

PROPOSAL FOR A GAME RECOMMENDATION SYSTEM BASED ON REVIEW KEYWORDS

JEONGMIN SEO, HYUNSUNG YOO, WOOSUK JUNG, WONKEUN LEE
AND MINJUNG KWAK*

Department of Industrial and Information Systems Engineering
Soongsil University

369 Sangdo-Ro, Dongjak-Gu, Seoul 06978, Korea

{ jm661155; cap2885; lwg1421; yhsung42 }@soongsil.ac.kr; *Corresponding author: mkwak@ssu.ac.kr

Received October 2022; accepted January 2023

ABSTRACT. *With the growth of e-commerce market, personalized recommendation systems have emerged as a critical success factor in various industries. In the video game industry, previous recommendation systems were intended for existing users and required historical data on user purchases and ratings. In this study, a game recommendation system based on customer reviews was developed. It makes it possible to serve new users even when the data about them are limited. In the proposed system, games are represented in the form of keywords extracted from customer reviews. Recommendations are made based on keyword similarity. The KeyBERT and Word2vec techniques are applied for the natural-language processing of customer reviews.*

Keywords: Game recommendation, Online review, Natural-language processing, KeyBERT, Keyword extraction

1. **Introduction.** Making personalized recommendations of products and services has become essential for success in numerous e-commerce industries. Various recommendation systems and engines have been developed, and their effectiveness has been proven empirically through many cases in the market. For instance, the system recommended approximately 70% of the video contents consumed on Netflix. In the case of Amazon, 35% of all sales come from recommended products [1].

This study developed a personalized recommendation system for the video game industry, one of the biggest entertainment industries that generates more than \$159 billion in revenue per year as of 2020 [2]. The video game market demands a good recommendation system to suggest a curated list of games for individual users that they may find interesting and want to play. Research on game recommendation systems has increased in recent years to serve their needs. However, most recommendation systems in the video game market adopt either user-based or item-based collaborative filtering intended for existing users and require cumulative data about the users who receive recommendations [3,4].

These filtering methods can only work if there is sufficient information about the user receiving the recommendation (e.g., purchase history and rating history) [3-6]. User-based collaborative filtering can work when there are sufficient data about the experiences of existing users; the more the data, the higher the accuracy. Item-based filtering algorithms also recommend each item through data from users who purchased similar products.

In this study, a game recommendation system based on customer reviews was developed. The system is intended for new users and can operate even if there are few data about the users. By conducting natural-language processing (NLP) of customer reviews, the proposed system extracts a set of keywords and key phrases corresponding to individual

games. When a new customer signs up, the system requires the customer to select a keyword describing the customer's tastes or interests; then, it recommends games relevant to the keyword.

The rest of this article is organized as follows. In Section 2, the procedure of the proposed review recommendation system is described. In Section 3, a case study is described. Section 4 concludes the paper and provides future research directions.

2. Method. Figure 1 shows the procedure for building the proposed game recommendation system based on review keywords. The procedure includes two parallel steps: 1) a step for extracting keywords and key phrases of a game using the KeyBERT algorithm and 2) a step for building a keyword-similarity calculator using the Word2vec algorithm.

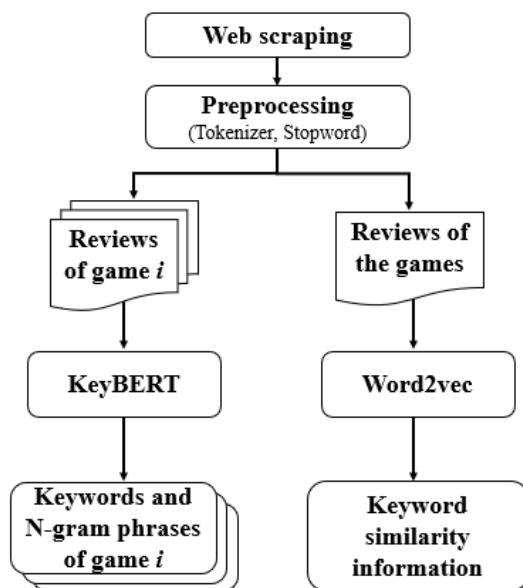


FIGURE 1. Procedure for building a game recommendation system

In this study, KeyBERT is used to summarize the reviews of a game in the form of keywords and key phrases. KeyBERT is a keyword extraction technique based on the bidirectional encoder representations from transformers (BERT) model developed by Google [7]. KeyBERT leverages the BERT model to identify keywords and key phrases that are most similar to a document [8]. In recent studies, it was reported that KeyBERT can outperform other keyword extraction techniques [9,10]. Because it considers the context of reviews, the keywords and key phrases extracted by KeyBERT better represent the reviews of a game [11].

Each keyword can have similar keywords usually positioned in a similar context in reviews. To identify such similar keywords of a given keyword, the Word2vec algorithm was adopted in this study. Word2vec is an NLP technique for word embedding and is useful for identifying semantic similarity between keywords [12]. When a particular word is entered, it is possible to measure the distance of the word from other words to identify words of higher similarity.

Figure 2 represents the procedure for operating the proposed recommendation system. When the user selects a particular keyword, the system returns a list of recommended games with the selected keyword and its similar keywords.

3. Main Results. The proposed system was demonstrated using 7487 reviews of 100 games. The reviews were scraped from Metacritic.com (Figure 3), a website aggregating user reviews in several entertainment genres, including film, music, and games.

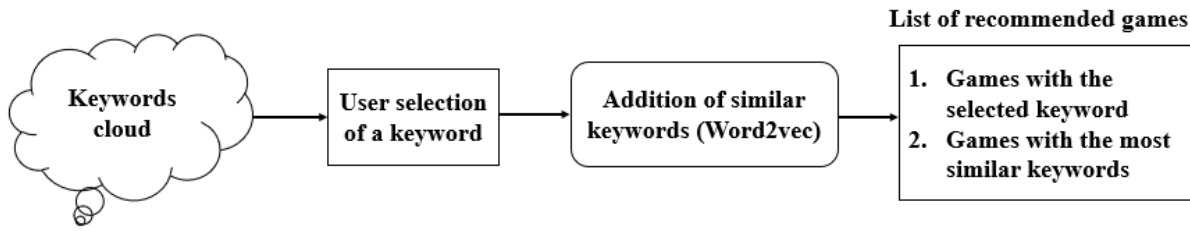


FIGURE 2. Procedure of operating the game recommendation system

FIGURE 3. Example: User reviews of the metacritic website (<https://www.metacritic.com>)

For each game, five three-word phrases were extracted using the KeyBERT algorithm. The extracted keywords representing the 100 games consisted of 582 different words. Table 1 shows a part of the results about different games.

The similarities between keywords were calculated by applying the Word2vec technique to the entire set of game reviews. Table 2 shows a part of the results. For instance, “soundtrack” is 95% similar to “atmosphere” and 94% similar to “art” throughout the review. In addition, “zombie” showed a 90% similarity to “survival” and 89% to “stealth”.

The game recommendation is based on the main keyword that the user selects. The system first recommends the games with the main keyword and then recommends the games with the most similar keywords. For example, if a user chooses the word “storyline”, the system recommends the games with the keyword “storyline” and the word “soundtrack” that is most similar to “storyline”, as shown in Figure 4. For instance, the system recommended “Final Fantasy XIV: Endwalker”, which includes “storyline” and “soundtrack” in the keywords, as shown in Table 1.

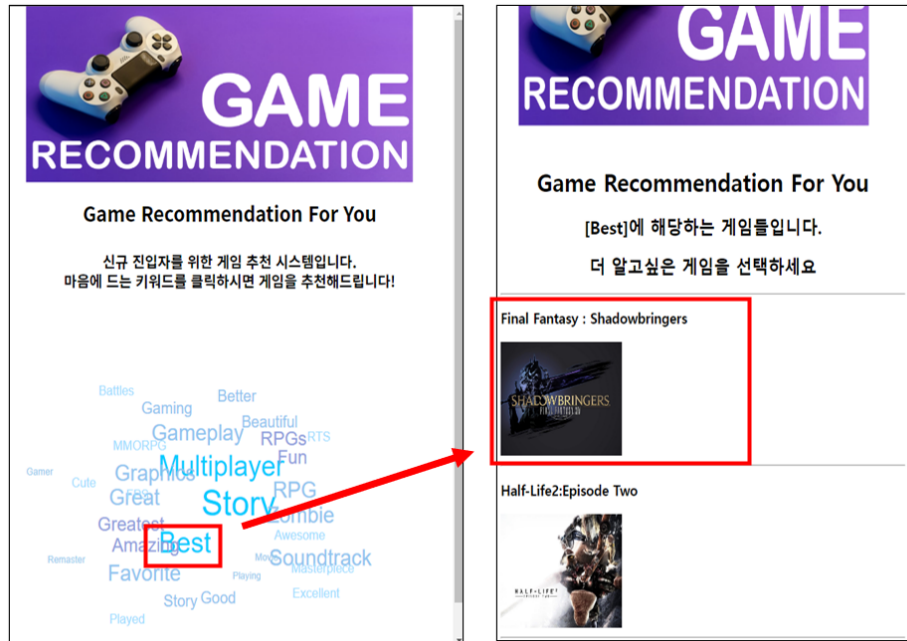


FIGURE 5. Prototype website

TABLE 3. Comparison of extracted keywords with the tags in steam

Keyword	Multiplayer	Soundtrack	RPG	Horror	Zombie
Number of games with the keyword	24	9	27	5	9
Number of games not listed in Steam	5	0	0	0	0
Number of games with the same keyword in Steam	17	8	25	5	9
Number of games without the keyword in Steam	2	1	2	0	0
Keyword similarity	89.47%	88.89%	92.59%	100%	100%

with manually assigned tags. In this study, tags posted by game producers and users in Steam (<https://store.steampowered.com>), one of the most popular game platforms, were used for the comparison. As shown in Table 3, the keyword similarity between the proposed system and Steam was calculated for five different keywords. For example, 24 games have the keyword “Multiplayer” according to the proposed system, and 17 games among them have the same keyword (tag) “Multiplayer” in Steam, while five games are not listed on the Steam website; thus, the similarity is $17 / (24 - 5) \times 100 = 89.47\%$. On average, the keyword similarity is 92.75%, which implies that the keywords extracted by the proposed system reflect the characteristics and strengths of a game well.

4. Conclusions. The review-based game recommendation system proposed is meaningful because it can be applied to new users. Unlike the existing recommendation system based on user-based filtering or item-based filtering, the system requires little information about the users. The user is only required to select a keyword that is most interesting to the user, and the system returns a list of games represented by the keyword. Keywords for each game are determined based on the experiences of game users. A set of keywords is automatically assigned to individual games based on customer reviews and used for recommending games.

One limitation of this study is that the proposed system needs more validation of the effectiveness of the recommendations. The quality of recommendations from a customer

viewpoint should be measured, and the performance of the proposed system should be compared with those of alternatives.

Acknowledgment. The authors gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

REFERENCES

- [1] K. Wang, T. Zhang, T. Xue, Y. Lu and S. G. Na, E-commerce personalized recommendation analysis by deeply-learned clustering, *Journal of Visual Communication and Image Representation*, vol.71, 102735, 2022.
- [2] S. Pashkov, *Video Game Industry Market Analysis: Approaches that Resulted in Industry Success and High Demand*, Master Thesis, Unit International Business Administration, Vaasa University of Applied Sciences, 2021.
- [3] S. M. Anwar, T. Shahzad, Z. Sattar, R. Khan and M. Majid, A game recommender system using collaborative filtering (GAMBIT), *Proc. of the 14th International Bhurban Conference on Applied Sciences and Technology (IBCAST)*, 2017.
- [4] H.-C. Yang and Z.-R. Huang, Mining personality traits from social messages for game recommender systems, *Knowledge-Based Systems*, vol.165, pp.157-168, 2019.
- [5] Q. Li, X. Liang, C. Su and Y. Wang, A game recommendation method based on machine learning, *2021 IEEE 3rd International Conference on Frontiers Technology of Information and Computer (ICFTIC)*, 2021.
- [6] Rajendran, D. P. Daniel and R. P. Sundarraaj, Using topic models with browsing history in hybrid collaborative filtering recommender system: Experiments with user ratings, *International Journal of Information Management Data Insights*, vol.1, no.2, 100027, 2021.
- [7] J. Devlin, M. W. Chang, K. Lee and K. Toutanova, BERT: Pre-training of deep bidirectional transformers for language understanding, *Proc. of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Minneapolis, MN, 2019.
- [8] M. Grootendorst, *KeyBERT: Minimal Keyword Extraction with BERT*, <https://maartengr.github.io/KeyBERT/index.html>, 2020.
- [9] P. Sharma and Y. Li, Self-supervised contextual keyword and keyphrase retrieval with self-labelling, *Preprints.org*, DOI: 10.20944/preprints201908.0073.v1, 2019.
- [10] M. Q. Khan, A. Shahid, M. I. Uddin, M. Roman, A. Alharbi, W. Alosaimi, J. Almalki and S. M. Alshahrani, Impact analysis of keyword extraction using contextual word embedding, *PeerJ Computer Science*, vol.8, e967, 2022.
- [11] A. Kosmopoulos, A. Liapis, G. Giannakopoulos and N. Pittaras, Summarizing game reviews: First contact, *SETN Workshops*, 2020.
- [12] K. W. Church, Word2Vec, *Natural Language Engineering*, vol.23, no.1, pp.155-162, 2017.