Journal of KIIT. Vol. 23, No. 3, pp. 39-50, Mar. 31, 2025. pISSN 1598-8619, eISSN 2093-7571 **39** http://dx.doi.org/10.14801/jkiit.2025.23.3.39

Movie Recommender System (BERT-More)

Hwi Chan Jeung*¹, Jun Pyo Jeon*², and Seung Yeop Lee**

Abstract

Large-scale rating data is required to recommend movies suitable for users on media platforms. However, since not all users can watch all movies and provide ratings, rating data is inherently sparse and acts as a major factor in the performance degradation of the recommender system. Previous studies have attempted to improve the recommendation performance by complementarily utilizing a user's profile information, reviews, and movie metadata to alleviate such data sparsity. However, existing research approaches have fundamental limitations in terms of data collection in that it is often difficult to obtain user profiles corresponding to personal information, and data written with ratings and reviews at the same time is even rarer. Therefore, this study proposes a BERT-More movie recommender system that can predict a high level of ratings by using only rating data, review data, and movie metadata written by different users to overcome the difficulty of collecting data. This model focuses on improving the performance of recommendations by effectively utilizing more accessible data.

요 약

미디어 플랫폼에서 이용자들에게 적합한 영화를 추천하기 위해서는 대규모의 평점 데이터가 필요하다. 그러 나 모든 이용자가 모든 영화를 시청하고 평점을 제공할 수는 없기 때문에, 평점 데이터는 본질적으로 희소성 을 가지며 이는 추천시스템 성능 저하의 주요 요인으로 작용한다. 선행 연구들은 이러한 데이터 희소성을 완 화하기 위해 이용자의 프로필 정보, 댓글, 그리고 영화의 메타데이터를 보완적으로 활용하여 추천 성능을 개선 하고자 시도하였다. 그러나 개인정보에 해당하는 이용자 프로필은 확보가 어려운 경우가 많으며, 평점과 댓글 이 동시에 작성된 데이터는 더욱 드물다는 점에서 기존 연구의 접근법은 데이터 확보 측면에서 근본적인 한계 를 가진다. 이에 본 연구는 데이터 확보의 어려움을 극복하고자, 서로 다른 이용자가 작성한 평점, 댓글, 그리 고 영화 메타데이터만을 활용하여 높은 수준의 평점 예측이 가능한 BERT-More 영화 추천시스템을 제안한다. 본 모델은 보다 접근 가능한 데이터를 효과적으로 활용함으로써 추천 성능을 향상시키는 데 중점을 둔다.

Keywords

BERT, data sparsity, movie rating, data collection

* MS candidate, Department of Media

- Communication, Pukyong National University
- ORCID1: https://orcid.org/0009-0002-7356-4926
- ORCID²: https://orcid.org/0009-0008-2347-2912
- ** Assistant Professor, Department of Media and Communication, Pukyong National University
- ORCID: https://orcid.org/0000-0001-7959-4005
- · Received: Dec. 16, 2024, Revised: Feb. 04, 2025, Accepted: Feb. 07, 2025
- · Corresponding Author: Seung Yeop Lee

Dept. of Media Communication, Pukyong National University, 45 Yongsoro, Namgu, Busan 48513, South Korea

Tel.: +82-51-629-4612, Email: yeopcp@pknu.ac.kr

I. Introduction

The advent of online streaming platforms and diverse digital media services has liberated users from fixed times and locations, allowing them to easily access a vast array of content [1]. However, the proliferation of media platforms has also led to a surge in content production, making it challenging for users to consume content efficiently [2]. In this context of information overload, recommender systems, which utilize users' preferences to provide personalized content recommendations, have become increasingly important [3].

Collaborative filtering performed well based on the similarity between items or users [4]. However, in actual online platform environments, the performance of collaborative filtering is limited and the application of recommender systems is often difficult due to the problem of deteriorating the performance of recommender systems due to data sparsity [5].

To address the issue of performance degradation of recommender systems due to the sparsity of user-movie rating data, a rating prediction model including reviews of movies or metadata including directors, actors, and plotlines of movies has been proposed [6]. However, in a real media environment, platform users are reluctant to provide their information [7], and only a few users write both reviews and ratings in the movie domain [8]. The models presented in previous studies have limitations in securing data. Therefore, this study proposes BERT-More, a recommender system that can be applied to real-world situations that do not utilize data that is difficult to obtain for model learning.

II. Related Work

2.1 BERT-CF

In the case of Item Based CF (IBCF), the existing CF model predicted the rating by using several methods (such as Cosine Similarity, Euclidean Distance, Jaccard Similarity, etc) to calculate the similarity of rating data after constructing a matrix with Movie-ID as rows, User-ID as columns, and rating as values. Hoang proposed a model applying the existing CF [5]. The CF model using BERT (BERT-CF) converts a user's profile and movie information into vectors through BERT, and then calculates the similarity by using the embedding of movie metadata and a user's profile instead of the Movie-ID and User-ID location in the movie-user matrix. Figure 1 outlines the framework of the BERT-CF model, highlighting the key processes involved.



Fig. 1. CF using BERT embedding model calculation example

2.2 BERT-MF

Similar to BERT-CF, BERT-Matrix Factorization (BERT-MF) utilizes embeddings of metadata [9]. Conventional MF calculates user preferences and item features as an inner product of vectors using latent features of users and items.

The initial value of the latent vector in MF is randomly initialized using Gaussian distribution and trained to gradually approach the actual rating during the learning process [10]. Figure 2 depicts the fundamental workflow of the BERT-MF model. BERT-MF also utilizes the Stochastic Gradient Descent (SGD), but for each item, the model further learns the additional information of the item by designating the pre-trained embedding vector as the initial latent vector.



Fig. 2. Examples of metadata embedding integration methods

2.3 BERT-HoRS

Figure 3 illustrates the operational mechanism of BERT-HoRS in detail. Recently, unlike the recommender system using metadata of users or items, BERT-HoRS, a model applied to a recommender system using data from users who wrote both ratings and reviews, has been proposed [11]. This is because if the data of users who have left two or more explicit preferences for one item is abundant, the data sparsity problem of that item can be alleviated [12].

BERT-HoRS converts User-ID and Hotel-ID into a Dense Vector through a User-Hotel Interaction Extractor. The above two vectors are vectors that explain the relationship between the user and the hotel, and are used as inputs for the Self-Attention Layer by concatenating the two vectors to minimize information loss, and finally, a vector reflecting the user's preference for the item is extracted.

The BERT-based ranking evaluation model and rating prediction model that have been proposed have two limitations. First, they failed to utilize all the data available for model training. BERT-CF [5] and BERT-MF [9] used only rating data and metadata, and BERT-HoRS used rating data and review data [11].



Fig. 3. BERT-HoRS model architecture

According to previous studies, models utilizing review or metadata were effective in alleviating the data sparsity of rating data, but they were not effective in predicting ratings.

Second, a model using review data and user profile that are difficult to obtain was proposed. It is very difficult to collect data in which a user has written both ratings and reviews for the same movie.

Table 1 shows the type of the MovieLens Dataset. Movie Lense dataset is one of the most widely used data, the dataset containing the most rating data has a data acquisition rate of only 6% [13]. In addition, Leaving a comment requires more effort and time [14] compared to rating, which only requires users to click the "like" button. The data acquisition rate is only 6% when a user has both rated and written a review for a movie.

Table 1. MovieLens data sparsity problem

Dataset	Sparsity level (%)					
ML-100k	$1 - \frac{100000}{943 * 1682} = 93.696\%$					
ML-1M	$1 - \frac{1000209}{6040 * 3900} = 95.754\%$					
ML-10M	$1\!-\!\frac{10000054}{71567*10681}\!=\!98.692\%$					
ML-20M	$1 - \frac{20000263}{138493 * 27278} = 99.470\%$					

III. Proposed Model

3.1 BERT-More

To overcome the limitations of existing models, BERT-More concatenates the User-ID, Movie-ID, the average rating given by each user, and the average rating received by each movie for use. When calculating the user similarity in cosine similarity-based CF, incorporating the average rating provides better insight into each user's pattern while accounting rating deviation [15]. This is particularly advantageous when dealing with user groups with various rating tendencies and contributes to enhancing the accuracy and prediction stability of the model.

The BERT-More model converts sentiment scores into consecutive values between 0 and 1 to utilize reviews written by different users. BERT is used in this process, and the sentiment score is information that additionally explains each movie. Previous studies on sentiment analysis [16] have mainly used a method of evaluating sentiment scores on a 5-point scale. In this study, although the sentiment score affects the prediction process of the model, a relatively low weight is given compared to the rating data. This ensures that the influence of sentiment scores is more constrained compared to rating data since reviews are written by different users, unlike ratings, which are directly assigned by a specific user to a movie. In fact, in previous studies, for more accurate rating prediction, the importance of the actual rating given by the user was reflected by normalizing the sentiment score or adjusting the score interval [17].

According to previous studies, actors actors have been recognized as one of the key attributes of movies [18]. In several recommender system studies, actor as an attribute is classified as a major factor that accurately reflects the user's interests when analyzing reviews. Additionally, individually computing the similarity between each attribute and the user has been shown to be effective in enhancing recommendation accuracy [19].

As a result, the movie recommendation model demonstrated the potential for providing more customized recommendations by considering not only the overall rating or sentiment score for a movie but also the user's preference for specific attributes. Therefore, in this study, easily accessible movie metadata was also utilized for rating prediction.

The BERT-More model proposed in this study aims to improve the accuracy of movie rating prediction by utilizing as much User-movie data as available. Previous studies adopted a method of combining multiple input vectors into one integrated vector using a vector combination method when combine different types of data.

This vector combination method was designed so that useful information was not lost while combining different types of data, which has the advantage of maximizing the performance of the predictive model [20]. This study also adopted an approach that concatenate data from different sources, combining various types of data into one vector, and then training the model on it.

Figure 4 is the structure of the BERT-More model proposed in this study. This model used BERT-based regression model to process text data and solve prediction problems. The model used BERT-base-uncased, which is a pre-trained model consisting of 12 Transformer layers and 768 hidden units per layer. Compared to BERT-large, BERT-base has fewer hidden layers and hidden units, so the inference speed is faster and the performance is better than BERT-small [21]. BERT can process up to 512 tokens, but for BERT-More, the maximum sequence was limited to 128 and the batch size was set to 32, increasing the learning speed and memory efficiency. The learning rate of BERT-base was originally 5e-5, but the learning rate was set to 2e-5 to ensure stable learning. In order to prevent overfitting, the dropout rate was set to 0.3 [22]. AdamW was used as an optimizer. AdamW is effective in preventing overfitting by independently applying weight decay as a variant of Adam optimizer [23]. Model training was conducted with a total of 3 epochs.

IV. Experiment

4.1 Data

The data used in this study are the MovieLens dataset and the Rotten Tomatoes dataset. In the data preprocessing process, users who had watched at least 25 movies and movies that were watched at least 25 times were selected and used. This is because 25 or more samples can be statistically considered as significant [24]. By applying this criterion, essential data for model training were secured while maintaining data quality.



Fig. 4. BERT-More model architecture

The data used in this study consist of 609 users, 618 movies, and 40,134 ratings, consisting of a sparse matrix in which about 90% are empty cells. This indicates that user-movie interactions is concentrated on specific movie, because there are many movie with no rating. The review data utilized in the experiment were sourced from Rotten Tomatoes, consisted of sentence-level reviews evaluated by movie critics for a particular movies. Additionally, metadata, including the movie title, director, genre, and actor, was also incorporated. In this study, the rating of a specific movie, the sentiment score derived from its review, and the movie metadata were concatenated into a single sentence. Table 2 presents an example of the experimental data. The model was then trained using concatenated sentence data comprising movie ratings.

Table 2. Experiment data example

Data types	Data example					
Movie title	Inception					
Director	Christopher Nolan					
Genre	Science Fiction, Thriller					
Actor	Leonardo Dicaprio					
Ratings	4.5					
Review	An emotionally resonant performance with a compelling storyline, masterfully executed					
Sentiment score	0.8					
Concatenated senctence	Rating: 4.5, Inception, directed by Christopher Nolan, is a Science Fiction and Thriller movie starring Leonardo DiCaprio. Review: An emotionally resonant performance with a compelling storyline, masterfully executed. Sentiment Score: 0.8					

4.2 Data sparsity experiment

This study aimed to utilize additional data to mitigate data sparsity and improve the accuracy of rating prediction. The total number of user and movie rating data obtained for rating prediction was 40,134, and the performance of the model was evaluated by gradually increasing the number of data. Specifically, the size of the dataset was increased in increments of 5,000 up to a maximum of 40,134. A total of eight experiments were conducted to assess whether the model could maintain stable performance even in a sparse data environment.

Table 3 presents a sparsity level similar to that of the MovieLens rating data. As shown in Table 1, the MovieLens dataset comprises multiple subsets with rating data accounting for approximately 1% and 6% of the total data. In this study, datasets with varying sparsity levels, ranging from approximately 1% to 9%, were constructed, and model performance was compared across different sparsity conditions..

Table 3.	Experiment	data	sparsity
----------	------------	------	----------

Dataset	Sparsity level (%)
5,000	$1 - \frac{5000}{609 * 618} = 98.672\%$
10,000	$1 - \frac{10000}{609 * 618} = 97.343\%$
15,000	$1 - \frac{15000}{609 * 618} = 96.050\%$
20,000	$1\!-\!\frac{20000}{609*618}\!=\!94.686\%$
25,000	$1\!-\!\frac{25000}{609*618}\!=93.358\%$
30,000	$1 - \frac{30000}{609 * 618} = 92.029\%$
35,000	$1 - \frac{35000}{609 * 618} = 90.701\%$
40,134	$1 - \frac{40134}{609 * 618} = 89.337\%$

4.3 Model stability experiment

Considering that the characteristics of the dataset differ partially from those used in existing CF or MF methods, this study examined whether the use of BERT embeddings has a significant effect on the dataset. The performance of the BERT-applied models (BERT-CF, BERT-MF) was compared with that of traditional collaborative filtering models (Conv-CF, Conv-MF).

BERT-HoRS demonstrated excellent performance when using ratings and review data written by the same user, but was included in the experiment to evaluate the performance on review data written by different users. In the case of BERT-More, if only the rating data is used, the detailed performance difference were analyzed by dividing it into the case of using a combination of ratings and sentiment scores, ratings and metadata, and ratings and sentiment scores. In addition, in this study, a random sampling technique was employed to measure the performance of the model after repeatedly separating the training data and evaluation data from the same dataset. This is because the random sampling technique ensures the consistent performance of the model and enhances the evaluation of the generalization ability more accurately by preventing the model from overfitting to specific data [25]. It is imperative to conduct additional trials to augment the dataset, as the imbalance in the case of rating data is pronounced [26], and the model's performance is contingent on the data cluster. Consequently, the entire dataset was randomly split into training data (75%) and evaluation data (25%), and this process was repeated five times.

4.4 Data sparsity experiment results

Conv-CF and Conv-MF refer to conventional CF and MF respectively. In the experiment, BERT-More was categorized into four variations: BERT-More (rating) when only rating data was used, BERT-More (r+s) when ratings and reviews were used, BERT-More (r+m) when ratings and metadata were used, and BERT-More (r+s+m) when all available data were utilized.

Figures 5 and 6 present graphs depicting MSE as a function of data size, where the x-axis represents data size and the y-axis represents MSE. Table 4 provides the MSE values corresponding to different data sizes for each model. Conv-CF exhibited a high MSE in data-scarce environments, but its performance gradually improved as the dataset size increased. Similarly,

BERT-CF also demonstrated a decreasing MSE trend as the data size grew. Notably, at around 25,000 rating data points, its performance was comparable to that of Conv-CF. However, in subsequent experiments, BERT-CF consistently outperformed Conv-CF. Both Conv-MF and BERT-MF exhibited a gradual decrease in MSE as data size increased. However, BERT-MF showed a performance plateau beyond 30,000 rating data points, with no significant further improvement.

Regarding the BERT-More model, the BERT-More (rating) model, which used only rating data, the BERT-More (r+s) model, which incorporated ratings and sentiment scores, and the BERT-More (r+s+m) model, which utilized ratings, reviews, and metadata, all consistently achieved excellent MSE results. In contrast. the BERT-More (r+m) model, which combined ratings and metadata, exhibited significant performance fluctuations and instability, recording the lowest performance among the BERT-More variants. Meanwhile, the BERT-More (r+s+m) model, which leveraged all available data, demonstrated the most stable and superior performance, with continuous improvements as data size increased.



	Conv-CF	BERT-CF	Conv-MF	BERT-MF	BERT -HoRS	BERT- More (rating)	BERT- More (r+s)	BERT- More (r+m)	BERT- More (r+s+m)
5,000	1.2959	0.9701	0.8599	0.6891	1.1038	0.0545	0.0132	0.0259	0.0567
10,000	0.9570	0.8569	0.8161	0.4338	1.1948	0.0305	0.0343	0.1732	0.0189
15,000	0.8716	0.7951	0.8077	0.4048	1.0843	0.0120	0.0219	0.1380	0.0248
20,000	0.8560	0.7889	0.7664	0.4169	1.0782	0.0371	0.0298	0.2392	0.0244
25,000	0.8050	0.8027	0.7706	0.4004	1.0291	0.0148	0.0231	0.1643	0.0248
30,000	0.8094	0.7962	0.7487	0.3917	0.9624	0.0361	0.0089	0.0025	0.0113
35,000	0.8322	0.8122	0.7456	0.3900	0.9616	0.0555	0.0153	0.1098	0.0149
40,134	0.8053	0.7813	0.7394	0.4107	0.9278	0.0191	0.0088	0.0133	0.0077

Table 4. Data sparsity experiment results

The superior performance of the BERT-based model with embeddings as the dataset size increases aligns with previous studies [27], which indicate that incorporating not only movie rating data but also additional embeddings from movie reviews and metadata significantly enhances the model's ability to predict ratings accurately.

4.5 Model stability experiment results

Figures 7 and 8 present graphs visualizing the MSE values from the experiments applying random where the x-axis represents sampling, randomly extracted data and the y-axis represents the MAE. Conv-CF showed relatively strong performance in the 2nd and 4th rounds, particularly in the 4th round, where it recorded a lower MAE than BERT-CF. However, overall, BERT-CF did not demonstrate a significant advantage over Conv-CF in terms of MAE.

Conv-MF exhibited fluctuating MAE values between 0.2 and 0.4, indicating a lack of stability. In contrast, BERT-MF outperformed Conv-MF in certain experiments, achieving а lower MAE than BERT-More(r+s), which only utilized ratings. Both the model combining ratings and sentiment scores and the model combining ratings and metadata (BERT-More(r+m)) demonstrated high performance with consistent results. Notably, the model incorporating all data (BERT-More(r+s+m)) achieved

the best performance, consistently recording the lowest MAE across all experiments.





Table 5 demonstrated that BERT-based models yielded stable and robust performance even when applied to data augmented through random sampling. Furthermore, these models effectively handled diverse and imbalanced datasets, maintaining consistency in predictive accuracy.

Random sampling	Conv-CF	BERT-CF	Conv-MF	BERT-MF	BERT -HoRS	BERT- More (rating)	BERT- More (r+s)	BERT- More (r+m)	BERT- More (r+s+m)
Random 1	0.8150	0.7818	0.4371	0.4321	0.9237	0.0883	0.0185	0.0089	0.0047
Random 2	0.8539	0.8199	0.4167	0.4152	0.9739	0.0476	0.0101	0.0076	0.0027
Random 3	0.8800	0.8452	0.4324	0.3858	0.9316	0.0537	0.0163	0.0086	0.0030
Random 4	0.8311	0.8873	0.4047	0.3633	0.9455	0.0779	0.0201	0.0118	0.0034
Random 5	0.8584	0.7714	0.4117	0.3524	0.9231	0.0446	0.0120	0.0087	0.0037

Table 5. Model stability experiment results

This stability highlights the model's ability to generalize across varying data distributions and underscores its adaptability in scenarios where data heterogeneity or class imbalance might otherwise impair performance. These findings further emphasize the potential of BERT-based approaches to deliver reliable outcomes under challenging experimental conditions.

V. Conclusion

This study presents a methodology to address data sparsity in movie recommender systems, making a significant academic contribution. While existing studies have primarily developed recommendation models by combining one or two of the following elements—user rating data, review data, and metadata —this study proposes the BERT-More model, which integrates all three.

In particular, the limitations of previous methods were effectively mitigated by introducing a novel regression approach that converts review data into sentiment scores, integrates them with metadata, and then utilizes them as input for BERT. Experimental results demonstrated that the BERT-More model, which applies this regression-based approach, outperforms the BERT-based collaborative filtering model (BERT-CF), the matrix factorization model (BERT-MF), and the comment-based model (BERT-HoRS) employing a multilayer perceptron (MLP) proposed in previous studies.

These findings indicate that beyond merely

leveraging BERT embeddings, the accuracy and efficiency of recommendation systems can be further improved by restructuring how embedded data is utilized. Moreover, this study confirms that a model relying solely on easily obtainable data (reviews and metadata) can achieve superior performance compared to traditional models.

The significance of BERT-More lies not only in its introduction of a novel academic approach to mitigating data sparsity but also in its demonstration of a methodology capable of achieving high performance even in environments with limited data availability.

This study also provides significant implications practical implications. The BERT-More model offers valuable insights into the construction and operation of the actual recommender system. In movie metadata-such as recommender systems, genre, director, and actor-was found to be a more enhancing influential factor in recommendation accuracy than review data. This suggests that even with limited resources, an efficient recommender system can be developed. Movie recommender systems fundamentally differ from product recommender systems for items such as hotels and home appliances. Whereas services such as hotel reservation applications are have limited consumption and infrequent changes, movies are everyday content that can be easily consumed [8].

This characteristic contributes to the sparsity of rating data. These characteristics highlight the importance of incorporating review data and metadata alongside rating data in modeling, underscoring the effectiveness of the BERT-More model proposed in this study. WWatcha continues to prioritize user rating data [28], while Amazon's approach of delivering personalized recommendations based on ratings and user-generated reviews remains widely adopted in industry [29]. However, in the movie field, it is rare for the same user to write both rating and review data. To resolve the issue, this study proposed an approach that utilizes sentiment scores extracted from reviews as separate metadata. which provides differentiated competitiveness from existing methods based on data from the same user.

However, a key limitation of this study is its difficulty in overcoming the cold start problem, as the proposed model does not incorporate user-specific attributes such as nationality, race, age, and occupation. Additionally, the analysis is limited to users who have watched at least 25 movies. Furthermore, the lack of individual weighting for each metadata element in the model's input may lead to overfitting. Addressing this issue will require further refinement of metadata weighting strategies.

References

- D. Menon, "Purchase and continuation intentions of over-the-top (OTT) video streaming platform subscriptions: A uses and gratification theory perspective", Telematics and Informatics Reports, Vol. 5, pp. 100006, Mar. 2022. https://doi.org/10.1016/j.teler.2022.100006.
- [2] C. Villavicencio, S. Schiaffino, J. A. Diaz-Pace, and A. Monteserin, "Group recommender systems: A multi-agent solution", Knowledge-Based Systems, Vol. 164, pp. 436-458, Jan. 2019. https://doi.org/10.1016/j.knosys.2018.11.013.
- [3] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions",

IEEE Transactions on Knowledge and Data Engineering, Vol. 17, Np. 6, pp. 734-749, Jun. 2005. https://doi.org/10.1109/TKDE.2005.99.

- [4] D. Valcarce, A. Landin, J. Parapar, and Á. Barreiro, "Collaborative filtering embeddings for memory-based recommender systems", Engineering Applications of Artificial Intelligence, Vol. 85, pp. 347-356, Oct. 2019. https://doi.org/10.1016/j.engappai. 2019.06.020.
- [5] B. N. M. Hoang, et al., "Using Bert Embedding to improve memory-based collaborative filtering recommender systems", 2021 RIVF International Conference on Computing and Communication Technologies (RIVF), Hanoi, Vietnam, pp. 1-6, Aug. 2021. https://doi.org/10.1109/RIVF51545.2021. 9642103.
- [6] B. Ray, A. Garain, and R. Sarkar, "An ensemble-based hotel recommender system using sentiment analysis and aspect categorization of hotel reviews", Applied Soft Computing, Vol. 98, pp. 106935, Jan. 2021. https://doi.org/10.1016/ j.asoc.2020.106935.
- [7] T. Bao and H. Kim, "Factors Influencing the Resistance of Recommendation System in OTT Service Focusing on the Case of Netflix", Korean Journal of Broadcasing & Telecommunications Research, No. 115, pp. 9-46, Jul. 2021.
- [8] S. Matrix, "The Netflix effect: Teens, binge watching, and on-demand digital media trends", Jeunesse: young people, texts, cultures, Vol. 6, No. 1, pp. 119-138, 2014. https://doi.org/10.3138/ jeunesse.6.1.119.
- [9] J. He and H. Hu, "MF-BERT: Multimodal fusion in pre-trained BERT for sentiment analysis", IEEE Signal Processing Letters, Vol. 29, pp. 454-458, Dec. 2021. https://doi.org/10.1109/LSP.2021.3139856.
- [10] H. J. Xue, et al., "Deep matrix factorization models for recommender systems", IJCAI, Vol. 17, pp. 3203-3209, Aug. 2017.
- [11] S. Park, X. Li, J. Y. Kim, and J. K. Lim,

"Development of a BERT-based Hotel Recommender System Reflecting User's Specific Features", Journal of Intelligence and Information Systems, Vol. 30, No. 1, pp. 139-158, Mar. 2024. https://doi.org/10.13088/jiis.2024.30.1.139.

- [12] S. Lee and J. Y. Choeh, "Predicting the helpfulness of online reviews using multilayer perceptron neural networks", Expert Systems with Applications, Vol. 41, No. 6, pp. 3041-3046, May 2014. https://doi.org/10.1016/j.eswa.2013.10.034.
- [13] V. Verma and R. K. Aggarwal, "Neighborhoodbased collaborative recommendations: an introduction", Applications of Machine Learning, pp. 91-110, May 2020. https://doi.org/10.1007/978-981-15-3357-0 7.
- [14] S. W. Lee, et al., "The future of The Internet Industry", Hanul, Seoul, 2019.
- [15] K. G. Suresh, "Building a Recommendation System with R", acorn+PACKT Technical Book, pp. 84-87, 2017.
- [16] O. Abiola, et al., "Sentiment analysis of COVID-19 tweets from selected hashtags in Nigeria using VADER and Text Blob analyser", Journal of Electrical Systems and Information Technology, Vol. 10, No. 1, Jan. 2023. https://doi.org/10.1186/s43067-023-00070-9.
- [17] S. Y. Yun and S. D. Yoon, "Item-Based Collaborative Filtering Recommendation Technique Using Product Review Sentiment Analysis", Journal of the Korea Institute of Information and Communication Engineering, Vol. 24, No. 8, pp. 970-977, Aug. 2020.
- [18] M. Saadati, S. Shihab, and M. S. Rahman, "Movie recommender systems: Implementation and performance evaluation", arXiv preprint arXiv:1909.12749, Sep. 2019. https://doi.org/10. 48550/arXiv.1909.12749.
- [19] M. Kumar, et al., "A movie recommender system: Movrec", International journal of computer applications, Vol. 124, No. 3, Aug. 2015.

- [20] R. Cao, X. Zhang, and H. Wang, "A review semantics based model for rating prediction", IEEE Access, Vol. 8, pp. 4714-4723, Dec. 2019. https://doi.org/10.1109/ACCESS.2019.2962075.
- [21] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding", NAACL-HLT. in Proc. Stroudsburg, PA. USA: Association for Computational Linguistics, Vol. 1, pp. 4171-4186, Jun. 2019. https://doi.org/10.18653/v1/N19-1423.
- [22] N. Srivastava, et al., "Dropout: a simple way to prevent neural networks from overfitting", The journal of machine learning research, Vol. 15, No. 1, pp. 1929-1958, Jun. 2014.
- [23] I. Loshchilov and F. Hutter, "Decoupled Weight Decay Regularization", arXiv preprint arXiv:1711.05101, Nov. 2017. https://arxiv.org/abs/ 1711.05101.
- [24] B. M. King and E. W. Minium, "Statistical reasoning in psychology and education", New York: Wiley, 2003.
- [25] D. Li, et al., "On item-sampling evaluation for recommender system", ACM Transactions on Recommender Systems, Vol. 2, No. 1, pp. 1-36, Mar. 2024. https://doi.org/10.1145/3629171.
- [26] J. W. Choi, S. G. Choi, and T. W. Kang, "Augmentation Strategies on Multi-Modal Lifelog Data for Smartphone Gait Authentication", KIIT, Vol. 22, No.11, pp. 39-45, Nov. 2024. http://dx.doi.org/10.14801/jkiit.2024.22.11.39.
- [27] W. S. Kang, S. W. Lee, and S. M. Choi, "A Matrix Factorization-based Recommendation Approach with SBERT Embeddings", KIIT, Vol. 21, No. 11, pp. 203-211, Nov. 2023. http://dx.doi.org/10.14801/jkiit.2023.21.11.203.
- [28] T. H. Park, "Watcha demonstate a high level of proficiency in selecting movie from the Netflix streaming platform", Insight Korea, pp. 40-43, Aug 2020.

[29] K. Wu and K. Chi, "Enhanced e-commerce customer engagement: A comprehensive three-tiered recommendation system", Journal of Knowledge Learning and Science Technology, Vol. 2, No. 3, pp. 348-359, 2023. https://doi.org/10.60087/jklst.vol2.n2.p359.

Authors

Hwi Chan Jeung



2023. 8 : BA degree, Department of Media Communication, Pukyong National University 2023. 9 ~ Present : MS candidate, Department of Media Communication, Pukyong National University

Research interests : Personalization Algorithm, Deep Learning, NLP, SNS

Jun Pyo Jeon



2024. 8 : BS degree, Department of Media Communication, Pukyong National University
2024. 9 ~ Present : MS candidate, Department of Media Communication, Pukyong National University

Research interests : NLP, Text Mining, SNS, Data Visualization

Seung Yeop Lee



2003. 2 : BS degree, Department of Electrical and Electronic Engineering, Yonsei University2005. 2 : MS degree, Department of Computer Science and Engineering, Seoul National University

2017. 2 : PhD degree, Graduate School of Information, Yonsei University

2022. 3 ~ Present : Assistant Professor, Department of Media Communication, Pukyong National University Research interests : Digital Media, SNS, Data Science, Personalization Algorithm