

Jury Member Report – Doctor of Philosophy thesis.

Name of Candidate: Evgeny Frolov

PhD Program: Computational and Data Science and Engineering

Title of Thesis: Low-rank models for recommender systems with limited preference information


Supervisor: Prof. Ivan Oseledets

Chair of PhD defense Jury: Prof. Andrzej Cichocki

Email: a.cichocki@skoltech.ru

Date of Thesis Defense: 19 September 2018

Name of the Reviewer: Marko Tkalčič

I confirm the absence of any conflict of interest	Signature:  Date: 16-08-2018
---------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------

The purpose of this report is to obtain an independent review from the members of PhD defense Jury before the thesis defense. The members of PhD defense Jury are asked to submit signed copy of the report at least 30 days prior the thesis defense. The Reviewers are asked to bring a copy of the completed report to the thesis defense and to discuss the contents of each report with each other before the thesis defense.

If the reviewers have any queries about the thesis which they wish to raise in advance, please contact the Chair of the Jury.

Reviewer's Report

--

Brief evaluation of the thesis quality and overall structure of the dissertation

In general, the thesis is of good quality. The language used in the thesis is generally clear. The candidate shows a good understanding of the subject and a good command of English language. He uses the terminology in a correct way. The research topic, low-rank approximation in recommender systems, is relevant and dealt in a coherent way. The main downsides of the chosen baselines, PureSVD and related approaches, are presented. As an answer to these downsides the candidate presents three methods for addressing different downsides. I recommend to improve the description of the problems addressed to emphasize the significance of the results.

The thesis is well structured. In the chapters 1-4, the candidate introduces the research topics, in the chapters 5-7, he presents the scientific contributions, in chapter 8 he presents a software framework. Finally, there is a conclusion chapter.

The formatting of the bibliography is inconsistent. Some entries start with the title and some with the authors. The DOI is missing in most of the bibliographic entries. I recommend to stick to a single bibliographic format (e.g. APA, IEEE, Chicago etc.).

Relevance of the topic of dissertation work to its actual content

The topic and the content of the thesis are coherent.

Relevance of the methods used in the dissertation

The methods used in the thesis are relevant to the domain of recommender systems. The novel methods proposed advance the performance of the state-of-the-art in the selected metrics. The chosen metrics do reflect the aspect of recommender systems that the proposed methods are supposed to address.

Scientific significance of the results obtained and their compliance with the international level and current state of the art

The candidate proposes three novel methods for addressing two main problems: (i) lack of usage of negative feedback in related work and (ii) lack of usage of side information in related work. The first model addresses problem (i), the second addresses problem (ii) while the third proposed model addresses both problems together.

The results show that the proposed methods advance the state-of-the-art and are hence significant. Furthermore, the results were published at two prestigious venues (a journal with a good impact factor and a selective conference), which confirms the significance of the results.

I would recommend to elaborate more on the downsides of existing work in order to justify and strengthen the scientific contributions.

Relevance of the obtained results to applications (if applicable)

The thesis does not address specifically applications. The experiments are done on offline data, which is in line with perspective taken during this research. Conducting online experiments in real application scenarios would be too big an effort that is out of scope.

Quality of publications

The candidate has published five publications. One is in a journal with a good impact factor (1.939), another at the RecSys 2016 conference (with a selective, 18% acceptance rate), a book chapter and two arXiv publications. The candidate has shown, especially with the journal paper and the RecSys paper, that he is qualified to present the research outcomes in a clear and persuasive way. For the two arXiv publications it would be useful to mention whether they are in the submission process to some other venues.

Summary of issues to be addressed before/during the thesis defense

The candidate presents the research topic in the chapters 1-4. In chapter 1 he provides an overview of the field and exposes the following challenges: cold start, missing values, usage of implicit feedback, evaluation and reproducibility aspects, real-time issues, context and features. I would recommend to conclude this chapter with specifying which of these challenges will be addressed in the remainder of the thesis.

In chapter 2 the candidate provides a thorough overview of factorization techniques, their advantages and limitations. Similarly, in chapter 3, the candidate gives an overview of tensor factorization techniques. I would recommend that each of these chapters concludes with the concrete downside that the thesis is addressing. Furthermore, for each research problem identified, the author should point to the relevant chapter (5/6/7) that addresses it. For example, the conclusion of chapter 3 is "A possible cure for this problem is to use TT/HT decomposition. In our opinion, this is a promising direction for further investigations." It is not entirely clear whether the author addresses this problem with the proposed method later on in the thesis or it is a proposal for future work based on an educated guess. After reading the next chapters the reader understands that chapters 2 and 3 do not expose the concrete issues addressed in the thesis. This is done in chapter 4. However, I would recommend to make it clearer.

In chapter 4, the candidate describes the problem of missing data, referred to as limited preference information. The author distinguishes between the local and global lack of preferences. However, the distinction is not very clear. Is the global a matter of global sparsity and local only user- or item-bound? Please, rephrase to make it clearer. The candidate argues that there is little related work that uses negative ratings. While this is true to a certain extent, there are techniques for eliciting negative feedback, such as the MinRating. I would recommend comparing active learning techniques for preference elicitation, e.g. M Elahi, F Ricci, N Rubens. A survey of active learning in collaborative filtering recommender systems. - Computer Science Review, 2016.

In Sec. 4.4 the candidate lays out the requirements for improvements over state-of-the-art methods, described in Sects 4.3.1 and 4.3.2. However, in these two sections, the author has briefly described the existing methods, but has failed to clearly identify the downsides of these methods. For example, stating that those methods do not use SVD-based techniques is not a valid argument for introducing SVD. The

remainder of the section 4.4 is similarly unconvincing. For example, the author states "Moreover, these methods focus on a particular subproblem." Why is this a downside? Please elaborate more on the downsides of existing work in order to justify and strengthen your scientific contributions.

In chapter 5 the candidate presents a tensor-based model that uses the "users, who dislike that item, do like these items instead" approach to take advantage of negative ratings as well as positive. The problem is well formulated. The limitations of existing solutions are laid out. The proposed approach moves from User x Item -> Rating to User x Item x Rating -> RelevanceScore, where RelevanceScore is the likelihood of observing a specific UxIxR triplet. He used the Movielens dataset and a couple of baseline algorithms. It is not clear why there is no FM among these as at the beginning of the chapter he states "This type of relations can be modelled with several methods, such as Factorization Machines [67] or other context-aware methods [42]." There are details missing on how baselines were implemented. E.g. how were negative ratings used in BPR and WRMF? How were Movielens data converted to pairwise (for BPR)? The candidate explains well why NDCG is not a good metric and why it had to be replaced.

In chapter 6, the candidate presents a hybrid model. He starts with the statement "To the best of our knowledge there were no attempts to build a hybrid SVD-based approach where interaction data and side information would be factorized jointly in a seamless way". Could you elaborate a bit more? Especially what do you mean by "seamless"? There are factorization approaches where additional data is injected as additional (not really latent) factors, e.g. Fernández-Tobías, I., Braunhofer, M., Elahi, M., Ricci, F., & Cantador, I. (2016). Alleviating the new user problem in collaborative filtering by exploiting personality information. *User Modeling and User-Adapted Interaction*, 26(2), 1–35. <https://doi.org/10.1007/s11257-016-9172-z>. The candidate introduces a matrix S that contains relations between items based on side information. I would recommend to introduce an example of side information in order to make figure 6.2 more understandable. At this point it is still not clear which side information is used and why that behaviour is observed. The candidate then explains how to compute the matrix S and how to use it to generate predicted ratings. In three experiments the candidate shows the benefits of the proposed hybrid model. The choice of the baselines is adequate (especially having FM). I would recommend to report also the computation times and adding a discussion on the trade-off between performance and accuracy.

In the seventh chapter, the candidate presents a model that combines the previous two models. The motivation is sound and the algebraic formulation as well. However, in the evaluation, the candidate assumes "User feedback is considered to be positive if the rating value is equal or above 4 (including 4.5 if it is present in data) with the highest rating being 5." This depends on the dataset. For example, Movielens does not instruct the users on what the stars mean. In user studies it is usually very clear what each rating means, but not in Movielens, so each user interprets this scale at his/her own will. For example, in Slovenian schools the rating system is from 1-5, where only 1 is negative and 2-5 are various degrees of positive. One would hence speculate that Slovenian users might treat 1.5 as the threshold between positive and negative. I would recommend to discuss what happens with the proposed model if the threshold between positive and negative ratings is shifted.

In chapter 8, the candidate presents the software framework he devised for implementing the models presented in the thesis. I would recommend to include a comparison between the main existing frameworks and the presented one. I believe a comparison table would suffice.

Furthermore, I would recommend to proofread the thesis. There are some language issues that should be addressed. A couple of examples:

- "Moreover, SVD algorithm has stable..." -> "Moreover, the PureSVD ..."

- "Inability to take that into..." -> "The inability to take that into..."
- "Moreover, PureSVD offers a number practical advantages..." -> "a number of practical"
- "6.1 Understanding limitations of SVD" -> "6.1 Understanding the limitations of SVD"
- "Many various software libraries" -> "Many software libraries" or "Various software libraries"

Provisional Recommendation

I recommend that the candidate should defend the thesis by means of a formal thesis defense

I recommend that the candidate should defend the thesis by means of a formal thesis defense only after appropriate changes would be introduced in candidate's thesis according to the recommendations of the present report

The thesis is not acceptable and I recommend that the candidate be exempt from the formal thesis defense