Only for You: Relevant Board Game Suggestions from BoardGameGeek.com

Sanni Isokirmo[1], Ahti Katiska[1], and Miika Mäki[1]

[1]*University of Helsinki*

September 22, 2025 4:11pm

Introduction

Play—coming together to engage in a common recreational activity—is a human universal [1] and essential part in human development [2, 3]. Board games are a popular and accessible form of play that bring family, friends and strangers alike together, and promote well-being across the life-span [4, 5].

But what if a group of people have outplayed the games they own and would like to find some new ones? There are at least 150.000 different board games out there [6] so finding a good match can be a time consuming process. A service that would suggest new board games based on the individual taste could come in handy on such occasions.

This technical report will outline the steps we made to create a recommendation system for new board games based on the user's ratings in BoardGameGeek.com (BGG) database as well as the technical details of setting up a webpage *LINK TO WEBSITE* through which users can find recommendations for new board games. To achieve this, we used non-negative matrix factorization (NMF) with a selected users' review scores to create tailored suggestions for new board games.

Data

We chose the biggest board game ratings database BoardGameGeek.com to fetch user ratings and board game metadata, such as playing mechanics and category. As it is openly available, it having a data access was simple and there were no data privacy issues to be addressed.

Since there are 2.7 million users and 150.000 boardgames in BGG [6, 7], fetching all of them would not be feasible with the API interface, at least within this project. Since many users have left no or few reviews, the data would be too sparse if we had taken a random sample. To get a compromise between training matrix sparsity, time constraints and selection bias, we decided to choose x guilds (online discussion groups) that represent different geographic regions, target audiences (teens, parents, seniors), genres as well as general and special interest groups. This way we could get X reviews from Y reviewers about Z games. The sampling strategy did bias

our distribution towards active users, but also allows for reliable training. Users with only few reviews would make the estimation computationally intensive and possibly unreliable.

We decided to load our data via the BGG XML API [8]. While the data is well preprocessed and clean, getting the data was no simple task.

First, the API has rate limits that we had to find out by trial and error. We used a base delay of .75 seconds and if the rate was limited, increased the delay exponentially and waited for the maximum of one minute. The maximum batch size for games was 20 and it took some time to find out why the game metadata coverage was so low on larger test runs.

Second, since the game metadata fetching was unreliable even with the exponential backoff time, we used a game metadata cachefile to make sure that if a game metadata had been fetched successfully once, it was stored on the project home directory.

Finally, we included only users with at least 25 reviews and fetched the most popular games first to avoid training data matrix sparsity.

Methods

Using review scores is the single best way of generating recommendations [9]. It seems counterintutive as one would think that categories, genres, sales or other available metadata would be equally helpful. Based on this empirical finding and the scope of the project, we chose to focus solely on the review scores.

Results

Discussion

We created a board game recommendation webpage based on a sample of 200.000 board game reviews from 1000 active BoarGameGeek.com users to help BoardGameGeek users—or anyone—find new boardgames that match their taste.

If this project was expanded, we would use the whole BGG database as a training data

References

[1] D. E. Brown, "Human Universals, Human Nature & Human Culture," *Daedalus*, vol. 133, no. 4, pp. 47–54, 2004, ISSN: 0011-5266. JSTOR: 20027944. Accessed: Sep. 18, 2025. [Online]. Available: https://www.jstor.org/stable/20027944.

[2] P. K. Smith, "Play: Types and Functions in Human Development," in *Origins of the Social Mind: Evolutionary Psychology and Child Development*, New York, NY, US: The Guilford Press, 2005, pp. 271–291, ISBN: 978-1-59385-103-3.

[3] A. D. Pellegrini, *The Role of Play in Human Development*. Oxford University Press, 2009, 289 pp., ISBN: 978-0-19-536732-4. Google Books: yO3zr_0iFYIC.

[4] L. Dell'Angela, A. Zaharia, A. Lobel, O. Vico Begara, D. Sander, and A. C. Samson, "Board Games on Emotional Competences for School-Age Children," *Games for Health Journal*, vol. 9, no. 3, pp. 187–196, Jun. 2020, ISSN: 2161-783X. DOI: 10.1089/g4h.2019.0050. Accessed: Sep. 18, 2025. [Online]. Available: https://www.liebertpub.com/doi/full/10.1089/g4h.2019.0050.

[5] S. Solway, "Wellness through board games," M.S. thesis, The Chicago School of Professional Psychology, 2011.

[6] WordsRated. "Board Games Statistics – WordsRated," Accessed: Sep. 18, 2025. [Online]. Available: https://wordsrated.com/board-games-statistics/.

[7] M. Didymus-True. "BoardGameGeek's support drive numbers fall for third year running - ," Accessed: Sep. 18, 2025. [Online]. Available: https://boardgamewire.com/index.php/2024/01/09/boardgamegeeks-support-drive-numbers-fall-for-third-year-running/.

[8] BBG. "BGG XML API2," BoardGameGeek, Accessed: Sep. 19, 2025. [Online]. Available: https://boardgamegeek.com/wiki/page/BGG_XML_API2.

[9] D. Epstein, *Range: Why Generalists Triumph in a Specialized World*. New York: Riverhead Books, 2021, 368 pp., ISBN: 978-0-7352-1450-7.

Appendix