Hybrid Collaborative fusion based product recommendation exploiting sentiments from implicit and explicit reviews

Zafar Ali Khan N, R Mahalakshmi

Department of Computer Science and Engineering
Presidency University, School of Engineering,
Itgalpur, Rajanakunte,Yelahanka, Bengaluru 560064, karnataka, India

Abstract—Product recommendation is an important feature of online shopping. The goal of the recommendation system is to recommend products with higher accuracy such that purchase success ratio are increased. User profile, product purchase history etc have been used to provide high quality recommendations. Product reviews is one of the important source for personalized recommendation. Typical collaborative recommendation systems are built upon user rating on products. But in many cases, these rating information are inaccurate or not available. There is also a problem of biased reviews which decreases the accuracy of recommendation systems. This work proposes a collaborative fusion based recommendation system mining the aspects information in the implicit and explicit reviews. The sentiments about different aspects in the reviews are translated to multi dimensional aspect ratings and these ratings information is fused with user profile and demographic attributes to provide high quality of recommendation. The proposed recommendation system has 3.79% lower RMSE, 4.51% lower MAE and 22% lower MRE compared to most recent collaborative filtering based recommendation system.

Keywords: Collaborative fusion, implicit reviews, aspect mining, sentiment analysis.

I. Introduction

With rapid availability of internet, E-Commerce and online shopping has gained wide acceptance. Product recommendation system is an important part of online shopping helping to find product of interest filtering out irrelevant products. Product recommendation system personalizes product view experiences based on purchase history, user profile and user history etc. The existing methods for product recommendation falls in three categories- content based filtering, collaborative filtering and hybrid filtering.

Content based filtering groups product based on similarity and recommends new products which are similar to user's past purchased products. Collaborative filtering methods groups the users based on their rating of past products and recommends new products which are recently purchased by similar users. Hybrid filtering methods combine both content based and collaborative filtering methods. An important challenge in collaborative filtering methods is the availability of user rating for the products. Many users prefer to use free form of text in form of reviews to express their opinions. The collective opinion is mixed as it has both positivesand negatives on different aspects of the product. Product reviews from users has become an important criteria for new users product purchase decisions in online shopping as past customers exchange their experiences about the product in those reviews. Multi criteria based product recommendation can be realized by mining those reviews and translating the implicit and explicit aspect sentiments to feature wise ratings. Several challenges exist in this realization of multi criteria based product recommendation

1. Segregation of aspects expressed in natural language implicitly in the reviews
2. Quantization of sentiments expressed in

natural language in the reviews
3. Recommendation fusing multi dimensional aspect oriented rating ,user profile attributes , product similarities, demographic information etc
4. Removal of bias from the reviews

This work proposed a hybrid collaborative fusion based product recommendation addressing the above mentioned challenges. The recommendation is based on fusion of multiple dimensions of users, products and reviews. The system applies natural language processing to quantize sentiments expressed implicitly and explicitly into ratings. Multi dimensional collaborative fusion on different attributes gathered from three different dimensions of user, product and reviews is used to rank the product and personalize the recommendation list for the user. Following are the important contributions of this work

1. Machine Learning model for classification of both implicit and explicit reviews is proposed in this work while most existing works are based only on explicit reviews
2. A model to analyze the aspect based reviews is proposed to quantize the degree of polarity of the review.
3. A method to detect biased reviews based on Histogram analysis is proposed in this work
4. Collaborative recommendation based multi dimensional feature fusion is proposed in this work.

II. Related Work

Buettner et al proposed a personality based product recommendation system which used social media data to predict user's personality. Machine learning approaches are applied to predict user personality traits based on social networking features. Product recommendation system uses the relationships between the personality based consumer preferences and the product characteristics to rank the products. The scale used for personality trait in this work did not have higher correlation with product characteristics as it could not model user interest and temporal changes in interests. Geng et al proposed a deep learning model for image recommendation. Latent features are learnt from user image using Deep Learning network and similarity of these features with new images is measured using Euclidean distance. The new images are then ranked based on distance and top N ranked images are recommended to the users. Though this concept applies only for images, user of deep learning for learning latent features is a salient feature which can be used for product recommendations. A memory based technique for group recommendation system is proposed by Ghazarian et al. Support vector based regression model is used to compute the similarity between the items. This work used Pearson Universal Kernel (PUK) function to model the similarities between the items. Use of Support vector regression is able to solve the data sparsity problem. The method works only for single dimension ratings and kernel function needs to be adapted for multi dimensional ratings. A two stage cascaded recommendation system using decision tree and collaborative filtering is proposed by Krzywicki et al for people recommendation in online dating services. At first stage collaborative filtering is applied and the recommendations from it are re-ranked using a decision tree critic. Due to this two stage cascaded recommendation, the success rate of match making improved. A important take away from this work, applying post re ranking procedures to collaborative filtering help to achieve better personalization. Zahra et al proposed a highly scalable k-means clustering based recommendation algorithm. A new centroid selection algorithm exploiting underlying data correlation structures provides better accuracy than random centroid selection. Y. Zhang et al proposed a new neural complementary recommender system called ENCORE which user the complementary item relationships and user preferences. A neural network model is built to learn the complex (non-linear) relationships between items for flexible and scalable complementary product recommendations. A mixture model approach for post purchase complementary product recommendation is proposed by H. Zhao. The mixture model is trained to learn latent prediction contexts, which are determined by user and item profiles, and then make open rate predictions accordingly.

Expectation Maximization (EM) algorithm is used to optimize the parameters of mixture model. A major problem in this method is that it could model the temporal features of user behavior. Huynh et al proposed a complementary recommendation system that learns visual cues in a unsupervised manner to calculate the co-occurrence distribution of items. A salient feature in this solution is that a conditional generative model is trained to produce multiple novel samples of complementary items (in the feature space) for a given query item. K. Zhao et al proposed a deep learning solution using Siamese Convolutional Neural Network architecture to learn style compatibility from the products. The deep learning model is able to find the related products based on style compatibility and recommend those related products to the users. The solution is trained on word model and could be extended for more sophisticated sentence models to be useful in real world environment. Barkan et al extended the item based collaborative filtering to work in the framework of neural word embedding. Item embedding is generated in latent space and using it the item to item relationship is inferred. Skip gram with negative sampling is the word embedding method used in this work. A salient feature in this solution is that item relationships can be learnt from unstructured product descriptions. Liu et al proposed a new user similarity model for collaborative recommendation which solves the cold start problems. The solution is able to increase the recommendation performance when only fewer ratings are available using the local context information of ratings and global preference of user behavior. A hybrid recommendation system combining content based, collaborative filtering and data mining techniques is proposed by FátimaRodrigues to solve the efficiency problems in recommendation for huge size of transactions. The customers are clustered and association rule mining in done for customers in same cluster to provide a more assertive and personalized recommendations. Cui et al extended the collaborative filtering recommendation for the case of implicit feedbacks. The implicit user observations are transformed into two paired magnitudes: preferences and confidence levels. For each user-item pair, this work derive from the input data an estimate to whether the user would like or dislike the item ("preference") and couple this estimate with a confidence level. This preference-confidence partition has no parallel in thewidely studied explicit-feedback datasets, yet serves a keyrole in analyzing implicit feedback. Latent factor algorithm is designed that directly addresses the preference-confidence paradigm. Unlike explicit datasets, here the model should take all user-item preferences as an input, including those which are not related to any input observation (thus hinting to a zero preference).

This is crucial, as the given observations are inherently biased towards a positive preference, and thus do not reflect well the user profile. However, taking all user-item values as an input to the model raises serious scalability issues –the number of all those pairs tends to significantly exceed the input size since a typical user would provide feedback only on a small fraction of the available items. Gai Li et al proposed a personalized ranking algorithm based on both implicit and explicit user feedback. The proposed MERRSVD++ algorithm optimizes the well-known evaluation metricExpected Reciprocal Rank (ERR) and is based on the newestxCLiMF model and SVD++ algorithm. Yuan Li et al proposed a new matrix factorization model named PSVD, which allows us to capture user's different preferences over different items flexibly in rating prediction. Specially, authors use a pair of preferences to represent the whole preference of user over items. Then the dual preferences are considered simultaneously in building the latent feature vector of user. Moreover, PSVD model allows users to adjust their own feature vector when selecting different products. An unified one class collaborative filtering approach is proposed in Zhiqiang Zhang to simultaneously optimize both rating and rank of recommended items. The proposed solution integrated Collaborative less is more filtering (CLMF) and probabilistic matrix factorization (PMF) approaches by sharing common latent features of users and items in CLMF and PMF. A collaborative recommendation algorithm with importance to tags is proposed Yuehua Dong. Type and frequency of use of the label reflect user preferences and preferences, in order to establish a new user preferences model for better mining and use implicit user feedback data will affect the degree of the label on the user to quantify, to establish a new method for similarity computation. A learning-to-rank recommender system is proposed by Babak Loni. It uses implicit feedback signals from multiple channels. The solution was on focused on Factorization Machines (FMs) with

Bayesian Personalized Ranking (BPR), a pairwise learning-to-rank method that allows to experiment with different forms of exploitation. Kommineni et al proposed a user based collaborative filtering approach for book recommendation. The authors worked on multiple factors like raining, feedback, management, reporting, configuration, and using it to offer useful information to the user in order to aid in decision-making and data item recommendations. Kantepe et al designed a recommendation system using autoencoders. Various deep learning training algorithms like Gradient descent, Rmsprop, Adaptive momentum etc were used in fine tuning the recommendation system. Fanca et all compared different machine learning algorithms for product recommendation. Janjarassuk et al used genetic algorithm for product recommendation. Wang et al used SVM with active learning for product recommendations.

III. PROPOSED SOLUTION

The architecture of the proposed hybrid collaborative fusion based recommendation system is given in fig 1. The proposed solution has five important processes

1. Review classification
2. Polarity Estimation
3. Bias removal
4. Collaborative recommendation
5. Multi criteria re-ranking

A. Review Classification

The reviews need to be classified into explicit or implicit before the next step of polarity estimation from the reviews. Naïve Baiyes classifier is trained to classify the reviews into explicit and implicit. Training set of labeled reviews is exploded using a novel sentence explosion algorithm to increase the volume of training set. From the reviews, nouns and adjectives are extracted and each review is converted to a word vector. Label is associated with each of the word vector and a Naïve Baiyes classifier is trained to classify a word vector to implicit or explicit. Sentence explosion helps to increase the accuracy of classification byincreasing the volume of training set. Sentence explosion is done by vocabulary replacement for the nouns and adjectives. For each noun and adjectives suitable synonyms are substituted and alternative sentences are created. The label for the exploded sentences is same as that of root sentence from which explosion is done.

The accuracy and loss of classification with and without sentence explosion for amazon review dataset is measured and given below

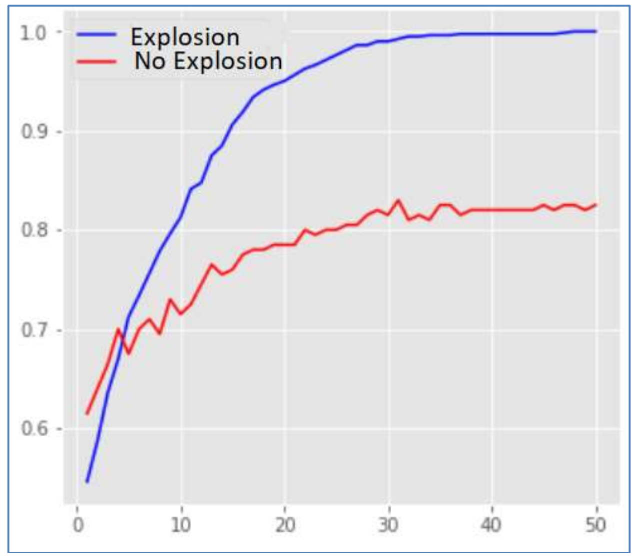


Figure 1 Accuracy comparison

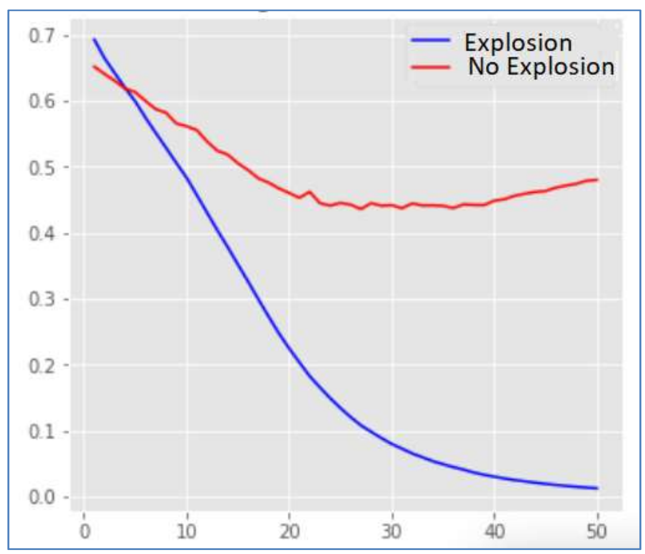


Figure 2 Loss Comparison

From the results, it can be seen that explosion has increased the accuracy compared with review classification without explosion.

B. Polarity Estimation

Polarity estimation is handled differently for explicit and implicit reviews. The aim of polarity estimation is to covert a review into aspect polarity vector (ASV) as below

$$ASV = \sum_{i=1}^{all\ sentence\ in\ review} < A_i , AP_i >$$

Where A is the aspect and AP is the polarity of the aspect. The polarity of aspect is one of the five levels (Most Negative, Negative, Neutral, Positive and Most Positive).

For explicit reviews, contextual polarity of sentence is mapped to one of the 5 polarity level. The noun in the sentence is mapped to aspect. For implicit reviews, the extraction of aspect is not so direct like explicit reviews. Work proposed in [19] does contextual polarity of sentence instead of lexicon-based polarity estimation. This work is adopted for polarity estimated in the proposed solution with some minor modifications. The work in [19] used MPQA (Multi-perspective Question Answering) corpus for annotating the sentiments, but for the proposed Amazon product review dataset is used. The annotations were only of three level of positive, negative and neural in [19], but the proposed solution five different levels are applied. Also, the sentence is exploded and the average of the score of each sentence is given as the final score. The result of using amazon corpus and sentence explosion in scoring is compared to work proposed in [19] and the result for accuracy and loss is given below

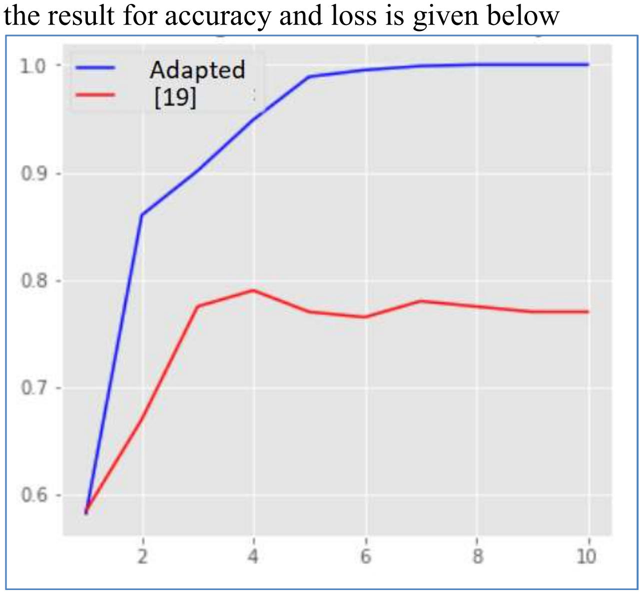


Figure 3 Accuracy of Polarity Estimation

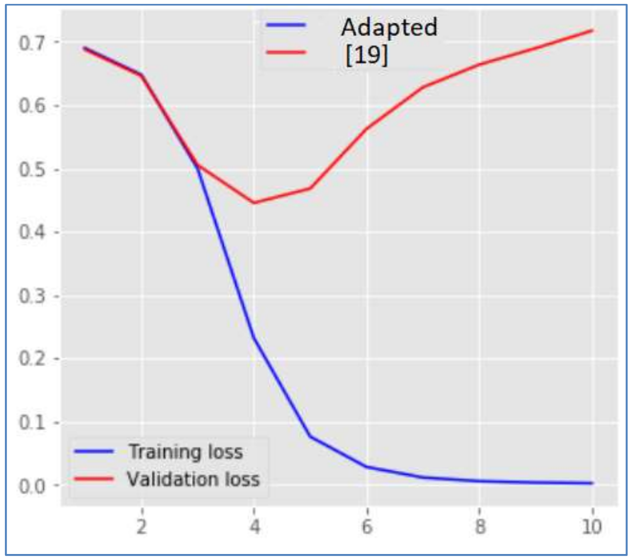


Figure 4 Loss of Polarity Estimation

The results show that accuracy is improved in the adapted solution compared to [19] because of using amazon corpus and sentence explosion in phrase analysis.

C. *Bias removal*

Since online reviews are becoming an important decision factor for purchase, false propagation through online reviews has become more rampant. Biased reviews are injected either to boost or degrade purchase decisions. These reviews must be detected and removed to increase the efficiency of the collaborative recommendation process. Histogram based anomaly detection is proposed in this work to remove the biased reviews. Anomaly analysis is based on deviation from normal behavior with consideration for temporal trend and customers review profile.

The customers who have given reviews are split into three histogram bins based on review count threshold.

1. Most Frequent (MF)
2. Frequent (F)
3. Not Frequent (NF)

The anomaly detection process is designed differentially for each of the customer group defined above. For the customers in MF category, all those reviews about all products are collected. The overall sentiment for each of the products with and without consideration for each of customers review is calculated separately. For each of customers in the group, polarity is calculated as

$$P_c = \frac{\prod_{i=1}^{N} PS_i \neq CS_i}{N}$$

$$PS_i = \frac{\sum_{leaving\ C} \sum_{j=1}^{T} SP_j}{T}$$

$$CS_i = \frac{\sum_{only\ C} \sum_{j=1}^{T} SP_j}{T}$$

$$SP = \frac{\sum_{for\ all\ aspect} AP_{asepct}}{total\ no\ of\ asepcts}$$

Where N is total number of products reviewed, T is the total number of reviews, SP is the sentiment polarity of the review. Sentiment polarity of the review is the average aspect polarity of all aspects in the review. The average polarity of each product for all customers is calculated. The customers who product polarity deviating from average product polarity by a threshold C_1 for more than 50% of the products is marked as suspect. The reviews of customers whose are marked as suspect is split into time units of one week. The customer polarity CS_i is calculated in the epoch time of each week. The customers are confirmed as suspectusing a trend function on sequence of customer polarity over epoch and confirmed suspects reviews are marked as bias and removed from further analysis.

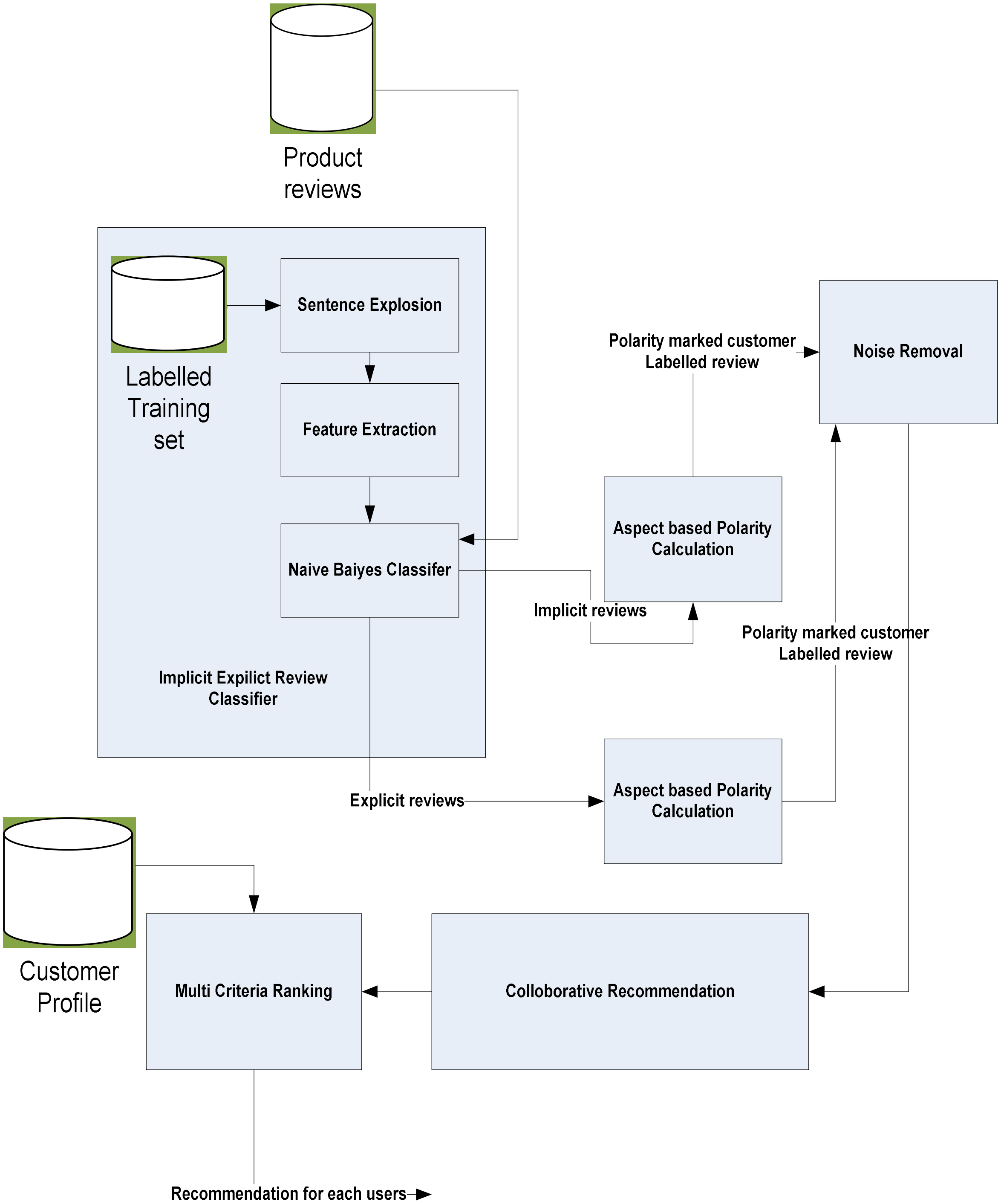


Figure 5 Collaborative Fusion Recommendation

The trend function for suspect confirmation is as follows

$$TF = \begin{cases} confirmed\ , & CS\ is\ monotically\ increasing \\ temporary, & more\ CS > average \\ not\ suspect, & more\ CS < average \end{cases}$$

The same processing is repeated for both Frequent and Not Frequent category with different values of threshold for C_1. The value of C_1 is selected as follows

$$C_1 = \begin{cases} 0.6 * AvgPolarity, & MF \\ 0.7 * AvgPolarity, & F \\ 0.8 * AvgPolarity, & NF \end{cases}$$

The performance of proposed histogram based bias removal is measured in terms of standard root mean square error(RMSE) and compared with User bias removal(UBR)I and II technique proposed in [20]. Amazon food review dataset, Amazon e-commerce electronics dataset and Amazon e-commerce movies dataset are used for evaluation.

Table 1 RMSE Comparison

Dataset	RMSE in UBR-I	RMSE in UBR-II	Histogram based
Food review	0.56	0.71	0.23
Electronics	0.86	0.9	0.27
Movies	0.87	0.87	0.25
Average	0.76	0.83	0.25

From the results, RMSE is very less compared to UBR-I and UBR-II. The RMSE value is reduced by 67% compared to UBR-I and 69% compared to UBR-II in the proposed Histogram based user bias removal.

D. *Collaborative recommendation*

A collaborative fusion of user relationship, product relationship and user to product relationship with matrix factorization is done to recommend the products. Collaborative fusion using all three dimensions of user relationship, product relationship and user to product relationship solves the problems related to cold start, scalability and sparsity. The recommendation system works for n users $u = \{u_1, u_2, \dots u_n\}$ and m products $v = \{v_1, v_2, \dots v_m\}$. User's polarity on a product in terms of average of all aspects polarity of product is given as rating matrix R. The rating matrix is usually spare and the aim of the recommendation system is to predict the value for unknown ratings R_{ij}.

In the proposed recommendation model we introduce two additional vectors for ease of rating prediction called latent vector and bias vector.

$U \in R^{n \times D}$ is the user specific latent vector

$V \in R^{m \times D}$ is the product specific latent vector

$bu - \{bu_1, bu_2, \dots bu_n\} \in R^n$ is the user bias vector

$bv - \{bv_1, bv_2, \dots bv_n\} \in R^m$ is the product bias vector

User to user and product to product relationship matrix is built. These relationship matrixes are fused using Matrix Factorization to identify the bias and latent features for users and products. The learnt bias and latent features are then used to predict the ratings.

A user to user relationship matrix S is a matrix of 0s and 1s and value of $S_{ij} = 1$represents the users i, j are similar than a threshold

A product to product relationship matrix C is a matrix of 0s and 1s and value of $C_{ij} = 1$ represents the products i, j are similar category.

The similarity between users is calculated using Jaccard measure of how the users have co-rated the products and how many products are co-rated. It is given as

$$sim(u_i, u_j) = \frac{\sum_{v_j \in R(u_i) \cap R(u_j)} \exp(-\lg|R(v_j)|)}{|R(u_i) \cup R(u_j)|}$$

The latent features and biases of the users and products are found by factorizing R. A low rank matrix factorization with bias approximates the R as

$R_{ij} \sim bu_i + bv_j + U_i^T V_j$

Singular Value Decomposition (SVD) is used to approximate the rating matrix R by minimizing the squared error between actual observed ratings and predicted estimation for the available ratings. The optimization function is given as

$$min_{U,V,bu,bv} L_1(R_{ij}, U_i, V_j, bu_i, bv_j) = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{m} I_{ij}(R_{ij} - bu_i - bv_j - U_i^T V_j)^2$$

I_{ij} is the indicator function whose value is 1 if user u_i rates item v_j. Through the optimization function, interpolation W is learnt which represents the influence of user $u_i$onu_j. The rating R_{ij} for unrated position is given as

$$R_{ij} = \sum_{u_{k \in S^{v_j}(u_i)}} W_{ik} \frac{R_{kj} - \overline{R_{uk}}}{\sqrt{S^{v_j}(u_i)}} + bu_i + bv_j + U_i^T V_j$$

Once all the unrated positions are predicted in R, the products having the predicted rating $R_{ij} \geq \beta$ are selected for recommendation.

The performance of ranking is measured in terms of mean absolute error(MAE) between the predicted and actual rating for different training to test split and plotted below. The performance is compared against Collaborative filtering (CF) without applying SVD.

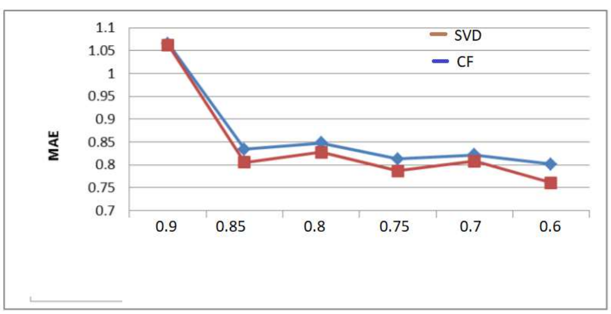


Figure 6 Ranking Error

The ranking error is less in the SVD based rating prediction compared to Collaborative filtering based rating prediction.

*E. **Multi criteria ranking***

The products selected for recommendation are ranked based on multiple criteria and the top K ranked products are recommended to the user. The criteria used for ranking are

1. Demography
2. Product Class of Customer
3. Age of the customer
4. Aspect preference

For each criterion, a criteria score is maintained for each user. Initial value is 0. For first time when products are to be recommended, four separate ranking list based on demography, product class, aspect preference and age of customer is shown to the user. For each product a internal criteria indicator for which the product is ranked and displayed is kept. The preference of products on these criteria is continuously tracked in terms of weights for the criteria and this weight bias is multiplied to the predicted rating. The products are then sorted in descending order of rating. The top K products from the descending order are provided as recommendation to the user.

Against each product that is recommended, a criteria preference indicator is kept which one of three values (1- Demography, 2 – Product class, 3- Age). When customer selects that product, it preference is added in exponential moving average model as

$C_t = \alpha * pref + (1 - \alpha) * C_{t-1}$

Where C_t is the criterion score for that criteria and $pref$ is one of four values (1, 2, 3and 4) based on product criteria selected.

Against each product that is given as recommendation, the average of aspect scores for each of product against whom the calculated user similarity greater thanβ is displayed.

The overall flow of the proposed algorithm is given in Figure 8 below

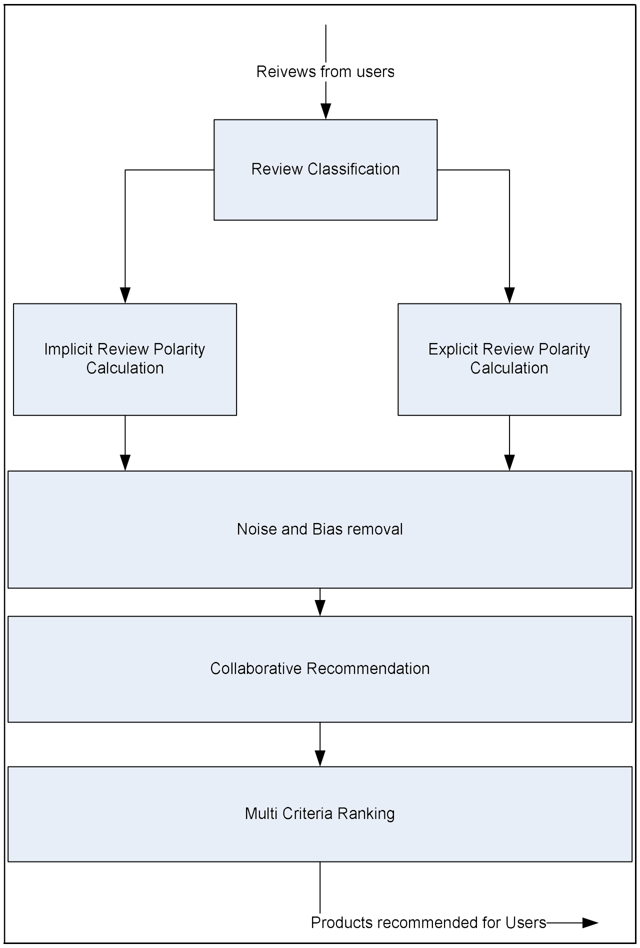


Figure 7 Process Flow

IV. RESULTS

The performance of the proposed Hybrid collaborative fusion recommendation is evaluated against Amazon product review dataset. The proposed hybrid recommendation algorithm is developed in Python 3.6 using sklearn module.There are 548,550 different products. The dataset includes various information for each product and we extract the ASIN, title and review information for each product. There are 7,593,244 unique reviews extracted. From all the review information data, we obtained customerID, review. By extracting user information from product review section, we have 1,555,170 unique users extracted, who gave rates and reviews to the 548K products. The grand average for user review rating is about 4.17. The dataset summary is below

Table 2 Dataset description

Parameter	Value
No of Products	548550
No of Reviews	7593244
No of users	1555170

The dataset was split to 80:20 ratio and recommendation was verified with 20% of data. The performance of the proposed solution is compared with

1. Matrix factorization model with dual preference for rating prediction [15]
2. One Class Collaborative Filtering based on Rating prediction [16]

The performance of the recommendation is measured in terms of

1. RMSE (Root Mean Square Error)
2. MSE (Mean Square Error)
3. MRR (Mean Reciprocal Rank)

All the above metrics quantify the difference between the predicted ratings and real one.

$$RMSE = \sqrt{\frac{\sum_{(u,v)\epsilon T}(\check{r}_{uv} - r_{uv})^2}{|T|}}$$

$$MAE = \frac{\sum_{(u,v)\epsilon T}|\check{r}_{uv} - r_{uv}|}{|T|}$$

$$MRR = \frac{1}{N}\sum_{u=1}^{N}\sum_{i=1}^{M}\frac{\overline{R_{ui}}}{R_{ui}}\prod_{j=1}^{M}(1 - Y_{uj}I(R_{uj} < R_{ui}))$$

Where is r_{uv} is the actual rating of user u on product v and $\check{r}_{uv}$ is the predicted rating of user u on product v and $|T|$ is the number of user item pair in the test set.

Table 3 Performance Comparison

Performance Indicator	**RMSE**	**MAE**	**MRR**
Hybrid Collaborative Fusion –	9.7790	7.2184	0.07

Proposed			
[15]	10.1640	7.5593	0.09
[16]	58.1064	53.7222	0.25
[11]	72.4569	69.1423	0.39
[17]	96.5463	80.456	0.45

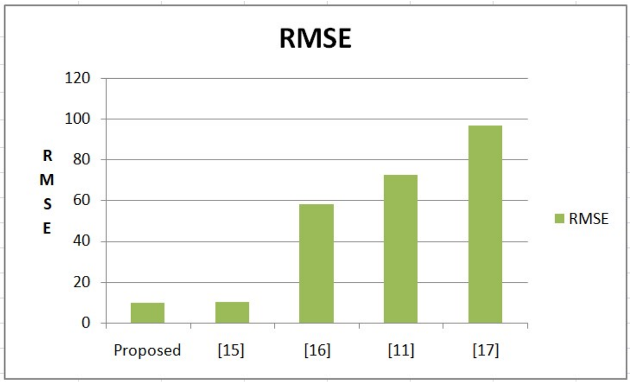


Figure 8 RMSE Comparison

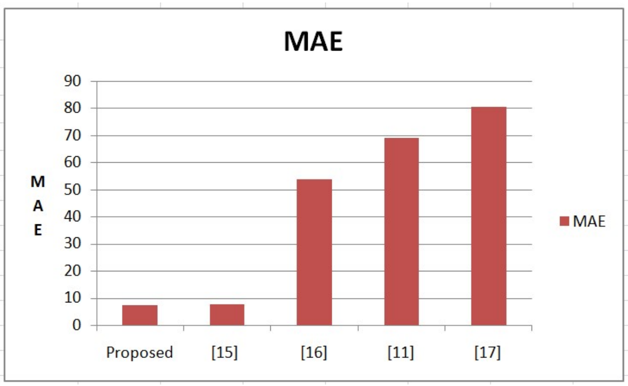


Figure 9 MAE Comparison

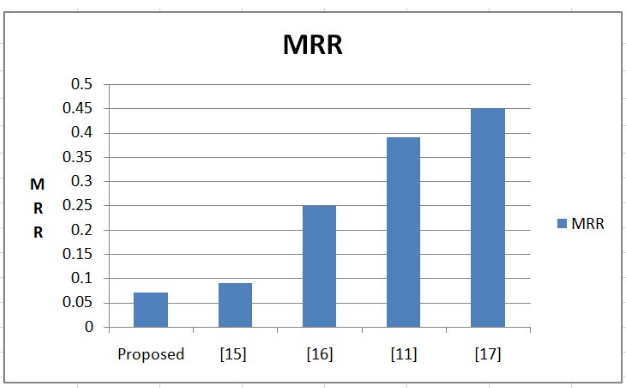


Figure 10 MRR Comparison

The RMSE value in the proposed solution is 3.79% less compared to [15] and 83.17% less compared to [16]. The MAE value in proposed solution is 4.51% less compared to [15] and 86.56% less compared to [16]. The MRR value in proposed solution is 22% less compared to [15] and 72% less compared to [16].

The RMSE values for 5-fold cross validation across three solutions is given below

Table 4 RMSE Comparison on 5 fold validation

	Hybrid Collaborative Fusion	[15]	[16]
1	1.30557698	1.40586561	1.41833305
2	1.34797047	1.41032484	1.39467324
3	1.30508411	1.38607694	1.37559504
4	1.29058551	1.355646	1.41717618
5	1.31637607	1.39993234	1.41828346

The MAE values for 5-fold cross validation across three solutions is given below

Table 5 MAE Comparison on 5 fold validation

	Hybrid Collaborative Fusion	[15]	[16]
1	1.0728638	1.15636132	1.41833305
2	1.10065138	1.14396937	1.39467324
3	1.05124118	1.38607694	1.37559504
4	1.03845935	1.12841074	1.41717618
5	1.0802465	1.09838052,	1.41828346

The RMSE value is measured for different number of criteria for different value K (no of products to recommend) and the result is below

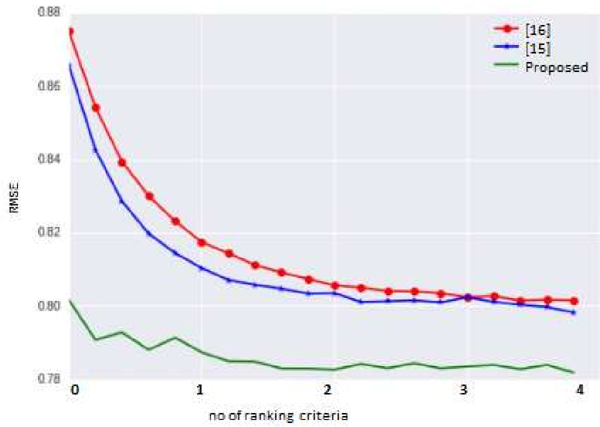


Figure 11 RMSE for K=15

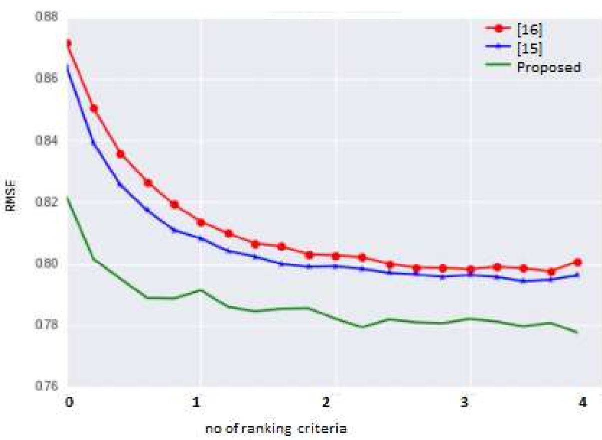
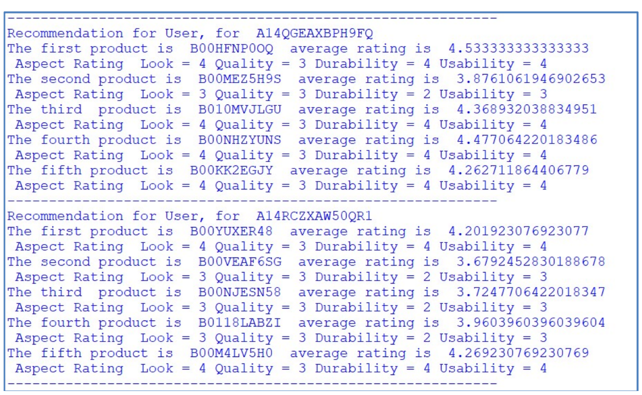


Figure 12 RMSE for K=10

As the number of re-ranking criteria increased, the RMSE value reduced by larger factor in the proposed solution compared to [16] and [15].

The snapshot of the recommendation by the system is given below



```
------------------------------------------------------------
Recommendation for User, for  A14QGEAXBPH9FQ
The first product is  B00HFNP0OQ  average rating is  4.533333333333333
 Aspect Rating  Look = 4 Quality = 3 Durability = 4 Usability = 4
The second product is  B00MEZ5H9S  average rating is  3.8761061946902653
 Aspect Rating  Look = 3 Quality = 3 Durability = 2 Usability = 3
The third  product is  B010MVJLGU  average rating is  4.368932038834951
 Aspect Rating  Look = 4 Quality = 3 Durability = 4 Usability = 4
The fourth product is  B00NHZYUNS  average rating is  4.477064220183486
 Aspect Rating  Look = 4 Quality = 3 Durability = 4 Usability = 4
The fifth product is  B00KK2EGJY  average rating is  4.262711864406779
 Aspect Rating  Look = 4 Quality = 3 Durability = 4 Usability = 4
------------------------------------------------------------
Recommendation for User, for  A14RCZXAW50QR1
The first product is  B00YUXER48  average rating is  4.201923076923077
 Aspect Rating  Look = 4 Quality = 3 Durability = 4 Usability = 4
The second product is  B00VEAF6SG  average rating is  3.6792452830188678
 Aspect Rating  Look = 3 Quality = 3 Durability = 2 Usability = 3
The third  product is  B00NJESN58  average rating is  3.7247706422018347
 Aspect Rating  Look = 3 Quality = 3 Durability = 2 Usability = 3
The fourth product is  B0118LABZI  average rating is  3.9603960396039604
 Aspect Rating  Look = 3 Quality = 3 Durability = 2 Usability = 3
The fifth product is  B00M4LV5H0  average rating is  4.269230769230769
 Aspect Rating  Look = 4 Quality = 3 Durability = 4 Usability = 4
------------------------------------------------------------
```

Figure 13 Recommendation

V. CONCLUSION

A Hybrid collaborative fusion based recommendation is proposed in this work. Machine learning based classification of reviews to implicit and explicit is done. Contextual polarity estimation is done for the sentences in the reviews and associated to the product aspects. Collaborative matrix factorization based rating prediction is done for the products not rated by the user. The products are then ranked with a multi attribute adaptive ranking technique. The performance of the proposed solution is tested in terms of recommendation accuracy, review classification accuracy and temporal adaptivity to criteria preferences. The proposed solution performed better than existing solution in all the metrics. The proposed work can be improved with deep learning based implicit review classification and polarity estimation.

REFERENCES

[1] Buettner, R. (2016). Predicting user behavior in electronic markets based on personality-mining in large online social networks: a personality-based product recommender framework. Electronic Markets, 1–19. doi:10 . 1007 / s12525-016-0228-z.

[2] Geng, X., Zhang, H., Bian, J., & Chua, T.-S. (2016). Learning image and user features for recommendation in social networks. (Vol. 11-18-December-2015, pp. 4274–4282). doi:10.1109/ICCV.2015.486.

[3] Ghazarian, S. & Nematbakhsh, M. (2015). Enhancing memory-based collaborative filtering for group recommender systems. Expert Systems with Applications, 42(7), 3801–3812. doi:10.1016/j.eswa.2014.11.042.

[4] Krzywicki, A., Wobcke, W., Kim, Y., Cai, X., Bain, M., Mahidadia, A., & Compton, P. (2015). Collaborative filtering for people-to-people recommendation in online dating: data analysis and user trial. International Journal of Human Computer Studies, 76, 50–66. doi:10.1016/j. ijhcs.2014.12.003.

[5] Zahra, S., Ghazanfar, M., Khalid, A., Azam, M., Naeem, U., & Prugel-Bennett, A. (2015). Novel centroid selection approaches for kmeans-clustering based recommender systems. Information Sciences, 320, 156–189. doi:10 . 1016/j.ins.2015.03.062.

[6] Y. Zhang, H. Lu, W. Niu, J. Caverlee, "Quality-aware neural complementary item recommendation", *Proceedings of the 12th ACM Conference on Recommender Systems ser. RecSys '18*, pp. 77-85, 2018, [online] Available: http://doi.acm.org/10.1145/3240323.3240368.

[7] H. Zhao, L. Si, X. Li, Q. Zhang, "Recommending complementary products in e-commerce push

notifications with a mixture model approach", *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval ser. SIGIR '17*, pp. 909-912, 2017, [online] Available: http://doi.acm.org/10.1145/3077136.3080676.

[8] C. P. Huynh, A. Ciptadi, A. Tyagi, A. Agrawal, L. Leal-Taixé, S. Roth, "Craft: Complementary recommendation by adversarial feature transform" in Computer Vision – ECCV 2018 Workshops, Cham:Springer International Publishing, pp. 54-66, 2019.

[9] K. Zhao, X. Hu, J. Bu, C. Wang, "Deep style match for complementary recommendation", *AAAI Workshops*, 2017, [online] Available: https://aaai.org/ocs/index.php/WS/AAAIW17/paper/view/15069.

[10] O. Barkan, N. Koenigstein, "Item2vec: Neural item embedding for collaborative filtering", *2016 IEEE 26th International Workshop on Machine Learning for Signal Processing (MLSP)*, pp. 1-6, Sep. 2016.

[11] H. Liu, Z. Hu, A. Mian, H. Tian, X. Zhu, "A new user similarity model to improve the accuracy of collaborative filtering", *Knowledge-Based Systems*, vol. 56, pp. 156-166, 2014.

[12] FátimaRodrigues, "Product Recommendation based on Shared Customer's Behaviour",Procedia Computer Science Volume 100, 2016, Elsevier 2016.

[13] Cui, H. & Zhu, M.. (2014). Collaboration filtering recommendation optimization with user implicit feedback. Journal of Computational Information Systems. 10. 5855-5862. 10.12733/jcis10758.

[14] Gai Li and Qiang Chen, "Exploiting Explicit and Implicit Feedback for Personalized Ranking," Mathematical Problems in Engineering, vol. 2016, Article ID 2535329, 11 pages, 2016.

[15] Yuan Li, and Kedian Mu, "Matrix Factorization Model with Dual Preferences for Rating Prediction," 2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC), pp. 364–372

[16] Zhiqiang Zhang, Gai Li, Liyang Wang, Qiang Chen, and Jincai Pan, "One-class collaborative filtering based on rating prediction and ranking prediction," Knowledge-Based Systems, vol. 124, pp. 46–54, 2017

[17] Yuehua Dong, and Xuelei Liang, "Personalized Recommendation Algorithm Based on Commodity Label," Computational Intelligence and Intelligent Systems, vol. 873, pp. 499–512, 2018.

[18] Babak Loni, Roberto Pagano, Martha Larson, and Alan Hanjalic, "Top-N recommendation with multi-channel positive feedback using factorization machines," ACM Transactions on Information Systems, vol. 37, no. 2, 2019.

[19] Wilson, Theresa & Wiebe, Janyce & Hoffmann, Paul. (2009). Recognizing Contextual Polarity: An Exploration of Features for Phrase-Level Sentiment Analysis. Computational Linguistics. 35. 399-433. 10.1162/coli.08-012-R1-06-90.

[20] Wadbude, Rahul & Gupta, Vivek & Mekala, Dheeraj & Karnick, Harish. (2018). User bias removal in review score prediction. 175-179. 10.1145/3152494.3152520.

[21] M. Kommineni, P. Alekhya, T. M. Vyshnavi, V. Aparna, K. Swetha and V. Mounika, "Machine Learning based Efficient Recommendation System for Book Selection using User based Collaborative Filtering Algorithm," *2020 Fourth International Conference on Inventive Systems and Control (ICISC)*, Coimbatore, India, 2020, pp. 66-71, doi: 10.1109/ICISC47916.2020.917122

[22] E. Kantepe, Z. A. Altikardes and H. Erdal, "Product Recommendation System with Explicit Feedback Using Deep Learning Methods," *2020 Innovations in Intelligent Systems and Applications Conference (ASYU)*, Istanbul, Turkey, 2020, pp. 1-6, doi: 10.1109/ASYU50717.2020.9259814.

[23] A. Fanca, A. Puscasiu, D. -I. Gota and H. Valean, "Recommendation Systems with Machine Learning," *2020 21th International Carpathian Control Conference (ICCC)*, High Tatras, Slovakia, 2020, pp. 1-6, doi: 10.1109/ICCC49264.2020.9257290.

[24] U. Janjarassuk and S. Puengrusme, "Product recommendation based on genetic algorithm," *2019 5th International Conference on Engineering, Applied Sciences and Technology (ICEAST)*, Luang Prabang, Laos, 2019, pp. 1-4, doi: 10.1109/ICEAST.2019.8802561.

[25] Xibin Wang,Zhenyu Dai,Hui Li and Jianfeng Yan, "A New Collaborative Filtering Recommendation Method Based on Transductive SVM and Active Learning",Dynamical Modeling, Analysis, and Control of Information Diffusion over Social Networks,2020