ENHANCED RATING PREDICTION BASED ON LOCATION AND FRIEND SET

[1]Santhosh Voruganti,[2]Thota DivyaGoutham,[3]Phanindranath Peddi,[4]karnati RamyaKrishna
[1,2,3] Department of Information Technology, CBIT,Hyderabad
[4]Department of MCA, Osmania University,Hyderabad

Abstract : As the mobile phones are getting more sophisticated with GPS (locations services). It is opening up the new research based on the consumer behavior based on location. The friend set from the social media can be mined for research on the consumer behavior based on the friend recommendations. Location and friend set are used for predicting the rating. We also reduce the error in prediction using the Sparsity reducing index. In this paper, We mine 1. Location weight, 2.Friends weight and 3. Sparsity reducing index and combining all the three factors, we predict the rating of particular service or hotel.

***IndexTerms* - Rating prediction; location; friend set; sparsity, location weight; Friends weight.**

I. INTRODUCTION

Whenever, we plan to eat or buy item or utilize an item. We take different factors into consideration. So, the plain average rating of the service or item could not be the only entity for recommending the people. The rating prediction is nothing but adding weight to the average rating based on certain factors. Rating prediction helps in ranking the product based on certain factors. Which in turn helps in recommending the product to people? The certain factors could be location of the service and recommendation by friend's mouth. Location plays an important role in choosing a service. How far is the service or the item is from the home? If a person travels a long distance for a particular service and gives high rating, this shows that it is very good and unique. So, the Rating for this particular service should be higher. Similarly, if he travels long distance and give less rating. It shows that service is really poor. So, the service which gets good rating from the people coming from the farthest location gives extra punch to the rating. Friends also play an important role in recommending the services. Generally, people take recommendations from multiple friends take a decision. So, if more friends give positive feedback then we tend to utilize the particular item. As the distance increases we also take the friend recommendations. For example, we are travelling to a new country we tend to take the recommendations from the friends in that location. Lesser the rating difference between the friends, Higher the rating. The feedback system is generally consisting of different attributes. For example, A restaurant will have attributes like rating for food, service, ambience and parking. Most of the people would not give the rating properly which causes the sparsity of data. The sparsity of the data may cause deviation in the results. So, it is very important to handle the sparsity of the data. The sparsity reducing index as the name suggests compensates the sparsity by adding a weight to the ratings.The paper is structured as follows. The designing of rating prediction systems using models which are presented in Section 2. The design concept of our model is briefed in Section 3. The implementation of our system in Section 4 and evaluation of our system in Section 5. The results of the system in Section 6 and conclusion of our work in Section 7.

II. LITERATURE REVIEW

A. Matrix Factorization in Social Networks.

It is an approach for recommending in social network. It stores the user's ratings and items for performs matrix factorization.Cold-start users are the new users without any friend's in the systems. They are the biggest challenge in the recommendation. Most of the social network user do not involve in the rating. Sparsity of the data is also important challenge in this method.

B. Personalized Recommendation Combining User Interest and Social Circle.

Most number of social users are participating in the ratings and reviews. The new technique like interpersonal influence and friend circle interest try to eliminate the cold start and sparsity problem of datasets. For cold start users, the interpersonal interest similarity solves the issue.

C. Summary

The cold start, situation where recommendation for the new users in the system without friend is difficult. This has to be addressed using different technique so that the new user should be able to get the recommendation right from joining based on the certain factors. The sparsity of the data should also be handled. Otherwise, the sparsity of data may cause huge deviation from the results.

MODEL OVERVIEW

The model consists of three main aspects location, friends and sparsity. The predicted rating must be based on the above three factors. We propose following terms which will help us predict the rating for the cold start users.

1. Location weight.
2. Friends weight.
3. sparsity reducing index.

The deviation from the average rating must be less because it the average rating is also a key aspect of the recommender system.

Figure 1: Design Overview of the model

A. Location Weight

Relevance between user ratings and distance between user location and item. X-axis-distance between social user's location and service location Y-axis ratings given by social users. Dots are actual data. Line is the normalize distribution using Gaussian model.

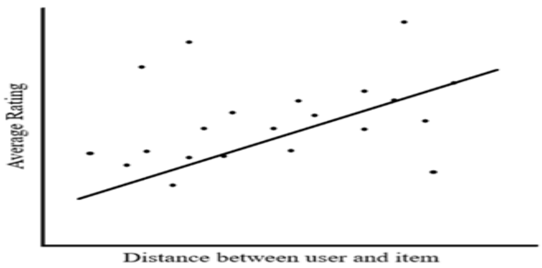

Figure 2: Sample graph for location weight

Gaussian probability distribution equation,

$f(x)=lw=\Sigma_i a_i * \exp(-((x-b_i)/c_i)^2)$

To solve the a, b, and c coefficients values we need to take following constraints. Applying linear Curve fitting,

$\Sigma Y=n\,a + b\,\Sigma X;$ $\Sigma XY= a\,\Sigma X + b\,\Sigma X*X;$

Applying Cramer's Rule on above equations,

$a=(\Sigma Y*\Sigma XX)-(\Sigma X * \Sigma XY)/(n*\Sigma XX)-(\Sigma X*\Sigma X);$

$b=(n*\Sigma XY)-(\Sigma Y * \Sigma X)/(n*\Sigma XX)-(\Sigma X*\Sigma X);$

Since, using Cramer's Rule $px + qy = m$; $rx + sy = n$ then,

$x=(ms-qn)/(ps-qr)$; $y=(pn-mr)/(ps-qr)$, c is the mean of X values.

B. Friends Weight

Relevance between rating differences in friends and Distances between friends. X-axis distance between friends. Y-axis rating difference (user1: 5, user2: 2 and rating difference= 5-2= 3) given by friends. Dots are actual data. Line is the normalize distribution using Gaussian model.

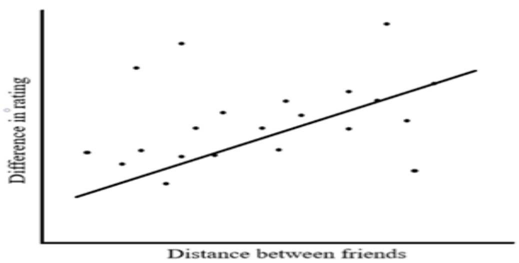

Figure 3: Sample graph for friends weight.

Gaussian probability distribution equation,

$f(x)=fw=\Sigma_i a_i * \exp(-((x-b_i)/c_i)^2)$

To solve the a, b, and c coefficients values we need to take following constraints.Applying linear Curve fitting,

$\Sigma Y=n\,a + b\,\Sigma X;$ $\Sigma XY= a\,\Sigma X + b\,\Sigma X*X;$

Applying Cramer's Rule on above equations,

$a=(\Sigma Y*\Sigma XX)-(\Sigma X * \Sigma XY)/(n*\Sigma XX)-(\Sigma X*\Sigma X);$

$b=(n*\Sigma XY)-(\Sigma Y * \Sigma X)/(n*\Sigma XX)-(\Sigma X*\Sigma X);$

Since, using Cramer's Rule $px + qy = m$; $rx + sy = n$ then,

$x=(ms-qn)/(ps-qr)$; $y=(pn-mr)/(ps-qr)$; c is the mean of X values.

C. Sparsity Reducing Index

User interest is a representative and prevalent factor in recommender system. It is necessary to represent user interest vector.

Ex: User 1 (u) participated for rating of some attributes of service: att 1, att2, att 3. User 2 (v) participated for rating of some attributes: att 1, att 3. Difference (u, v) / total (u, v) in this case is 2/3 sri =Σ difference (u, v) / total (u, v);

D. Rating prediction

It is the combination of following, Rating prediction, R*= lw + fw +sir;

E. Root mean square error

The RMSD [11] is the standard deviation of the differences between predicted rating and observed rating in a sample data.

Where, y_i=original rating,(y_i) ̂=Predicted rating and n=number of items.

III. IMPLEMENTATION

A. Hardware The hardware specification of the system to build and compare the models is as follows; CPU- Intel core I5 , RAM- 8GB, 100 GB of free disk space

B. Software : Windows 7 or higher with Apache tomcat, MySQL server, Java JDK SE 8 and Editplus 5. It is a JavaScript web application. A web browser like Mozilla Firefox, google chrome, Microsoft edge or opera web is required**.**

EVALUATION : In order to evaluate the implementation of our model, we have built Webapp and checked the results The item details and geo location are added as shown in the below figure.

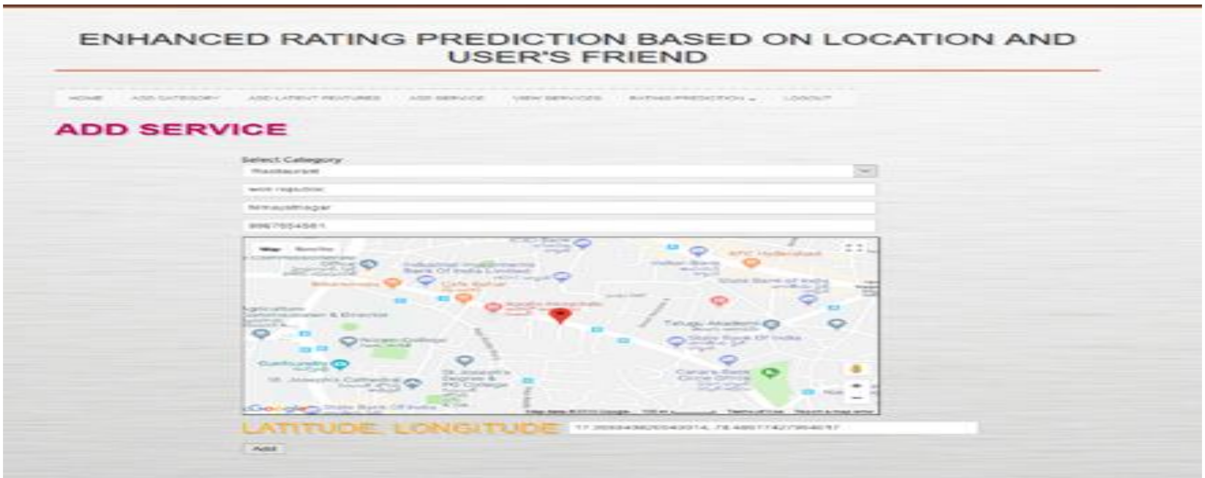

Figure 4: Adding Service item

User ratings and Location details are collected as shown in figure 5 and 6.User rating can be divided into sub categories for rating the item. For example, Restaurants can divide the rating into sub categories like food, ambiance and service.

Figure 5 :User rating an item

Figure 6 : Collecting the location of the user

User can maintain a friend circle by the request mechanism. The friends ratings can be used for predicting the friends' weights.

Distance between friends	**Difference rating**	**Friends weight**
Increase	Increase	Positive
Increase	decrease	Negative

Figure 7 : Friends can be added by request mechanism.

After collecting all the data, the ratings are predicted for each item by calculating location weight, friend weight and sparsity index is calculated for each service item. The Following results and observations were obtained.

IV. Results :

LOCATION WEIGHT

As the rating increases with distance, the location weight also increases. Higher ratings by people staying far from the item's geographical location adds more weight to the item.As the rating decreases with increasing distance, the location weight decreases. A lower rating by people staying far from the item's geographical location reduces the weight to the item.

FRIEND'S WEIGHT

As the rating Difference increases with distance, the location weight also increases. High rating difference among the nearby friends signifies the poor performance by the item. Hence, reducing the weight of the item.Higher rating difference among friends far apart, adds more weight because the one among them likes the item and might recommend it to the other friends.

We have taken a small survey among friends

Root mean square error = 0.232. The system predicts the Rating with a root mean square error of less than 1. The signifies that the weights added to the rating are always far less than one. The weight signifies the role of location and friends in the recommendation without deviating much from the Average rating of the item. The root mean square error for the location weight is far less than 1. The root means square error for friends' weight is also far less than 1.

Distance	Rating	Location weight
Increase	Increase	Positive
Increase	decrease	Negative

Service	Original rating and weights		Rating prediction
McDonalds	Avg.	3.27	3.46
	lw	0.01	
	fw	0.02	
	sri	0.16	
Burger King	Avg.	3.00	3.13
	lw	-0.06	
	fw	0.03	
	sri	0.16	
KFC	Avg.	3.11	3.42
	lw	0.05	
	fw	0.09	
	sri	0.16	

V. DISCUSSION

Though the proposed model predicts the rating as expected, there is a limitation,

Rating Prediction Beyond the Rating Range

The predicted rating may go above the upper limit of the rating range or may go below the lower limit of the rating range. In this case we have to strictly bound to the rating range. i.e. if the rating prediction is below 0, then rating prediction should be considered as 0. Similarly, if rating prediction is more than 5 then we must consider it as 5.

VI. CONCLUSION AND FUTURE SCOPE

In this work, we present an approach for rating prediction which can overcome sparsity of the data the cold-start problem in the recommender system. The three factor which influence the user are location (Distance from his/her home), friends (friend recommendation are also important role in finding the better service) and sparsity of data has to be reduced to obtain a perfect result with less root mean square error. The deployed system was able predict the ratings with a root mean square error less than 1.

In the future, most of the websites can implement this design for recommending the services based on the location and social friends.

REFERENCES

[1] C. Cheng, H. Yang, I. King, M. R. Lyu, "Fused Matrix Factorization with Geographical and Social Influence in Location-Based Social Networks," *AAAI'12*, 2012.

[2] R. Salakhutdinov, and A. Mnih, "Probabilistic matrix factorization," NIPS, 2008.

[3] Y. Koren, R. Bell, and C. Volinsky, "Matrix factorization techniques for recommendation systems," Computer, pp. 30-37, Aug.2009.

[4] X. Qian, H. Feng, G. Zhao, and T. Mei, "Personalized Recommendation Combining User Interest and Social Circle," *IEEE Trans. Knowledge and Data Engineering*, vol. 26, no. 7, pp. 1763–1777, 2014.

[5] X.-W. Yang, H. Steck, and Y. Liu, "Circle-based recommendation in online social networks," KDD'12, pp. 1267-1275, 2012.

[6] H. Feng, and X. Qian, "Recommendation via user's personality and social contextual," ACM CIKM, 2013.

[7] G. Adomavicius, and A. Tuzhilin, "Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions," IEEE Transactions on Knowledge and Data Engineering, pp. 734-749, Jun. 2005.

[8] M. Jamali, and M. Ester, "A matrix factorization technique with trust propagation for recommendation in social networks," ACM RecSys, 2010.

[9] G. Zhao, X. Qian, and H. Feng, "Personalized Recommendation by Exploring Social Users' Behaviours," In Proc. MMM, 2014.

[10] J. J. Levandoski, M. Sarwat, A. Eldawy, M. F. Mokbel, "LARS: A Location-Aware Recommender System," ICDE'12, 2012.