

# Movie Recommendation System

Neslihan Bulut


# Recommendation Systems

CS533: Information Retrieval x collaborative filtering - Google x www.stanford.edu/~mori x Amazon.com: Recommen x


← → C [https://www.amazon.com/gp/yourstore/home?ie=UTF8&ref\\_=topnav\\_ys](https://www.amazon.com/gp/yourstore/home?ie=UTF8&ref_=topnav_ys) ☆ K ≡

Your Amazon.com


Amazon Instant Video - Movies Page 1 of 17




The Wolf of Wall ...  
Leonardo DiCaprio  
★★★★☆ (886)  
\$3.99 to rent  
[Why recommended?](#)




Saving Mr. Banks  
Emma Thompson  
★★★★☆ (709)  
\$4.99 to rent  
[Why recommended?](#)




Inside Llewyn Davis  
Justin Timberlake  
★★★★☆ (215)  
\$3.99 to rent  
[Why recommended?](#)



About Time  
Domhnall Gleeson  
★★★★☆ (792)  
\$4.99 to rent  
[Why recommended?](#)




American Hustle  
Christian Bale  
★★★★☆ (772)  
\$3.99 to rent  
[Why recommended?](#)



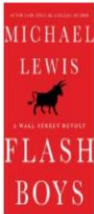
Pi  
Sean Gullette  
★★★★☆ (816)  
\$0.00 Prime  
\$2.99 to rent  
[Why recommended?](#)

See all recommendations in Amazon Instant Video - Movies


Kindle eBooks Page 1 of 17




difference  
BERNARDITA JINA




MICHAEL LEWIS  
FLASH BOYS




ALGORITHMS  
THOMAS H. COOPER



What and What Not to Do  
Liz Taylor



Quiet  
SUSAN CAHILL




The Trial  
By Franz Kafka

Amazon.com: Why is this recommended for you? - Google Chrome

[https://www.amazon.com/gp/yourstore/rate-this-asin/ref=pd\\_ys\\_sf\\_s\\_16386](https://www.amazon.com/gp/yourstore/rate-this-asin/ref=pd_ys_sf_s_16386)

amazon.com [Help](#) | [Close window](#)


Recommended for You




[Pi](#)  
Amazon Instant Video ~ Sean Gullette  
(January 7, 2014)  
\$0.00 Prime  
\$2.99 to rent  
\$9.99 to buy  
[See all buying options](#)

Rate this item  
☒ ★★★★★  
☐ I own it  
☐ Not interested

Because you watched...



[Requiem for a Dream](#) (Amazon Instant Video)  
Saved  
☒ ★★★★★  
☐ Don't use for recommendations



[The Fountain](#) (Amazon Instant Video)  
Saved  
☒ ★★★★★  
☐ Don't use for recommendations

[Help](#) | [Close window](#)

★★★★☆ (887)      ★★★★★ (726)      ★★★★★ (215)

# Collaborative Filtering

- ▶ Predictions on how well a user will like an item
- ▶ Has a training data of items, users and user votes
- ▶ Neighborhood methods are used ; find similar users and make predictions based on their votes

# Netflix Data



# Netflix Data

- ▶ 3M votes for the training data
- ▶ 100K votes for testing the accuracy of the algorithm

# Algorithms: Collaborative Filtering (Breese et al.)

- ▶  $v_{i,j}$  = vote of user  $i$  on item  $j$
- ▶  $I_i$  = items for which user  $i$  has voted
- ▶ Mean vote for  $i$  is

$$\bar{v}_i = \frac{1}{|I_i|} \sum_{j \in I_i} v_{i,j}$$

- ▶ Predicted vote for “active user”  $a$  is weighted sum

$$p_{a,j} = \bar{v}_a + \kappa \sum_{i=1}^n \underbrace{w(a,i)}_{\text{weights of } n \text{ similar users}} (v_{i,j} - \bar{v}_i)$$

normalizer  $\nearrow$

# Algorithms: Collaborative Filtering (Breese et al.) - Cont.

- ▶ Pearson correlation coefficient (Resnick '94, Grouplens):

$$w(a, i) = \frac{\sum_j (v_{a,j} - \bar{v}_a)(v_{i,j} - \bar{v}_i)}{\sqrt{\sum_j (v_{a,j} - \bar{v}_a)^2 \sum_j (v_{i,j} - \bar{v}_i)^2}}$$

- ▶ Most similar users are selected to predict scores for unrated movies
  - ▶ *Empirical Analysis of Predictive Algorithms for Collaborative Filtering*, Breese, Heckerman, Kadie

# Evaluation

- ▶ Test and training sets from the Netflix prize
- ▶ Compare the predictions from the algorithm with the actual vote data from the test set



# Thanks for listening!

► Any questions?