Proposed Model for Emotions Based Recommender Systems Using Reviews

Veepu Uppal[1], Rajesh Kumar Singh[2*]

[1] Manav Rachna University, Faridabad, INDIA,
[2]Shaheed Udham Singh college of Engineering,Tangori, INDIA

Corresponding Author: simranbhandol@gmail.com

Abstract – Information Analysis and extraction is difficult due to huge amount of data on the Internet. Recommender Systems provide efficient and useful information for user according to their preferences. Large numbers of research have been accomplished on Emotion based Recommender systems Techniques. These techniques extract the human emotions for any items from reviews. In this paper we summarize the existing techniques to extract emotions from reviews written by users for different items and propose a new method to design a dynamic search engine which will extract the online reviews and recommend items of different category on the basis of user search. Further our proposed technique will recommend items to user by combination of online reviews and ratings of product too. The spam reviews will be identified and removed.

*Keywords -*Recommender system, emotions, collaborative, content based reviews.

I. INTRODUCTION

On the Internet huge amount of data is available. It is very difficult task to extract the meaningful information by analysing this large amount of data. Recommender systems search through this wide range of data to extract the useful information and services for users.

Information is continuously changing and increasing in internet, the recommender systems cope up with this dynamic information to fetch useful contents. Recommender Systems refer the user's behaviour, interest and their past history of searching to recommend the items for the users. The recommender system which provide efficient recommendations reduce the cost of information extraction and have been proved as   better decision making process sources.

Recommender System: Recommender system is a content refining system that predicts the users' preference to an item. It suggests the items to the users. The Applications of Recommender Systems are widely usedin:
Enjoyment- recommendation for films, songs, TV shows etc.
Content–recommendation of books, newspapers, documents, web sources.
Merchandising- recommendation for customers to buy products.
Services - recommendation of travel services, experts for consultation.

A. *Types of Recommender System:*
  1) Collaborative Recommender System**:** This Recommender Systems recommend items based on the similarities either between User and User similarity or Item and Item Similarities. Collaborative recommender system recommends items to user on the basis of their liking in past and the similarities between users who have same interest as them. The collaborative recommender system is of further two types:

- *User based Collaborative System provides recommendations based on* the similarities between users. Similarities are computed using k nearest neighbour algorithm. Weights are associated with the items according to their importance. The items having maximum weights are recommended to similar users. [1].

- *Item based Recommender system* computes the similarities between items. The new items are rated on the basis of average of rating provided by user to similar old items. The similarity between items is computed by three methods:
  o *Cosine-based Similarity:* In this technique items are considered as vectors. Similarities between different items are calculated by **cosine** of angles between items. The similarity between g, i is calculated by the cosine of angle as given below:

$$Sim(g,i) = \cos\left(\overrightarrow{g},\overrightarrow{i}\right) = \frac{\overrightarrow{g} \cdot \overrightarrow{i}}{\left\|\overrightarrow{g}\right\|_2 * \left\|\overrightarrow{i}\right\|_2}$$

o *Correlation-based Similarity* In this case, similarity between two items is computed by Pearson-r correlation $Corr_{j,k}$. U is the set of users. In this method we first have to isolate the cases where users rated both j and k items.

$$Sim(j,k) = \frac{\sum_{u \in U}\left(R_{u.j} - \overline{R_j}\right)\left(R_{u,k} - \overline{R_k}\right)}{\sqrt{\sum_{u \in U}\left(R_{u,j} - \overline{R_j}\right)^2}\sqrt{\sum_{u \in U}\left(R_{u,k} - \overline{R_k}\right)^2}}$$

$R_{u,j}$ is rating of user u for item j, $\overline{R_j}$ is the average rating of jth item.

o Adjusted Cosine Similarity**:** Computation of similarity between items using basic cosine and Correlation-based methods in item-item similarity has one drawback that these methods do not consider the differences between the rating scales of users. This method similarity overcomes the drawback of not considering the differences between the rating scales of users by subtracting the average ratings of corresponding user from rating of that corresponding user for co rated items. [2]. The similarity between two items b and c by Adjusted Cosine is computed as follows:

$$Sim(b,c)$$
$$= \frac{\sum_{u \in U}\left(R_{u,b} - \overline{R}_u\right)\left(R_{u,c} - \overline{R}_u\right)}{\sqrt{\sum_{u \in U}\left(R_{u,b} - \overline{R}_u\right)^2}\sqrt{\sum_{u \in U}\left(R_{u,c} - \overline{R}_u\right)^2}}$$

$\overline{R}_u$ is the average of uth user's ratings.

o Recommendation Prediction Computation: The rating predicted by user m for item n is computed by dividing the sum of ratings of user m for all similar items N by sum of all similarities between corresponding item with other similar items. Each rating of user m is multiplied by similarity of given item and other similar items.

$$P_{m,n} = \frac{\sum_{all\ similar\ items\ P}\left(S_{n,P} * R_{m,P}\right)}{\sum_{all\ similar\ items\ P}\left(|S_{n,p}|\right)}$$

$P_{m,n}$ is prediction of rating of user m for item n. $S_{n,P}$ is Similarity between item n and all similar items P

Advantages:
▪ It is Easy to implement.
▪ Scalable
▪ No need to consider the content of the items being recommended

Disadvantages:

▪ Dependency on user ratings.
▪ They cannot provide recommendations for new users and items.

2) Content Based Recommendation System: These techniques recommend the items to users that have similar features as the items liked by user in the past history. Items' features are stored and are referred while recommending new items to the user on the basis of content correlations. User profile is created based on user's past ratings and the corresponding item is compared to user profile to determine the similarity between items liked by user in past and that particular item.

Advantages:
▪ *User Independence:* In this the user is independent of other users. The rating provided by active user is used to find recommendations. To find recommendations there is no role of similarity between users.
▪ *Transparency:* The features or descriptions of items can provide the explanation that caused an item to occur in the list of recommendations. Those features can be used as indications to decide whether to trust the system or not.
▪ *New Item:* Content-based recommenders are capable of recommending items not yet rated by any user. Item similarities or user similarities are not required to rate the new item.

Disadvantages:
▪ *Limited Content Analysis: It* requires Descriptions of features of items and users.
▪ *Specialization:* Users are restricted to getting recommendations similar to items already defined in their profiles [3].

3) Hybrid Recommendation System**:** This technique is combination of different recommendation techniques to avoid the limitations of other recommender system. The combination of different techniques of recommender systems provides better recommendations in term of accuracy. Types of hybrid recommender systems are:

• Weighted Hybridization: Weighted hybridization integrates the scores from each technique. P-tango is weighted hybrid recommendation technique. Different recommendations from content and collaborative are combined in this technique. Initially equal weights are given to each technique then weights are adjusted when recommendations are confirmed.
• Switching Hybridization: Switching hybrid technique switches to another methods depending upon the rating of recommendations produced. It is not bound to one method. The main drawback of

switching technique is its complexity due to switching between different methods.

- **Cascade Hybridization:** This technique provides the recommendations in iterative way. The recommendations produced by one method are filtered by another method. The recommendations list generated by first technique is refined by second recommendation technique. [1].

B. *Privacy Issues of Recommendation System*-A shilling attack of recommendations is where the recommendations are manipulated for a specific purpose. The recommendation attack is of two types

1) Push Attack: This attack decreases the rating of items except the items which it wants to make more recommended.
2) Nuke Attack: attack increases the rating of items except the item which it wants to make less recommended. [5].

II. LITERATURE SURVEY

Micheal j.Pazzami et.al discussed contents based recommender systems. Structured items are represented as database table and unstructured items such as newspapers, web pages are represented as vectors of weights. Term-frequency time's inverse document frequency a system to incrementally refine the recommendations based on iterated queries. The documents are represented as vectors of Weight. This technique inverse document frequency has been used to calculate the weight of an item in the document. User Profile is created and updated dynamically on the basis of their preferences. Different methods are used to create the user profile depending upon the type of contents.

Simon Philip, P.B. Shola proposed a technique to recommend the research papers based on Past rating of active user and his taste. The representation of research Paper's features such as Title, keywords, Publisher, authors have been done as vectors of weights using Term-frequency times inverse document frequency. The weight associated with each term representation its frequency in document. Then cosine similarity method is used to find distance between vectors to checks how much similar is a research paper to the papers liked by user in the past.

Emotions: Emotions can be defined as mental states that consists of thinking and feelings, mind-set changes, behaviours, and inclinations to act[6].Research is being done to identify, categorize and process the human emotions. Emotions can be categorised in to dimensional model and discrete model. Dimensional models, map overlapping emotions on few dimensions like valence and arousal. Discrete models are based on distinct emotions categories [7]. W.Parrott list of emotions divides emotions in two categories: primary and secondary .At the first level there are six basic

emotions: love, sad, joy, anger, fear, surprise. Emotions at second level are derived from Primary emotions like cheerful follows from joy. Where cheerful is secondary emotion and joy is primary emotion. Miller extended the emotion list by adding synonyms from WorldNet [8].

Emotional Recommender System:
Emotional recommender system recommends items to user as per their emotions. Emotions of user are extracted from reviews written by user for any entity. Positive or negative emotions are identified from reviews. Apart from this emotions of people are extracted from different social sites like face book, twitter.

Data extracted from unstructured text such as reviews and comments can be used to compare and link the different area or domains. The approach in[9] introduced topic modelling of semantic space in multiple domains. It designed a common space for the corpus in all different areas. The relationships between words of different domains which are non-identical but have same semantic meaning are identified using cosine similarity. The work in [10] further enhances the topic modelling approach by using Point wise Mutual Information (PMI) and Explicit Semantic Analysis (ESA) for computing topic similarity. Review and comments written by user are the rich sources of emotions these emotions can be used for decision making in different recommender system. The approach in [11] detects user's interest and suggests movies of different domains accordingly. A new recommendation technology based on Inductive Learning which consists of content preference and credit preference by user to rate movies is proposed in [12]. Shi et al. proposed a mood oriented movie recommendation system based on similarities between movies, which uses joint matrix factorization model [13]. It adds tags in dataset which in turn is used for factorizing the user-item rating matrix.

Filipa Peleja [14] they proposed a recommendation system to recommend the TV shoes to user on the basis of Media Rating and reviews written by uses .User's reviews are the best indication of their satisfaction so they integrated the reviews with rating .the work is done in two parts: in part one the explicit free text is converted into ratings and then all the ratings are integrated to build a matrix and recommendations are computed from this matrix by matrix factorization method.The reviews are represented as opinion words .Each opinion word is categorised as 0 or 1 means positive or negative. Ratings from 0 to 5 are assigned to each opinion word. The sigmoid function is applied to find the overall rating of review. The proposed method has been proved better than the traditional recommendation method those only consider the media rating but not the user comments.

Guanliang Chen [15] proposed a recommendation system on the basis of context dependent preferences for an item integrated with context independent preference of items.

Firstly the relation between aspect level preferences and context is generated from reviews to derive the context dependent preferences. "Room is nice to stay with family". In this review item is hotel. Aspect is room and the context is family. Proposed work extracts the context dependent emotions of user from reviews. The derived preferences those were dependent upon context were then combined with preferences independent on contexts. Both were inferred from reviews. They proposed a linear regression method to generate the recommendations from integrated context dependent and independent preferences .They tested their method on two dataset one on restaurant data and second on hotel reviews and their method has been proved better from previously suggested approaches.

Lihua Sun [16] they utilized users' reviews and proposed a restaurant recommender system.
Uncertain user opinion has been analyzed by uncertain theory to determine the review strength and polarity. Distance based approach has been used to find similar reviews by combining strength and polarity of reviews. This system differs from other system because it analyzes the similarity between the users' opinions to remove the uncertainty in reviews.

Zhijun Yan[17] proposed a method called EXPRS (An Extended Page Rank algorithm enhanced by a Synonym lexicon) which extracts the product feature form the users' reviews the algorithm works in different steps firstly it eliminates the useless characters from review sentence using word segmentation. Then dependency relationships between head and modifiers of review are generated. For example "price is k", Price is head and k is modifier. Using Node rank algorithm a graph is formed where nodes are Cartesian product of features and sentiment words present in reviews. If any sentiment word of a node is associated with any feature of another node a link is establish between these two nodes. It also uses synom expansion to extract the hidden features from reviews.

A. Razia Sulthana [18] in their research they proposed an Ontology and Context Based Recommendation System (OCBRS) to found the opinion of user from their reviews. This system extracts the context from the review. Different fuzzy rules are used to categories the review in categories on the basis of context. Further an ontology of product features is made. Initially it is made by general features but when reviews are repeatedly processed by fuzzy rules new features are extracted and added in ontology. Ontology is increased after each iteration and acts as repository of context.

III. PROPOSED WORK

A framework has been proposed in this paper (as per Fig. 1) to design a dynamic recommender system which will recommend multiple categories of products to users by combination of reviews and online rating. It will eliminate the spam reviews using ontology of rules for different items.

A. *Data Acquisition*
First we will get all data of various categories from Flipkart and Amazon using Ant colony Optimization algorithm.

B. *Data Pre-processing*
Then in second phase we will separate Html tags from products reviews as well as features of Products.

C. *Design Ontology*
Then Apply Genetic algorithm to make Rules based Ontology based on different product categories.

D. *Spam Identification*
Then we will pass each review of product in different ontology rules and identity the spam if found then truncate that reviews of product.

E. *Recommendation*
Then apply sentimental analysis on product reviews based on that we will identity emotions of user accordingly products will recommend to them.
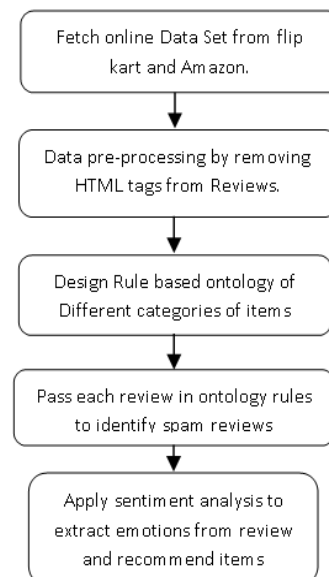

Fig. 1 Proposed Recommender system framework

IV. CONCLUSION

The proposed Recommender System will work on dynamic online reviews. It will work for multiple categories of items. Every time when user will search for an item it will extract the updated reviews for that particular item and recommend the user who searched for according to items ratings. Spam reviews will be identified and eliminated. This framework will be implemented using Java and results will be compared with the existing techniques.

REFERENCES

[1] F.O.IsinkayeY.O.FolajimiandB.A.Ojokoh*"Recommendation systems: Principles, methods and evaluation",*Egyptian Informatics Journal,Volume 16, Issue 3, November 2015, pp: 261-273

[2] Badrul Sarwar, George Karypis, Joseph Konstan, and John Riedl "*Item-Based Collaborative Filtering Recommendation Algorithms*"Proceedings:10th international conference on World Wide Web, Pages 285-295 Hong Kong, Hong Kong — May 01 - 05, 2001 ACM New York, NY.

[3] Micheal j.Pazzami and Daniel Bilsus "*Content based Recommendation System*"Book The adaptive web Springer-Verleg Berlin, Heidlbergpp: 325-341. ISBN:978-3-540-72078-2,2007.

[4] Simon Philio and P.B Shola "*A Paper Reccomender System Based on the Past Rating of the User*"International journal of Advanced Computer Technology, Volume 3,Issue 6,pp 40-46, Dec 2015.

[5] Dhoha Almazro, Ghadeer Shahatah, Lamia Albdulkarim, Mona Kherees, Romy Martinez and William Nzoukou "*A Survey Paper on Recommender Systems*"Published 2010 in ArXiv.

[6] R. W. Picard, "*Affective computing: challenges*" International Journal of Human-Computer Studies, Vol. 59, No. 1, pp. 55--64, 2003.

[7] Russel JA (1980) A circumplex model of affect. J Pers Soc Psychol 39:1161–1178. doi:10.1037/h0077714

[8] Miller GA (1995) WordNet: a lexical database for English. Commun ACM 38(11):39–41

[9] Kumar A, Kumar N, Hussain M, et al "*Semantic clustering-based cross-Domain recommendation*" Proc Computational Intelligence and Data Mining (CIDM): pp: 137–141, 2014.

[10] SaraswatM, Chakraverty S,Mahajan N, Tokas N (2016) "*On using reviews and comments for cross domain recommendations and decision making*". In: Computing for Sustainable Global Development (INDIACom), 2016 3rd International Conference on IEEE, p 3656–3659

[11] Winoto P and Tang TY*"The role of user mood in movie recommendations"*Expert Systems with Applications,Elsevier Volume 37, Issue 8, August 2010, Pages 6086-6092

[12] Peng Li,Seiji Yamada "*A Movie Recommender System based on Inductive Learning*" IEEE Conference on Cybernetics and Intelligent Systems Singapore, 1-3 December, 2004.

[13] Shi Y, Larson M and Hanjalic A"*Mining mood-specific movie similarity with matrix factorization forcontext-aware recommendation*" Proc of the workshop on context-aware movie recommendation, pp: 34–40, 2010.

[14] Peleja, Filipa, et al. "*A recommender system for the TV on the web: integrating unrated reviews and movie ratings*" Multimedia systems 19.6 (2013): 543-558.

[15] Chen, Guanliang, and Li Chen. "*Augmenting service recommender systems by incorporating contextual opinions from user reviews*" User Modeling and User-Adapted Interaction 25.3 (2015): 295-329

[16] Sun, Lihua, Junpeng Guo, and Yanlin Zhu. "*Applying uncertainty theory into the restaurant recommender system based on sentiment analysis of online Chinese reviews*." *World Wide Web* (2018): 1-18.

[17] Zhijun Yan, Meiming Xing, Dongsong Zhang,Baizhang Ma (2015) "*Exprs: An Extended Page rank method for product feature extraction from online consumer reviews*" Information & Management 52.7 (2015): 850-858.

[18] Sulthana, A. Razia, and Subburaj Ramasamy. "*Ontology and context based recommendation system using Neuro-Fuzzy Classification*" Computers & Electrical Engineering (2018).