Deep Hotel Recommender System Using Aspect-based Sentiment Analysis of Users' Reviews

Alper Ozcan
Akdeniz University
Computer Engineering Department
Antalya, Turkey
Email: alperozcan@akdeniz.edu.tr

Bilgehan Emiral
Istanbul Technical University
Computer Engineering Department
Istanbul, Turkey
Email: emiral16@itu.edu.tr

Ayse Betul Cetin
Istanbul Technical University
Computer Engineering Department
Istanbul, Turkey
Email: cetina16@itu.edu.tr

***Abstract*—The value of search engines has increased with the digitalization of industries such as e-commerce and entertainment. There are numerous areas where recommendations are used and many different methods and technologies have been used to develop recommender systems. Deep learning-based methods are most effective as they improve the system's overall performance by obtaining state-of-the-art predictive accuracy. However, collaborative filtering is the most famous technique used in the recommneder systems that rely on the user-item interaction. Besides, Aspect-Based Sentiment Analysis (ABSA) is a text analysis method that divides opinions into aspects and determines the sentiment associated with each. Unlike previous works, in this paper, we combine aspect-based sentiment analysis with multi-criteria collaborative filtering using deep learning. In this study, we use polarities obtained by aspect-based sentiment analysis from user reviews combined with the overall rating and sub-ratings in a deep neural network model to predict user-item interactions. The model is composed of three parts: The first part includes Aspect-based Sentiment Analysis (ABSA) which is applied to the reviews, and polarities are extracted, then the final dataset is generated. In the second part, the user and item embeddings are concatenated and given as input to the Polarities and Sub-ratings Deep Neural Network (DNN) to predict the sub-ratings and polarities. Finally, in the third part, we give obtained criteria array as input to the Overall Score DNN to predict the overall score. Based on polarities the proposed approach leverages aspect-based user's opinion. The obtained results indicate that proposed approach performs significantly better on the TripAdvisor dataset.**

I. Introduction

The importance of search engines has increased tremendously with the digitalization of countless industries such as e-commerce and entertainment. Consequently, there are numerous areas where the recommendation is widely used. The variety of data that can feed recommendation systems is constantly increasing. In addition to product descriptions and information, user profile features, ratings, and reviews also contains important and useful data that can guide a recommendation system. Such data is used to identify similar users or similar products and recommend products based on these similarities. Recommendation systems are categorized as Collaborative, Content-based, Knowledge-based and hybrid in the literature. In the Content-based recommendation systems, items recommended by the system are determined according to similarity with the ones that the user liked in the past [1]. On the other hand, Collaborative filtering depends on similarities between users and makes predictions according to those similarities. Collaborative filtering approaches do not require any prior knowledge of users or items to produce recommendations since they rely on interactions between them [2]. Besides, the knowledge-based method pursues a knowledge-based approach to generating a suggestion, reasoning about which items satisfy the user's criteria using knowledge about users and products [3]. Combining these approaches is also possible and is known as hybrid systems. Deep learning has been reviving the recommendation architectures significantly and it is bringing more opportunities to improve the performance of recommendation models. The main factor of this success comes from capturing the non-linear and non-trivial user-item connections [4].

In a recommendation system, the use of features deduced from reviews and ratings strengthens the system. Such features are crucial to predict the rating of a user for a particular item. For hotel recommendation, the user's travel purpose is substantial, such as business, touristic and vacation, food preferences, room preferences, what he/she expects from a hotel, and whom he/she goes with (family, couple, single). Generally, multi criteria ratings are used to represent more complex representations, e.g., average rating of a specific item by users. Sub-ratings are ratings given to the item by the user for a certain category. In this study, the high null rate in the sub-ratings given by the users was a source of motivation for us to produce features from their reviews. Detailed data about the user's approach to the hotel is obtained from the reviews. For this reason, aspect-polarity pairs extracted from the reviews is used in the proposed model in order to enrich the collaborative filtering fed by the simple traditional rating. By aspects, we consider attributes or components of an entity (a food or staff, in this work). The polarity values represent the user's positive or negative interaction for that aspect.

In this paper, we propose a novel solution for improving collaborative filtering model performance that combines aspect sentiment analysis and multi-criteria recommendation with deep learning. We use a deep neural network to discover the relationship between overall rating and sub-ratings with aspect polarities.

The main contributions of this article are summarized as follows:

- A novel deep multi-criteria collaborative filtering approach using aspect-based sentiment analysis of users' reviews was developed for hotel recommendation.
- The proposed model has demonstrated better results as the novel architecture is implemented that integrates aspect-based sentiment analysis and multi-criteria ratings.

The rest of this paper is organized as follows: In Section II, we analyze the related work. We introduce our ABSA method, recommendation model, and settings in detail in Section III. Section IV presents the dataset and experimental evaluations. Finally, in Section V, we provide a conclusion and introduce our plans for future work.

II. Related Work

Many research papers have been published so far on recommendation systems. This section discusses relevant studies for multi-criteria recommendation and aspect-based sentiment analysis.

A. Recommendation Model

Neighborhood-based techniques employ user-item assessments to calculate the parallels among user-based [5] and item-based [6] to provide recommendations directly. In model-based techniques, user-item assessments are utilized to develop a prediction model, which is subsequently used to provide recommendations. Model-based techniques include Support Vector Machines [7], Bayesian Clustering [8], Latent Semantic Analysis [9] and Singular Value Decomposition [10]. Studies [11]–[13] consider the user's preference for certain product features to estimate more specific recommendations.

B. Multi-criteria Recommendation

Memory-based and model-based strategies are two main types of multi-criteria recommendation techniques based on their utility function. The similarity can be computed in advance methods and conservative methods using memory-based techniques. The advanced techniques utilize multidimensional distance metrics directly to calculate the distance between multi-criteria scores [14]. Conservative techniques utilize average [14], worst-case [14], and weighted sum of individual similarities [15] to combine traditional similarities values determined separately on each criterion into a single similarity.

The model based approach [14], predicts the unknown ratings with the assumption of item's overall rating is independent of other ratings. It generates a direct relationship between the overall rating and multi-criteria rating. Moreover, approaches like Probabilistic Support Vector Regression [16], Modeling [17], Multi-linear Singular Value Decomposition [18], and Genetic Algorithm [19] uses multidimensional distance metrics to find direct distance between multi-criteria scores [14].

Even though only few studies have been proposed on integrating deep learning into a collaborative filtering model. Deep learning for recommender systems has recently received much notice. Some researchers have developed deep learning-based recommendation models, that have shown capability to capture complex relationships [20], resulting in improved recommender system performance.

While the early literature on recommendation largely focused on raw explicit feedback [21] [22], recent attention is increasingly shifting to data [23] [24]. For instance, extracted data from user reviews, which is processed using aspect-based sentiment analysis techniques and produce polarity of users on specific aspects such as presented in this paper.

C. Aspect-based Sentiment Analysis

Akhtar et al. [25] analyze the hotel reviews to extract common aspects. The study categorizes the reviews into predetermined categories, which are the aspects that frequently recur in the reviews dataset. Our model is based on such analysis of aspect classification. Some publications explore Aspect-Based Sentiment Analysis for user reviews and using ABSA in recommendation systems [26]–[28]. Musto et al. [29] describe a method for justifying the recommendations given by a recommendation algorithm by automatically extracting relevant and distinguishable properties of the recommended item from user reviews. It claims that the users prefer review-based arguments over other explanation strategies. Bauman et al. [30] propose a method called the Sentiment Utility Logistic Model (SULM) that uses sentiment analysis of user reviews. This study demonstrates some experiments in which users who selected the items according to their recommendations for the most valuable aspects had better experiences, as measured by the rating. Liu et al. [31] propose a novel Multilingual Review-aware Deep Recommendation Model (MrRec) that mainly consists of a Multilingual aspect-based sentiment analysis module (MABSA) and a Multilingual recommendation module for rating prediction. Fukumoto et al. [32] present a collaborative filtering method for hotel recommendation incorporating user choices by using aspect-based sentiment analysis.

III. Method

A. Aspect-Based Sentiment Analysis (ABSA)

Aspect-based sentiment analysis operations were performed for each sentence of each review. It contains the following steps:

- Breaking the whole review into sentences by using the Natural Language Toolkit tokenization package and by using sent_tokenization method.
- Splitting the sentences into opinion units: Words in a review contains different importance with each aspect. Most of the times a user's reviews contains multiple aspect for a specific item. To explain it through an example sentence, for instance, we should perform the following operations for the sentence "I liked the room, but the food was horrible." First of all, it is necessary to find the sentence's predicate. In this sentence, "like" and "was" are the verbs we want to find. There is the verb "like" at the beginning, but there is a verb in the

continuation of the sentence, which is "was". If there is more than one verb as in the example, the model looks at the rest of the sentence to find it. Thus, it detects more than one emotion in a sentence. By looking at the properties and positions of the tokens, it determines the domain of verbs in the sentence. In this way, sentences are divided into judgments according to different aspects. These judgments are called opinion units.

- Word Extraction stage: Operations include removing stop words (e.g., at, on, in), removing words containing less than three letters, removing numbers written in a typeface.
- Lemmatization stage: Process of distinguishing verbs and nouns by identifying a word's position in a sentence using NLTK.
- Detection of aspects phase: Aspects are found by identifying nouns in the sentence. The most frequently used nouns are designated as aspects. We used aspects that [25] categorized. For instance, if the aspect from the sentence is "Railway", "View", "Station", "Airport" or "Distance", the polarity of the "Location" category is obtained. To give another example, if the aspect of the sentence is "Drink", "Breakfast", or "Food", the polarity of the "Meal" category is obtained.
- The polarity was found using the SentimentIntensityAnalyzer method in the VaderSentiment library. For instance, this method helps classify polarity as positive if the word "great" is used for a particular aspect and negative if "bad" is used.
- Adding polarity phase, polarities created and extracted feature column according to aspects. If a user has made more than one review for a hotel and two of their reviews have two different polarities for an aspect, they are averaged. The information on how many times the user has reviewed the hotel has also been added as a feature since the number of visits to the hotel is an important indicator in the interaction between the user and the hotel.
- Polarity features are given between negative (-1) and positive (+1). Finally, each extracted polarity was normalized into the interval [0,1].

B. Model

The proposed model is based on the architecture implemented in [2]. As shown in Fig. 1, the model is composed of 3 parts: The first part includes aspect-based sentiment analysis. ABSA is applied to the reviews, then polarities are extracted, and the final dataset, fed by polarity features, is obtained.

In the second part, we get the concatenated author and hotel embeddings and feed them as input to the Polarities and Sub-ratings DNN. Since only IDs of users and items are used as a feature, this model uses absolute collaborative filtering. The embedding vectors of users and items are first set to random values, and then the values are changed during model training to minimize the loss function. A deep neural network is used to predict a user's criteria ratings and polarities on an item in this step.

Equation 1 shows the output of a hidden layer l. While W represents the weight matrix, b stands for bias vector.

$$h_i = ReLU(W_i h_{l-1} + b_l) \tag{1}$$

The output layer formulated as Equation 2. The output of Polarities and Sub-ratings DNN is prediction of criteria as represented in Equation 3. In the output layer, we predict the user criteria array using Eqs. 2,3 where L is the number of layers.

$$y_c = ReLU(W_L h_{L-1} + b_L) \tag{2}$$

$$y_c = [c_1, c_2, c_3, .., c_k]^T \tag{3}$$

Finally, in the third part, we normalize the continuous features $(c1, c2, c3, .., ck)$ with the normalization formula in Equation 4.

$$x' = (x - x_{min})/(x_{max} - x_{min}) \tag{4}$$

Then, we give obtained criteria array $(c1, c2, c3, .., ck)$ as input to Overall Score DNN to predict the overall score y that represents the probability of the user choosing the hotel shown in Equation 5.

$$y = f(c_1, c_2, c_3, .., c_k) \tag{5}$$

The criteria array contains the sub-ratings given by the user to the hotel and the polarity values obtained by ABSA. The relationship between the overall score and the criteria array is learned in this part. The common and obvious assumption in the aggregation-function-based method is that the overall rating could be assessed using an aggregate function of multi-criteria ratings.

The hidden layer sizes, learning rate, batch size, and embedding vector size used are determined for both DNNs according to the results of experiments in Section IV-A.

IV. Experimental Results

A. Settings

We used Keras [1] with TensorFlow [2] for the neural network model implementation. On the other hand, we used Natural Language Toolkit (NLTK) [3] for aspect-based sentiment analysis.

In Sub-ratings and Polarities DNN, the output layer has 20 neurons that is equal to the number of the sub-ratings plus polarities and the number of reviews. In the Overall score DNN, the output layer has one neuron for the overall score representing the value between 0 and 1, showing the estimation of the probability of the user choosing the hotel. Experiments were conducted to determine the number of hidden layers. The results of architectures with different hidden layer sizes are shown in Table I. Mean Absolute Error (MAE) was

[1] https://keras.io/
[2] https://www.tensorflow.org/
[3] https://www.nltk.org/

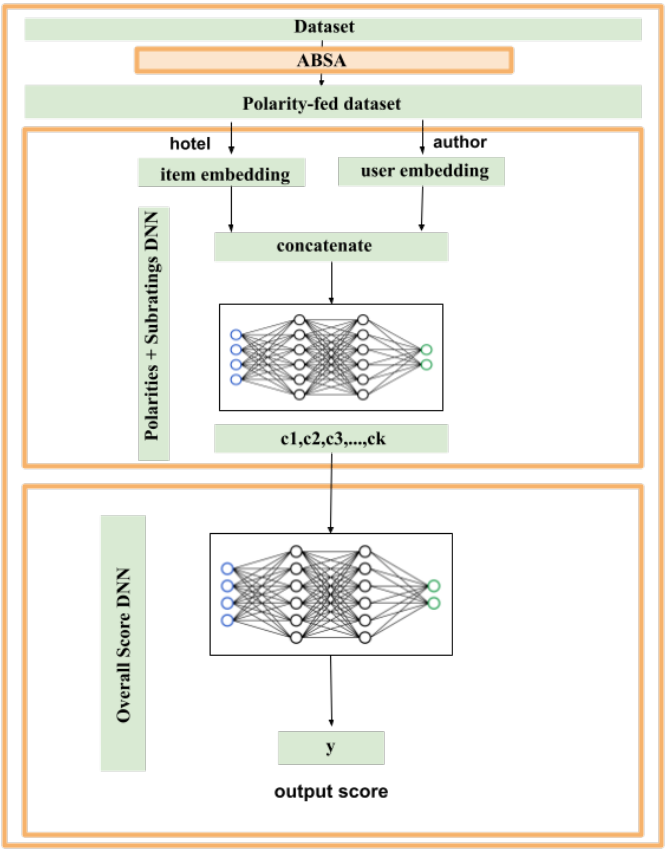

Fig. 1. The architecture of the proposed model: Deep Neural Networks and ABSA

calculated by estimating the overall score. Selected learning rate, optimizer, batch, epoch, and embedding size based on the results of experiments shown in Table II. Rectified Linear Units (ReLU) was chosen as the activation function since it produces the most desired effect [34].

TABLE I
HIDDEN LAYER SIZES

Hidden Layers	MAE	Elapsed Time (s)
[16, 8]	0.11192	195.5212
[32, 16, 8]	0.10101	217.5914
[64, 32, 16, 8]	0.10161	217.5524
[128, 64, 32, 16, 8]	0.18150	277.5289
[256, 128, 64, 32, 16, 8]	0.26197	238.8719
[512, 256, 128, 64, 32, 16, 8]	0.17999	277.4658
[1024, 512, 256, 128, 64, 32, 16]	**0.10247**	**343.1469**
[2048, 1024, 512, 256, 128, 64, 32]	0.12090	337.6681

B. Dataset

In this study, the model was evaluated with a multi-criteria rating dataset called HotelRec [35], containing 50 million reviews. This dataset is based on TripAdvisor and is a very large-scale, convenient data for a hotel recommendation system. Each review contains the user name and hotel URL, the date, the overall rating, the summary of the review, the whole

TABLE II
HYPERPARAMETER SEARCH SPACE AND BEST HYPERPARAMETERS

Hyperparameter	Search Space	Selected Value
Learning Rate	[0.05-0.001]	0.001
Optimizer	[Adam, Adamax, Rmsprop]	Adam
Embedding Size	[32-512]	256
Batch	[8-128]	32
Epoch	[8-128]	32

review text, and the sub-ratings. The dataset of 50 million reviews was split into 11 pieces for analysis, and each part had about 5 million comments. For these pieces, we performed the following operations, respectively. First of all, rows with user columns undefined or none were removed. As a result, about 3 percent of data has been reduced. After that, reviews made in the last ten years were selected. Finally, users with ten or more reviews were selected for the model. As a result, approximately 1% of the original data is suitable for the model.

593,000 rows were extracted from 50 million reviews. This filtered dataset includes 24,890 unique users, 178,096 unique hotels, and 31191 hotel locations.

There are null values for sub-ratings (Service, Location, Value, Cleanliness, Sleep, Rooms, FrontDesk, Business) in the hotel dataset, and almost all of the ratings given for FrontDesk and Business are null. Therefore, the data for these two features (FrontDesk and Business) were ignored and not examined. In addition to these, the number of reviews made by users to the hotel was added as a feature since it could be useful for the model. Figure 2 shows the distribution of users according to the number of reviews they make and the average of the overall ratings they give. This result shows us that as the total number of reviews made by the user increases, the average rating decreases slightly.

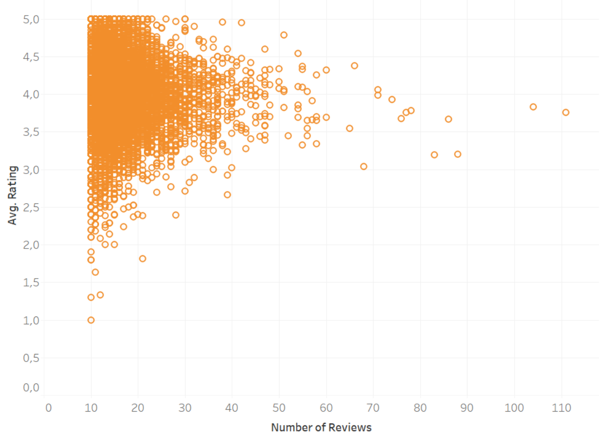

Fig. 2. The Average Rating and Number of Reviews Given by the User

On the other hand, Table III shows the distribution of ratings and Table IV shows null rates and mean values of ratings. As can be seen in these results, the null rate is relatively high in sub-ratings. In addition, the overall rating distribution is not very distinctive.

TABLE III
OVERALL RATING AND SUB-RATINGS DISTRIBUTIONS

	Overall	Service	Location	Value	Cleanliness	Sleep	Rooms
1	1.7%	1.5%	0.3%	0.9%	0.6%	0.8%	0.7%
2	4.6%	2.7%	1.0%	2.1%	1.3%	1.6%	2.0%
3	19.3%	11.3%	6.2%	8.8%	5.6%	6.6%	9.2%
4	40.7%	23.3%	14.5%	15.2%	14.1%	14.2%	15.0%
5	33.7%	32.7%	20.5%	15.5%	20.8%	17.1%	14.6%

TABLE IV
RATING PROPERTIES

	Overall	Service	Location	Value	Cleanliness	Sleep	Rooms
Null Rate:	0.0%	28,5%	57,5%	57,5%	42,5%	40,2%	58,4%
Mean:	4.0014	4.1617	4.2659	3.9971	4.2493	4.1198	3.9796
Normalized Mean:	0.7399	0.5946	0.3723	0.3479	0.3704	0.3406	0.3382

Finally, polarity feature columns were created according to aspects in the dataset as explained in Section III. Total polarity is the general polarity that emerges from the whole review. If there is no sentence about an aspect in the review, a neutral (0) is given for that aspect. Mean values of extracted polarity columns are shown in Table V. These values were extracted in interval [0,1].

TABLE V
ABSA: EXTRACTED POLARITIES

	Mean
Total Polarity	0.65327
Value Polarity	0.42399
Location Polarity	0.49672
Service Polarity	0.50180
Meal Polarity	0.47155
Facility Polarity	0.48178
Room Polarity	0.52255
Quality Polarity	0.46807
Staff Polarity	0.51647
Surrounding Polarity	0.48803

C. Evaluation metrics

Since we make predictions between 0-1, we can interpret 0.5 and above as the user prefers that hotel. However, since the recommendation system also requires a ranking, an ordered list can be suggested to the user based on this score. For this reason, we used MAE for the normalized rating prediction and the F-score evaluation based on the recall and precision values.

1) Mean Absolute Error (MAE):

$$MAE = \frac{1}{|N|} \sum_{r'_{ui} \epsilon N} |r_{ui} - r'_{ui}| \tag{6}$$

where r'_{ui} is predicted score of user u for item i, r_{ui} is the actual score, and $|N|$ is size of test set.

2) F-Score: F1 is calculated as Eq. 7. Precision is calculated as recommended and relevant items divided by recommended items. Recall is calculated as recommended and relevant items divided by relevant items. The definition of relevant item is based on its true rating r_{ui} which is greater than a specific threshold. If the estimated rating r'_{ui} is greater than the given threshold in its k-neighbours then that item is considered as a recommendation to the user. We set the threshold 0.5 since predictions in interval 0-1.

$$F1 = 2 \times \frac{(precision \times recall)}{(precision + recall)} \tag{7}$$

D. Comparisons with the state-of-the-art methods

The Surprise library is used [36] to compare singular recommendation methods e.g., SVD, SVD++, Baseline, and KNN. In addition, to directly predict the score a singular neural network is used with only multi-criteria ratings to show the difference between the version with added polarities. To check the statistical authenticity of the achieved results for proposed method the Wilcoxon Signed Rank Test [37] is applied with the significance value of 0.05. The 10-cross validation results in terms of MAE and F1-Score indicated in that Table VI, illustrate that multi-criteria ratings and ABSA features are better than the baseline models. The results shown in the table indicate that our approach achieved 3% better results than the previous baseline and state of the art methods.

TABLE VI
METHOD COMPARISONS

Method	**MAE**	**F1-Score**
BaselineOnly	0.164395	0.728009
SVD++	0.168171	0.723819
KNNBaseline	0.171455	0.730498
KNNWithZScore	0.175365	0.753533
KNNWithMeans	0.175581	0.752269
KNNBasic	0.177417	0.819961
SVD	0.177629	0.723647
NormalPredictor	0.244793	0.639061
CoClustering	0.403692	0.759395
Single Rating DNN	0.154805	0.774335
Multi-criteria Ratings DNN	0.102469	0.794236
Proposed model	**0.099484**	**0.823616**

V. CONCLUSION AND FUTURE WORK

In this paper, we proposed a deep recommender system using aspect-based sentiment analysis of users' reviews for hotel recommendations. To the best of our knowledge, this is the first study that merged reviews and extracted polarity for a deep multi-criteria collaborative filtering model. The proposed model discovers the relationship between overall rating and sub-ratings with aspect polarities. The model is composed of 3 parts: the first part includes aspect-based sentiment analysis which is applied to the reviews, then polarities are extracted, and polarity features feed the final dataset is obtained. In the second part, we get the author and hotel embeddings, concatenate them, and give them as input to a Polarities and Sub-ratings Deep Neural Network (DNN) model to predict the sub-ratings and polarities. Finally, in the third part, we give obtained criteria array as input to another DNN to predict the overall score. The proposed model was evaluated with a multi-criteria rating and review dataset called HotelRec, based on TripAdvisor, and the results are shown in detail.

Future work will improve the dataset with location and season information. Location information can be extracted and classified from the hotel URL. However, there may be problems in extracting the locations since the place names in the data are not very proper. In addition, users can be categorized as domestic - international travelers according to their activities. Traveler-type information such as single, couple, family, business, and friends could benefit. With such extra data, the system can be enriched, and the number of features can be increased. In future work, we also plan to improve the performance of the deep learning model by studying different deep learning methods such as Recurrent Neural Network (RNN), Generative Adversarial Networks (GAN), Convolutional Neural Network (CNN), and Autoencoder or other feature representation methods.

References

[1] Ricci, F., Rokach, L., & Shapira, B. (2011). Introduction to Recommender Systems Handbook. *Recommender Systems Handbook*, 1–35. https://doi.org/10.1007/978-0-387-85820-3_1

[2] Nassar, N., Jafar, A., & Rahhal, Y. (2020). A novel deep multi-criteria collaborative filtering model for recommendation system.*Knowledge-Based Systems, 187*, 104811. https://doi.org/10.1016/J.KNOSYS.2019.06.019

[3] Burke, R., "Knowledge-based recommender systems." Encyclopedia of library and information systems 69.Supplement 32 (2000): 175-186.

[4] Tao, Z., Wei, Y., Wang, X., He, X., Huang, X., & Chua, T. S. (2020). MGAT: Multimodal Graph Attention Network for Recommendation. *Information Processing & Management, 57*(5), 102277. https://doi.org/10.1016/J.IPM.2020.102277

[5] Herlocker, J. L., Konstan, J. A., Borchers, A., & Riedl, J. (2017). An Algorithmic Framework for Performing Collaborative Filtering. *ACM SIGIR Forum, 51*(2), 227–234. https://doi.org/10.1145/3130348.3130372

[6] Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2001). Item-based collaborative filtering recommendation algorithms.*Proceedings of the 10th International Conference on World Wide Web, WWW 2001,* 285–295. https://doi.org/10.1145/371920.372071

[7] Grčar, M., Fortuna, B., Mladenič, D., & Grobelnik, M. (2006). kNN Versus SVM in the Collaborative Filtering Framework. *Data Science and Classification, 251–260.* https://doi.org/10.1007/3-540-34416-0_27

[8] Breese, J. S., Heckerman, D., & Kadie, C. (2013). *Empirical Analysis of Predictive Algorithms for Collaborative Filtering*. https://arxiv.org/abs/1301.7363v1

[9] Hofmann, T. (2003). Collaborative Filtering via Gaussian Probabilistic Latent Semantic Analysis. *SIGIR Forum (ACM Special Interest Group on Information Retrieval), SPEC. ISS.*, 259–266. https://doi.org/10.1145/860435.860483

[10] Bell, R., Koren, Y., & Volinsky, C. (2007). Modeling relationships at multiple scales to improve accuracy of large recommender systems. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 95–104. https://doi.org/10.1145/1281192.1281206

[11] Zhang, K., Wang, K., Wang, X., Jin, C., & Zhou, A. (2015). Hotel recommendation based on user preference analysis. *Proceedings - International Conference on Data Engineering, 2015-June,* 134–138. https://doi.org/10.1109/ICDEW.2015.7129564

[12] Zhou, W., & Han, W. (2019). Personalized recommendation via user preference matching. *Information Processing & Management, 56*(3), 955–968. https://doi.org/10.1016/J.IPM.2019.02.002

[13] Zhang, J., Peng, Q., Sun, S., & Liu, C. (2014). Collaborative filtering recommendation algorithm based on user preference derived from item domain features.*Physica A: Statistical Mechanics and Its Applications, 396,* 66–76. https://doi.org/10.1016/J.PHYSA.2013.11.013

[14] Adomavicius, G., & Kwon, Y. O. (2007). New recommendation techniques for multicriteria rating systems. *IEEE Intelligent Systems, 22*(3), 48–55. https://doi.org/10.1109/MIS.2007.58

[15] T.Y. Tang, G. Mccalla, The pedagogical value of papers: A collaborative filtering based paper recommender, J. Digit. Inf. 10 (2009).

[16] Drucker·, H., Burges, C. J. C., Kaufman, L., Smola··, A., & Vapoik, V., *Support vector regression machines*, Adv. Neural Inf. Process. Syst. 9 (1) (1997) 155–161.

[17] Sahoo, N., Krishnan, R., Duncan, G., & Callan, J. (2012). The halo effect in multicomponent ratings and its implications for recommender systems: The case of yahoo! movies. *Information Systems Research, 23*(1), 231–246. https://doi.org/10.1287/ISRE.1100.0336

[18] Hassan, M., & Hamada, M. (2017). A Neural Networks Approach for Improving the Accuracy of Multi-Criteria Recommender Systems. *Applied Sciences 2017, Vol. 7, Page 868, 7*(9), 868. https://doi.org/10.3390/APP7090868

[19] Hassan, M., & Hamada, M. (2018). Genetic Algorithm Approaches for Improving Prediction Accuracy of Multi-criteria Recommender Systems. *International Journal of Computational Intelligence Systems, 11*(1), 146–162. https://doi.org/10.2991/IJCIS.11.1.12

[20] Fu, M., Qu, H., Yi, Z., Lu, L., & Liu, Y. (2019). A Novel Deep Learning-Based Collaborative Filtering Model for Recommendation System. *IEEE Transactions on Cybernetics, 49*(3), 1084–1096. https://doi.org/10.1109/TCYB.2018.2795041

[21] Salakhutdinov, R., Mnih, A., & Hinton, G. Restricted boltzmann machines for collaborative filtering. In ICDM, pages 791–798, 2007.

[22] Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2001). Item-based collaborative filtering recommendation algorithms. *Proceedings of the 10th International Conference on World Wide Web, WWW 2001*, 285–295. https://doi.org/10.1145/371920.372071

[23] Liang, D., Charlin, L., McInerney, J., & Blei, D. M. (2015). Modeling User Exposure in Recommendation. *25th International World Wide Web Conference, WWW 2016*, 951–961. https://doi.org/10.1145/2872427.2883090

[24] He, X., Zhang, H., Kan, M. Y., & Chua, T. S. (2017). Fast Matrix Factorization for Online Recommendation with Implicit Feedback. SIGIR 2016 - *Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 549–558. https://doi.org/10.1145/2911451.2911489

[25] Akhtar, N., Zubair, N., Kumar, A., & Ahmad, T. (2017). Aspect based Sentiment Oriented Summarization of Hotel Reviews. *Procedia Computer Science, 115*, 563–571. https://doi.org/10.1016/J.PROCS.2017.09.115

[26] Zhang, Y., Du, J., Ma, X. et al. Aspect-Based Sentiment Analysis for User Reviews.*Cogn Comput 13*, 1114–1127 (2021). https://doi.org/10.1007/s12559-021-09855-4

[27] Bai, P., Xia, Y., & Xia, Y. (2020). Fusing Knowledge and Aspect Sentiment for Explainable Recommendation. *IEEE Access, 8*, 137150–137160. https://doi.org/10.1109/ACCESS.2020.3012347

[28] Zhang, Y., Lu, H., Jiang, C., Li, X., & Tian, X. (2021). Aspect-Based Sentiment Analysis of User Reviews in 5G Networks.*IEEE Network, 35(4)*, 228–233. https://doi.org/10.1109/MNET.011.2000400

[29] Musto, C., Lops, P., De Gemmis, M., & Semeraro, G. (2019). Justifying recommendations through aspect-based sentiment analysis of users' reviews.*ACM UMAP 2019 - Proceedings of the 27th ACM Conference on User Modeling, Adaptation and Personalization*, 4–12. https://doi.org/10.1145/3320435.3320457

[30] Bauman, K., Liu, B., & Tuzhilin, A. (2017). Aspect Based Recommendations: Recommending Items with the Most Valuable Aspects Based on User Reviews. *ACM Reference Format, 9.* https://doi.org/10.1145/3097983.3098170

[31] Liu, P., Zhang, L., & Gulla, J. A. (2021). Multilingual Review-aware Deep Recommender System via Aspect-based Sentiment Analysis. *ACM Transactions on Information Systems (TOIS), 39(2).* https://doi.org/10.1145/3432049

[32] Fukumoto, F., Sugiyama, H., Suzuki, Y., & Matsuyoshi, S. (2014). Exploiting Guest Preferences with Aspect-based Sentiment Analysis for Hotel Recommendation. *Communications in Computer and Information Science, 553*, 34–49. https://doi.org/10.1007/978-3-319-25840-9_3

[33] Gojali, S., & Khodra, M. L. (2016). Aspect based sentiment analysis for review rating prediction. *4th IGNITE Conference and 2016 International*

Conference on Advanced Informatics: Concepts, Theory and Application, ICAICTA 2016. https://doi.org/10.1109/ICAICTA.2016.7803110
[34] Fuentes, O., Parra, J., Anthony, E., & Kreinovich, V. (2020). Why Rectified Linear Neurons Are Efficient: A Possible Theoretical Explanation. *Studies in Computational Intelligence, 835*, 603–613. https://doi.org/10.1007/978-3-030-31041-7_34
[35] Antognini, D., & Faltings, B. (2020). HotelRec: a Novel Very Large-Scale Hotel Recommendation Dataset. *LREC 2020 - 12th International Conference on Language Resources and Evaluation, Conference Proceedings*, 4917–4923. https://arxiv.org/abs/2002.06854v1
[36] Hug, N. (2020). Surprise: A Python library for recommender systems. *Journal of Open Source Software*, 5(52), 2174. https://doi.org/10.21105/JOSS.02174
[37] Fix, E., & Jr, J. L. H. (1955). Significance Probabilities of the Wilcoxon Test., 26(2), 301–312. https://doi.org/10.1214/AOMS/1177728547