Opinion Based Memory Access Algorithms using Collaborative Filtering in Recommender Systems

J. Sangeetha, V. Sinthu Janita Prakash

Abstract –In recent years, the online shopping and the online advertisement businesses is growing in a vast way. The reason behind this growth is, the peoples are not having sufficient time for go for a shop. Without seeing the quality of the product directly, the people are ready to buy the product by seeing the other user recommendation of the particular product. This leads an interest / the need to develop the researcher an innovative recommendation framework. Based on the opinion prediction rule, the huge size of words and the phrases which are presented in the unstructured data is modified as a numerical values. The sale of the particular product in an online shopping is depends on its description of the quality, the review of the customer. Based on the positive and negative polarity, an Inclusive Similarity-based Clustering (ISC) is proposed to cluster the extracted related keywords from the user reviews. To evaluate the strength, weakness of the product, estimate the respective features, as well as the opinions, the Improved Feature Specific Collaborative Filtering (IFSCF) model for the feature with aspect opinion is proposed. Finally the complete feedback of the product is estimated by propose the Novel Product Feature-based Opinion Score Estimation process. The main challenge in this recommendation system is the fault information estimation of the reviews and the unrelated recommendations of the bestelling or the better quality product. To neglect these issues, an Enhanced Feature Specific Collaborative Filtering Model based on temporal (EFCFM) is proposed in the recommendation system. Hence the developed EFCFM method is investigated by comparing along with the existing methods in terms of subsequent parameters, precision, recall, f-measure, MAE and the RMSE. The outcome shows that the developed EFCFM approach predicts the best product and produce the accurate recommendation to the customers.

I. INTRODUCTION

A lot of review based recommender framework have been invented, along with the aim of integrating the truthful data in customer generated text base reviews into the user to process the recommendation. (Chen, Chen, & Wang, 2015) Generally the improved text analysis as well as the opinion mining systems are engage the different types of extraction for review the textural elements, for example topics which are all taken for the discussion, opinions in the multi-faceted nature, contextual basis information, relative opinions, and it includes the emotions of the reviewers. (Ravi & Ravi, 2015) Now a days, among the word wide the peoples are very eager to show and convey their feedback regarding on the daily routines and the world issues on web. This is because of the development and the vast usage of the social media.

Thus, the social media is provided the immense contribution and also act as the transparent platform to express the reviews among the world. To permit the customer to express their point of opinion, (Ravi & Ravi, 2015) the electronic Word of Mouth (eWOM) statements which are presented in the web are established in the service industry as well as the business. In the last decade, the researchers, academician, peoples and service businesses are occupied majority place in the sentiment analysis, also termed as the opinion mining, which is utilized to extract and can investigate the public moods and the reviews.

In different kinds of online shopping services, the importance of the recommendation is well being recognized. From the experienced users, the explicit and the implicit feedbacks are extracted by the use of mostly used recommendation algorithms. Every day, in online shopping the new buyer is joining. He/ she usually not having the explicit ratings as well as the slight implicit behavior. At the time of facing the problem of new user recommendation, the organized systems may enquire the buyer to intimate the priority on the basis of particular features of the product for example, the brand of the camera, cost, resolution, etc. Though, the restriction of existing works is that the determination is required from the side of the buyer should be more. Furthermore, the more number of buyers are not ready to express their whole priority because of the insufficient knowledge about the domain of the product, even though they have participated in the conversation with the system. The main issue is the way of prediction of the missing preferences on unspecified features of the buyer’s textual content. The existing classical model-based and the memory-based algorithms in the collaborative filtering are not construct the collaborative relations between the users and without the feedback of the users. Usually, the content based approaches, also lacks in the accurate finding of the similar products when the preference of the users are only specified about the subset of features.

The main objectives of this proposed approach are listed as follows:

- To improve the clustering model using feature specific collaborative filtering techniques
- To estimate the Score based on the keywords in review and the features of the product.
- To extract the features based on the strength and weakness of the features.
- To estimate the product strength and weakness score of the product and categorize the product
- Review is recommended based on the Time and Location.

Revised Manuscript Received on October 05, 2019.
Ms. Sangeetha, J. Research Scholar, Dept. of e.sc, Cauvery college for Women, Bharathidasan University, Trichy, India. Email: sangeetha@research@gmail.com

Dr. Sinthu Janita, Professor, Dept. of e.sc, Cauvery college for Women, Bharathidasan University, Trichy, India.
Opinion Based Memory Access Algorithms using Collaborative Filtering in Recommender Systems

To estimate the Score based on the keywords in review and the features of the product.

To extract the features based on the strength and weakness of the features.

To estimate the product strength and weakness score of the product and categorize the product aspect and its opinions by using the Improved Feature Specific collaborative Filtering Model.

To evaluate the average score of product strength and weakness and shows the better recommendation based on a Novel Product Feature-based Opinion Score Estimation process.

The paper is organized as follows: The detailed description of the related works regarding recommendation system, filtering techniques, and the rating prediction technique is discussed in section II. The implementation process of the Innovative Product Recommendation Framework (IPRF) for proficient opinion prediction and rating estimation, the implementation process of an Enhanced Feature Specific collaborative filtering Model (ECFCM) for web service recommendation in section III. The comparative analysis of proposed ECFCM technique with prevailing approaches and qualified investigation of the proposed IPRF model is compared. Lastly, section V gets concluded the proficiency of the proposed IPRF approach and the proposed ECFCM approach.

II. RELATED WORKS

In this section, the review of the various recommendation system, filtering technique, and the score estimation methods are presented with their merits and demerits. (Chen et al., 2015) provided a complete summary of the way of exploitation of the review elements to improvise the level of recommendation based on content, collaborative filtering, and the ranking approach of product based on the standard content. By two various types of the existing surveys are classified such as review based user profile building and review based product profile building. The Multi-faceted opinions are additionally to develop the weight/value priorities that users place on the specified aspects. The advantages of each branch of work were deliberated by means of development of algorithm and the evaluation procedure of proposed algorithm. Additionally, the various development according to this survey in future, may encourage the researchers to work related to this area. (Ravi & Ravi, 2015) Presented the report on the sentiment analysis. This report have been portrayed the reviews which are presented more than one hundred published articles in last ten years. These articles are related to the required tasks, methods, and the sentiment analysis application. There was a need of various sub-tasks for performs the sentiment analysis, for accomplish the different way of methods and approaches. These surveys have been organized on the sub task basis which has to be performed, natural language processing methods, sentiment analysis applications, and the machine learning. This survey also included the open problem and the comparison table of huge number of articles. (Liu, Liu, Zhang, Kim, & Gao, 2016) proposed an unsupervised innovative method to create a chief development. The method was based on the lifelong learning context as well as the implementation along with two procedures of recommendations which are mainly established based on the similarity of semantic and the feature relations correspondingly. The information was exposed from the huge set of review reports. The experimental outcomes have been shown that the overwhelmed performance of AER. For the future work, the author considered to discover the way of recommendation has to be used, by the aim to develop the various extraction approaches. The author concluded that there is a possibilities to select the various extraction algorithm in AER. (Wang, Pan, & Chen, 2013) proposed a MoBRec, which was an innovative two-stage user similarity based approach. This recommendation has combined the data which was provided by the gradually varying aspects of the mobile framework and also the implicit reviewer feedback revealing the preferences of users. (Hu, Liang, Kuang, & Honavar, 2018) The MoBRec was focused on the cluster which was containing the users those who are searching the most relevant topics by means of related features and the implicit responses. This MoBRec also have been employed an innovative approach for the strong approximation of the users preferences. In the scalability and the recommendation quality, the MobRec outperformed when comparing with the existing methods. (Vagliano, Monti, & Morisio, 2018) proposed an innovative recommendation method established based on the semantic annotation of objects which was stated in the user reviews and also the information presented in the data on web. In this proposed work, the author was compared recommender system along with two baseline algorithms and also with the existing Linked Data based method. Additionally our proposed system have been provided added various recommendations regarding the former methods in terms of attaining the superior precision over the Linked Data based approach. (Eirinaki, Gao, Varlamis, & Tserpes, 2018) focused on large-scale recommender structures which was taken the benefit of the characteristics. These characteristics are fundamental of social network and have been focused on the variability as well as the instability of bonds in social. Additionally challenge the problems associated in the size and change of speed in social graphs, scalability testing of existing recommender systems and also presented the solutions which was taken this proposed recommender systems to the next stage. (Kuang et al., 2018) designed a framework for CPS services, which was used for the prediction of QoS. The author also proposed a modified QoS prediction method which was based on the repetitive and the collaborative filtering based on location. This approach estimated the user’s reputation by using the Dirichlet probability distribution. This distribution has been used to identify the untrusted users and progress their variable information. These have been used to improve the estimation of similarity users and services. At last, the information from the users those who are presented in the geographical neighbors. These services are attached to calculate the unknown QoS standards. The experiments have been carried out by using the real time datasets that shown that this approach overwhelmed with the other existing approached by means.
of robustness, effectiveness and the precision. (Yao, Wang, Sheng, Benatallah, & Huang, 2018) proposed a probabilistic matrix factorization method along with the regularization implicit correlation. This approach have been used to solve the issues which are associated in the recommendation and also aimed to enhance the recommendation multiplicity. To uncover the latent correlations among APIs, a latent variable model have been developed by examining the co-invocation patterns. Further explored the relations of groupings the method. Also demonstrated the efficiency of the method by accompanying the wide ranging on a real dataset crawled from Programmable Web. (Li, Cao, Wu, Huang, & Buyya, 2018) proposed an innovative optimization approach named the discrete quantum inspired shuffled frog leaping (DQSFL) algorithm. These have been based on the quantum information theory and the shuffled frog leaping algorithm. According to the co-evolution of the quantum frog colony, the DQSFL algorithm utilized the quantum movement equations to search for the finest location. The experiments have been demonstrated that the CF recommendation approach according to the DQSFL which was effectively solved the data on rating. The sparseness issues in the resemblance estimation approach to improvise the precision of the rating score prediction, and also it was provided the superior recommended outcomes than the existing CF algorithms.

III. IPRF: INNOVATIVE PRODUCT RECOMMENDATION FRAMEWORK

This division deliberates the execution particulars of the proposed Innovative Product Recommendation Framework (IPRF) for achieving an accurate recommendation about the product reviews via the utilization of collaborative filtering along with the opinion score approximation method. The workflow of the proposed IPRF is shown in Fig. 1.

### Fig.1 Over all Frame Work

Primarily, the data is extracted in the native form and are loaded for processing; here the data is the customer product reviews. By applying POS tagger, slicing, stop word removal, and the normalization, the reviews are pre-processed. From the each and every user review, the keywords are extracted. Based on the Inclusive Similarity based Clustering (ISC) method, the extracted keywords are clustered. Then the user reviews are clustered as groups according to the polarity. Additionally, the Improved Feature Specific Collaborative Filtering (IFSCF) models are developed for the purpose of group the user reviews as feature specific clusters. Then the detection of polarity is carried out and based on the Novel Product Feature based Opinion Score Estimation algorithm, the overall opinion of the products is determined. Lastly, the user query is pre-processed and the user’s requirement is analyzed with the help of extracted keywords. The overall opinion about the product and the feature-specific opinion of the product is specified as the query result.
A. Dataset
In this work, the customer product reviews composed in the field of opinion mining, sentiment analysis, and detection of opinion spam dataset from the social media is utilized for attain the effective product opinions. In the Customer Review Dataset (CRD) is included of five kinds of product reviews for example digital camera 1, digital camera 2, phone, MP3 player, and the DVD player.

B. Preprocessing based Document Clustering
The preprocessing is a very significant step since it can improve the accuracy result of a Clustering based Inclusive Similarity (ISC) algorithm. Therefore, it is essential to preprocess the data reasonably. The following steps are performed in the preprocessing. They are,
- Stop words removal
- PoS Tagging
- Slicing
- Rank of each review extraction from the dataset
- Normalization

C. Preprocessing based Document Clustering
The preprocessing is a very significant step since it can improve the accuracy result of a Clustering based Inclusive Similarity (ISC) algorithm. Therefore, it is essential to preprocess the data reasonably. The following steps are performed in the preprocessing. They are,
- Stop words removal
- PoS Tagging
- Slicing
- Rank of each review extraction from the dataset
- Normalization

The mathematical derivation of normalization as,
$$Nor' = \left( \frac{Nor - Nor_{min}}{Nor_{max} - Nor_{min}} \right) \times (q - p) + p \quad (1)$$

Where,
- $Nor'$ represents min-max normalized data one
- $(Nor_{max}, Nor_{min})$ represents the pre-defined boundary
- $(p, q)$ represents the new interval
After that, validate the extracted keywords with the SentiWordNet database which shows two numerical scores ranging from 0 to 1, each indicating the synset’s positive and negative bias polarity. The mathematical derivation of cumulative distance measure as,
$$CD = dist (i) \times dist (j) \quad (2)$$

Where $i$ and $j$ represent 1 to $n$ of each review. Then, estimate the inclusive similarity as,
$$IS = cum.dis (i, j) \times n / dist (i) \times dist (j) \quad (3)$$

D. Memory Management Model
To check the similarity of the user query along with the cache memory the memory management model (MMM) (Sangeetha, Prakash, & Bhuvaneswari, 2017) is utilized. If it’s similar, then the results is retrieved from the cache memory otherwise it retrieve the result from the knowledge database. The execution time and the process time are minimized based on the optimal memory access.

E. Improved Features Specific Collaborative Filtering (IFSCF)
The proposed Improved Feature Specific Collaborative Filtering (IFSCF) model for the feature specific clusters are used to evaluate the product strength and weakness and predict the corresponding aspects and its opinions. To obtain the overall opinion score and the recommendation initialize the $A_{Op}$ and $U_{Q}$. Then the $U_{Q}$ requested query is preprocessed to extract the relevant query and remove the unwanted informations. Then the extracted features is characterised into features and its view to categorize the positive features and the negative features. Then evaluate the polarity score of positive and negative features to retrieve the rank of the scores. At last, attain the strength and weakness score of the product and then express the proposed product is recommend or not.

F. OPINION SCORE ANALYSIS BASED RECOMMENDATION
Finally, the opinion score based product recommendation is carried out after calculating the score values for both the user review and product property. Here a product is recommended based on the opinion score based analysis.

II. INNOVATIVE PRODUCT Recommendation Framework (IPRF)
Algorithm:
Input : Customer Review
Output: Recommendation
Step 1: Apply Preprocessing
Step 2: Calculate distance between reviews
Step 3: Construct clustering for the positive and negative reviews.
Step 4: Extract Features in the reviews & assign Index
Step 5: User will give the query about product
Step 6: Query will be preprocessed and extract the word
Step 7: Find out the aspect and opinion of the user
Step 8: Calculate the average score of positive and negative reviews
Step 9: Produce the result based on the score
The above algorithm describe about the overall product Recommendation Framework. This will be more helpful to the new user. They decide to purchase the product or not.

IV. PERFORMANCE ANALYSIS
This section shows the investigational of the proposed IPRF and the ability of the diverse fundamental approaches such as collaborative filtering approaches as well as the explicable recommendation approaches. The effectiveness of this proposed technique is analyzed using the parameters like Precision, Recall, F-Measure, RMSE and MAE. Also the performance of the proposed techniques is compared with several existing techniques. Thus, the expertise is measured by examining the several parameters such as RMSE, MSE, precision, recall, and F-measure.

Precision:
The precision which is defined as the positive predictive rate is the proportion of retrieved reviews which are suitable to the user query. It is mentioned as,
Recall:
A recall in data recovery is the ratio of the user reviews which is suitable to the effectively retrieved user query.

\[ R = \frac{\text{relevant reviews}}{\text{relevant user query}} \]  

F-Measures:
The F-Measure is defined as the biased mean of the precision and recall of the opinion calculation according to the query of the user.

\[ F - M = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \]  

RMSE:
The RMSE is the root mean square deviation (RMSD) which is frequently used the quantity of the variation between the values achieved through an organization as well as the values that are essentially perceived from the background exhibited.

To combine them into a distinct amount of predictive power, these distinct dissimilarities are used which is called as residuals, and the RMSE functions.

\[ \text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(X_{\text{exp}} - X_{\text{est}})^2}{n}} \]  

MAE:
The MAE measures usually the standard size of the mistakes in an arrangement of expectations, lacking of the behavior consideration.

\[ \text{MAE} = \frac{\sum_{i=1}^{n}|X_{\text{exp}} - X_{\text{est}}|}{n} \]  

Here \( X_{\text{exp}} \) signifies the experimental results and \( X_{\text{est}} \) denotes the estimated variable at time \( i \).

The performance of all collaborative filtering methods (Ren, Liang, Li, Wang, & de Rijke, 2017) in terms of MAE and the RMSE values are validated.

Average Precision:
The average prediction is the approximation of overall product opinion amongst the various types of user queries (Lei, Qian, & Zhao, 2016). The Fig. 3 signifies the average precision rate. The proposed EFCFM method is compared with the traditional Sentiment based Rating prediction (RPS) approach. The plots 4, 5 & 6 demonstrates the comparison of proposed EFCFM model with existing collaborative filtering (CF) model among the various increasing sparsity methods. It is established that the proposed EFCFM technique precision, recall, and the F-measure value is greater than the basic ClubCF & RPS in all sparsity measures. Thus, the proposed EFCFM method establishes the superior performance.
Opinion Based Memory Access Algorithms using Collaborative Filtering in Recommender Systems

V. CONCLUSION

This paper the Innovative Product Recommendation Framework (IPRF) is proposed. Based on the positive and negative polarity, an Inclusive Similarity-based Clustering (ISC) is proposed to cluster the extracted related keywords from the user reviews. The Enhanced Feature Specific Collaborative Filtering (EFCFM) model for the feature specific cluster is used to evaluate the strength, weakness of the product, estimate the respective features, as well as the opinions. Finally the complete feedback of the product is estimated by propose the Novel Product Feature-based Opinion Score Estimation (NPF-OSE) process. The main challenge in this recommendation system is the fault information estimation of the reviews and the unrelated recommendations of the best selling or the better quality product. To neglect these problems, an Enhanced Feature Specific Collaborative Filtering Model (EFCFM) is developed in the web based service recommendation system. Thus the developed EFCFM method is examined by compared with the traditional approaches in terms of subsequent parameters, precision, recall, F-measure, MAE and the RMSE. The results show that the attained EFCFM approach produced the superior accuracy recommendation to the customers. This framework will be more helpful for the customer to take decision for purchase or not.

REFERENCES:

1. Chen, L., Chen, G., & Wang, F. (2015). Recommender systems based on user reviews: the state of the art. User Modeling and User-Adapted Interaction, 25(2), 99-154.
2. Eirinaki, M., Gao, J., Varlamis, I., & Tserpes, K. (2018). Recommender systems for large-scale social networks: A review of challenges and solutions: Elsevier.
3. Hu, J., Liang, J., Kuang, Y., & Honavar, V. (2018). A user similarity-based Top-N recommendation approach for mobile in-application advertising. Expert Systems with Applications.
4. Kuang, L., Yu, L., Huang, L., Wang, Y., Ma, P., Li, C., & Zhu, Y. (2018). A Personalized QoS Prediction Approach for CPS Service Recommendation Based on Reputation and Location-Aware Collaborative Filtering. Sensors, 18(5), 1556.
5. Lei, X., Qian, X., & Zhao, G. (2016). Rating prediction based on social sentiment from textual reviews. IEEE Transactions on Multimedia, 18(9), 1910-1921.
6. Li, W., Cao, J., Wu, J., Huang, C., & Buyya, R. (2018). A collaborative filtering recommendation method based on discrete quantum-inspired shuffled frog leaping algorithms in social networks. Future Generation Computer Systems.
7. Liu, Q., Liu, B., Zhang, Y., Kim, D. S., & Gao, Z. (2016). Improving Opinion Aspect Extraction Using Semantic Similarity and Aspect Associations. Paper presented at the AAL.
8. Najafabadi, M. K., Mahrin, M. N. r., Chuaprat, S., & Sarkan, H. M. (2017). Improving the accuracy of collaborative filtering recommendations using clustering and association rules mining on implicit data. Computers in Human Behavior, 67, 113-128.
9. Ravi, K., & Ravi, V. (2015). A survey on opinion mining and sentiment analysis: tasks, approaches and applications. Knowledge-Based Systems, 89, 14-46.
10. Ren, Z., Liang, S., Li, P., Wang, S., & de Rijke, M. (2017). Social collaborative viewpoint regression with explainable recommendations. Paper presented at the Proceedings of the Tenth ACM International Conference on Web Search and Data Mining.
11. Sangeetha, J., Prakash, V. S. J., & Bhuvaneswari, A. (2017). Dual Access Cache Memory Management Recommendation Model Based on User Reviews.
12. Vaglano, I., Monti, D., & Morisio, M. (2018). SemRevRec: A Recommender System based on User Reviews and Linked Data.
13. Wang, F., Pan, W., & Chen, L. (2013). Recommendation for new users with partial preferences by integrating product reviews with static specifications. Paper presented at the International Conference on User Modeling,
Adaptation, and Personalization.

14. Yao, L., Wang, X., Sheng, Q. Z., Benatallah, B., & Huang, C. (2018). Mashup Recommendation by Regularizing Matrix Factorization with API Co-Invocations. IEEE Transactions on Services Computing.

**AUTHORS PROFILE**

Ms. J. Sangeetha Completed M.C.A., M.Phil., from Bharathidasan University, Trichy. She is a Research Scholar & Associate professor (C.Sc) in Cauvery College For Women, Affiliated to Bharathidasan University, Trichy. She has 19 years of teaching experience. Her research interests are data mining, Artificial Intelligence. She has published many articles in Scopus Springer, IEEE xplore, and peer reviewed journals.

Dr. Sinthu Janita pursued Master of Computer Application and Ph.D from Bharathidasan University. She is currently working as a Professor and Head in the PG & Research Department of Computer Science, Cauvery College for Women, Trichy. She is a member of IEEE, Computer Society of India. She has published more than 35 papers in reputed international journals including Thomson Reuters (SCI & Web of science) and Conferences including IEEE, Elsevier, ACM, Springer and it’s also available online. Her main research work focuses on Communication Networks, Network Security, Big Data Analytics, Data Mining, and Mobile Communications. She has 23 years of teaching experience and 14 years of Research Experience.

|    | ClubCF | RPS | FSCF | EFCFM |
|----|--------|-----|------|-------|
| 02-04 | 0.74   | 0.71 | 0.791 | 0.821 |
| 04-06 | 0.75   | 0.72 | 0.83  | 0.84  |
| 06-08 | 0.78   | 0.75 | 0.83  | 0.85  |
| 08-10 | 0.74   | 0.73 | 0.85  | 0.86  |