Sequence Recommender with Item and User Embeddings

Lara Thompson
Principle Data Scientist @ Salesforce
lara.thompson@salesforce.com

April 6, 2023

Abstract

Many recommendations rely on the sequence of interactions between users and items. User tastes and abilities evolve in time; for items that rely on a skill level, such as bouldering or children's reading, a good sequence of items (boulder problems or books) can be key in skill development. In this paper, we explore common models from natural language processing applied to a skill progression based space of times: the CBOW model of word2vec (Mikolov et al., 2013), a simple recurrent model that accounts for sequence ordering, and a transformer model that attentively accounts for sequence ordering. The recommendation objective can be either next item or the sequence of items that best lead up to a final "goal" item. User and item features generally improve the accuracy of the sequence model. A transformer model that dually trains item and user embeddings and leverages item features can predict next item with accuracy nearing 25% for BC boulder problems (compared to using collaborative filtering).

1 Introduction

The consensus is that collaborative filtering will generally outperform content-based recommendation [1]. However, it is only applicable when usage data is available. Collaborative filtering suffers from the cold start problem: new items that have not been consumed before cannot be recommended. Additionally, items that are only of interest to a niche audience are more difficult to recommend because usage data is scarce

2 Prior Literature

This section can make extensive use of your lit review prose. (Arora et al., 2017)

3 Data

Likely to be very detailed if the datasets are new or unfamiliar to the community, or if familiar datasets are being used in new ways. Includes prior work on them, statistics, and a collection protocol.

4 Model

Flesh out your own approach, perhaps amplifying themes from the 'Prior lit' section.

5 Methods

The experimental approach, including descriptions of metrics, baseline models, etc. Details about hyperparameters, optimization choices, etc., are probably best given in appendices, unless they are central to the arguments.

Explicitly define the metrics, even the common ones (or at least reference them). Be clear about how the data is split for assessment.

6 Results

A no-nonsense report of what happened.

7 Analysis

Discussion of what the results mean, what they don't mean, where they can be improved, etc. These sections vary a lot depending on the nature of the paper. (For papers reporting on experiments with multiple datasets, it can be good to repeats Methods/Results/Analysis in separate (sub)sections for each dataset.)

8 Conclusion

Quickly summarize what the paper did, and then chart out possible future directions that anyone might pursue.

Known Project Limitations

For this section, imagine that your reader is a well-intentioned NLP practitioner who is seeking to make use of your data, models, or findings as part of a separate scholarly project, deployed system, or some other kind of real-world intervention. What should such a person know about your work? Especially important here are limitations and biases that might affect this person, their findings, their experiment participants, or the users of their product or service. The idea is that what you say here will be taken into consideration but this well-intentioned user, leading to better outcomes for everyone.

Authorship Statement

I worked on this project solo including the literature review, the iterative implementation and experimentation, result gathering and final writing up. I had many fruitful conversations with Alistair Fraser, also taking the course, and Michael Nathe about my design choices and result analysis.

References

Sanjeev Arora, Yingyu Liang, and Tengyu Ma. 2017. A simple but tough-to-beat baseline for sentence embeddings. *5th International Conference on Learning Representations, ICLR 2017*.

Tomas Mikolov, Kai Chen, Gregory S. Corrado, and Jeffrey Dean. 2013. Efficient estimation of word representations in vector space. In *International Conference on Learning Representations*.