The 58th Annual Meeting of the Association for Computational Linguistics

ACL 2020

Review Report

Title: Heterogeneous User-Interest Transfer Learning for Cold-Start News Recommendation
Authors: Guangneng Hu, Yu Zhang and Qiang Yang
Status: Rejected

MetaReview

Comments: This work tackles the cold-start problem in recommender systems for new users - a challenging problem within research and industry. The reviewers liked the application of transfer learning in this problem, but believed that the gains in performance are marginal compared to other transfer learning methods, and although the experiments are thorough, the paper's contributions are marginal.

Review #1

What is this paper about, what contributions does it make, what are the main strengths and weaknesses?

This paper aims to address the cold start problem for news recommendation systems. Specifically, they study the cold start problem of new users who are frequent consumers of one category of news (e.g., sports) but new users in another category (e.g., sports). This paper proposes a model that uses a translator-based transfer-learning strategy for learning a representation mapping from a source domain to a target domain. The use of transfer learning is not new but the application for this domain provides a good improvement over existing recommender systems that are not targeted at cold-start users. The experiments seem fairly robust although many of the baselines are not recommender systems targeted at cold-start users. Overall, this is a nice paper that is well-written and enjoyable to read.

The strengths of this paper are:

- The cold start problem is an important problem in recommender systems, not just news recommenders. There is good potential for extending this approach to other domains of items.
- The proposed TrNews architecture is described in sufficient detail and offers a good improvement over a good range of recommendation algorithms from a popularity-based one to more advanced ones based on deep neural networks.
- There is a nice analysis of the different parameters of this model, such as the number of shared users, consumption history length, shared word embeddings, etc.
- The paper is well-written and easy to follow.

The weaknesses of this paper are:

- Transfer learning itself is not new although this specific application to news item recommendation for cold-start user may be. The improvement over existing transfer learning techniques is quite marginal and less than 1% in most cases.
- There is a good selection of baseline recommendation algorithms but they are not designed to target the cold-start user recommendation problem. It is expected

that these algorithms will not perform well for cold-start users so a fairer evaluation would be compared against similar algorithms for cold-start recommendation, such as SSCDR (Kang et al., 2019) that was cited and [1].

I wish the authors all the best as they proceed with this interesting work.

[1] Li, Jingjing, Mengmeng Jing, Ke Lu, Lei Zhu, Yang Yang, and Zi Huang. "From zero-shot learning to cold-start recommendation." In Proceedings of the AAAI Conference on Artificial Intelligence, vol. 33, pp. 4189-4196. 2019.

Reasons to accept

There is an interesting use of transfer learning to solve the cold-start recommendation problem. Please refer to my detailed comments.

Reasons to reject

Some of the recommendation algorithms for cold-start users are not included in the evaluation. Please refer to my detailed comments.

Overall Recommendation: 3.5

Review #2

What is this paper about, what contributions does it make, what are the main strengths and weaknesses?

Overall, this paper mainly uses transfer learning to address the cold start problem of news recommendation. It analyzes the user browsing behaviors on both source and target domain for news recommendation. A translator is built for the purpose of transfer learning. It makes some contributions to the area, but in a limited manner.

Strength:

First, two base networks are used to learn the information of the source domain and the target domain. Through an autoencoder, the user information of the source domain is transferred to the target domain. The experiment is relatively sufficient. From the experimental results, it has a certain effect on cold start problem, which is better than baselines. Although the base network for target / source domain is similar to the network structure mentioned in (Wu et al., 2019a), an innovative translator built with a variant of autoencoder realizes cross domain cold start.

Weakness: However, this autoencoder is actually more like a MLP, which is similar to the transfer-learning structure of EMCDR. According to the description in this paper, when comparing the transfer-learning methods such as EMCDR and DDTCDR, it only replaces the translator in TrNews model with the transfer-learning strategies of these methods, which is not directly compared with these models, so lacks some comparative experiments.

It seems I do not see the temporal factor incorporated into the model, which is very important for user-interest modeling. Users' interests are always changing and recent interests might take higher weights instead of equal ones to the past.

Wish authors good luck!

Reasons to accept

The application of transfer learning on cold-start news recommendation.

Reasons to reject

Performance evaluation with comparison to state-of-the-art baselines is lack.

Overall Recommendation: 3.5

Review #3

What is this paper about, what contributions does it make, what are the main strengths and weaknesses?

This work tackles the cold-start problem in recommender systems for new users - a challenging problem within research and industry. Given a user who has read articles on a source domain S, the goal is to recommend him/her articles in a new domain T, for which there is no prior knowledge about the user's preferences. The authors propose a neural architecture with three components for users in a given training set: (a) a source network, which learns a scoring function of the user and the news items based on the articles that he/she has read on S in the past; (b) a target network, which learns a scoring function over the domain T in a similar fashion; (c) a translator that learns to map the scores of S into those of T. The three models are trained on the fly and then, given a user in the test set with prior knowledge over S, the translator converts his/her representation in S into the representation in T, so that a score can be assigned to every article in the domain T. The authors compare their approach against several competing methods, showcasing clear cut gains in performance. They also compare against competing transfer learning approaches.

The paper is well-written, the methods are well-supported and there is a detailed experimentation and comparison against existing (a) recommender systems and (b) transfer learning approaches. My major concern is that, despite demonstrating the effectiveness of the proposed approach against competing (a) recommender systems, the same does not hold for the comparison against (b) transfer learning approaches, where the performance looks quite similar. Also, it is not straight-forward how the proposed model could be applied when dealing with multiple source/target domains. Finally, a few mistakes -- that can easily be corrected -- are scattered throughout the document and some details of the modelling decisions need some further clarification (see my comments later in this review).

Reasons to accept

- The proposed method is well-supported.
- There is a detailed evaluation and comparison against competing methods and transfer learning approaches.
- There are clear gains in performance against competing recommender systems.

Reasons to reject

- There are only marginal gains in performance against existing transfer learning methods (i.e., different translators), thus questioning the contribution in this aspect of the paper.
- It is not very clear how the proposed method could be evaluated on multiple domains - would we need one network for each domain?

Overall Recommendation: 4

Questions for the Authors(s)

- General: There are only two domains considered in this work (a source and a target). How could the proposed model handle the transfer learning task across multiple domains (either as sources or as targets)?
- Section 3.2/News encoder: How are the words in the news items represented? (i.e., how were the word embeddings generated initially? What type of word embeddings did you use?)

- Section 4.1/Dataset: Which are the two domains? (Finance/Tech/Science/Sports/…?)
- Section 4.1/Evaluation: Why sample only 99 negative examples? How robust is the performance of the tested models across several runs with different samples? Why not considering all of the news items that have not been read by a user as negative examples?

Typos, Grammar, and Style

- Scan throughout the document and correct the splitting of words in two lines (e.g., differen-t, user-s, tw-o, etc.)
- Section 4: There is a space between “Cheetah Mobile\footnote{…}” and the “,” following it
- Section 5: “We also shows”, “In future works”: remove final “s”