International Journal of Modern Computer Science (IJMCS)
Volume 5, Issue 3, June, 2017

ISSN: 2320-7868 (Online)

Shop-Type Recommendation System

Ms. Pooja C. Sohale
Department of Computer Engineering
VPKBIET, Baramati,
Pune, India
E-mail: poojasohale@gmail.com

Mr. Santosh A. Shinde
Department of Computer Engineering
VPKBIET, Baramati,
Pune, India
E-mail: shinde.meetsan@gmail.com

Abstract: In business world, it is important to determine the appropriate type of shop (e.g. restaurant, fashion, coffee) for newly opened store. Shop type recommendation system suggests a suitable shop type using social media and location-based services (LBS) data. User generated data from social media and LBS provide rich information about individual consumption experiences and various shop attributes. Shop type recommendation is useful for making investment and urban planning decisions. This system extracts the location and commercial features from the data to describe the characteristics of shops. The feature fusion module incorporates all the features into singular value decomposition (SVD) matrix factorization model for parameter learning. This system develops a novel bias learning matrix factorization algorithm with feature fusion for shop type recommendation.

Keywords: Shop-Type, Location-based Services, Singular Value Decomposition, Matrix Factorization

I. Introduction

In today's world of technology, various types of recommendations are provided to the user. By using such recommendation techniques user's time can be saved and also performance of system can be improved. Shop-Type recommendation system is one such recommendation technique which saves users effort and time. Shop-Type recommendation system suggests an appropriate Shop-Type for newly opened store. Suppose any investor wants to establish a business in a city or a town, first of all he/she has to decide which type of shop will earn more profit. In real life, the following situation would occur: an investor has an available store space, but he does not know which type of shop is more profitable. Traditionally, the investors usually consult with professional business analytic teams. They conduct field surveys and provide suggestions according to their observed estimation. However, traditional methods cannot follow dynamics of business mobility and geographical change. As social media and LBS are widely used, the user generated data from these platforms provide rich information not only about individual consumption experiences, but also about shop characteristics. Shop information includes the shop type, location, reviews, per capita consumption, and so on. Users also give reviews about the shop whether they like it or not which provides an indicator of the popularity of real-world shops. On the other hand, LBS provide address geocoding API and traffic information, which can help us obtain information around a particular shop. Therefore, determining the most suitable shop type which pursues to maximize the number of customers served becomes feasible by leveraging user preference and location context extracted from social media and LBS. In this system heterogeneous data from social media and LBS are used to address the shop-type recommendation problem, i.e., which type of shop is suitable to establish at a given location. Specifically, when given a set of candidate types for an opened shop, systems aim is to predict the most promising type from the set in terms of attracting as many customers as possible. To achieve this goal, the collected data is required to be analyzed by extracting location and commercial features to model the shop and type preference, respectively. Then explicit preferences into the bias learning matrix factorization method are integrated for shop popularity prediction.

II. Literature Review

Nowadays, people are aware about various types of recommendation systems such as products, movies, articles recommendations and so on. Zheng et al. [4] present a survey of various recommendation techniques in a location-based social network on the basis of locations and location related contents shared by users on various social media services. They have done a comprehensive survey by analyzing the data source used, the methodology employed to generate a recommendation, and the objective of the recommendation. In personalized recommendation systems, context plays an important role. Baltrunas et al. [8] introduce a collaborative filtering method based on tensor factorization which allows integration of contextual information. In the proposed model, called Multiverse Recommendation, different types of context are used as a dimension in the representation of data as a tensor. Yu et al. [10] introduce personalized travel package recommendations to help users to make travel plans. They suggest a heuristic search-based travel route planning algorithm to generate travel packages. Ali Abbasi et al. [2] suggest item recommendations to users according to their preferences and activity history. Recommendation method is based on items location, distance between those items and utilizes a Markov-based approach which can be easily applied to implicit datasets. Baltrunas et al. [3] and Salakhutdinov et al. [9] suggest some matrix factorization techniques to generate recommendations.

Du et al. [5] attempt to achieve intelligent business recommendation by exploring location-based check-ins or reviews provided by users. Fu et al. [6] suggest a geographic ClusRanking method to evaluate the value of the estate by using the estate related mobile data to give suggestions for home buyers. Some researchers have done several work in shop recommendation systems. Karamshuk et al. [7] explore the usefulness of various features on the popularity of retail stores and exploits the features to recognize the optimal store placement using a dataset collected from Foursquare in New York. Yu et al. [11] recommend an approach to identify the optimal location by investigating the user check-in records. They solve the problem of ranking areas by popularity of a business category.

III. SYSTEM ARCHITECTURE

Fig. 1 shows the system architecture of shop type recommendation system.

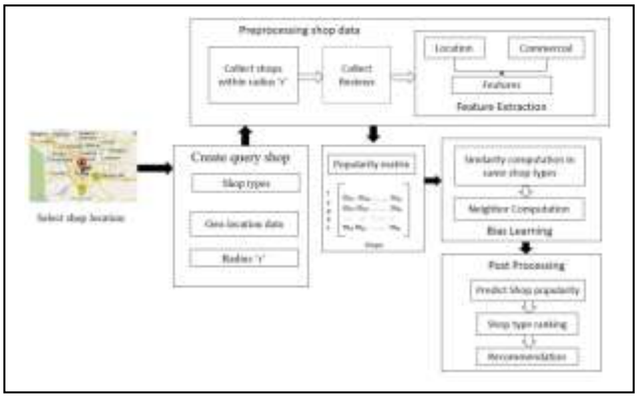

Figure 1. Architecture of shop type recommendation system

A. Feature Extraction

This system considers the relation of the features with the location and then distribute the features into two categories: location features and commercial features. Location features are closely related to spatial characteristics, and commercial features are factors extracted from relationships between the shop types in the neighborhood.

1) Location Features:

Location features are features related to the spatial and geographical characteristics of shop. The geographical information of a shop has a significant influence on its business. It includes distance to downtown, traffic accessibility, shop diversity and human density.

(a) Distance to Downtown

Downtown refers to the center of a city or the central business district. By calculating the distance to downtown, how the distance to downtown influences the shop popularity is evaluated. The distance to downtown is calculated as follows, which returns a value corresponding to the closeness of the distance to downtown:

$$D_s = \frac{1}{log d_s} \qquad (1)$$

Where, d_s = value of distance to downtown.

D_s = effect of distance to downtown on shop popularity.

(b) Traffic Accessibility

Traffic and transportation are very important characteristics of location. Traffic accessibility is one of the key concerns of consumers when they go shopping. Traffic accessibility refers to consumer's ability to reach shops, i.e., the quality and cost of vehicle options available to reach given shops. Traffic accessibility is calculated as follows:

$$TA_s = \frac{log_2(N_{bus}(s,r)+1)}{log_2(d_{bus})} + \frac{log_2(N_{sub}(s,r)+1)}{log_2(d_{sub})} \qquad (2)$$

Where,

$N_{bus}(s,r)$ = bus stop number in the region.

$N_{sub}(s,r)$ = subway number in the region.

d_{bus} = minimum distance to bus stop.

d_{sub} = minimum distance to subway station.

(c) Shop Diversity

Diversity means differences, and shop diversity implies the different types of shops in the area, which can be used to measure the areas characteristics. Shop diversity is calculated as follow:

$$D_{diver} = -\sum_{t \in T} \frac{N_t(s,r)}{N(s,r)} \times log \frac{N_t(s,r)}{N(s,r)} \qquad (3)$$

Where,

$N_t(s,r)$ = denote the number of shops, which are of the same type with the given type t.

$N(s,r)$ = the total number of shops in the bounding circle with a radius of r.

(d)Human Density

Human density refers to the visiting frequency of people in an area in a period of time. In general, when a shop is located in a crowded area, it can attract many more customers. The human density is formalized as:

$$HD_s = \log \sum_{s \in R} R(s) \qquad (4)$$

Where R(s) is the review number of the shops in the neighborhood.

2) Commercial Features:

Besides location features, several features about the relationships of the shops in the neighborhood are also extracted. Feature which shows the relationship between shops in the neighborhood are known as commercial features. Competitiveness and complementarities are two important commercial features which plays important role in shop type recommendation. While deciding the shop type of a given place, the competitiveness among shops of the same type and the complementarities among shops of different types are need to be consider.

(a) Competitiveness

Generally, different shops of the same type in one area will compete with each other. Competition between shops can benefit some of them. Given a shop of type t, its competitiveness can be measured as follows:

$$C_s = \frac{N_{ts}(s,r)}{N(s,r)} \quad (5)$$

Where,

$N_{ts}(s,r)$ = the number of shops with the same type t in the bounding circle with a radius of r and $N(s,r)$ = the total number of nearby shops.

(b) Complementarities

In general, several shops of different types co-located in one area have the potential to achieve mutually beneficial win-win cooperation through complementarities, by sharing opportunities through collaboration. It is similar to the complementarity between two products which can be bundled and reinforce the benefit of each other, e.g., coffee shop and bread shop. Given a shop of type t, its complementarities are measured as follows:

$$\rho_{t\to t^*} = \frac{2N_{set}(t,t^*)}{N_T(N_T-1)} \quad (6)$$

Where,

$N_{set}(t,t^*)$ = the appearance number of the pair-wise type t and t^*, N_T = the number of all shop types, $N_T(N_T-1)/2$ = the total combination number of all the shop types and $\rho_{t\to t^*}$ = indicates the appearance possibility of shop type pair t and t^* combined number in all combinations.

B. Shop Type Recommendation

In this segment, how exactly shop type recommendation system will recommend a particular shop type for a newly opened store is described. The algorithm for shop-type recommendation is suggested. It presents the utilization of matrix factorization method for popularity prediction. Then it combines the extracted features into the matrix factorization method.

(1) Matrix Factorization

Matrix factorization based methods have been widely used to model the user-item relationship for recommendations [2]. In a resemblance with user-item recommendation, matrix factorization method is adopted for shop-type recommendation. SVD i.e. singular value decomposition [8] is one of the matrix factorization methods that can be applied in this shop type recommendation system. Singular value decomposition is a type of matrix factorization method which results in 1× 1 matrix. The popularity P_{st} can be factorized with S_s and T_t, where S_s and T_t are vectors of dimension k, which represent the shop latent features (i.e., the preference of shop s in the latent space) and the type latent features (the preference of type t in the latent space).

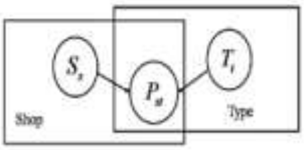

Figure 2. Matrix Factorization [1]

As shown in Figure 2, the popularity P_{st} is decided by the user-mediated interaction between the shop and the type. Specifically, this model can be learned by estimating the value of S_s and T_t. And the predicted popularity, denoted by P_{st}, is defined as:

$$P_{st} = S_s^T T_t$$

(2) *Feature Fusion Matrix Factorization:*

Location features and commercial features are integrated into SVD based on the basic matrix factorization, which are denoted as G-CoSVD. Shop preference and type influence of a given location are influenced by location features and commercial features separately. G denotes location features and Co denotes commercial features.

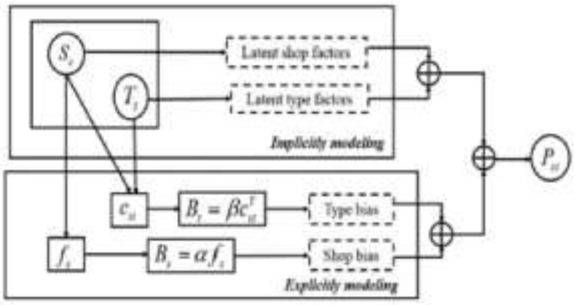

Figure 3. Feature Fusion Matrix Factorization

Figure 3 shows the fusion of location and commercial feature with matrix factorization. In this system G-CoSVD is used to co-factorize the shop popularity by modeling the

$$P_{st} = u + B_S + B_t + S_S^T T_t \quad (7)$$

shop and type bias, based on which the popularity of a shop can be predicted as follows:

Where u is a constant and tends to be 0.5, for when a new shop opens in a specific location, the probability of whether it will be popular or not is about 50%. B_s and B_t indicate the shop and type effects, respectively, which are calculated by location features and commercial features. In this way, location features f_s=<D_s,TA_s,D_{diver},HD_s> and commercial features c_{st}=<C_s,$\rho_{t\to t}$*> integrated into the popularity prediction.

Table 1 specifies the mathematical notations used in feature fusion matrix factorization. Specifically, in G-CoSVD, the bias for each shop $b_s \in B_s$ is defined as $b_s = \alpha f_s^T$, and the bias for each type $b_t \in B_t$ is defined as $b_t = \beta c^T_{st}$, based on which the above formula can be transformed as follows:

$$p_{st} = u + \sum \alpha_s f_s + \sum \beta_t c_{st} + S_S^T T_t \quad (8)$$

Table 1. Mathematical Notations

Symbol	Dimension	Description
S	I × N	Shop Latent Feature vector
T	I × M	Type Latent Feature vector
B_s	N × M	Shop bias
B_t	N × M	Type bias
f_s	1 × K	Location Feature
c_{st}	1 × P	Commercial Feature
A	1 × K	Parameter of location feature
B	1 × P	Parameter of commercial feature

This Model can be learned by estimating the values in ∧(S, T, α, β). After obtaining all the parameters in ∧, when given a specific type n, the shop popularity can be predicted accordingly. By using this predicted popularity for each shop type, recommendation results are generated. After predicting the shop popularities of each type, the candidate shop types are needed to rank by popularity in descending order and then recommend the top N shop types to the location. In this shop type recommendation system, shop types having popularity greater than 5 will be considered as the top recommendation result.

IV. Result Analysis

In shop-type recommendation system, the shop related information and geographic information is collected by using Google Business API and Google Map API. When user selects any location from map, its latitude and longitude values are fetched. Then by entering certain radius value, shops information within that circular area is collected. Shop information includes name of shop, type, address, reviews and ratings given by various users to those shops. Table 2 shows the shop information collected by using Google APIs.

Table 2. Shop Information

Sr. No.	Name	Type	Address	Review	Rating
1	Hotel Basera	Lodging	Bajirao Road, Pune	5	3.7
2	Bata	Shoe store	Chintamani kelkar road, Pune	1	4
3	Manish Clinic	Health	Kasaba Peth, Pune	0	0
4	Siddharth Hotel	Restaurant	Shivaji Nagar, Pune	5	4.7
5	Sardar Cycle	Bicycle store	Near Phadke Haud, Pune	5	3.6

In shop-type recommendation system, after collecting shop related information, sentiment analysis is done by using collected reviews and then location features and commercial features are extracted. Table 3 shows the feature values extracted for each shop.

Table 3. Feature Extraction

Name	Rating	Sentiment	Down Town Distance	Traffic Accessibility	Shop Diversity	Human Density	Competitiveness	Complementarities
Hotel Basera	3.7	1	1.54 km	2.14	0.13	0.21	0.1	0
Bata	4	1	1.84 km	1.59	0.2	0.04	0.15	0
Manish Clinic	0	0	1.59 km	1.82	0	0	0	0.05
Siddharth Hotel	4.7	-1	1.43 km	0.77	0.13	0.21	0.1	0
Sardar Cycle	3.6	-1	1.47 km	1.2	0.2	0.21	0.15	0

As shown in table 2 and table 3 shops information is collected and feature values are calculated for each shops within the circular area of given location. By using this collected data, type-feature matrix is formulated. Based upon this type-feature matrix popularity for each type of shop is calculated. Shop-type with popularity greater than 5.0 is recommended to the user of this system. Table 4 shows the type-feature matrix values with popularity calculation.

Table 4. Type-Feature Matrix with popularity calculation

Type \ Feature	Rating	Sentiment	Down Town Distance	Traffic Accessibility	Shop Diversity	Human Density	Competitiveness	Complementarities	Popularity
Lodging	2.8	1	1.64	1.99	0.13	0.1	0	0	6.79
Health	0	0	1.59	1.82	0	0	0	0.05	1.46
Shoe store	4	1	1.84	1.59	0.2	0.04	0.15	0	7.71
Bicycle Store	3.6	-1	1.47	1.2	0.2	0.21	0.15	0	2.46

Popularity value for each shop type is calculated by considering the sentiment analysis value as matrix factorization factor. Shop-type with popularity greater than 5 will be recommended to user. As shown in table 4, Lodging and Shoe-store are the shop-types recommended to user of this shop-type recommendation system.

1) Efficiency measurement: To measure the efficiency of a shop type recommendation system, precision and recall values are required to be found out.

Table 5. Result Analysis with Precision and Recall

Location	#Type	#Shops	#Recommendation	#Correctly Predicted Recommendation	Precision	Recall
Shivaji Nagar, Pune	7	16	5	4	0.8	0.71
Deccan Gymkhana, Pune	6	14	4	3	0.75	0.67
Koregaon Park, Pune	7	18	6	4	0.67	0.85
Kothrud, Pune	11	15	7	5	0.71	0.7
Viman Nagar, Pune	8	12	6	4	0.67	0.75

In information retrieval systems, precision is the fraction of retrieved instances that are relevant and recall is the fraction of relevant instances that are retrieved. For a given system, the precision and recall values are calculated by using:

$$Precision = \frac{\#CorrectlyPredictedRecommendations}{\#Recommendations}$$

$$Recall = \frac{\#Recommendation}{\#Types}$$

Table 5 shows the results analysis with precision and recall calculation. Figure 4 shows the graph of precision versus recall.

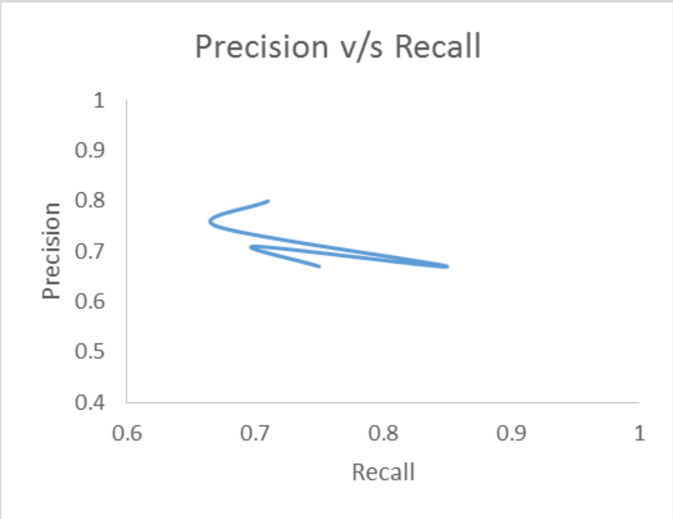

Figure 4. Precision v/s Recall Graph

V. Conclusion

In given shop type recommendation system, an appropriate shop type for a newly opened store is recommended. A bias learning feature fusion matrix factorization algorithm is suggested to predict shop popularity, based on which the optimal type for a given shop is recommended. When an investor plans to open a new store in the location s and if there are n candidate types, then shop-type recommendation can be done in following way: The location features f_s and the commercial features C_{st} of each candidate type are required to be captured first. For a candidate type t, we input the set of feature values to the feature fusion matrix factorization model. Using the bias learning matrix factorization with feature fusion algorithm, the initial matrix values are recalculated. The value of the given type t and shop s in the matrix is the predicted popularity P_{st}. In this way, for each shop type, firstly, predict its popularity in the given location. Afterward, the candidate shop types are ranked by popularity in descending order and recommend the top N shop types to the location. This shop type recommendation system is useful for making investment decisions and urban planning.

Acknowledgment

This is to acknowledge and thank all the individuals who played defining role in shaping this work. I avail this opportunity to express my deep sense of gratitude and whole hearted thanks to Prof. Santosh A. Shinde for giving his valuable guidance, inspiration and encouragement to complete this task.

References

[1] Zhiwen Yu, Miao Tian, Zhu Wang, And Bin Guo, “Shop-Type Recommendation Leveraging the Data from Social Media and Location-Based Services,” in ACM Transactions on Knowledge Discovery from Data,Vol. 11, No. 1, Article 1, Publication date: July 2016.

[2] Ali Abbasi, Amin Javari, Mahdi Jalili, and Hamid R. Rabiee, “Enhancing precision of Markov-based recommenders using location information,” in Proceedings of the International Conference on Advances in Computing, Communications and Informatics (ICACCI), IEEE 2014.

[3] Linas Baltrunas, Bernd Ludwig and Francesco Ricci “Matrix factorization techniques for context aware recommendation,” in Proceedings of ACM Conference on Recommender Systems, pages 301304, 2011.

[4] Jie Bao, Yu Zheng, David Wilkie, and Mohamed Mokbel “Recommendations in location-based social networks: A survey,” in GeoInformatica, 525565, 2015.

[5] Rong Du, Zhiwen Yu, Tao Mei, Zhitao Wang, Zhu Wang, and Bin Guo “Predicting activity attendance in event-based social networks: Content, context and social influence,” in Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing, 2014.

[6] Yanjie Fu, Hui Xiong, Yong Ge, Zijun Yao, Yu Zheng, and Zhi-Hua Zhou “Exploiting geographic dependencies for real estate appraisal: A mutual perspective of ranking and clustering,” in Proceedings of ACM SIGKDD, 10471056, 2014.

[7] Dmytro Karamshuk, Anastasios Noulas, Salvatore Scellato, Vincenzo Nicosia, and Cecilia Mascolo “Geo-spotting: Mining online locationbased services for optimal retail store placement,” in Proceedings of ACM SIGKDD, 793801, 2013.

[8] Alexandros Karatzoglou, Xavier Amatriain, Linas Baltrunas, and Nuria Oliver “Multiverse recommendation: N-dimensional tensor factorization for context-aware collaborative filtering,” in Proceedings of ACM Conference on Recommender Systems, 7986, 2010.

[9] Ruslan Salakhutdinov and Andriy Mnih “Bayesian probabilistic matrix factorization using Markov chain Monte Carlo,” in Proceedings International Conference on Machine Learning. AAAI Press, 880887, 2008.

[10] Zhiwen Yu, Huang Xu, Zhe Yang, and BinGuo “Personalized travel package with multi-point-of-interest recommendation based on crowd sourced user footprints,” in IEEE Trans. Human-Mach. Syst. 46, 2016.

 ISSN: 2320-7868 (Online)

[11] Zhiyong Yu, Daqing Zhang, and Dingqi Yang in "Where is the largest market: Ranking areas by popularity from location based social networks," in Proceedings of International Conference on Ubiquitous Intelligence and Computing (UIC) IEEE, 157162, 2013.

BIOGRAPHIES

Pooja C. Sohale: Received B.E. degree in Computer Science and Engineering from SGBAU Amravati University. Now pursuing M.E. degree in Computer Engineering at VPKBIET, Baramati, Savitribai Phule Pune University.

Santosh A. Shinde: Received his B.E. degree in computer engineering (First class with Distinction) in year 2003 from Pune University and M.E. degree in computer engineering (First class with Distinction) in year 2010 from Pune University. He has 14 years of teaching experience at undergraduate and postgraduate level. Currently, he is working as Assistant Professor in Department of Computer Engineering of VPKBIET, Baramati, Savitribai Phule Pune University. His autobiography has been published in Marquis Who's Who, an International Magazine of Prominent Personalities of the World in the year 2012. He is also a Life Member of IACSIT and ISTE professional bodies. His research interests are Digital Image Processing and Web Services.