The 6th International Conference on Interdisciplinary research on Computer science, Psychology, and Education (ICICPE' 2022)
December 26-28, 2022. Pattaya, Thailand.

Text-only Keyword Recommendation: No Longer Have Cold-Start Problem

Jungseob Lee[1], Hyeonseok Moon[1], JeongBae Park[2*]

[1] Dept. of Computer Science and Engineering, Korea University
[2] Human-inspired Artificial Intelligence Research, Korea University
{omanma1928, glee889, insmile}@korea.ac.kr

Abstract. The cold start-problem is a chronic problem in terms of providing a recommendation system. For this reason, we propose a recommendation system based on base properties that can alleviate the cold start problem. In order to verify the proposed methodology, a qualitative evaluation is conducted based on the Korean movie review dataset collected by ourselves.

Keywords: Recommendation system, Cold-start problem, Text recommendation

1 Introduction

The recommendation system is a technology that is practically helpful to the end-user because it can provide personalized items through the user's history [1].

However, since collaborative filtering data operates based on a user's individual ranking history, a personalized data set in the service is necessary, and public data or external data cannot be used. This is a cold-start problem that, when modeling a specific system based on collaborative filtering, not only must a pre-established service exist to obtain user history data, but also the design is possible after securing many users.

Therefore, this study proposes a base-property-based recommendation system that is not affected by the cold start problem by applying a text content-based filtering method. Since this is a text search system that reflects the latent vector that reflects the user's preference that the ranking-based search system cannot solve, it has the advantage of being free from the cold start problem.

2 Proposed Method

Feature extraction extracts feature using a user's score and text pair. In this study, the key to feature extraction is to extract keywords that can describe items in a sentence or indicate a review score. There are two methods for keyword extraction: dynamic extraction, which is a method of dynamically extracting keywords using KeyBERT [2] among deep learning models, and static extraction, which extracts keywords by defining them in advance to fit the domain. Filter texts under the hyperparameter review score threshold and sentence length, extracted text based on the keyword is extracted.

The extracted keyword is not a simple frequency counting but an item index that reflects the weight of the review score and the number of appearances of each keyword to reflect the review score.

* Corresponding Author

The table search extracts the keywords of the entered sentences and searches for the keyword with the highest score among the generated keywords in the counting matrix. The top-k rankings are searched by sorting non-zero occurrence keywords in descending order.

3 Experiments

Table 1. Recommendation performance in Naver Movie reviews.

Model	Acc@3
TF-IDF	24.6
BM25	15.0
Key-rec (ours)	**41.7**

Experimental results are rounded to two decimal places. Our proposed method shows that it is possible to build a text recommendation system that overwhelms the existing search-based algorithms.

4 Conclusion

Our study aims to solve the chronic problem of the recommendation system, cold start. In addition, it was possible to build an efficient recommendation system not related to cold start through text and a table that reflects the ranking score.

5 Acknowledgement

This work was supported by Institute of Information & communications Technology Planning & Evaluation(IITP) grant funded by the Korea government(MSIT) (No. 2020-0-00368, A Neural-Symbolic Model for Knowledge Acquisition and Inference Techniques). This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2022-2018-0-01405) supervised by the IITP(Institute for Information & Communications Technology Planning & Evaluation).

References

1. P. Kumar and R. S. Thakur, "Recommendation system techniques and related issues: a survey," International Journal of Information Technology, Vol. 10, No. 4, pp. 495–501, 2018
2. M. Grootendorst, "Keybert: Minimal keyword extraction with bert," Internet]. Available: https://maartengr. github. io/KeyBERT/index. html, 2020.