IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 1 November 2022, accepted 18 November 2022, date of publication 28 November 2022, date of current version 12 December 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3225190

RESEARCH ARTICLE

Analyzing the Impact of Components of Yelp.com on Recommender System Performance: Case of Austin

SEUNGWOO LEE[1], HYUNJOO SHIN[2], ILYOUNG CHOI[1,3], AND JAEKYEONG KIM[1,2]
[1]Department of Big Data Analytics, Kyung Hee University, Seoul 02447, South Korea
[2]Graduate School of Business Administration, Kyung Hee University, Seoul 02447, South Korea
[3]BK21 Four Program Educational Research Group, Kyung Hee University, Seoul 02447, South Korea

Corresponding author: JaeKyeong Kim (jaek@khu.ac.kr)

This work was supported in part by the BK21 FOUR (Fostering Outstanding Universities for Research) funded by the Ministry of Education (MOE), South Korea; and in part by the National Research Foundation of Korea (NRF).

ABSTRACT As people's demand for eating out is steadily increasing, the number of restaurants is continuously increasing, and catering industry platforms such as Yelp, Open Table, and Zomato provide basic information and evaluation information of restaurants and restaurant recommendation services suitable for users. Existing research on recommending restaurants mainly uses only evaluation information to find neighbors, and the use of user and restaurant information is still in its infancy. In addition, there is little study on how various types of input information affect the performance of the recommender system. This study examines the influence of three component information provided by Yelp.com on the performance of the recommender system using various real restaurants, reviews, and users dataset provided by Yelp.com. For this purpose, Two Phase Experiment was designed, and restaurant data located in Austin, Texas, USA, which has the largest number of review data, was collected. As a result of the experiment, elite status, the cumulative number of reviews, price range and average rating of restaurants could improve the recommendation performance.

INDEX TERMS Restaurant recommender systems, yelp, review data, experimental design.

I. INTRODUCTION

The ICT(Information and Communication Technology) has developed rapidly, and mobile devices are spreading worldwide. So that people generate various types of information, such as location, text, voice, anywhere and anytime, which helps people make decisions in their daily lives. For example, people who want to buy a new laptop can easily collect information about various brands' laptops to make optimal decisions. However, it makes the decision-making process more complex as the range of options extends [1], [2]. It causes some people want something to solve, not help, their decision-making problems directly. They do not want several laptop information from different brands anymore, but the one recommended to buy. Some global platforms, such as Amazon, YouTube, and Netflix, have provided recommendation services to resolve people's requirements [3], [4], [5].

As people's income levels rise, the demand for eating out and the number of restaurants has been constantly increasing. Restaurant platforms such as Yelp, Open Table, and Zomato provide inclusive information and reviews on restaurants for people to support decision-making to visit the restaurants preferred [6]. Furthermore, such platforms collect the reviewers' information and develop recommender systems(RS) with information on restaurants, reviews, and reviewers to provide recommendation services. RS refers to a system that recommends restaurants using information filtering techniques with a user's history of visiting or experiencing restaurants. It has been developed to resolve a user's decision-making problems from information overload. Traditional approaches to RS are based on Contents Based Filtering(CB) and Collaborative Filtering(CF). CF is well known and has been widely used.

The associate editor coordinating the review of this manuscript and approving it for publication was Fabrizio Marozzo.

CF recommends products purchased or services experienced by similar users to a target user.

Most recommendation services are based on users' explicit information on items, such as ratings or stars. However, in recent years, in addition to explicit information about products or services such as ratings and star ratings, as multimedia information such as reviews and photos increases, the recommender system must consider implicit information about users and items in order to achieve better performance. Therefore, some studies on RS considering implicit information are actively being conducted. However, these studies aim to develop new systems to improve the accuracy or diversity of RS without knowing which information could improve the performance [7]. It causes the systems to be restrained in explaining to people why the recommendations are suggested to them. Also, once they know how the information affects the performance of the recommender system, restaurant managers will focus on that specific information to improve customer satisfaction. In addition, restaurant reservation platform managers will also have a perspective on what information they need to acquire and manage more in the future. Therefore, this study aims to measure the impact of diverse information on RS and analyze it to determine which information positively affects the performance of RS.

We collected the dataset from Yelp.com, a global platform for restaurants, in the city of Austin because it has the most reviews in the United States. And due to the impact of COVID-19, data is too scarce after 2020, so 10 years of data from 2010 to 2019 were directly web crawled. It contains 4,259 users, 4,469 restaurants, and 196,934 reviews. We designed Two Phase Experiment Framework to measure the impact of information on the performance of RS more precisely. In phase 1, we develop a revised CF adding a filtering neighborhood step to traditional CF. Each feature of users and reviews information is put into the new step, and the accuracy of RS based on all features is calculated. The accuracy derived from traditional CF is compared to the accuracy of revised CF with each feature. When the accuracy improves considering a specific feature, we infer it affects the performance of RS positively. As the same, we develop another revised CF adding a filtering recommendation candidate step to traditional CF and compare the accuracy according to restaurant features in phase 2. As a result of the experiment, elite status of reviewers, the cumulative number of reviews, price range and average rating of restaurants could improve the RS performance.

II. LITERATURE REVIEW

A. RECOMMENDER SYSTEM

As the decision-making process has been getting more complex, people want some support to help their determination, especially related to purchasing. RS tends to be seen as the most widely used and powerful solution for supporting their purchasing decisions. Numerous global platforms have been applying RS to their own business to improve customer satisfaction. We cannot imagine Youtube or Netflix without recommendation services anymore.

The basic idea of RS is to figure out one's preference and suggest items or services used by others with similar preferences because past interests or tendencies often present good directions for future choices. One source to extract people's preferences is interactions between users and items, called feedback. It is possible to create a similar group of users interested in similar items by utilizing the set of feedback. The feedback made from the group can be utilized to recommend to individuals belonging to that [8]. We call this group neighborhood.

One of the most popular RS is CF, which uses ratings as feedback that multiple users have already evaluated to predict ratings that have not yet been evaluated. The recommendation process of CF can be specified into data representation, neighborhood formation, and recommendation generation. In data representation, a user profile is developed by a set of transactions of m users for n products as input data of CF. In neighborhood formation, establishing neighborhoods is conducted by calculating similarities based on input data between users. In recommendation generation, a list of the top N recommendations purchased by neighbors but not by a target user is suggested.

Following the studies for RS, users' explicit feedback is regarded as their preferences. Ratings, as explicit feedback, indicate the overall satisfaction of users with an item, including food price, quality, and so forth. Many RS studies extract a user preference based on ratings [9]. On the other side, when a user logs into a platform, the user's preferences are extracted by answering some queries. Miao et al. [10] developed an RS using queries that ask users to select their preferred price range and food type when logging into the system. Nowadays many RSs extract users' implicit preference analyzing their log history, review data, and so on [1], [2]. Therefore, most current RSs exploit implicit information to reduce users' burden and identify more accurate preferences.

B. RESTAURANT RECOMMENDER SYSTEM

RS is flexible and applicable enough to be employed in all types of business. The restaurant business is the one in that RS studies have been conducted. RS in the restaurant business is especially beneficial for first time travelers to unfamiliar cities. As the same, it identifies users' preferences for restaurants and suggests recommendations that best match their preferences.

The current studies for restaurant RS have been exploiting implicit information to extract users' preferences. Unlike explicit feedback, implicit feedback does not literally indicate the preferences [11], [12], [13], [14]. For example, it cannot be assumed that a user prefers a restaurant because he has visited it several times before. Nevertheless, implicit feedback has significant advantages in terms of diversity and amount of data to discover preferences [15].

It is a review that restaurant studies have been used actively. DeacPetruşel and Limboi [16] applied Senti-WordNet to reviews to estimate user emotions for different restaurant features and suggested recommendations considering attractiveness, relevance, and popularity of restaurants. Asani et al. [12] proposed RS that extracts food preferences from reviews and compares them with restaurant menus. So that the RS could recommend restaurants and specific menus served by them. Moreover, ICT growth makes studies start to exploit other information, not only reviews. Sarkar et al. [17] attempted to derive insights using deep learning technology and RS for numerous reviews collected from Zomato, a global restaurant platform. He extracted price range and menu items from the reviews, and compared several models that reflected them with RS approach. Zeng et al. [18] and Gupta et al. [19] proposed RS that suggests recommendations using the located information in real time.

RS studies exploiting Yelp dataset have been conducted with various approaches. One of the traditional approaches is sentiment analysis based on user reviews. After deep-learning technology arose, it was developed by incorporating embedding methods, such as Word2Vec. Also, some approaches have emerged by applying social network analysis and clustering techniques to RS. Zhang et al. [20] proposed a CB model for cultural restaurants with restaurant reviews and sentiments of them. The model analyzes the weights and sentiments of various aspects within reviews at the sentence level. Cultural restaurants could be suggested with some areas to improve. Sawant [21] developed a personalized RS using Network-Based-Inference CF Algorithm. The recommendation problems are presented as a graph projection. User ratings weigh the edges from the user to the business. Bathla et al. [22] calculated Recop Score by applying sentiment analysis and a logistic regression model to user reviews and user similarities based on the score. Zhao et al. [23] proposed a recommender model that can extract sentiments and contextual information by converting user reviews into multidimensional vectors to resolve the limitations of traditional topic modeling. Eskandanian et al. [24] devised a pre-filtering clustering approach for group users with similar tolerance for diversity of recommendations. It is a methodology to personalize diversity by independently performing CF on different clusters by the degree of diversity in the user profile. LR et al. [25] developed CREPMF, a two-stage clustering-based matrix factorization model, to resolve the limitation that recommendations through existing social network analysis do not sufficiently reflect user preferences. The first clustering is applied to users, and the second is applied to items according to the rating.

The Yelp website provides a lot of information about reviewers and restaurant characteristics, including review data, so many recommender system studies have been conducted. However, most existing studies on restaurant RS (based on Yelp dataset) have been conducted about how to reflect implicit information in RS, but few studies on which information contributes to improving recommendation performance. It causes RS performance to be improved but non-explainable at once. Therefore, this study aims to measure the impact of diverse implicit information on RS and analyze it to define which information positively affects RS.

III. EXPERIMENTAL DESIGN

A. DATASET

The dataset on Yelp.com, a global restaurant platform, has been used to conduct the Two Phase Experiment in this study. It is composed of three sub-datasets on users, restaurants, and reviews. Each sub-dataset includes numerous features, such as the number of reviews commented and elite user status on the user dataset, as shown in Figure 1. The user sub-dataset includes username, location, and the cumulative number of reviews, as shown in Figure 1(a). The restaurant sub-dataset consists of restaurant name, average rating, the number of reviews, and price range, as shown in Figure 1(b). The review sub-dataset includes the username, elite status, and the number of votes for the review, as shown in Figure 1(c).

We have filtered the dataset generated from 2010 to 2019 before the COVID19 virus spread to derive more accurate results. 4,509 restaurants in Austin, Texas, have been selected for the experiment. RS fundamentally helps users make future decision-makings based on their past histories, some studies split the dataset considering time not randomly [26], [27]. Ji et al. [28] pointed out, learning from interactions that happen in future contradicts to the problem definition of RS. Paraschakis et al. [29] argued that time-aware evaluation, where dataset are divided chronologically is meaningful comparing the two different approaches to splitting dataset, random split and chronological split. However, some studies still put the models learn user interactions with iPhone 14 to predict the possibility to buy iPhone 13. For enhancing reasonability of this study from the previous studies, we chronologically divided the dataset, the past eight years(from 2010 to 2017) dataset as training sets and the latest two years(from 2018 to 2019) as test sets. The training dataset contains 151,297 reviews, 81.44% of the total dataset, and the test dataset contains 34,447 reviews, 18.56% of the total dataset. Data sparsity problems could limit the performance of RS, which refers to the insufficiency of data about users' preferences [30]. To ease this problem, we filtered out users who wrote at least 20 reviews [31], [32], [33]. The statistics of the final dataset we used are shown in [Table 1].

TABLE 1. Descriptive statics of the dataset.

#Users	#Restaurants	#Reviews
8,246	4,509	226,106

B. TWO PHASE EXPERIMENTAL DESIGN

Recommendations have been generated by only the user's quantitative preferences, such as ratings or stars, on CF. However, this study redevelops CF to consider users, restaurants, and reviews. Furthermore, we compare its performance with the traditional CF's to reveal which aspects could improve

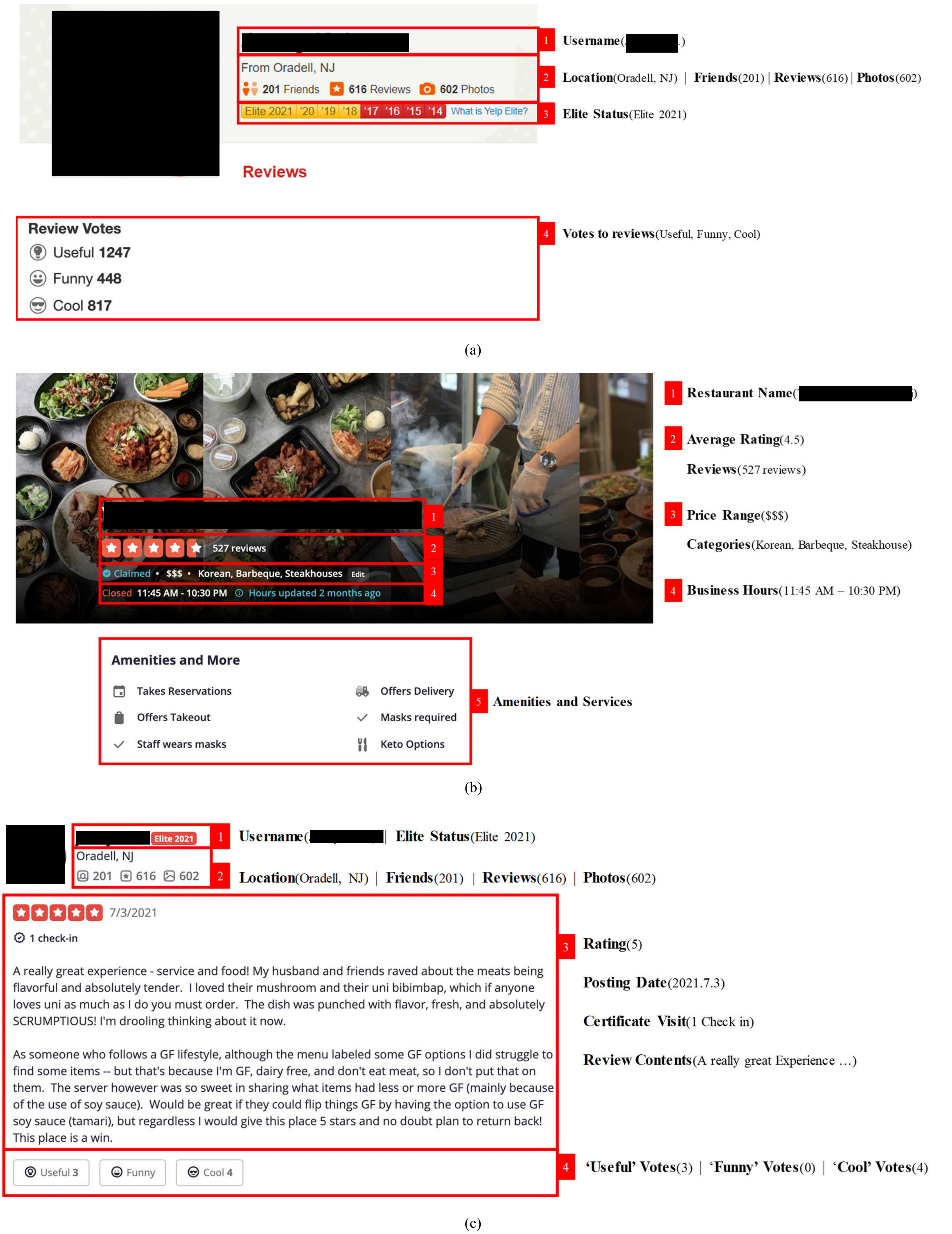

FIGURE 1. (a) Components of User Sub-Dataset (b) Components of Restaurants Sub-Dataset. (c) Components of Review Sub-Dataset.

the recommendation performance. This experiment was conducted in two stages. In the first step, information from users and reviews is used to filter out similar neighbors in the CF process. In the second step, the restaurant information is taken into account to derive the candidate's final recommendation. There are several types of CF, but this experiment uses UBCF(User Based CF) because it needs to find similar neighbors using user and review information in the first phase. Since the purpose of this study is not to increase the performance of the recommender system, but to investigate what type of information about user characteristics affect the performance of the recommender system, we use UBCF. UBCF calculates the affinity similarity between a target user and other users, and predicts the Visiting Likelihood Score(VLS) of a target's visit to a restaurant based on the similarity.

The formula calculating similarities between users a and b is shown in (1).

$$SIM\ (a, b) = \frac{\sum_{i\in R}\left(r_{a,i} - \overline{r_a}\right) \cdot \sum_{i\in R}\left(r_{b,i} - \overline{r_b}\right)}{\sqrt{\sum_{i\in R}\left(r_{a,i} - \overline{r_a}\right)^2} \cdot \sqrt{\sum_{i\in R}\left(r_{b,i} - \overline{r_b}\right)^2}} \tag{1}$$

where i means the entire set of restaurants, $r_{a,i}$ means the rating given to the restaurant i by user a, and $\overline{r_a}$ means the average rating given to the entire restaurant by user a [1], [2]. After the similarities between users are calculated by (1), N of neighbors with the most similar preferences are selected. VLS for restaurant i of user a is calculated by (2).

$$VLS\ (a, i) = \overline{r_a} + \frac{\sum_{b\in N_a}\left(r_{b,i} - \overline{r_b}\right) \cdot SIM\ (a, b)}{\sum_{b\in N_a} SIM\ (a, b)} \tag{2}$$

N_a is the set of neighbors of user a, $r_{b,i}$ is the rating given to restaurant i by neighbor b, $\overline{r_b}$ is the average rating given to the entire restaurant by neighbor b, and $SIM(a,b)$ is similarity score between user a and neighbor b.

C. PHASE 1 EXPERIMENT

Recommendations have been generated by only the user's quantitative preferences, such as ratings or stars, on traditional CF. However, this study desires to measure the impact of features from users, restaurants, and reviews information on RS and reveal which features could improve the performance. To measure the impact of features more precisely, we developed two different revised RSs considering features; one is for features of users and reviews, and the other is for features of restaurants. That is because the features of users and reviews are closer to a factor of users, and the features of restaurants are closer to a factor of items in RS. Therefore, we developed Two Phase Experiment Framework shown in Figure 2. Each phase experiment is conducted independently.

The process of phase 1 experiment identifies which features related to users and reviews positively influence RS's performance. Traditional CF selects neighbors with similarities calculated by quantitative preferences. In this phase, we develop Neighbor-Restricted CF(NRCF), limiting the range of neighbor candidates with each feature of users and reviews. For example, the candidates must be Elite(Certificate Reviewer) when considering the feature of elite status. In the end, we compare the accuracy of traditional CF and NRCF. When NRCF considering a feature of elite status performs better than traditional CF, the feature is regarded as a positive feature for RS.

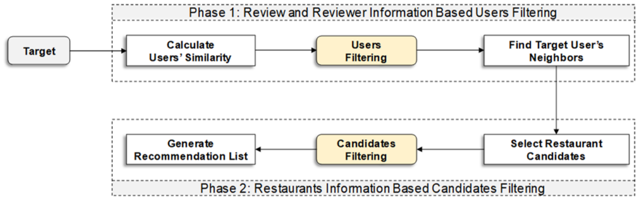

FIGURE 2. **Framework of two phase experiment.**

The users' features considered in this phase are elite status and activity. Recency and lengthiness are features when it comes to reviews. Elite refers to a reviewer certificated by Yelp.com with activities on the platform, and activity is the number of reviews posted from the date a user joined Yelp.com. Recency refers to reviews posted in the last three years(2015-2017), and lengthiness means the number of characters in a user's reviews. The framework of the first phase is shown in Figure 3.

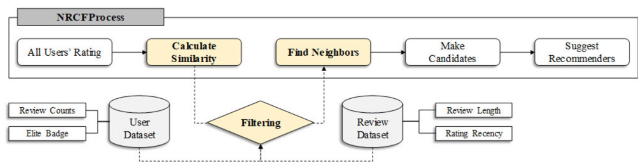

FIGURE 3. **Framework of NRCF.**

D. PHASE 2 EXPERIMENT

UBCF recommends restaurant candidates in order of expected ratings when target users visit based on neighbors' restaurant ratings. In phase 2, we develop Double-Candidates CF(DCCF) that recommends final restaurant candidates by CB method among restaurant candidates based on UBCF. CB is an RS that calculates and recommends restaurants similar to those visited by users in the past. It is applied to filter final recommendation list to a user from pre recommendation list, similar to his/her favorite restaurants' features, such as price range and average ratings. In this experiment, we define favorite restaurants that rated 5 out of 5. We want to test whether the performance of RS is improved when a restaurant is finally recommended based on which of several characteristics average price of a restaurant, average rating of a restaurant, etc.) of restaurants that users have visited in the past.

The second phase experiment is compared with UBCF generating final recommendations based on only VLS to identify which aspects of restaurants influence the performance of RS. The restaurant aspects we use in this phase are price range, categories, average ratings, and cumulative review counts.

The framework of the two stage experiment is shown in Figure 4.

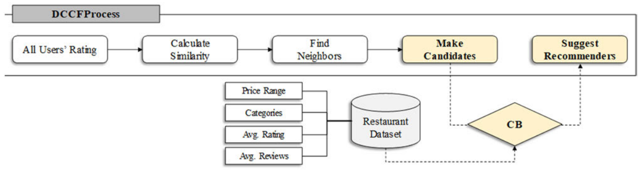

FIGURE 4. Framework of DCCF.

The second phase consists of the following five steps. In step 1, users' ratings derive similarities between users by Pearson Correlation Coefficients. In step 2, neighbors are filtered based on their similarities in descending order. In step 3, VLS is calculated from the similarities between a target user and his/her neighbors, generating candidates for recommendation. In step 4, pre-recommendations are selected out from the candidates based on VLS. And CB is applied to pre-recommendations for completing final recommendations considering a particular restaurant aspect. In the final step, F1 score evaluates the performance with the user's visiting history.

We have varied the number of recommendations from 2 to 20 to see the tendency of change in the performance.

E. EVALUATION

There are several methods to evaluate RS performance. The criteria to determine a proper method for assessing RS depends on the type of data and the purpose of assessment. Prediction accuracy is fit for continuous data, and classification accuracy is for categorical data. In the first phase, this study evaluates the performance with predictive accuracy to know the differences between predictive ratings by NRCF and actual ratings of a user to restaurants. However, in the second phase, classification accuracy is computed because a set of recommendations generated by DCCF is compared with a set of restaurants a user actually visited.

Mean Absolute Error(MAE) and Root Mean Square Error(RMSE) are computed in the first phase, which are widely used to evaluate the predictive accuracy of RS. Both measures evaluate predictive performance by calculating the difference between the actual rating and the predicted one. The smaller the value means, the higher the predictive accuracy. The MAE gives the same weight regardless of the magnitude of the error between the actual and predicted ratings. However, RMSE gives a relatively high weight a large error between the actual and predicted ratings [34], [35]. In the second phase, the performance was calculated by F1 score, which is a balanced weighted average between precision and recall. It is a tradeoff between Precision and Recall. As the number of recommendations increases, Recall improves while Precision worsens. Therefore, this study used F1 score to compensate for the tradeoff. A High F1 Score means a high prediction ability for the recommender system [36], [37], [38], [39], [40].

MAE and RMSE are obtained from the entire dataset, but the F1 score was tested by dividing the data from 2010 to 2019 into a training set and a test set, as described in the data set.

IV. EXPERIMENTAL RESULT

A. PRIOR EXPERIMENT FOR VARIABLE BASELINE

Before two phase experiments, a prior experiment should be conducted for some ambiguous features. In terms of the feature, Elite Status, we filter the neighbors who are Elite. However, the feature of the cumulative number of reviews is a bit more complicated than others. When filtering users who posted 'many' reviews, the criterion of 'many' is ambiguous.

So, a prior experiment was conducted by varying the criteria - mode, median, and mean of review counts. The value with the highest recommendation performance was used as the criterion for the feature of cumulative reviews. Figure 5 is a distribution of the cumulative reviews. The mode value is 40, the median value is 116 and the mean value is 265.52.

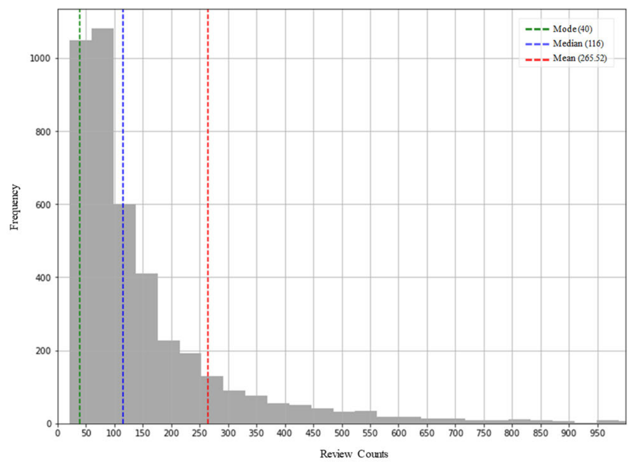

FIGURE 5. Distribution of the cumulative reviews.

Descriptive statistics on the three values are shown in [Table 2]. The number of users who wrote reviews over the mode value was 3,889(91.31%). 1,823(42.80%) users wrote reviews over the median value, and 659(15.48%) users wrote reviews over the average value.

TABLE 2. Descriptive statics of review counts.

	MODE	MEDIAN	MEAN
VALUE (x)	40	116	265.52
COUNTS ($x \leq$)	3,889	1,823	659
RATIO ($x \leq$)	91.31%	42.80%	15.48%

The prior experiment was developed to reveal which criterion could draw the best performance in RS. For that, we applied the phase 1 experiment to the feature of review counts by setting the size of neighbors from 10 to 100. The experimental results are shown in Figure 6. "Existing" means the traditional UBCF, and "Mean" refers to CF considering only neighbors who wrote reviews over mean value. "Mode" and "Median" refer to CF with neighbors who wrote over mode and median values.

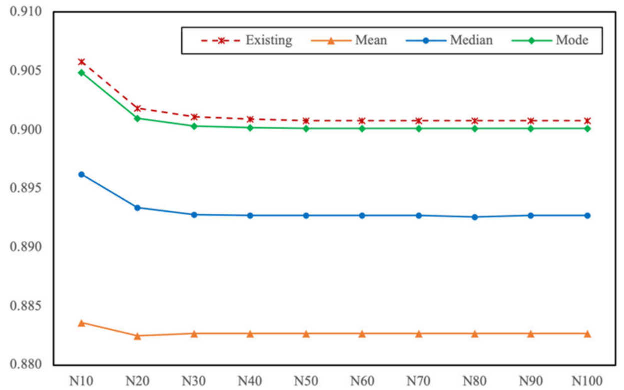

FIGURE 6. Result of prior experiment(MAE).

As a result of the experiment, "Mean" showed the best RS performance. Therefore, users who wrote reviews over the mean value would be considered in the post-experiment in terms of review counts. The prior experiment had been conducted through the above process when the feature criterion is required.

B. PHASE 1 EXPERIMENTAL RESULT

The experiment in phase 1 demonstrates a feature that improves RS performance belonging to users and reviews information. It was tested and compared by varying the size of neighbors from 2 to 30. While the traditional CF-based recommender system selects neighbors by deriving the similarity between the target user and all users, this experiment derives the neighbor by limiting the range of users based on user and review information. The results of phase 1 experiment are shown in Figure 7 (MAE) and Figure 8 (RMSE) with [Table 3] and [Table 4]. "Existing" is a traditional UBCF that only considers ratings; "Elite" is NRCF restricting neighbors to users being elite; "Lengthiness" is NRCF restricting neighbors with an average length of reviews; "Activity" is NRCF with the number of reviews posted from the date a user joined the platform; "Recency" is NRCF with reviews posted in the last three years (from 2015 to 2017). Among the four features proposed in phase 1, "Elite" and "Activity" belong to user information, and "Lengthiness" and "Recency" belong to reviews information.

The experimental results show that the performance could be improved in all neighborhood sizes compared to UBCF when reflecting elite status and cumulative number of reviews. All of them belong to user information. Contrariwise, it was not found that the performance was improved when reviews information was considered. Also, this result shows how much the performance can be improved. Also, the similarity of the MAE and RMSE result graphs shows that a large error between the actual and predicted ratings does not occur for a given number of neighbors.

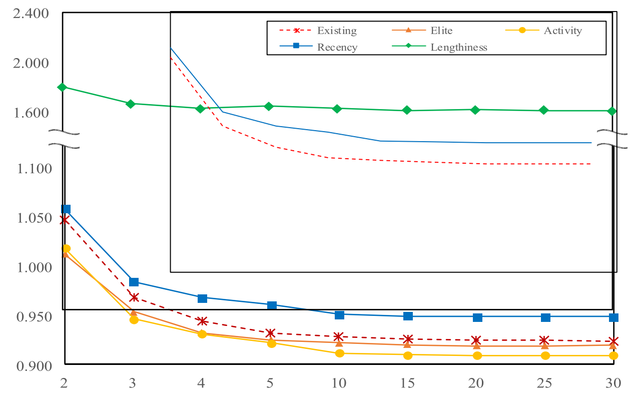

FIGURE 7. Results of phase 1 experiment(MAE).

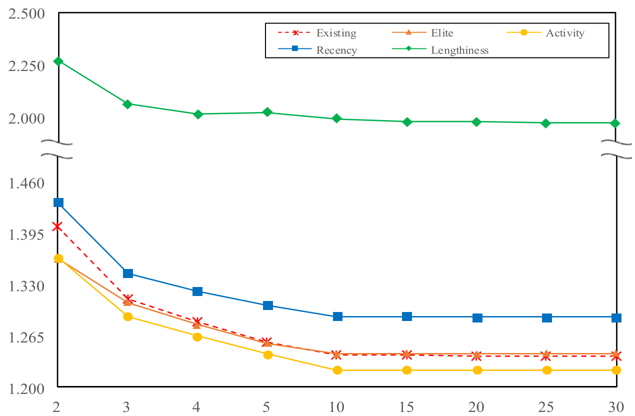

FIGURE 8. Results of phase 1 experiment(RMSE).

TABLE 3. Results of phase 1 experiment(MAE).

#NEIGHBORS	EXISTING	ELITE	ACTIVITY	RECENCY	LENGTHNIESS
2	1.047	1.012	1.018	1.058	1.800
3	0.968	0.953	0.946	0.984	1.661
4	0.944	0.932	0.931	0.968	1.624
5	0.931	0.925	0.922	0.961	1.642
10	0.929	0.922	0.912	0.951	1.626
15	0.926	0.920	0.910	0.949	1.611
20	0.925	0.919	0.909	0.948	1.613
25	0.924	0.919	0.909	0.948	1.608
30	0.924	0.920	0.910	0.948	1.607

C. PHASE 2 EXPERIMENTAL RESULT

The experiment in phase 2 demonstrated a feature that positively affects the RS performance belonging to restaurant information. The accuracy was calculated by varying the size

TABLE 4. Results of phase 1 experiment(RMSE).

#NEIGHBORS	EXISTING	ELITE	ACTIVITY	RECENCY	LENGTHNIESS
2	1.403	1.363	1.435	1.365	2.270
3	1.311	1.307	1.344	1.289	2.063
4	1.284	1.280	1.322	1.265	2.014
5	1.258	1.256	1.304	1.242	2.025
10	1.242	1.242	1.290	1.222	1.995
15	1.241	1.243	1.290	1.222	1.979
20	1.239	1.242	1.289	1.222	1.979
25	1.240	1.243	1.289	1.222	1.975
30	1.240	1.243	1.289	1.222	1.974

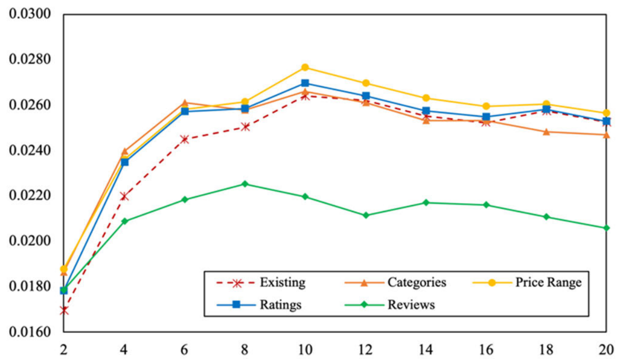

FIGURE 9. Result of phase 2 experiment(F1 Score).

of recommendations from 2 to 20. The second result is shown in Figure 9 and [Table 5]. "Existing" refers to CF in which recommendations are made in the order of VLS in the candidates. "Categories" considers restaurant categories, such as Mexican restaurants, Korean barbecue restaurants, Chinese restaurants, Japanese restaurants, and so on. It is a method of constructing a final recommendation list by classifying restaurants that the target user has visited in the past in the candidate list for the target user, and filtering the restaurants corresponding to the most visited restaurant categories. This is a method in which CB is applied to UBCF, meaning that even a restaurant with a high VLS value is excluded from the final recommendation list if it does not correspond to the category of restaurant frequented by the target user. In the same way, "Price Range" is a method of constructing a final recommended list of restaurants with the same price range as the restaurant price visited by the target user. "Ratings" is a method of constructing a final recommendation list with restaurants similar to the average rating of restaurants visited by the target user. "Reviews", refers to a method of composing a final recommended list of restaurants having a number of reviews similar to the cumulative number of reviews of restaurants visited by the target user.

As a result of the experiment, when CF considered the price range or average rating, the performance improved regardless of the size of recommendations. In the case of categories of restaurants, the performance improved when the

TABLE 5. Results of phase 2 experiment(F1).

#ITEMS	EXISTING	CATEGORIES	PRICE RANGE	RATINGS	REVIEWS
2	0.0169	0.0236	0.0188	0.0178	0.0178
4	0.0220	0.0285	0.0236	0.0235	0.0209
6	0.0245	0.0280	0.0258	0.0257	0.0218
8	0.0250	0.0275	0.0261	0.0258	0.0225
10	0.0264	0.0264	0.0276	0.0270	0.0220
12	0.0262	0.0254	0.0270	0.0264	0.0211
14	0.0255	0.0244	0.0263	0.0257	0.0217
16	0.0252	0.0235	0.0259	0.0255	0.0216
18	0.0257	0.0225	0.0260	0.0258	0.0211
20	0.0252	0.0217	0.0257	0.0253	0.0206

recommendation size was less than 10. When restaurant information is additionally reflected in the process of selecting a candidate list for recommendation, it is possible to additionally consider restaurant characteristics not included in the user's quantitative preference information, suggesting that the performance of the recommender system can be improved.

V. CONCLUSION

A. DISCUSSION

Most of the current RS studies focus on developing new recommendation algorithms for improving the performance recommendations such as accuracy or diversity. However, they have not focused on which features or aspects could improve or worsen the performance of RS. The existing CF-based RS, which aims to improve the performance, cannot explain why the recommendation was made to the user who received the recommendation list. Therefore, this study aims to measure the impact of numerous implicit information on RS and analyze it to determine which information positively affects RS.

We collected the dataset for the experiment from Yelp.com, a global review platform specializing in restaurants. The restaurants located in Austin, Texas, with the most significant number of reviews, were filtered out for the experiment. Moreover, ten years of data were used, from 2010.01.01 to 2019.12.31, before the Covid19 pandemic.

A Two Phase Experiment was designed to measure the impact of users, restaurants, and reviews information on RS performance. The first phase experiment was conducted to see the change in performance according to the features of users and reviews information. As a result, when recommendations reflect elite status and the cumulative number of reviews, the RS performance improves in all neighbors' sizes compared to traditional CF. Conversely, reviews information did not help the performance to improve.

The second phase experiment was conducted to confirm the change in performance according to the features of restaurant information. As a result, CF considering the restaurant price range and average rating, improved the performance regardless of the size of recommendations compared to traditional CF. According to the results of the two experiments, users and restaurants information were vital to improving the RS performance. On the contrary, reviews information did not affect the improvement of the RS performance. It presents that some implicit information includes additional user preferences, not included in quantitative information such as ratings or stars. The additional preferences could improve the RS performance rather than considering only quantitative information.

B. THEORETICAL CONTRIBUTIONS AND PRACTICAL IMPLICATIONS

The theoretical implications of this study are as follows.

First, to improve the performance of the recommendation algorithm, filtering only influential and meaningful data is required along with research to develop a new recommendation algorithm. However, there have been few studies on how changes in input data affect recommendation performance in the recommender system research so far. Previous studies focus on enhancing recommendation performance by developing new algorithms, but this study focused on applications based on customer behavior data and restaurant data. Prior to this research, we proposed a methodology to investigate the effect of review consistency and helpfulness on recommendation performance [38]. We focused on customers who have written helpful and consistent reviews to select influential and representative neighbors. We evaluated the performance of the proposed methodology using several real-world Amazon review data sets for experimental utility and reliability. The experimental results confirmed that the recommendation performance was excellent when a neighbor was selected who wrote consistent or helpful reviews more than when neighbors were selected for all customers. This research extended our previous research. Therefore, these studies can expand the scope of recommender system research.

Second, this study is an exploratory study trying to measure the impact on RS performance using users, restaurants, and reviews information in Yelp.com. Compared to our previous research [38] this study presents a guideline for follow-up research on the impact of distinct information on RS. A specific experimental framework is presented for follow-up researchers on how to reflect information in CF. More specifically, this study is the first study to measure the impact of users, restaurants, and reviews information at once. This study made it possible to measure and compare the impacts of information step by step by our suggested two-phase experiment framework.

These results provide e-commerce companies with the following practical implications.

First, existing restaurant information providers are developing recommender systems using all customer behavior data and restaurant information. This is because we believe that a lot of data can improve recommendation performance. However, our experiments show that too much customer, review, and restaurant-related information can reduce recommendation performance. This study investigated whether all customer behavior and restaurant related factors affected recommendation performance. Therefore, restaurant recommender system developers should consider more options. Knowing which factors in the input data affect the performance of the recommender system can provide guidance when designing future customer interfaces.

Second, a restaurant manager or individual restaurant website designer needs to know that the factors affecting the performance of the recommender system are the factors that customers consider more important. When managing customers or when designing individual restaurant homepages, it is necessary to emphasize the factors that customers consider important. Global e-commerce websites apply deep learning and artificial intelligence technologies for personalized recommendation services. However, most individual restaurant websites are challenging to apply such technologies due to development costs and lack of technical human resources. Experimental results of these studies can provide guidelines for individual restaurant website developers.

C. LIMITATIONS AND FUTURE STUDY

Since this study is an exploratory study for measuring the impact of information on RS, the following limitations exist.

First, it is impossible to verify any other information through this study not provided by Yelp.com. For example, this study could not verify whether a user's demographic features could affect RS performance. Therefore, it is necessary to conduct follow-up studies using others from different platforms, such as Open Table and Zomato, not included in Yelp.com.

Second, the experimental results cannot be generalized since the experimental subject is limited to Austin, Texas. According to existing studies, there are distinct differences in food culture between countries and races. Therefore, additional experiments should be conducted on whether the same features affect RS in other countries and regions.

Third, the features of information exploited in this study are restricted. For example, this study did not cover various facilities or services in restaurants, such as allowance for pets, acceptance of credit cards, and so forth. Therefore, it is necessary to examine features that influence RS among numerous additional facilities in future studies.

Lastly, recent studies have shown that customers are more receptive to recommendations by adding an explanation function to the recommender system [39], [40]. When designing an explanatory recommender system platform, it is necessary to study whether customers will accept the recommendation if an explanation function is added focusing on important factors affecting the importance of the recommender system.

REFERENCES

[1] H. K. Kim, Y. U. Ryu, Y. Cho, and J. K. Kim, "Customer-driven content recommendation over a network of customers," *IEEE Trans. Syst., Man, Cybern. A, Syst. Humans*, vol. 42, no. 1, pp. 48–56, Jan. 2011.

[2] Y. H. Cho and J. K. Kim, "Application of web usage mining and product taxonomy to collaborative recommendations in e-commerce," *Expert Syst. Appl.*, vol. 26, no. 2, pp. 233–246, Feb. 2004.
[3] J. Bennett and S. Lanning, "The Netflix prize," in *Proc. KDD Cup Workshop*, San Jose, CA, USA, 2007, pp. 3–6.
[4] A. S. Das, M. Datar, A. Garg, and S. Rajaram, "Google news personalization: Scalable online collaborative filtering," in *Proc. 16th Int. Conf. World Wide Web (WWW)*, Banff, AB, Canada, 2007, pp. 271–280.
[5] G. Linden, B. Smith, and J. York, "Amazon. com recommendations: Item-to-item collaborative filtering," *IEEE Internet Comput.*, vol. 7, no. 1, pp. 76–80, Jan. 2003.
[6] C. Zhang, H. Zhang, and J. Wang, "Personalized restaurant recommendation method combining group correlations and customer preferences," *Inf. Sci.*, vols. 454–455, pp. 128–143, Jul. 2018.
[7] F. Zhang, N. J. Yuan, K. Zheng, D. Lian, X. Xie, and Y. Rui, "Exploiting dining preference for restaurant recommendation," in *Proc. 25th Int. Conf. World Wide Web*, Montréal, QC, Canada, Apr. 2016, pp. 725–735.
[8] C. Aggarwal, "Neighborhood-based collaborative filtering," in *Recommender Systems: The Textbook*. New York, NY, USA: Springer, 2016, pp. 29–70.
[9] J. Castro, F. J. Quesada, I. Palomares, and L. Martínez, "A consensus-driven group recommender system," *Int. J. Intell. Syst.*, vol. 30, no. 8, pp. 887–906, Apr. 2015.
[10] X. Miao, Y. Gao, G. Chen, H. Cui, C. Guo, and W. Pan, "SI2P: A restaurant recommendation system using preference queries over incomplete information," in *Proc. 42nd Int. Conf. Very Large Data Bases*, New Delhi, India, 2016, pp. 1509–1512.
[11] G. Adomavicius and A. Tuzhilin, "Context-aware recommender systems," in *Recommender Systems Handbook*. New York, NY, USA: Springer, 2011, pp. 217–253.
[12] E. Asani, H. Vahdat-Nejad, and J. Sadri, "Restaurant recommender system based on sentiment analysis," *Mach. Learn. with Appl.*, vol. 6, Dec. 2021, Art. no. 100114.
[13] A. Danylenko and W. Lowe, "Context-aware recommender systems for non-functional requirements," in *Proc. 3rd Int. Workshop Recommendation Syst. Softw. Eng. (RSSE)*, Zurich, Switzerland, Jun. 2012, pp. 80–84.
[14] S. Kulkarni and S. F. Rodd, "Context aware recommendation systems: A review of the state of the art techniques," *Comput. Sci. Rev.*, vol. 37, Aug. 2020, Art. no. 100255.
[15] X. Zhang, H. Luo, B. Chen, and G. Guo, "Multi-view visual Bayesian personalized ranking for restaurant recommendation," *Appl. Intell.*, vol. 50, no. 9, pp. 2901–2915, Apr. 2020.
[16] M. Deac-Petrusel and S. Limboi, "A sentiment-based similarity model for recommendation systems," in *Proc. 22nd Int. Symp. Symbolic Numeric Algorithms Sci. Comput. (SYNASC)*, Sep. 2020, pp. 224–230.
[17] A. Sarkar, A. Baksy, and V. Kirpalani, "Analysis of zomato services using recommender system models," in *Proc. Int. Conf. Intell. Technol. (CONIT)*, Karnataka, India, Jun. 2021, pp. 1–5.
[18] Zeng, F. Li, H. Liu, J. Wen, and S. Hirokawa, "A restaurant recommender system based on user preference and location in mobile environment," in *Proc. 5th IIAI Int. Congr. Adv. Appl. Informat. (IIAI-AAI)*, Kumamoto, Japan, 2016, pp. 55–60.
[19] R. Gupta, I. Pandey, K. Mishra, and K. Seeja, "Recommendation system for location-based services," in *Applied Information Processing Systems*. New York, NY, USA: Springer, 2022, pp. 553–561.
[20] S. Zhang, M. Salehan, A. Leung, I. Cabral, and N. Aghakhani, "A recommender system for cultural restaurants based on review factors and review sentiment," in *Proc. 24th Americas Conf. Inf. Syst.*, New Orleans, LA, USA, 2018, pp. 1–5.
[21] S. Sawant, "Collaborative filtering using weighted bipartite graph projection: A recommendation system for yelp," in *Proc. Social Inf. Netw. Anal. Conf.*, Los Angeles, CA, USA, vol. 38, 2013.
[22] G. Bathla, P. Singh, S. Kumar, M. Verma, D. Garg, and K. Kotecha, "Recop: Fine-grained opinions and sentiments-based recommender system for industry 5.0," *Soft Comput.*, vol. 2021, pp. 1–10, Dec. 2021.
[23] S. Zhao, S. Han, R. Meng, D. He, and D. Zhang, "Learning semantic representation from restaurant reviews: A study of yelp dataset," in *Proc. iConf.*, Wuhan, China, 2017, pp. 1–4.
[24] F. Eskandanian, B. Mobasher, and R. Burke, "A clustering approach for personalizing diversity in collaborative recommender systems," in *Proc. 25th Conf. User Modeling, Adaptation Personalization*, Bratislava, Slovakia, Jul. 2017, pp. 280–284.
[25] L. R. Divyaa, A. Tamhane, and N. Pervin, "A clustering based social matrix factorization technique for personalized recommender systems," in *Proc. Amer. Conf. Inf. Syst. (AMCIS)*, New Orleans, LA, USA, 2018, pp. 1–10.
[26] M. Naumov, D. Mudigere, H.-J. M. Shi, J. Huang, N. Sundaraman, J. Park, X. Wang, U. Gupta, C. J. Wu, A. G. Azzolini, and D. Dzhulgakov, "Deep learning recommendation model for personalization and recommendation systems," May 2019, *arXiv:1906.00091*.
[27] D. Peng, S. J. Pan, J. Zhang, and A. Zeng, "Learning an adaptive meta model-generator for incrementally updating recommender systems," in *Proc. 15th ACM Conf. Recommender Syst.*, New York, NY, USA, Sep. 2021, pp. 411–421.
[28] Y. Ji, A. Sun, J. Zhang, and C. Li, "A critical study on data leakage in recommender system offline evaluation," *ACM Trans. Inf. Syst.*, vol. 2022, pp. 1–27, Oct. 2022.
[29] D. Paraschakis, "Recommender systems from an industrial and ethical perspective," in *Proc. 10th ACM Conf. Recommender Syst.*, Boston, MA, USA, Sep. 2016, pp. 463–466.
[30] J. Bobadilla, F. Ortega, A. Hernando, and A. Gutiérrez, "Rec-ommender systems survey," *Knowl.-Based Syst.*, vol. 46, pp. 109–132, Jul. 2013.
[31] H. Al-bashiri, M. A. Abdulgabber, A. Romli, and H. Kahtan, "An improved memory-based collaborative filtering method based on the TOPSIS technique," *PLoS ONE*, vol. 13, no. 10, Oct. 2018, Art. no. e0204434.
[32] M. Elahi, F. Ricci, and N. Rubens, "A survey of active learning in collaborative filtering recommender systems," *Comput. Sci. Rev.*, vol. 20, pp. 29–50, May 2016.
[33] J. Kim, I. Choi, and Q. Li, "Customer satisfaction of recommender system: Examining accuracy and diversity in several types of recommendation approaches," *Sustainability*, vol. 13, no. 11, p. 6165, May 2021.
[34] D. H. Park, H. K. Kim, I. Y. Choi, and J. K. Kim, "A literature review and classification of recommender systems research," *Expert Syst. Appl.*, vol. 39, no. 11, pp. 10059–10072, Sep. 2012.
[35] X. Su and T. M. Khoshgoftaar, "A survey of collaborative filtering techniques," *Adv. Artif. Intell.*, vol. 2009, pp. 1–19, Aug. 2009.
[36] W. B. Frakes and R Baeza-Yates, *Information Retrieval: Data Structures and Algorithms*. Hoboken, NJ, USA: Prentice-Hall, 1992.
[37] G. G. Chowdhury, *Introduction to Modern Information Retrieval*. New York, NY, USA: McGraw-Hill, 1983.
[38] J. Park, X. Li, Q. Li, and J. Kim, "Impact on recommendation performance of online review helpfulness and consistency," *Data Technol. Appl.*, vol. 2022, pp. 1–23, Sep. 2022.
[39] J. B. Daher, A. Brun, and A. Boyer, "A review on explanations in recommender systems," Ph.D. dissertation, Dept. Complex Syst., Artif. Intell. Robot., Univ. Lorraine, Leopold, Nancy, 2017.
[40] J. Vig, S. Sen, and J. Riedl, "Tagsplanations: Explaining recommendations using tags," in *Proc. 14th Int. Conf. Intell. User Interfaces*, Orlando, FL, USA, Feb. 2009, pp. 47–56.

SEUNGWOO LEE received the B.S. degree in management from Kyung Hee University, Seoul, South Korea, in 2021, where he is currently pursuing the M.S. degree in big data analytics. His research interests include the development of explainable recommender systems, bigdata analysis with hospitality industry, and natural language processing.

HYUNJOO SHIN received the Master of Business Administration (M.B.A.) degree and the Ph.D. degree in business administration (major in management information systems) from Kyung Hee University. He is currently the CEO of Lexia Company Ltd. His research interests include big data analysis, machine learning, and recommender systems, which have been published in the *Asia Pacific Journal of Information Systems*.

ILYOUNG CHOI received the B.S. degree in economics and the M.S. and Ph.D. degrees in MIS from Kyung Hee University. He is currently a Researcher Fellow at BK21 Four Program Educational Research Group, Kyung Hee University. His current research interests include personalized service, big data analysis, social network analysis, and business analytics. He has published numerous papers, which have appeared in *Information Technology and Management*, *International Journal of Information Management*, and *Online Information Review*.

JAEKYEONG KIM received the B.S. degree in industrial engineering from Seoul National University, Seoul, South Korea, and the M.S. and Ph.D. degrees in industrial engineering from the Korea Advanced Institute of Science and Technology, Seoul. From 2006 to 2007, he was a Visiting Scholar at the School of Management, The University of Texas at Dallas. He is currently a Professor at the School of Management and the Dean of the Department of Big Data Analytics, Kyung Hee University, Seoul. He is also the Head of the BK21 Four Program Educational Research Group. His research interests include recommender systems, big data analytics, machine learning, and deep learning applications. He has participated in several industrial projects with high-technology companies and nonprofit research and development centers in Korea. He has published more than 70 papers in international articles, such as IEEE Transactions on Systems, Men, and Cybernetics, IEEE Transactions on Services Computing, *International Journal of Information Management*, *Journal of Electronic Commerce Research*, *Information & Managemen*, and *International Journal of Information Management*.