern of people's behavior and use it to predict what else they might want or like.

Applications

- What to buy?
  - E-commerce, book, movies…
- What to watch?
- Which job to apply to?
- Who you should be friends with?
- Personalize your experience on the web
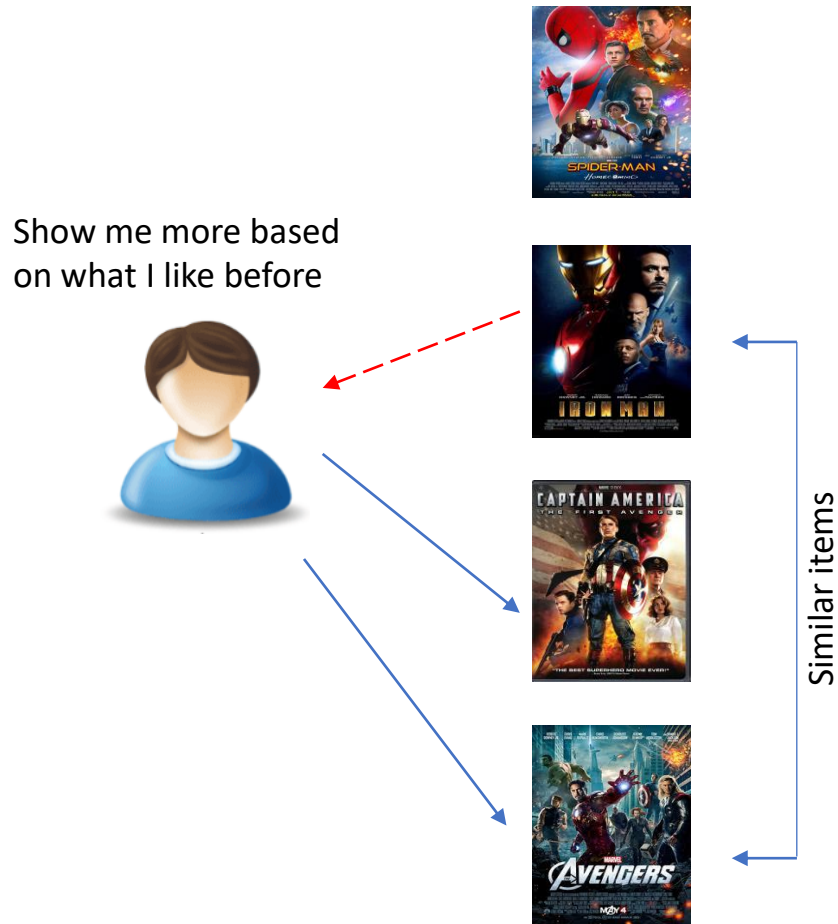 - News platforms, new personalization

Advantages of recommender systems

- Broader exposure
- Possibility of continual usage or purchase of products
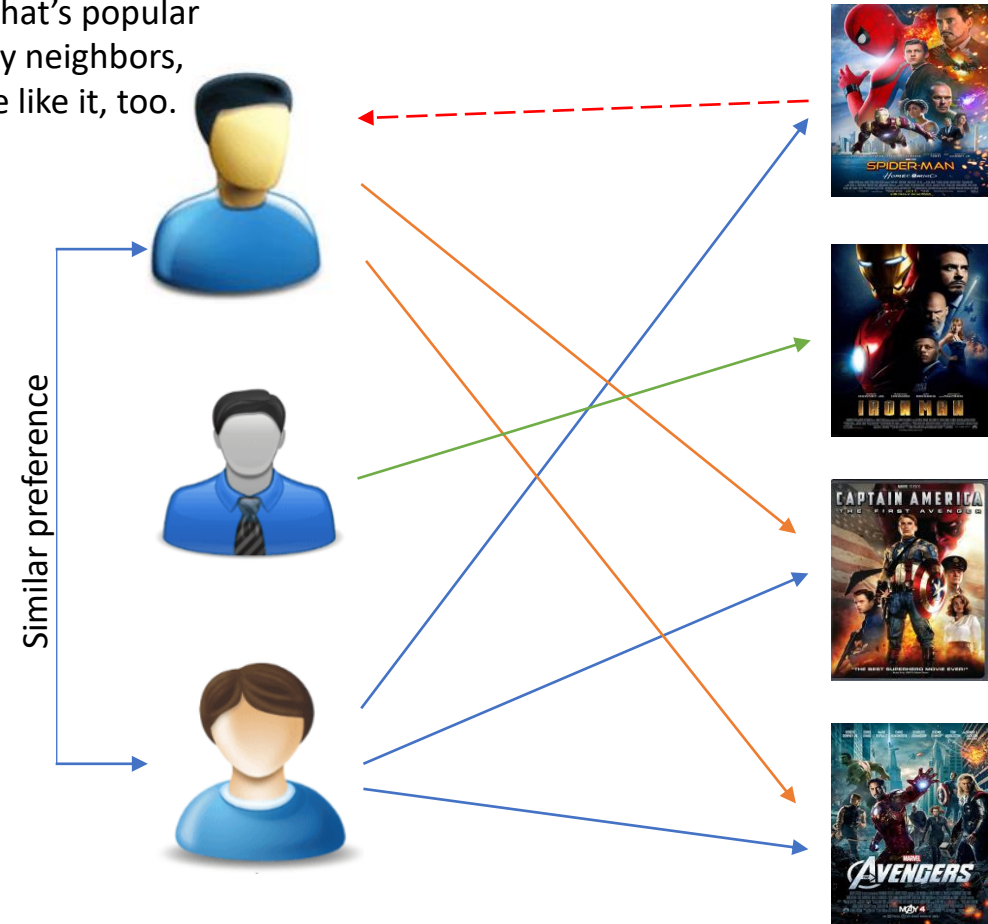- Provides better experiences

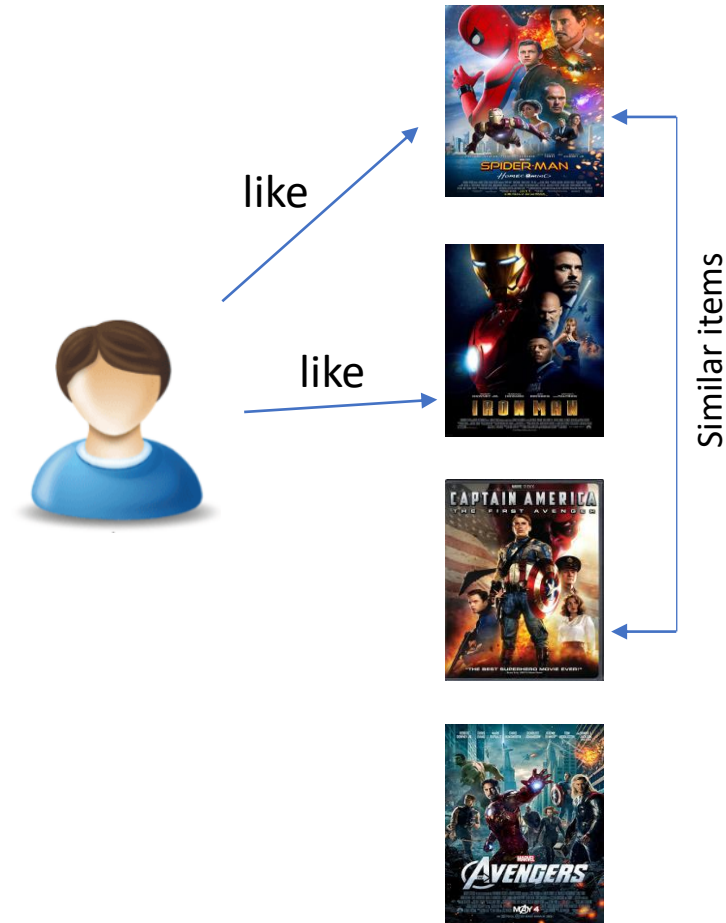# Two types of recommender systems

Content based

Show me more based on what I like before

Similar items

Collaborative filtering

Tell me what's popular among my neighbors, I might be like it, too.

Similar preference

Implementing recommender systems

- ## Memory-based
  - Uses the entire user item dataset to generate a recommendation
  - Uses statistical techniques to approximate users or items
    e.g, Pearson correlation, Euclidean distance, etc

- ## Model-based
  - Develops a model of users in an attempt to learn their preferences
  - Models can be created using machine learning technique like regression, classification, clustering, etc.

Content-based recommendation systems

like

like

Similar items

Content-based recommendation systems

Spider man   (Adventure, super hero)

2

10   Iron man   (Comedy, Adventure, super hero, Sci-Fi)

8   Captain American   (Comedy, super hero)

Hitchhiker's guide to the galaxy   (Comedy, super hero, Sci-Fi)

Avenger   (Super hero)

The dark knight rises   (comedy, Super hero)

Weighing the genres

		Comedy	adventure	super hero	Sci-Fi
	2	0	1	1	0
	10	1	1	1	1
	8	1	0	1	0

Weighing the genres

		Comedy	adventure	super hero	Sci-Fi
2		0	1	1	0
10	X	1	1	1	1
8		1	0	1	0

Weighing the genres

	Comedy	adventure	super hero	Sci-Fi
2	0	1	1	0
10	1	1	1	1
8	1	0	1	0

X

=

	Comedy	adventure	super hero	Sci-Fi
	0	2	2	0
	10	10	10	10
	8	0	8	0
User profile	18	12	20	8

Weighing the genres

	Comedy	adventure	super hero	Sci-Fi
Spider-Man	0	2	2	0
Iron Man	10	10	10	10
Captain America	8	0	8	0
User profile	0.3	0.2	0.33	0.16

User profile

	Comedy	adventure	super hero	Sci-Fi
2	0	1	1	0
10	1	1	1	1
8	1	0	1	0

X

=

Candidate movies for recommendation
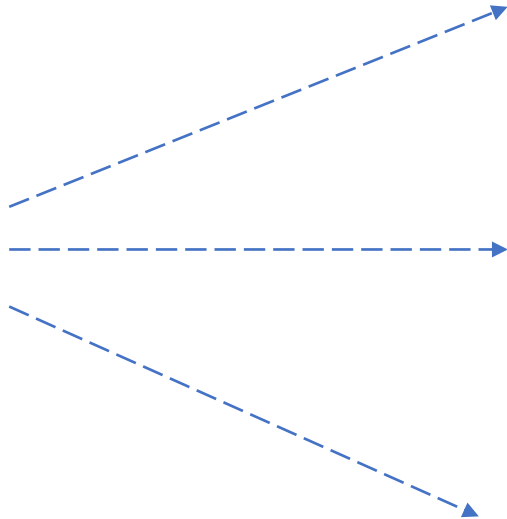
	Comedy	adventure	super hero	Sci-Fi
	1	1	0	1
	0	0	1	0
	1	0	1	0

Finding the recommendation

0.3   0.2   0.33   0.16

X

	Comedy	adventure	super hero	Sci-Fi
	1	1	0	1
	0	0	1	0
	1	0	1	0

=

	Comedy	adventure	super hero	Sci-Fi
	0.3	0.2	0	0.16
	0	0	0.33	0
	0.3	0	0.33	0

Σ

0.66
0.33
0.63

Content-based recommendation systems

Spider man (Adventure, super hero)

(Comedy, Adventure, super hero, Sci-Fi)

Iron man

2

10

(Comedy, super hero)

8

Captain American

6.6

Hitchhiker's guide to the galaxy (Comedy, super hero, Sci-Fi)

3.3

(Super hero)

Avenger

6.3

(comedy, Super hero)

The dark knight rises

Content-based recommendation systems

Spider man (Adventure, super hero)

 (Comedy, Adventure, super hero, Sci-Fi)

2

10 Iron man

Captain American (Comedy, super hero)

8

6.6 Hitchhiker's guide to the galaxy (Comedy, super hero, Sci-Fi)

3.3 Avenger (Super hero)
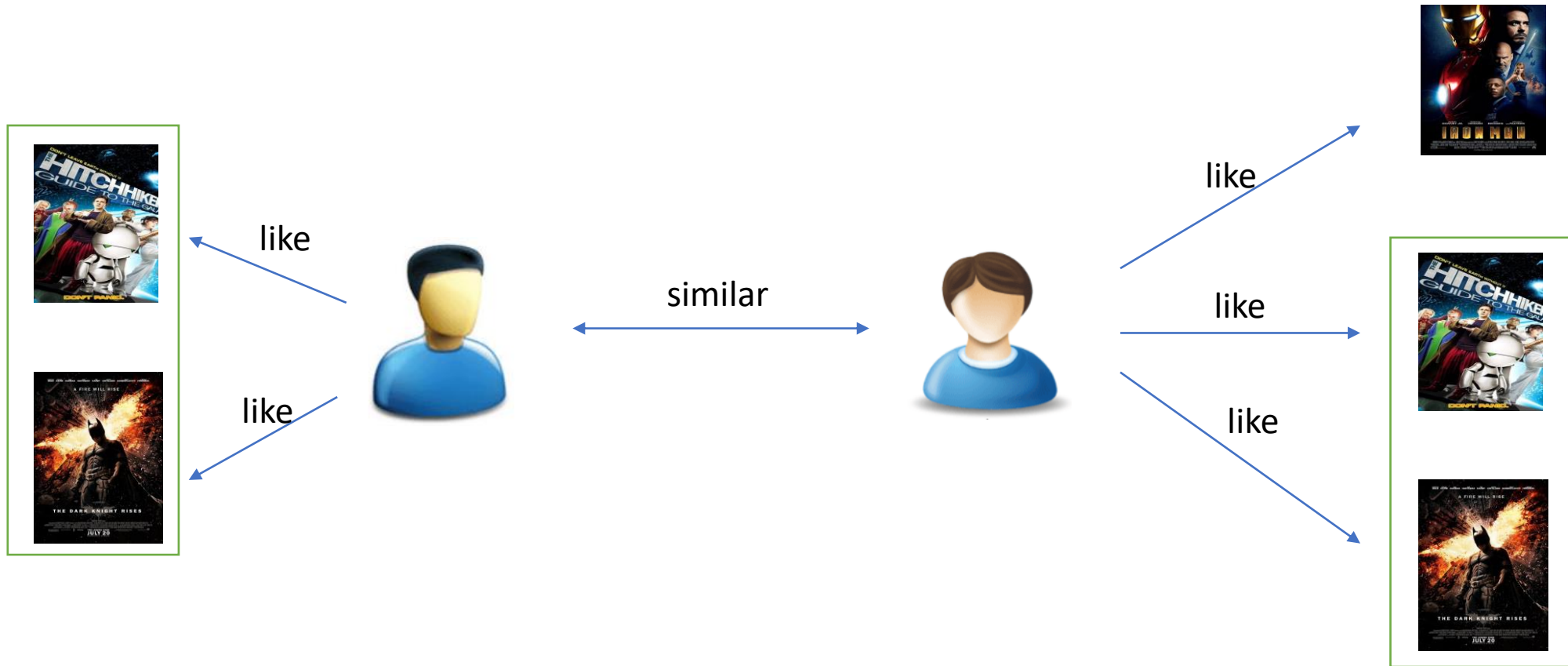
6.3

The king (drama)

The dark knight rises (comedy, Super hero)

16

Collaborative filtering

- User-based collaborative filtering
  - Based on users' neighborhood

- Item-based collaborative filtering
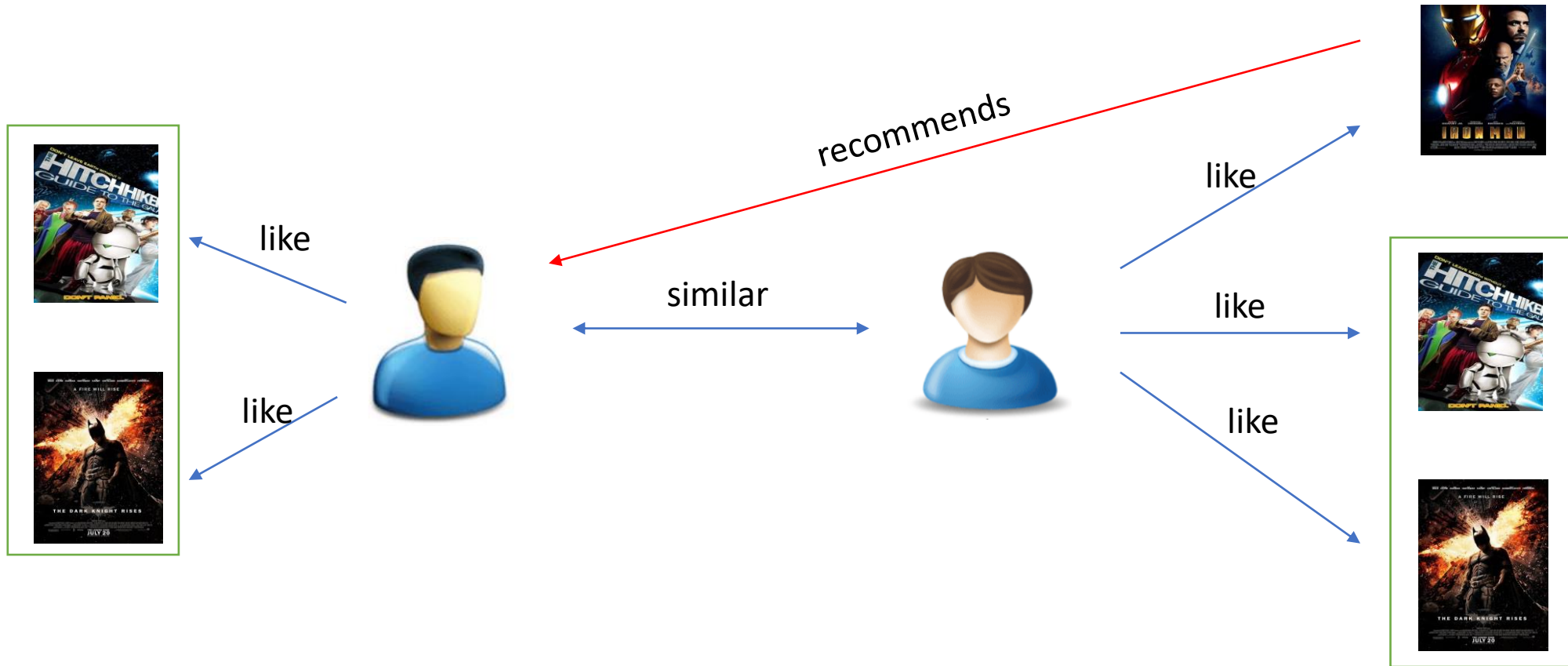 - Based on items' similarity

Collaborative filtering

- User-based collaborative filtering

Collaborative filtering

- User-based collaborative filtering

User ratings matrix

	Spider-Man	Iron Man	Captain America	Hitchhiker's Guide	Avengers
	9	6	8	4	
	2	10	6		8
	5	9		10	7
Active user		10	7	8	

Rating matrix

Learning the similarity weights

	Spider-Man	Iron Man	Captain America	Hitchhiker's Guide	Avengers
	9	6	8	4	
	2	10	6		8
	5	9		10	7
Active user	?	10	7	8	?

Rating matrix

Learning the similarity weights

	Spider-Man	Iron Man	Captain America	Hitchhiker's Guide	Avengers
	9	6	8	4	
	2	10	6		8
	5	9		10	7
	?	10	7	8	?

0.4

0.9

0.7

Rating matrix

Creating the weighted ratings matrix

Rating Matrix Subset × Similarity Matrix = Weighted Ratings Matrix

	Spider-Man	Avengers
Woman	9	
Man (blond)	2	8
Man (tie)	5	7

Similarity Index

	Similarity
Woman	0.4
Man (blond)	0.9
Man (tie)	0.7

	Spider-Man	Avengers
Woman	3.6	
Man (blond)	1.8	7.2
Man (tie)	3.5	4.9

Creating the weighted ratings matrix

Rating and weight matrix

Σ

Sum-weighted rating

Weighted Ratings Matrix

Creating the weighted ratings matrix

Sum similarity matrix

Σ

			Weight sum
		12.1	(0.9+0.7)
	8.9		(0.4 + 0.9+0.7)

0.4
0.9
0.7

Rating and weight matrix

Σ

Sum-weighted rating

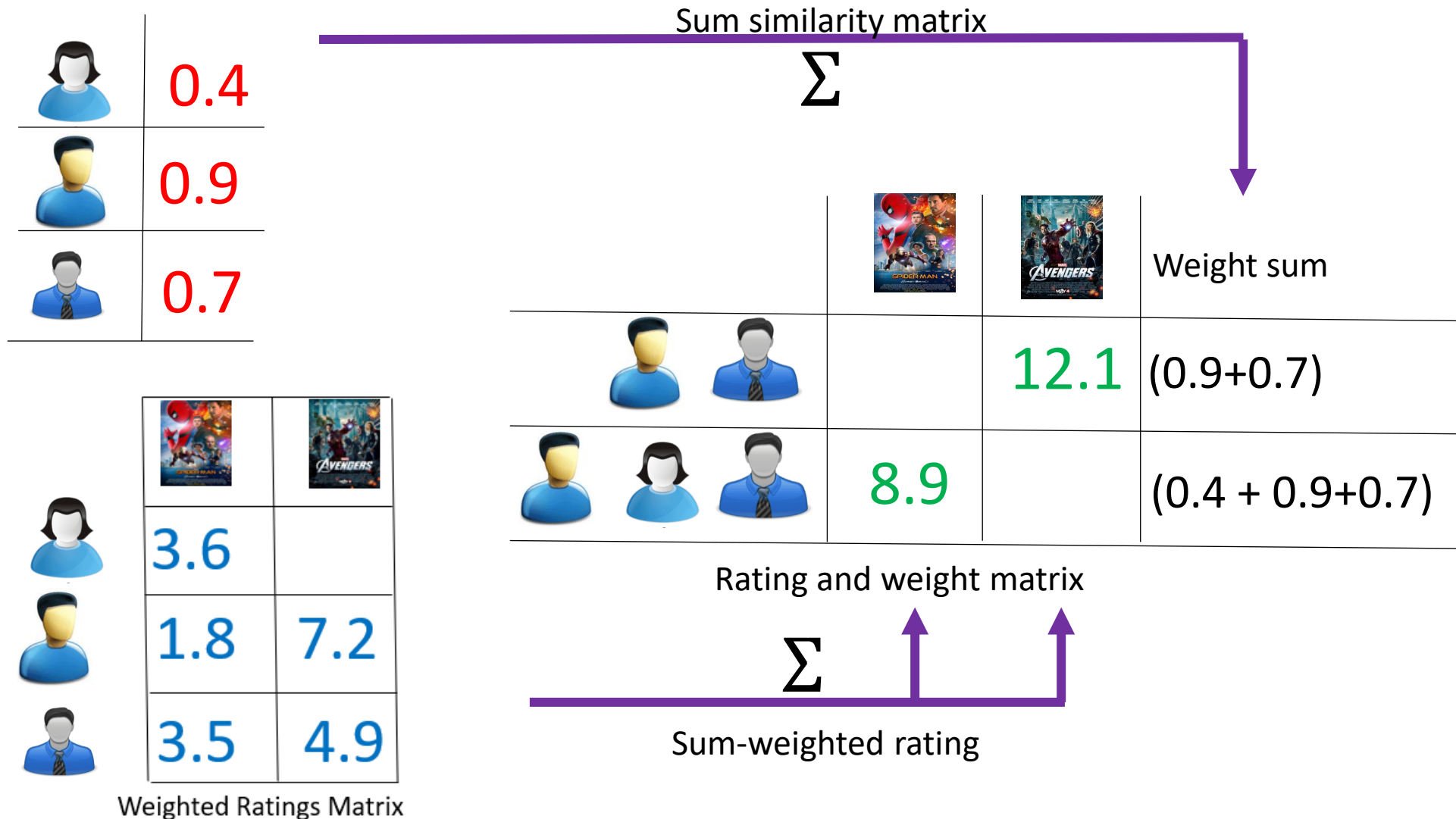|	
3.6	
1.8	7.2
3.5	4.9

Weighted Ratings Matrix

Creating the weighted ratings matrix

Sum similarity matrix

Σ

			Weight sum
		12.1	(0.9+0.7)
		8.9	(0.4 + 0.9+0.7)

Rating and weight matrix

Σ

Sum-weighted rating

Weighted Ratings Matrix

	3.6	
	1.8	7.2
	3.5	4.9

	0.4	
	0.9	
	0.7	

	4.4	7.5

Recommendation matrix

Collaborative Filtering

User-based

Item-based

Similar user

Similar items

27
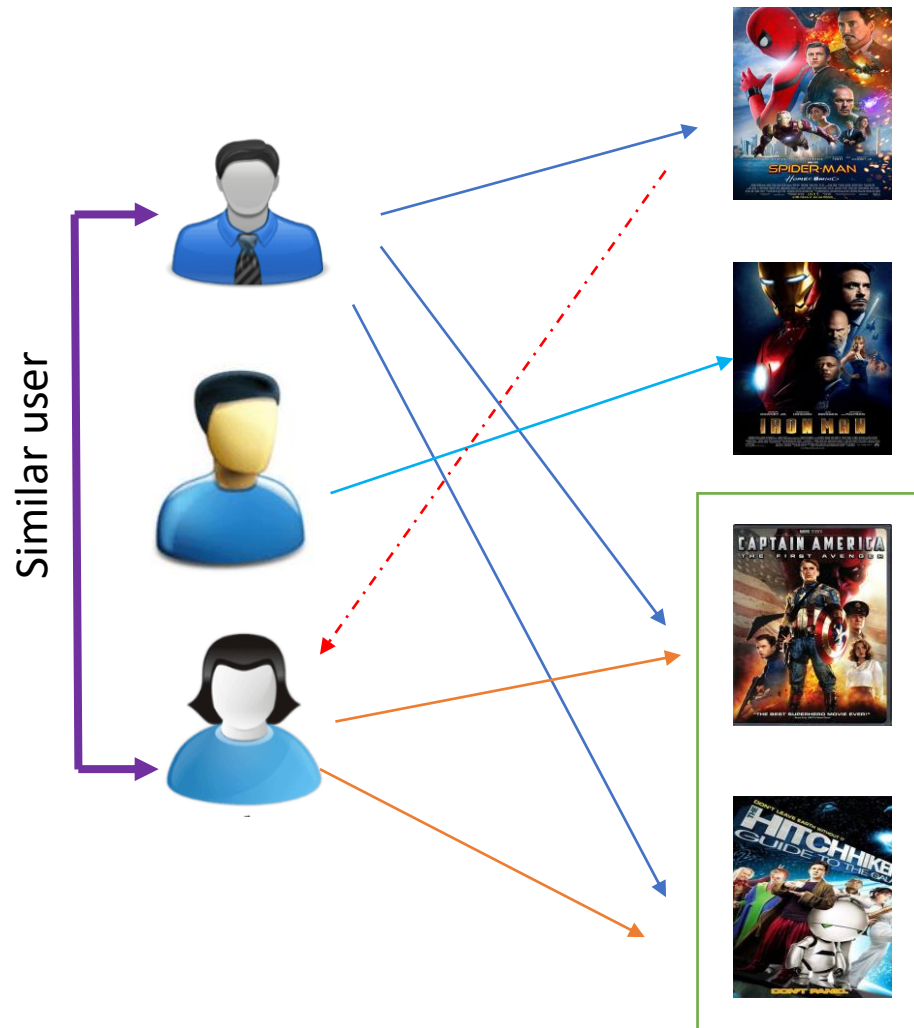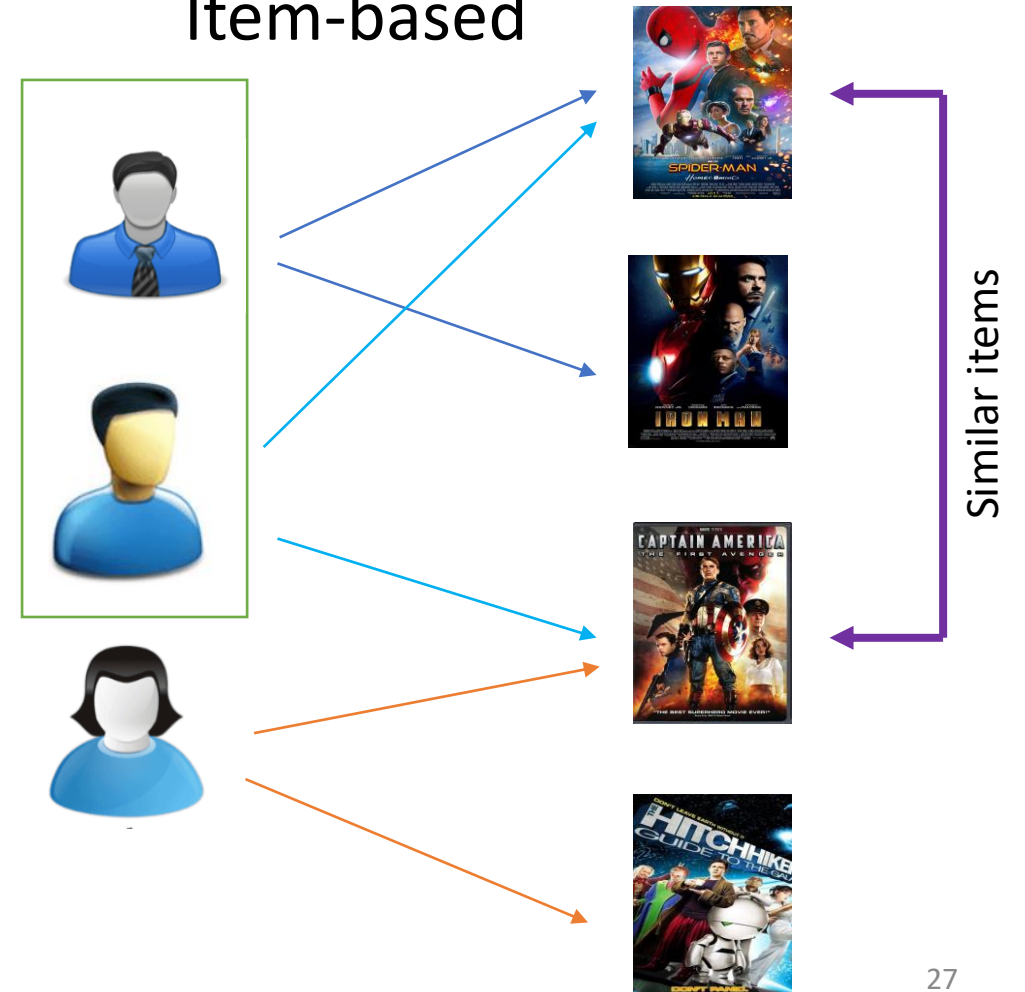
Challenges of collaborative filtering

- Data sparsity
  - Users in general rate only a limited items

- Cold start
  - Difficulty in recommendation to new users or new item

- Scalability
  - Increase in number of users or item