ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

A New Personalized Product Recommender System Using Stacking and Memory Based Collaborative Algorithm

A.Venugopal[1], Nisha[2]

Assistant Professor, Department of Computer Science, Sree Narayana Guru College, Coimbatore, Tamil Nadu, India[1]

M. Phil Scholar, Department of Computer Science, Sree Narayana Guru College, Coimbatore, Tamil Nadu, India[2]

ABSTRACT: Online shopping is one of the major emerging trends among busy people. Nowadays social networks have better popularity and advent nature, so more and more users wish to share their opinions about their experience in the form of making reviews, ratings and blogs. Many Recommender System (RS) have followed the above factors for optimal suggestion to shop. Even though, the results are optimal and reliable, the system suffers from cold start and sparsity problem. Some of the social factors have been used in RS; still they have not been fully measured. The personalization can't be determined with only social similarity, this also need to be identified by their personalized attributes. While considering the other attributes and factors, this created a dimensionality problem. To overcome the challenges in the RS, the current work proposed a superfine model named as PRAS (Personalized Recommender and Alert System), which combines users profile, their intrapersonal and interpersonal interest similarity along with the interpersonal influence. This helps to overcome the huge cold start challenge and NP hard problem. This will be worthwhile for the user to select appropriate service from the large shopping domain. The PRAS identifies and analyzes user behavior to investigate the value of stable recommendations. And it utilizes the memory based collaborative filtering and an ensemble stacking algorithm for optimal RS in online shopping. As like RS, the AS performs the alert process, when the similar product found in the shopping domain. This leads the e-commerce domain to be more effective and reliable.

KEYWORDS: Data Mining, Recommender system, Collaborative System, Stacking, Online shopping,

I. INTRODUCTION

Data mining is the process of retrieving patterns and useful information from huge data. In the context of social networks, many data mining techniques are more applicable. The most used data mining techniques in the field of social networks are: prediction, evaluation and recommendation. With the enormous amount of data stored in databases, files and other repositories, it is increasingly important if not necessary to develop powerful means for analysis and perhaps interpretation of such data and for the extraction of interesting knowledge that could help in decision making. Data Mining is known as Knowledge Discovery in Databases (KDD) refers to the nontrivial extraction of implicit previously unknown and potentially useful information from data in databases, while data mining and knowledge discovery in databases are frequently treated as synonyms data mining is actually part of the knowledge discovery process.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

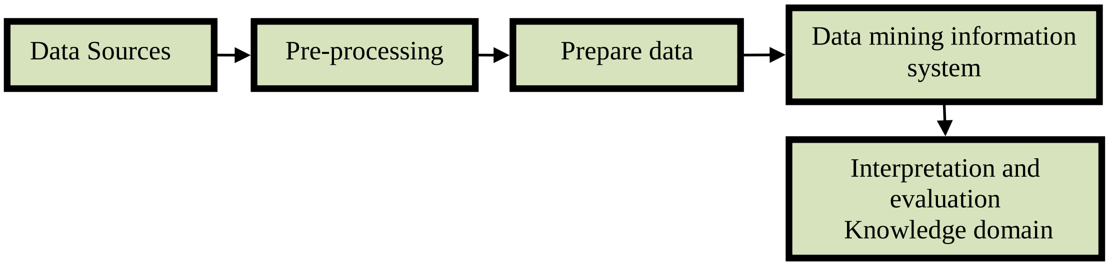

Fig 1.0 Steps in the Evolution of Data Mining

Data mining, the extraction of hidden predictive information from large databases is a powerful new technology with great potential to help focus on the most important information in their data warehouses shows in fig1.0. Data mining tools predict future trends and behaviors allowing users to make proactive knowledge driven decisions. The automated prospective analyses offered by data mining move beyond the analyses of past events provided by retrospective tools typical of decision support systems. Data mining techniques can be implemented rapidly on existing software and hardware platforms to enhance the value of existing information resources and can be integrated with new products and systems as they are brought online.

In the recent years, the Web has undergone a tremendous growth of Internet access, various types of social applications are available (e.g., online shopping, searching images, audios, and videos) have become ubiquitously available anytime [1]. But the amount of information available on the Internet has become immense and is still growing at an unbelievably fast rate. The emergence of online shopping networks (e.g., Amazon, Flipkart) has further boosted the volume of online information resources, since these technologies enable online users to freely view, buy and share products related to all types. On one hand, the abundance of online information may virtually guarantee that users are able to find what they are looking for. On the other hand, this same abundance also makes the useful information difficult to find, a problem referred to as "information overload". Two major Internet techniques, to wit, information search and recommendation, have been developed to help online users handle the information overload problem.

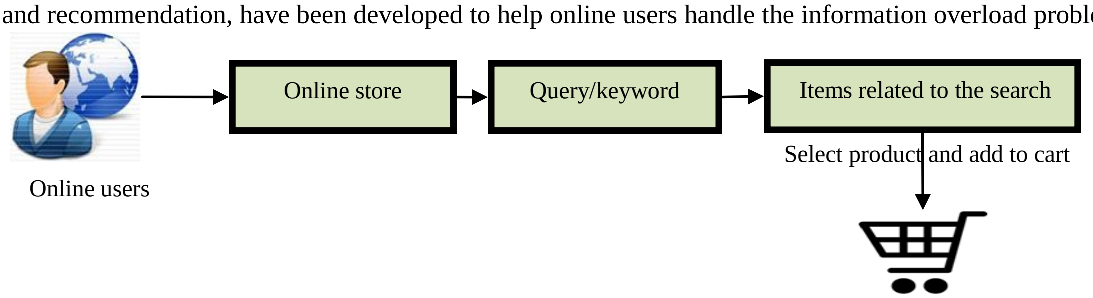

Fig: 2.0 online shopping processes

In the search case, illustrated in Fig 2.0, users actively express their information needs by submitting queries to the website (engine), and then the system tries to find the items in the collection that best match the queries and user selection.

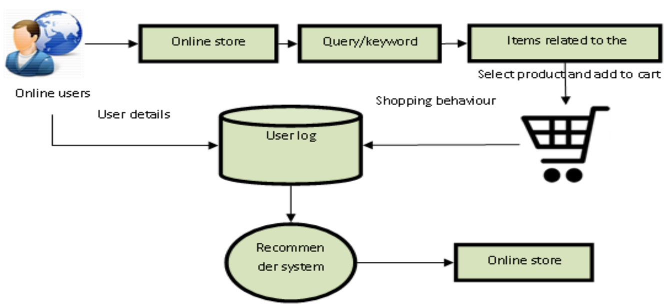

Fig: 3.0 online shopping process with recommender system

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

In the recommendation case, illustrated in Fig.3.0, recommenders intend to provide people with suggestions of products they will appreciate, based upon their past preferences, history of purchase, or demographic information.

II. RELATED WORK

A. RECOMMENDATION SYSTEM AND TECHNIQUES:

Recommendation systems (RS) apply data mining techniques and prediction algorithms to predict user's interest on information and products among the tremendous amount of available items. Recommendation systems are software agents that extract the interests and preferences of individual consumers and make recommendations accordingly. RS have the potential to help and improve the quality of the decisions consumers make while searching for and selecting products online. With the tremendous growth of e-commerce, the information overload problem created; due to this problem users are not able to effectively search items on the web. In the electronic world, RS has introduced the need for information filtering techniques that are use to help users by filter out information in which they are interested in.

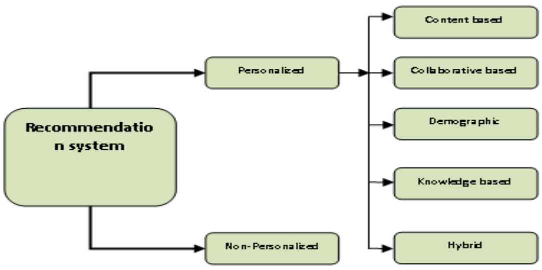

Fig: 4.0 recommendation system types

Recommendation systems are one of the approaches for the shopping recommendation which is based on providing possible items of interest to a user instead of the user to go searching for them. RS changed the way as the websites communicate with their users. Instead of providing a static feel for the users, in shopping items searching, this provides potential suggestions which increases communication to provide a higher experience. RS recognize recommendations autonomously for individual users based on past browsing history, profiles, rating and other reviews given to item and this also considers the other users behavior too. List of recommendations in given in fig 4.0. The branches of RS are described below.

i. PERSONALIZED RECOMMENDATION:

Personalized recommendation enables the online customization or suggestion of data in any format that is relevant to each and every user based on the users implicit and explicit behavior and tastes and explicitly provided details [2]. Personalized recommendation engines are classified into five types depends on their approach to recommendation [3]:

1. Content-Based Filtering:

Content-based recommendation method is based on the information about item content and ratings a user has given to items. This technique combines these ratings to profile of the user's interests based on the features of the rated items. The recommendation engine then can find items with the preferred in the past as illustrated in Fig 5.0. The recommendations of a content-based system are based on individual information and ignore contributions from other users.

 DOI: 10.15680/ijircce.2015. 0308098

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

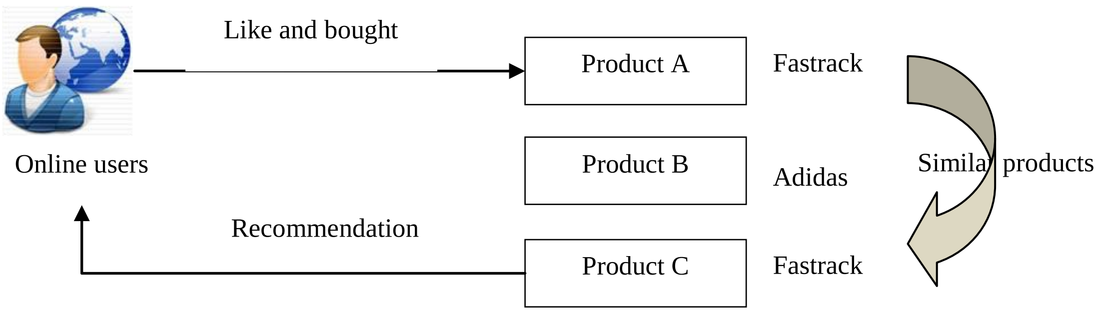

Fig. 5.0 Content-Based Recommendation

2. Collaborative Filtering:

Collaborative filtering technique based on users history in the form of rating given by the user to an item as their information source [5]. It can be accomplished by making relation between the users or between items. Collaborative filtering is categorized into three types: user-based, item-based, model-based. User-based Approach makes recommendation based on the interest of the user having the similar taste. It correlates user as per the rating given to the items. From the Fig. 6.0, first user related to third user instead of second because the rating given by third user is quite similar to the first one. That's why item 3 is recommended to the user as it's the only remained item.

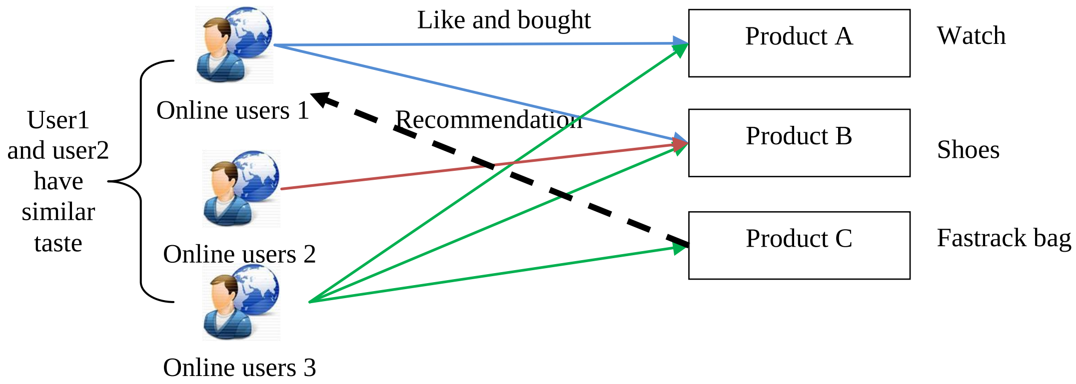

Fig. 6.0 User-Based Collaborative Filtering

3. Item-based:

Item-based Approach is based on the items as the user rated items similarly are probably similar. From Fig. 7.0, 2nd and 3rd user rated item 1 and 3 so it assumes that item 1 and 3 are become similar. As 1st user like item 1, item 3 is recommended.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

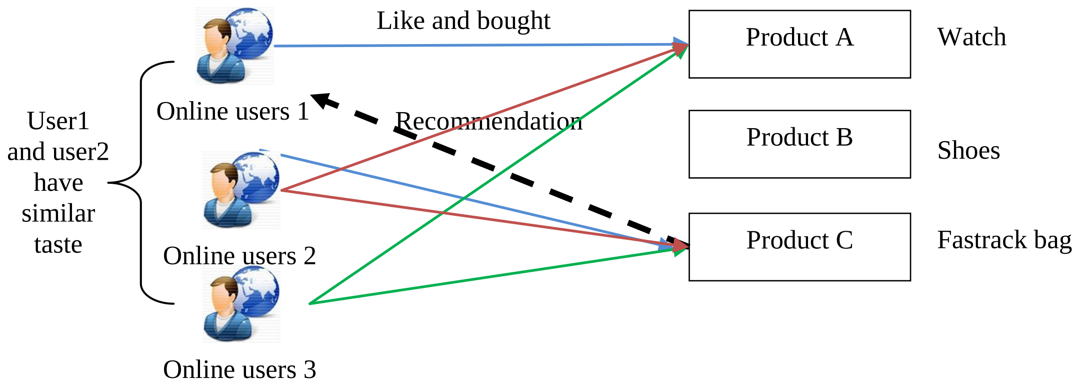

Fig. 7.0 Item-Based Collaborative Filtering

4. Demographic:

Demographic recommendation technique uses information about user only. The demographic types of users include gender, age, and knowledge of languages, disabilities, ethnicity, mobility, employment status, home ownership and even location. The system recommends items according to the demographic similarities of the users.

5. Knowledge-Based Filtering:

Knowledge based recommendation system is based on the explicit knowledge about item classification, user interest and recommendation standard (which item should be recommend in which feature) [6]. It is an alternative approach to the collaborative filtering and content-based filtering. 5. Hybrid approach Hybrid approach is a combination of all above types [8].

III. PROPOSED SYSTEM

Our proposal provides a study of developing an optimal RS and AS based on inters personal similarities. Our proposal introduces a new superfine algorithm which combines the Machine learning and Recommender system to overcome the service selection challenge in online shopping domain, where creates the service recommendation and similarity measure problems. Our system initially analysed the drawbacks of the existing RS and develops two different approaches based on collaborative filtering and clustering algorithm and the other one is Alert System, which is for capturing and suggesting the optimal service from large user log set. And finally the system applies a new memory collaborative filtering technique to eliminate ineffective products from the result set.

Contributions:

We proposed a personalized recommendation system which combining user profile, intrapersonal and interpersonal interest similarity along with the interpersonal influence. The thing of user individual interest makes associations between user and product with hidden features. The followings are the contributions of the proposed work.

- Identifying and analyzing user behavior to investigate the value of stable recommendations.
- The proposed RS utilizes interpersonal and intra personal similarities from the user log.
- Creates a memory based collaborative filtering for online shopping product recommendation.
- Identifying user's usage patterns and how many users accepted the recommendations.
- Appling a new recommender system with computational overhead reduction.
- This has be applied an ensemble approach called stacking. Stacking (sometimes called *stacked generalization*) involves training a learning algorithm to combine the predictions of several other learning algorithms.

The main contributions of this paper are summarized above. Such community factors make connections between user and other user using the same hidden feature vectors. So this proposes a Recommendation and alert approach by enforcing user personal interests. Identifying personal unique interest is modeled to get an accurate model for the above mentioned cold start problem, here the cold start user and user with very few friends and rated items can able to get the

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

appropriate suggestion and recommendations. The influence of the hidden factors in the recommendation system with performance metrics has been modeled.

IV. P-RAS APPROACH

The chapter explains about the proposed **P-RAS** approach on personalized shopping product recommendation based on the **P-RAS** model, which follows a combinatorial recommendation strategy and has the ability to combine many possible constraints that exist in the real-world scenarios. The probabilities associated with various state changes are called **transition probabilities.**

P-RAS are used for combinatorial optimization in which an optimal solution is required over a discrete search area. This helps to overcome the problem of decision making problem where the search-space of product selection grows faster than exponentially as the size of the user log grows, which makes an extensive search for the optimal solution infeasible. Several systems suffered from decision making and cold start problem, where set of products and user log may result in several multidimensional combinatorial problems.

A. Improved Stacking ensemble Algorithm:

Stacked algorithm is an improved method of using a high-level model to combine lower level models to attain maximum predictive, replacing accuracy. A stacking algorithm is a mapping taking as follows; {a training set of s pairs $\{x_k \in T\ n\ ,\ y_k \in T\}\ 1 \leq k \leq S$ together with a question ∈ T n } into {this ∈ T}. (Full generality would have predict the rating R ∈ T s not T. this can predict, replace and generalize rating guesses in R m with the CF model of m separate generalizes making guesses in R.

- **Mean rule:** $\mu j(\mathbf{x})=1T\sum Tt=1dt,j(\mathbf{x})$
- **Sum rule:** $\mu j(\mathbf{x})=\sum Tt=1dt,j(\mathbf{x})$ (provides identical final decision as the mean rule)
- **Weighted sum rule:** $\mu j(\mathbf{x})=\sum Tt=1wtdt,j(\mathbf{x})$ (where wt is the weight assigned to the tth classifier ht according to some measure of performance)
- **Product rule:** $\mu j(\mathbf{x})=\prod Tt=1dt,j(\mathbf{x})$
- **Maximum rule** $\mu j(\mathbf{x})=\max t=1,\cdots,T\{dt,j(\mathbf{x})\}$
- **Minimum rule** $\mu j(\mathbf{x})=\min t=1,\cdots,T\{dt,j(\mathbf{x})\}$
- **Median rule** $\mu j(\mathbf{x})=\mathrm{med}t=1,\cdots,T\{dt,j(\mathbf{x})\}$
- **Stacking mean rule** $\mu j,\alpha(\mathbf{x})=(1T\sum Tt=1dt,j(\mathbf{x})\alpha)1/\alpha$
  - $\alpha\rightarrow-\infty\Rightarrow$ Minimum rule
  - $\alpha\rightarrow\infty\Rightarrow$ maximum rule
  - $\alpha=0\Rightarrow$ Geometric mean rule
  - $\alpha=1\Rightarrow$ Mean rule

The improved stacking algorithm in RS involves with the following steps, this initially performs the splitting process, which splits the training dataset into n number of disjoint datasets. Here the user log details will be segmented according to their profile. The next step is training process, which involves the learning process from base learners.

Algorithm:

Input: user log, user profile, product details

Output: predicted ratings and suggestion

Steps:

1. Split the training set T into two disjoint sets n[3]=S.split(T).
2. Train ti several base learners P on the first part n[1].
3. Test the base learners ti=P(t) on the second part.
4. Using the predictions from step 3 as the inputs, and the correct responses as the outputs, train a higher level learner.

In P-RAS other user's interest and log will be taken as training data for another user. After the successful training phase the system performs the test phase of individual interest measures, all suggestions and recommendations are taken in this phase. Finally this predicts the rating and user interested products with high correlation.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

The basic principle of **P-RAS** is to pursue Neighboring Solutions whenever it encounters a local optimum which is known is interpersonal similarities. And performs returns back to previously visited solutions, which is prevented by the temp memories which is known as RS list. priority list helps to record the recent history of the product search by the user.

B. User Interest Factor

The similarity value calculation between users in the same category by means of interpersonal and intra personal as well as product based is important task in P-RPS. The system proposed the effectiveness of P_RPS model with consideration of individual preference, interpersonal influence and intra personal influence. The system considers the independence of user interest in the e shopping domain. It means this can recommend items based on user interest at a certain extent this also utilizes user's association with the items to train the hidden feature vectors in stacking algorithm, especially for the existing and more expecting users. The system also considers the Interest circle inference technique. As per the stacking algorithm this segments the social network into several sub-networks and each of them correspond to particular item collection. To overcome the cold start users who has a few rating records from the ulog then the ratings of their associated user's interest in the same category to link user interest products.

C. Individual Interest measure:

Due to the individuality especially users with huge rating records in e shopping domain, users usually choose products all by themselves with little influence others. To provide the product recommendation without affecting individuality for experienced users, the system proposed an optimal personalized recommendation system. The significance of user and item depends on the relevance of user interest *Tu* and item topic *Ti* to a certain domain this takes several attributes such as product category, company name, price and offers. This denoted the relevance of user T's personal interest to the category of item *i* in the RAS model by *RASu,i*

$$RASu,i = Sim(Tu,Ti).$$

This performs the similarity measure by fine filtered attributes. The derived products should satisfy the personal interest and as well as social influence without affecting their attribute consideration.

For each attribute in the attribute list (Ai) € sim (user1, user2, Ai)

From the above equation the similarity has been calculated between two users with the consideration of any particular attribute such as price, i.e. this measures whether the average price spent by user 1 is similar to user2, based on this the system suggests products to user 1 from user 2 log.

V. RESULTS

In this section measure the performance of the each iteration then measure the results the overall time of the P-RAS. Shopping recommendation accuracy is evaluated by comparing the packages which assigned by the similarity measures. In this section, through statistic analysis, the system finds the average time taken for each process in P-RAS. This is initially calculates the individual times such as profile based, interpersonal based and intra personal based. This shows the impact of the amount of user information (user's number of rated items and number of friends) to the accuracy of the proposed model and compared models in **e-shopping.**

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

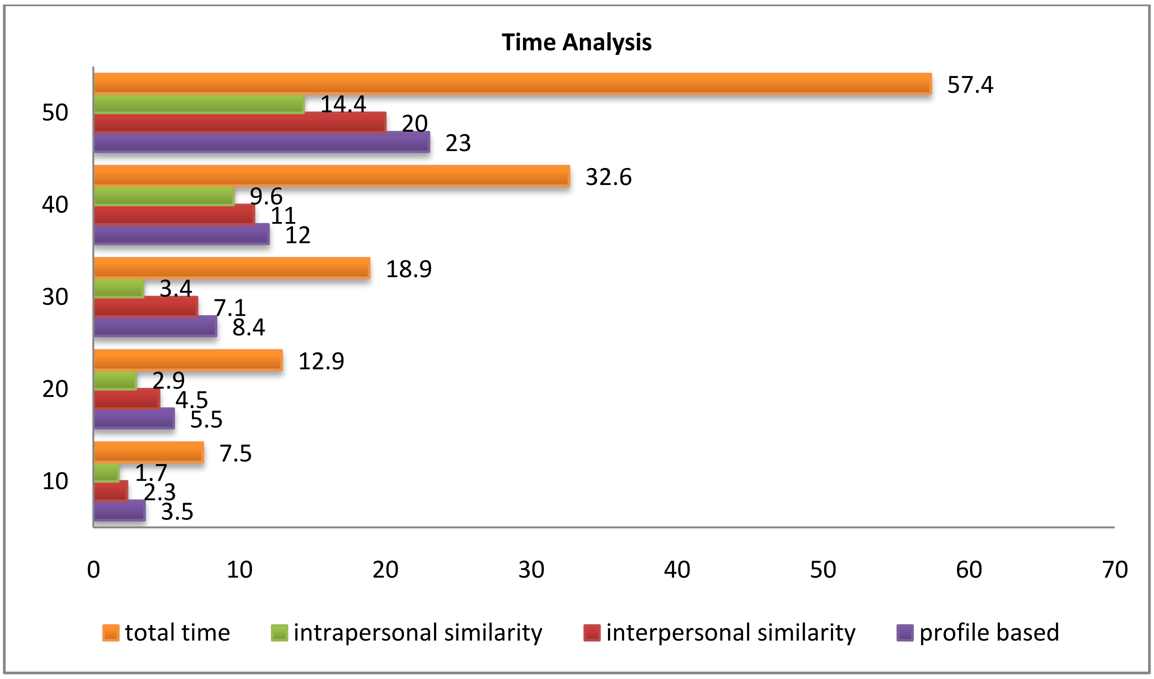

Fig 8.0: Time analysis

From the above chart fig 8.0 the system shows the time taken at each stage of P-RAS. The number of product is specified from 10 to 50. For each process, the time calculated and finally the system provides the overall time taken by the P-RAS . This shows for 5 Products the system takes 57.4 mille second time.

VI. CONCLUSION AND FUTURE WORK

A personalized recommendation and alert approach has been proposed by combining several real time factors known as profile based, personal interest based, interpersonal interest similarity, intra personal similarity and product influence with various attributes. This overcomes the cold start and recommendation problem, which is common in recommender system. To overcome the above issues, the system implemented improved stacking ensemble algorithm with memory based collaborative filtering techniques. This approach is implemented in an e-shopping dataset with relevant similarity measurement phases. This measures individuality of rating items with the reference of experienced users with various factors. At present the personalized recommendation model in the literature only takes interpersonal relationship and user's historical rating records. In the proposed P-RAS, the system takes the area information and other attributes to recommend more personalized and real-time items to the users.

REFERENCES

[1] X.-W. Yang, H. Steck, and Y. Liu, "Circle-based recommendation in online social networks," in *Proc. 18th ACM SIGKDD Int. Conf. KDD*, New York, NY, USA, Aug. 2012, pp. 1267–1275.
[2] M. Jiang *et al.*, "Social contextual recommendation," in *Proc. 21st ACM Int. CIKM*, New York, NY, USA, 2012, pp. 45–54.
[3]] M. Jamali and M. Ester, "A matrix factorization technique with trust propagation for recommendation in social networks," in *Proc. ACM Conf. RecSys*, Barcelona, Spain, 2010, pp. 135–142.
[4] R. Salakhutdinov and A. Mnih, "Probabilistic matrix factorization," in *Proc. NIPS*, 2008.
[5] [5] Y. Koren, R. Bell, and C. Volinsky, "Matrix factorization techniques for recommender systems," *Computer*, vol. 42, no. 8, pp. 30–37, Aug. 2009.
[6] M. E. Tipping and C. M. Bishop, "Probabilistic principal component analysis
[7] M. Jahrer, A. Toscher, and R. Legenstein. "*Combining predictions for accurate recommender systems*". KDD'10, pp. 693-702, 2010.
[8] H. Ma, I. King, and M. R. Lyu. "*Learning to recommend with social trust ensemble*". SIGIR'09, 2009.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

[9] X. Yang, Y. Guo and Y. Liu. "*Bayesian-inference based recommendation in online social networks*". The 30th Annual IEEE International Conference on Computer Communications (INFOCOM), 2011.

[10] J. Huang, X. Cheng, J. Guo, H. Shen, and K. Yang. "*Social recommendation with interpersonal influence*". In Proceedings of the 19th European Conference on Artificial Intelligence (ECAI), pp.601–606, 2010.

[11] L Sharma, A Gera "*A Survey of Recommendation System: Research Challenges*" International Journal of Engineering Trends and Technology, 2013

[12] GS Sadasivam "*A Survey on Personalized Recommendation Techniques"*

[13] P. Cui, F. Wang, S. Liu, M. Ou, S. Yang, and L. Sun. "*Who should share what? item-level social influence prediction for users and posts ranking*". SIGIR, pp. 185-194, 2011.

[14] H.R. Kim and P.K. Chan. "*Learning Implicit User Interest Hierarchy for Context in Personalization*". In Proc. of International Conference on Intelligent User Interface, 2003.

[15] O.Phelan, K. McCarthy, and B. Smyth. "*Using twitter to recommend real-time topical news*." In Proceedings of the third ACM conference on Recommender systems, pp. 385-388. ACM, 2009.

[16] Y Chen, C Wu, M Xie, X Guo "*Solving the Sparsity Problem in Recommender System*

[17] H Ma, D Zhou, C Liu, MR Lyu, I King. "*Recommender Systems with Social Regularization*". Fourth ACM international conference on Web search and data mining, 2011.

[18] A. Rahman, and S. Hailes, "*Supporting trust in virtual communities*".In Proceedings of the 33rd Hawaii International Conference on System Sciences. Maui, HW, USA (2000).

[19]] A Bellogín, P Castells, I Cantador. "*Self-adjusting Hybrid Recommenders Based on Social Network Analysis"*. 34th international ACM SIGIR conference on Research and development in Information Retrieval, 2011.

[20] Xavier Amatriain "*Mining Large Streams of User Data for Personalized Recommendations*" ACM SIGKDD Explorations, 2012.