Social Networks & Recommendation Systems

XIII. Recommendation Systems.

Grzegorz Siudem

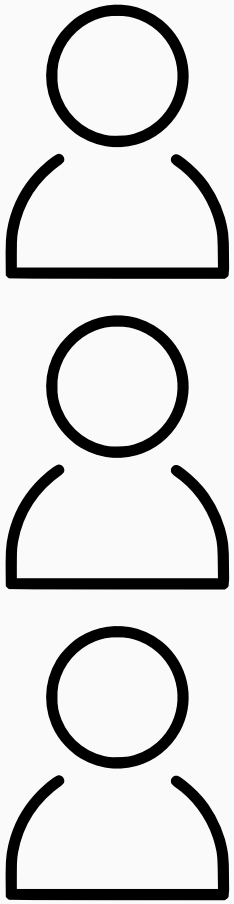
Warsaw University of Technology

Warsaw University of Technology

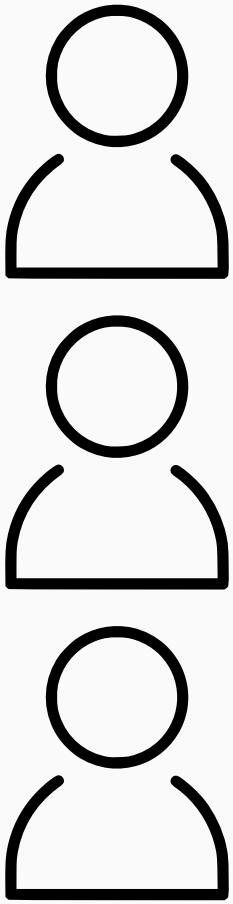
MSc program in Data Science has been developed
as a part of task 10 of the project
„NERW PW. Science - Education - Development - Cooperation”
co-funded by European Union from European Social Fund.

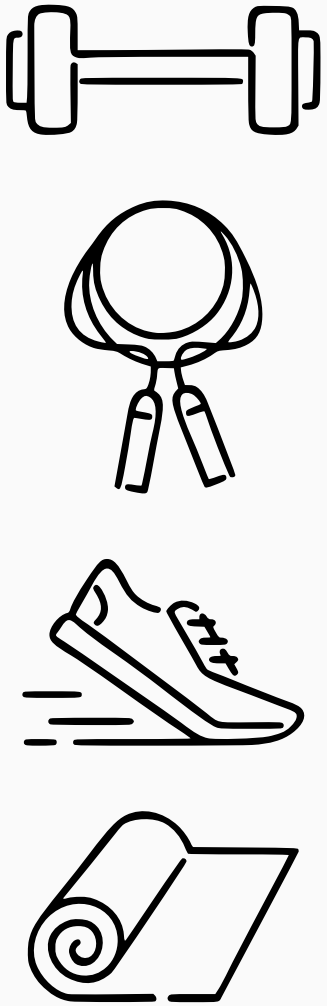
Lecture

Recommendation Systems

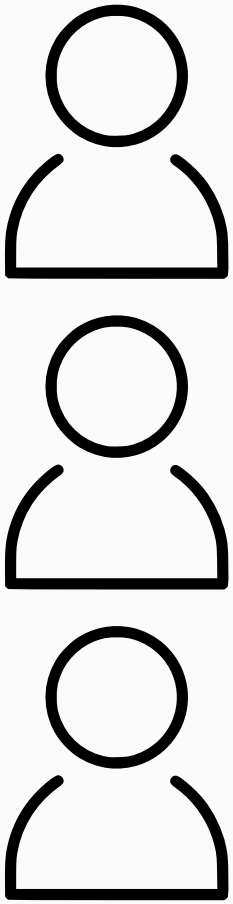


Recommendation Systems

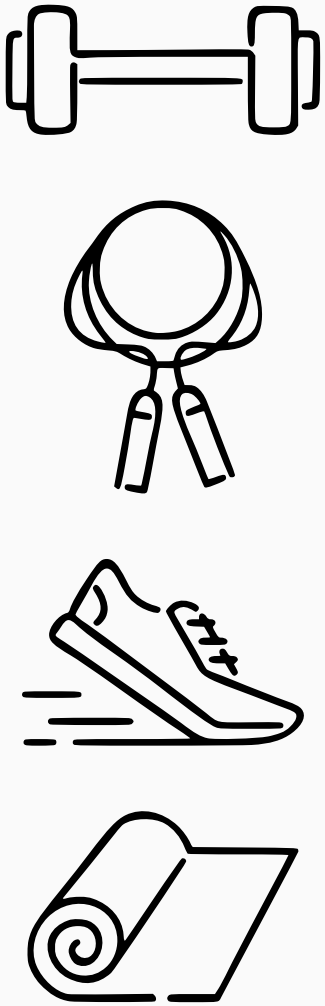




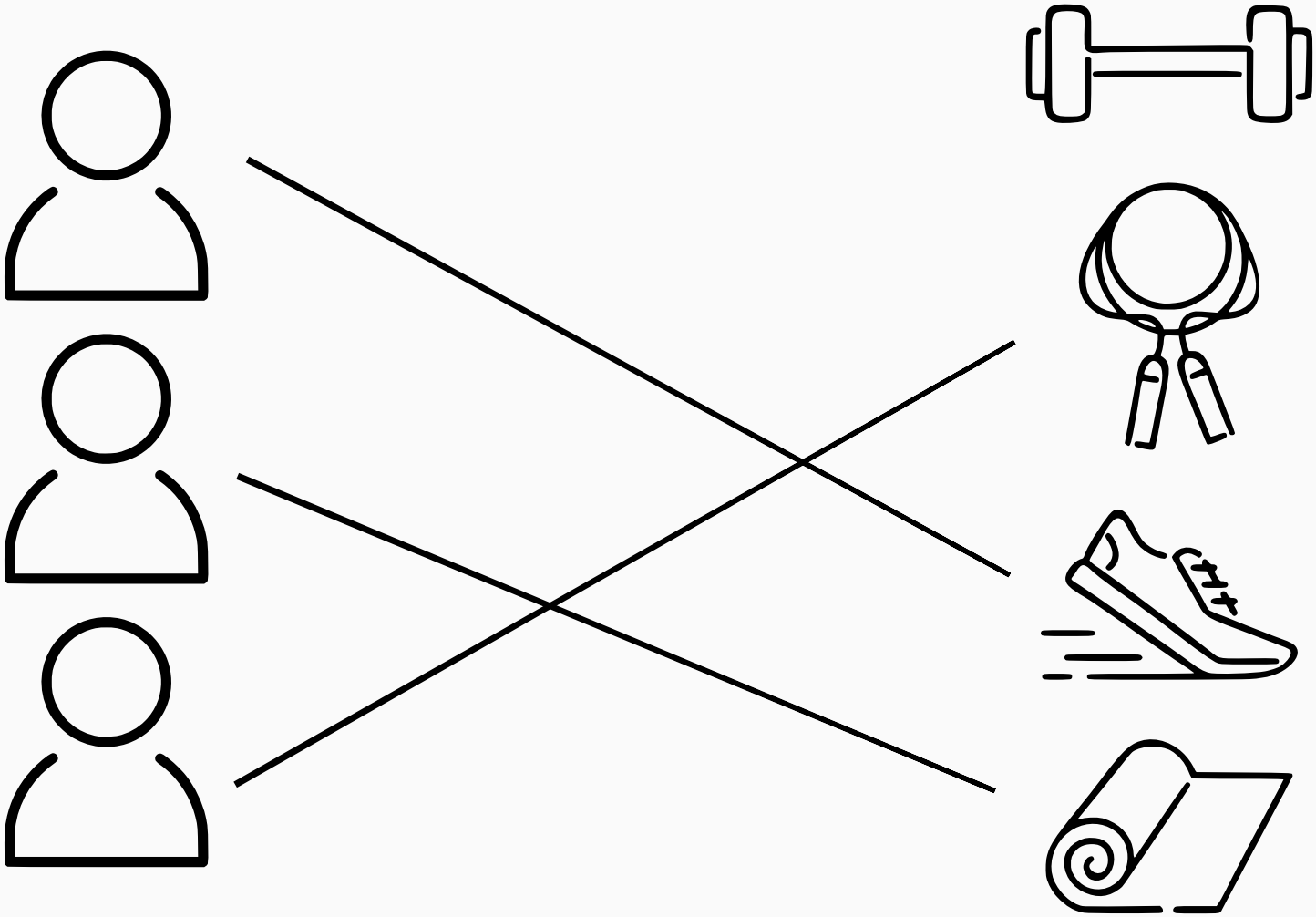
Recommendation Systems



?



Recommendation Systems



Recommendation Systems: applications

What can be recommended?

- movies on Netflix,
- books on Amazon,
- music at Spotify,
- games at GOG/Steam,
- content on Social Networks (Twitter, Facebook),
- matching on online dating services,
- news on the informational pages,
- other products in online shops/services.

Recommendation Systems: definitions

Formal definition

The goal of a recommender system is to generate meaningful recommendations to a collection of users for items or products that might interest them.[1]

Less formally

We are looking for the best matching user-item pairs.

From networks perspective

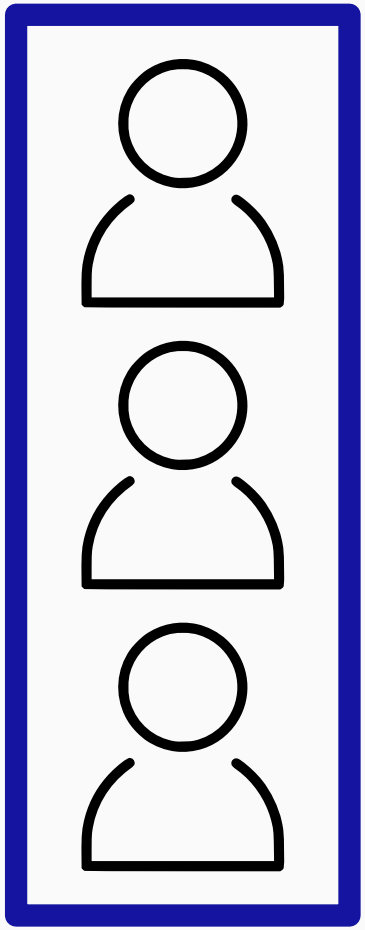
Which is optimal edges configuration in the bipartite graph (user-item)?

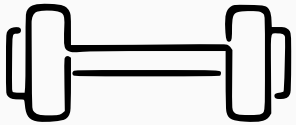
What does it mean *optmial*?

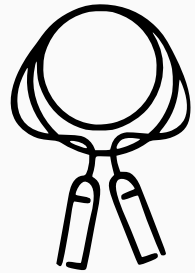
[1] Encyclopedia of Machine Learning

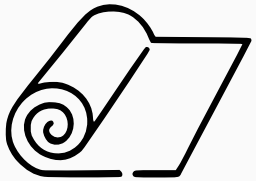
How to measure goodness of recommendation?

Ideally we would ask every user...



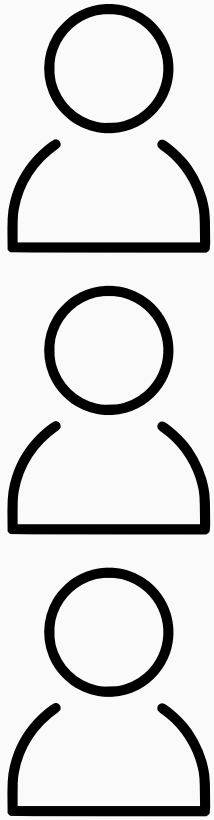


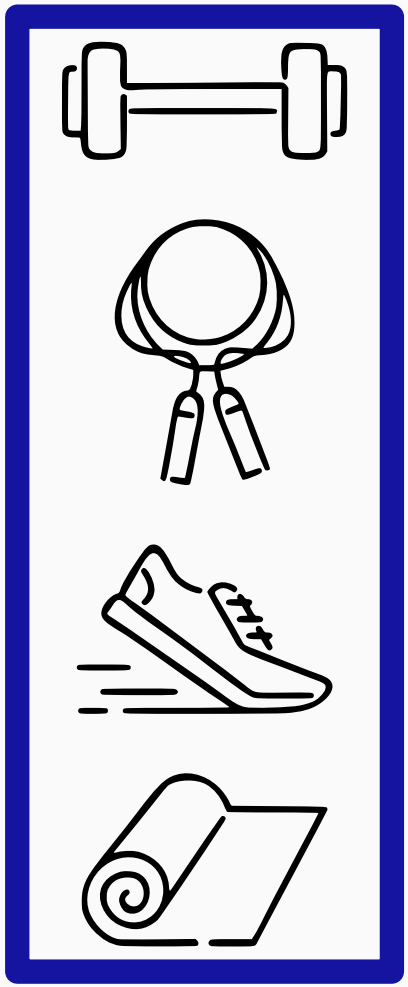




How to measure goodness of recommendation?

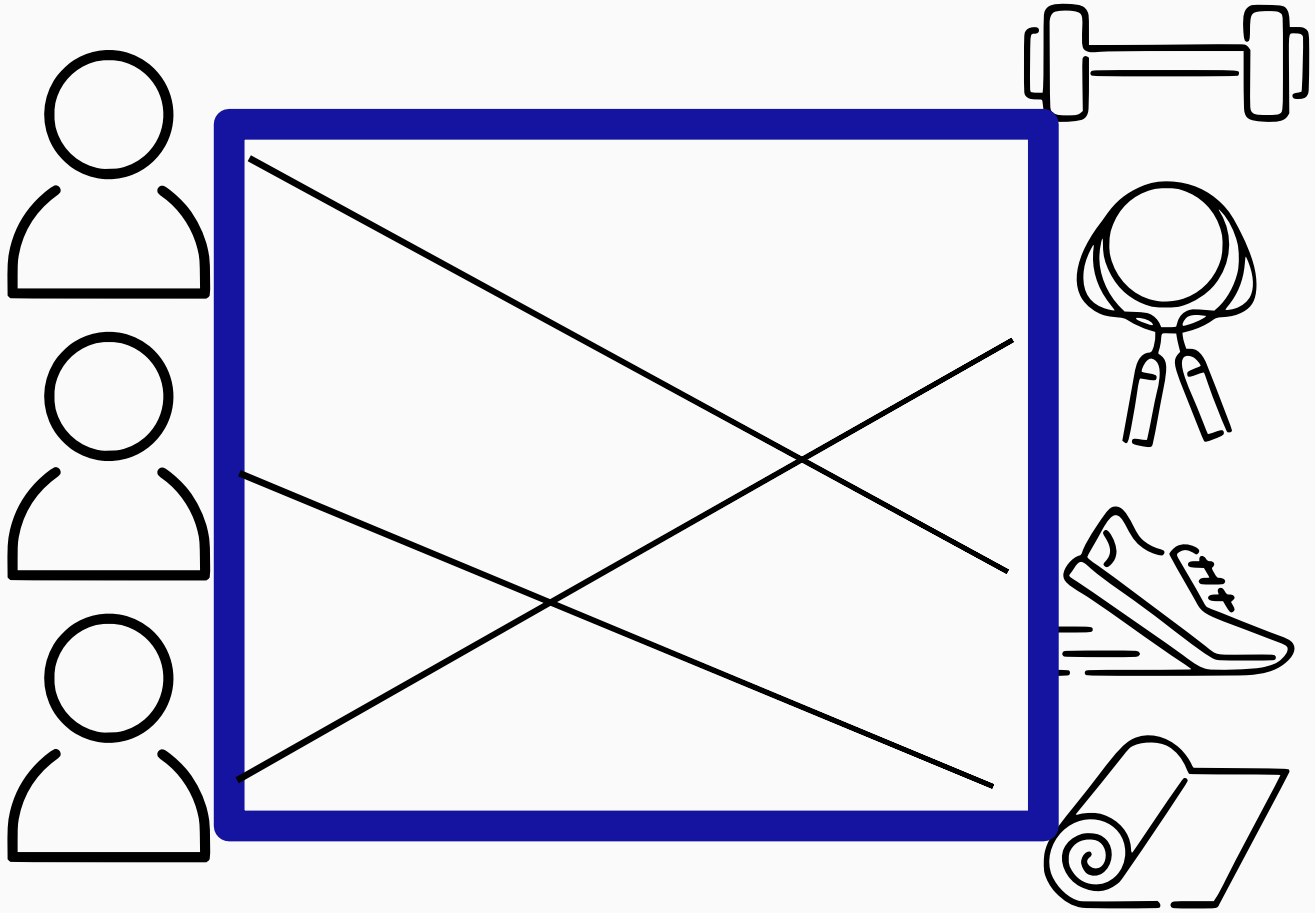
We could also focus on the items.





How to measure goodness of recommendation?

What with already known connections?



How to measure goodness of recommendation?

Recommendation Systemsare algorithms and techniques that try to predict what the user wants.

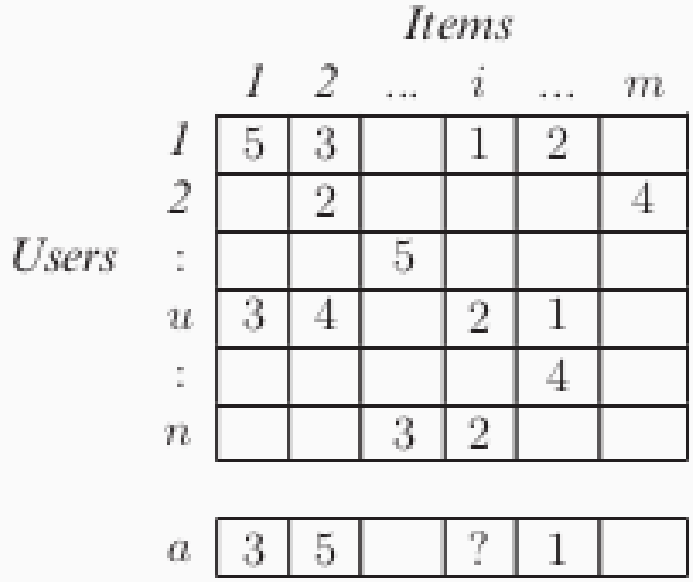
How to measure goodness of recommendation?

Recommendation Systemsare algorithms and techniques that try to predict what the user wants.

$$MAE = \frac{\sum_{u,i} |p_{u,i} - r_{u,i}|}{N}$$

$$RMSE = \sqrt{\frac{\sum_{u,i} |p_{u,i} - r_{u,i}|^2}{N}}$$

Recommendation Systems – problem formulation



Recommender Systems. Figure 1. User ratings matrix, where each cell $r_{u,i}$ corresponds to the rating of user u for item i. The task is to predict the missing rating $r_{a,i}$ for the active user a

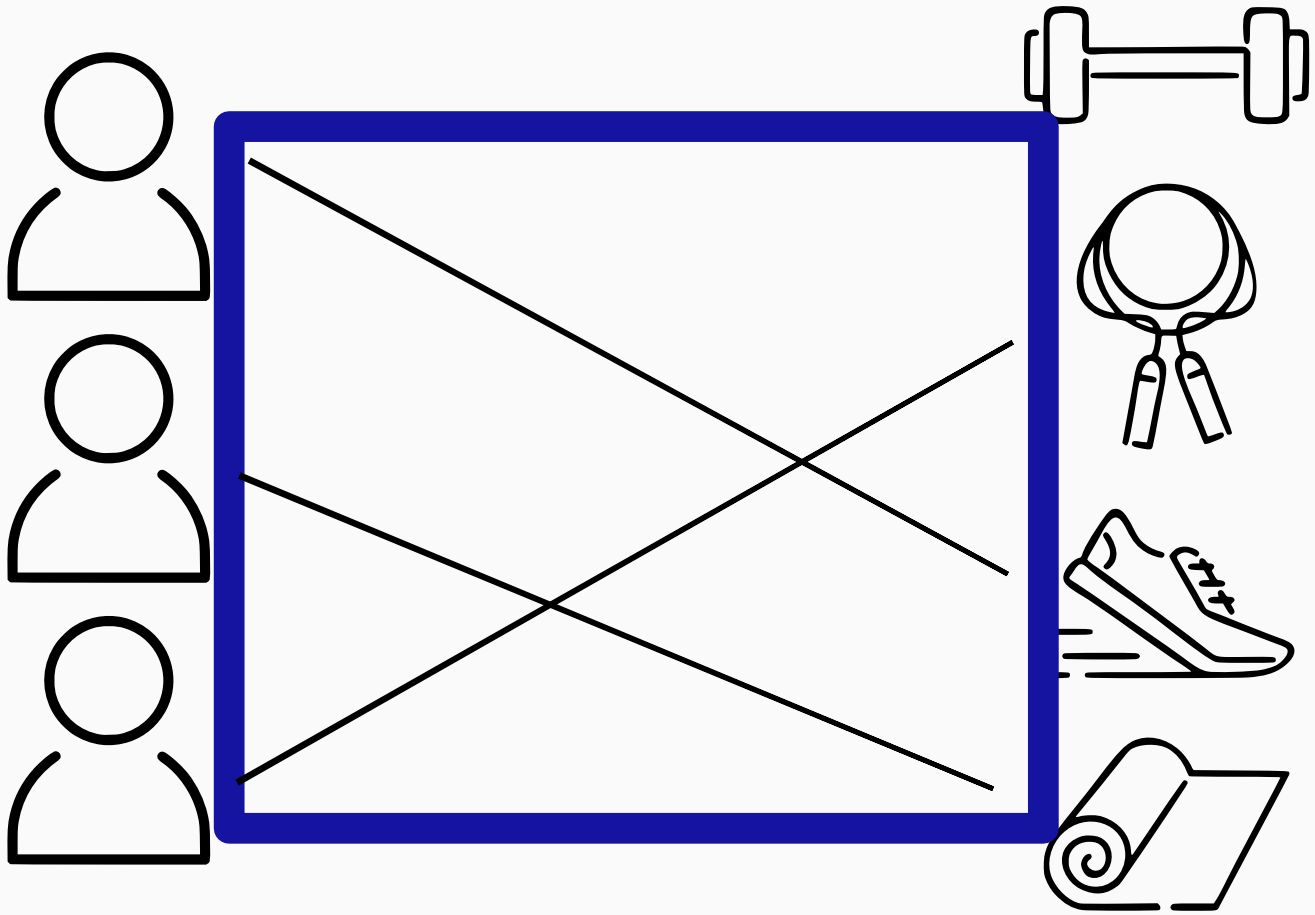
Recommendation Systems

There are different approaches:

- Collaborative Filtering – based on the historic choices of all users,
- Content-based recommending – based on user and product characteristics,
- Hybrid approaches.

I recommend to read `http://www.prem-melville.com/publications/recommender-systems-eml2010.pdf`

Collaborative Filtering



Collaborative Filtering

Collaborative filtering we divide into following sub-types:

- *neighborhood-based*, where we look for users similar to the analyzed one, and then, based on their selections.
- *model-based* approach, in which we estimate the parameters of the model characterizing users, and we make the recommendation based on the model.

Neighborhood-based collaborative filtering

We develop a prediction according to the following scheme:

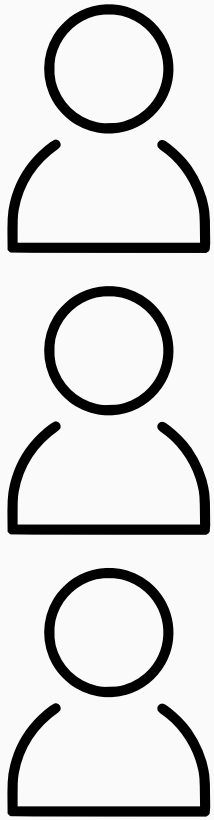
- We calculate the weights for each user according to their similarity (measured e.g. by the Pearson or Spearman or Kendall coefficient or the mean square error, etc.) to the selected user.
- We choose k nearest neighbors.
- We get the prediction as a weighted average of the neighbors' responses.

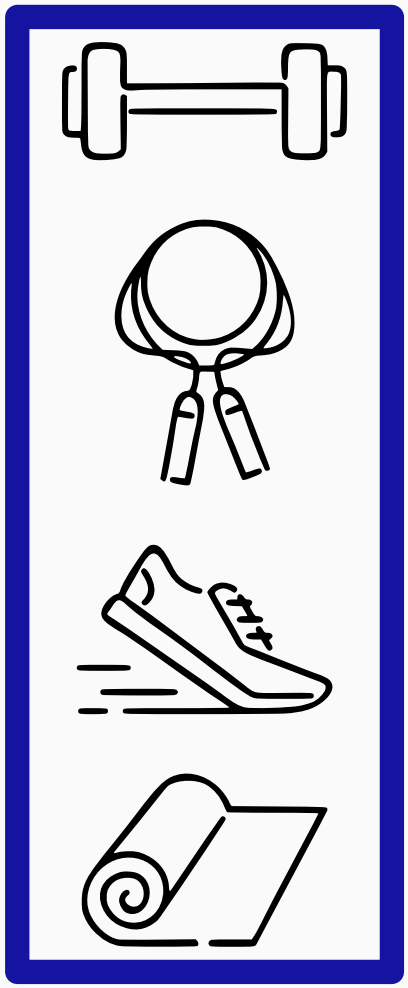
Model-based collaborative filtering

Statistical approaches:

- we try to build an (item-) classifier for each user.
- linear regression with regularization.
- SVM classifiers.
- matrix factorization methods.

Content-based recommending



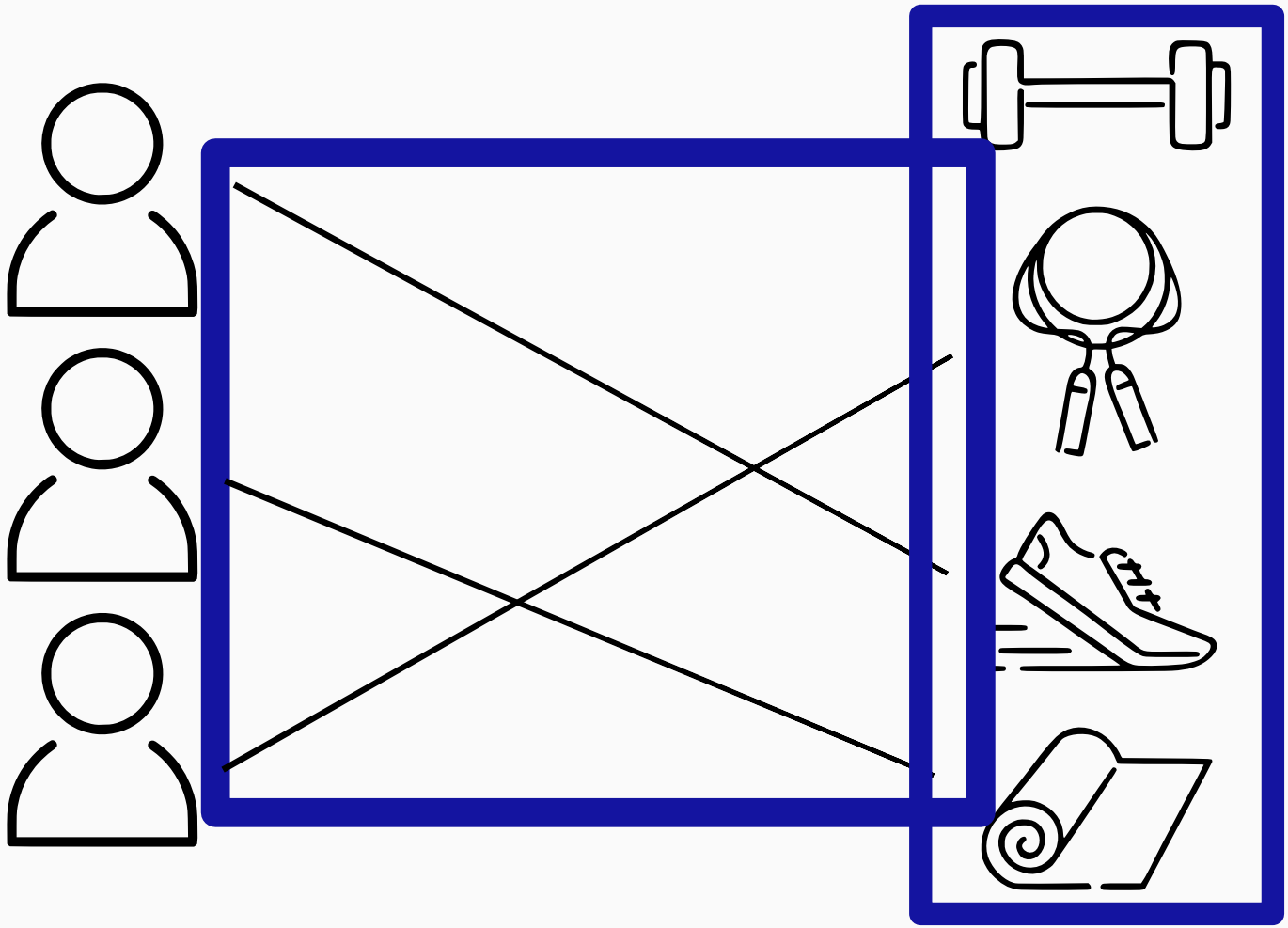


Content-based Recommending

Content-based Recommending methods are classification problems thus we can use

- naive Bayesian classificators,
- k-nearest neighbors algorithms,
- decision trees,
- neural networks,
- etc.

Hybrid methods



Challenges and Limitations

- Sparsity (very high item-to-user ratio),
- The Cold-Start Problem (aka first-rater-problem),
- Frauds.

Thank you for your attention!

Warsaw University of Technology

MSc program in Data Science has been developed
as a part of task 10 of the project
„NERW PW. Science - Education - Development - Cooperation”
co-funded by European Union from European Social Fund.