

Recommendation Systems



Akash Khandelwal, Avijit Saha, Mohit Kumar, Vivek Mehta

Happy Festival Season!!!



Recommendations - Overview



Flipkart  [Signup](#) | [Login](#)

So, what are you wishing for today? SEARCH CART 5

[ELECTRONICS](#) [MEN](#) [WOMEN](#) [BABY & KIDS](#) [HOME & FURNITURE](#) [BOOKS & MEDIA](#) [AU](#)

Puma Striped Men's Polo T-Shirt

 ★ ★ ★ ★ 1 Write a REVIEW Add to WISHLIST

Select Color  Select Size  SIZE CHART

 Change

Rs. 1,499
List Price
(Free delivery)

ADD TO CART BUY NOW

Unlock mega savings with Great Offers OPEN IN APP

Easy Returns Brand New 100% Original Pay Securely

SOLD BY **WS Retail** 4.2 / 5  [?](#)

DELIVERED BY ? **CASH ON DELIVERY** ?
Mon, 19th Oct: **FREE** ?
Available

30 day Exchange Guarantee. ?

Flipkart 

SEARCH  CART 3 

So, what are you wishing for today?

ELECTRONICS MEN WOMEN BABY & KIDS HOME & FURNITURE BOOKS & MEDIA AUTO & SPORTS OFFER

Home > Mobiles & Accessories > Mobiles > Motorola Mobiles > Moto G (3rd Generation) (Black, 16 GB)

Moto G (3rd Generation) (Black, 16 GB)



★★★★★ 8314 2,273 REVIEWS   

WARRANTY
1 year manufacturer warranty for Phone and for in the box accessories

Color  

Storage  

 [View Compatible Accessories](#)

 Available with 1 Seller at 560035 [Change](#)

Rs. 12,999
List Price

EMI starts from Rs. 631  

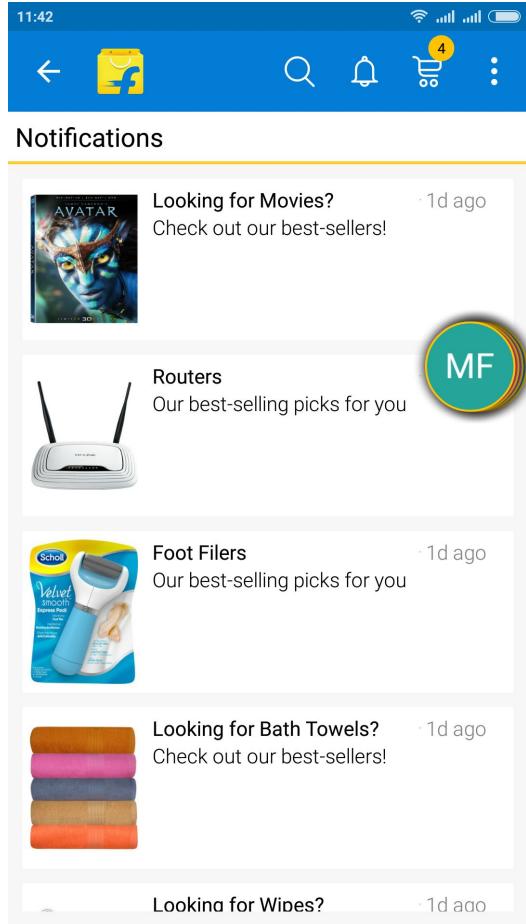
DELIVERED BY: ?  
Sat, 17th Oct: FREE 

CASH ON DELIVERY 
Available

30 day Replacement Guarantee. 

ADD TO CART **BUY NOW**

Recommendation Overview



- Related product recommendation (Flipkart)
- Movie recommendation (Netflix)
- Web page ranking (Google)
- Social recommendation (Facebook)
- News content recommendation (Yahoo)
- Priority inbox & spam filtering (Google)
- Online dating (OK Cupid)
- Computational Advertising (Yahoo)
- Event recommendation



- Type of Methods
 - Collaborative filtering
 - Content based
 - Hybrid
- Collaborative Filtering
 - Neighbourhood methods
 - Matrix factorization



- Slides are prepared by leveraging internet heavily
 - <http://www2.research.att.com/~volinsky/papers/ieeecomputer.pdf>
 - <http://www.slideshare.net/xamat>
 - http://alex.smola.org/teaching/berkeley2012/slides/8_Recommender.pdf
 - <http://blog.comsysto.com/2013/04/03/background-of-collaborative-filtering-with-mahout/>
 - [Simon funk's blog](#)



- User activity

- Movies
 - Ratings, reviews, movies rented, watched
- E-commerce
 - Browse, wishlist, share, ratings, etc.



Collaborative
filtering

- Item details

- Movies
 - Actors, Directors, Genre, etc.
- E-commerce
 - Product attributes, images, description



Content
Based



Hybrid

CF - Neighbourhood Methods

Neighbourhood Method



CUSTOMERS WHO VIEWED THIS PRODUCT ALSO VIEWED



GAS Printed Men's Round Neck T-Shirt



GAS Printed Men's Round Neck T-Shirt



GAS Printed Men's Round Neck T-Shirt



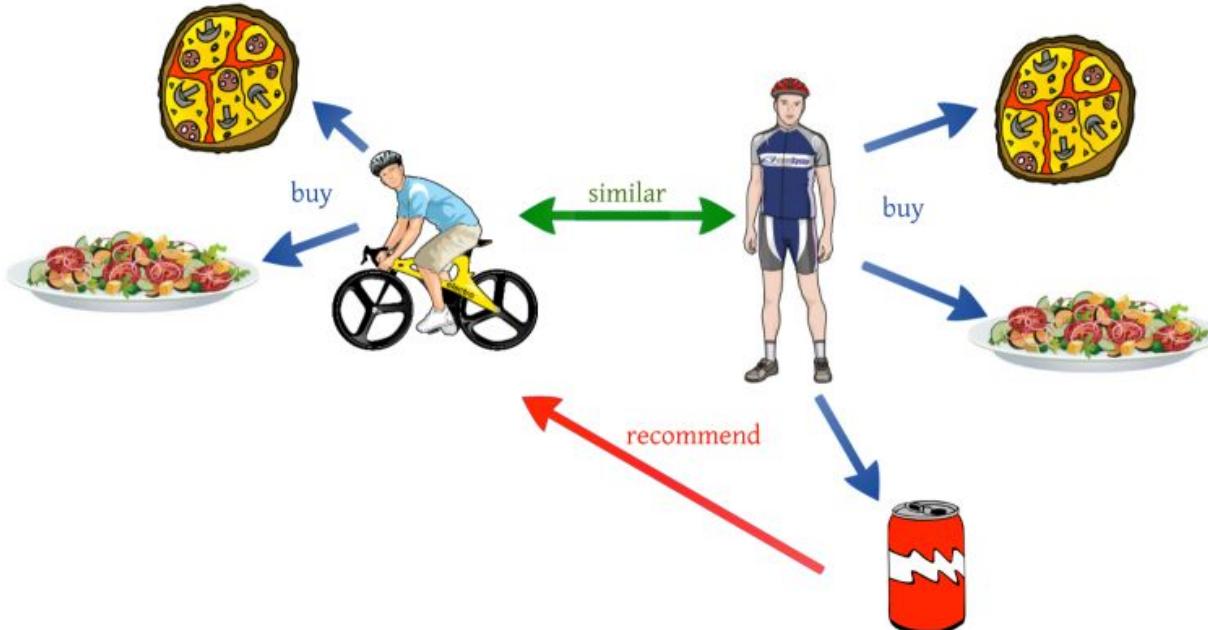
GAS Printed Men's Round Neck T-Shirt



Gas Printed Men's Round Neck T-Shirt



Neighbourhood Method



Neighbourhood Method - Setup



		Users				
		3	2	1		
Items	3					
	4	4	?	5	3	
	1	2	2		4	
	2		2		4	

(user, user) similarity



- $r_{ui} = ?$
- For user u , find similar users (u')
- Find likeness of item i for similar users.

$$r_{u,i} = \frac{1}{N} \sum_{u' \in U} r_{u',i}$$

$$r_{u,i} = k \sum_{u' \in U} \text{simil}(u, u') r_{u',i} \quad k = 1 / \sum_{u' \in U} |\text{simil}(u, u')|$$

$$r_{u,i} = \bar{r}_u + k \sum_{u' \in U} \text{simil}(u, u') (r_{u',i} - \bar{r}_{u'})$$



- Pearson correlation

$$\text{simil}(x, y) = \frac{\sum_{i \in I_{xy}} (r_{x,i} - \bar{r}_x)(r_{y,i} - \bar{r}_y)}{\sqrt{\sum_{i \in I_{xy}} (r_{x,i} - \bar{r}_x)^2 \sum_{i \in I_{xy}} (r_{y,i} - \bar{r}_y)^2}}$$



- Similar to (user, user) similarity
- Find users who have liked a product
- Find other products liked by these users



- Pros:
 - Intuitive and easy to explain to user
 - Easy to setup
 - Produces good-enough results
- Cons:
 - Scalability
 - Sparseness -> coverage -> cold start

Matrix factorization



4	?	?	?
?	?	?	5
?	1	?	2
3	?	?	?
?	2	?	?

$5 * 4$

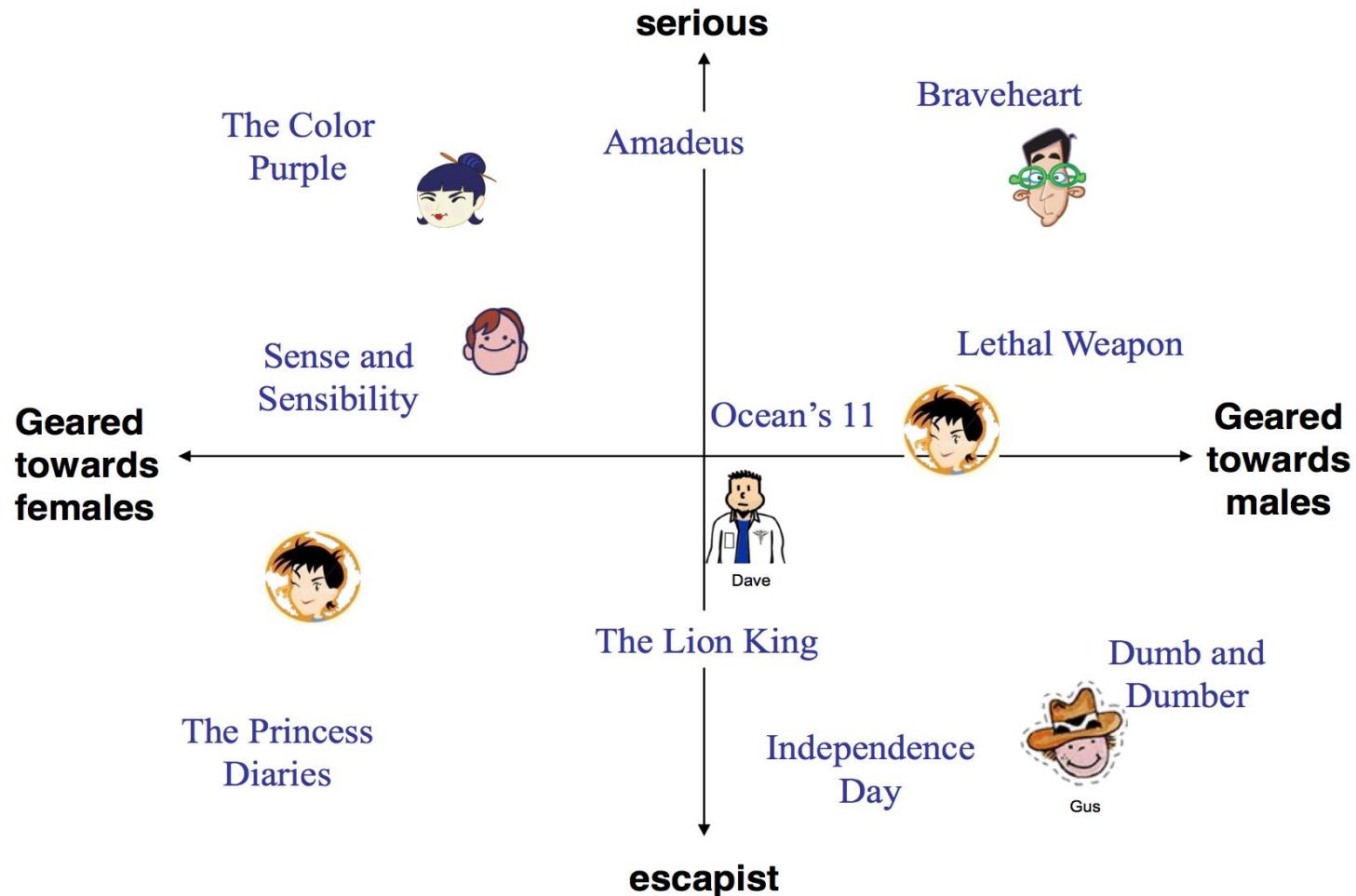
=

$5 * 2$

\times

$2 * 4$

$$\hat{r}_{ui} = q_i^T p_u$$



Matrix factorization



- Similar to SVD, but missing values!!
- Overfitting is one major problem
- SGD for learning latent factors
- Objective function:

$$\min_{q^*, p^*} \sum_{(u,i) \in \kappa} (r_{ui} - q_i^T p_u)^2 + \lambda (||q_i||^2 + ||p_u||^2)$$

Regularization Constant

Stochastic gradient descent



- Compute error

$$e_{ui} \stackrel{\text{def}}{=} r_{ui} - q_i^T p_u.$$

- Adjust the factors

$$q_i \leftarrow q_i + \gamma \cdot (e_{ui} \cdot p_u - \lambda \cdot q_i)$$

$$p_u \leftarrow p_u + \gamma \cdot (e_{ui} \cdot q_i - \lambda \cdot p_u)$$

Learning rate

Adding Biases

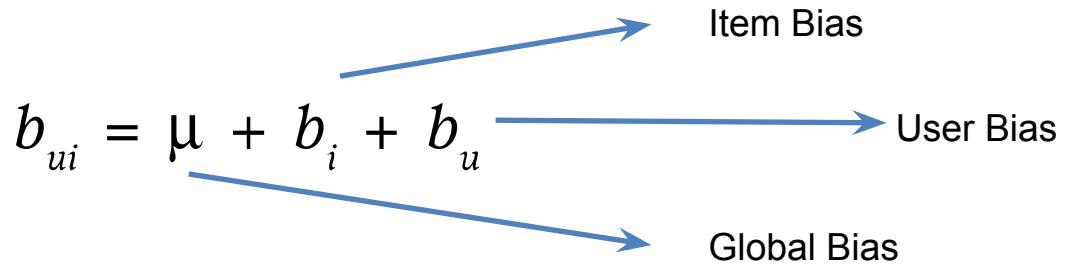


$$b_{ui} = \mu + b_i + b_u$$

Item Bias

User Bias

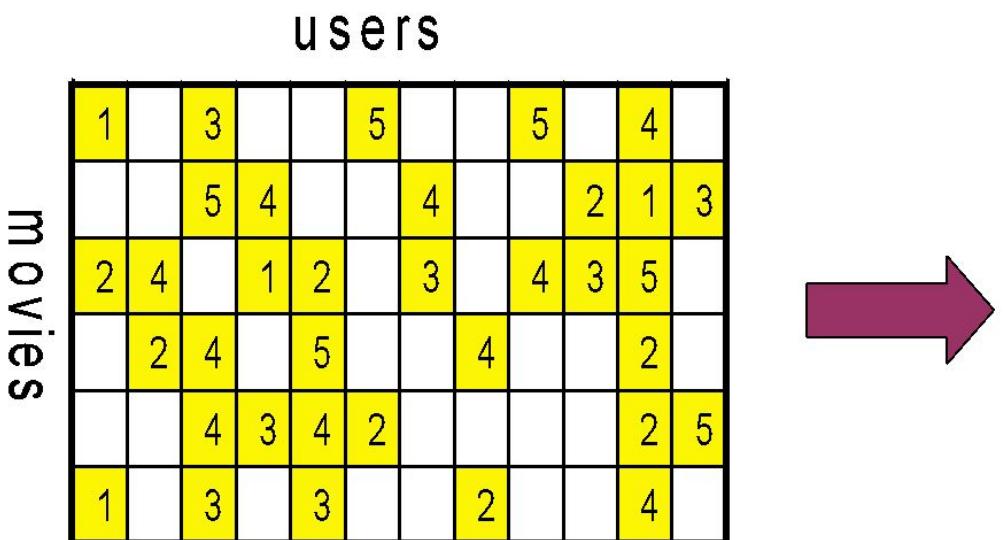
Global Bias



$$\hat{r}_{ui} = \mu + b_i + b_u + q_i^T p_u$$

$$\min_{p^*, q^*, b^*} \sum_{(u, i) \in \mathcal{K}} (r_{ui} - \mu - b_u - b_i - p_u^T q_i)^2 + \lambda (\|p_u\|^2 + \|q_i\|^2 + b_u^2 + b_i^2)$$

Implicit feedback



users

1	0	1	0	0	1	0	0	1	0	1	0
0	0	1	1	0	0	1	0	0	1	1	1
1	1	0	1	1	0	1	0	1	1	1	0
0	1	1	0	1	0	0	1	0	0	1	0
0	0	1	1	1	1	0	0	0	0	1	1
1	0	1	0	1	0	0	1	0	0	1	0

movies



- Objective function with implicit feedback formulation

$$\min_{p^*, q^*, b^*} \sum_{(u, i) \in \mathcal{K}} c_{ui} (r_{ui} - \mu - b_u - b_i - p_u^T q_i)^2 + \lambda (\|p_u\|^2 + \|q_i\|^2 + b_u^2 + b_i^2)$$



Thank You