International Journal of Innovative Research in Science and Engineering
Vol. No.2, Issue 09, September 2016
www.ijirse.com
ISSN (O) 2454-9665
ISSN (P) 2454-0663

KNOWLEDGE AND DENSITY BASED FUZZY-NEURO MODEL FOR COLLABORATIVE FILTERING RECOMMENDER SYSTEMS

P. Prabhu[1], M. Mohamed Sirajudeen[2], B. Balaji[3]

[1]*Assistant Professor – DDE / Research Scholar,*
Department of Computer Applications, Alagappa University, Karaikudi, Tamilnadu, India
[2]*Assistant Professor / Research Scholar, Department of Information Technology,*
Faculty of Informatics, University of Gondar, Ethiopia.
[3]*Student, Department of EEE, Velammal Institute of Technology, Chennai, India*

ABSTRACT

Recommender systems are one of the information filtering systems which provides business intelligence to the active users on e-commerce by more suggesting suitable items through personalized web page. User's explicit knowledge about the items, preferences, density of neighbours and recommendation criteria play vital role in providing accurate recommendations. This paper introduces knowledge and density based Fuzzy-Neuro model for personalized recommender system. Several experiments have been carried out to evaluate the performance of the model with different factors and evaluation measures. The experimental results show that the proposed model outperforms the conventional models.

Keywords *Data mining, DBSCAN clustering, e-commerce, fuzzy logic, neural network, personalization.*

I. INTRODUCTION

Due to the lot of e-business transactions over the internet leads to the huge volume of data with information overload problem. Searching of information from this large data is difficult to the active users on the web. Web mining is extracting meaningful information from World Wide Web. Web mining can be classified as web content, structure and usage mining. Personalization is one of the examples of web content mining. By using personalization, the web page with best suitable information for the particular active user in online can

ISSN (O) 2454-9665
ISSN (P) 2454-0663

constructed. Recommender systems are one of the business intelligence systems which provide methods for suggesting best matching items to the active users.

The major significance of this paper is to reduce the search time of items to the active users in e-business websites. Since this research work focus on knowledge of the user such as profile and preferences more accurate recommendations can be generated. The learning behaviour of neural network based models can handle large volume of user's data. The introduction of fuzzy logic in this model solves the uncertainty behaviour of users in online. This model is used to improve cross sale of business by recommending more items to the end-users. It also improves the loyalty of business by linking the relationship between user and recommender system and change the visitor of the websites become the buyers. This research work is also necessary to the IT professionals for designing and developing an efficient personalized webpage to the active users.

II. RELATED WORKS

Research in any field requires a highly structured review and study of related literature. A critical analysis of related literature will provide information on what has previously been done in the relevant area this will lead to new approaches and investigation.

S. Lee [1] investigated the effect of similarities Cosine and Mean Square Difference (MSD) and the number of co-rated items onto the recommendation performance. Further, it suggests a proper range of similarities and the ratio of the number of co-rated items yielding the best performance, through extensive experiments.

Wang, W. et.al. [2] proposed a new CF new information entropy-driven user similarity measure model to measure the relative difference between ratings. A Manhattan distance-based model is developed to address the fat tail problem by estimating the alternative active user average rating. As a result of the introduction of the new similarity measure and average rating estimation this approach provides better recommendations.

Zan Wanga, et.al.,[3] proposed a hybrid model-based movie recommendation system which utilizes the improved K-means clustering and genetic algorithms (GAs) to partition transformed user space. PCA data reduction technique is used to dense the movie population space which could reduce the computation complexity in intelligent movie recommendation as well. The experiment result shows that this approach can provide high performance in terms of accuracy, reliable and personalized movie recommendations when compared with the existing methods using Movielens dataset.

It was observed from the research literature that, these RS algorithms suffer from challenges and issues like scalability, sparsity, cold-start problem and accuracy and still needs improvements for producing more efficient recommendations. The identified lacking has been in the field of knowledge-based Collaborative Filtering (CF) recommendation as extension which suggest recommendations based on knowledge of active users, consumers and technology providers.

III. KNOWLEDGE AND DENSITY BASED FUZZY-NEURO MODEL (KBDFNM)

KBDFNM model recommends items based on knowledge about user's requirements and interests using fuzzy logic and neural networks. The user profile and preferences can be any knowledge structure that maintains this inference. It is a hybrid approach which combines the features of memory and model based collaborative filtering. This model use data mining techniques such as DB scan clustering and BPN classification to generate recommendations.

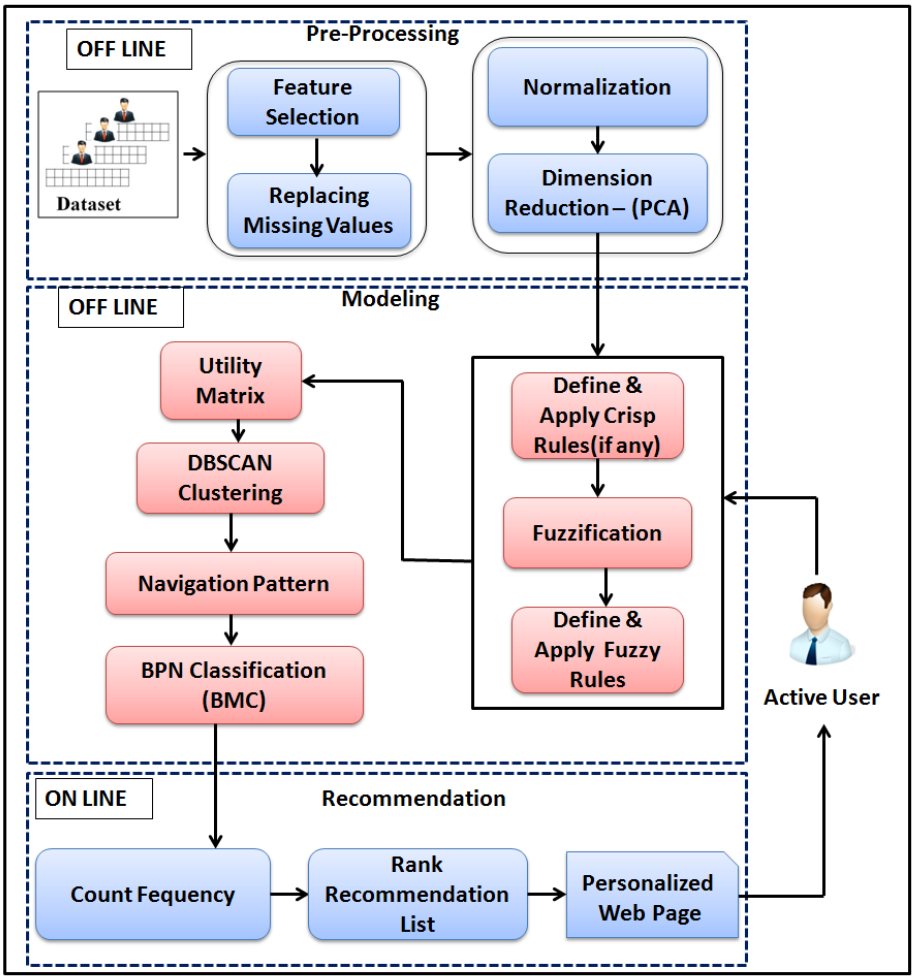

Figure 1. Knowledge and Density Based Fuzzy-Neuro Model Architecture (KBDFNM)

3.1 Methodology

The KBDFNM model works in three phases. First phase is pre-processing works in offline where feature selection, dimensionality reduction, missing values handling and normalization activities are carried out. In the second phase model is constructed offline. During this model construction phase the knowledge based fuzzy and crisp rules are defined and applied on utility matrix for information filtering. The DB Scan clustering method is applied to cluster filtered utility matrix to form neighbourhood of existing users. The active users Best Matching Cluster (BMC) is predicted using Back Propagation Neural Network (BPN). The business intelligence is obtained in third phase (recommendation) in on-line by identifying items from the BMC of users. These items are ranked according to its frequency count in descending order. The top-N items from this ranked list are selected as recommendation list.

3.2 Algorithm KBDFNM

Input: Training Dataset D and Test dataset TD.
The number of clusters *k*.
N = Potential number of recommendation.

Output: Recommendation List $\{I_1, I_2, \ldots, I_n\}$ of *Top-N* items.

// Phase I: Pre-processing

Select relevant features using Feature Selection.
Replace missing values using NN.(if any)
Perform Normalization.
Do Dimension Reduction using PCA.

// Phase II: Model Construction

Define KB crisp rules (if any) for user preferences and/or profiles.
//Formulate the fuzzy rules (FR).

- Define the linguistic variables and terms (initialization).
- Construct the membership functions (initialization).
- Construct the rule base (initialization).

Convert crisp input data to fuzzy values using membership functions.
// Fuzzification of attributes/user preferences.

$compute\ \mu_s(a_j)$ where $j = 1$ to m.

// Fuzzification of samples/user profile features.

$compute\ \mu_s(p_j)$ where $j = 1$ to m .

Define KB fuzzy rules for user preferences and/or profiles
Apply rules and generate the resultant dataset.
If New User **// Cold-Start Problem**

Register and Login.
Go to Step 2 of Phase III.
Else
Clustering of utility matrix using DBSCAN algorithm
//Predicting Best Matching Cluster (BMC) of active user.
For each Active user in TD do
Find the Best Matching Cluster (BMC) using RBPN.
End
TMAE ← Evaluate Matching Cluster
End If
//Phase III: Recommendation
For each Active user in TD **do**
Identify items from Best Matching Cluster (BMC) of users.
Calculate the frequency count and rank the items.
Select and recommend *top-N* items.
End

Algorithm Fuzzy-Neuro and Density Based Model (KBDFNM).

The time complexity of KBDFNM algorithm is O(n^2)+O(W^3)+ O(m^2+m) where n is number of train data, w is the weight and m is the number of test data.

IV. EXPERIMENTAL RESULTS

KBDFNM model has been experimentally simulated and evaluated using MovieLens 100k[10] and Jester[11] real-world datasets with different active users profile and preferences. The various factors with range of dimensions, clusters, *top-N* value, similarity and test/train size is considered to evaluate the performance of this model. To evaluate recommendation we use two metrics widely used in the information retrieval (IR) community namely recall and precision [4].

$$\text{Recall} = \frac{\text{size of hit set}}{\text{size of test set}} = \frac{|\text{test} \cap \text{top} - \text{N}|}{|\text{test}|}$$

$$\text{Precision} = \frac{\text{size of hit set}}{\text{size of top} - \text{N set}} = \frac{|\text{test} \cap \text{top} - \text{N}|}{\text{N}}$$

The standard F1 metric that gives equal weight to recall and precision and are computed as follows:

$$\text{F1} = \frac{2 * \text{Recall} * \text{Precision}}{(\text{Recall} + \text{Precision})}$$

We compute F1 for each individual active user and calculate the average value to use as our metric.

The results of KBDFNM model is compared with conventional k-Nearest Neighbour (k-NNBM) Approach[5][6][7][2] using MovieLens and Ant Based Recommender System (ARS) [8], k-nearest neighbour based Mean Squared Distance (MSD-CMB)[9] Jester datasets with commonly compared top-N value as 10. The Table 1 shows the comparison of Mean F1 measure between KBFNM model and conventional models using MovieLens dataset for *top-N*=10.

TABLE 1. Mean F1 of KBDFNM with conventional methods using MovieLens.

Algorithm	k-NNBM	k-NNBM(w)	GAC	k-NNBM(PCC)	KBDFNM
F1 Measure	0.44	0.53	0.60	0.66	0.77

When compared with k-NNBM, k-NNBM (w), GAC and k-NNBM (PCC) algorithms 11% increase in F1-measure is obtained using KBDFNM model. Therefore the performance of the KBDFNM is more significant than the conventional methods. The Table 2 shows the comparison of mean F1 measure between KBDFNM model and conventional models using Jester dataset for *top-N*=10.

TABLE 2. Mean F1 of KBDFNM with conventional methods using Jester.

Algorithm	ARS	MSD-CMB	KBDFNM
F1 Measure	0.49	0.68	0.89

The results listed in the Table 2 shows that, when compared with ARS and MSD_CMB algorithms KBDFNM model outperformed using Jester dataset. Therefore the performance of the KBDFNM is more significant than the ARS and MSD_CMB methods since it gives more recommendation accuracy measured in terms of F1 measure.

V. CONCLUSION

Recommender Systems (RS) play a important role in providing suitable recommendations to the end-users. In this paper a new knowledge and density based recommender systems model has been proposed. This method is experimentally evaluated using real-world datasets and compared their performance with conventional systems. This new KBDFNM model provides better accuracy in providing recommendations to the active user. The optimization techniques such as Particle Swam Optimization can be used to improve the performance.

ISSN (O) 2454-9665
ISSN (P) 2454-0663

REFERENCES

[1]. S.Lee, The Effect of Neighbour Selection in Collaborative Filtering Systems, Proceedings, Int'l Conf. Information and Knowledge Engineering | IKE'15, Dept. of Computer Education, Gyeongin National Univ. of Education, Anyang, Korea pp. 62-65, 2015.

[2]. Wang, W., Zhang, G. and Lu, J., Collaborative Filtering with Entropy-Driven User Similarity in Recommender Systems. Int. J. Intell. Syst., 30: 854–870. doi:10.1002/int.21735,2015.

[3]. Zan Wanga, Xue Yub,Nan Fengb, Zhenhua Wangc, An improved collaborative movie recommendation system using computational intelligence, Journal of Visual Languages & Computing Volume 25, Issue 6, Pages 667–675, December 2014.

[4]. Badrul Sarwar, George Karypis, Joseph Konstan, and John Riedl, Analysis of Recommendation Algorithms for Ecommerce,ACM,EC’00,October 17-20,2000.

[5]. Deepa Anand, Enhancing Accuracy of Recommender Systems Through Various Approaches to Local and Global Similarity Measures, PhD Thesis, Jawaharlal Nehru University, July 2011.

[6]. Hema Banati and Shikha Mehta, A Multi-Perspective Evaluation of MA and GA for Collaborative Filtering Recommender System, International journal of computer science & information Technology (IJCSIT) Vol.2, No.5, pp. 103-122,October 2010.

[7]. Panagiotis Adamopoulos and Alexander Tuzhilin, Recommendation Opportunities: Improving Item Prediction Using Weighted Percentile Methods in Collaborative Filtering Systems, ACM 978-1-4503-2409-0/13/10, RecSys’13, October 12–16, 2013.

[8]. P.Bedi, R.Sharma, H.Kaur, Recommender system Based on collaborative behaviour of Ants, Journal of Artificial Intelligence, ISSN 1994-5450,Asian Network for Scientific Information, 2 (2):pp.40-55,2009.

[9]. Symeonidis, P. Alexandros Nanopoulos, Apostolos N. Papadopoulos, Yannis Manolopoulos, Collaborative recommender systems: Combining effectiveness, Expert Systems with Applications, doi:10.1016/j.eswa.05.013, 2007.

[10]. The MovieLens dataset. Available: http://www.movielens.umn.edu.

[11]. The Jester data set. Available: http://goldberg.berkeley.edu/jester-data/.