Hotel Recommendation Based on **Hybrid Model**

Keyword: Collaborative filtering , Content-based, SVD, Hierarchical clustering, Decision Tree

Jing Wang[1], Jiajun Sun[1], Zhendong Lin[1]

[1]*Deparment of Civil and Environmental Engineering*

Stanford University

Introduction

A Hybrid Model is developed for hotel recommendation. The benchmark approach in the industry relies on solely hotel popularity. This method does not provide personalized hotel recommendation. The Hybrid Model combines user preferences and hotel location specific popularity. Recommendation on returning user considers their historical booking history. While for new user, a decision tree is applied to classify them into our existing user clusters.

Hybrid Model :

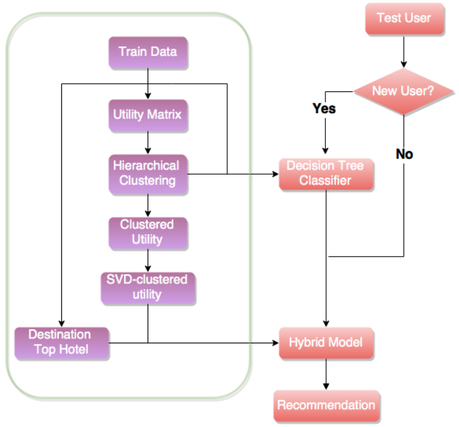
The Hybrid Model includes information on both user preference and hotel location specific popularity.

$$M^{(i)} = [m_{i1}, m_{i2}, \dots m_{ik}, \dots \mathrm{m_{in}}]$$

$$D^{(i)} = [d_{j1}, d_{j2}, \dots d_{jk}, \dots \mathrm{d_{jn}}]$$

$$R_{ik} = M_k^{(i)} D_k^{(j)}$$

M: utility matrix
D: hotel ranking matrix for different destinations
m_{ik}: rating utility matrix
d_{jk}: rating in hotel ranking matrix in destination j
i: user_id
j: destination_id
k: hotel cluster_id



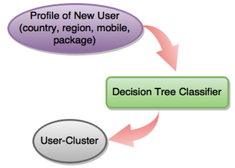
Methods

Utility Matrix:
Giving for each user-item pair, a value that represents what is known about the degree of preference of that user for that item.

Hierarchical Clustering
Begin with every points in its own cluster. Larger clusters will be constructed by combining two smaller clusters

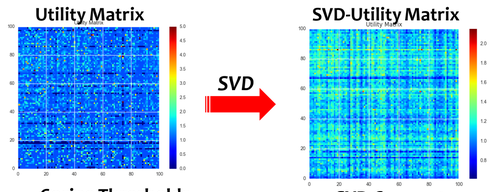
SVD(UV-decomposition):
SVD seeks a low-rank matrix $X = UV^T$ to reduce the dimension and fill the some unrated items.

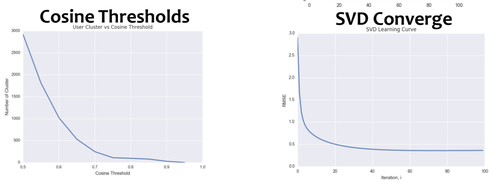
Decision Tree Classifier:.
We applied the method of decision tree to classify new users into our user clusters according to their user profiles.



Dataset Head

user_location_country	user_location_region	user_location_city	user_id	is_mobile	is_package	srch_destination_id	is_booking	hotel_cluster
66	258	5888	8798	0	0	25946	0	73
66	258	5888	8798	0	0	25946	0	73
66	258	5888	8798	0	0	25946	0	96
66	258	5888	8798	0	0	25946	0	48
66	258	5888	8798	0	0	25946	0	42

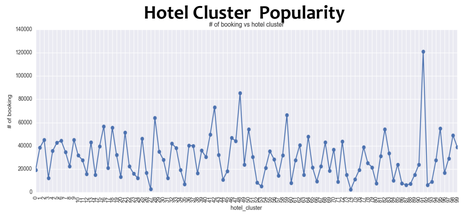




$$cosine\ similarity = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}}$$

$$\mathrm{U_i} := \mathrm{U_i} + \alpha\left(\left(\mathrm{M_{ij}} - \mathrm{V_j^T U_i}\right)\mathrm{V_j} - \lambda \mathrm{U_i}\right)$$
$$\mathrm{V_j} := \mathrm{V_j} + \alpha\left(\left(\mathrm{M_{ij}} - \mathrm{U_i^T V_j}\right)\mathrm{U_i} - \lambda \mathrm{V_j}\right)$$

Hotel Cluster Popularity

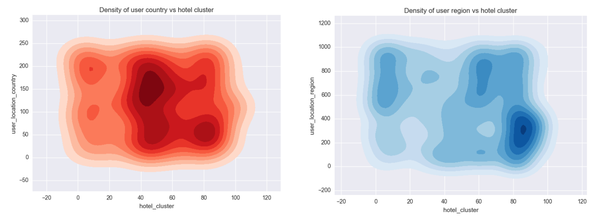


Result & Conclusion

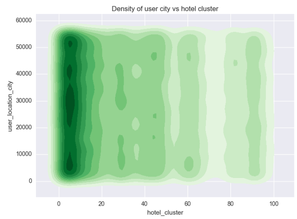
Hybrid model results in prediction with 53.6% accuracy on testing data - 4% improvement on content-base model. This result is consistent with our hypothesis: both user preference and hotel popularity are vital in recommendation system. Currently, the leading team on Kaggle hotel recommendation competition achieved 51.7% accuracy. The benchmark content-based model has 49.8% accuracy. Therefore, this Hybrid Model is very likely to get the first place in this competition.

User Profile Detection

Country and region – **CORRELATE** with users preference.

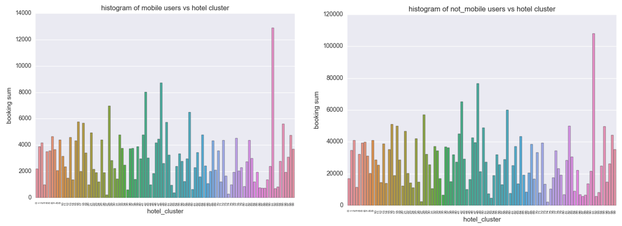


City - **LITTLE** effects on users preference.

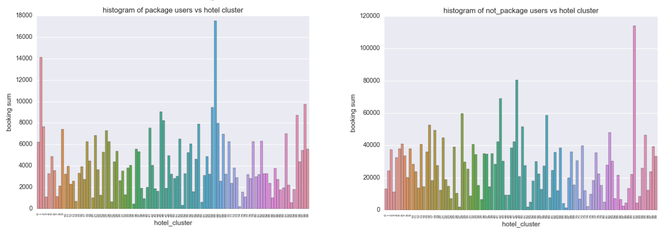


According to ontology model, user profile data can be used to detect the preference of users. Based on current data form we have, we draw several figures to show what factors in users profile affect the preference.

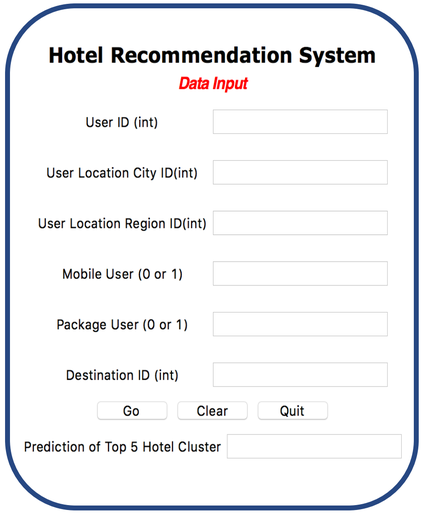
Mobile Users - **NO** obvious preference



Package Users - **OBVIOUS** preference



GUI Design



Accuracy Comparison

Accuracy	Hybrid Model	Content-based
Train	61.29%	51.03%
Test	53.62%	51.74%

We compared the outcomes of predicted top 5 hotels clusters that a coming user might book between single content-based model and hybrid model. Combined with utility matrix in collaborative filtering, the hybrid model reached more accuracy not matter in training or testing to predict the booking outcomes of users.

Reference

[1] Ba Q, Li X, Bai Z. Clustering collaborative filtering recommendation system based on svd algorithm[C]//Software Engineering and Service Science (ICSESS), 2013 4th IEEE International Conference on. IEEE, 2013: 963-967.
[2] Huming G, Weili L. A hotel recommendation system based on collaborative filtering and rankboost algorithm[C]//2010 second international conference on multimedia and information technology. IEEE, 2010: 317-320.
[3] Meng C, Cheng Y, Jiechao C, et al. A Method to Solve Cold-Start Problem in Recommendation System based on Social Network Sub-community and Ontology Decision Model[C]//3rd International Conference on Multimedia Technology (ICMT-13). Atlantis Press, 2013.
[4] Salakhutdinov R, Mnih A, Hinton G. Restricted Boltzmann machines for collaborative filtering[C]//Proceedings of the 24th international conference on Machine learning. ACM, 2007: 791-798.