

PAPER • OPEN ACCESS

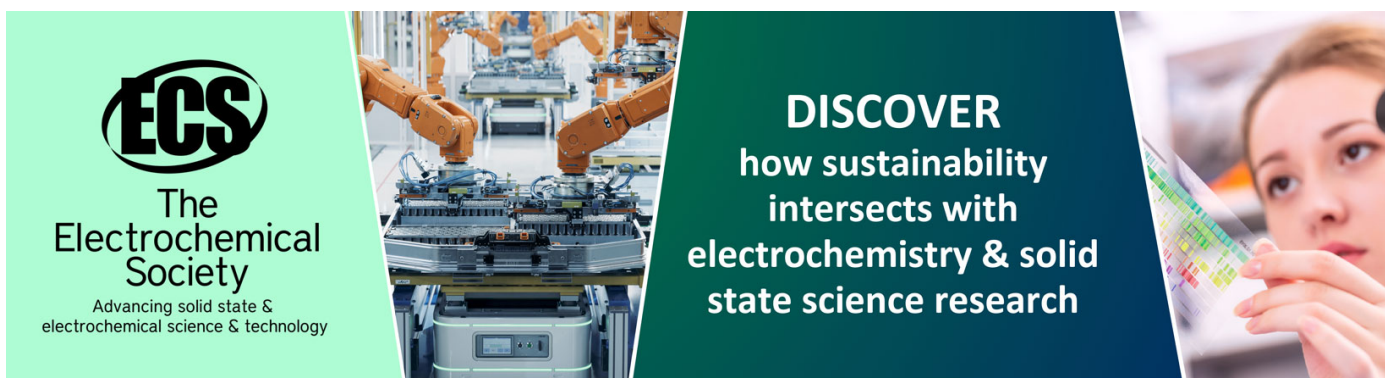
Hotel Content-Based Recommendation System

To cite this article: Kristian Wahyudi *et al* 2020 *J. Phys.: Conf. Ser.* **1485** 012017

View the [article online](#) for updates and enhancements.

You may also like

- [Research on Key Technologies of Personalized Education Resource Recommendation System Based on Big Data Environment](#)
Yunli Cheng and Xu Bu
- [Online Book Recommendation System using Collaborative Filtering \(With Jaccard Similarity\)](#)
Avi Rana and K. Deeba
- [IOT-Based Personalized products recommendation system](#)
Shili Mohamed, Kaouthar Sethom and Ahmed J. Obaid



Hotel Content-Based Recommendation System

**Kristian Wahyudi ,Johanes Latupapua, Ritchie Chandra, Abba Suganda
Girsang**

Computer Science Department, BINUS Graduate Program-Master of Computer
Science, Bina Nusantara University, Jakarta, Indonesia 11480.

Email : kristian.wahyudi@binus.ac.id, johanes.latupapua@binus.ac.id,
ritchie.chandra@binus.ac.id, agirsang@binus.edu

Abstract. Content based recommendation system tries to recommend items similar to those a given user has likely in the past, whereas systems designed according to the collaborative recommendation paradigm identify users whose preferences are similar to those of the given user and recommend items they have liked. The proposed recommendation system is discussing about Data finitihotel recommendations for a place across U.S. The selected process for the recommendation is calculating the rating of hotel categories on the city. By calculating hotel categories from several city by combining two features namely categories and taken from the city of the selected hotelcategory as well. The result of process will be recommended hotel with highest ratingto the user.

1. Introduction

Recommender system have developed in response to a manifest need: helping people deal with the world of information abundance and overload. Further it has become clear that they can link people with other people who share their interests, not just with relevant information [1]. Recommender systems are utilized in a variety of areas and are most commonly recognized as playlist generators for video and music services like Netflix, YouTube and Spotify, product recommenders for services such as Amazon, or content recommenders for social media platforms such as Facebook and Twitter. Current needs for content-based recommendation system have been increasing. Here the system has function to generate content recommendation according to database features.

The amount and varies type of information circling around internet, has proven to be a steep challenge to determine what user desire, especially that fit user needs. Content-based recommendation systems try to recommend items similar to those a given user has likely in the past, whereas systems designed according to the collaborative recommendation paradigm identify users whose preferences are similar to those of the given user and recommend items they have liked [2].

Research about recommendation hotel have been conducted before such as [6][7][8]. Saga et al [6] proposed the recommendation can be gained from the sales record as it contains user preference relation among hotels. Zhang et al [7] also used the Content-based hotel recommendation but combined with collaborative filtering. Zhao et al [8] also used the combined with collaborative filtering to get hotel recommendation.



Our content-based recommendations is about Datafiniti hotel recommendations for a place in the U.S. The dataset used on this case was obtained from Kaggle. Referring to only one table presented, the data will be selected that contained the rating of a place. Our dataset has 10000 records x 26 columns where the data to be used as a measurement reference is the hotel name, city, hotel categories, and rating.

2. Related Work

During last years, researches in recommendation system have been increasing. Wang et. al [3] propose an efficient hybrid model based on chi-square feature selection and softmax regression. It is a recommender system on computer science publications referred to as the Publication Recommender System (PRS). This system is based on a new content-based filtering (CBF) recommendation model using chi-square and softmax regression which are combined to construct a real-time online system. Their test results show that, the system can achieve an accuracy of 61.37%.

Madhushree, B [1] propose a recommendation system, combining user and item information, namely User Based Recommender which uses preferences of their other similar users for recommendation of research paper. Item Based Recommender uses User's Item visit history for recommendation. Four algorithms are implemented in each category by using standard dataset. A recommender that could predict all preferences exactly would present all items ranked by researcher. The, to evaluate precision and recall which are two popular decision support metrics from information retrieval. In this recommender system, they judge relevance by counting recommendation hits: if a withheld item appears in a top-n recommendation list [4][5]. These measures are also done per user and averaged over all users.

3. Proposed Method

This system offers for users the hotels which are not rated or stayed. The calculation from similarity of two user based on two features, category and city. First, the dataset is cleaned as it might contains null value or duplicate data. Then, the dataset is pre-processed, especially the category and city feature. Two features are changed to be "one hot encoding" to be able to calculate similarity between two users. This similarity feature is generated from the hotel which users already stayed. This similarity indicates the preference of users. After gaining the similarity, the hotel with score will be offered to users. Generally, the offered hotels are offered sorting by the highest to lowest rating. This system is represented as an algorithm as shown Fig.1

```

procedure item-based hotel recommendation system
1:   Data Cleansing
2:   Encode Feature Category and city
3:   for all hoteli which useru has no preference do
4:       for all userv which has preferences for hoteli do
5:           Select Category and City feature
6:           Compute similarity s between useru and userv
7:       end for
8:       Add userv's preference for hoteli, weighted by s, to a running average.
9:   end for
10:  return top hotels, ranked by weighted average
endprocedure

```

Figure 1. Algorithm of Proposed Method

4. Analysis Results

4.1. Dataset

The dataset was Data finiti hotel database taken from Kaggle. The dataset consists of 10000 rows of record. Each record represents one user that gave rating and comments on a hotel that has been rented by that user and that row consist information such as, the address of the hotel, the user's username, the hotel latitude, the hotel longitude, etc. There are 1433 different hotels found in the dataset. Also, there are 9222 users found in the dataset.

4.2. Analysis Performance

To evaluate the performance, the dataset will be filtered on those 9222 users on the 10000 rows of record that has reviewed or stayed more than 2 hotels which only left 108 users. A user that has reviewed less than 2 hotels can't be computed to get the similarity to another user. From each 108 user, select 2 features, categories and city to get similarity between two users. After that, the result will be all hotels that will be recommended with its own rating. Performance will be calculated between rating score difference of recommended hotel and the hotel that user already rated in the past.

For example, choose a user, called user A. User A has rated or stayed 5 hotels in the past. Two first hotel will be chosen as an input. Select categories and city, also rating on those two records to calculate the recommendation. After obtaining the calculation, filter only on last three records to compare on the real data. Finally, calculate the rating score difference between recommended hotels and the real hotel data.

Based on test that have been done, the system will produce a recommendation and comparing to the real data to make confusion matrix as show in Fig. 2. From the confusion matrix, precision, recall, and accuracy can be obtained. The Precision value is 88,9%, accuracy is 85%, and recall is 80%.

		Actual	
		TP	P
Prediction	TP	40	5
	P	10	45

Figure 2. Confusion matrix

5. Conclusions

Content-based filtering is one of the common methods in building recommendation systems, beside collaborative and hybrid. Content based has functioned to retrieve features from the content of item, sometimes they are words which describe the item. This method is best suited to situations where there is known data on an item (name, location, description, etc.), but not on the user. Content-based recommenders treat recommendation as a user-specific classification problem and learn a classifier for the user's likes and dislikes based on product features.

In our research using U.S Data finiti data, every new review from user will impact and update rating for one location. Content based runs and generate new ratings.

References

- [1] B. Madhushree, "A Novel Research Paper Recommendation System," *International Journal of Advanced Research in Engineering and Technology*, vol. VII, no. 1, pp. 7-16, 2016.
- [2] F. Ricci, L. Rokach, B. Shapira and P. Kantor, *Recommender Systems Handbook*, New York:

Springer US, 2011.

- [3] D. Wang, Y. Liang, D. Xu, X. Feng and R. Guan, "A Content-based Recommender System for Computer Science Publications," *Elsevier Knowledge Based System*, vol. 157, pp. 1-9, 2018.
- [4] T. Joshi, "Improved Product Ranking For Recommendation Systems," *IOSR Journal of Computer Engineering*, vol. 18, no. 2, pp. 18-23, 2016.
- [5] Wikipedia, "Wikipedia Recommender System," Wikipedia, 25 August 2019. [Online]. Available: https://en.wikipedia.org/wiki/Recommender_system.
- [6] R. Saga, Y. Hayashi, and H. Tsuji, "Hotel recommender system based on user's preference transition," in *2008 IEEE International Conference on Systems, Man and Cybernetics*, 2008, pp. 2437–2442.
- [7] K. Zhang, K. Wang, X. Wang, C. Jin, and A. Zhou, "Hotel recommendation based on user preference analysis," in *2015 31st IEEE International Conference on Data Engineering Workshops*, 2015, pp. 134–138.
- [8] F. Zhao, F. Yan, H. Jin, L. T. Yang, and C. Yu, "Personalized mobile searching approach based on combining content-based filtering and collaborative filtering," *IEEE Syst. J.*, vol. 11, no. 1, pp. 324–332, 2015.