

Received 5 December 2023, accepted 26 December 2023, date of publication 2 January 2024,  
date of current version 12 January 2024.

Digital Object Identifier 10.1109/ACCESS.2023.3349291

## RESEARCH ARTICLE

# An Attentive Aspect-Based Recommendation Model With Deep Neural Network

SIGEON YANG<sup>1</sup>, QINGLONG LI<sup>1</sup>, HAEBIN LIM<sup>1</sup>, AND JAEKYEONG KIM<sup>1,2</sup>

<sup>1</sup>Department of Big Data Analytics, Kyung Hee University, Seoul 02447, South Korea

<sup>2</sup>School of Management, Kyung Hee University, Seoul 02447, South Korea

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

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). (Sigeon Yang and Qinglong Li contributed equally to this work.)

**ABSTRACT** With the growth of the internet and e-commerce, online reviews have become a prevalent and rich source of information for personalized recommendations. Review text typically contains user preferences regarding various aspects of items or services. However, most previous recommendation models that used online reviews focused on general opinions, without considering the various aspects of the review text. Although previous approaches have effectively reflected the overall preferences embedded in reviews, capturing the preferences for various aspects is limited. This paper proposes an attentive aspect-based recommendation model with a deep neural network (AARN), which can capture user preferences regarding various aspects embedded in review text. The proposed model uses an advanced aspect-based sentiment analysis (ABSA) method. Unlike previous approaches that extract overall preferences, ABSA offers the advantage of extracting preferences for various aspects embedded in review texts. An attention mechanism is then applied to measure the unique attention weights for each aspect. Extensive experiments were conducted on real-world online review datasets, showing that the proposed model outperformed baseline models in effectively predicting user ratings for target items. Furthermore, this study demonstrated the influence of the ABSA method on personalized and explainable recommendations.

**INDEX TERMS** Aspect-based sentiment analysis, attention mechanism, deep learning, online review, recommendation model, user preference.

## I. INTRODUCTION

With the advancement of information and communication technology (ICT) and popularization of smartphones, the e-commerce market has grown rapidly [1]. Furthermore, booking services through online platforms has become common in various service industries, such as hotels and restaurants [2]. However, owing to the rapid development of online platforms, many users face information overload when exploring items [3]. Moreover, users in an online environment often experience uncertainty in finding options that are suitable for their personal preferences among the numerous options provided by websites because it is challenging to confirm items and services in advance [4]. To address these

issues, there is increasing demand for personalized recommender systems to assist users in their decision-making process [5].

Collaborative filtering (CF) has shown excellent performance and is one of the most popular models used in recommendation studies [6]. However, this model has a significant limitation in that it uses rating information as its only information source, which results in a data sparsity problem [4], [7]. Additionally, rating information only represents users' overall preferences for target items, limiting the ability to understand user behavior motivations [8].

Several studies have proposed hybrid recommender systems that utilize online reviews as additional information to address the data sparsity problem [9], [10]. For example, Ghasemi and Momtazi [11] proposed a novel methodology that combines rating similarity from collaborative filtering

The associate editor coordinating the review of this manuscript and approving it for publication was Yin Zhang<sup>1</sup>.

and text similarity from reviews using language models such as document-to-vector (Doc2Vec) and long short-term memory (LSTM). Furthermore, Liu et al. [12] reported that applying a multi-embedding method with advanced language models, such as a bidirectional encoder representations from transformers (BERT) and a robustly optimized BERT pretraining approach (RoBERTa), improves recommendation performance when incorporating review text. Such an approach achieves excellent performance compared to recommendation models that use ratings as the only information source but ignore the fact that the review text typically includes user preferences for various aspects of an item or service [13].

Many scholars have reported that users have different preferences for various aspects of an item, which is an essential information source for user behavior motivation [14], [15], [16]. However, most previous recommendation models using online reviews only considered the overall opinions without considering the various aspects of the review text [17]. Although these approaches may reflect the general preferences of users, capturing the preferences of various aspects is limited [14]. Here, the preference represents a user assigning more attention and weight to a specific aspect. For example, suppose a user prefers the price aspect over others. In this case, the user may give a high rating at a reasonable price even though other aspects of the item are unsatisfactory. Previous approaches have limitations in capturing such aspect-specific preferences embedded in the review texts.

To address this research gap, this paper proposes an attentive aspect-based recommendation model with a deep neural network (AARN), which can capture user preferences for various aspects embedded in review text. The proposed model adopts an aspect-based sentiment analysis (ABSA) method with BERT to extract user preferences from various aspects. ABSA is one of the core methods in natural language processing, extracting sentiments for different aspects embedded in the review text [18]. In addition, the proposed model utilizes an attention mechanism to capture unique attention weights for users and items.

Users typically evaluate various aspects of the items before providing an overall rating [4]. As individual users pay different attention to various aspects, the overall rating is influenced by the attentional weight assigned to each aspect [8]. Consequently, the proposed model extracts unique attentional weights by establishing the relationship between the overall rating and the evaluation of each aspect.

The proposed model consists of three neural networks in the following sequence. First, the proposed model extracts the latent factors of the user and item to learn a high-level interaction representation. Subsequently, ABSA is applied to capture user preferences for various aspects embedded in the review text. Second, the attentional weights for the user and item are generated using the attention mechanism. Finally, the proposed model predicts the final overall rating using a multi-layer perceptron (MLP) network by combining the unique representations of users and items. To evaluate the

performance of the proposed model, we conducted comprehensive experiments on the Amazon and Yelp datasets for comparison with baseline models. The results showed that the proposed model outperformed the baselines with an improved prediction performance.

In this study, we aim to answer the following research questions (RQ):

- RQ 1: Does the proposed AARN perform better than the other baseline models?
- RQ 2: How does the attention mechanism contribute to the recommendation performance?
- RQ 3: Does an ABSA effectively capture user preference representation embedded in the review text?
- RQ 4: How do different hyperparameters affect the recommendation performance?

The main contributions of this study are as follows.

- This study proposes a novel recommendation model that considers user preferences for various aspects embedded in review texts and considers the importance of each aspect. Thus, the proposed model reflects user preferences for various aspects of the decision-making process.
- This paper proposes a personalized recommendation model that combines an advanced ABSA method with an attention mechanism. This is meaningful as a pioneer to integrate a BERT-based ABSA with a recommendation model.
- Through extensive experiments using real-world online review datasets, this study demonstrated that applying the ABSA method leads to superior recommendation performance. This paper also introduces a novel experimental design comparing the impact of various text embedding methods on personalized recommendations.

The remainder of this paper is organized as follows. First, a review of related work is introduced in Section II. Next, the problem definition is presented in Section III. The proposed model is described in detail in Section IV. Section V presents the experimental design. In Section VI a series of experimental results and discussions are presented. Finally, conclusions and future work are discussed in Section VII.

## II. RELATED WORK

### A. ONLINE REVIEW-BASED RECOMMENDATION

Collaborative filtering, the most widely used recommendation algorithm, predicts a target user's preferences by finding neighbors with similar preferences based on the user's historical ratings [19], [20]. However, the well-known data sparsity problem occurs because CF-based models use rating information as the only input source [14]. Furthermore, rating information only represents the user's overall evaluation of items, limiting the ability to understand the specific preferences of users [21]. To address this issue, advanced CF models have been proposed to leverage the rich information implied by review texts. For example, Li et al. [22] presented a data-filtering approach to improve the performance of a CF model based on a helpfulness score of online reviews.

Topic modeling was used to extract common topics from the review texts. This approach typically combines topics in reviews with latent factors of users and items from a CF-based model [23]. For example, Duan et al. [7] proposed a solution to the data sparsity problem by clustering reviews using topic modeling and sentiment analysis and proposed a review-based matrix factorization (RMF) model that combines topic clusters with matrix factorization. Ghasemi and Momtazi [11] proposed a novel CF approach that combines rating similarity from CF and review text similarity from review texts.

With the remarkable development of deep learning, increasing efforts have been made to incorporate the rich information embedded in review texts into recommendations. A deep cooperative neural network (DeepCoNN) was the first attempt to represent the integration between users and items based on their review sets [9]. This model applies a convolutional neural network (CNN) technique to review sets to derive word embeddings and extract the unique features of users and items. Following these efforts, numerous deep learning-based text analysis techniques have been applied to aggregated review sets to analyze user preferences [6], [24]. For example, Nam [25] demonstrated the efficiency of using BERT embeddings to describe textual reviews.

In particular, Cao et al. [8] proposed a novel review semantics-based model (RSBM) using individual review texts, unlike previous approaches, which considered all review sets as long documents. This approach is significant because it incorporates a neural collaborative filtering model [26] and deep-learning-based text analysis technique for individual reviews. Likewise, Liu et al. [12] proposed a mobile application recommendation model that improves the quality of latent factors of users and items by applying multi-embedding techniques to individual review texts. Review texts have become an essential source of information for efficiently capturing user preferences. However, this approach does not fully utilize the aspect-level user preference information embedded in review text [13]. Therefore, this paper proposes a novel advanced recommendation model that can capture user preferences for various aspects embedded in review text.

## B. ASPECT-BASED SENTIMENT ANALYSIS

With the rapid development of deep learning, natural language processing (NLP) has undergone significant advancements. Sentiment analysis, which is a core NLP task, typically extracts general opinions represented in documents or review texts with positive, negative, or neutral polarity [27]. Previous studies have applied sentiment analysis to extract users' general opinions about items from review texts and reflect them in the recommendation process [10], [28], [29]. For example, Asani et al. [28] adopted lexicon-based sentiment analysis using the SentiWordNet dictionary and applied it to a restaurant recommender system. Recently, aspect-based sentiment analysis (ABSA) has been widely applied in various academic fields, with the aim of deriving detailed user opinions [15], [30], [31].

ABSA is a fine-grained opinion analysis technique used for NLP tasks [31], [32]. While conventional sentiment analysis focuses on extracting a general opinion embedded in text, ABSA aims to predict granular opinions for various aspects embedded in the text [18]. For example, in the case of review text "the restaurant is expensive, but the food is fantastic," the opinion for each aspect is extracted more specifically, such as 'negative' for the price aspect and 'positive' for the food aspect. In early studies, conventional text analysis techniques, such as topic modeling and word clustering, were applied for ABSA [14], [33]. However, because the pretrained language model BERT has shown significant improvement in NLP, most ABSA approaches have dealt with how to effectively utilize BERT [34]. Major studies have constructed an input dataset combining key aspects with sentences for the sentence-pair classification task of BERT [32], [35]. In particular, Zeng et al. [18] extracted local context features (LCF) of text and combined them with a pretrained BERT language model, which is a more stable and precise method for aspect-level sentiment classification.

Recently, advanced ABSA methods have been applied in various academic fields. For example, Li et al. [31] employed BERT-based ABSA to measure aspect-specific sentiment scores from review data on Yelp.com and used them as input data to perform a restaurant survival analysis. Mehra [30] also applied BERT-based ABSA with emotion analysis to identify travelers' behavioral intentions. Hajek et al. [15] proposed a novel approach to detect fake reviews using an advanced ABSA while considering the effects of product types on e-commerce platforms. However, to the best of our knowledge, studies that combine advanced ABSA with recommendation models are lacking. This paper proposes a novel recommendation model that applies BERT-based ABSA to extract user preferences for various aspects embedded in review text and combines it with a recommendation model.

## C. ATTENTION MECHANISM

The key idea of an attention mechanism is to learn the assigned attentive weights for various aspects by normalizing the weight scores and summing them to 1 [36]. Higher weight scores indicate that the corresponding aspects are more important than other aspects. Recommendation models adopt an attention mechanism for user review texts [37]. Most reviews contain user preferences regarding various aspects of the target item or service [14]. Here, an aspect, also called a feature or attribute, refers to a component of a certain item or service [5].

In the recent literature, attention mechanisms have been applied to identify user preferences regarding various aspects of items [8], [24], [38]. For example, Lai and Hsu [39] extracted user and item features from textual reviews using an attention mechanism and applied them to a recommendation model. Guan et al. [38] demonstrated the importance of modeling the interactions between different aspects and the

effectiveness of the attention mechanism in capturing users' varied attentions towards unique aspects. Lai and Tseng [17] presented a method for deriving user and item feature matrices by combining sentence-level latent Dirichlet allocation (LDA), word embedding, and an attention mechanism.

Several studies have adopted CNN techniques to extract aspects from the structured semantics of review texts and applied an attention mechanism to the key aspects. For example, Cao et al. [8] pointed out that users typically assess different aspects of the target items in their decision-making process and proposed a rating prediction model integrating a CNN and attention mechanism. Da'u et al. [13] utilized a deep CNN technique to extract aspects from review text and generate aspect ratings by computing the sentiment polarities of users for various aspects. Zhang et al. [24] proposed an explainable recommendation model by applying a CNN to extract features from review texts, which enabled the attention mechanism to track the keywords in the original review text using CNN kernels. Unlike previous approaches that used topic modeling or CNNs, the current study adopted ABSA to extract user preferences for various aspects embedded in review texts. Accordingly, this study captured both user and item attentive representations.

### III. PROBLEM DEFINITION

Let  $D$  be a collection of reviews and ratings for a given dataset. The dataset  $D$  is composed of four sets: users ( $U$ ), items ( $V$ ), ratings ( $R$ ), and textual reviews ( $T$ ). We represent each user-item interaction as a tuple,  $(u_i, v_j, r_{i,j}, t_{i,j})$ , consisting of review text  $t_{i,j} \in T$  written by user  $u_i \in U$  for item  $v_j \in V$  with corresponding rating  $r_{i,j} \in R$ . Each dataset has predefined key aspects  $A = \{a_1 a_2, \dots, a_k\}$ , such as 'food', 'service', and 'price' for a restaurant dataset. An overview of the research process is illustrated in Figure 1. In this figure,  $u_i$ ,  $v_j$ ,  $t_{i,j}$ , and  $r_{ij}$  are depicted as 'User ID', 'Item ID', 'Review text', and 'Rating', respectively. The main objective of the current study is to develop a model that predicts the overall rating  $r_{i,j}$  based on these input data.

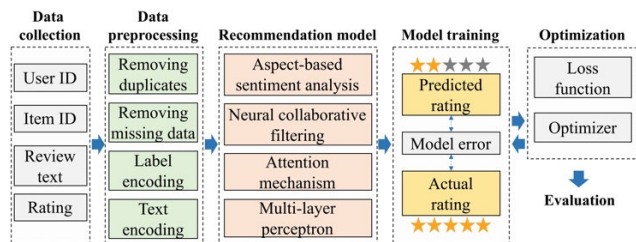


FIGURE 1. An overview of the research process.

The input data from collected dataset are preprocessed before being fed into the recommendation model. The proposed recommendation model utilizes four major techniques: aspect-based sentiment analysis, neural collaborative filtering, an attention mechanism, and multi-layer perceptron. These models work together to generate personalized rating

predictions. Then, the model is optimized using a loss function and an optimizer to minimize the observed errors between the predicted and actual ratings. Finally, the performance is evaluated using the trained model.

### IV. AARN FRAMEWORK

The proposed AARN model consists of three major phases, as illustrated in Figure 2. First, in the latent factor phase, user and item latent factors are generated through each embedding layer. Advanced ABSA is then applied to extract aspect-specific sentiment scores from the review text. These sentiment scores are converted into attentional weights for each user and item using an attention mechanism. In the next phase, latent factors and attentional weights are integrated. Here, two novel representations are introduced to predict the final rating scores: user-attentive and item-attentive representations. Finally, the user- and item-attentive representations are fed into the rating prediction phase. The overall framework of the proposed model represents the decision-making process by which a user evaluates various aspects before giving a final rating to a particular item.

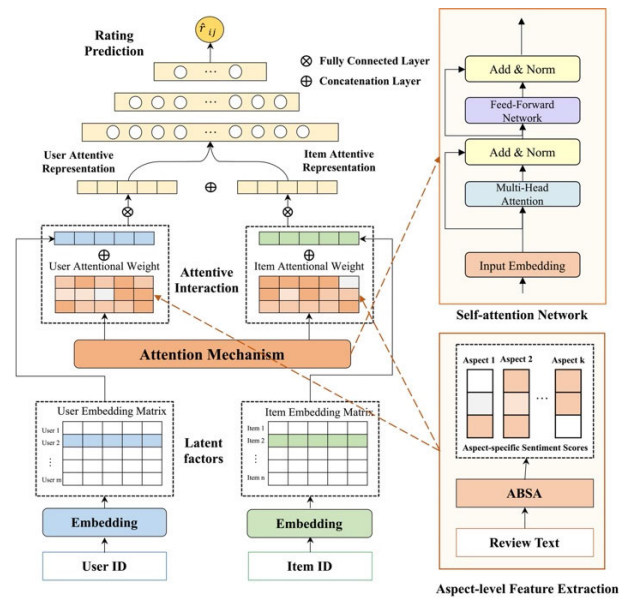


FIGURE 2. Framework of AARN.

#### A. THE LATENT FACTORS OF USERS AND ITEMS

The proposed model derives the latent factors of users and items from the interaction tuple  $(u_i, v_j, r_{i,j}, t_{i,j})$ . The unique IDs of user  $u_i$  and item  $v_j$  are converted to one-hot encodings and then projected to  $d$ -dimensional dense vectors through embedding layers. Here,  $d$  represents the dimension of latent factors and is set to the optimal value to capture the collaborative interaction between users and items [26]. The user latent factor  $p_u$  and item latent factor  $q_v$  are defined as:

$$p_u = P^T o_i^U \quad (1)$$

$$q_v = Q^T o_j^V \quad (2)$$

where  $P \in \mathbb{R}^{m \times d}$  is the embedding matrix of users,  $Q \in \mathbb{R}^{n \times d}$  is the embedding matrix of items,  $O_i^U$  denotes the one-hot encoding of  $u_i$  and  $O_j^V$  denotes the one-hot encoding of  $v_j$ . Here,  $m$  and  $n$  are the numbers of unique users and items in the datasets, respectively.

Therefore, each row value of the embedding matrices represents a specific user or item. In Figure 2, the unique values of the user/item embedding matrix are initialized as random float numbers and optimized through model training. Note that in the pure neural collaborative filtering model, only ID information is used to describe a user and item [26]. In contrast, this study used review texts written by users as additional information to learn rich representations [2].

### B. ASPECT-LEVEL FEATURE EXTRACTION

In this phase, aspect-based sentiment analysis is applied to the review text  $t_{i,j}$  to extract granular sentiment scores from the set of key aspects  $A$ . This study adopted BERT-based ABSA [18]. This pipeline model provides a pretrained language model to make predictions for aspect-based sentiment analysis tasks and has been widely used in various academic fields. Fine-tuning of the model was not performed in this study because it was already pretrained for various datasets [30], [31].

The number of aspects  $k$  depends on the datasets  $D$ . The output of the ABSA model consists of three sentiment scores for each aspect: positive, negative, and neutral. A higher score indicates stronger sentiment. The aspect-specific sentiment scores  $S_{i,j}$  are concatenated as

$$S_{i,j} = [s_1 : s_2 : \dots : s_k] \quad (3)$$

where  $[\cdot]$  represents concatenation of the vectors, and each  $s_1, s_2, \dots, s_k$  is a vector of three sentiment scores corresponding to each aspect,  $a_1 a_2, \dots, a_k$ . The concatenated layer is applied without any transformations to avoid information loss. Therefore, the size of the vector  $S_{i,j}$  comprises  $k \times 3$  (number of aspects  $\times$  three sentiment scores).

### C. ATTENTIVE INTERACTION REPRESENTATIONS

In this subsection, we introduce the attention mechanism for extracting the user/item-attentive representations. Intuitively, the attention weights of the aspects represent this relationship, as illustrated in Figure 3. If the user gives a high rating (e.g., a rating of 5), the user may have more attention weight on the aspects with a positive sentiment score and less attention weights on the aspects with a negative sentiment score. In contrast, if the user gives a low rating (e.g., a rating of 1), the use may have more attention weights on the aspects with a negative sentiment score and less attention weights on the aspects with a positive sentiment score.

A content-based memory-like attention layer is applied to extract the attention weight from the vector  $S_{i,j}$  [36]. First, a parameterized function  $F_k$  is adopted for mapping input  $X_h$  from  $h$  to  $k$  dimensions.

$$B_{h,k} = F_k(X_h) \quad (4)$$

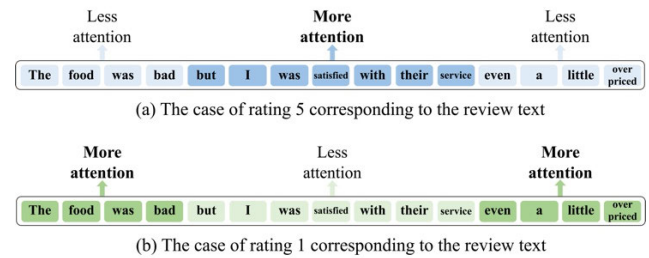


FIGURE 3. Illustration of relation between attention weights and ratings.

In practice, we adopt a two-layer feed-forward network with an activation function as follows:

$$F_k(X_h) = W_{2,k}(\sigma(W_{1,h}(X_h) + b_{1,h})) + b_{2,k} \quad (5)$$

where  $W_{1,h} \in \mathbb{R}^{h \times h}$  and  $W_{2,k} \in \mathbb{R}^{h \times k}$  are the weight matrices, and  $\sigma$  is the rectified linear unit (ReLU) activation function as follows:

$$ReLU(x) = \max(0, x) \quad (6)$$

The attention mechanism is applied to calculate the attentional weight  $z_k$  of a particular aspect  $k$  from the aspect-specific sentiment scores  $S_{i,j}$ :

$$z_k = \text{Softmax}(F_k(S_{i,j})) G(S_{i,j}) \quad (7)$$

where  $G(X_h)$  is a transformation function that matches the raw input values with the attentional probabilities. The final attentional weight for a particular aspect  $k$  is obtained by normalizing attentional weights from all aspects with the SoftMax function, which can be interpreted as the importance of the  $k$ -th aspect in the dataset:

$$\text{Softmax}(z_k) = \frac{\exp(z_k)}{\sum_{j=1}^K \exp(z_j)} \quad (8)$$

Finally, the attention mechanism computes the attention weight matrix of all aspects  $Z$ . The proposed model derives the user's attentional weight  $Z_u$  and item's attentional weight  $Z_v$ . These two vectors are concatenated with  $p_u$  and  $q_v$  and calculated as  $d$ -dimensional vectors  $e_u$  and  $e_v$ , respectively, through a dense layer:

$$e_u = \sigma(W_u^d([p_u : Z_u]) + b_u) \quad (9)$$

$$e_v = \sigma(W_v^d([q_v : Z_v]) + b_v) \quad (10)$$

The  $d$ -dimensional vectors  $e_u$  and  $e_v$  are the user-attentive and item-attentive representations in Figure 2, respectively.

### D. RATING PREDICTION

In the rating prediction phase, this study applied an MLP as the final rating prediction network. The unique representation vectors  $e_u$  and  $e_v$  are passed as inputs of the MLP. The output is then obtained as follows:

$$\begin{aligned} \text{output} &= \text{Sigmoid}(W_L(\sigma_{L-1}(W_{L-1} \dots \sigma_1(W_1[e_u : e_v] + b_1)) \\ &\quad + b_{L-1}) + b_L) \end{aligned} \quad (11)$$

A linear transformation is applied to the rating scores to ensure that the final predicted rating remains between the minimum and maximum observed ratings [3] as follows:

$$r' = (r - r_{min}) / (r_{max} - r_{min}) \quad (12)$$

All datasets used in this study had an  $r_{min}$  of 1 and  $r_{max}$  of 5. That is, the integer values [1], [2], [3], [4], [5] were normalized to real numbers [0, 0.25, 0.5, 0.75, 1.0]. Therefore, the sigmoid function, which converts the final predicted result value to a float-type number in the range [0:1], was adopted as the activation function for the last dense layer to optimize the model.

$$\text{Sigmoid}(x) = \frac{1}{1 + e^{-x}} \quad (13)$$

Finally, the linear transformation in (14) is applied to restore the scale of the ratings at the top of the proposed model. Note that (14) follows the reverse order of (12):

$$\hat{r}_{i,j} = r_{min} + output \times (r_{max} - r_{min}) \quad (14)$$

To optimize the model, we defined the objective function of the deep neural network as (15), where  $\theta$  denotes all parameters of the model,  $\hat{Y}_{i,j}$  is the final output,  $Y_{i,j}$  is the actual value, and  $M, N$  are the numbers of users and items in the dataset, respectively.

$$\text{Min}_{\theta} - \frac{1}{N} \sum_{i=1}^M \sum_{j=1}^N (y_i \cdot \log(\hat{y}_{i,j}) + (1 - y_{i,j}) \cdot \log(1 - \hat{y}_{i,j})) \quad (15)$$

## V. EXPERIMENTS

### A. DATASETS

This study used three different datasets to cover diverse domains: Yelp restaurant, Amazon laptop, and Amazon tablet. The Yelp restaurant dataset is provided by Yelp.com and contains six sub-datasets, such as user reviews, business attributes, and user information [31]. The datasets were downloaded in February 2023. We extracted user IDs, business IDs, overall ratings, and corresponding review texts. We then filtered out users and business IDs with more than five review texts. To eliminate potential regional biases, we segmented the dataset based on the city in which each restaurant was located. Because the city of Philadelphia had the highest number of reviews in the original dataset, we selected it as our focus. The restaurant dataset contains five key aspects [18], [30], [31], [34].

Amazon review datasets are publicly accessible from Amazon Product Review data and were downloaded in February 2023 [40]. We used the processed 5-core datasets, which guaranteed that every user and item ID has a minimum of five reviews. The products were divided into 29 main categories, and we extracted the review data for the laptop and tablet products lines from the 'Electronics' category. We collected the user IDs, product IDs, overall ratings, and corresponding review texts. Each tablet and laptop dataset has eight key aspects [18], [34], [35].

Finally, we filtered out review data with less than 90% length for model optimization. In the case of Amazon laptop data, the median word count in the reviews was 64 words, while the maximum review had 4,924 words. Table 1 provides comprehensive statistics of the final datasets. For all datasets, the overall ratings fell within the range of 1–5, with 5 being most preferred and 1 least preferred. Each dataset was randomly divided into training, validation, and test sets in a ratio of 8:1:1 [12].

TABLE 1. Statistics of the datasets.

| Datasets        | Number of reviews | Number of users | Number of items |
|-----------------|-------------------|-----------------|-----------------|
| Yelp restaurant | 598,499           | 37,465          | 14,342          |
| Amazon tablet   | 62,062            | 51,884          | 1,378           |
| Amazon laptop   | 53,052            | 45,221          | 2,567           |

### B. EVALUATION METRICS

To evaluate the performance of the proposed model, we used two evaluation metrics for rating prediction: mean absolute error (MAE) and root mean square error (RMSE), which are widely used to evaluate recommendation performance [41]. The RMSE metric, which calculates the error between the actual and predicted values in a squared format, provides more weight to large errors and less weight to smaller ones; thus, it is generally preferred over MAE. However, because these metrics are sensitive to outliers, the MAE metric can also be preferred depending on the distribution of the data [42].

$$MAE = \frac{1}{N} \sum_{n=1}^N |r_{i,j} - \hat{r}_{i,j}| \quad (16)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{n=1}^N (r_{i,j} - \hat{r}_{i,j})^2} \quad (17)$$

where,  $N$  represents the size of the test dataset.  $r_{ij}$  and  $\hat{r}_{i,j}$  represent the actual and predicted ratings, respectively. Because both metrics indicate the degree of error, lower values suggest a better predictive performance of the model.

### C. BASELINE MODELS

To verify the effectiveness of the proposed model, we compared its performance with that of popular baseline models. These baselines are categorized into three groups: (1) Recommendation models based solely on ratings (i.e., CF and PMF), (2) recommendation models using review text (i.e., HFT and RARV2), (3) recommendation models using the attention approach (i.e., NARRE and RSBM), and recommendation models using aspect-based approach (i.e., A3NCF and UCAM). A brief description of the baseline models is provided below:

- CF [20]: The collaborative filtering is one of the successful recommendation algorithms among early models.

The items are recommended based on the nearest neighbors by measuring the similarities based on the user's rating histories.

- PMF [43]: The probabilistic matrix factorization, a variant of the matrix factorization model, is effective for sparse and imbalanced rating data. The model decomposes a rating matrix into latent structures based on Gaussian prior distribution.
- HFT [23]: The hidden factors and hidden topic model employ LDA technique to extract hidden topics from the aggregated reviews of users and items. This model uses the fusion of the hidden topics and hidden factors from matrix factorization.
- NARRE [44]: The neural attentional rating regression with review-level explanations model is a recommendation model which can provide highly useful reviews by introducing a neural attention mechanism. The authors adopt matrix factorization to obtain user/item representations from ratings. Then a CNN model learns the importance of reviews for better results at the final prediction layer.
- RARV2 [12]: This recommendation model improves the representation of review texts by using a multi-embedding method. The text embeddings from BERT and RoBERTa are used as auxiliary information in the deep matrix factorization method.
- RSBM [8]: The review semantics-based model extracts the semantics of user's evaluation in the review using a CNN. Then it uses an attention network to represent users' evaluation action on each aspect for the final rating prediction.
- A3NCF [45]: A novel aspect-aware recommender model designed to capture user's varying attentions to different aspects of items. This model utilizes topic modeling to extract user's preferences and item's characteristics for the personalized recommendations.
- UCAM [46]: This recommendation model uses a deep neural network, utilizing unstructured textual information derived from reviews. It concatenates user-item interactions with the corresponding review information. The review representation is generated using BERT-based ABSA, extracting aspect-specific sentiments.

#### D. HYPERPARAMETER SETTINGS

This study used the TensorFlow framework with the Python programming language for the experiment. The experiments were conducted in a computerized environment with an Intel®Core™i9-10900K CPU, 128 GB memory, and NVIDIA GeForce RTX 3090 GPU. All experimental results are given as mean values of 5 experimental runs. Additionally, the random seed was set to 42 to ensure consistent performance.

The validation set was used to determine the optimal hyperparameter values for the proposed model. The batch size was

set to 64 from {32, 64, 128, 256, 512, 1024}. A learning rate of 0.001 was applied from {0.001, 0.0001, 0.00001}, and an adaptive moment estimation (Adam) optimizer [24] was used. An embedding size of 64 was set to represent the latent factors of users and items from {8, 16, 32, 64, 128}. The number of dense layers for the MLP was set to 3 from {1, 2, 3, 4}, including the embedding layer; thus, the size of each layer in the proposed model decreases in the order of 128 → 64 → 32 [26].

A detailed sensitivity analysis is presented in Section VI-D. The hyperparameters for the baseline models were applied as suggested by the original papers, and the parameters not mentioned were set to be the same as in the proposed model.

## VI. RESULTS AND DISCUSSION

### A. PERFORMANCE COMPARISON (RQ 1)

Tables 2 presents the performance comparisons of MAE and RMSE for the three datasets, with the results of AARN highlighted with bold font. The performance comparison showed the following major observations. First, the proposed model consistently outperformed all baseline models. These results suggest that the proposed model can effectively predict user's preferences because it considers various aspects embedded in review texts.

Second, the conventional models that uses rating information as only source of information (i.e., CF and PMF) exhibited lower performances than other baseline models. These results suggest that the data sparsity problem can be alleviated by using reviews as a supplement to ratings. For example, HFT extracts hidden topics from review texts using topic modeling and combines them with user/item latent factors. The lower MAE of HFT shows that the quality of the user/item latent factor is enhanced by integrating review features.

Third, among the baseline models that use review information, HFT and A3NCF showed comparatively lower performances. These models utilize conventional text analysis to represent review information. These results suggest that advanced text analysis techniques, such as CNN and BERT, contribute to improved recommendation performance when representing review text. Additionally, RSBM uses an individual review as input for the recommendation; the authors pointed out the problem that previous approaches consider aggregations of reviews as long documents because different users have different preferences for items. Similarly, RARV2 applies multi-embedding techniques to an individual review written by a user on a particular item and uses this as additional information. The superior performances of RSBM, RARV2, and AARN suggest that the approach of using individual reviews as additional information is effective for personalized recommendations.

Finally, the proposed AARN, showed superior performance compared to baseline models that use aspect-specific information (i.e., NARRE, RSBM, A3CNF, and UCAM).

**TABLE 2.** Performance comparison with baseline models.

| Dataset         | Model         | MAE          | RMSE         |
|-----------------|---------------|--------------|--------------|
| Yelp restaurant | CF            | 0.933        | 1.214        |
|                 | PMF           | 0.989        | 1.076        |
|                 | HFT           | 0.829        | 1.105        |
|                 | RARV2         | 0.564        | 0.751        |
|                 | NARRE         | 0.827        | 1.072        |
|                 | RSBM          | 0.542        | 0.763        |
|                 | A3NCF         | 0.994        | 1.224        |
|                 | UCAM          | 0.596        | 0.795        |
|                 | <b>AARN</b>   | <b>0.486</b> | <b>0.654</b> |
|                 | Amazon tablet | CF           | 1.067        |
| PMF             |               | 1.041        | 1.044        |
| HFT             |               | 1.095        | 1.344        |
| RARV2           |               | 0.694        | 0.833        |
| NARRE           |               | 0.946        | 1.317        |
| RSBM            |               | 0.543        | 0.800        |
| A3NCF           |               | 1.032        | 1.363        |
| UCAM            |               | 0.650        | 0.983        |
| <b>AARN</b>     |               | <b>0.475</b> | <b>0.728</b> |
| Amazon laptop   |               | CF           | 1.291        |
|                 | PMF           | 1.003        | 1.008        |
|                 | HFT           | 0.901        | 1.297        |
|                 | RARV2         | 0.655        | 0.809        |
|                 | NARRE         | 1.023        | 1.373        |
|                 | RSBM          | 0.578        | 0.799        |
|                 | A3NCF         | 0.992        | 1.307        |
|                 | UCAM          | 0.767        | 0.997        |
|                 | <b>AARN</b>   | <b>0.509</b> | <b>0.764</b> |

These results indicate that advanced ABSA is more effective in capturing user preferences than general text analysis techniques for personalized recommendations. Although UCAM uses review representations generated by BERT-based ABSA, it simply concatenates this valuable information to the prediction layer. On the other hands, AARN combines advanced ABSA with an attention mechanism to explicitly analyze the aspect-specific preferences embedded in the review texts

### B. MODEL COMPONENTS ANALYSIS (RQ 2)

Table 3 presents the results of the ablation analysis of the attention mechanism in the proposed model. “Without attention” refers to a model excluding the attention mechanism from the proposed model. In this model, aspect-specific sentiment scores from ABSA are directly concatenated with the user/item latent factors. The experimental results demonstrate the impact of the attention mechanism. The lower MAE and RMSE suggest that the attention mechanism is more effective than simply applying fully connected layers to capture the aspect-specific information from review texts.

**TABLE 3.** The results of the component analysis.

| Dataset         | Without attention |       | With attention |       |
|-----------------|-------------------|-------|----------------|-------|
|                 | MAE               | RMSE  | MAE            | RMSE  |
| Yelp restaurant | 0.501             | 0.662 | 0.486          | 0.654 |
| Amazon tablet   | 0.490             | 0.733 | 0.475          | 0.727 |
| Amazon laptop   | 0.526             | 0.773 | 0.509          | 0.765 |

### C. TEXT EMBEDDING EFFICIENCY ANALYSIS (RQ 3)

This section presents a novel experimental design to verify that the ABSA used in the proposed model captures detailed information from review texts better than other text analysis techniques. According to the literature, users typically express different preferences on various aspects [4], [8], [17]. This tendency is particularly noticeable in reviews given a rating of 3, which is the median value on the rating scale of [1:5]. Therefore, extracting user preferences from such reviews is challenging. Meanwhile, user preferences are well identified in reviews with extreme ratings, such as ratings of 1, 2, 4, or 5 [47].

This study investigated the performance of capturing detailed information regarding various aspects by separately evaluating only the ratings of 3 in the test dataset. The following text analysis techniques were adopted for comparison with the proposed model: Doc2Vec [11], BERT [10] and RoBERTa [12]. Note that Doc2Vec, BERT, and RoBERTa only derive single text embeddings from reviews; therefore, an attention mechanism is not applied, and all other settings for the recommendation process are the same. Table 4 presents a comparison of the MAE for the original and divided test datasets.

**TABLE 4.** Impact of granular text embedding (MAE).

| Text Embedding method | The performance for the original test set | The performance for the divided test set |
|-----------------------|---|--|
| Doc2Vec               | 0.692                                     | 0.990                                    |
| BERT                  | 0.585                                     | 0.878                                    |
| RoBERTa               | 0.576                                     | 0.823                                    |
| Proposed              | 0.509                                     | 0.604                                    |

The recommendation model applying Doc2Vec shows a significant increase in MAE (43.06%) when predicting only the rating of 3 in the original test dataset. This result implies that the model performs excellent prediction for the extreme ratings (i.e., 1, 2, 4, and 5) in the test dataset but performs poorly for the rating of 3 in the test dataset. BERT and RoBERTa also showed acceptable MAEs for the original test dataset, but their performances significantly deteriorated with a rating of 3 in the test dataset (50.09% and 42.88%, respectively). This suggests that single-dimensional embeddings are limited in reflecting diverse user preferences. The proposed model showed a slight decrease on the rating of 3 test dataset but reported robust results (18.66%) compared to other text



analysis techniques. These results suggest that more detailed preferences can be derived by using ABSA. Therefore, even if a user evaluates certain aspects positively and others negatively, the proposed model is effective in capturing diverse user preferences.

#### D. IMPACT OF HYPERPARAMETERS (RQ 4)

Figure 4 presents the results of the sensitivity analysis. In Figure 4(a), “MLP-k” indicates that the generating prediction phase consists of “k” hidden layers, excluding the input layer, and the size of each layer is halved as the layers deepen [26]. The inputs of the MLP, i.e., user/item attentive representations, have the same layer size as the user/item latent factors. As shown in (a), the MAE increases as the number of layers increases because overfitting occurs. All datasets showed optimal results for MLP-2. Therefore, the proposed model applies two hidden layers for the rating prediction.

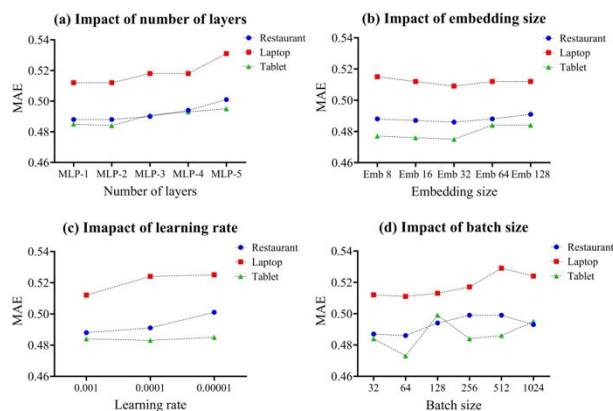


FIGURE 4. Sensitivity analysis of hyperparameters.

In (b), Impact of embedding size, MAE decreases as the embedding size increases because the latent factors of the users and items can be expressed more precisely [2]. However, overfitting also occurs starting from Emb-64 in all datasets, and MAE begins to increase. Consequently, the optimal embedding size for the latent factors is set to 32. Therefore, the prediction generation phase consists of an input layer and two hidden layers, and the layer dimensions are in the order of  $128 \rightarrow 64 \rightarrow 32$ .

The learning rate, which updates the model weights, is set to ensure optimal convergence of the neural networks to the best result during training [12]. Figure 4 (c), Impact of learning rate, shows that the MAE is lowest at 0.001, and MAE increases as the learning rate decreases. This suggests that the convergence of the proposed model to the optimal result is limited when the learning rate is set too low. The batch size was optimized at 64 for all three datasets, as shown in (d), Impact of batch size. This is consistent with the literature, which claims that neural networks are affected by noise when the batch size is too large or small [48].

## VII. CONCLUSION

This study proposed a novel recommendation model designed to capture diverse aspects of user preferences. In contrast to conventional approaches that primarily focus on users' general preferences, the proposed approach considers detailed preferences for various aspects embedded in the review text. Notably, this study pioneers the integration of a BERT-based ABSA with an attention mechanism in a recommendation model. Comprehensive experimental results showed that the proposed model significantly outperformed other recommendation models.

Moreover, practical implications could be suggested from our findings. The application of ABSA to evaluate aspect-specific sentiments provides richer and more granular insights into user behavior patterns compared to alternative text embedding techniques. Recommendations derived from these granular preferences are more likely to align with users' interests, thereby supporting more informed decision-making. This method can also offer users a reason for making recommendations by mapping the main aspects of users and items. Online platforms can construct personalized recommendation pages using the proposed method. Thus, users can search for their desired items with significantly lower search efforts and communication costs, and webpage administrators can encourage users to write more online reviews by providing improved recommendations.

This study has several limitations that should be addressed in future research. The proposed approach uses reviews as additional information. However, it is possible to consider other information, such as the location or demographic information of users. Future research could enhance personalization to capture the unique preferences of users by using richer information. Furthermore, this study conducted experiments using datasets from the online platforms Yelp and Amazon. However, future research could extend the dataset domain. Finally, the proposed model uses an MLP-based neural network architecture to capture the latent interactions between users and items; however, future research could consider structuring predictive modeling with more advanced algorithms, such as graph methods. Developing a model that suits the domain while improving its performance is important for future works.

## REFERENCES

- [1] C.-L. Liao and S.-J. Lee, “A clustering based approach to improving the efficiency of collaborative filtering recommendation,” *Electron. Commerce Res. Appl.*, vol. 18, pp. 1–9, Jul. 2016.
- [2] Q. Li, D. Jang, D. Kim, and J. Kim, “Restaurant recommendation model using textual information to estimate consumer preference: Evidence from an online restaurant platform,” *J. Hospitality Tourism Technol.*, vol. 14, no. 5, pp. 857–877, Aug. 2023.
- [3] P. Kumar and B. Bhasker, “DNNRec: A novel deep learning based hybrid recommender system,” *Expert Syst. Appl.*, vol. 144, Apr. 2020, Art. no. 113054.
- [4] L. Sun, J. Guo, and Y. Zhu, “A multi-aspect user-interest model based on sentiment analysis and uncertainty theory for recommender systems,” *Electron. Commerce Res.*, vol. 20, no. 4, pp. 857–882, Dec. 2020.

- [5] Y. Ma, G. Chen, and Q. Wei, "Finding users preferences from large-scale online reviews for personalized recommendation," *Electron. Commerce Res.*, vol. 17, no. 1, pp. 3–29, Mar. 2017.
- [6] Z. Zhu, M. Yan, X. Deng, and M. Gao, "Rating prediction of recommended item based on review deep learning and rating probability matrix factorization," *Electron. Commerce Res. Appl.*, vol. 54, Jul. 2022, Art. no. 101160.
- [7] R. Duan, C. Jiang, and H. K. Jain, "Combining review-based collaborative filtering and matrix factorization: A solution to rating's sparsity problem," *Decis. Support Syst.*, vol. 156, May 2022, Art. no. 113748.
- [8] R. Cao, X. Zhang, and H. Wang, "A review semantics based model for rating prediction," *IEEE Access*, vol. 8, pp. 4714–4723, 2020.
- [9] L. Zheng, V. Noroozi, and P. S. Yu, "Joint deep modeling of users and items using reviews for recommendation," in *Proc. 10th ACM Int. Conf. Web Search Data Mining*, Cambridge, U.K., Feb. 2017, pp. 425–434.
- [10] A. L. Karn, R. K. Karna, B. R. Kondamudi, G. Bagale, D. A. Pustokhin, I. V. Pustokhina, and S. Sengan, "Customer centric hybrid recommendation system for e-commerce applications by integrating hybrid sentiment analysis," *Electron. Commerce Res.*, vol. 23, no. 1, pp. 279–314, Mar. 2023.
- [11] N. Ghasemi and S. Momtazi, "Neural text similarity of user reviews for improving collaborative filtering recommender systems," *Electron. Commerce Res. Appl.*, vol. 45, Jan. 2021, Art. no. 101019.
- [12] Y.-H. Liu, Y.-L. Chen, and P.-Y. Chang, "A deep multi-embedding model for mobile application recommendation," *Decis. Support Syst.*, vol. 173, Oct. 2023, Art. no. 114011.
- [13] A. Da'ud, N. Salim, I. Rabiui, and A. Osman, "Recommendation system exploiting aspect-based opinion mining with deep learning method," *Inf. Sci.*, vol. 512, pp. 1279–1292, Feb. 2020.
- [14] L. Qiu, S. Gao, W. Cheng, and J. Guo, "Aspect-based latent factor model by integrating ratings and reviews for recommender system," *Knowl.-Based Syst.*, vol. 110, pp. 233–243, Oct. 2016.
- [15] P. Hajek, L. Hikkerova, and J.-M. Sahut, "Fake review detection in e-commerce platforms using aspect-based sentiment analysis," *J. Bus. Res.*, vol. 167, Nov. 2023, Art. no. 114143.
- [16] C. Yang, X. Yu, Y. Liu, Y. Nie, and Y. Wang, "Collaborative filtering with weighted opinion aspects," *Neurocomputing*, vol. 210, pp. 185–196, Oct. 2016.
- [17] C.-H. Lai and K.-C. Tseng, "Applying deep learning models to analyze users' aspects, sentiment, and semantic features for product recommendation," *Appl. Sci.*, vol. 12, no. 4, p. 2118, Feb. 2022.
- [18] B. Zeng, H. Yang, R. Xu, W. Zhou, and X. Han, "LCF: A local context focus mechanism for aspect-based sentiment classification," *Appl. Sci.*, vol. 9, no. 16, p. 3389, Aug. 2019.
- [19] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions," *IEEE Trans. Knowl. Data Eng.*, vol. 17, no. 6, pp. 734–749, Jun. 2005.
- [20] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl, "Item-based collaborative filtering recommendation algorithms," in *Proc. 10th Int. Conf. World Wide Web*, Apr. 2001, pp. 285–295.
- [21] Q. Li, X. Li, B. Lee, and J. Kim, "A hybrid CNN-based review helpfulness filtering model for improving e-commerce recommendation service," *Appl. Sci.*, vol. 11, no. 18, p. 8613, Sep. 2021.
- [22] X. Li, Q. Li, and J. Kim, "A review helpfulness modeling mechanism for online e-commerce: Multi-channel CNN end-to-end approach," *Appl. Artif. Intell.*, vol. 37, no. 1, Jan. 2023, Art. no. 2166226.
- [23] J. McAuley and J. Leskovec, "Hidden factors and hidden topics: Understanding rating dimensions with review text," in *Proc. 7th ACM Conf. Recommender Syst.*, Oct. 2013, pp. 165–172.
- [24] T. Zhang, C. Sun, Z. Cheng, and X. Dong, "AENAR: An aspect-aware explainable neural attentional recommender model for rating prediction," *Expert Syst. Appl.*, vol. 198, Jul. 2022, Art. no. 116717.
- [25] L. N. H. Nam, "Incorporating textual reviews in the learning of latent factors for recommender systems," *Electron. Commerce Res. Appl.*, vol. 52, Mar. 2022, Art. no. 101133.
- [26] X. He, L. Liao, H. Zhang, L. Nie, X. Hu, and T.-S. Chua, "Neural collaborative filtering," in *Proc. 26th Int. Conf. World Wide Web (WWW)*, vol. 2017, pp. 173–182.
- [27] T. Ali, B. Marc, B. Omar, K. Soulaïmane, and S. Larbi, "Exploring destination's negative e-reputation using aspect based sentiment analysis approach: Case of Marrakech destination on TripAdvisor," *Tourism Manage. Perspect.*, vol. 40, Oct. 2021, Art. no. 100892.
- [28] E. Asani, H. Vahdat-Nejad, and J. Sadri, "Restaurant recommender system based on sentiment analysis," *Mach. Learn. Appl.*, vol. 6, Dec. 2021, Art. no. 100114.
- [29] R. L. Rosa, G. M. Schwartz, W. V. Ruggiero, and D. Z. Rodríguez, "A knowledge-based recommendation system that includes sentiment analysis and deep learning," *IEEE Trans. Ind. Informat.*, vol. 15, no. 4, pp. 2124–2135, Apr. 2019.
- [30] P. Mehra, "Unexpected surprise: Emotion analysis and aspect based sentiment analysis (ABSA) of user generated comments to study behavioral intentions of tourists," *Tourism Manage. Perspect.*, vol. 45, Jan. 2023, Art. no. 101063.
- [31] H. Li, B. X. B. Yu, G. Li, and H. Gao, "Restaurant survival prediction using customer-generated content: An aspect-based sentiment analysis of online reviews," *Tourism Manage.*, vol. 96, Jun. 2023, Art. no. 104707.
- [32] C. Sun, L. Huang, and X. Qiu, "Utilizing BERT for aspect-based sentiment analysis via constructing auxiliary sentence," in *Proc. Conf. North Amer. Chapter Assoc. Comput. Linguistics*, Minneapolis, MN, USA, 2019, pp. 380–385.
- [33] Y. Luo, L. Tang, E. Kim, and X. Wang, "Finding the reviews on yelp that actually matter to me: Innovative approach of improving recommender systems," *Int. J. Hospitality Manage.*, vol. 91, Oct. 2020, Art. no. 102697.
- [34] H. Xu, B. Liu, L. Shu, and P. S. Yu, "BERT post-training for review reading comprehension and aspect-based sentiment analysis," in *Proc. Conf. North Amer. Chapter Assoc. Comput. Linguistics*, Minneapolis, MN, USA, 2019, pp. 2324–2335.
- [35] M. Hoang, O. A. Bihorac, and J. Rouces, "Aspect-based sentiment analysis using bert," in *Proc. 22nd Nordic Conf. Comput. Linguistics*, Turku, Finland, 2019, pp. 187–196.
- [36] Y. Tay, D. Bahri, D. Metzler, D.-C. Juan, Z. Zhao, and C. Zheng, "Synthesizer: Rethinking self-attention for transformer models," in *Proc. 38th Int. Conf. Mach. Learn.*, 2021, pp. 10183–10192.
- [37] Y. Hou, N. Yang, Y. Wu, and P. S. Yu, "Explainable recommendation with fusion of aspect information," *World Wide Web*, vol. 22, no. 1, pp. 221–240, Jan. 2019.
- [38] X. Guan, Z. Cheng, X. He, Y. Zhang, Z. Zhu, Q. Peng, and T.-S. Chua, "Attentive aspect modeling for review-aware recommendation," *ACM Trans. Inf. Syst.*, vol. 37, no. 3, pp. 1–27, Mar. 2019.
- [39] C.-H. Lai and C.-Y. Hsu, "Rating prediction based on combination of review mining and user preference analysis," *Inf. Syst.*, vol. 99, Jul. 2021, Art. no. 101742.
- [40] J. Ni, J. Li, and J. McAuley, "Justifying recommendations using distantly-labeled reviews and fine-grained aspects," in *Proc. Conf. Empirical Methods Natural Lang. Process. 9th Int. Joint Conf. Natural Lang. Process.*, 2019, pp. 188–197.
- [41] T. Chai and R. R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE)?—Arguments against avoiding RMSE in the literature," *Geosci. Model Develop.*, vol. 7, no. 3, pp. 1247–1250, Jun. 2014.
- [42] M. Lee, W. Kwon, and K.-J. Back, "Artificial intelligence for hospitality big data analytics: Developing a prediction model of restaurant review helpfulness for customer decision-making," *Int. J. Contemp. Hospitality Manage.*, vol. 33, no. 6, pp. 2117–2136, Aug. 2021.
- [43] A. Mnih and R. R. Salakhutdinov, "Probabilistic matrix factorization," in *Proc. 20th Int. Conf. Neural Inf. Process. Syst.*, Vancouver, BC, Canada, 2007, pp. 1257–1264.
- [44] C. Chen, M. Zhang, Y. Liu, and S. Ma, "Neural attentional rating regression with review-level explanations," in *Proc. World Wide Web Conf.*, Lyon, France, 2018, pp. 1583–1592.
- [45] Z. Cheng, Y. Ding, X. He, L. Zhu, X. Song, and M. Kankanhalli, "A3NCF: An adaptive aspect attention model for rating prediction," in *Proc. 27th Int. Joint Conf. Artif. Intell.*, Stockholm, Sweden, 2018, pp. 3748–3754.
- [46] M. Unger, A. Tuzhilin, and A. Livne, "Context-aware recommendations based on deep learning frameworks," *ACM Trans. Manage. Inf. Syst.*, vol. 11, no. 2, pp. 1–15, May 2020.
- [47] Y. Zhao, X. Xu, and M. Wang, "Predicting overall customer satisfaction: Big data evidence from hotel online textual reviews," *Int. J. Hospitality Manage.*, vol. 76, pp. 111–121, Jan. 2019.
- [48] J. Janke, M. Castelli, and A. Popović, "Analysis of the proficiency of fully connected neural networks in the process of classifying digital images. Benchmark of different classification algorithms on high-level image features from convolutional layers," *Expert Syst. Appl.*, vol. 135, pp. 12–38, Nov. 2019.



**SIGEON YANG** received the B.S. degree in hospitality management from Sejong University, in 2020. He is currently pursuing the M.S. degree in big data analytics with Kyung Hee University. His research interests include text mining, recommender systems, big data analytics, and machine learning.



**HAEBIN LIM** received the B.S. degree in big data from Daegu University, in 2023. He is currently pursuing the M.S. degree with Kyung Hee University. His research interests include recommender systems, big data analysis, and natural language processing.



**QINGLONG LI** received the B.S. degree in business administration from Kyung Hee University and the M.S. degree from the Department of Big Data Analytics, Kyung Hee University, where he is currently pursuing the Ph.D. degree. His current research interests include text mining, big data analysis, recommender systems, and machine learning.



**JAEKYEONG KIM** received the B.S. degree in industrial engineering from Seoul National University and the M.S. and Ph.D. degrees in management information systems (MIS) from the Korea Advanced Institute of Science and Technology (KAIST). He is currently a Professor with the School of Management, Kyung Hee University. He has published several articles, which have appeared in *IEEE TRANSACTIONS ON SYSTEMS, MAN AND CYBERNETICS*, *IEEE TRANSACTIONS ON SERVICES COMPUTING*, *Information and Management*, *International Journal of Information Management*, *Applied Artificial Intelligence*, *Artificial Intelligence Review*, *International Journal of Human-Computer Studies*, *Expert Systems with Applications*, and *Group Decision and Negotiations*. His current research interests include business analytics, recommender systems, big data analysis, and deep learning.

...