RECOMMENDER SYSTEMS WITH SOCIAL REGULARIZATION

Pulkit Kariryaa, Anubhav Kumar
Advisor: Prof. Piyush Rai
Indian Institute of Technology, Kanpur

Introduction

Recommender system is a information filtering system which is used to predict user responses to various options. Traditional approaches to recommender systems are based on collaborative filtering methods but make no use of social data available. In this project we tried to use the social graph available for a user to improve the obtained results we used matrix factorization using gradient descent with social regularization

Methodology

We used low rank matrix factorization which basically factorizes the user item matrix to predict missing values

$$R \approx U^T V = R^*$$

Average based method

$$E_{ij}^2 = (R_{ij} - R_{ij}^*)^2 + \frac{\mu_1}{2}\|U\|^2 + \frac{\mu_2}{2}\|V\|^2 + \frac{\gamma}{2}(\|U_i - \frac{\sum sim(i,f) \times U_f}{\sum sim(i,f)}\|^2)$$

User based method

$$E_{ij}^2 = (R_{ij} - R_{ij}^*)^2 + \frac{\mu_1}{2}\|U\|^2 + \frac{\mu_2}{2}\|V\|^2 + \frac{\gamma}{2}(\|\sum sim(i,f)(U_i - U_f)\|^2)$$

we used gradient descent to minimize E_{ij}^2 and complete the matrix

Similarity Function

We implemented Pearson State Coefficient method to find similarity function:

$$sim(i,f) = \frac{\sum((R_{ij} - \overline{R_i})(R_{fj} - \overline{R_f}))}{\sqrt{\sum(R_{ij} - \overline{R_i})^2}\sqrt{\sum(R_{fj} - \overline{R_f})^2}}$$

Dataset

Initially we tried to use epinions dataset provided on the website, but that dataset was quite large so our algorithm was taking huge time to work. So our supervisor provided us a dataset he had which contains a matrix of 1000 users vs 1500 items and a graph of social connection between different users.

In our experminets we used 50% data for the training and other 50% for testing.

Conclusion and Future work

- Recommender system with social regularization performed better than the normal method

- Instead of running this method for all friend we can selectively use this social data to get more accurate results

- Similar methods can also be applied to items as well for giving suggestions

Results

We used MAE(Mean Error) and RMSE(Root Mean Square Error) to show effects of our method by varying various parameter
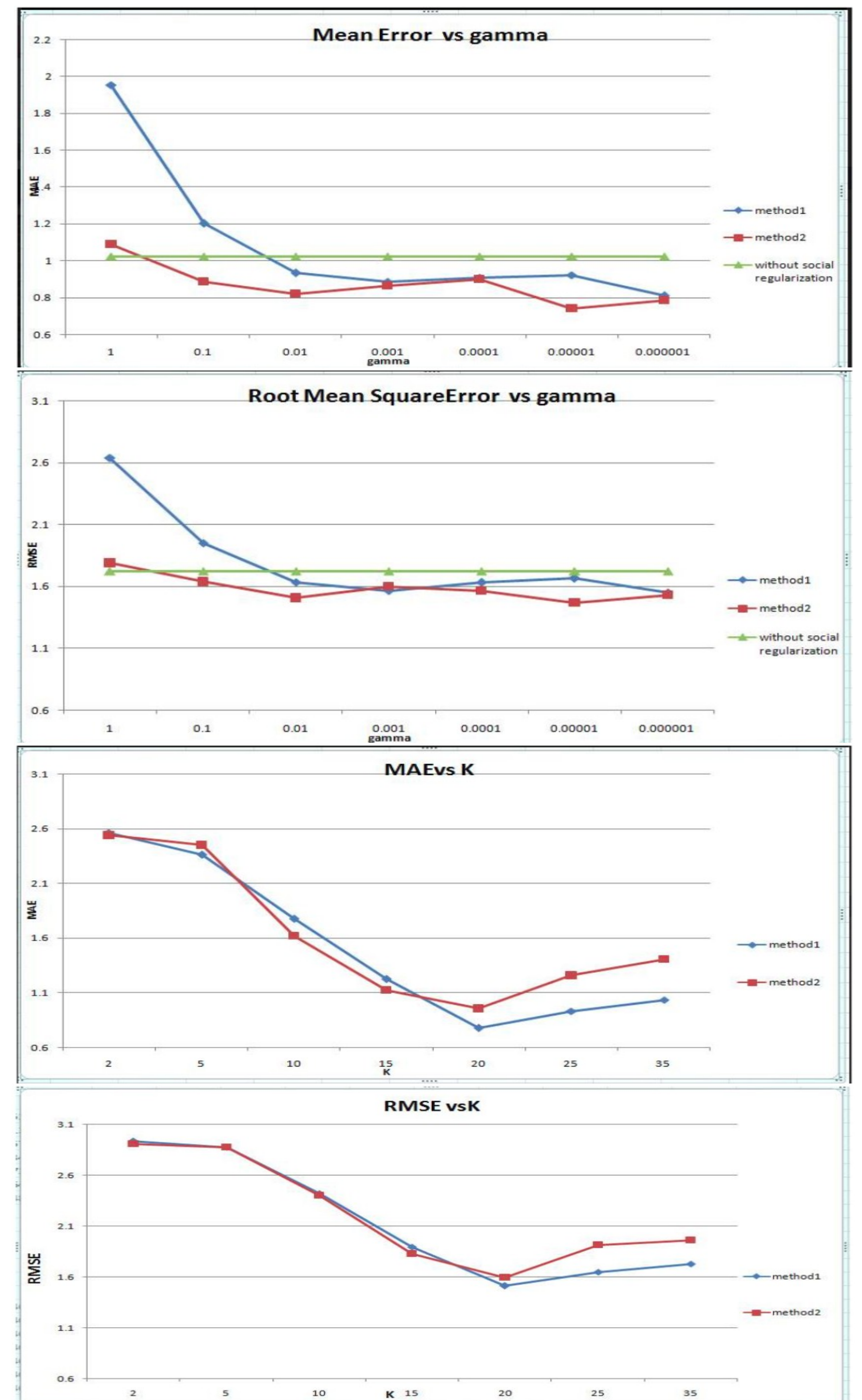
Our Work

- We used similarity between users and items to update the ratings but the change in result was not significant

$$R_{ij}^* = R_{ij}^* + \delta(U_i V_j^T - 0.5)$$

- We tried using a comibined apporach in which we used normal method for users who rated many items and user based method for other users but we couldn't get signifincant change in result as the data on which we worked is too sparse.

- We used the similarity between the user's friend and the item for factorization but still the change in result was not significant on our dataset

$$\sum_{f \in F_i} simVU(j,f)\|U_i - U_f\|^2$$

References

- http://www.quuxlabs.com/blog/2010/09/matrix-factorization-a-simple-tutorial-and-implementation-in-python/

- Hao Ma , Dengyong Zhou , Chao Liu , Michael R. Lyu , Irwin King, Recommender systems with social regularization, Proceedings of the fourth ACM international conference on Web search and data mining, February 09-12, 2011, Hong Kong, China

- http://www.public.asu.edu/~jtang20/datasetcode/truststudy.htm