Deep Learning based Trust-Aware Recommender for Social Networks

Pooja Dinkar Shinde, Dr. Sunil Rathod

[1, 2]*Department of Computer Engineering, Dr. D. Y. Patil School of Engineering, Lohegoan, Pune, India*

Abstract: Due to the rapid growth of information available on the WWW (World Wide Web) and the rapid introduction of new web services, information overload becomes a critical problem overwhelms the users. It also becomes an unprecedented challenge to determine how to help users find the information they are interested in, and this issue is attracting the attention in both research and application areas. In this paper, we propose a new recommendation technique for the trust aware recommendation in social networks based on the Deep Learning (DL). Here, we firstly find out the per-train initial value of the parameter for this we used a deep auto-encoder. The community detection algorithm based on trust relations in social networks is proposed for the revamp the MF (Matrix Factorization) model with social trust together and community effect. The benefit of this system is that it deals with the cold start users.
Keywords: Social Network, Recommendation Techniques, Deep Learning, Cold Start problem, and Matrix Factorization.

I. INTRODUCTION

Now-a-days the recommender strategy plays an important role in advising of the items that the user is interested in. That is why they use the previous interaction records system in order to extract the user's taste, and they provide the correct list of items in the users, which music to listen to, which items to buy or read any book/news. Most of the social networking sites are used to become directly interconnected to share all kind of information like professional profile, videos, icons and news etc. A user contacts other users for the information sharing between them on social media sites like Facebook, Twitter, Weibo etc. With the number of friends increasing, amount of information sharing is also increasing and because of this, information overloads problems are caused and they can be solved by using the trust about recommended techniques.

Although Recommender Systems (RSs) have widely been studied in both the academia and industry, some important problems still remain such as sparsity problems, cold start problems and trustworthiness problem -

1) *Cold-Start Problem:* A cold start is an item that is rated by a user who has expressed a no or some rating, or by a small number of users. Because there is not enough evaluation data, a similar-based approach would not be able to find the user or item next to them recently, and would use the traditional recommendation algorithm.

2) *Sparsity Problem:* Users typically rate or experience only a small subset of the available items. As a result, the available rating density of RSs is often less than 1% [2]. Due to this data sparsity has a lot of difficulties with the collaborative filtering methods of similar users of the system. Therefore, the predicted quality of RS may be significantly limited.

3) *Trustworthiness Problem:* Traditional RSs usually lack the ability to distinguish user creditability, that is, the reliability of user ratings. In fact, the user's trust is more determined by other business users rather than their friend's voice. Moreover, there will always be spam users who give fake ratings for malicious purposes. Apparently, the rating recommends is excluding.

Due to the problem of finding new solutions to the trust-based approach they use recommendations of the social network [3]. Figure-1 shows a set of trust-aware recommendation techniques to illustrate key ideas for a trust-aware recommendation technique that allows users to rate items based on their own personal experience.
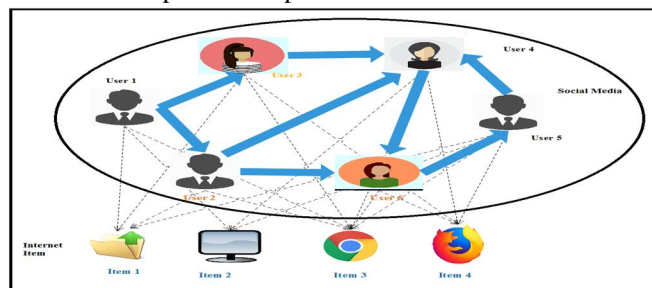

Figure1: Trust-Aware Social Recommendation Example

TABLE I. MATRIX OF USER RATING

Items/ Users	Item 1	Item 2	Item 3	Item 4
User 1	*****	***	*	****
User 2	***	**		*****
User 3	*	*****	**	*
User 4	*****		*****	**
User 5		*****	***	*****
User 6	***	****	**	

In this paper study about the related work is done in section II, whereas the proposed approach modules description, mathematical modeling, algorithm and experimental setup are discussed in section III. Finally, we provide a conclusion in section IV.

II. LITERATURE REVIEW

Here we discuss the literature review of existing techniques:

F. Alqadah, Ju. Hu and H. F. Alqadah, they propose Collaborative Filtering techniques for the top-n recommendation task they used bi-clustering neighborhood approach [1]. The result shows that the proposed methods generate a better recommendation than the existing modern algorithm on sparse data. Performance infrastructure Biclustering neighborhood (BCN) is evaluated on five real datasets, such as PayPal, dataset, lastfm dataset dataset lastfm_friends and delic_bookmark dataset. The performance of BCN methods is compared with three different algorithms, namely SLIM (sparse linear methods), Item CF (joint filtering), and WRMF (weighted matrix factorization regulation).

A. Javari and M. Jalili, they proposed Probabilistic techniques to resolve Diversity-Accuracy Dilemma in the RS (Recommendation System) [2]. The proposed system of recommendations there are two models: the First-the maximization of accuracy and the second-specified in the list of recommendations to the tastes of users. The recommended technique is based on the Markov model. For this experiment, they use two real-world datasets, such as Movielens and Netflix. These datasets are divided into test and training datasets.

Y. Rao, N. Zhang and H. Zou they proposed Topology-based ensemble model for combining users own taste and his trusted users/friends testes [3]. Experimental analysis performed on Flixster, Epinions, and Ciao for better results. The proposed method is applicable not only to the ensemble based on the model, but also to the methods of recommendation of the ensemble based on the instance. The hidden key theme model (SLTM) is used to build a connection between feelings and words.

Gu. Xu, Do. Wang and Sh. Deng they propose emotional aware recommendation approach to incorporate an emotional context into music recommendation. They use a Chinese data set of the twitter service [4]. The performance of this technique has improved in terms of accuracy, recall, hit rate, and F1 rating. The proposed method is compared with two existing methods, such as fine-grained emotion from microblogs and rough information about emotions. In this they first music service connected to microblogs services such as Sina Weibo and Twitter for collecting more listing records. In this they increase the accuracy and efficiency of musical techniques recommendations.

Shui Guang Deng Jian Wu and Zhao Hui Wu, they proposed techniques for providing personalized service recommendation to individual's user, for this they used trust relationship between services and users [5]. The proposed recommendation method is based on a collaborative filtering approach, such as a trust-based service recommendation. The TSR methods are compared with the other five approaches based on the ratio of TSR, QoS at response rate, QoS at trough output, hits and PAGERANK.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.177
Volume 7 Issue VI, June 2019- Available at www.ijraset.com

S. Dhillon, Hsiang Fu Yu, and I. S. Dhillon they propose techniques for the solving the large scale matrix factorization problems in recommendation system [6]. In this, they used optimization techniques such as the cyclic coordinate worthy approach (CCD), and it is applicable to large-scale problems such as the maximum entropy model, NMF problems, sparse inverse covariance estimation, and linear SVM. The CCD approach updates a single variable simultaneously while keeping the other fixed. The proposed system is compared with ALS, SGD and CCD++ in large-scale data. The dataset used for the experiment is Movielens 10m, Movielens 1m, Yahoo Music, and Netflix.

Ha. Park, M. Ishteva, and R. Kannan they used matrix factorization approach, such as the Bounded Matrix Factorization (BMF) approach for the rating matrix [7]. They test the performance of the proposed algorithm on a real dataset, such as Netflix, and compare the modern algorithm with SVD++, SGD, BIOS-SVD, and ALSWR.

Shuiguang Deng, Guandong Xu and Longtao Huang they propose recommendation method for services based on social networks [8]. To begin with, The Matrix decomposition method was used in order to evaluate the reliability among users in social networks. Next, an extended random walk algorithm is proposed and recommended results are obtained. In order to evaluate the accuracy of the algorithm, experiments on real-world data sets were carried out, and it was shown that the quality of the recommendations and the speed of the method were improved compared with the existing algorithms.

Kuan Zhang, Ee Peng Lim, and David Lo compact lossless representation based on the concept of closed patterns is extracted so as not to explode the number of mining antagonistic communities [9]. In addition, the main memory is not enough for large data sets that use the current variation algorithm. The extensibility of the approach or synthetic data sets of various sizes is mined in various parameters. Case studies on Amazon, Epinions, and the Slashdot data set further demonstrate the efficiency and efficiency of this approach in extracting antagonistic communities from social interactions.

Balazs Hidasi and Domonkos Tikk they present a general initialization framework that preserves the similarity between entities (users / items) in the creation of initial feature vectors [10]. They also used initialization framework to combine the MF algorithm. Various similarity functions, different contexts, and similar concepts of metadata databases are experimented. The evaluation is performed on two implicit variants: the MovieLens10M dataset and the four actual implicit databases. They also show that the initialization significantly improves the performance of the MF algorithm most ranking measures.

III. PROPOSED APPROACH

A. Problem Statement

To develop techniques such as trust-based approach for recommendation in social networks using MF (Matrix Factorization) for trust-aware social recommendations and to differentiate the community effect in users trusted friendships.

B. Proposed System Overview

This paper mainly focuses on the Trust-Based recommendations; Memory-based approaches have largely figured on integrating trust into recommendations. The most common RSs cause users to issue trusted statements to other users. We used a Matrix Factorization approach for trust-aware recommendation in social networks; called DLMF (Deep Learning based Matrix Factorization). Matrix factorization characterizes both the items and the users by vectors of factors deducted from item rating patterns. High correspondence between the item and the user factors leads to a final recommendation. Matrix factorization models map both the users and the items to a joint latent factor space of dimensionality f, such that the user-item interactions are modeled as inner products in that space. Each of the item, say i, is associated with a vector $q_i \in \mathbb{R}^f$, and each of the user, say u, is associated with a vector $p_u \in \mathbb{R}^f$. Now, for a given item i, the elements of q_i will measure the extent to which the item possesses those factors, positive or negative. And for a given user u, the elements of p_u will measure the extent of interest that the user has in items that are high on the corresponding factors - positive or negative. The resulting dot product, $q_i^T p_u$, will capture the interaction between user u and item i—the user's overall interest in the item's characteristics. This approximates user u's rating of item i, which is denoted by r_{ui}, will lead to the estimate

$$\hat{r}_{ui}, = q_i^T p_u \qquad (1)$$

The major challenge is computing the mapping of each item and user to factor the vectors $q_i, p_u \in \mathbb{R}^f$. After the completion of the mapping by the recommender system, it can easily estimate the rating a user will give to any item by using Equation 1.

The advantage of matrix factorization is the fact that it can incorporate implicit feedback, information that is not directly given but can be derived by analyzing user behavior. Using this strength we can estimate if a user is going to like an item that (he/she) is unaware of and if that estimated rating is high, we can recommend that item to the user.
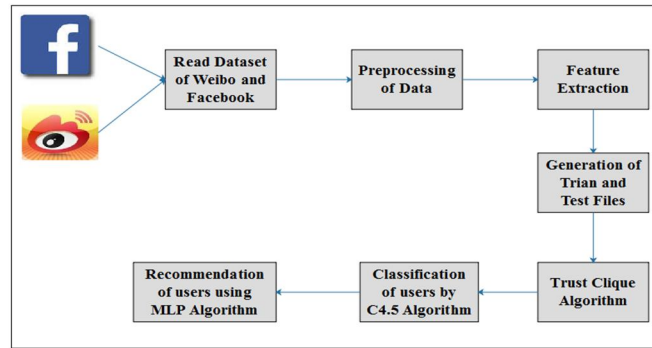
Figure 2: Proposed System Architecture

In this system architecture we have following modules:
1) *Read Dataset:* Firstly, we read from the Facebook and Weibo website. In this weibo.com dataset contains data of Chinese social media websites which contains 500+ users' data and for comparison we also used facebook.com dataset which contains 500+users.
2) *Preprocessing Data:* In this preprocessing stage unwanted data or irrelevant data is removed, this data are removed using stop words and stemming operations.
a) *Stop Words:* They are frequently occurring and insignificant words in a language that help construct sentences but do not represent any content of the documents.
b) *Stemming:* Stemming refers to the process of reducing words to their stems or roots.
3) *Feature Extraction: In* this feature selection the part of original features is sorted and here takes essential part of data. The module takes features as input and feature selection method is applied using this algorithm. It increases learning speed of classifier.
4) *Deep Autoencoder:* The determination of the initial user and item latent feature vectors should obviously be aligned with the learning algorithm applied. The learning of user latent and item matrix is a minimization of the loss function between the observed ratings R and the predicted ratings. Due to the high dimensionality of user and item, there may exist a set of local optimal values in the whole loss function space. How to select an appropriate initialization of user and item latent feature vector will undoubtedly influence the convergence of these matrices. To obtain the user and item latent feature in a low-dimensional space, it would be easier to deal with the initialization difficulty. To do this, here utilize a deep learning method, i.e. deep autoencoder, which is an efficient approach to nonlinear dimensionality reduction. Deep autoencoder is quite suitable for the initialization phase, which can abstract the high-dimensional rating data to latent features.
5) *Trust Clique Algorithm:* In this n-trust-clique is used to identify the cliques in the social network which is based on the trust relationship between the users. Here, n-clique is used to detect overlapping communities. It guarantees good performance. It is a typical concept for n-cliques in social science. Given a network N, an n-clique is a maximal sub-graph in which the largest distance of each pair of nodes is number greater than n and equation shows the distance is defined in the original network:
$$\text{distance}(D_i, D_j) \leq \forall D_i, D_j \in N$$
6) *Recommendation of Users using C4.5 MLP Algorithm:* Here, used recommend on the basis of extracted features from dataset like flowers counts, friend count and following count.

C. Algorithm
1) *Algorithm 1:* C4.5 Algorithm
Process
a) Check for the below base cases:
i) All the samples in the list belong to the same class. When this happens, it simply creates a leaf node for the decision tree saying to choose that class.
ii) None of the features provide any information gain. In this case, C4.5 creates a decision node higher up the tree using the expected value of the class.
iii) Instance of previously-unseen class encountered. Again, C4.5 creates a decision node higher up the tree using the expected value.

b) For each attribute a, find the normalized information gain ratio from splitting on

c) Let a best be the attribute with the highest normalized information gain.

d) Create a decision node that splits on a_best.

e) Recur on the sublists obtained by splitting on a_best, and add those nodes as children of node.

2) *Algorithm 2:* MLP(Multi-layer perceptron) Process:

For classification, we use Multi-layer Perceptron classifier. MLP works in two propagation feed - forward propagation and back propagation. So here, the proposed algorithm uses feed forward propagation to recommend users on the basis of extracted features from dataset like flowers counts, friend count and following count.

a) *Input:* Query

b) *Dataset:* weibo.com user's data and facebook.com user's data.

c) *Process:*

i) Create test files using dataset.

ii) Take query as input & creating training file related to query.

iii) Pass training & testing file to multilayer perceptron.

iv) Matching query on test data assigning class

v) If query match then

vi) class =1   // find related user

vii) else

viii) class =0 // non-related users

ix) Stored relevant users   who has class = 1 in matched   with users list

x) Ends

3) *Algorithm 3:* Trust Cliques Algorithm:

a) *Input:* Clique size n, iteration limitation l, User set U, user trust matrix T ;

b) *Ensure:*

i) Clique's collection Communities;

ii) Communities = U, n_c =| Communities|;

iii) While, n_c = 1 or number of iteration < l do;

iv) For i = , n_c do;

v) C_i = the ith cliques in Communities;

vi) J = the set of C_i 's connected vertices;

vii) ΔCol_{max} = 0;

viii) for each $v_j \in J$ do;

ix) $C'_i = C_i \cup v_j$ do;

x) If C'_i satisfies the first condition of n-trust-clique;

xi) Then

xii) Calculate ΔCol for the change;

xiii) If $\Delta Col > \Delta Col_{max}$ then

xiv) $\Delta Col_{max} = \Delta Col$

xv) $C_t = C'_i$;

xvi) End if;

xvii) End if;

xviii) End if;

xix) If$\Delta Col_{max} > 0 \vee Random(0,1) >$ probthen;

xx) Delete clique C_i to Communities;

xxi) Store current partition result Communities and the corresponding Col;

xxii) End if;

xxiii) End if;

xxiv) n_c = |Communities|;

xxv) End while;

xxvi) Return optimal Communities with the maximum value of Col;

D. Mathematical Model

Let S be solutions of system

S = {I, O, P, Fc, Sc};

Where;

I = Input (Weibo and Facebook dataset);

O = Output (Recommendation of users);

P = Process

Sc = Success Conditions;

Fc = Failures Conditions;

Process

1) P1 = {I}; Read dataset of Weibo and Facebook dataset;
2) P2 = {P1}; Perform preprocessing to achieve data redundancy;
3) P3 = {P2}; Features extraction requires creating training and testing files. Features like followers count from Weibo dataset and friends count from Facebook dataset.
4) P4 = {P3}; Trust Clique algorithm;
5) P5 = {P4}; Perform classification algorithm to classify users into clod-start of existing users using C4.5 algorithm.
6) P6 = {P5}; Perform deep NN algorithm multi-perceptron model to provide accurate recommendation of users.
7) Sc = {success of system when appropriate users   is recommendation by MLP}
8) Fc = {Failures of system when wrong recommendation is provided to users}

IV. RESULTS AND DISCUSSION

A. Experimental Setup

The system is built using Java framework on Windows platform. The Net bean IDE is used as a development tool. The system doesn't require any specific hardware to run; any standard machine is capable of running this application.

B. Dataset

Here we used dataset weibo.com dataset of Chinese social media websites that is of which contains 500+ users' data and for comparison we also used facebook.com dataset which contains 500+users.

C. Experimental Result

Table II shows that comparison between existing system C4.5 algorithm and proposed system MLP algorithm. The proposed technique is more accurate than the existing techniques.

Table II. Accuracy comparison of c 4.5 and mlp algorithm

Algorithm	Accuracy in %
C4.5	85
MLP	86

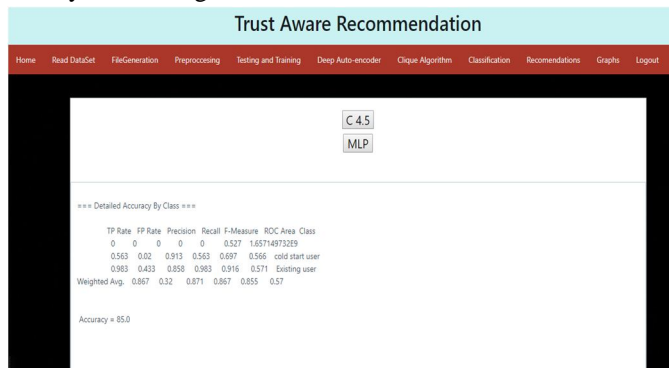The following Figure-3 shows the accuracy of C4.5 algorithm.

Figure 3: Accuracy of C4.5 Algorithm

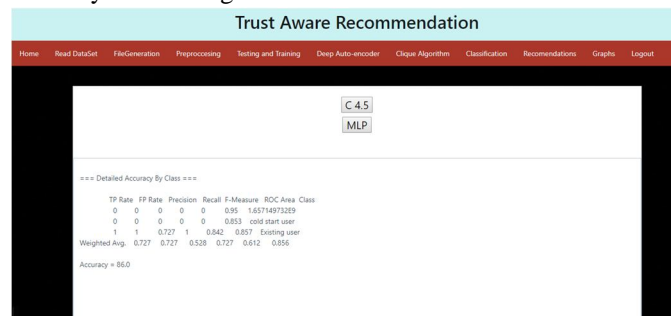The following Figure-4 shows the accuracy of MLP algorithm.

Figure 4: Accuracy of MLP Algorithm

Figure 5 shows the graph of accuracy comparison between existing system C4.5 algorithm and proposed system MLP algorithm. Result graph shows that the proposed system is more accurate than the existing system.
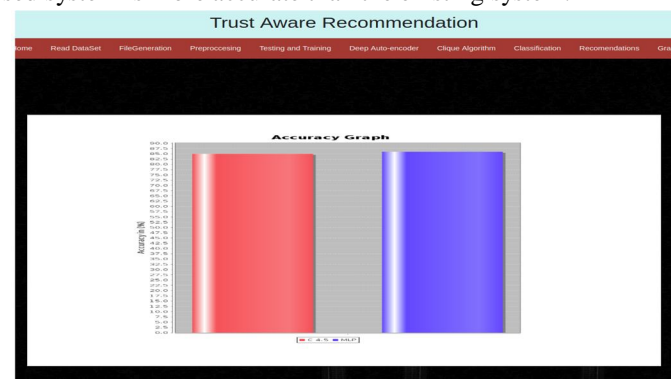

Figure 5: Accuracy Comparison Graph of C4.5 and MLP Algorithm

V. CONCLUSION AND FUTURE SCOPE

In this paper, we present trust aware recommendation in social networks based on Deep Learning (DL). Here, we use a deep auto encoder function to learn initial values of hidden user functions and elements first and then using the results obtained to train the minimization of the target function. A deep auto encoder is used to examine the initial values of the MF (Matrix Factorization) models in the first training phase after that it make the second training phase to ensure the best quality of recommendations. In this MF model, we consider the users characteristics and their trusted friends' recommendations. We provide prediction by using MLP algorithm; it recommends users by their interest.

REFERENCES

[1] F. Alqadah, Ju. Hu and H. F. Alqadah, "Biclustering Neighborhood Based Collaborative Filtering Method for Top-n Recommender Systems", Knowledge Information System, Springer, [2015].

[2] A. Javari and M. Jalili, "A Probabilistic Model to Resolve Diversity Accuracy Challenge of Recommendation Systems", Knowledge Information System, volume-44, no.3, [2015].

[3] Y. Rao, N. Zhang and H. Zou, "Adaptive Ensemble with Trust Networks and Collaborative Recommendations", Knowledge Information System, volume. 44, No. 3, PP. 663–688, [2015].

[4] Gu. Xu, Do. Wang and Sh. Deng, "Exploring User Emotion in Microblogs for Music Recommendation", Expert Systems with Applications, Volume-42, Page No.:9284–9293, Elsevier, [2015].

[5] Shui Guang Deng Jian Wu and Zhao Hui Wu, "Trust based personalized service recommendation a network perspective", JOURNAL OF COMPUTER SCIENCE AND TECHNOLOGY, Volume-29, Page No-69–80, Springer, [2014].

[6] S. Dhillon, Hsiang Fu Yu, and I. S. Dhillon, "Parallel Matrix Factorization for Recommender Systems", Knowledge and Information Systems, Springer, [2014].

[7] Ha. Park, M. Ishteva, and R. Kannan, "Bounded Matrix Factorization for Recommender System", Knowledge Information System, Springer, [2014].

[8] Shuiguang Deng, Guandong Xu and Longtao Huang, "Social Network Based Service Recommendation With Trust Enhancement"; Expert Systems with Applications, Volume 41, Issue 18, Pages 8075-8084, , Elesiver-15, [2014].

[9] Kuan Zhang, Ee Peng Lim, and David Lo, "Mining Indirect Antagonistic Communities from Social Interactions"; Knowledge and Information Systems, Volume 35, Issue 3, pp 553–583, Springer, [2013].

[10] Balazs Hidasi and Domonkos Tikk, "Initializing matrix factorization methods on implicit feedback databases"; Universal Computer Science, vol. 19, no. 12, [2013].

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)