



Review report of a final thesis

Reviewer: Rodrigo Augusto da Silva Alves, Ph.D.
Student: Bc. Jan Šafařík
Thesis title: Visual evaluation of recommender systems
Branch / specialization: Knowledge Engineering
Created on: 23 May 2022

Evaluation criteria

1. Fulfillment of the assignment

- ▶ [1] assignment fulfilled
- [2] assignment fulfilled with minor objections
- [3] assignment fulfilled with major objections
- [4] assignment not fulfilled

In this master thesis, Jan Šafařík proposed a novel visualization and evaluation framework for recommender systems (named Repsys). Repsys is a platform that supports a comprehensive analysis that exceeds typical accuracy metrics. For instance, Repsys can reveal the diversity of the items in recommendation strategies. His master's thesis is pertinent to both the recommender systems research community and final users. It is also in line with the present endeavor of the European commission and the machine learning community to produce not just accurate but also explicable models.

2. Main written part 95/100 (A)

The text is easy to understand and well structured. I could only find feel typos that does not compromise the quality of the text.

3. Non-written part, attachments 100/100 (A)

The systems have a complex structure based on the best programming practices. The code is well written and easy to read. The author also chose the correct program languages, related to the recommend system community.

4. Evaluation of results, publication outputs and awards 100/100 (A)

The thesis has already result in a system that is (free) available in the author's online repository. The author also claim that he will submit it to a demo track of the ACM

Recommender Systems Conference. Particularly, I share the opinion that is a perfect venue for this work.

The overall evaluation

98_{/100} (A)

The thesis has a significant topic and the software is already freely available online. Considering the standards for a master's thesis, it is well-written and meticulously diagrammed. I emphasize that the figures and presentation of the results are of outstanding quality. It is particularly significant because this thesis is about visual evaluation.

Questions for the defense

Would be possible an extension repsys framework in a scenario of sequential recommendation?

Metrics in recommender systems, as cited in the work, are very heterogenous. For instance, one might recommend cross sell items instead of similar items. It is not clear how RepSys could be adapted for customized metrics. Could you explain (if/how) to fulfill this important requirement in a online setting (e.g. users creating their own metrics)?

Instructions

Fulfillment of the assignment

Assess whether the submitted FT defines the objectives sufficiently and in line with the assignment; whether the objectives are formulated correctly and fulfilled sufficiently. In the comment, specify the points of the assignment that have not been met, assess the severity, impact, and, if appropriate, also the cause of the deficiencies. If the assignment differs substantially from the standards for the FT or if the student has developed the FT beyond the assignment, describe the way it got reflected on the quality of the assignment's fulfilment and the way it affected your final evaluation.

Main written part

Evaluate whether the extent of the FT is adequate to its content and scope: are all the parts of the FT contentful and necessary? Next, consider whether the submitted FT is actually correct – are there factual errors or inaccuracies?

Evaluate the logical structure of the FT, the thematic flow between chapters and whether the text is comprehensible to the reader. Assess whether the formal notations in the FT are used correctly. Assess the typographic and language aspects of the FT, follow the Dean's Directive No. 52/2021, Art. 3.

Evaluate whether the relevant sources are properly used, quoted and cited. Verify that all quotes are properly distinguished from the results achieved in the FT, thus, that the citation ethics has not been violated and that the citations are complete and in accordance with citation practices and standards. Finally, evaluate whether the software and other copyrighted works have been used in accordance with their license terms.

Non-written part, attachments

Depending on the nature of the FT, comment on the non-written part of the thesis. For example: SW work – the overall quality of the program. Is the technology used (from the development to deployment) suitable and adequate? HW – functional sample. Evaluate the technology and tools used. Research and experimental work – repeatability of the experiment.

Evaluation of results, publication outputs and awards

Depending on the nature of the thesis, estimate whether the thesis results could be deployed in practice; alternatively, evaluate whether the results of the FT extend the already published/known results or whether they bring in completely new findings.

The overall evaluation

Summarize which of the aspects of the FT affected your grading process the most. The overall grade does not need to be an arithmetic mean (or other value) calculated from the evaluation in the previous criteria. Generally, a well-fulfilled assignment is assessed by grade A.