Article

Improving Matrix Factorization Based Expert Recommendation for Manuscript Editing Services by Refining User Opinions with Binary Ratings

Yeonbin Son and Yerim Choi *

Department of Industrial and Management Engineering, Kyonggi University, 154-42 Gwanggyosan-ro, Yeongtong-gu, Suwon-si 16227, Korea; ybson@kgu.ac.kr

* Correspondence: yrchoi@kgu.ac.kr

Received: 21 April 2020; Accepted: 11 May 2020; Published: 14 May 2020

Abstract: As language editing became an essential process for enhancing the quality of a research manuscript, there are several companies providing manuscript editing services. In such companies, a manuscript submitted for proofreading is matched with an editing expert through a manual process, which is costly and often subjective. The major drawback of the manual process is that it is almost impossible to consider the inherent characteristics of a manuscript such as writing style and paragraph composition. To this end, we propose an expert recommendation method for manuscript editing services based on matrix factorization, a well-known collaborative filtering approach for learning latent information in ordinal ratings given by users. Specifically, binary ratings are utilized to substitute ordinal ratings when negative opinions are expressed by users since negative opinions are more accurately expressed by binary ratings than ordinal ratings. From the experiments using a real-world dataset, the proposed method outperformed the rest of the compared methods with an RMSE (root mean squared error) of 0.1. Moreover, the effectiveness of substituting ordinal ratings with binary ratings was validated by conducting sentiment analysis on text reviews.

Keywords: research manuscript editing service; editing expert recommender system; collaborative filtering; matrix factorization; user feedback; ordinal ratings; binary ratings

1. Introduction

Language editing has become one of the fundamental processes for journal submission of a research manuscript, which is an article conveying researcher's academic achievements and opinions. As most international journals require researchers to submit their manuscripts in English, there are possibilities of grammatical errors, syntactical errors and expression ambiguities for non-native English speakers. Moreover, writing a research manuscript is difficult even for native English speakers since the manuscripts need to be logically organized to clearly convey complex ideas [1].

Thus, a number of companies provide language editing services for research manuscripts [2–4]. The editing services in such companies are usually processed through the following steps. When a user asks a company to proofread his/her manuscript, a matching manager finds an appropriate editing expert from its expert pool for the manuscript. Then, the selected expert is asked to proofread the manuscript. When the expert finishes proofreading, the edited version of the manuscript is returned to the user. In particular, the matching between a manuscript and an expert is performed manually by a human manager through identifying the research field of the manuscript and comparing it with the expert areas of editing experts.

The current manual matching system has several drawbacks. First, it is time-consuming and costly since a human manager finds an appropriate editing expert for a manuscript by comparing one by one. Second, the matching is often subjective as the selection is done based on the manager's knowledge

and experience. Third, there is a possibility of human errors and inconsistencies. More importantly, the inherent characteristics of a manuscript, such as literary style and paragraph composition, cannot be considered as sufficient knowledge and a lot of time is required to identify these characteristics.

We learned from the analysis of the review texts collected from a manuscript editing service that the major inadequacies mentioned by users are closely related to the aforementioned inherent characteristics of a manuscript. A co-occurrence network of keywords shown in Figure 1 is constructed by analyzing 275 review texts with ordinal ratings under 3 (out of 5) among a total of 1928 reviews. Note that ordinal rating is a numerical expression of a user's opinion on an item such as a product or a service in a predefined scale, normally one to five (a higher number indicates a more positive opinion) [5] and review text is a brief comment written by a user to express his/her opinion on an item. In Figure 1, keywords and their frequencies are respectively represented by nodes and their sizes, and the thickness of an edge indicates the cosine similarity [6] between the two connected nodes. An edge between two keywords is thicker when the words more frequently co-occur. Three negative terms—'shortage', 'insufficiency', and 'error'—and keywords closely connected with the terms are highlighted to explore the users' negative opinions. According to the network, it can be found that the keywords related to the inherent characteristics of a manuscript such as 'composition', 'structure', 'context', and 'content' were frequently mentioned with the negative terms.

Figure 1. Co-occurrence network of keywords constructed by analyzing negative review texts (ordinal ratings under 3) collected from a manuscript editing service. The size of a node and the thickness of an edge respectively represent the frequency of a keyword and the co-occurrence frequency of the two connected keywords.

It has been demonstrated that the inherent characteristics of items can be successfully exploited for user opinion inference by using matrix factorization (MF) [7], which is one of the popular collaborative filtering methods for recommender systems [8]. Zheng et al. [9] adopted MF for music recommendation in order to consider the latent features of music, and Yin et al. [10] utilized MF for service recommendation to investigate the implicit associations among users and services. MF learns the characteristics of users and items from a matrix composed of the ordinal ratings (called feedback rating matrix in the rest of manuscript), which are further used to approximate unknown ratings in the feedback rating matrix [11]. Specifically, the feedback rating matrix is factorized into user and item latent matrices, and the factorized matrices are multiplied to build an estimated feedback rating matrix that does not have unknown ratings [12].

However, MF usually suffers from a data sparsity problem [13–15], which arises when the number of ordinal ratings is small and the ratings are concentrated to certain users or items. The problem is even worse for the ordinal ratings in manuscript editing services as the number of manuscripts that a researcher has written is usually small and not all manuscripts receive an editing service. Moreover, a huge portion of users do not leave their feedback in the form of ordinal ratings.

Several attempts were made to address the data sparsity problem of ordinal ratings by additionally using indirect user feedback such as review texts and behavior logs. For example, Jiang et al. [16] and Chu et al. [17] utilized review texts for a car recommendation and item recommendation, respectively, and Lian et al. [18] adopted users' visit frequencies for a location recommendation. However, extracting user opinions from the indirect feedback is usually costly and difficult as it requires a number of complex processes such as data collection and preprocessing [19].

On the other hand, there are a few studies that tried to utilize a type of direct user feedback called binary ratings to supplement ordinal ratings. Binary ratings are an indication of a user's opinion in a binary form, whether a user liked the service or not. Previous research reported that the binary ratings provide intuitive and accurate information about a user's opinion [20–22]. Moreover, binary ratings are relatively easy to collect compared to other types of user feedback, such as ordinal ratings and review texts, since users can express their opinions by simply clicking a button—like or dislike [23]. For this reason, binary ratings were exploited with ordinal ratings simultaneously to infer a user's opinion in previous studies. Pan and Yang [24] proposed a factorization method which finds a latent matrix that has information depending on both ordinal and binary ratings and tried to estimate unknown feedback ratings by utilizing the matrix. In addition, Pan et al. [22] suggested transfer by mixed factorization (TMF) to incorporate binary ratings with ordinal ratings by adding an extra component to the conventional MF.

Most manuscript editing services collect both ordinal and binary ratings in order to monitor service quality and provide a function for users to favor or exclude a certain editing expert. Figure 2 shows screenshots of two web pages, where users can leave their feedback on the received services. A user can leave a binary rating—like or dislike—for an editing expert on the page shown in Figure 2a. When a user selects the like button for an expert, the expert will be assigned priority for future services. In addition, a user can give an ordinal rating with a review text to an expert at the page shown in Figure 2b, which can be reached by clicking 'leave feedback' button in Figure 2a.

Order	Service	Total	Status	Action
#38358 Jan 27, 2018	Academic/Dissertation Editing Service Editor : James	\$115.55	Completed	Documents Download Leave Feedback

(a) Web page for leaving binary ratings

(b) Web page for leaving ordinal ratings and review text

Figure 2. Screenshots of two web pages, where users leave their feedback, (a) binary ratings, (b) ordinal ratings and review texts, on the received editing services.

While exploring user feedback logs collected from a manuscript editing service, which are composed of ordinal ratings, binary ratings and review texts, we were able to observe the limitations of ordinal ratings on expressing negative opinions. Figure 3a,b show boxplots representing the distributions of sentiment scores (x -axis) of review texts according to the ordinal and binary ratings (y -axis), respectively. There was a positive correlation between binary ratings and sentiment scores as shown in Figure 3b, while no significant correlation was found between ordinal ratings and sentiment scores as shown in Figure 3a.

Figure 3. Boxplots representing distributions of review texts' sentiment scores according to (a) ordinal ratings and (b) binary ratings.

For in-depth analysis, we explored review texts given by users who clicked dislike button while giving ordinal ratings of 4 or 5. In most cases, review texts contained negative opinions as shown in Table 1. For instance, users *D* and *E* expressed their dissatisfaction by clicking the dislike button and criticized expert's misunderstanding of manuscripts in the review texts while leaving 4-point ordinal ratings. This implies that ordinal ratings may have a bias in expressing a user's negative opinion. Therefore, it is possible to enhance the performances of editor recommendation through refining user opinions by not just incorporating ordinal and binary ratings together but utilizing them selectively, making a feedback rating matrix for performing MF to have more reliable information.

Table 1. Examples of user feedback logs composed of binary ratings, ordinal ratings, and review texts, where the category of binary ratings are dislike while the those of ordinal ratings are 4 or 5.

User	Binary Rating	Ordinal Rating	Review Text
A	Dislike	5	Your editing result is not what I intended.. I was a little disappointed T.T
B	Dislike	4	I always got the results I asked for 3-4 hours ago, but in this time, i got the edited manuscript just on time. It was later than i expected, so I felt uncomfortable. Also, requests to convert all sentences to colloquialism were not reflected.
C	Dislike	4	There is a sentence edited differently from my intention. The editing expert proofread the manuscript to his style, and i was disappointed.
D	Dislike	4	If the editing expert concentrated in considering the sentence structure in the whole paragraph when he edited, it is possible to reflect my intention all. Experts only changed sentences with awkward expressions.
E	Dislike	4	Understand the context and structure of the manuscript first. And then edit.
F	Dislike	4	It seems like the editor used translator. I asked the editing service for conveying the meaning of a sentence accurately. Not a literal translation.

To this end, we propose an MF based expert recommender system for manuscript editing services. MF is adopted to explore the inherent characteristics of a manuscript such as writing style and paragraph composition, which are difficult for humans to detect. Moreover, ordinal and binary ratings are selectively utilized to refine user opinions and alleviate data sparsity problem. The two types of user feedback are combined in various ways to maximize the recommendation performances.

Specifically, the proposed method is composed of three steps. First, a feedback rating matrix is constructed by combining ordinal and binary ratings. Second, user opinions are inferred by performing MF on the feedback rating matrix. Lastly, the optimal editing expert is selected for a user based on the result of the second step.

The rest of paper is organized as follows. The proposed method is introduced and explained in Section 2. Section 3 presents experiment settings and results. In Section 4, guidelines for the application of the proposed method to real-world services are suggested, and the paper is concluded in Section 5.

2. Matrix Factorization Based Editing Expert Recommender System

2.1. Problem Definition

The proposed method attempts to recommend the optimal editing expert to a user based on the inferred user's opinion by analyzing previously collected user feedback logs to editing experts. We denote a set of users by $U = \{u_1, \dots, u_i, \dots, u_{n_u}\}$, where u_i and n_u respectively represent the i -th user and the total number of users, and a set of editing experts by $E = \{e_1, \dots, e_j, \dots, e_{n_e}\}$, where e_j and n_e respectively indicate the j -th editing expert and the total number of editing experts.

Two types of user feedback, ordinal and binary ratings, are utilized for the user opinion inference. Ordinal ratings given by U to E compose an ordinal rating matrix $O = [o_{i,j}]$, where $o_{i,j}$ refers to the ordinal rating given by u_i to e_j and is an integer ranging from 1 to 5 in 1 point interval.

Binary ratings given by U to E compose a binary rating matrix $B = [b_{i,j}]$, where $b_{i,j}$ indicates whether u_i liked the editing result provided by e_j or disliked it. $b_{i,j}$ is 1 when u_i clicked the like button for e_j or -1 when u_i clicked the dislike button for e_j as described in Equation (1).

$$b_{i,j} = \begin{cases} 1 & \text{if } u_i \text{ clicked the like button for } e_j \\ -1 & \text{if } u_i \text{ clicked the dislike button for } e_j \\ \emptyset & \text{if } u_i \text{ did not rate } e_j, \end{cases} \quad (1)$$

where $\emptyset$ indicates null. Note that $o_{i,j}$ and $b_{i,j}$ are null when u_i did not rate the editing service provided by e_j .

A feedback rating matrix $F = [f_{i,j}]$ is constructed by combining O and B , where $f_{i,j}$ is a feedback rating indicating the degree to which u_i prefers e_j ranging from 1 to 5. When u_i did not give any ratings to e_j , $f_{i,j}$ is null. The dimension of O , B , and F is $n_u \times n_e$. However, there are many unknown $f_{i,j}$ as there are limited number of feedback from U to E . In summary, given O and B , we try to find the optimal editing expert e_{j^*} for u_{i^*} who requests an editing service by using the estimated feedback rating matrix $\hat{F}$ which does not contain null elements and is approximated by analyzing F .

2.2. Overview

The proposed method is composed of three steps as illustrated in Figure 4. First, F is constructed by utilizing O and B . Second, unknown feedback ratings in F are estimated by performing MF. Lastly, the optimal editing expert is recommended to a user who requests the editing service based on the results from the previous step.

Figure 4. Overview of the proposed non-negative matrix factorization based editing expert recommender system.

2.3. Constructing Feedback Rating Matrix

In this section, the process of constructing F by combining O and B is described in detail. Diverse approaches can be adopted for the process as it is possible for u_i to leave both ordinal and binary ratings to e_j . Moreover, there exist numerous criteria for determining the type of user feedback to employ preferentially. Therefore, we suggest four different combination methods, $c_o(o_{i,j}, b_{i,j})$, $c_b(o_{i,j}, b_{i,j})$, $c_l(o_{i,j}, b_{i,j})$, and $c_d(o_{i,j}, b_{i,j})$, for inferring $f_{i,j}$ to identify the optimal approach by which the user's positive and negative opinions are accurately expressed. Note that the four methods are differentiated by the type of user feedback they utilize when both $o_{i,j}$ and $b_{i,j}$ exist, and when only one type of user feedback exists, the existing rating is employed in all methods.

By $c_o(o_{i,j}, b_{i,j})$, $o_{i,j}$ is employed preferentially over $b_{i,j}$ when both $o_{i,j}$ and $b_{i,j}$ exist as described in Equation (2).

$$c_o(o_{i,j}, b_{i,j}) = \begin{cases} o_{i,j} & \text{if } \exists o_{i,j}, b_{i,j} \\ o_{i,j} & \text{if } \exists o_{i,j}, \nexists b_{i,j} \\ \delta_b(b_{i,j}) & \text{if } \nexists o_{i,j}, \exists b_{i,j} \\ \emptyset & \text{otherwise,} \end{cases} \quad (2)$$

where $\delta_b(b_{i,j})$ is a function that transforms binary ratings into 5-point scale as shown in Equation (3).

$$\delta_b(b_{i,j}) = \begin{cases} 5 & \text{if } b_{i,j} = 1 \\ 1 & \text{if } b_{i,j} = -1. \end{cases} \quad (3)$$

The idea behind $\delta_b(b_{i,j})$ is as follows. e_j will be recommended to u_i in future services if u_i clicked like button to e_j , while e_j will be excluded for u_i if u_i clicked dislike button to e_j . The former case indicates that u_i is extremely satisfied with the service provided by e_j , and, thus, $b_{i,j}$ is transformed into 5, which represents the most positive opinion. The latter case implies that u_i is absolutely unsatisfied with the service provided by e_j , and, thus, $b_{i,j}$ is transformed into 1, which is the most negative opinion.

By $c_b(o_{i,j}, b_{i,j})$, $b_{i,j}$ is employed preferentially over $o_{i,j}$ when both $o_{i,j}$ and $b_{i,j}$ exist as described in Equation (4).

$$c_b(o_{i,j}, b_{i,j}) = \begin{cases} \delta_b(b_{i,j}) & \text{if } \exists o_{i,j}, b_{i,j} \\ o_{i,j} & \text{if } \exists o_{i,j}, \nexists b_{i,j} \\ \delta_b(b_{i,j}) & \text{if } \nexists o_{i,j}, \exists b_{i,j} \\ \emptyset & \text{otherwise.} \end{cases} \quad (4)$$

Next, $c_d(o_{i,j}, b_{i,j})$ and $c_l(o_{i,j}, b_{i,j})$ are approaches that employ $b_{i,j}$ preferentially over $o_{i,j}$ when both $o_{i,j}$ and $b_{i,j}$ exist like $c_b(o_{i,j}, b_{i,j})$. However, they employ $b_{i,j}$ over $o_{i,j}$ only when $b_{i,j}$ is a certain value,

like or dislike. Specifically, by $c_d(o_{i,j}, b_{i,j})$, $b_{i,j}$ is selected over $o_{i,j}$ only when $b_{i,j}$ is -1 (dislike), and $o_{i,j}$ is employed otherwise as shown in Equation (5).

$$c_d(o_{i,j}, b_{i,j}) = \begin{cases} \delta_d(o_{i,j}, b_{i,j}) & \text{if } \exists o_{i,j}, b_{i,j} \\ o_{i,j} & \text{if } \exists o_{i,j}, \nexists b_{i,j} \\ \delta_b(b_{i,j}) & \text{if } \nexists o_{i,j}, \exists b_{i,j} \\ \emptyset & \text{otherwise,} \end{cases} \quad (5)$$

where $\delta_d(o_{i,j}, b_{i,j})$ determines the utilization of $b_{i,j}$ according to the value of $b_{i,j}$ as described by Equation (6).

$$\delta_d(o_{i,j}, b_{i,j}) = \begin{cases} o_{i,j} & \text{if } b_{i,j} = 1 \\ \delta_b(b_{i,j}) & \text{if } b_{i,j} = -1. \end{cases} \quad (6)$$

On the contrary, by $c_l(o_{i,j}, b_{i,j})$, $b_{i,j}$ is selected over $o_{i,j}$ only when $b_{i,j}$ is 1 (like), and $o_{i,j}$ is employed otherwise as shown in Equation (7).

$$c_l(o_{i,j}, b_{i,j}) = \begin{cases} \delta_l(o_{i,j}, b_{i,j}) & \text{if } \exists o_{i,j}, b_{i,j} \\ o_{i,j} & \text{if } \exists o_{i,j}, \nexists b_{i,j} \\ \delta_b(b_{i,j}) & \text{if } \nexists o_{i,j}, \exists b_{i,j} \\ \emptyset & \text{otherwise,} \end{cases} \quad (7)$$

where $\delta_l(\cdot, \cdot)$ is defined as Equation (8).

$$\delta_l(o_{i,j}, b_{i,j}) = \begin{cases} \delta_b(b_{i,j}) & \text{if } b_{i,j} = 1 \\ o_{i,j} & \text{if } b_{i,j} = -1. \end{cases} \quad (8)$$

We propose $c_d(o_{i,j}, b_{i,j})$ to integrate the findings of the user feedback log analysis in Section 1 that binary ratings reflect user's negative opinion more accurately than ordinal ratings. Thus, $c_d(o_{i,j}, b_{i,j})$ selects $b_{i,j}$ when $o_{i,j}$ could not precisely reveal the user's opinion, which is when $b_{i,j} = -1$ (dislike). In other words, $c_d(o_{i,j}, b_{i,j})$ chooses more effective user feedback between ordinal and binary ratings according to the value of binary ratings, like or dislike. $c_l(o_{i,j}, b_{i,j})$ is additionally implemented to validate the effectiveness of $c_d(o_{i,j}, b_{i,j})$.

2.4. Estimating Feedback Ratings by Performing Matrix Factorization

In this step, we try to infer the unknown feedback ratings of F since it is necessary to know user's opinion on every editing expert to find the optimal expert for the user. MF decomposes a feedback rating matrix with null values into a user matrix, which represents the latent factors of the users, and an editing expert matrix, which represents the latent factors of the experts. Then, a feedback rating matrix that is closest to the real one can be obtained through an optimization process by estimating the values in the user matrix and the editing expert matrix, and the unknown values in the original feedback rating matrix can be filled [25]. In other words, MF is adopted to build an estimated feedback rating matrix denoted by $\hat{F}$, whose element $\hat{f}_{i,j}$ is not null for all i, j . Specifically, $\hat{F}$ is approximated by multiply the two low-rank matrices obtained by factorizing F as shown in Equation (9).

$$\hat{F} \approx RQ^T, \quad (9)$$

where $R \in \mathbb{R}^{n_u \times k}$ and $Q \in \mathbb{R}^{n_e \times k}$ is latent user and expert feature matrices, whose row vectors r_i and q_j represents the k -dimensional latent feature vectors of u_i and e_j , respectively. The parameter k controls the rank of the factorization and indicates the dimension of latent space for representing

the characteristics of users and editing experts [7]. k must be a positive integer and smaller than the minimum of n_u and n_e [12].

MF maps both users and editing experts to a k -dimensional latent feature space, and $\hat{f}_{i,j}$ is modeled as the inner products of feature vectors, r_i and q_j , in the space as described in Equation (10).

$$\hat{f}_{i,j} = r_i \cdot q_j^T. \quad (10)$$

$\hat{F}$ is estimated by minimizing the objective function in Equation (11).

$$L = \sum_{i=1}^{n_u} \sum_{j=1}^{n_e} l_{i,j} (f_{i,j} - \hat{f}_{i,j}) + \lambda \left(\sum_{i=1}^{n_u} \|r_i\|^2 + \sum_{j=1}^{n_e} \|q_j\|^2 \right), \quad (11)$$

where $l_{i,j}$ is an indicator which is 1 if u_i has rated e_j , and 0 otherwise. $\lambda(\sum_{i=1}^{n_u} \|r_i\|^2 + \sum_{j=1}^{n_e} \|q_j\|^2)$ is a regularization term for preventing overfitting, and λ is a parameter for controlling the strength of the regularization [26]. Equation (11) can be solved by using the stochastic gradient descent, and details are presented in Reference [11].

2.5. Recommending Optimal Editing Expert

Lastly, the optimal editing expert for a user is determined based on $\hat{F}$. Specifically, when u_{i*} asks for manuscript editing, a matching manager explores E and recommends e_{j*} who fits u_{i*} best based on $\hat{F}$. The goal is to find the index of the optimal editing expert j^* for u_{i*} , and the optimality can be inferred by comparing the estimated feedback ratings of u_{i*} .

Thus, the editing expert recommendation process for u_{i*} is as follows. First, the i^* -th row of $\hat{F}$ is extracted. Then, the elements in the row, $\hat{f}_{i*,j}$ for all j is compared. Lastly, the index of the optimal editing expert j^* whose estimated feedback rating $\hat{f}_{i*,j^*}$ was the highest is selected by Equation (12).

$$j^* = \operatorname{argmax}_j \hat{f}_{i*,j}. \quad (12)$$

Thus, e_{j^*} is recommended to u_{i*} as the optimal editing expert.

3. Experiment

3.1. Dataset

For the experiments, we utilized user feedback logs collected from a manuscript editing company called ‘Essayreview’ [2] to demonstrate the effectiveness of the proposed MF based editing expert recommendation method. There are three types of user feedback in the logs, ordinal ratings, binary ratings, and review texts. Ordinal ratings were drawn from a 5-point scale where the ratings are integer values ranging from 1 to 5, and binary ratings were selected from two options—like and dislike. Ordinal ratings and binary ratings were utilized to evaluate the recommendation performances of the proposed method, and review texts were used to validate the ability of the proposed method in refining the user’s opinion by conducting sentiment analysis.

Table 2 shows the summary of the collected user feedback logs. The number of ordinal ratings, binary ratings, and review texts were respectively 1326, 202, and 179. As there are logs that do not contain all three types of user feedback, we present the combination frequencies, which are the number of logs having diverse combinations of user feedback. The number of user feedback logs composed of both ordinal and binary ratings was 180 while that composed of ordinal or binary ratings was 1348. In addition, the number of logs composed of both binary rating and review text was 179, and that composed of either binary ratings or review texts was 202. There were a total of 179 logs, in which all three types exist, and 1348 logs contained at least one type.

Table 2. Summary of user feedback logs utilized in the experiments.

Type	Category	Frequency	Combination Frequency (Intersection, Union)	
Ordinal rating	1	32	(180, 1348)	-
	2	46		
	3	77		
	4	293		
	5	878		
	Total	1326		(179, 1348)
Binary rating	Like	164	(0, 202)	
	Dislike	38		
	Total	202		
Review text		179	-	

Note that only active users and editing experts were used in the experiments. We define an active user as a person who used the manuscript editing service and left feedback at least once and an active editing expert as a person who has received more than one feedback. The number of active users and experts were 854 and 94, respectively.

3.2. Settings

Four types of experiments were conducted to observe the performances and characteristics of the proposed method. First, the collected user feedback logs were explored to identify the data sparsity problem. Second, the impact of the proposed method's parameters on the recommendation performances were investigated to determine the optimal values of the parameters. Third, we evaluated the performances of the proposed method and compared them with the state-of-the-art method. Lastly, recommendation results were closely observed to validate the effectiveness of the proposed method in refining user opinions. Specifically, we performed sentiment analysis, the task of identifying user's opinion inherent in a text [27] to compare the results with those of the proposed method.

The collected user feedback logs were partitioned into training, validation, and test set, where the training set and validation set were respectively utilized for building a feedback rating matrix and for selecting the optimal parameters of the proposed method and the test set was used for the performance evaluation of the proposed method using the feedback rating matrix. Among two hundred of the most recent logs, the oldest one hundred of the logs were selected as the validation set and the most recent one hundred of the logs were selected as the test set. Note that the feedback ratings of the test set generated by performing the proposed method were indicated as 'estimated feedback ratings' while those before performing as 'original feedback ratings'. All experiments were repeated thirty times, and the results were averaged to minimize the randomness.

As evaluation measures, we adopted root mean squared error (RMSE) and mean absolute percentage error (MAPE), which are two of the most widely utilized measures for rating estimation [28,29]. RMSE is defined as the square root of the average squared difference between the actual and the estimated feedback ratings as shown in Equation (13).

$$RMSE = \sqrt{\frac{\sum_{(i,j) \in T} (f_{i,j} - \hat{f}_{i,j})^2}{|T|}}, \quad (13)$$

where T is a set of index pairs (i, j) in the test set and $f_{i,j}$ and $\hat{f}_{i,j}$ respectively indicate the original and the estimated feedback ratings from u_i to e_j . In addition, MAPE is defined as the average percentage

of the absolute difference between the original and the estimated feedback ratings divided by the absolute value of the original feedback ratings, and it is calculated by Equation (14).

$$MAPE = \frac{100}{|T|} \sum_{(i,j) \in T} \frac{|f_{i,j} - \hat{f}_{i,j}|}{|f_{i,j}|}. \quad (14)$$

The smaller values of RMSE and MAPE indicate better recommendation performances.

We compared the recommendation performances of the proposed method with TMF [22], which exploits ordinal and binary ratings simultaneously. Specifically, TMF estimates the unknown feedback ratings by utilizing additional user latent matrix generated by analyzing binary ratings. When calculating the additional matrix, two parameters— w_l and w_d —were used to control the weights of like and dislike information in binary ratings. The comparison between the proposed method and TMF can provide an opportunity to demonstrate the effectiveness of substituting ordinal ratings with binary ratings compared to the simultaneous utilization of the two ratings.

For the performance comparison, nine methods, which are differentiated in the way of utilizing user feedback, were implemented. The methods are grouped into two categories, ones that utilize only ordinal or binary ratings, denoted by O or B , and ones that exploit both ratings including the aforementioned TMF and the proposed method. Specifically, there are three types of TMF denoted by TMF(1,1), TMF(1,0), and TMF(0,1), which respectively indicate TMF with a parameter pair (w_l, w_d) of (1,1), (1,0), and (0,1). In addition, four types of the proposed method, $F(c_o)$, $F(c_b)$, $F(c_d)$, and $F(c_l)$, which utilize feedback rating matrices constructed by using the four functions, $c_o(o_{i,j}, b_{i,j})$, $c_b(o_{i,j}, b_{i,j})$, $c_l(o_{i,j}, b_{i,j})$, and $c_d(o_{i,j}, b_{i,j})$, respectively, were implemented. The numbers of non-null elements in the feedback rating matrices utilized for O and B were 1326 and 202, respectively, and those of TMFs and Fs were the same—1348. Note that we only adopted the basic approach of MF and tried to focus on the comparison of feedback rating incorporation approach.

For the optimal parameter selection, we considered two parameters— λ and k —in the proposed method. λ is a regularization parameter which helps prevent overfitting as shown in Equation (11) and k is the dimension of the latent space for performing MF. We investigated the effects of λ and k on the performances of the proposed method. Experiments were conducted on the validation set with diverse λ s, 0.2, 0.02, and 0.002, and diverse k s ranging from 10 to 90 at intervals of 10. According to the results, we determined the optimal values of parameters for the rest of the experiments.

For sentiment analysis, we employed a lexicon-based approach [30], where the sentiment score of a review text is determined as the average of the sentiment scores of the words composing the review text. The sentiment score of a word is assigned according to a lexicon of words, where words are annotated with sentiment scores between -2 (negative) to 2 (positive). Thus, when the sentiment score of a review text is close to 2 , it means the review is positive, while when it is close to -2 , it indicates the review is negative. The publicly-available deep learning based morphological analyzer for Korean called Khaiii [31] was utilized to extract the morpheme tags for the words in review texts. Among 23 morphemes, only the words corresponding to the four morphemes that have lexical meanings such as noun, adjective, positive copulas, and negative copulas were used. Additionally, the publicly-available sentiment lexicon for Korean named KOSAC (Korean sentiment analysis corpus) [32] was employed to assign the sentiment scores of the words in review texts.

3.3. Experiment Results

In this section, the experimental results of the proposed method are presented. In detail, we report results on data exploration, parameter selection, performance evaluation, and performance validation.

3.3.1. Data Exploration

We explored the user feedback logs utilized in the experiments in order to demonstrate the sparseness of ordinal ratings. In Figure 5a,b, the frequencies of users and editing experts are illustrated

as bar charts according to the number of ordinal ratings that the users gave and the experts received, respectively.

Figure 5. Visualization of the ordinal ratings utilized in the experiments.

From the exploration, it can be concluded that the ordinal ratings are sparse and concentrated to a small number of users and editing experts. Specifically, about 70% of the users left ordinal ratings for one time, and 21% of the users gave ordinal rating for two times as shown in Figure 5a. In conclusion, about 90% of the users left ordinal ratings on the received editing services to a maximum of two times. Moreover, only 5% of the editing experts received more than 50 ordinal ratings, and up to 70% of the editing experts received ordinal ratings less than 10 times as shown in Figure 5b.

3.3.2. Parameter Selection

The impact of parameters k and λ of the proposed method on the recommendation performances was investigated to determine the optimal values. Figure 6 shows the recommendation performances of the proposed methods according to the iteration number (x -axis) in terms of RMSE (y -axis) for diverse k s ranging from 10 to 90 at intervals of 10. Each graph in Figure 6 consists of three lines that plot the performances of the proposed method with diverse λ s, 0.2, 0.02, and 0.002.

Through the comparison of the nine graphs in Figure 6, it was observed that the RMSE tends to decrease as k increased. The RMSE decreased rapidly until k reached 30, and it became steady as k became larger than 80. In terms of λ , overfitting was observed when λ was 0.02 and 0.002. The performance of the method stopped improving and started to degrade at a certain iteration number. Specifically, overfitting clearly occurred for $\lambda = 0.02$ and $k \geq 30$ and for $\lambda = 0.002$ and $k \geq 80$ while when $\lambda = 0.2$, no overfitting was evident for every k . Thus, we set k and λ to 80 and 0.2, respectively, for the rest of the experiments.

3.3.3. Performance Evaluation

The recommendation performances of the proposed method were evaluated qualitatively and quantitatively. For the qualitative evaluation, we visualized the original and estimated feedback rating matrices and compared them to show the resemblance. For the quantitative evaluation, the RMSE and MAPE of the proposed and compared methods were evaluated to show the superior performances of the proposed method and the effectiveness of the selective utilization of ordinal and binary ratings.

Feedback rating matrices were illustrated as heatmaps in Figure 7. Figures 7a–c, respectively represent an original feedback rating matrix, a matrix whose elements belong to the test set were changed to null, and an estimated matrix by using the proposed method. In each graph, x and y -axis respectively indicate the users and editing experts in the test set, and a box represents the feedback rating constructed by using the user feedback log given by the corresponding user to the corresponding expert. A box filled with white color indicates null, meaning that the corresponding user did not rate the corresponding expert, and the darker the color of a box is, the greater the corresponding ordinal

rating is. Overall, the heatmap in Figure 7c is similar to that in Figure 7a, implying that the proposed method was successful at inferring user opinions.

Figure 6. Recommendation performances of the proposed method with diverse λ s, 0.2, 0.02, and 0.002, and k s ranging from 10 to 90 at intervals of 10 in terms of RMSE according to the iteration number.

Table 3 and 4 respectively show the recommendation performances of the considered nine methods in terms of RMSE and MAPE and the paired sample t-test results comparing the performances of the methods in terms of p -value. Among the methods using a single type of user feedback, O outperformed B . The result was foreseen since the number of non-null elements in the feedback rating matrix of O was greater than that of B and ordinal ratings usually contain abundant information compared to binary ratings.

Between O , B and the methods using both ratings, TMFs and F s, the latter performed much better than the former with the minimum of 34% reduction in RMSE. Particularly, RMSE and MAPE of $F(c_o)$, which utilizes both ratings and prefers ordinal ratings, were 0.17% and 4.23% while those of O , which only uses ordinal ratings, were 0.26% and 6.49%, where p -value between the performances of $F(c_o)$ and O was 0.0052. Since the p -value was smaller than the significance level of 0.05, it is possible to say that the performance gains are statistically significant. It demonstrates that the utilization of binary ratings in addition to ordinal ratings has a positive effect on the recommendation performances. Moreover, $F(c_b)$ performed better than $F(c_o)$, implying that substituting ordinal ratings with binary ratings was effective as it reduces the effect of bias inherent in ordinal ratings, and this result was considered statistically rigorous with the p -value of 0.0044.

Figure 7. Heatmap representation of the feedback rating matrices, where x and y -axis respectively represent the users and editing experts belong to the test set.

Table 3. Recommendation performances of the proposed method and compared methods with respect to the type of the utilized user feedback in terms of RMSE and MAPE.

Type	Method	Performance	
		RMSE	MAPE (%)
Ordinal rating only	O	0.2573 ± 0.0006	6.4974 ± 0.0067
Binary rating only	B	0.4263 ± 0.0027	10.7651 ± 0.0039
Ordinal and binary rating	Mixed	TMF(1,1)	0.1381 ± 0.0009
		TMF(1,0)	0.1744 ± 0.0014
		TMF(0,1)	0.1317 ± 0.0008
	Substituted	$F(c_o)$	0.1677 ± 0.0008
		$F(c_b)$	0.1201 ± 0.0005
		$F(c_l)$	0.1149 ± 0.0006
		$F(c_d)$	0.0999 ± 0.0012

It was interesting that the methods emphasizing the negative user opinions (dislike) showed better performances than those emphasizing positive ones (like) or both. $F(c_d)$ which substituted

ordinal ratings with binary ratings when a user disliked an editing expert outperformed the rest of the methods with an RMSE of 0.0999, and $F(c_o)$, which used ordinal ratings preferentially, showed the worst performance among F s with an RMSE of 0.1677. In addition, $F(c_d)$ showed better performances than $F(c_l)$ with an RMSE of 0.1149, where the p -value of 0.0023 smaller than 0.05 indicates that there is significant performance improvement for $F(c_d)$. A similar phenomenon was observed in the performance results of TMFs as well. The performances of TMF(0,1), where dislike information was adopted, were superior to TMF(1,0), where like information was adopted, and the performances of TMF(1,1), where both like and dislike information was adopted, were better than those of TMF(1,0). They were statistically significant since the p -value between TMF(0,1) and TMF(1,0) was 0.0039, and between TMF(1,1) and TMF(1,0) was 0.0035.

Moreover, it was noticeable that when both ordinal and binary ratings were given, substituting ordinal ratings with binary ratings enhances the performances more than incorporating the two ratings together. This can be supported by the fact that the performances of $F(c_d)$ were superior to TMF(0,1), and it is statistically approved with a p -value of 0.0012. Both methods utilized ordinal and binary ratings, but $F(c_d)$ only utilized binary ratings when the binary rating was dislike while TMF(0,1) incorporated the dislike information of binary ratings with ordinal ratings. The conclusion that utilizing ordinal ratings degrades the opinion inference performance especially for negative ones conforms with the problem of ordinal ratings raised in Section 1, that negative opinions are not correctly expressed in ordinal ratings.

Further analyses were conducted to show the trends in the recommendation performances with diverse k s as illustrated in Figure 8a,b. The performances of all nine methods enhanced as k increased, and the performance ranks remained the same except for $F(c_b)$ and $F(c_d)$. $F(c_b)$ showed the best performances when k was smaller than 40, but $F(c_d)$ overtook $F(c_b)$ with k larger than 50.

Figure 8. Performances of the nine compared methods in terms of RMSE and MAPE according to various k s.

Table 4. Results of paired sample t-tests on the comparison performances for main conclusions.

Main Conclusion	Comparison	p -Value
Using both binary and ordinal ratings is more effective for inferring user opinion than only using ordinal ratings.	$F(c_o) > O$	0.0052
Binary ratings provide user opinion more accurately than ordinal ratings.	$F(c_b) > F(c_o)$	0.0044
User's negative feedback is more effective for inferring user opinion than positive one.	$F(c_d) > F(c_l)$	0.0023
	$TMF(0,1) > TMF(1,0)$	0.0039
	$TMF(1,1) > TMF(1,0)$	0.0035
Substitution is more effective than simultaneous utilization.	$F(c_d) > TMF(0,1)$	0.0012

3.3.4. Performance Validation

To validate the effectiveness of the proposed method in refining a user's opinion, the recommendation results of the test set were investigated in detail. First, we observed the changes in frequencies of like and dislike between the original and the estimated feedback ratings. Then, we conducted sentiment analysis on the review text of the test set and examined the sentiment scores of review texts according to the original and estimated feedback ratings.

Figure 9a,b show the frequency distributions of user feedback logs of two binary rating categories—like and dislike—in the test set according to the original and estimated feedback ratings, respectively. It was observed that the frequencies of dislike (filled box) whose ratings were 1 and 2 increased in Figure 9b compared to the original feedback ratings in Figure 9a, while the portion of dislike (filled box) whose ratings were 4 and 5 decreased in Figure 9b compared to the original feedback ratings in Figure 9a. For more precise analysis, the percentages of user feedback logs with respect to the ordinal (feedback) ratings and the binary ratings were shown in Table 5. The percentage, where the ordinal ratings were 1 or 2 and the binary ratings were dislike, increased from 39% (original ordinal ratings) to 76% (estimated feedback ratings), while the percentage, where the ordinal ratings were 4 or 5 and the binary ratings were dislike, decreased from 37% to 13%. This implies that the proposed method was successful at refining user opinions, especially for negative ones, by substituting ordinal ratings with binary ratings when the binary rating was dislike for building a feedback rating matrix.

Figure 9. Frequency distributions of user feedback logs with respect to the two categories of binary ratings, like (empty) and dislike (filled), according to the (a) original ordinal ratings and (b) estimated feedback ratings.

Table 5. Percentage of user feedback logs of two binary rating categories, like and dislike, when the original ordinal ratings and estimated feedback ratings were 1 or 2 and 4 or 5.

		Binary Rating			
		Like		Dislike	
		Original	Estimated	Original	Estimated
Ordinal (feedback) rating	1, 2	7.69%	10.65%	39.53%	76.00%
	4, 5	80.77%	83.56%	37.21%	12.50%

We assumed that if the proposed method performed successfully, the estimated feedback ratings would be positively correlated with the sentiment of user opinions contained in review texts. In other words, the sentiment score should be high (or low) when the feedback rating is high (or low). Thus, the sentiment analysis mentioned in Section 3.2 was conducted on the review text of the user feedback logs

in the test set, and compared the results between original and estimated feedback ratings. Figure 10 shows the surface plots illustrating the frequency distributions of user feedback logs according to the sentiment score (x -axis) and the original and estimated feedback ratings (y -axis).

Figure 10. Surface plots illustrating the frequencies of user feedback logs according to the sentiment scores of review texts (x -axis) and the (a) original ordinal ratings or (b) estimated feedback ratings (y -axis).

By the assumption, a surface plot having an uplifted surface near the diagonal from $(-2, 1)$ to $(2, 5)$ and caved-in near the vertices, $(2, 1)$ and $(-1, 5)$, represents the optimal estimation of feedback ratings. In Figure 10a, it can be observed that the rear surface of the diagonal protruded and the front of the diagonal caved in, and Pearson correlation coefficient between the original ordinal ratings and the sentiment score was 0.46. On the other hand, there was a noticeable uplift near the diagonal in the case of the estimated feedback ratings as shown in Figure 10b, and the correlation coefficient was 0.67. The increased correlation indicates that the proposed method appropriately refined the user's opinion by using binary ratings.

4. Guideline for Applications

4.1. Requirements

In order to implement the proposed method for real-world editing expert services, the proposed method should be divided into two processes, online and offline, as it is computationally demanding and time-consuming to execute every time the editing service is requested. The estimation process of $\hat{F}$ can be done in advance as an offline process, while the recommendation should be carried out in real-time as an online process. The process is described in Algorithm 1.

Algorithm 1 Editing expert recommendation process

INPUT: User feedback logs (users, experts, user feedback), a user u_{i*} .

OUTPUT: The optimal editing expert e_{j*} to be recommended with u_{i*} .

Training phase (Offline)

(Step 1: Constructing a feedback rating matrix F)

(Step 2: Performing MF on F to estimate $\hat{F}$)

Test phase (Online)

(Step 3: Recommending the optimal editing expert based on $\hat{F}$)

An expert is selected whose feedback rating is highest in the i -th row of $\hat{F}$.

For the real-world implementation, it is important to find the maximum possible number of unknown values in a feedback rating matrix that does not degrade performance as a previous study reported that the performance of a method is not fully dependent on the amount of data but on the

ratio of unknown values over known values [33]. We denoted the ratio of the unknown values over the known values by ρ as Equation (15).

$$\rho = N_u / N_k, \quad (15)$$

where N_u and N_k respectively indicates the number of unknown and known values in a feedback rating matrix. Therefore, experiments were conducted using diverse ρ s to find the maximum possible number.

Figure 11 shows the results in terms of RMSE (y -axis) according to ρ s (x -axis). We can observe that RMSE was steady when ρ was smaller than 4. However, RMSE increased rapidly as ρ got bigger. Therefore, it is recommended for companies utilizing the proposed method to maintain ρ under 4.

Lastly, we evaluated the time duration for performing expert recommendation by using the proposed method. We utilized a server with 3.5 GHz Intel Core i7 and repeatedly ran the online process of the proposed method for 20 times under the same environment. As a result, the average service time for online process was 0.78 seconds which was short enough for actual implementation.

Figure 11. Experiment result according to various ρ s in terms of RMSE.

5. Conclusions

In this paper, we proposed an MF based editing expert recommendation method which utilizes ordinal and binary ratings given by users to editing experts. MF is adopted to explore the inherent characteristics of a manuscript and latent information in user opinions to address the drawbacks of the current manual matching process in manuscript editing companies. Specifically, binary ratings were utilized in addition to ordinal ratings to alleviate the sparsity in ordinal ratings and refine user's negative opinions.

Experiments on a real-world dataset collected from a manuscript editing service were conducted to evaluate the recommendation performances of the proposed method and validate its capabilities. Two conclusions can be drawn from the results. First, the recommender system utilizing both ordinal and binary ratings outperformed the method utilizing only ordinal ratings, which implies that the use of binary ratings can successfully address the data sparsity problem. Second, in terms of constructing a feedback rating matrix, a method, where binary ratings are employed to substitute ordinal ratings when a user left a negative opinion (dislike), outperformed the rest. This implies that the negative opinions of users are more accurately expressed by binary ratings than ordinal ratings.

Our future research directions are as follows. Experiments using diverse datasets will be conducted to improve the robustness of performance evaluation results as the experiments utilized review texts written in Korean and collected from a specific service. The proposed framework can be applied to other languages with minor adaptations such as changing POS tagger and sentiment dictionary. Moreover, we plan to investigate the adaptation of the selective use of binary ratings over ordinal ratings to various methods of MF and to validate the effectiveness using more general datasets. Next, we plan to extend the proposed method to utilize additional types of user feedback, such as review texts, to enrich users' opinions. We hope our finding on the effectiveness of the selective use of

binary ratings over the ordinal rating when a user showed a negative opinion can be the beginning of more in-depth research.

Author Contributions: Conceptualization, Y.S. and Y.C.; data curation, Y.S. and Y.C.; formal analysis, Y.C.; funding acquisition, Y.C.; methodology, Y.S. and Y.C.; software, Y.S.; supervision, Y.C.; validation, Y.S.; visualization, Y.S.; writing—original draft preparation, Y. S. and Y.C.; writing—review and editing, Y.S. and Y.C. All authors have read and agree to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: This work was supported by Kyonggi University Research Grant 2017.

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

References

- Shiyab, S.; Lynch, M.S. Can Literary Style Be Translated? *Babel* **2006**, *52*, 262–275. [CrossRef]
- Professional Proofreading & Editing. Available online: <http://essayreview.co.kr> (accessed on 10 February 2020).
- Academic English Editing and Manuscript Proofreading. Available online: <http://www.enago.co.kr> (accessed on 10 February 2020).
- Professional English and Academic Editing Service by Experts. Available online: <http://www.editage.co.kr> (accessed on 10 February 2020).
- Resnick, P.; Varian, H.R. Recommender Systems. *Commun. ACM* **1997**, *40*, 56–58. [CrossRef]
- Singhal, A. Modern Information Retrieval: A Brief Overview. *IEEE Data Eng. Bull.* **2001**, *24*, 1–9.
- Nathan, S.; Jason, R.; Tommi, J. Maximum-Margin Matrix Factorization. In *Proceedings of the 17th International Conference on Neural Information Processing Systems*; MIT Press: Vancouver, BC, Canada, 2005; Volume 17, pp. 1329–1336.
- Bobadilla, J.; Alonso, S.; Hernando, A. Deep Learning Architecture for Collaborative Filtering Recommender Systems. *Appl. Sci.* **2020**, *10*, 2441. [CrossRef]
- Zheng, E.; Kondo, G.Y.; Zilora, S.; Yu, Q. Tag-Aware Dynamic Music Recommendation. *Expert Syst. Appl.* **2018**, *106*, 244–251. [CrossRef]
- Yin, Y.; Chen, L.; Xu, Y.; Wan, J. Location-Aware Service Recommendation with Enhanced Probabilistic Matrix Factorization. *IEEE Access* **2018**, *6*, 62815–62825. [CrossRef]
- Koren, Y. Collaborative Filtering with Temporal Dynamics. In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*; ACM Press: Paris, France, 2009; pp. 447–456.
- Salakhutdinov, R.; Mnih, A. Probabilistic Matrix Factorization. In *Proceedings of the 17th International Conference on Neural Information Processing Systems*; MIT Press: Red Hook, NY, USA, 2007; pp. 1257–1264.
- Yuan.; Zahir.; Yang. Modeling Implicit Trust in Matrix Factorization-Based Collaborative Filtering. *Appl. Sci.* **2019**, *9*, 4378. [CrossRef]
- Pirasteh, P.; Hwang, D.; Jung, J.J. Exploiting Matrix Factorization to Asymmetric User Similarities in Recommendation Systems. *Knowl.-Based Syst.* **2015**, *83*, 51–57. [CrossRef]
- Ma, H.; King, I.; Lyu, M.R. Effective Missing Data Prediction for Collaborative Filtering. In *Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*; ACM Press: Amsterdam, The Netherlands, 2007; pp. 39–46.
- Jiang, C.; Duan, R.; Jain, H.K.; Liu, S.; Liang, K. Hybrid Collaborative Filtering for High-Involvement Products: A Solution to Opinion Sparsity and Dynamics. *Decis. Support Syst.* **2015**, *79*, 195–208. [CrossRef]
- Chu, P.M.; Mao, Y.S.; Lee, S.J.; Hou, C.L. Leveraging User Comments for Recommendation in E-Commerce. *Appl. Sci.* **2020**, *10*, 2540. [CrossRef]
- Lian, D.; Zheng, K.; Ge, Y.; Cao, L.; Chen, E.; Xie, X. GeoMF++: Scalable Location Recommendation via Joint Geographical Modeling and Matrix Factorization. *Acm Trans. Inf. Syst.* **2018**, *36*, 1–29. [CrossRef]
- Hu, K.; Gui, Z.; Cheng, X.; Qi, K.; Zheng, J.; You, L.; Wu, H. Content-Based Discovery for Web Map Service Using Support Vector Machine and User Relevance Feedback. *PLoS ONE* **2016**, *11*. [CrossRef] [PubMed]
- Jin, X.; Wang, C.; Luo, J.; Yu, X.; Han, J. LikeMiner: A System for Mining the Power of ‘Like’ in Social Media Networks. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*; ACM Press: San Diego, CA, USA, 2011; pp. 753–756.

21. Yang, S.H.; Long, B.; Smola, A.; Sadagopan, N.; Zheng, Z.; Zha, H. Like Like Alike: Joint Friendship and Interest Propagation in Social Networks. In *Proceedings of the 20th International Conference on World Wide Web*; ACM Press: Hyderabad, India, 2011; pp. 537–546.
22. Pan, W.; Xia, S.; Liu, Z.; Peng, X.; Ming, Z. Mixed Factorization for Collaborative Recommendation with Heterogeneous Explicit Feedbacks. *Inf. Sci.* **2016**, *332*, 84–93. [[CrossRef](#)]
23. Pan, W.; Liu, N.N.; Xiang, E.W.; Yang, Q. Transfer Learning to Predict Missing Ratings via Heterogeneous User Feedbacks. In *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence*; AAAI Press: Barcelona, Catalonia, Spain, 2011; pp. 2318–2323.
24. Pan, W.; Yang, Q. Transfer learning in Heterogeneous Collaborative Filtering Domains. *Artif. Intell.* **2013**, *197*, 39–55. [[CrossRef](#)]
25. Koren, Y.; Bell, R.; Volinsky, C. Matrix Factorization Techniques for Recommender Systems. *Computer* **2009**, *42*, 30–37. [[CrossRef](#)]
26. He, X.; Zhang, H.; Kan, M.Y.; Chua, T.S. Fast Matrix Factorization for Online Recommendation with Implicit Feedback. In *Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval*; ACM Press: Pisa, Italy, 2016; pp. 549–558.
27. Feldman, R. Techniques and Applications for Sentiment Analysis. *Commun. ACM* **2013**, *56*, 82–89. [[CrossRef](#)]
28. Chen, H.; Fu, J.; Zhang, L.; Wang, S.; Lin, K.; Shi, L.; Wang, L. Deformable Convolutional Matrix Factorization for Document Context-Aware Recommendation in Social Networks. *IEEE Access* **2019**, *7*, 66347–66357. [[CrossRef](#)]
29. Xin, L.; Mengchu, Z.; Yunni, X.; Qingsheng, Z. An Efficient Non-Negative Matrix-Factorization-Based Approach to Collaborative Filtering for Recommender Systems. *IEEE Trans. Ind. Informatics* **2014**, *10*, 1273–1284. [[CrossRef](#)]
30. Taboada, M.; Brooke, J.; Tofiloski, M.; Voll, K.; Stede, M. Lexicon-Based Methods for Sentiment Analysis. *Comput. Linguist.* **2011**, *37*, 267–307. [[CrossRef](#)]
31. Khaiiii. Available online: <https://github.com/kakao/khaiiii> (accessed on 10 February 2020).
32. Kim, M.; Jang, H.Y.; Jo, Y.M.; Shin, H. KOSAC: Korean Sentiment Analysis Corpus. *Inf. Comput.* **2013**, *40*, 650–652.
33. Serruya, M.; Fellows, M.; Paninski, L.; Donoghue, J.; Hatsopoulos, N. Robustness of Neuroprosthetic Decoding Algorithms. *Biol. Cybern.* **2003**, *88*, 219–228. [[CrossRef](#)] [[PubMed](#)]

© 2020 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 (<http://creativecommons.org/licenses/by/4.0/>).