

Semestral work variants

RecSIS: enhanced course recommenders

RecSIS is a recent research project completed by Michal Hercik and Michal Medek. It tries to support students of MFF in better decisions on the organization of their study in MatFyz - especially in how to select the courses that would fit them the best. At the current stage, the basic functionality is already implemented, relatively broad datasets are ready, but the course recommendations are only rudimentary. You may be the person who's gonna change that.

Task: Implement an alternative course recommendation service for RecSIS
[the problem you all know]

Starting points:

- <https://github.com/michalhercik/RecSIS/tree/ndbi021/main> (repository)
- <https://acheron.ms.mff.cuni.cz:42050/> (running instance)
- https://cunicz-my.sharepoint.com/:v/g/personal/81411247_cuni_cz/EU_hcMQ1UdZGp-dalKCIPjMBdNFykbUNIJcSlwCf4k9dpw?e=im8lb9 (video instructions)

Main contact: Michal Hercik, hercik77@seznam.cz

Name(s)	Notes

Controlling multi-objective RS: initialization

In one direction of our research, we focused on controllable multi-objective recommender systems. We introduced a slider-based UI that enables users to set their current propensity towards receiving relevant, novel, diverse, etc recommendations. These values need to be initialized in some way, which we did based on the previous user interactions (preference elicitation or previously selected items). Nevertheless, based on the results of the user study, we appear to have overstated the relative importance of beyond-accuracy objectives at the expense of relevance. Can you learn from our mistakes and improve on our initial estimations?

Task: Construct a method for unbiased initialization of user's propensities towards individual objectives based on the preference elicitation.
[well-defined dataset with baselines available]

Starting points:

- General context: Accuracy and beyond-accuracy perspectives of controllable multi-objective recommender systems, <https://www.sciencedirect.com/science/article/pii/S0306457325002080>
- Paper with dataset to use & baselines: SM-RS 2.0: User-perceived Qualities of Single- and Multi-Objective Recommender Systems, <https://dl.acm.org/doi/10.1145/3754459> [Task 1]
- Dataset: <https://osf.io/wsakx>

Additional contact: Patrik Dokoupil, patrik.dokoupil@matfyz.cuni.cz

Name(s)	Notes

Enhance EasyStudy's Preference Elicitation with SAE

How can we maximize the quality of preference elicitation while minimizing its cost? That's a decades-long question, where you've already seen some early approaches to this issue. Nevertheless, SAE may bring some fresh air to this area. In particular, we may try to utilize their natural capability to provide a reasonable segmentation of the latent space - both in terms of item and user representations. Each user/item only activates a few neurons and their activation patterns may be clustered towards some broader segments (these can be labeled w.r.t. some available metadata, but are emerging directly from the observed interactions). As such, we may select candidates to maximize coverage of the topical space first, and then divide the next supplied items between those serving to confirm our prior beliefs about users and those aiming to explore previously unobserved parts of the latent space. This can be naturally combined with some 2D recommendation presentation & labeling (if, for example, you prefer to work in a small group).

Task: Enhance EasyStudy with a preference elicitation based on SAE representation [visual exploration available]

Starting points:

- SAE usage: SAGEA: Sparse Autoencoder-based Group Embeddings Aggregation for Fairness-Preserving Group Recommendations, <https://dl.acm.org/doi/abs/10.1145/3705328.3759322>
- SAE usage: From Knots to Knobs: Towards Steerable Collaborative Filtering Using Sparse Autoencoders [not yet published], <https://anonymous.4open.science/r/knots-to-knobs>

Additional contacts: Martin Spisak, masp430@gmail.com

Name(s)	Notes

Understanding the meaning of serendipity

In our past research, we collected a decent volume of user study data on how users perceive different objectives in recommender systems. Probably the most intriguing is the notion of **serendipity**, i.e., unexpected yet relevant recommendations. Nevertheless, the concept itself is still eluding our grasp, and it is hard to predict how the users gonna react to actual recommendations. This (and other) tasks are nicely wrapped in a resource paper together with a few simple baselines, so you can test whether you can beat us in serendipity estimation:-)

Task: Try to improve the prediction of how serendipitous the recommendations would appear.

[join the company of great scientists who failed in this task:-)]

Starting points:

- Paper with dataset to use: SM-RS 2.0: User-perceived Qualities of Single- and Multi-Objective Recommender Systems, <https://dl.acm.org/doi/10.1145/3754459> [Task 5.4]
- Dataset: <https://osf.io/wsakx>

Additional contact: Patrik Dokoupil, patrik.dokoupil@matfyz.cuni.cz

Name(s)	Notes

Enhance EasyStudy with another domain

Currently, EasyStudy can work with two pre-built domains: books and movies. Nevertheless, some user studies could benefit from different use-cases and thus, a broader set of datasets would be beneficial. In particular, we would like to focus on **the fashion** and **news** domains, but we are open to other suggestions. In both cases, your task would be to identify a suitable dataset, possibly define how to extend it to cover all our needs (including, e.g., metadata to display and visual representation), and incorporate it into EasyStudy. The fashion domain may require an additional baseline recommender based on visual similarity, while for the news domain, being able to serve up-to-date articles (e.g., based on some RSS feeds) would be of great interest.

Task: Expand EasyStudy with support for the fashion domain

Task: Expand EasyStudy with support for the news domain
[straightforward task]

Starting points:

- EasyStudy: <https://dl.acm.org/doi/10.1145/3604915.3610640>,
<https://github.com/pdokoupil/EasyStudy>

Main contact: Patrik Dokoupil, patrik.dokoupil@matfyz.cuni.cz

Name(s)	Notes