Journal of Information Technology Management

Print ISSN: 2008-5893

Online ISSN: 2423-5059

A Grouping Hotel Recommender System Based on Deep Learning and Sentiment Analysis

Fatemeh Abbasi

Postdoctoral researcher, Department of Social and Economic, Alzahra University, Tehran, Iran. ORCID: https://orcid.org/0000-0002-6918-9679. E-mail: f.abbasi@alzahra.ac.ir

Ameneh Khadivar*

*Corresponding author, Associate Professor, Department of Social and Economic, Alzahra University, Tehran, Iran. ORCID: 0000-0003-4620-1409. E-mail: a.khadivar@alzahra.ac.ir

Mohsen Yazdinejad

Ph.D. Student, Faculty of Computing Engineering, University of Isfahan , Iran. ORCID: 0000-0001-7805-6344 E-mail: mohsen.yazdinejad@gmail.com

Abstract

Recommender systems are important tools for users to identify their preferred items and for businesses to improve their products and services. In recent years, the use of online services for selection and reservation of hotels have witnessed a booming growth. Customer' reviews have replaced the word of mouth marketing, but searching hotels based on user priorities is more time-consuming. This study is aimed at designing a recommender system based on the explicit and implicit preferences of the customers in order to increase prediction's accuracy. In this study, we have combined sentiment analysis with the Collaborative Filtering (CF) based on deep learning for user groups in order to increase system accuracy. The proposed system uses Natural Language Processing (NLP) and supervised classification approach to analyze sentiments and extract implicit features. In order to design the recommender system, the Singular Value Decomposition (SVD) was used to improve scalability. The results show that our proposed method improves CF performance.

Keywords: Grouping recommender systems; Sentiment analysis; Deep learning; Singular Value Decomposition (SVD).

DOI: 10.22059/jitm.2019.289271.2402

Introduction

Nowadays, a huge amount of information is produced by the users of the social media. Consequently, people face some problems in the optimal use of this information for their decision-making purposes. The first problem is related to the individuals who need to find their intended content in a massive volume of information. The other problem relates to online businesses and their great demand for the recommendation of their products on the web environment (Alahmadi & Zeng, 2015). On the other hand, this information greatly helps the users to find the items and services based on their preferences (Ravi & Vairavasundaram, 2016). The fact that users and customers have a lot of alternatives makes the selection of the best choice a difficult decision. In this situation, the conditions for the filtering of the information and its personalization for each specific user become significantly important. The recommender systems are one of the modern tools for the recommendation of items to the users based on their needs and preferences. Nowadays, the recommender systems are used in various domains, e.g. movies and books. Due to the expansion of the tourism industry in the last decade, the hotel and travel recommender systems have also attracted researchers as well as companies that provide related services. Sometimes, deciding on one residence from among many available alternatives can truly perplex a tourist (Ebadi & Krzyzak, 2016). The classic recommender systems use ratings information history in order to provide recommendations. In recent years, methods such as the use of users' online reviews in order to design recommender systems have been suggested. In these systems, the NLP is used to analyze users' online reviews in order to find their sentiments (Alahmadi & Zeng, 2015).

Nonetheless, the majority of available recommender systems use a simple recommendation provision method that is based on the explicit features of the guests' profiles such as the users' rating of the hotels (Gavalas & Kenteris, 2011). Users' reviews comprise a valuable source to tell us about their preferences. The users' opinions and sentiments show the implicit aspects and features of the items and improve the quality of the users' profiles (María, 2017). Studies that explore the users' opinions using NLP tools are the most preferred trends in the text analysis literature. Nowadays, sentiment analysis has turned into an important tool for decision making. Hundreds of thousands of users are influenced by the positive or negative reviews about goods and services provided by other users of the social media. In April 2013, ninety percent of the customers' decisions have been dependent on the positive/negative reviews on the social media (Hussein, 2018).

In this study, collaborative filtering is used to design the recommender system. As shown in Figure 1 and Figure 2 sentiment analysis and recommender system have attached more attention in recent years. In Figure 1 the results of the VOSviewer analysis are presented. Due to network visualization combination of sentiment analysis and recommender system by deep learning are emerging area in recent years. VOSviewer is a software tool for creating maps based on network data and for visualizing and exploring these maps which is used network analysis (Waltman & van Eck, 2017).

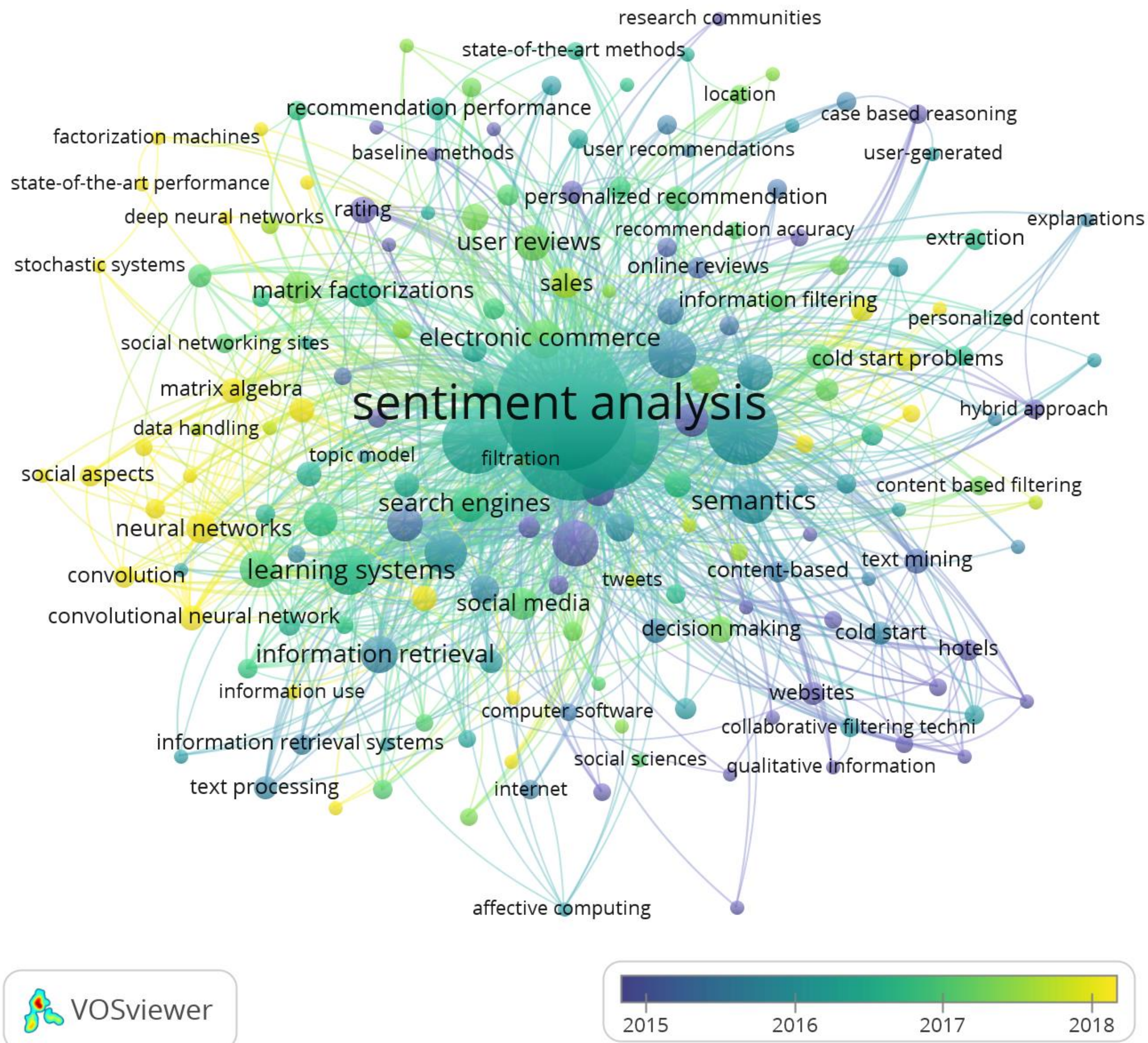

Figure 1. Term map of full text of sentiment analysis and recommender system

Figure 2 shows the trend research in sentiment analysis and recommender system. The main constituents of this system are the users, items, and the items-users matrix. As each user has reviewed only one hotel, the matrix sparse is very high. In order to solve this problem, a combination of grouping recommender system and deep learning is used in designing the recommender system. We have suggested a framework for the combination of the sentiment analysis and matrix factorization to design the grouping recommender system that helps improve the performance of the recommender systems. The key point in this research is the extraction of the item features using the machine learning algorithms and embedding the features in the collaborative recommender system.

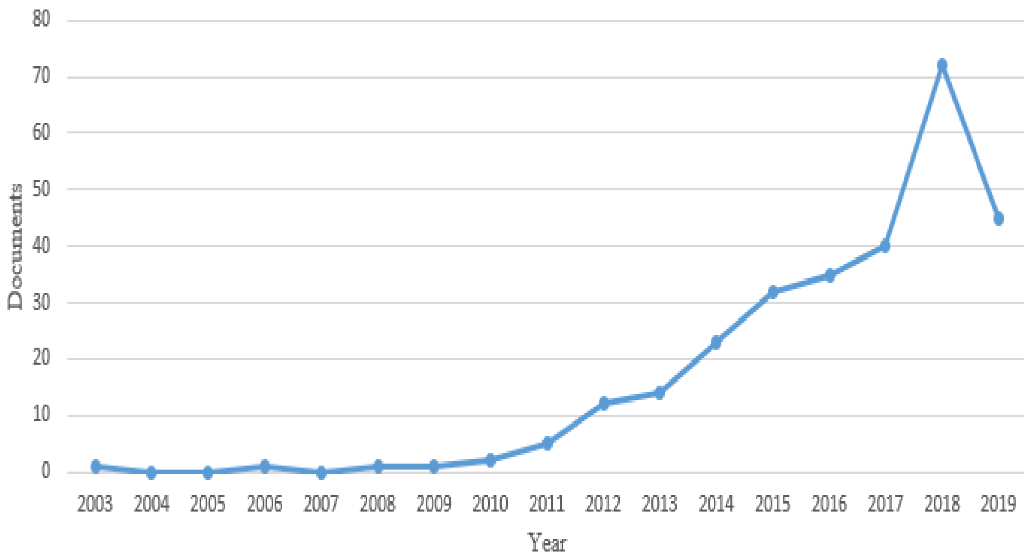

Figure 2. Trend of research in sentiment analysis and recommender system

Related Works

Sentiment Analysis

Sentiment analysis is a technique that has been used in recent years to understand the perceptions and the characteristics of the user reviews (Ribeiro, Araújo, Gonçalves, Benevenuto, & Gonçalves, 2015). Sentiment analysis is a domain of research that analyzes the opinions, sentiment, evaluations, and attitudes of the people to the entities such as products, services, organizations, individuals, issues, events, topics and their characteristics (Liu, 2012). In the sentiment analysis, Natural Language Processing (NPR), text analysis, and computational techniques are used to automatically extract and classify sentiments from user reviews. Sentiment analysis is used in various domains such as marketing, selecting and buying books or reserving hotels (Agarwal, Mittal, Bansal, & Garg, 2015).

Different methods have been suggested for sentiment analysis so far. These can be classified into three main approaches: machine learning-based methods, lexicon-based methods and hybrid methods (Pandey, Rajpoor, & Saraswat, 2017). Figure 3 shows the sentiment analysis method.

The supervised machine learning methods tries to classify and predict the sentiments of the text. The supervised learning approach principally performs better than the lexicon-based approach due to the consideration of effective characteristics and elements (Wang, 2017). The supervised machine learning methods is a type of learning in which the input and output are clear and there is a supervisor that feeds the learner with some information; this way, the system tries to learn an input-to-output function. There are two sets of texts in a machine-learning classification: the training set and testing set. The training set is used by an automatic classification unit to learn the main features of the texts, while the testing set is used for the validation of the automatic classification unit (Medhat, Hassan, & Korashy, 2014).In unsupervised machine learning, the training set include unlabeled data. In this method we use specific algorithms for the clustering of sentiments (Liu, 2015).

Lexicon-based methods can be divided into corpus-based and dictionary-based methods. The corpus-based methods use the relationship between words (the syntactic or co-occurrence relationship) to determine sentiment in a large corpus. The dictionary-based methods act based on the analysis of polarity of each word. These methods can be used for sentiment classification (Liu, 2012).Corpus-based method consider syntactic pattern of words. This method uses statistical and semantic techniques to find the polarity of the sentiment words. The next method is the hybrid that combines the supervised machine learning and lexicon-based methods (Ravi & Ravi, 2015). This method uses a lexicon to find polarity orientation of sentiment words and transform this lexicon into supervised learning based methods (Wang, 2017).

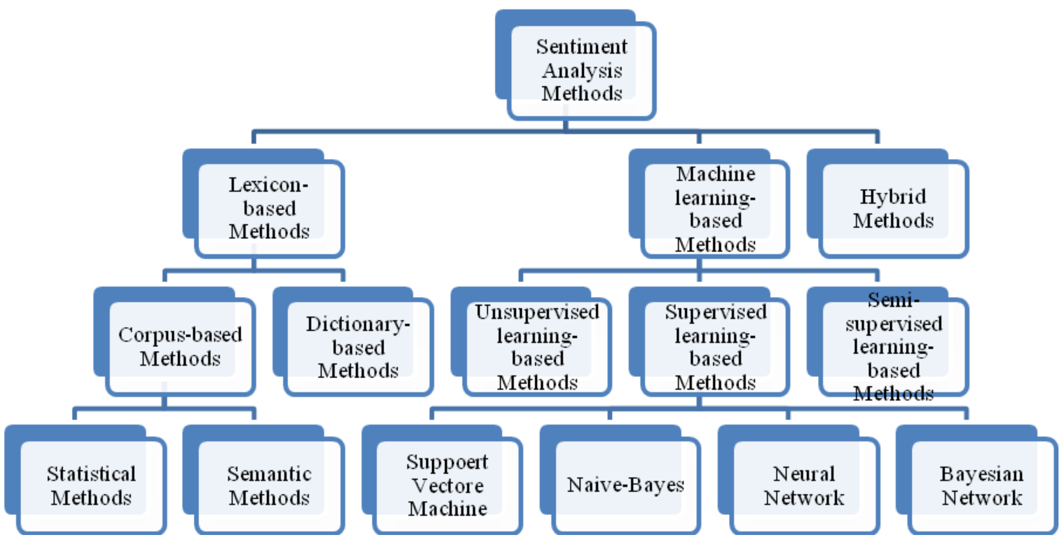

Figure 3. Sentiment Analysis Methods (Pandey, Rajpoor, & Saraswat, 2017)

Recommender systems

Recommender systems have been increasingly used in various domains such as film, music, book, and hotel recommendation. With the increase in social media usage, the use of the recommender system, especially the use of recommender systems in the grouping recommender systems have been augmented. The grouping recommender systems suggest products and services based on group preferences. Moreover, the grouping recommender systems solve the cold start problem in the recommender systems (Dara, Chowdary, & Kumar, 2019). The recommender systems identify and recommend products and services based on users' interest. Recommender systems are divided into three classes: collaborative filtering recommender system, content-based recommender system, and hybrid recommender system (Sulthana & Ramasamy, 2019).

The Content-based method is based on the earlier information of the user and the attributes allocated to each product. The purpose of content-based filtering is to suggest items which are similar to the items that users liked in the past (Leimstoll & Stormer, 2007; Chenni, Bouda, Benachour, & Zakaria, 2015). Collaborative filtering is one of the most important approaches to the development of recommender systems. This approach derives from the fact that the selection and purchase of items and products is based on the experience and reviews of other people. This method is based on the premise that the users who have similar tastes give similar ratings (Das, Sahoo, & Datta, 2017) .The collaborative filtering methods are divided into three classes, which have been summarized in Table 1 below (Bokde, Girase, & Mukhopadhyay, 2015). Hybrid methods combine collaborative and content-based methods to create a more robust framework. Through the combination of recommender systems methods, we can cover the disadvantages of one with the advantages of the others (Das, Sahoo, & Datta, 2017). Table 1 shows the overview of collaborative filtering techniques.

Table 1. Overview of Collaborative Filtering Techniques (Bokde, Girase, & Mukhopadhyay, 2015)

Collaborative Filtering Techniques	Representative Algorithm	Advantages	Limitations
Memory-Based Collaborative Filtering (Neighborhood Based)	- User-Based CF - Item-Based CF	- Easy implementation - New data can be added easily and incrementally - Need not consider the content of items being recommended	- Are dependent on human ratings - Cold star problem for new user and new item - Sparsity problem of rating matrix - Limited scalability for large datasets
Model-Based Collaborative Filtering	- Stop-One CF - Dimensionality Reduction Matrix (Matrix Factorization), e.g. SVD	- Better addresses the sparsity and scalability problem - Improve prediction performance	- Expensive model building - Trade-off between the prediction performance and scalability - Loss of information in dimensional reduction technique (SVD)
Hybrid Collaborative Filtering	- Combination of Memory-Based and Model-Based CF	- Overcome limitations of CF such as sparsity and grey sheep - Improve prediction performance	- Increase complexity ad expense for implementation

In the memory-based approach, the user-items ratings matrix is used. These systems are time-aware and often use similarity measures to assess the degree of closeness between the users and the items (Adomavicius & Tuzhilin, 2005). The model-based systems rely on the users' rating database to construct the model. These raw data are pre-processed offline and then the user model is developed. The system relies on the developed model to recommend the items. If any new information is obtained about a user, the model is outdated (Su & Khoshgoftaar, 2009) .Two important issues in the collaborative systems are the cold start and data sparsity. Cold start occurs when a new recommender system enter the system. In this situation, there is no information available about the users' ratings of the items. In other words, there is no information about the interaction between the users and the items (Schafer, Frankowski, Herlocker, & Sen, 2007). Data sparsity occurs when users are only interested in certain items and only rate them. Therefore, most of the cells of the item-user matrix are null (Huang, Chen, & Zeng, 2004). In order to solve this problem, many researchers have used dimensionality reduction methods that rely on matrix factorization (Sarwar, Karypis, Konstan, & Riedl, 2000).

One of the most popular methods in collaborative filtering is the matrix factorization, which calculates the latent factors of the users and items by analyzing the item-user matrix (Xiao & Shen, 2019). Matrix factorization emphasizes the factorization of the rating matrix into two dimensions: user latent vectors and item-latent vectors. These techniques involve the extraction of attributes using latent variables in order to describe the latent reasons for the co-occurrence data. Approaches such as Probabilistic Matrix Factorization (PMF) and Singular Value decomposition (SVD) are used in matrix factorization (Wei, He, Chen, Zhou, & Tang, 2016).

Mnih and Salakhutdinov (2008) have proposed the Probabilistic Matrix Factorization (PMF) algorithm. In this approach, the user latent features and item latent features are used to fit the ratings (Zhang, Luo, Zhang, & Wu, 2018). Value Decomposition (SVD) is used to fit the optimal model in order to relate the users and the items. In this method, we aim at identifying the model parameters including the user and item latent factors though user ratings. The method is used to predict the relationship between users and items based on implicit preferences (Bhavana, Kumar, & Padmanabhan, 2019).

In this method, the matrix $A_{n,m}$ is factorized into the three matrixes of $A = U\Sigma V^T$. The U and V matrixes are the two orthogonal matrixes with the n×n and m×m dimensions, respectively. Σ is the singular orthogonal matrix with m×n dimensions and non-negative elements. In this method, the null values of the users-items matrix are filled with each user's rating average so as to make the extraction of the latent significant relationships possible. One of the features of SVD is the reduction of matrix rank, which is calculated by the following formula:

$$(1) \quad SVD(A_k) = U_k \Sigma_k V_k^T$$

Where U_k and V_k are matrixes with m×k and n×k dimensions are combined with the 1st k columns in matrixes U and K, respectively. Matrix Σ_k is the diagonal sub-matrix of Σ with k×k dimensions. The closest approximation of matrix A with reduce rank k is A_k (Bokde, Girase, & Mukhopadhyay, 2015).

Prediction is a numerical value that the recommender system returns for the user rating *U* of the item *I*. Often, the Mean Score Error (MSE) or Root Mean Score Error (RMSE) is used as the criterion to evaluate the accuracy of predicted value. This criterion is applied to the items that have been previously rated by the user. The value of this criterion is calculated through the following formula. *P* represents the predicted rating and *a* is the true rating (Cremonesi, Turrin, Lentini, & Matteucci, 2008).

$$(2)\quad RMSE = \sqrt{\frac{\sum_{i=1}^{n}(p_i - a_i)^2}{n}}$$

Background

Divyashree et al. (2017) have had a look at the user reviews of the hotels and analyzed the positive and negative words in this regard. Using the J48 algorithm, they set out to classify the positive and negative words on the TripAdvisor website and showed the top ten hotels (Divyashree, Santhosh Kumar, & Majumdar, 2017). Geetha et al. (2017) found a relationship between sentiment analysis and users' rating of the hotels. They found out that users' sentiments comprise a good measure of their ratings of the hotels (Divyashree, Santhosh Kumar, & Majumdar, 2017). With regard to sentiment analysis, Hou et al. (2019) offered the semantic association analysis approach. The results of their study reveal that the level of services offered by tour guides has significant effects on the tourists' satisfaction (Hou, Cui, Meng, Lian, & Yu, 2019).

Through the identification of the features related to the suggested item which have been automatically obtained through the exploration of users' reviews, Musto et al. (2019) offered their recommendation algorithm. The results of their study demonstrated that users prefer justifications which are based on reviews over other methods of explanation (Musto, Lops, Gemmis, & Semeraro, 2019). Abdul Hassan and Abdulwahhab (2019) offered a recommendation method which was based on review analysis. They used supervised, lexicon-based techniques to analyze the sentiments. Their proposed system extracts features using sentiment analysis. The accuracy level of their proposed sentiment analysis method was very high (Abdul Hassan & Abdulwahhab, 2019). Hernández-Rubio et al. (2019) offered a recommendation method which was based on three principles: the identification of the references to item features in the users' reviews, the classification of the sentiment orientations that exist in the reviews, and the use of information related to these features. They provide a useful list of resources for the exploration of reviews and recommendation purposes; for example, domain aspect vocabularies and domain-dependent, aspect-level lexicons (Hernández-Rubio, Cantador, & Bellogín, 2019).

Sun et al. (2018) proposed a new recommendation algorithm based on time cost. In order to specify the sentiment polarity and the strength level of the sentiments, they applied uncertain sets and statistics – founded on the uncertainty theory – to obtain UD-KM (Sun, Guo, & Zhu, 2018). Catalin-Mihai and Ziegler (2018) presented a hotel recommender system which works based on the users' reviews. The system uses interactive mechanisms to improve the satisfaction of the users (Catalin-Mihai & Ziegler, 2018). Zheng et al. (2018) proposed a new recommendation system through context clustering. The system works based on matrix factorization techniques and sentiment analysis (Catalin-Mihai & Ziegler, 2018). Silamai et al. (2017) devised a recommender system for tourists who do not have a specific plan in the destination city. Their system was based on the touristic information of the destination city and could be evaluated by the guests' experience (Silamai, Khamchuen, & Phithakkitnukoon, 2017). Another recommender system is the one offered by Takuma et al (2016). Their system shows a list of hotels derived from the preferences of the guests. In this method, the guests' preferences are obtained from their reviews (Takuma, Yamamoto, Kamei, & Fujita, 2016). Sharma et al. (2015) offered a multi-criteria recommendation system for the hoteling sector based on the data they had obtained from TripAdvisor website. The system recommends the most appropriate hotels based on the guests' preferences and the reviews of other guests (Sharma, Bhatt, & Magon, 2015). Zhang et al. (2015) provided a new framework for hotel recommendation that was founded on the combination of collaborative filtering and content analysis. This system also uses the analysis of user preferences as a provider of extra information (Zhang, Wang, Wang, Jin, & Zhou, 2015).

Levi et al. (2012) demonstrated that a content-based inquiry can ease the construction of cold start recommendation systems. Their study also indicated that the preprocessing of large sets of texts can be used to find the common features of groups (Levi, Mokryn, Diot, & Taft, 2012). Saga et al. (2008) used sales records to make a hotel recommender system. The system is comprised of two steps: the selection of the hotels to be recommended through target-node and tracking-node selection, and then ranking the available alternatives based on in-degree and out-degree in preference transition network (Saga, Hayashi, & Tsuji, 2008).

Materials and Methods

This study focuses on improving the recommender system performance using the results of sentiment analysis of user reviews. The purpose of this study is the sentiment analysis of user reviews and recommendations based on explicit and implicit feedbacks of the hotel guests.

In order to design the recommender systems, two datasets were used, namely the user ratings that express the guests' explicit feedback on the hotels, and the sentiment analysis of the users' reviews that represent the guests' implicit feedback about 32 hotels in five Iranian cities.

With regard to the purpose of the study, this research is a developmental study. At the same time, due to its advantages for the companies that provide online services, this study is an applied research, too. Because its findings help the online customers and users decide on their needed products based on the experiences of other people, and aids the managers and decision makers of these businesses to improve their performance and make more informed decisions. The research population of the study is comprised of the user reviews on social media about Iranian hotels. The methodology of the study shows in Figure 2.

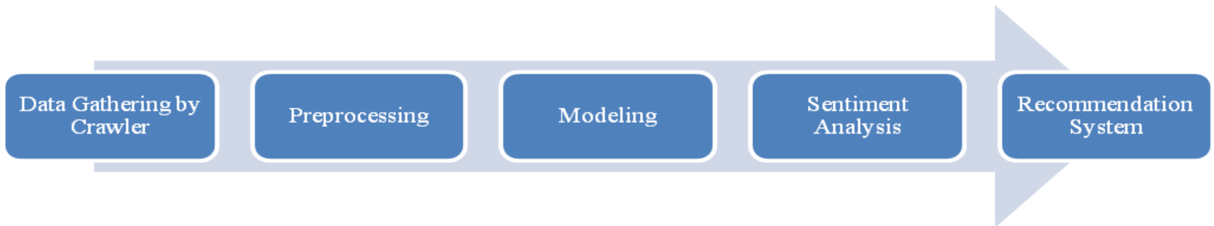

Figure 4. Methodology of proposed system

Data Gathering

Machine learning methods require a lot of data for prediction. Therefore, we have used the crawler software Octaparese (www.octoparse.com) to extract data. This software has been designed to extract data from webpages. Using this web crawler software, nearly seven thousand reviews were extracted from iranhotel.com website.

The collected data is related to users' reviews about the hotels of Mashhad, Isfahan, Shiraz, and Yazd. After the settings were tuned in the software, required data was extracted and saved in the .xls file format. The collected data includes information about the source city, the destination city, hotel name, hotel rating, comment title, the commenter name, the comment text, travel type, length of stay in the hotel, room type, and user rating. A sample of the data is shown in Table 2.

Table 2. A sample of the collected data

Hotel name	Hotel rating	Commenter name	Source city	Destination city	Comment title	Comment text	Travel type	Length of stay	Room type	User rating
Hotel A	5 star	Ahmadi	Mashhad	Kish	Good	Excellent staff	Family Travel	4 nights	3 bed room	5

Pre-processing

Since people use short and quick messages in their reviews, the data is not clean. Therefore, it is necessary to pre-process the data to prepare it for modeling (Ebadi & Krzyzak, 2016). Pre-processing is one of the main steps in text mining, Natural Language Processing (NLP) and Information Retrieval (IR). The characters, words, and sentences are identified in this step (Kannan & Gurusamy, 2014). Pre-processing consists of techniques such as the removing stop words, stemming, and part of speech (POS) tagging (Vijayarani, Ilamathi, & Nithya, 2015). The pre-processing of Persian texts is in some ways different from the processing of English texts. In English, all letters and words are written separately with clear rules. In Persian, however, some letters are connected while some are separate; moreover, some words are integrated while some others are divided into two or more parts using spaces or half-spaces. During the pre-processing stage, the punctuation marks, letters, spaces between words, acronyms, etc. in Persian texts are turned into their standard format with no changes in meaning (Zamani &Sorkhpour, 2014). In order to pre-process and normalize the texts in this study, the following steps were taken:

- Removal of stop words: In Persian texts, highly frequent words such as the conjunctions and propositions do not have semantic values, and so, are removed in this step;
- Removal of special characters: In this step, characters such as tags, html, @, comma, question mark, semi-column, and any other undesirable character are removed;
- Removal of hyphen that is used to connect words;
- Removal of space and half-space used in the text;
- Replacement of English quotation marks with the Persian ones: English and Persian quotation marks are different. In English texts, two small vertical lines (") are used for quotations, while in Persian, Guillemets (« ») are used;
- Review and correction of Hamza sign: for instance, the character ۀ is changed to هی and the character أ is changed to ا;
- Connecting the separately written comparative (تر) and superlative (ترین) signs to the end of the adjective;
- Correction of the space between the plural sign (ها) and the nouns as well as related word endings (های), (هایی), (هایم), (هایت), (هایش);
- Correction of the space between morphemes (می), (نمی), (درمی), (برمی), and (بی) before words;
- Replacement of some repetitious characters with one character;
- Changing the Arabic and English numbers with Persian ones;
- Correction of the spaces between punctuation marks so that these marks attach to the words before them and have a space with the words that follow them;
- Correction of the space within () [] {} “” «» signs;
- Correction of (ک) character and changing it into its Persian format;
- Correction of (ی) character and changing it into its Persian format;

Term Frequency Inverse Document Frequency (TF-IDF): Feature extraction technique is used to extract the main features in text classification, information retrieval, topic identification, and document summarization. The main methods are Term Frequency Inverse Document Frequency (TF-IDF), Information Gain (IG), and Chi-square Statistics (CHI) (Chakraborty, 2013).

In order to calculate this index, the term frequency (the importance of a term in a document) is combined with inverse document frequency (the importance of the term in all documents) to make clear the weight of each word in each document. Using this method, the terms that exist in a smaller set of documents are given more weights and the words that exist in most of the documents receive smaller weights. Consequently, words with smaller weights are excluded from analysis (Manning, Raghavan, & Schütze, 2009).

$$W_t = TF(d,t)\ IDF(t) \quad (3)$$

Based on this equation, the highest weight is obtained when the term t appears in numerous cases in a document or in a limited set of documents. On the contrary, the lowest weight is obtained when the term t appears a few times in the document or in many documents. Moreover, the least weight is obtained when a term appears in all documents. The similarity between the documents is calculated by the following formula, which is equivalent to the cosine similarity method (Appel, Fuchs, Doll´ar, & Perona, 2013; Ertek, Tapucu, & Arın, 2012) .

$$(4)\ \text{sim}\ (d1, d_2) = \frac{d1.d_2}{|d1\ |d_2\ |}$$

This study uses TF-IDF at the character and word level. The results show that TF-IDF has a higher precision at the character level. In order to create the final model of sentiment analysis, the TF-IDF has been used at the character level.

Modeling

After changing the unstructured data to structured data, the extracted terms and features were used for modeling. In this study, Multinomial Naive Bayes (MNB) and Linear Support Vector Classification (Linear SVC) from Python NLTK library have been used for modeling.

Multinomial Naive Bayes (MNB) classifier is extensively used for text classification. It is based on the Bayes learning principles and assumes that the term distribution in the documents is created by certain parametric models (Zhao, et al., 2016).

Linear Support Vector Classification: (LSVC) is the supervised learning method. It is a fairly new method that uses higher efficiency rates compared to the previous methods. The basis of this method is the linear classification of data in which it is tried to select a line with higher confidence interval (Mitchell, 2015).

Evaluation of sentiment analysis model

In order to evaluate the model, we have used the confusion matrix. This matrix shows the performance of the related algorithms. Table 2 illustrates confusion matrix based on which the validation is done (Hossin & Sulaiman, 2015). Table 3 shows confusion matrix of the classifier algorithms.

Table 3. Confusion Matrix of the Classifier Algorithms

		Predicted Class	
Actual Class		**Positive**	**Negative**
	Positive	True Positive (TP)	False Negative (FN)
	Negative	False Positive (FP)	True Negative (TN)

True positive (TP): some data that have been truly identified as the positive group.
False positive (FP): some data that have been falsely identified as the positive group.
False Negative (FN): some data that have been falsely identified as the negative group.
True Negative (TN): some data that have been truly identified as the negative group.

The value of the matrix is always between 0 and 1; as the value gets closer to 1, the performance of the recommended method is better (Sokolova & Lapalme, 2009).

Evaluation criteria

(5) $$\frac{TP}{FN+TP} = \text{Recall}$$

(6) $$\frac{TP}{TP+FP} = \text{Precision}$$

(7) $$\frac{2*Recall*Precision}{Recall+Precision} = \text{measure F}$$

Measures for evaluating classifier algorithms:

(8) $$\frac{TN+TP}{TN+FN+TP+FP} = \text{Accuracy Rate}$$

(9) $$\frac{\boldsymbol{FN+FP}}{\boldsymbol{TN+FN+TP+FP}} = 1 - \text{Accuracy Rate} = \text{Error Rate}$$

One of the methods for the evaluation of classification algorithms is the K-Fold Cross Validation. In order to reduce modeling errors, the K-Fold Cross Validation is used to separate the training and testing data. In this method, all data is divided into “k” equal parts. Out of this k sub-data, each time is taken for testing and the k-1 remaining sub-data is used for training. As a result, this process is repeated k times, in a way that each of the k parts is used for testing only once and each time one accuracy rate is calculated for the constructed model. Finally, the average score of these k times testing is taken as the final estimation (Jung & Hu, 2015) .The best value considered for k in scientific reports is 10 (P. Bradford & E. Brodley, 2001) Based on the previous studies, the value of k was set to 10 in the study at hand, the model was run 10 times, and the average score was taken as the accuracy rate.

Sentiment analysis

After evaluating the model, the sentiment analysis is applied to seven thousand reviews collected from Iranhotel website, and the opinion of every user about every hotel is made clear.

Results

After the pre-processing of the data and changing the unstructured data to the structured data, words and terms were extracted for modeling. In this stage, Linear Support Vector Classifier (Linear SVC) and Multinomial Naive Bayes (MNB) were used for modeling, and the classifier with higher precision was adopted as the selected model for sentiment analysis. As the content of the following table shows, the Linear SVC with a precision rate of 0.865 was selected for sentiment analysis. In Table 3, the evaluation results for each of the classifier are presented. Table 4 shows the evaluation of classifier algorithms

Table 4. Evaluation of Classifier Algorithms

Classifier	**Error Rate**	**Accuracy-Rate**	**F Measure**	**Recall**	**Precision**
Linear Support Vector Classification (Linear SVC)	0.135	0.865	0.905	0.893	0.917
Multinomial Naive Bayes (MNB)	0.141	0.859	0.89	0.873	0.911

In the next step, the selected model was applied to seven thousand reviews of the hotel guests in the five cities of Mashhad, Isfahan, Shiraz, and Yazd. All in all, the users' reviews about 32 hotels of these five cities have been collected. A sample of the analysis applied to this data is provided in Table 4. As this table shows, the sentiment analysis presents the implicit reviews of the users about the hotels that have been obtained through the analysis of the textual reviews. The value 0 represents negative review and the value 1 shows the positive review by the users. The user rating column illustrates the explicit reviews of the users that have been designated by the users through the use of scores 1, 2, 3, 4, or 5. Table 5 shows a sample of sentiment analysis.

Table 5. A Sample of the Sentiment Analysis of Users' Reviews about the Hotels

User's review	**Sentiment analysis of the review (positive or negative)**	**User's rating of the hotel (out of 5)**
Reception did not do their job well enough	0 (Negative)	4
Everything was good except for the photography team who did not don't treat the guests properly	1 (Positive)	4
Everything was perfect, the only problem was that the rooms were slightly cold	0 (Negative)	5
Everything was good	1 (Positive)	4
Despite the pictures on the internet, the quality was not good enough and the food wasn't good either	0 (Negative)	1

In this stage, a particular item is recommended to the user based on sentiment analysis. The main factors of the recommender systems are Users (U), Items (I), and the interactions between them. This interaction is shown by the matrix F (U, I) in which each cell includes data about the interaction. For instance, every cell shows the bought/visited or rated item. The size of the matrix depends on the number of users and items.

Data analysis revealed that in these sets of data, every user has reviewed only one hotel. This has brought about a very sparse ratings matrix. In order to overcome this problem, a combination of grouping recommender system and deep learning architecture was used to provide the recommendation. The grouping recommender system recommends the items to a group of users based on their preferences (Dara, Chowdary, & Kumar, 2019) .

In this study, the combination of "source city – travel type" is selected as the group of users, and the recommendation is made to all users who belong to the same source city-travel type group. For instance, certain suggestions are made to all members of the group "Mashhad-Family travel" based on their preferences. Moreover, we mapped every user and item on some separate vector space. The length of the vectors is fixed. In order to attain this purpose, the Singular Value Decomposition (SVD) of the F matrix is used. These are vectors in which SVD embedding is comprised of users and items. In this method, users and items are embedded separately. In this stage, the item-user matrix is constructed using the SVD of the user and item dense embedding.

In this study, a recommender system evaluation is done based on the three datasets, i.e. the user ratings of the hotels (a number from 1 to 5), the score resulting from the sentiment analysis of the user's review (a number between -1 and 1), and the combination of the user's rating and sentiment analysis score. In order to combine the users' ratings and the score resulting from sentiment analysis, the latter was normalized and rewritten in a scale of 0 to 5. Then, to combine the two sets of scores, the averaging method was used.

In this system, the negative score means that the user does not like that item. Therefore, this item is omitted from the list of recommendations to that user. In this stage, the precision of the results of the recommender system are compared using RMSE criterion. Figure 5 shows the RMSE rate.

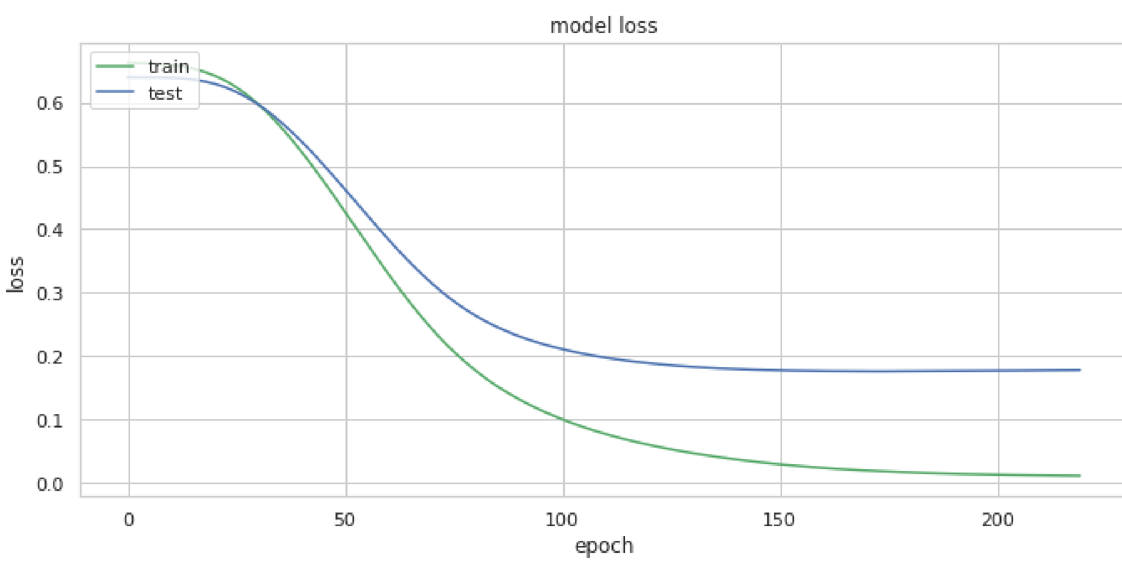

Figure 5. MSE Rate When the Combination of User Rating and Sentiment Analysis Is Used

Table 6 shows the evaluation of the recommender system.

Table 6. Evaluation the Recommender System based on Three Types of Data

Used data/criterion	Epoch	RMSE (test data)	RMSE (train data)
User ratings	183	0.2195	0.0255
Scores resulting from sentiment analysis	180	0.2717	0.0295
Combination of user rating and sentiment analysis score	175	0.1760	0.0184

Table 6 shows that the use of sentiment analysis has positive effects on the performance of recommender system and reduces MSE to a considerable degree.

Conclusions

In this article, we combined the matrix factorization method as a deep learning approach with sentiment analysis in order to implement a recommender system. Such an approach turns the user preferences – which are expressed using natural, unstructured language texts – to numerical scales that are comprehensible by collaborative filtering algorithms. The study at hand provides a step by step review of the suggested framework. Moreover, it discusses the use of sentiment analysis in the suggested system. The purpose of such an analysis is to understand how the sentiment analysis advantages can be used in the suggested framework.

As Table 5 reveals, the textual reviews show the guest' real sentiments and implicit opinions about the hotels. However, the rating that user chooses from the available numbers shows the guests' explicit opinion and might be incongruent with the reality. This can originate from the users' lack of true understanding of the ratings and shows that the ratings cannot be considered as an appropriate basis for the extraction of the guests' true opinions and the designing of the recommender systems. The results of sentiment analysis can be regarded as an appropriate input for the recommender systems. We conclude that words entail sentiments and advantages and their analysis can have significant effects on the improvement of online businesses. As the item-user matrix is sparse, this study has used a grouping recommender system and deep learning to design the system. The designed system works well in the domain of hotel recommendation, but can be applied to other domains in which the users' rating information and users' reviews are available. In the study, the collaborative recommender system is used for system designing. It is suggested to future researchers to use the hybrid approach to design the system. Moreover, the future systems might take into account the time factor in order to introduce more precision into the recommendations.

References

Abdul Hassan, A., & Abdulwahhab, A. (2019). Location aspect based sentiment analyzer for hotel recommender. *Iraqi Journal of Science, 60*(1), 143-156.

Adomavicius, G., & Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge & Data Engineering*, 734-749.

Agarwal, B., Mittal, N., Bansal, P., & Garg, S. (2015). Sentiment analysis using common-sense and context information. *Computational intelligence and neuroscience*, 30.

Alahmadi, D., & Zeng, X.-J. (2015). ISTS: Implicit social trust and sentiment based approach to recommender systems. *Expert Systems With Applications*, 8840–8849.

Appel, R., Fuchs, T., Doll´ar, P., & Perona, P. (2013). Quickly boosting decision trees - Pruning underachieving features early. *Proceedings of the 30th International Conference on Machine Learning, Atlanta.* Georgia: JMLR:W&CP.

Bhavana, P., Kumar, V., & Padmanabhan, V. (2019). Block based singular value decomposition approach to matrix factorization for recommender systems. *arXiv preprint, arXiv:1907.07410*.

Bokde, D., Girase, S., & Mukhopadhyay, D. (2015). Matrix factorization model in collaborative filtering algorithms: A survey. *Procedia Computer Science, 49*, 136-146.

Catalin-Mihai, B., & Ziegler, J. (2018). Designing interactive visualizations of personalized review data for a hotel recommender system. *RecTour* .

Chakraborty, R. (2013). Domain keyword extraction technique: a new weighting method based on frequency analysis. *Computer Science & Information Technology*, 109-118.

Chenni, O., Bouda, Y., Benachour, H., & Zakaria, C. (2015). A content-based recommendation approach using semantic user profile in e-recruitment. *International Publishing Switzerland*, 23-32.

Cremonesi, P., Turrin, R., Lentini, E., & Matteucci, M. (2008, November). An evaluation methodology for collaborative recommender systems. In 2008 International Conference on Automated Solutions for Cross Media Content and Multi-Channel Distribution (pp. 224-231). IEEE.

Dara, S., Chowdary, C., & Kumar, C. (2019). A survey on group recommender systems. *Journal of Intelligent Information Systems*, 1-25.

Das, D., Sahoo, L., & Datta, S. (2017). A survey on recommendation system. *International Journal of Computer Applications*, 0975 – 8887.

Divyashree, N., Santhosh Kumar, K., & Majumdar, J. (2017). Opinion mining and sentimental analysis of tripadvisor.in for hotel reviews. *International Research Journal of Engineering and Technology (IRJET)*, 1462-1467.

Ebadi, A., & Krzyzak, A. (2016). A hybrid multi-criteria hotel recommender system using explicit and implicit feedbacks. *International Journal of Computer and Information Engineering, 10*(8), 1377- 1385.

Ertek, G., Tapucu, D., & Arın, I. (2012). *Text mining with rapidminer.* USA: John Wiley & Sons.

Gavalas, D., & Kenteris, M. (2011). A web-based pervasive recommendation system for mobile tourist guides. *Personal and Ubiquitous Computing, 15*(7), 759-770.

Geetha, M., Singha, P., & Sinha, S. (2017). Relationship between customer sentiment and online customer ratings for hotels - An empirical analysis. *Tourism Management*, 43-54.

Hernández-Rubio, M., Cantador, I., & Bellogín, A. (2019). A comparative analysis of recommender systems based on item aspect opinions extracted from user reviews. *User Modeling and User-Adapted Interaction*, 381–441.

Hossin, M., & Sulaiman, M. (2015). A review on evaluation metrics for data classification evaluations. *International Journal of Data Mining & Knowledge Management Process (IJDKP), 5*(2). doi:10.5121/ijdkp.2015.5201

Hou, Z., Cui, F., Meng, Y., Lian, T., & Yu, C. (2019). Opinion mining from online travel reviews: A comparative analysis of Chinese major OTAs using semantic association analysis. *Tourism Management, 74*, 276-289.

Huang, Z., Chen, H., & Zeng, D. (2004). Applying associative retrieval techniques to alleviate the sparsity problem in collaborative filtering. *ACM Transactions on Information Systems (TOIS), 22*(1), 116-142.

Hussein, D.-D. (2018). A survey on sentiment analysis challenges. *Journal of King Saud University-Engineering Sciences*, 330-338.

Jung, Y., & Hu, J. (2015). A K-fold averaging cross-validation procedure. *Journal of Nonparametric Statistics*, 167-179. doi:10.1080/10485252.2015.1010532

Kannan, S., & Gurusamy, V. (2014). Preprocessing Techniques for Text Mining.

Leimstoll, U., & Stormer, H. (2007). Collaborative recommender systems for online shops. *Americas Conference on Information Systems (AMCIS).* Retrieved May 10, 2019, from http://aisel.aisnet.org/cgi/viewcontent.cgi?article=1666&context=amcis2007.

Levi, A., Mokryn, O., Diot, C., & Taft, N. (2012). Finding a needle in a haystack of reviews: cold start context-based hotel recommender system. *In Proceedings of the sixth ACM conference on Recommender systems* (pp. 15-122). ACM.

Liu, B. (2012). *Sentiment analysis and opinion mining (synthesis lectures on human language technologies).* Williston: Morgan & Claypool Publishers.

Liu, B. (2015). *Sentiment analysis: Mining opinions, sentiments, and emotions.* Cambridge University Press.

Manning, C., Raghavan, P., & Schütze, H. (2009). *An introduction to information retrieval.* Cambridge: Cambridge University Press.

María, H. (2017). *Aspect-based sentiment analysis and item recommendation.* Universidad Autónoma De Madrid, Madrid.

Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, 5(4), 1093-1113.

Mitchell, T. (2015). Generative and discriminative classifiers: Naive Bayes and logistic regression. In T. Mitchell, *Machine Learning.*

Mnih, A., & Salakhutdinov, R. (2008). Probabilistic matrix factorization. *Advances in neural information processing systems*, 1257–1264.

Musto, C., Lops, P., Gemmis, M., & Semeraro, G. (2019). Justifying recommendations through aspect-based sentiment analysis of users reviews. *UMAP '19 Proceedings of the 27th ACM Conference on User Modeling, Adaptation and Personalization*, 4-12.

P. Bradford, J., & E. Brodley, C. (2001). The effect of instance-space partition on significance. *Machine Learning*, 269-286.

Pandey, A., Rajpoor, D., & Saraswat, M. (2017). Twitter sentiment analysis using hybrid cuckoo search method. *Information Processing and Management*, 764–779.

Ravi, K., & Ravi, V. (2015). A survey on opinion mining and sentiment analysis: Tasks, approaches and applications. *Knowledge-Based Systems*, 9603-9611.

Ravi, L., & Vairavasundaram, S. (2016). A collaborative location based travel recommendation system through enhanced rating prediction for the group of users. *Computational Intelligence and Neuroscience*. doi: http://dx.doi.org/10.1155/2016/1291358

Ribeiro, F., Araújo, M., Gonçalves, P., Benevenuto, F., & Gonçalves, M. (2015). A benchmark comparison of state-of-the-practice sentiment analysis methods. *arXiv preprint, arXiv:1512.01818.*

Saga, R., Hayashi, Y., & Tsuji, H. (2008). Hotel recommender system based on user's preference transition. *2008 IEEE International Conference on Systems Man and Cybernetics.* IEEE.

Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2000). Application of dimensionality reduction in recommender system-a case study: DTIC Document. *Technology Report*.

Schafer, J., Frankowski, D., Herlocker, J., & Sen, S. (2007). Collaborative filtering recommender systems. *In The adaptive web*, 291-324.

Sharma, Y., Bhatt, J., & Magon, R. (2015). A multi-criteria review-based hotel recommendation system. *In 2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing* (pp. 687-691). IEEE.

Silamai, N., Khamchuen, N., & Phithakkitnukoon, S. (2017). TripRec: trip plan recommendation system that enhances hotel services. *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers.* ACM.

Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing and Management*, 427-437.

Su, X., & Khoshgoftaar, T. (2009). A survey of collaborative filtering techniques. *Advances in artificial intelligence* .

Sulthana, A., & Ramasamy, S. (2019). Ontology and context based recommendation system using neuro-fuzzy classification. *Computers and Electrical Engineering*, 498-510.

Sun, L., Guo, J., & Zhu, Y. (2018). Applying uncertainty theory into the restaurant recommender system based on sentiment analysis of online Chinese reviews. *World Wide Web*, 83-100.

Takuma, K., Yamamoto, J., Kamei, S., & Fujita, S. (2016). A hotel recommendation system based on reviews: What do you attach importance to? *2016 Fourth International Symposium on Computing and Networking (CANDAR).* IEEE.

Vijayarani, S., Ilamathi, M., & Nithya, M. (2015). Preprocessing techniques for text mining - An overview. *International Journal of Computer Science & Communication Networks*, 7-16.

Waltman, L., & van Eck, N. (2017). *VOSviewer Manual.* Retrieved May 15, 2019, from Bhttps://www.vosviewer.com/documentation/Manual_VOSviewer_1.6.6.pdf

Wang, Z. (2017). *The evaluation of ensemble sentiment classification approach on airline services using twitter.* Dublin: Dublin Institute of Technology.

Wei, J., He, J., Chen, K., Zhou, Y., & Tang, Z. (2016). Collaborative Filtering and Deep Learning Based Recommendation. *Expert Systems With Applications*, 29-39.

Xiao, T., & Shen, H. (2019). Neural variational matrix factorization for collaborative filtering in recommendation system. *Applied Intelligence*, 49(10), 3558–3569. doi:https://doi.org/10.1007/s10489-019-01469-6

Zamani, M., Sorkhpour, B. (2014). Data mining of Persian texts with knowledge management attitude. *7th National Conference and First International Knowledge Management Conference*, Tehran.

Zhang, K., Wang, K., Wang, X., Jin, C., & Zhou, A. (2015). Hotel recommendation based on user preference analysis. *2015 31st IEEE International Conference on Data Engineering Workshops.* Seoul, South Korea: IEEE.

Zhang, L., Luo, T., Zhang, F., & Wu, Y. (2018). A recommendation model based on deep neural network. *IEEE Access*, 9454-9463.

Zhao, L., Huang, M., Yao, Z., Su, R., Jiang, Y., & Zhu, X. (2016). Semi-supervised multinomial naive bayes for text classification by leveraging word-level statistical constraint. *Proceeding of the Thirtieth AAAI Conference on Artificial Intelligence*, (pp. 2877-2883).

Zheng, X., Luo, Y., Sun, L., Zhang, J., & Chen, F. (2018). A tourism destination recommender system using users' sentiment and temporal dynamics. *Journal of Intelligent Information Systems, 51*(3), 557–578.

Bibliographic information of this paper for citing:

Abbasi, Fatemeh; Khadivar, Ameneh, & Yazdinejad, Mohsen (2019). A grouping hotel recommender system based on deep learning and sentiment analysis. *Journal of Information Technology Management*, 11(2), 59-78.