

Article

A BERT-Based Multi-Criteria Recommender System for Hotel Promotion Management

Yuanyuan Zhuang and Jaekyeong Kim *

School of Management & Department of Big Data Analytics, KyungHee University, 26, Kyungheedaero, Dongdaemun-gu, Seoul 02453, Korea; zhuangyy1122@gmail.com

* Correspondence: jaek@khu.ac.kr

Abstract: Numerous reviews are posted every day on travel information sharing platforms and sites. Hotels want to develop a customer recommender system to quickly and effectively identify potential target customers. TripAdvisor, the travel website that provided the data used in this study, allows customers to rate the hotel based on six criteria: Value, Service, Location, Room, Cleanliness, and Sleep Quality. Existing studies classify reviews into positive, negative, and neutral by extracting sentiment terms through simple sentimental analysis. However, this method has limitations in that it does not consider various aspects of hotels well. Therefore, this study performs fine-tuning the BERT (Bidirectional Encoder Representations from Transformers) model using review data with rating labels on the TripAdvisor site. This study suggests a multi-criteria recommender system to recommend a suitable target customers for the hotel. As the rating values of six criteria of TripAdvisor are insufficient, the proposed recommender system uses fine-tuned BERT to predict six criteria ratings. Based on this predicted ratings, a multi-criteria recommender system recommends personalized Top-N customers for each hotel. The performance of the multi-criteria recommender system suggested in this study is better than that of the benchmark system, a single-criteria recommender system using overall ratings.

Keywords: recommender system; multi-criteria recommender systems; BERT; natural language processing; review data; text mining; hotel promotion



Citation: Zhuang, Y.; Kim, J. A BERT-Based Multi-Criteria Recommender System for Hotel Promotion Management. *Sustainability* **2021**, *13*, 8039. <https://doi.org/10.3390/su13148039>

Academic Editors:
Elena Carvajal-Trujillo and
David Castilla-Espino

Received: 24 June 2021
Accepted: 15 July 2021
Published: 19 July 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The tourism industry has brought about many changes due to the development of information technology and the Internet. Today, all categories of hotels employ online travel agents (OTAs) or booking platforms to diversify their sales channels and reach out to more potential customers [1]. Online travel agencies have numerous hotels registered, and numerous reviews are posted every day, resulting in information overload, which puts pressure on customers to make a choice. To solve these problems and provide customers with better services, hotel recommender systems have been introduced by major travel agencies, thereby reducing the user's decision-making time and efforts [2].

To enhance personalization capabilities, recommender systems are widely applied in many multimedia platforms targeting media products to specific customers [3]. Due to the recent overflow of recommender systems, many customers treat non-detailed, non-personalized recommendation services like old spam emails. Therefore, from the point of view of the hotel, it is necessary to accurately identify and promote to the customers who may visit the hotel. From the customer's point of view, rather than receiving promotions from numerous hotels, they want to be recommended only from hotels that are appropriate for the customer. Therefore, through personalized recommendation, it is possible to effectively promote the hotel through the recommendation of available customers at the hotel, as well as to increase the customer's order rate and to help increase the recognition and credibility of the hotel.

Most shopping sites encourage users to write review text for products they purchase. Such review information is very useful for understanding users' preferences and items' characteristics and enhances the personalized recommendation ability of a website [4]. Likewise, user-generated content, especially online hotel reviews, is a potentially rich source of customer information regarding opinions and sentiment [5]. From the decision-making perspective, the display of user-generated reviews improves both the awareness of the hotel and the attitudes towards the hotel when forming an opinion [6,7]. And the use of review text has been shown to improve the accuracy of rating prediction for users and items, especially when there are few ratings [8–10]. Therefore, when the rating is insufficient, the cold start problem can be alleviated by using the users' preference in the review and the characteristics of the item, and it is possible to provide an explanatory recommendation result to the users [11].

Accordingly, more and more studies are being conducted to apply review text analysis to the recommender system field. Text analysis is an important area in the field of natural language processing (NLP). In order to overcome various limitations of natural language processing, the recently announced BERT [12] or T5 [13] is a pre-trained transformer for language modeling [14]. These pre-trained language models have shown very strong performance in NLP as well as information retrieval, recommender systems, dialogue systems and other fields [14]. BERT, a pre-trained language model in Deep Learning Algorithms, has been widely studied and applied, and has achieved state-of-the-art performance in many NLP tasks [12,15,16].

Recently, many recommender systems are using review data for recommendation [17]. However, although there are many studies on Opinion Mining, which extracts more accurate opinions or emotions from reviews, there are not many studies on how to predict more accurate ratings using reviews in collaborative filtering (CF) [18]. In addition, many existing review-based recommender systems simply classify reviews into positive, negative, and neutral through sentimental analysis, and recommend based on this. This is regarded as a recommender system using a single criteria using text. The recommender system based on such a single criteria reflects the user's overall evaluation of the product, but does not reflect the preference for the detailed aspect of each product. That is, the sensibility of each attribute expressed by the users in the review or the specific characteristics of the product are ignored.

Since online reviews generally contain multiple opinions on multiple aspects, some recent studies have begun to predict detailed attribute ratings instead of one overall rating [19]. Moreover, more and more sites now allow users to rate products in multiple dimensions, so it becomes necessary to develop recommender systems which utilize this type of additional rating information to potentially improve the accuracy of recommendations. Multi-criteria based CF can provide accurate recommendations by considering users' preferences in various aspects, which can be a more appropriate choice for users [20]. The multi-criteria recommender system can ensure a more sophisticated understanding of the user's preferences by considering the knowledge of fundamental properties that induce the users to select a specific item [21]. Therefore, additional information about each user's preference can help to accurately model the user's preference, which can lead to more accurate recommendations [22].

As shown in Figure 1, on TripAdvisor. Available online: <https://www.tripadvisor.ca/> (accessed on 1 November 2020), the one of the largest world's best travel website, customers can rate a hotel based on six criteria: Value, Service, Location, Rooms, Cleanliness, and Sleep Quality. Therefore, in the case of TripAdvisor, the use of additional information including these customers preferences can help to accurately model customers preferences and lead to more accurate recommendations [22].

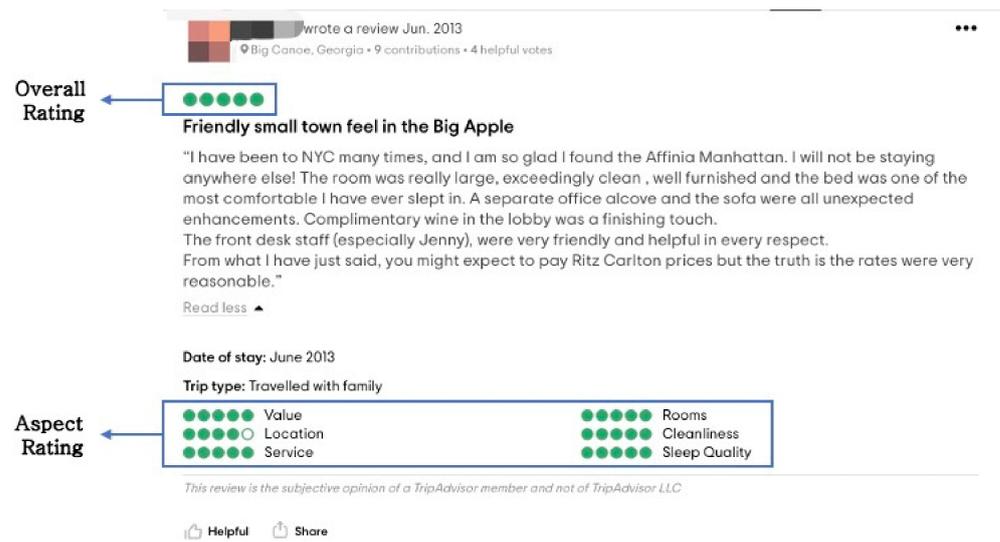


Figure 1. Example of hotel review on TripAdvisor.

Our observations on TripAdvisor show that most customers have written an overall rating and review of a hotel, but either rate only a few of the six aspect ratings or none of the six attributes. Due to this, a cold start problem may occur due to insufficient customers' attribute rating data, and accurate recommendations cannot be made. Therefore, a model that can predict an overall rating or six aspect ratings is needed.

The research purpose of this study is twofold. The first is to develop a model that can predict an overall rating or six aspect ratings based on review data and solve the problem of insufficient attribute rating data. The second is to develop a multi-criteria recommender system for hotels based on the predicted multi aspect ratings, not only to improve the recommendation performance, but also to increase the hotel's promotional efficiency.

This study performs fine-tuning the BERT model using review data with rating labels on the TripAdvisor site. We use this model to predict the overall rating and six attribute ratings from reviews. Top-N customers with the highest rating are recommended to the hotel through multi-criteria collaborative filtering (CF) using the rating predicted by the BERT model. The experimental results showed that there is insignificant difference between the performance of the single criteria recommender system using the overall rating value estimated by suggested BERT model and the performance of the recommender system using the overall rating value input by users. In addition, the performance of the proposed multi-criteria recommender system is better than that of the single-criteria recommender system. The hit ratio of the multi-criteria recommender system is improved to 6.19%, and NDCG improves to 7.08% compared to single criteria benchmark system.

2. Related Work

2.1. Multi-Criteria Recommender Systems

Recommender systems assist customers to find information or products they need among an overwhelming number of possibilities [23–26]. Collaborative filtering (CF) is one of the most successful methods in the recommender system and uses the past preferences of a group of users to recommend products or predict the preferences of other users [27]. In [28], proposed an item-based collaborative filtering (Item-base CF) recommendation algorithm, which identifies relationships between different items by analyzing a user-item matrix.

In addition, consumer reviews of products, namely reviews, opinions and shared experiences, are powerful sources of information about consumer preferences and can be used in recommender systems [29]. Therefore CF-based recommender systems in recent years has resulted in a paradigm shift, moving away from systems that are solely based on the ratings' matrix to systems that incorporate user generated free-text reviews in

the recommendation process as well [30]. However, most existing methods focus on the word or phrase level in the review, extract emotional terms or emotional phrases through simple sentiment analysis, and classify the reviews as positive, negative or neutral. These methods often fail to capture the whole context of the reviews, and cannot fully understand what the reviewer wants to express. Therefore, many recommender systems use context-independent embeddings methods when analyzing reviews, such as Word2Vec, Paragraph Vectors, etc.

Baek & Chung [31] proposed the multimedia recommendation method using Word2Vec-based social relationship mining. This is to analyze users with a similar tendency on the basis of the keywords related to multimedia content and sentiment words of comments, to build a trust relationship, and to recommend multimedia. Alexandridis et al. [32] proposes a recommender system named ParVecMF, namely a paragraph vector-based matrix factorization recommender system. The paragraph vector model [33] used in this study is an extension of the Word2Vec model, which presents a distributed representation of words in vector space [34]. The use of the Paragraph Vectors model permits the discovery of similarity in context of documents that use different words [32]. In this study, a novel approach of combining user reviews, in the form of neural embeddings, and ratings in probabilistic matrix factorization has been presented. Alexandridis et al. [30] present a new technique of incorporating reviews into collaborative filtering matrix factorization algorithms. The important contribution of this study is among the first to effectively account for word order & context, as well as document context at the same time, through the combination of paragraph vectors and CF matrix factorization in a unified learning approach.

In the respect of recommendation process, multi-attribute ratings can provide more information about users' preferences and products in various aspects than overall ratings, which represents the user's opinion on the entire item [35]. The multi-criteria recommender system based on multi-attribute ratings can ensure a more sophisticated understanding of the user's preferences by considering the knowledge of the fundamental properties that induce the users to select a specific item [21].

The rating function in the multi-criteria recommender system is defined as follows.

$$Users \times Items \rightarrow R_0 \times R_1 \times \dots \times R_k, \quad (1)$$

where R_0 is the overall rating and R_i is the rating for each attribute criteria i ($i = 1, \dots, k$) [36].

A lot of research has been done on multi-criteria recommender systems so far. In [36], two new recommendation techniques are proposed for multi-criteria rating systems, a similarity-based approach and aggregation-function-based approach. In similarity-based method, it can be divided into a method of counting the traditional similarity from a single criteria and a method of calculating the similarity using Multidimensional Distance Metrics. The accuracy of this multi-criteria recommendation method is at least comparable to or better than that of the single-criteria recommender system [36]. Therefore, in this study, a multi-criteria recommender system is used in the recommendation process based on the similarity-based method in [36]. Nie et al. [37] proposes a method to automatically predict the weights of various aspects when constructing the overall rating using the Tensor Factorization method, and the main idea is to use Constrained Optimization to predict users, items and aspects. Wang et al. [18] uses the movie domain as a case study to capture users' opinions on various attributes in the review text, and propose a framework that can use that information to increase the effectiveness of CF. Their recommendation process is carried out by predicting ratings through opinion mining, that is, extracting attribute terms and opinions.

Most of these multi-criteria recommender system studies focus on improving the recommendation accuracy through multi-criteria recommendation based on the ratings evaluated by users. A limitation of these existing studies is known that recommendations cannot be made if the ratings are insufficient. For TripAdvisor, these methods are not applicable because six attribute rating data is sparse, so it is necessary to predict six

attribute ratings to provide good recommendation. Another limitation is that simply extracting the attribute term from the review text using the opinion mining method is not helpful for recommendation. The opinion mining method does not consider context and cannot accurately understand user preferences, and thus may reduce the accuracy of rating prediction. Therefore, unlike previous studies, this study focuses on improving the recommendation accuracy through BERT-based predicting the multi-criteria rating values by analyzing the context in the review text well, and multi-criteria recommender system which accurately understand the user's preference.

2.2. BERT

Natural language processing methods include various tasks such as machine translation [38], question answering [39], and sentiment analysis [40]. In recent years, pre-trained models such as ELMo [41], BERT [12], and GPT-3 [42] perform fine-tuning after a large amount of text pre-training and NLP performance greatly improves.

BERT (Bidirectional Encoder Representations from Transformers) is an NLP model, designed to perform fine tuning using labeled text for various NLP tasks after pre-training deep bidirectional representation from unlabeled text [12]. BERT is pre-trained in a large corpus. For the pre-training corpus, BERT model use the BooksCorpus (800 M words) [43], and English Wikipedia (2500 M words) [12]. The success of BERT on NLP tasks lies mainly in the English language domain, as the main BERT models is trained on English [12]. The BERT model is one of the most popular models in the recent NLP field. The BERT model is mainly divided into two stages: pre-training and fine-tuning [12]. Pre-training mainly consists of two unsupervised tasks: Masked language model (MLM) and Next sentence prediction (NSP). In fine tuning, Transformer's Self-Attention mechanism allows BERT to model multiple downstream tasks by replacing appropriate inputs and outputs. When fine-tuning, we first initialize the BERT model with pre-trained parameters, and then use all parameters for end-to-end fine-tuning. Fine-tuned BERT can be used in downstream operations such as summary and relation extraction.

In this study, we decided to choose New York City as a research area. This is one of most international destinations in the world. Thus, there is high probability that reviews posted on TripAdvisor include those which are delivered in English by native speakers, in English by non-native speakers, and in other languages. Detection of sentiments in multilingual environment is extremely complex. Even using one language like English by both of mentioned groups (native speakers, and non-native speakers) in analysis is complicated. Therefore, this study uses the BERT model because this is an NLP model that is known to achieve the most advanced performance [12]. In addition, the BERT model is helpful to better analyze the context of the review text and predict more accurate attribute ratings after analyzing user preferences. The BERT model has been pre-trained in a large number of corpora, so we have reason to believe that the BERT model can solve the emotion detection problem in a multilingual environment.

3. A Multi-Criteria Customers Recommender System

This study suggests a multi-criteria customer recommender system with fine tuned BERT, which predicts the six-criteria ratings (Value Rating, Service Rating, Location Rating, Room Rating, Cleanliness Rating, and Sleep Quality), and overall rating from the review data in travel website.

The proposed model consists of three stages: 'data collection', 'BERT fine tuning', and 'multi-criteria recommendation', as shown in Figure 2.

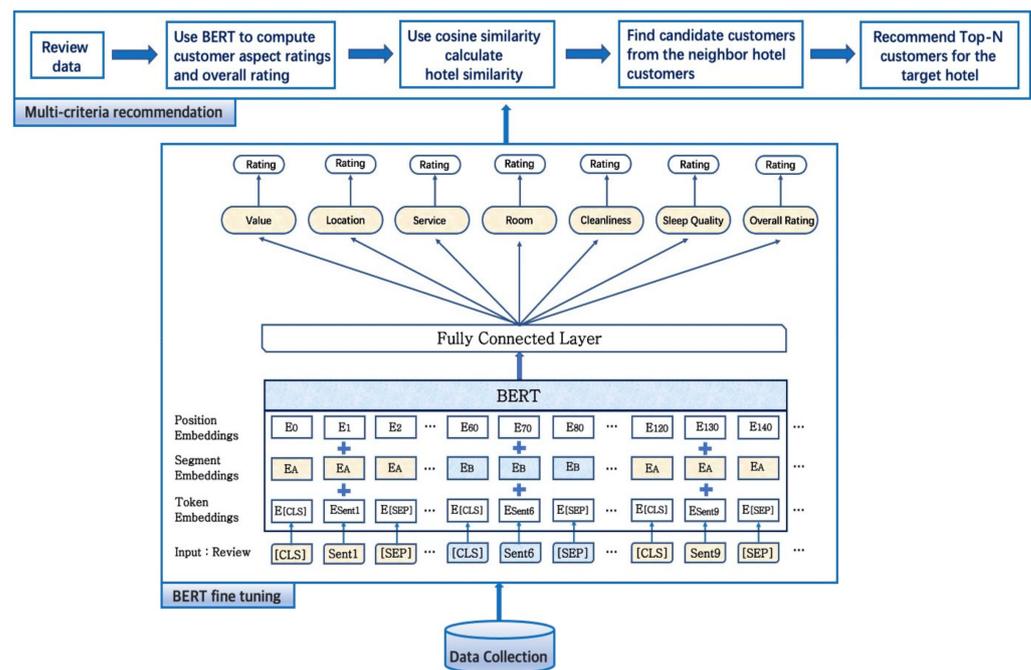


Figure 2. A multi-criteria recommender system model.

3.1. Data Collection

In order to build the multi-criteria recommender system, reviews of 4-star and 5-star hotels within 3 km of Central Park in New York City, USA were manually collected on TripAdvisor website. TripAdvisor is the world’s largest travel site, and it provides overall ratings, and 6 attribute ratings (Value, Service, Location, Room, Cleanliness, and Sleep Quality) per each hotel. The overall rating and the six attribute ratings consist of a five-point scale. The collected data set is summarized on the following Table 1.

Table 1. Data set summary.

Data	Sum
Review	132,744
User	103,075
Hotel	63
Overall rating	132,744
Value rating	66,238
Location rating	65,124
Service rating	90,485
Rooms rating	65,367
Cleanliness rating	66,450
Sleep quality rating	60,553

The collected data set is divided into two parts. When fine-tuning the BERT model, we need labeled data, so we use review data including one or more attribute ratings out of six attribute ratings. Review data that does not include any of the remaining six attribute ratings is used for recommendations with the predicted rating value of fine-tuned BERT model.

3.2. Fine Tuning BERT Model

In order to apply the BERT model to the rating prediction task, fine tuning is performed by introducing a fully connected layer in the final hidden state corresponding to the [CLS] input token according to the method in [12]. Regression is performed in 7 dimensions at the same time on the input reviews to calculate the final predicted ratings P1 to P7. That

is, it predicts the overall rating and six attribute ratings such as Value, Service, Location, Rooms, Cleanliness, and Sleep Quality.

After performing the preprocessing procedure, the data is split and transmitted to the BERT model to train the model. The data input to the BERT model are Token, Mask, and Rating. Here, the token is a review after encoding, and the length of the token does not exceed 512. The mask is divided into 0 and 1, where 1 is the unmasked token, i.e., the original review token, and 0 is the masked token, i.e., less than 512 tokens are [PAD] filled tokens. The rating is a label, namely the six attribute ratings and the overall rating.

The fine-tuning process consists of three parts as shown in Figure 3. The first part is the Embedding Layer. In this layer, the review is encoded using a pre-trained 'bert-base-uncased' model, and the outputs are the Final Hidden State and the Pooler Output. The second part is the Rectified Linear Unit (ReLU), where the ReLU is an activation function commonly used in artificial neural networks. The output from the embedding layer is passed to the ReLU. The third part is the fully connected layer. In order to predict the rating, the output of the ReLU layer is transmitted to the fully connected layer, and finally 7 ratings such as P1 to P7 are output.

Finally, the loss value of the model is calculated using the MSE (Mean Square Error) loss function.

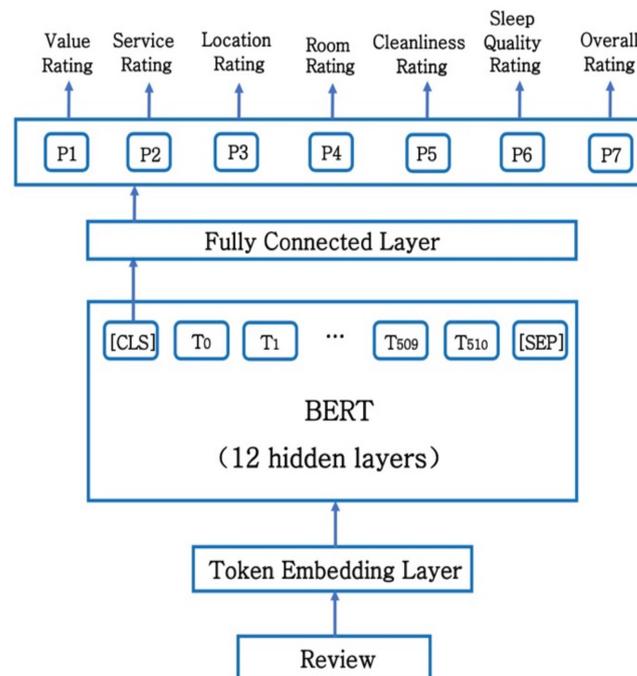


Figure 3. BERT Fine Tuning Process.

3.3. Multi-Criteria Recommendation Process

The customer recommendation process is divided into five steps as shown in Figure 4. The method proposed in this study is called MC-CF (BERT). Here, MC means Multi-Criteria, and CF means that the entire process is based on the CF method. And the six aspect ratings and the overall rating are all predicted by the fine-tuned BERT.

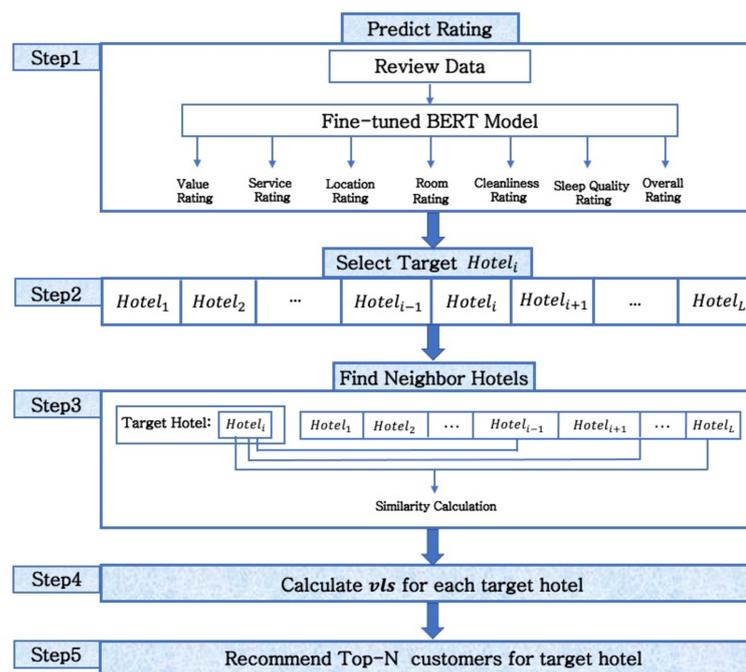


Figure 4. Recommendation process (In this figure, i is the target hotel, the range of i is $1 \leq i \leq L$ and L is the total number of hotels, vls is the visiting likelihood score.)

The first step is to collect review data and pre-process it. In the collected review data, users often do not give ratings for six attributes. Therefore, we predict six attribute ratings and overall rating using a fine-tuned BERT model.

The second step is to select a target hotel for customer promotion. In this study, in order to evaluate the proposed system and benchmark system, all 62 hotels are selected as target hotels in turn.

The third step is to find similar neighbor hotels for each target hotel. The method of finding neighbor hotels is to calculate the similarity between the target hotel and the rest of the hotels, and then select k hotels with the highest similarity as neighbor hotels. In this study, cosine similarity as shown in Equation (2) below is calculated using six attribute ratings and overall rating. Since data were collected from a total of 62 hotels in this study, experiments are performed to measure the accuracy of recommendation while changing k , the number of neighbor hotels from 2 to 10 in order.

$$\text{sim}(i, j) = \cos(\vec{i}, \vec{j}) = \frac{\vec{i} \cdot \vec{j}}{\|\vec{i}\|_2 \|\vec{j}\|_2}, \quad (2)$$

Here, ' \bullet ' indicates vector dot-product operation. And i is the target hotel, and j is the other hotel. The range of cosine similarity values is between -1 and 1 . The closer the value is to 1 , the more similar the two hotels are, and the closer the value is to -1 , the less similar the two hotels are.

The fourth step is to calculate the likelihood of customers in neighbor hotels to visit the target hotel. The visiting likelihood score is calculated as shown in Equation (3)

$$vls(i, h) = \frac{\sum_{j \in N_h} p_{ij} \text{Sim}(h, j)}{\sum_{j \in N_h} \text{Sim}(h, j)}, \quad (3)$$

Here, $vls(h, i)$ means the visiting likelihood score of customer h for target hotel i , j means neighbor hotel, and N_h means is the hotel h 's neighbor hotel set. p_{hj} is 1 if customer h has visited neighbor hotel j , otherwise it is 0.

In the fifth step, Top-N customers are recommender for the target hotel, where N is set from 5 to 15 in our experiments. That is, the Top-N customers with the highest visiting likelihood score for each target hotel are recommended.

And SC_CF(BERT) and SC_CF were proposed in this study as benchmark systems for comparison with the proposed methodology. For SC_CF(BERT), the rest of the process is the same, except that only the overall rating value predicted by BERT is used. The reason for introducing SC_CF (BERT) as a benchmark system is to compare the performance of multi-criteria recommendation process and that of single criteria recommendation process. SC_CF was introduced as another benchmark system, which uses the hotel overall rating entered by the customer directly, and the rest of the process are the same as MC_CF(BERT). Since item-based CF was used in this study, SC_CF corresponds to a general CF benchmark system.

4. Experiments and Results

4.1. BERT Finetuning

The BERT model is fine-tuned using labeled data, that is, a total of 90,950 review data containing one or more attribute ratings out of six attribute ratings. The review data set is divided into a training set, a test set and a development set at a ratio of 8:1:1.

The preprocessing is divided into 5 steps. First, because there are many missing values among the six attribute ratings, the missing values are indicated as -1 . Second, all reviews are separated into sentences based on the period. Third, all reviews are encoded using the encoder included in the BERT model. Fourth, a [CLS] tag is inserted before each sentence, and a [SEP] tag is inserted after each sentence. Finally the input token length is set to 512. BERT accepts only up to 512 tokens as input and outputs a sequence representation [44]. Therefore, according to the method of truncating the text in [44], only the preceding 512 tokens (including [CLS] and [SEP] tokens) are reserved, and when there are less than 512 tokens, it is filled with [PAD] tokens, that is, Padding Tokens.

In our experiments, we use Python's Pytorch framework and the case-insensitive 'bert-base-uncased' model. The number of epochs is 5, the Adam optimizer is used, the learning rate is $2e-5$, the model is trained using the mini-batch method, and the batch size is set to 16. The hidden size is 768, the maximum token embedding is 512, the number of attention heads is 12, and the number of hidden layers is 12.

BERT fine tuning results are as shown in Figure 5. At Epoch 1, the Train Loss and Test Loss is 0.324324, and 0.363528 respectively. At Epoch 5, Train Loss and Test Loss is 0.186292, and 0.342638 respectively. At Epoch 6, the loss values drops to 0.171123, but the Test Loss is increased to 0.346432, that is, overfitting problem occurs, so fine tuning is stopped at epoch 5.

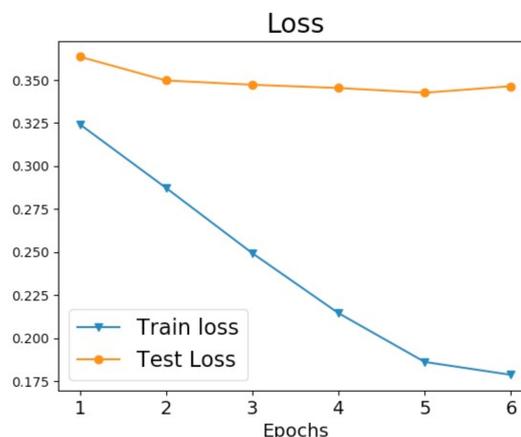


Figure 5. BERT Fine Tuning Loss Value.

Tables 2 and 3 are examples of ratings predicted by the BERT model. The score in the table is the actual rating evaluated by the customers, and the Predicted Score is the rating predicted by BERT. −1 means a missing value. From Tables 2 and 3, you can see that the rating value predicted by fine-tuned BERT based on the review data is quite accurate.

Table 2. Example 1 of rating predicted by BERT model.

	Value	Service	Location	Room	Cleanliness	Sleep Quality	Overall Rating
Score	−1	−1	5	5	−1	5	5
Predicted Score	4.7221	4.9681	4.8926	4.9356	4.9586	4.9460	4.9188
Actual Review Example	[CLS] I have stayed in a number of new york hotels in recent years [SEP] the london was a new discovery, and definitely the best [SEP] it is excellent value, particularly for a stay of a week or more [SEP] lovely large quiet rooms and really helpful service [SEP] the location is outstanding too [SEP]						

Table 3. Example 2 of rating predicted by BERT model.

	Value	Service	Location	Room	Cleanliness	Sleep Quality	Overall Rating
Score	3	4	3	2	3	2	2
Predicted Score	2.9435	4.0990	3.3371	2.3204	3.1560	1.9090	2.2416
Actual Review Example	[CLS] had to change rooms due to a ceiling leak [SEP] bed was drenched [SEP] moved me at midnight to a low floor and heard every noise nyc has to offer [SEP] price is good for the location [SEP] get a high floor and don't expect compensation if you have a problem [SEP]						

4.2. Experimental Design

The collected data set used for recommendation contains user names, hotel names, and user reviews. It contains 41,794 reviews, 35,410 users, and 63 hotels. In order to improve the accuracy of recommendations, users and hotels with fewer than 5 reviews were deleted. The final data set contains 2279 reviews, 340 users, and 62 hotels. In order to evaluate the performance of the proposed model, three experiments are conducted as shown in Figure 6.

Experiment 1. Experiment 1 compares the recommendation accuracy of a single criterion CF (SC_CF) using the overall ratings actually evaluated by the customers, and a single criterion CF using the overall ratings predicted by the BERT model proposed in this paper (SC_CF(BERT)). The purpose of Experiment 1 is to measure how the actual rating value evaluated by customers and the rating value predicted by the fine-tuning BERT model presented in this paper affect the recommendation accuracy. If the difference is insignificant, it shows that there is no difference in the accuracy of the recommendation service and the predicted rating value by the BERT model replace the actual rating value.

Experiment 2. Experiment 2 compares the recommendation accuracy of the multi-criteria CF (MC_CF(BERT)) and single criteria CF (SC_CF(BERT)) using ratings predicted by the BERT model proposed in this paper. The purpose of experiment 2 is to compare the recommendation accuracy when using the overall rating value and the recommendation accuracy when using the overall rating value plus the 6-criteria rating values.

Experiment 3. Experiment 3 compares the recommendation accuracy of SC_CF and MC_CF(BERT). The purpose of experiment 3 is to compare the performance of the recommender system proposed in this study (MC_CF(BERT)) with that of a general CF benchmark system (SC_CF).

To evaluate the performance of the proposed recommender system, we use the leave-one-out method widely used in [45–48] to create a cross-validation data set, where the test set is the target hotel and the training set is the remaining hotels other than the target hotel.

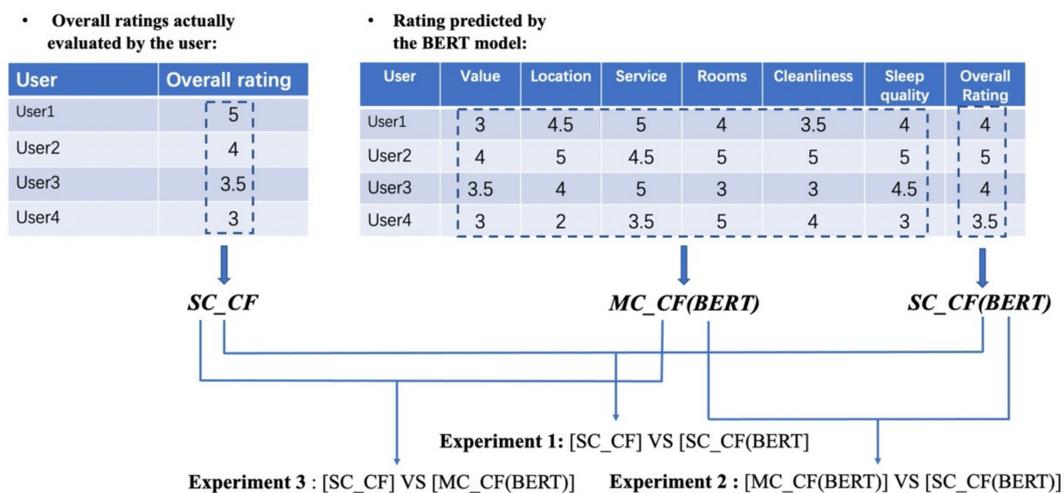


Figure 6. Experimental Design.

4.3. Evaluation Metrics

In this paper, the accuracy of Top-N recommendation list is evaluated using two metrics: Hit Ratio (HR) and Normalized Discounted Cumulative Gain (NDCG) measure. HR @ N checks whether the test set is in the Top-N list, and NDCG @ N places more weight on high-ranked users than other users on the Top-N list [48].

If a test users appears in the recommended user list, it is considered a hit. The calculation method of the hit ratio is the same as Equation (4) [49].

$$HR@N = \frac{\text{Number of Hits @N}}{|GT|}, \quad (4)$$

where, $|GT|$ is Top-N, that is, the number of recommended users, and *Number of Hits @N* is the number of users belonging to the test set in the Top-N recommendation list of each hotel, that is, the number of users who have already visited the target hotel among users recommended to the hotel.

Because HR is an evaluation index based on recall, it cannot reflect the accuracy of accurately obtaining the highest ranking, which is very important in many practical applications [48]. In order to solve this problem, the NDCG is used to give higher importance to the higher ranking results, and the marginal score utility is used to score the lower rankings in turn. The calculation method is the same as Equation (5).

$$NDCG@N = \sum_{i=1}^N \frac{2^{r_i} - 1}{\log_2(i + 1)}, \quad (5)$$

where N is the number of recommendations, and r_i is the hierarchical relevance of the user at position i [48]. The experiment uses a simple binary correlation, if the user is in the test set, $r_i = 1$, otherwise it is 0 [48].

In three experiments, two evaluation metrics are calculated for each test set, that is, each target hotel, and the average score is reported.

4.4. Experimental Results

Table 4 shows the summary of Hit Ratio results of SC_CF, SC_CF(BERT), and MC_CF(BERT). To determine the optimal number of neighbors, we performed several experiments setting k the number of neighbors from 2 to 10. The size of recommendation list is set from 5 to 15 and the Hit Ratio results of Top-5, Top10, and Top15 are shown in the table. The largest value among hit ratios in each method is indicated in bold. MC_CF

showed a higher hit ratio than SC_CF and SC_CF(BERT), but it can be seen that there is little difference between SC_CF and SC_CF(BERT).

Table 4. Hit Ratio Results (Top-N = 5, 10, 15).

Metric	Methods	k = 2	k = 3	k = 4	k = 5	k = 6	k = 7	k = 8	k = 9	k = 10
HR@5	SC_CF	0.255	0.276	0.281	0.279	0.277	0.28	0.278	0.273	0.284
	SC_CF(BERT)	0.281	0.269	0.279	0.257	0.254	0.281	0.278	0.27	0.264
	MC_CF(BERT)	0.3	0.325	0.333	0.3	0.303	0.297	0.31	0.28	0.289
HR@10	SC_CF	0.163	0.18	0.187	0.189	0.191	0.191	0.174	0.166	0.168
	SC_CF(BERT)	0.19	0.183	0.171	0.195	0.18	0.189	0.183	0.167	0.172
	MC_CF(BERT)	0.221	0.234	0.25	0.239	0.233	0.235	0.228	0.21	0.208
HR@15	SC_CF	0.154	0.157	0.163	0.155	0.153	0.154	0.147	0.141	0.14
	SC_CF(BERT)	0.176	0.167	0.149	0.162	0.153	0.163	0.159	0.161	0.165
	MC_CF(BERT)	0.183	0.181	0.198	0.194	0.208	0.217	0.197	0.19	0.183

The largest value among hit ratios in each method is indicated in bold.

Table 5 shows the NDCG results of MC_CF(BERT), SC_CF and SC_CF(BERT) proposed in this study. Regardless of the number of recommended customers, the accuracy of MC_CF(BERT) is higher than that of SC_CF and SC_CF(BERT). It can be seen that the NDCG value of SC_CF is slightly higher than the value of SC_CF(BERT), but the difference is insignificant. However, in the Hit Ratio value, it can be seen that SC_CF(BERT) is slightly higher than SC_CF, but the difference is below the significance level.

Table 5. NDCG Results (Top-N = 5, 10, 15).

Metric	Methods	k = 2	k = 3	k = 4	k = 5	k = 6	k = 7	k = 8	k = 9	k = 10
NDCG@5	SC_CF	0.641	0.671	0.68	0.663	0.68	0.644	0.655	0.679	0.688
	SC_CF(BERT)	0.601	0.638	0.626	0.621	0.59	0.577	0.653	0.642	0.653
	MC_CF(BERT)	0.694	0.665	0.656	0.632	0.635	0.66	0.654	0.634	0.63
NDCG@10	SC_CF	0.54	0.567	0.57	0.578	0.583	0.572	0.545	0.559	0.533
	SC_CF(BERT)	0.498	0.532	0.503	0.499	0.474	0.505	0.523	0.527	0.537
	MC_CF(BERT)	0.606	0.587	0.579	0.573	0.549	0.572	0.56	0.555	0.543
NDCG@15	SC_CF	0.502	0.522	0.519	0.517	0.52	0.508	0.5	0.504	0.486
	SC_CF(BERT)	0.484	0.504	0.48	0.49	0.467	0.482	0.487	0.488	0.514
	MC_CF(BERT)	0.538	0.554	0.559	0.549	0.54	0.569	0.55	0.536	0.523

The largest value among hit ratios in each method is indicated in bold.

Figure 7 shows the hit ratio results of MC_CF(BERT), SC_CF, and SC_CF(BERT). MC_CF(BERT), the recommender system proposed in this paper, showed the highest accuracy in most cases, and the accuracy tends to decrease as the number of recommendations increases. At HR@5, MC_CF(BERT) reached the highest accuracy of 0.3333. At HR@13, MC_CF(BERT) improved the most to 6.01% in accuracy than SC_CF. At least at HR@7, MC_CF(BERT) improved the accuracy by 2.95% compared to SC_CF.

Figure 8 shows the NDCG results of MC_CF(BERT), SC_CF, and SC_CF(BERT). MC_CF(BERT) showed the highest accuracy in most cases, and the accuracy tends to decrease as the number of recommendations increases. At NDCG@5, MC_CF(BERT) reached the highest accuracy of 0.6935. At NDCG@15, MC_CF(BERT) improved the most to 4.69% in accuracy than SC_CF.

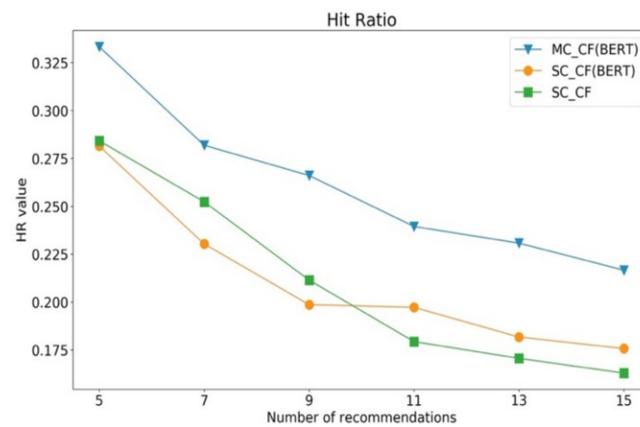


Figure 7. Evaluation of Hit Ratio where Top-N ranges from 5 to 15.

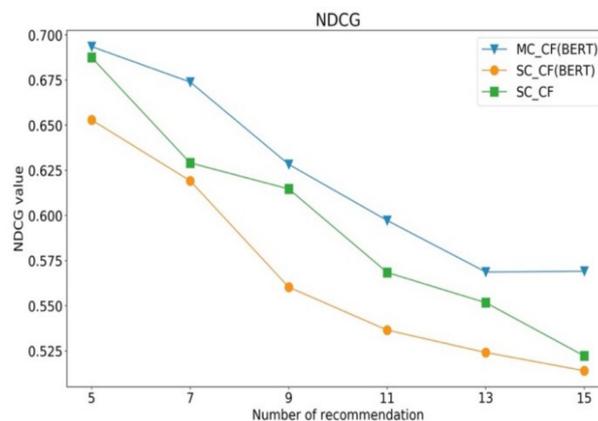


Figure 8. Evaluation of NDCG where Top-N ranges from 5 to 15.

In order to verify the experimental results, ANOVA analysis was performed on the final results. Table 6 shows the ANOVA result for the Hit Ratio value, and Table 7 shows the Multiple Comparison result for the Hit Ratio value. According to ANOVA results, the explainable variation in the total variation (0.065) was 0.019, the variation due to sampling error was 0.046, and the variances were 0.01 and 0.002, respectively. F value is 6.218, and p value is 0.006, which is less than 0.05, rejecting the null hypothesis. Therefore, in the case of Hit Ratio metric, the three methods have significant differences.

Table 6. ANOVA result (Hit Ratio).

	SS	df	MS	F -Value	p -Value
intergroup	0.019	2	0.010		
group-in	0.046	30	0.002	6.218	0.006
all	0.065	32			

Multiple comparison test was performed to more specifically analyze the difference between SC_CF, SC_CF (BERT) and MC_CF (BERT). Scheffe is used to perform multiple means comparison tests because there is no significant difference in equal variances, i.e., overall variance, over the course of the Test for homogeneity of variance analysis. Looking at the Table 7, in the case of Experiment 1, the p value between SC_CF and SC_CF (BERT) is 1, which is much greater than 0.05, so there is no difference between the two methods in the 0.05 significance probability level. In the case of Experiment 2, the p value between MC_CF (BERT) and SC_CF (BERT) is 0.017 and less than 0.05, so there is a difference between the two methods. In the case of Experiment 3, the p value between MC_CF (BERT) and SC_CF

is 0.018 and less than 0.05, so there is a difference between the two methods. Therefore, in the case of hit ratio, the multi-criteria recommender system proposed in this paper shows higher accuracy than the single criteria CF. And there is insignificant difference in the accuracy of the recommendation service, so the predicted rating value by the BERT model can replace the actual rating value.

Table 7. Multiple Comparison result (Hit Ratio).

	Models	Mean Difference (I-J)	SE	<i>p</i> -Value
SC_CF	SC_CF(BERT)	0.0003	0.0167	1.000
	MC_CF(BERT)	−0.0508691 *	0.0167	0.018 *
SC_CF(BERT)	SC_CF	−0.0003	0.0167	1.000
	MC_CF(BERT)	−0.0512113 *	0.0167	0.017 *
MC_CF(BERT)	SC_CF	0.0508691 *	0.0167	0.018 *
	SC_CF(BERT)	0.0512113 *	0.0167	0.017 *

* The mean difference is significant at the 0.05 level.

To summarize the experimental results again, in the case of hit ratio, the accuracy of the multi-criteria recommender system proposed in this paper is improved to a maximum of 6.01% compared to the single-criteria item-based CF. In the case of NDCG, the accuracy of the multi-criteria recommender system proposed in this paper is improved by up to 4.69% compared to the single-criteria item-based CF. Through these results, it can be said that the multi-criteria recommender system proposed in this study achieves the purpose of improving the recommendation accuracy. The reason is that the multi-criteria recommender system derives a rating by analyzing the customer's preference from six attributes through customer reviews, and when calculating the similarity between hotels, the attribute rating and the overall rating given by joint customers between hotels to the two hotels are calculated. In this way, by considering more detailed preferences, it is possible to find the most similar neighbor hotels, and to more accurately predict the customer's rating for the hotel, thereby the recommendation accuracy is improved.

5. Discussion

Most developed recommender system recommends hotels suitable for customers to support personalized service to customers, but in this study, a recommender system is developed that recommends customers to hotels, that is, helps the hotel's customer promotion campaign. Campaign management is the planning, execution, tracking, and analysis of a marketing initiative; sometimes centered on a new product launch or an event. Marketing shows its importance in every kind of tourist and hotel industry, since it presents itself as a tool that contributes to better management of hotel operations also help in defining appropriate strategies for their development [50]. Lambin (2000) feels that marketing and promotion are sufficiently important for the hotel, so it is necessary to develop techniques and strategies for promoting products and hotel services that could reach the market. Hotel's campaigns normally involve multiple pushes to potential buyers through email, social media, surveys, etc. The use of email marketing creates the opportunity to offer any potential interested guest to arrive at the right time at the minimum cost [51]. The main advantage of e-mail marketing is in its personalization—the message is made for a specific user, and if that person finds the offer interesting, it often results in the purchase without having to compare it with other competitors. Therefore, in the process of advertising push, a recommender system is needed to recommend customers with the largest intention to visit the hotel, so as to help the hotel improve the efficiency of campaign management.

A long time ago, merchants did not need to operate activities, but only needed to produce good products or service, and customers would come to buy them. But now, in addition to the need for high-quality products or service, we also need to manage

customers carefully to make more customers become loyal customers. In the past, many hotels have used discounts to attract more customers in their marketing activities, but now the competition among hotels is fierce. Deepening competition among hotels, hotel industry leads to the fact that the object of the competition is not only on accommodation rates, but also new kinds of proposals to stimulate and motivate consumers, quality, variety of programs to encourage repeat customers' discounts, bonuses, etc. [52]. In modern conditions of hotel complexes and the scope of their activities run into fierce competition, which allow them to seek out and apply for new ways, methods and techniques for the implementation of its services on the market.

Companies usually target customers who have purchased their own products, while insurance companies need to manage activities for target customers who have purchase intentions. The same is true for hotels. They need to manage activities for customers who are interested in visiting, and they need to decide to whom to push activity information when doing promotional activities. People nowadays receive a lot of invalid advertisements every day, which cause trouble to their lives, so most people ignore these advertisements. Therefore, it is necessary for the hotel to select customers who will frequently visit the hotel in the future, recommend to them, provide information, and manage them. Therefore, the customers selected in this study who have the greatest intention to visit the hotel, whether they are customers who have visited or who have not visited, will continue to manage and allow them to visit the hotel again and become frequent customers.

In order to help hotels with campaign management, this study uses deep learning methods, that is, using the BERT model and developed recommender system to recommend target customers for the hotel's promotion management. For hotels, this is a "starting study" of selecting and recommending customers who are most likely to visit their hotel. Many previous studies used natural language processing model to analyze customer review data through sentiment analysis, but they could not well identify the context of customers in the evaluation, so they could not correctly understand customer comments, and it was difficult to quantify them. BERT is one of the best-performing natural language processing models in recent years. Although there are many studies that use the BERT model for downstream tasks, this research uses BERT model to analyze customer reviews and predict the overall rating and six aspect ratings, and recommend customers for the hotel, which is of great significance in the field of recommender system studies. Therefore, this study will be helpful as a reference for future studies that seek to manage customers or promote company promotion through analysis of existing customer review data, even when customers are reluctant to directly input evaluation values or do not have evaluation values.

6. Conclusions

In this research, we proposed a multi-criteria recommender system, which recommends appropriate target customers for hotels using fine-tuned BERT model. To fit the recommendations in the hotel domain, we collected hotel reviews, overall ratings, and aspect ratings (Value, Service, Location, Room, Cleanliness, Sleep Quality) data from TripAdvisor, which is the world's largest travel website. Multi-criteria evaluation allows more accurate evaluation of user preference, but since TripAdvisor users write reviews but do not write rating values for six aspects, a lack of evaluation values for 6 aspects occurs. To overcome the insufficient aspect rating values, this study uses fine-tuned BERT to analyze the review data to predict overall rating and six aspect ratings. Then, using the predicted rating values, Top-N customers list with the highest visiting likelihood score is recommended to the hotels through the proposed multi-criteria recommender system. The recommendation performance of the proposed model is verified using Hit Ratio and NDCG as evaluation metrics. The results of the experiments showed that the recommendation performance of the proposed multi-criteria recommender system, MC_CF(BERT) is better than that of the single criteria CF, namely SC_CF and SC_CF (BERT). And, there is no difference in performance between SC_CF(BERT) and SC_CF at the 5% significance level,

so it can be seen that there is no difference in performance even using the rating value estimated using fine-tuned BERT.

From the hotel's point of view, the multi-criteria recommender system selects a promotion target customers more effectively by predicting the ratings based on the reviews proposed in this study. From the customer's point of view, a hotel that matches the customer's preference is recommended through a personalized recommendation, which may increase interest in the hotel and increase the intention to order.

There are three major contributions of this study. The first is to analyze customers preferences from the point of view of the hotel, and to recommend the most suitable target customers for the hotel, which helps to increase the efficiency of the hotel's promotional activities and the accommodation order rate. Second, we propose a BERT-based model that can predict ratings through reviews suitable for the hotel area and solve the problem of insufficient overall ratings and attribute ratings. Finally, we propose a multi-criteria recommender system that recommends customers to hotels and improve recommendation performance and accuracy.

The limitation and further research area of this study are summarized as follows. Since the recommendation system proposed in this study is based on item-based CF, all reviews with less than 5 reviews were removed and experiments are carried out. Therefore, the problem of not being able to recommend a customer who has not written a review or a customer who has written less than 5 reviews, that is, the cold start problem, remains. A traditional solution to this is to run promotional campaigns for customers who visit the hotel a lot. Regarding this problem, in the process of pre-processing the collected data, the data was greatly reduced. Therefore, how to use the BERT model to solve the cold start problem and this data shortage problem remains a subject for future research.

Second, the rating prediction model is based on hotel domain and English review data, so it is difficult to apply or use it for other domains and languages. Therefore, it is necessary to propose a model that can be used more broadly with more domain data and languages in future work.

Finally, the same weight is given to all attribute ratings when predicting the overall rating in the recommendation process of this study. However, each user has different preferences for each attribute, so there are not many cases in which different attributes are treated the same. Therefore, when predicting the overall rating in future work, it is necessary to make predictions considering the user's preference for different attribute ratings.

Author Contributions: Conceptualization, Y.Z. and J.K.; methodology, Y.Z. and J.K.; data curation, Y.Z.; writing—original draft preparation, Y.Z.; writing—review and editing, J.K.; supervision, J.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the BK21 FOUR (Fostering Outstanding Universities for Research) funded by the Ministry of Education(MOE, Korea) and National Research Foundation of Korea(NRF), and the Industrial Strategic Technology Development Program (20009050) funded by the Ministry of Trade, Industry, and Energy (MOTIE, Korea).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available because TripAdvisor shall own and retain all right to any data, information and other content provided by TripAdvisor.

Acknowledgments: We would like to thank Il-Young Choi for supporting statistical analysis. Thanks also to Qian Guo for her help in crawling and preprocessing TripAdvisor data.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Schuckert, M.; Liu, X.; Law, R. Insights into Suspicious Online Ratings: Direct Evidence from TripAdvisor. *Asia Pac. J. Tour. Res.* **2016**, *21*, 259–272. [CrossRef]
- Choi, I.Y.; Ryu, Y.U.; Kim, J.K. A Recommender System Based on Personal Constraints for Smart Tourism City. *Asia Pac. J. Tour. Res.* **2019**, *26*, 440–453. [CrossRef]
- Kim, J.; Choi, I.; Li, Q. Customer Satisfaction of Recommender System: Examining Accuracy and Diversity in Several Types of Recommendation Approaches. *Sustainability* **2021**, *13*, 6165. [CrossRef]
- Seo, S.; Huang, J.; Yang, H.; Liu, Y. Interpretable Convolutional Neural Networks with Dual Local and Global Attention for Review Rating Prediction. In Proceedings of the Eleventh ACM Conference on Recommender Systems, Como Italy, 27–31 August 2017; pp. 297–305.
- Li, H.; Ye, Q.; Law, R. Determinants of Customer Satisfaction in the Hotel Industry: An Application of Online Review Analysis. *Asia Pac. J. Tour. Res.* **2013**, *18*, 784–802. [CrossRef]
- Vermeulen, I.E.; Seegers, D. Tried and Tested: The Impact of Online Hotel Reviews on Consumer Consideration. *Tour. Manag.* **2009**, *30*, 123–127. [CrossRef]
- Hwang, J.; Park, S.; Woo, M. Understanding User Experiences of Online Travel Review Websites for Hotel Booking Behaviours: An Investigation of a Dual Motivation Theory. *Asia Pac. J. Tour. Res.* **2018**, *23*, 359–372. [CrossRef]
- Kim, D.; Park, C.; Oh, J.; Lee, S.; Yu, H. Convolutional Matrix Factorization for Document Context-Aware Recommendation. In Proceedings of the 10th ACM Conference on Recommender Systems, Boston, MA, USA, 15–19 September 2016; pp. 233–240.
- Ling, G.; Lyu, M.R.; King, I. Ratings Meet Reviews, a Combined Approach to Recommend. In Proceedings of the 8th ACM Conference on Recommender Systems, Silicon Valley, CA, USA, 6–10 October 2014; pp. 105–112.
- McAuley, J.; Leskovec, J. Hidden Factors and Hidden Topics: Understanding Rating Dimensions with Review Text. In Proceedings of the 7th ACM Conference on Recommender Systems, Hong Kong, China, 12–16 October 2013; pp. 165–172.
- Zhang, Y.; Lai, G.; Zhang, M.; Zhang, Y.; Liu, Y.; Ma, S. Explicit Factor Models for Explainable Recommendation Based on Phrase-Level Sentiment Analysis. In Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval, Gold Coast QLD, Australia, 6–11 July 2014; pp. 83–92.
- Devlin, J.; Chang, M.-W.; Lee, K.; Toutanova, K. Bert: Pre-Training of Deep Bidirectional Transformers for Language Understanding. *arXiv* **2018**, arXiv:1810.04805.
- Raffel, C.; Shazeer, N.; Roberts, A.; Lee, K.; Narang, S.; Matena, M.; Zhou, Y.; Li, W.; Liu, P.J. Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer. *arXiv* **2019**, arXiv:1910.10683.
- Penha, G.; Hauff, C. What Does BERT Know about Books, Movies and Music? Probing BERT for Conversational Recommendation. In Proceedings of the Fourteenth ACM Conference on Recommender Systems; Available online: <https://recsys.acm.org/recsys20/>. (accessed on 22 September 2020).
- Yang, Z.; Dai, Z.; Yang, Y.; Carbonell, J.; Salakhutdinov, R.; Le, Q.V. Xlnet: Generalized Autoregressive Pretraining for Language Understanding. *arXiv* **2019**, arXiv:1906.08237.
- Liu, Y.; Ott, M.; Goyal, N.; Du, J.; Joshi, M.; Chen, D.; Levy, O.; Lewis, M.; Zettlemoyer, L.; Stoyanov, V. Roberta: A Robustly Optimized Bert Pretraining Approach. *arXiv* **2019**, arXiv:1907.11692.
- Dong, X.; Ni, J.; Cheng, W.; Chen, Z.; Zong, B.; Song, D.; Liu, Y.; Chen, H.; de Melo, G. Asymmetrical Hierarchical Networks with Attentive Interactions for Interpretable Review-Based Recommendation. In Proceedings of the AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 February 2020; Volume 34, pp. 7667–7674.
- Wang, Y.; Liu, Y.; Yu, X. Collaborative Filtering with Aspect-Based Opinion Mining: A Tensor Factorization Approach. In Proceedings of the 2012 IEEE 12th International Conference on Data Mining, Brussels, Belgium, 10–13 December 2012; pp. 1152–1157.
- Wang, H.; Lu, Y.; Zhai, C. Latent Aspect Rating Analysis on Review Text Data: A Rating Regression Approach. In Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, 25–28 July 2010; pp. 783–792.
- Nilashi, M.; bin Ibrahim, O.; Ithnin, N.; Sarmin, N.H. A Multi-Criteria Collaborative Filtering Recommender System for the Tourism Domain Using Expectation Maximization (EM) and PCA-ANFIS. *Electron. Commer. Res. Appl.* **2015**, *14*, 542–562. [CrossRef]
- Shambour, Q.; Lu, J. A Hybrid Multi-Criteria Semantic-Enhanced Collaborative Filtering Approach for Personalized Recommendations. In Proceedings of the 2011 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology, Lyon, France, 22–27 August 2011; Volume 1, pp. 71–78.
- Lakiotaki, K.; Tsafarakis, S.; Matsatsinis, N. UTA-Rec: A Recommender System Based on Multiple Criteria Analysis. In Proceedings of the 2008 ACM Conference on Recommender Systems, Lausanne, Switzerland, 23–25 October 2008; pp. 219–226.
- Adomavicius, G.; Huang, Z.; Tuzhilin, A. Personalization and recommender systems. In *State-of-the-Art Decision-Making Tools in the Information-Intensive Age*; InformsPubsOnLine: Catonsville, MD, USA, 2008; pp. 55–107.
- Adomavicius, G.; Tuzhilin, A. Toward the next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions. *IEEE Trans. Knowl. Data Eng.* **2005**, *17*, 734–749. [CrossRef]
- Resnick, P.; Varian, H.R. Recommender Systems. *Commun. ACM* **1997**, *40*, 56–58. [CrossRef]

26. Kim, H.K.; Kim, J.K.; Ryu, Y.U. Personalized Recommendation over a Customer Network for Ubiquitous Shopping. *IEEE Trans. Serv. Comput.* **2009**, *2*, 140–151. [CrossRef]
27. Choi, I.Y.; Kim, J.K.; Ryu, Y.U. A Two-Tiered Recommender System for Tourism Product Recommendations. In Proceedings of the 2015 48th Hawaii International Conference on System Sciences, Kauai, HI, USA, 5–8 January 2015; pp. 3354–3363.
28. Sarwar, B.; Karypis, G.; Konstan, J.; Riedl, J. Item-Based Collaborative Filtering Recommendation Algorithms. In Proceedings of the 10th international conference on World Wide Web, Hong Kong, 1–5 May 2001; pp. 285–295.
29. Aciar, S.; Zhang, D.; Simoff, S.; Debenham, J. Recommender System Based on Consumer Product Reviews. In Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence (WI 2006 Main Conference Proceedings)(WI'06), Hong Kong, China, 18–22 December 2006; pp. 719–723.
30. Alexandridis, G.; Tagaris, T.; Siolas, G.; Stafylopatis, A. From Free-Text User Reviews to Product Recommendation Using Paragraph Vectors and Matrix Factorization. In Proceedings of the Companion Proceedings of The 2019 World Wide Web Conference, San Francisco, CA, USA, 13–17 May 2019; pp. 335–343.
31. Baek, J.-W.; Chung, K.-Y. Multimedia Recommendation Using Word2Vec-Based Social Relationship Mining. *Multimed. Tools Appl.* **2020**, *1*–17. [CrossRef]
32. Alexandridis, G.; Siolas, G.; Stafylopatis, A. ParVecMF: A Paragraph Vector-Based Matrix Factorization Recommender System. *arXiv* **2017**, arXiv:1706.07513 2017.
33. Le, Q.; Mikolov, T. Distributed Representations of Sentences and Documents. In Proceedings of the International conference on machine learning. *PMLR* **2014**, *32*, 1188–1196.
34. Mikolov, T.; Chen, K.; Corrado, G.; Dean, J. Efficient Estimation of Word Representations in Vector Space. *arXiv* **2013**, arXiv:1301.3781.
35. Liu, L.; Mehandjiev, N.; Xu, D.-L. Multi-Criteria Service Recommendation Based on User Criteria Preferences. In Proceedings of the Fifth ACM Conference on Recommender Systems, Chicago, IL, USA, 23–27 October 2011; pp. 77–84.
36. Adomavicius, G.; Kwon, Y. New Recommendation Techniques for Multicriteria Rating Systems. *IEEE Intell. Syst.* **2007**, *22*, 48–55. [CrossRef]
37. Nie, Y.; Liu, Y.; Yu, X. Weighted Aspect-Based Collaborative Filtering. In Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval, Gold Coast Queensland, Australia, 6–11 July 2014; pp. 1071–1074.
38. Zhu, J.; Xia, Y.; Wu, L.; He, D.; Qin, T.; Zhou, W.; Li, H.; Liu, T.-Y. Incorporating Bert into Neural Machine Translation. *arXiv* **2020**, arXiv:2002.06823.
39. Yang, W.; Xie, Y.; Lin, A.; Li, X.; Tan, L.; Xiong, K.; Li, M.; Lin, J. End-to-End Open-Domain Question Answering with Bertserini. *arXiv* **2019**, arXiv:1902.01718.
40. Gao, Z.; Feng, A.; Song, X.; Wu, X. Target-Dependent Sentiment Classification with BERT. *IEEE Access* **2019**, *7*, 154290–154299. [CrossRef]
41. Peters, M.E.; Neumann, M.; Iyyer, M.; Gardner, M.; Clark, C.; Lee, K.; Zettlemoyer, L. Deep Contextualized Word Representations. *arXiv* **2018**, arXiv:1802.05365.
42. Brown, T.B.; Mann, B.; Ryder, N.; Subbiah, M.; Kaplan, J.; Dhariwal, P.; Neelakantan, A.; Shyam, P.; Sastry, G.; Askell, A. Language Models Are Few-Shot Learners. *arXiv* **2020**, arXiv:2005.14165.
43. Zhu, Y.; Kiros, R.; Zemel, R.; Salakhutdinov, R.; Urtasun, R.; Torralba, A.; Fidler, S. Aligning Books and Movies: Towards Story-like Visual Explanations by Watching Movies and Reading Books. Available online: https://www.cv-foundation.org/openaccess/content_iccv_2015/papers/Zhu_Aligning_Books_and_ICCV_2015_paper.pdf. (accessed on 22 September 2020).
44. Sun, C.; Qiu, X.; Xu, Y.; Huang, X. How to Fine-Tune BERT for Text Classification? In *China National Conference on Chinese Computational Linguistics*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 194–206.
45. Bayer, I.; He, X.; Kanagal, B.; Rendle, S. A Generic Coordinate Descent Framework for Learning from Implicit Feedback. In Proceedings of the 26th International Conference on World Wide Web, Perth, Australia, 3–7 April 2017; pp. 1341–1350.
46. Rendle, S.; Freudenthaler, C.; Gantner, Z.; Schmidt-Thieme, L. BPR: Bayesian personalized ranking from implicit feedback. *arXiv* **2012**, arXiv:1205.2618.
47. He, X.; Du, X.; Wang, X.; Tian, F.; Tang, J.; Chua, T.S. Outer product-based neural collaborative filtering. *arXiv* **2018**, arXiv:1808.03912.
48. Lee, J.W.; Choi, M.; Lee, J.; Shim, H. Collaborative Distillation for Top-N Recommendation. In Proceedings of the 2019 IEEE International Conference on Data Mining (ICDM), Beijing, China, 8–11 November 2019; pp. 369–378.
49. He, X.; Chen, T.; Kan, M.Y.; Chen, X. Trirank: Review-aware explainable recommendation by modeling aspects. In Proceedings of the 24th ACM International on Conference on Information and Knowledge Management, Melbourne, Australia, 18–23 October 2015; pp. 1661–1670.
50. Pereira, L.; Almeida, P. Marketing and Promotion in the Hotel Industry: A Case Study in Family Hotel and Hotel Group. *Tour. Hosp. Int. J.* **2014**, *2*, 92–105.
51. Batinić, I. The Role and Importance of Internet Marketing in Modern Hotel Industry. *J. Process. Manag. New Technol.* **2015**, *3*, 34–38.
52. Goryushkina, N.Y.; Shkurkin, D.V.; Petrenko, A.S.; Demin, S.Y.; Yarovaya, N.S. Marketing Management in the Sphere of Hotel and Tourist Services. *Int. Rev. Manag. Mark.* **2016**, *6*, 777–780.