

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 1, January 2018

Combat data Sparsity using Spatial Hierarchical Collaborative Deep Learning for user POI Recommendation

Rajnandini Jamdade¹, Jyoti kharat²

P.G. Student, Department of Computer Engineering, JSPM Narhe Technical Campus, Pune, Maharashtra, India¹

Associate Professor, Department of Computer Engineering, JSPM Narhe Technical Campus, Pune, Maharashtra, India²

ABSTRACT: Point-of-interest recommendation has become an important way to help people discover attractive and interesting places, especially when they travel out of town. User performance may very important with respect to the geographical reason with different urban area cultures, environment. For solving these all issue I have to propose deep learning and proposed a special aware hierarchical collaborative deep learning algorithm. In this model jointly perform deep representation learning for point of interest from heterogeneous feature and hierarchical additive information learning from special aware individual performance. the collective preferences of the public in a given target region and the personal preferences of the user in adjacent regions are exploited in the form of social regularization and spatial smoothing. To deal with the multimodal heterogeneous features of the point of interests, we introduce a late feature fusion strategy into our special aware hierarchical collaborative deep learning method model.

KEYWORDS: POI recommendation, Spatial-Aware User Modeling, Deep Learning, DBN, Collaborative Filtering, Cold Start.

I. INTRODUCTION

In location based social network user can post their current location in the form of check in data on social network and share their experience on it. Using user check in data to make personalized point of interest recommendation in location based social network is used for helping users to explore new point of interests and regions and for facilitating location social network. In personalized point of interest recommendation is more essentially important and useful whenever user travel to unfamiliar area. In this scenario recommendation task is referred to out of town user. This paper is helpful for good recommendation system for home town and out of home town by effectively leveraging their check in data. There are three problems for challenging such problem.

A. Special Dynamics for user preferences

In special dynamics for user preferences Users tend to have different preferences when they travel in regions with different urban compositions and cultures, as noted in user preferences learned from their check-in history in one region (e.g., home town) may not always be effective for making recommendations in another region.

B. Data Sparsity

The number of POIs visited by an individual user is rather small compared to the total number of POIs in an LBSN, which results in a very sparse user-POI matrix. Moreover, the travel locality indicates that most of the users check-ins happen in their local regions (e.g., home cities). This observation aggravates the data sparsity problem further for out-of-town POI recommendations.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 1, January 2018

C. Cold Start Problem

In point of interests recommendation system cold start is very usual problem. point of interests that have not received any check-ins are called cold-start point of interests. Similarly, users who have not rated any point of interests are called cold-start users. Now days there are lots of methods are developed to make POI recommendations for both cold-start users. Now days there are lots of methods are developed to make POI recommendations for both hometown and out-of-town users. Now a day deep learning models have demonstrated great potential for learning effective representations and delivered the state-of-the-art performance in traditional recommendation tasks. We propose a novel POI recommendation framework to jointly perform spatial-aware personal preference modeling and POI semantic representation learning from versatile POI auxiliary information.

II. RELATED SURVEY

J. Bao, Y. Zheng, and M. F. Mokbel have introduced present a location based and preference-aware recommender system that offers a particular user a set of venues within a geospatial range with the consideration of both: 1) User personal preferences, which are automatically learned from her location history and 2) Social opinions, which are mined from the location histories of the local experts.[1]

T. Chen, W. Zhang, Q. Lu, K. Chen, Z. Zheng, and Y. Yu. Svdfeature introduced SVDFeature, a machine learning toolkit for feature-based collaborative filtering. SVD Feature is designed to efficiently solve the feature-based matrix factorization. The feature-based setting allows us to build factorization models incorporating side information such as temporal dynamics, neighborhood relationship, and hierarchical information.[2]

P. Covington, J. Adams, and E. Sargin described the system at a high level and focus on the dramatic performance improvements brought by deep learning. The paper is split according to the classic two-stage information retrieval dichotomy: first, we detail a deep candidate generation model and then describe a separate deep ranking model.[3]

R. Das, M. Zaheer, and C. Dyer replaced LDAs parameterization of topics as categorical distributions over opaque word types with multivariate Gaussian distributions on the embedding space. This encourages the model to group words that are a priori known to be semantically related into topics.[4]

G. Ferenc, M. Ye, and W.-C. Lee introduced the issues in making location recommendations for out-of-town users by taking into account user preference, social influence and geographical proximity.[5]

G. E. Hinton, S. Osindero, and Y.-W. Teh introduced the each training case consists of an image and an explicit class label, but work in progress has shown that the same learning algorithm can be used if the labels are replaced by a multilayer pathway whose inputs are spectrograms from multiple different speakers saying isolated digits.[6]

G. E. Hinton and R. R. Salakhutdinov presented the efficient learning and inference algorithms for this model, and show how a Monte-Carlo based method, Annealed Importance Sampling, can be used to produce an accurate estimate of the log-probability the model assigns to test data.[7]

D. Kim, C. Park, J. Oh, S. Lee, and H. Yu proposed a novel context-aware recommendation model, convolutional matrix factorization that integrates convolution neural network (CNN) into probabilistic matrix factorization. Consequently, ConvMF captures contextual information of documents and further enhances the rating prediction accuracy.[8]

S. Li, J. Kawale, and Y. Fu introduced a deep learning model for collaborative filtering that tightly couples matrix factorization based collaborative filtering with deep learning algorithm namely marginalized denoising auto-encoders (mDA).[9]

III. PROPOSED SYSTEM

In this proposed system, the main purpose is to offer high quality recommendations for both home-town and out-of-town users by effectively leveraging their check-in data. In this system I have to used deep learning models have demonstrated great potential for learning effective representations and delivered the state-of-the-art performance in traditional recommendation tasks and proposed a special aware hierarchical collaborative deep learning algorithm. In

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 1, January 2018

this model jointly perform deep representation learning for point of interest from heterogeneous feature and hierarchical additive information learning from special aware individual performance. the collective preferences of the public in a given target region and the personal preferences of the user in adjacent regions are exploited in the form of social regularization and spatial smoothing. To deal with the multimodal heterogeneous features of the point of interests, we introduce a late feature fusion strategy into our special aware hierarchical collaborative deep learning model.

A. Proposed System Architecture

Architecture introduces, users can post their physical locations in the form of a check-in and share their life experiences on the location based social network such as Yelp, foursquare. Using user check-in data to make personalized POI recommendation in location based social network is crucial for helping users to explore new POIs and regions. In our system user can upload dataset (yelp, foursquare) and user can apply spatial-aware hierarchical collaborative deep learning on review rating of location history, visited POI, and checkings. The model jointly performs deep representation learning for POIs from heterogeneous features and hierarchically additive representation learning for spatial-aware personal preferences. By using location based review and rating record our model will generate user POI recommendation.

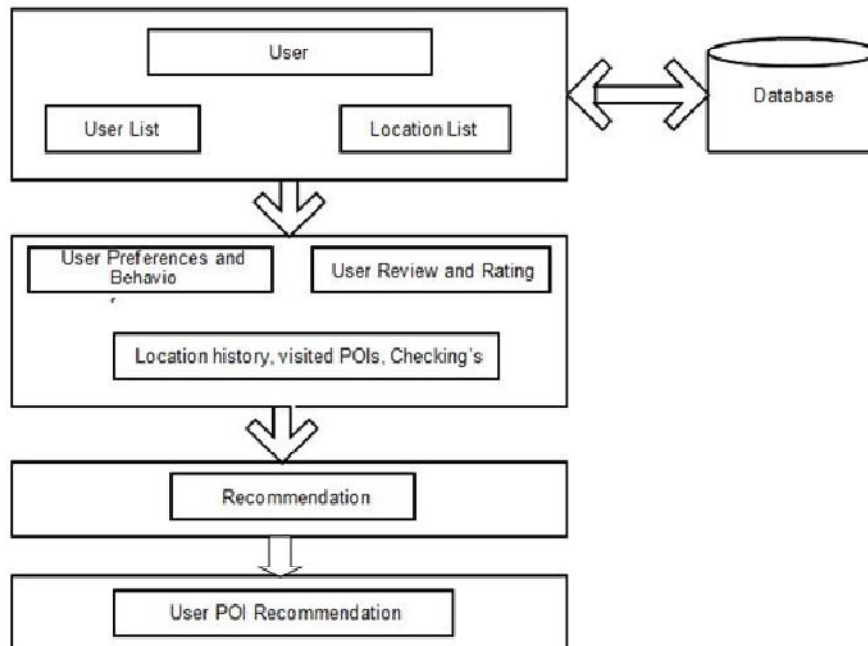


Fig. 1. Proposed Architecture

A. Mathematical Formulae:-

Let S be a system having Input(I), user(u) and POI(v).

$$S=I,U,V$$

where,

I is a input YELP dataset.

U is the number of user.

Copyright to IJIRSET

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 1, January 2018

V is the user point of interest.

$F = \{F1\}$

where,

F is the set of functions used for SH-CDL.

F1 is a function for SH-CDL.

Here we consider two latent factor for user and their POI interest (p_u, q_v).

The probabilistic generative model is listed as follow-

$$1) \text{Sample } p_u | \rho_u \sim N(0, \rho_u^2 \mathbf{I})$$

$$2) \text{Sample } q_v | \rho_v \sim N(0, \rho_v^2 \mathbf{I})$$

$$3) \text{Sample } C_{u,v} | \rho_u, \rho_v \sim N(g(\rho_u^T \rho_v), \rho_u^2 \rho_v^2 \mathbf{I})$$

Then,

1) Calculate the log-likelihood of probabilistic matrix factorization...

$$\text{LMF} = \sum_{u \in U} \sum_{v \in V} \|C_{u,v} - g(\rho_u^T \rho_v)\|^2 + \lambda_u / 2$$

$$\text{LMF} = \sum_{u \in U} \|\rho_u\|^2 + \lambda_v / 2 \sum_{v \in V} \|q_v\|^2$$

2) RBMs and Their Generalizations

$$E(x, y; \psi) = - \sum_{i=1}^k \sum_{j=1}^f w_{ij} x_i y_j - \sum_{i=1}^k b_i x_i - \varepsilon \sum_{j=1}^f a_j y_j$$

Where,

$\Psi = \{W, a, b\}$ are the model parameters and $\varepsilon = \sum_i X_i$ is the total number of words in a document.

3) Deep Belief Network

$$P(x, h | L; \psi) = \pi^{l-2} P(h^m | h^{m+1}) P(h^L, h^{L-1})$$

4) Model inference of SH-CDL

$$\frac{\sigma_{\text{LSH-CDL}}}{\sigma_{\Psi}} = \frac{\sigma_{\text{LSH-CDL}}}{\sigma_{\text{DBN}(X_v, \Psi)}} \frac{\sigma_{\text{DBN}(X_v, \Psi)}}{\sigma_{\Psi}}$$

5) POI Recommendation

$$S(q, v) = g(\Theta^T s q, r q, q v)$$

B. SH-CDL Algorithm

Step 1) Input: $q = (u_q, r_q, s_q)$

Where,

u_q =user query,

r_q =targeted region,

s_q =indicator.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 1, January 2018

Step 2) Select the top-k with highest ranking scores as recommendations.

$$S(q; v) = g(P^T uq; rqqv)$$

Step 3) As a latent factor vector qv is approximate to hidden feature vector $DBN(Xv, \psi)$.

The ranking score of the cold start POI can be computed as:

$$S(q, v) = g(P^T uq, rqq DBN(Xv, \psi))$$

Step 4) As for cold start user SH-CDL can make accurate recommendation according to collective preferences of the crowd with the same role Sq at target region rq .

$$S(q, v) = g(\Theta^T sq, rqqv)$$

IV. EXPERIMENTAL RESULTS

I have studied the impact of the spatial pyramids height and the regularization parameter for personal preference vectors on the Yelp dataset.

λ_u	OUT-OF-TOWN SCENARIO		HOME-TOWN SCENARIO	
	Ac@1	Ac@10	Ac@1	Ac@10
0.01	0.092	0.166	0.150	0.263
1	0.097	0.170	0.155	0.272

Fig. 2. Recommendation Accuracy on Yelp Dataset

λ_u is another important hyper-parameter in SH-CDL. The larger the λ_u more similar the personal preferences $\rho\mu, r$ are to the collective preferences of the crowd with the same role $\theta_{s,r}$ and the more strongly $\rho\mu, r$ is regularized. When λ_u becomes extremely small, our SH-CDL would degenerate to SH-CDL-T2 where the collective preferences are not exploited. To study its impact, we test the performance of SH-CDL model by varying the value of λ_u from 0.01 to 100, and the results are presented in Table 2. From the results, we observe that the recommendation accuracy of SH-CDL first increases with the increasing λ_u and then begins to decrease.

We first choose POIs with less than 10 check-ins as cold-start POIs, and then select users with at least one cold-start check-ins as test users. For each test user, we first choose her check-ins associated with cold-start POIs as the test set, and the remaining check-ins as the training set. Our aim is to measure whether the marked-off cold-start POIs in the test set can be accurately recommended to the right users in the top-k results.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 1, January 2018

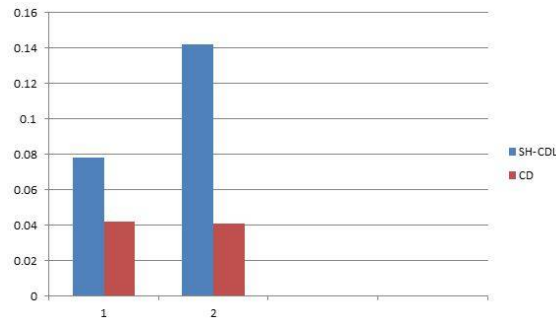


Fig 3. Out-of-town Recommendation for Yelp

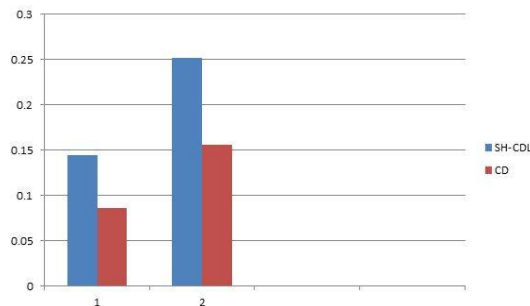


Fig 4. Home-town Recommendation for Yelp

V. CONCLUSION

We developed a POI recommendation model SH-CDL to jointly perform deep representation learning for POIs from heterogeneous features and hierarchically additive representation learning for spatial-aware personal preferences to overcome the challenges of the spatial dynamics of user preferences, cold start and data sparsity. Social regularization and spatial smoothing technologies were developed to overcome data sparsity in the spatial-aware dynamic user preference modeling. To deal with the multimodal heterogeneous features, we extended the DBN to MDBN by introducing a late feature fusion strategy.

REFERENCES

1. J. Bao, Y. Zheng, and M. F. Mokbel. Location-based and preference-aware recommendation using sparse geo-social networking data. In SIGSPATIAL, pages 199208, 2012.
2. T. Chen, W. Zhang, Q. Lu, K. Chen, Z. Zheng, and Y. Yu. Svdfeature: A toolkit for feature-based collaborative filtering. J. Mach. Learn. Res., 13(1):36193622, Dec. 2012.
3. P. Covington, J. Adams, and E. Sargin. Deep neural networks for youtube recommendations. In RecSys, pages 191198, 2016.
4. R. Das, M. Zaheer, and C. Dyer. Gaussian lda for topic models with word embeddings. In ACL, pages 795804, 2015.
5. G. Ferenc, M. Ye, and W.-C. Lee. Location recommendation for outoftownusers in location-based social networks. In CIKM, pages 721726, 2013.
6. G. E. Hinton, S. Osindero, and Y.-W. Teh. A fast learning algorithm for deep belief nets. Neural computation, 18(7):15271554, 2006.
7. G. E. Hinton and R. R. Salakhutdinov. Replicated softmax: an undirected topic model. In Advances in neural information processing systems, pages 16071614, 2009.
8. D. Kim, C. Park, J. Oh, S. Lee, and H. Yu. Convolutional matrix factorization for document context-aware recommendation. In RecSys, pages 233240, 2016.
9. S. Li, J. Kawale, and Y. Fu. Deep collaborative filtering via marginalized denoising auto-encoder. In CIKM, pages 811820, 2015.
10. X. Li, G. Cong, X.-L. Li, T.-A. N. Pham, and S. Krishnaswamy. Rankgeofm: Ranking based geographical factorization method for pointof interest recommendation. In SIGIR, pages 433442, 2015.