

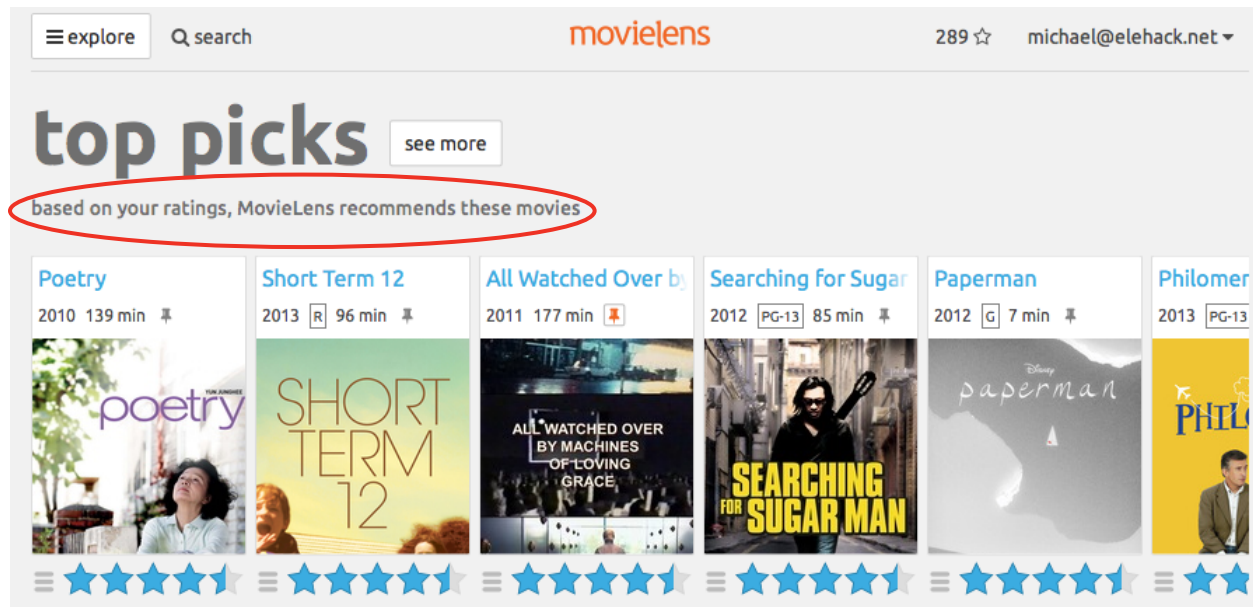
Practical Data Science – Tute 10 / Week 11

PDS TEACHING TEAM

Activity 1 – Recommender Systems

- Give a high-level explanation of what a recommender system is, and what it does.

It is a system that attempts to **recommend products** which a **user** would probably **like**, based on **his/her preferences profile** extracted from their consuming/rating history.

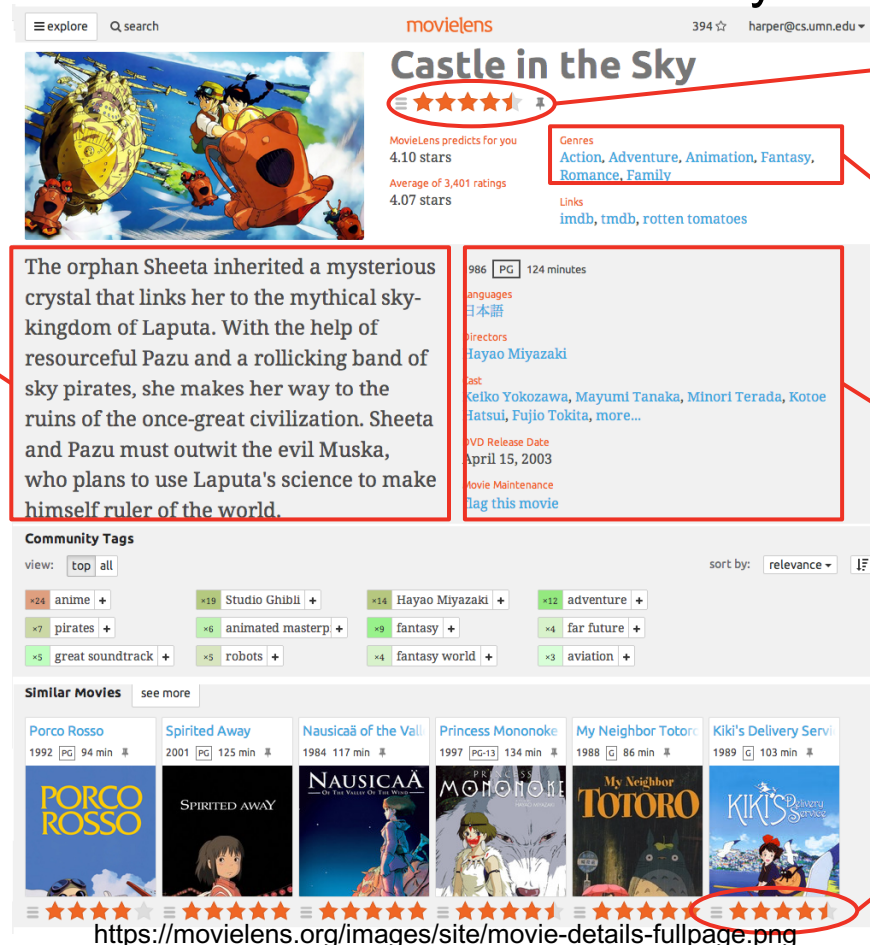


<http://md.ekstrandom.net/talks/2014/txstate-recsys-research/movielens.png>

Activity 1 – Recommender Systems

- User preference profiles are generally extracted from two sources of information in their historical data. What are they?

Content
Movie description



Ratings
Movie Rating

Content
Movie Genre

Content
Movie metadata

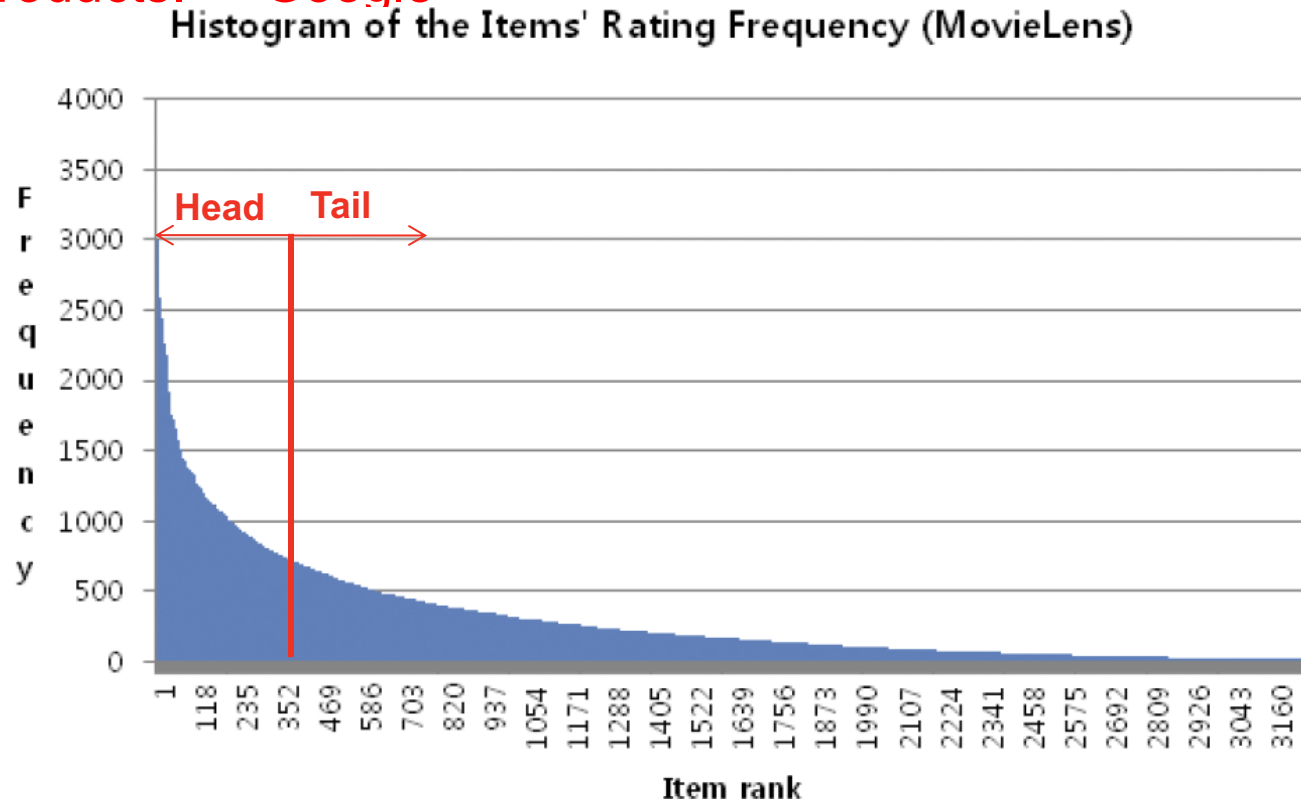
Ratings
Similar Movies
Ratings

<https://movielens.org/images/site/movie-details-fullpage.png>

Activity 1 – Recommender Systems

➤ What is the Theory of Long Tail?

“(in retail and marketing) used to refer to the **large number** of products that **sell in small** quantities, as contrasted with the **small number** of **best-selling** products.” – Google

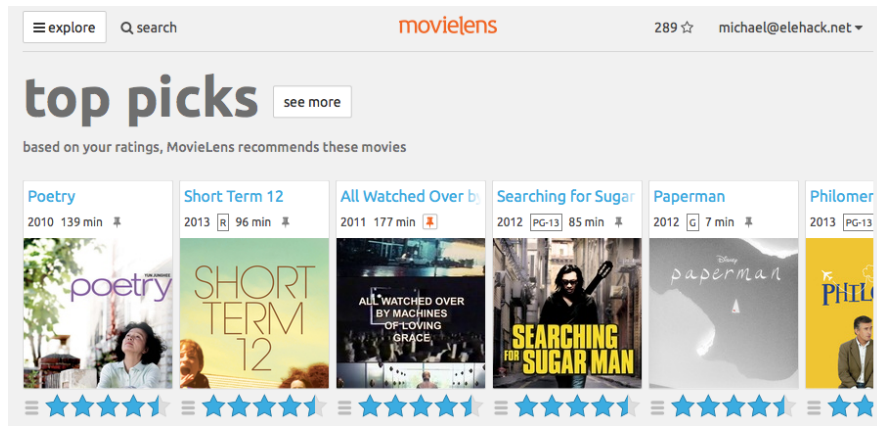


<http://doi.ieeecomputersociety.org/cms/Computer.org/dl/trans/tk/2013/08/figures/ttk20130819041.gif>

Activity 1 – Recommender Systems

- What is the purpose of recommender systems?

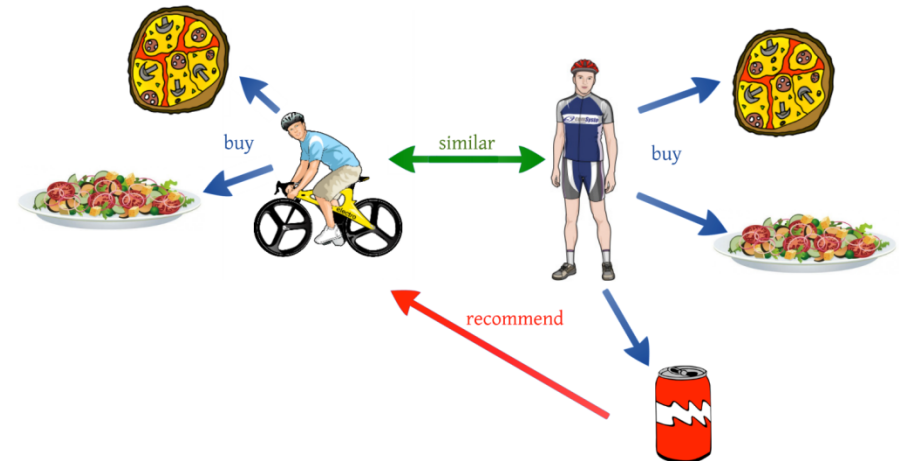
User perspective



<http://md.ekstrandom.net/talks/2014/txstate-recsys-research/movielens.png>

For a particular **user**, recommend **items** he/she may be interested in.

Item perspective



<http://dataconomy.com/wp-content/uploads/2015/03/Beginners-Guide-Recommender-Systems-Collaborative-Filtering.png>

For a particular **item**, identify potential **users** who may like it.

Activity 1 – Recommender Systems

- What is a user-item rating matrix? How are users represented?

Items. e.g. films

| Users | Spider-Man | 2012 | The Godfather | The Social Network |
|-------|------------|------|---------------|--------------------|
| Bob | 3 | 2 | 4 | 5 |
| Cindy | ∅ | 3 | 5 | 4 |
| Paul | 2 | 3 | 4 | ∅ |
| David | 3 | ∅ | 4 | 1 |

rating No rating

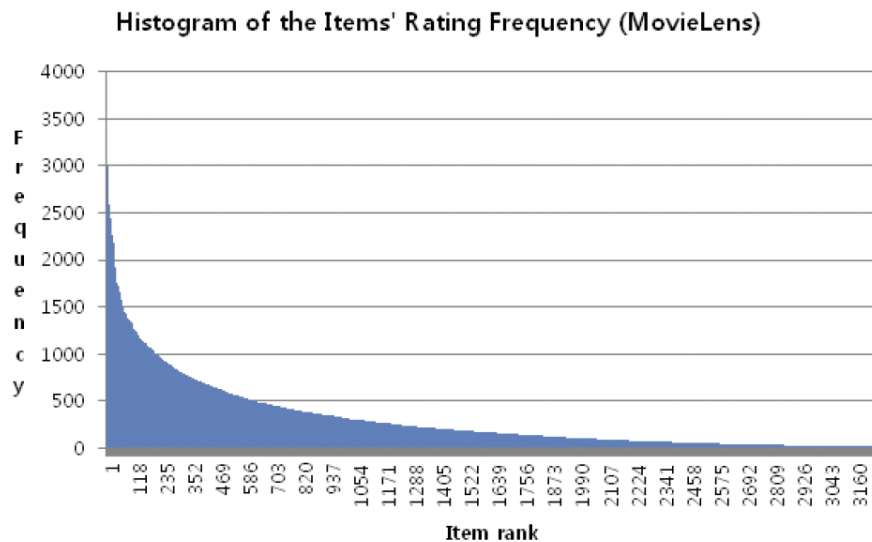
David did not watch '2012' ?
David does not like '2012' ?

A long vector with each cell value representing the users' rating on the corresponding item (e.g. movie).

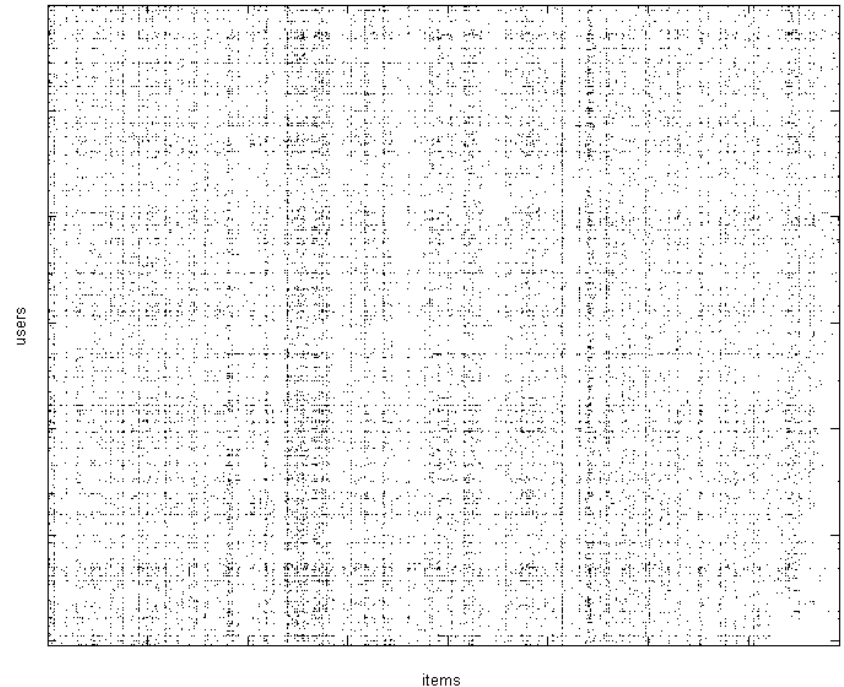
Activity 1 – Recommender Systems

➤ What does Data Sparsity mean?

Insufficient information on a user's rating history



<http://doi.ieeecomputersociety.org/cms/Computer.org/dl/trans/tk/2013/08/figures/ttk20130819041.gif>

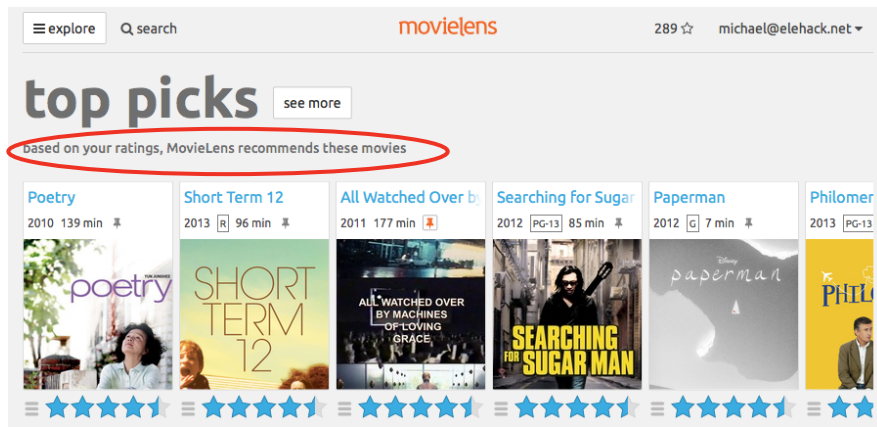


Activity 1 – Personalised Recommendation

➤ What is a personalised recommendation technique?

Recommended items are generated based on a user's rating/consuming history.

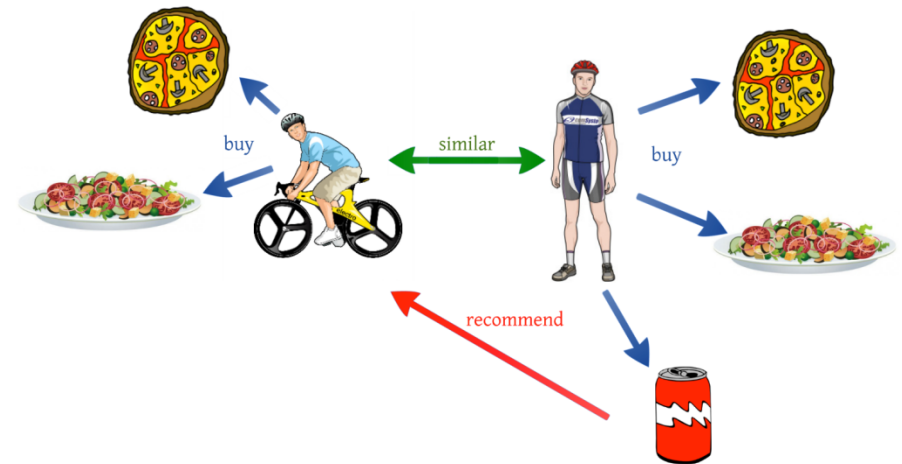
Content-based



<http://md.ekstrandom.net/talks/2014/txstate-recsys-research/movielens.png>

Estimate the rating on an unknown item based on the content of that item and the contents of other rated items.

Collaborative filtering



<http://dataconomy.com/wp-content/uploads/2015/03/Beginners-Guide-Recommend-Systems-Collaborative-Filtering.png>

Predict the rating on an unknown item based on the ratings on that item by other similar users.

Activity 1 – Personalised Recommendation

➤ What is the New User Problem?

Refers to the difficulty of making recommendations to a new user who has no rating history – NO USER PROFILE

How to solve this problem?!

Activity 2 – Non personalised Recommendation

- What is a non-personalised recommendation technique?

Recommended items are generated without considering a user's rating/consuming history.

TopPop

Recommends items with the highest popularity (**largest number of ratings**).

MovieAvg

Recommends top-N items with the **highest average rating**.

Example on imdb (explore the “Sort by” box)

<http://www.imdb.com/chart/moviemeter?sort=nv,desc&mode=simple&page=1>

Questions?