International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 03 Issue: 03 | Mar-2016 www.irjet.net p-ISSN: 2395-0072

Online Recommendation System Based on Reviews and Ratings

Ramkrishna Gaikwad[1], Ramdas Darwatkar[2], Nehal Jangawali[3], Harish Rukumpure[4] ,

Prof. Dipti Gogawale[5]

[1]*Ramkrishna Gaikwad, Dept. Of Computer Engineering, APCOER, Maharashtra, India*
[2]*Ramdas Darwatkar, Dept. Of Computer Engineering, APCOER, Maharashtra, India*
[3]*Nehal Jangawali, Dept. Of Computer Engineering, APCOER, Maharashtra, India*
[4]*Harish Rukumpure, Dept. of Computer Engineering, APCOER, Maharashtra, India*
[5]*Prof. Dipti Gogawale, Dept. of Computer Engineering, APCOER, Maharashtra, India*

Abstract - Now a days customers express their opinions and *reviews of various online shopping products by the means of ratings and number of varied comments, which becomes tedious to understand the actual quality of the product. How to mine these review data to understand customers choice and make recommendations is crucial to merchants and researchers. Companies like Amazon use their large amounts of data to give recommendation to users. Based on the similarities among items ,systems can give prediction for a new items ratings. Recommender systems uses the user, item, and ratings information to predict how other users will like a particular item. To reach this goal, systems require to parse a lot of data and collect information, sometimes from different resources, and predicts how the user will like the product or item. The proposed system will identify and analyze orientation of product reviews and according to that system will display the rating of particular product. The orientation of product reviews collected by crawler from e-commerce websites is determine using Maximum Entropy Algorithm. Most systems like Amazon, Flipkart, Snapdeal, and others suggest things to users based on similarities among users, products, or both. This will make those systems more efficient from a users perspective. Commercial and trading systems gain trust and benefits using such systems if they successfully know what users want at what time and where. The system will collect information from different websites and then use the stored data in the Maximum Entropy algorithm for text classification providing more accurate results.*

***Key Words*: Product Rating, Product Reviews, Maximum Entropy.**

1. INTRODUCTION

With the increasing demand for e-commerce, more and more products are sold on the Internet Websites, and more and more people are buying products online. This has lead to the development of recommended systems, which collect user preferences or other information and use it to provide recommendations that are likely to be more suited to user needs. Recommendation systems are well known for their use on e-commerce Websites, one where they use input about a customer's significance to generate a list of recommended items. E-commerce, Shopping online is becoming more and more popular. When we need to take decisions about whether to purchase a product or not, the opinions of others become important. The enhancement of new web technologies enables us to freely express our reviews and opinions for various products we have buy which leads to a serious problem. When it comes to an online shopping site, such as Amazon, Flipkart, Snapdeal this problem is about how to help its customers find the precise products that they will purchase. Users express opinions are important because whenever we need to make a decision we want to hear others suggestions. The system will collect information from different websites and then use the stored data in the Maximum Entropy algorithm and for text classification which provides more specific and accurate results. In our text classification scheme, Maximum Entropy measures the conditional distribution of class label given a document. The proposed system will identify and analyse information of product reviews and according to that reviews the system will display the positive and negative rating of product which is more and more useful for users to get information of that product.

2. RELATED WORK

In past few years, there has been an increasing amount of literature on recommendation systems, which focuses on providing recommendations for products to be used by a users or customers.

There are two major methods which are used in

recommendation system to give suggestions.

(i) Personalized recommendation system

(ii) Non-personalized recommendation system.

In Personalized recommendation systems the products can be recommended based on the top sellers of a site. In Non-personalized recommendation systems products are recommend to customers based on what other users have said about the products[1].

Recommender systems can provide an efficient structure to deal with the information overload problem in Personal Learning Environments[3].

Classification is another approach to perform review mining. A multi-label text classification model is proposed by Ganu et al where each sentence is labeled with corresponding topic and sentiment[5].

Recommender Systems have proved in past few years to be a beneficial means for coping with the information overload problem. Ultimately a RS addresses this event by pointing a user towards new, not-yet-experienced items that may be relevant to the user's current task[11].

3. MAIN IDEA OF SYSTEM

Our work is related to opinion mining, or also called as sentiment analysis on free text documents. The need of effective information retrieval and its implementation have became very essential for easy access of relevant information. Recommendation System provides suggestions for items to be used by a users or customers. Online Recommendation Systems opened up new possibilities for the users to know the detail information of items without visiting shops. Our System is Non personalized recommendation system which recommend products to customers based on what previous users have said about the products. The suggestions are independent of the users, so each customer gets the same recommendations. There has been computing positive ratings from overall ratings in e-commerce feedback comments or reviews. In e-commerce systems buyers leave the feedback and ratings, they express some disappointment and negativeness in free text feedback comments often towards specific aspects of transactions. The product reviews expressed by customer on e-commerce websites is collected by using crawler and the database of product reviews is created. When one of the user of the system wants to know about particular product, then the reviews of that product are given to review analyser model. The model created using Maximum Entropy Algorithm can analyse the orientation of all product reviews. And by using output generated by model, rating of particular product is displayed on screen.

3.1 Data Collection

The number of customer reviews available online is growing tremendously as online shopping becomes more popular. It is impossible to collect these online customer reviews manually. As review blogs sites and social networking sites emerge, it is also becoming more difficult to define what a customer reviews are. We have to use web crawling techniques to extract reviews.

3.1 Sentence Spliter

A customer review mostly comprises of several sentences. It is not uncommon to see that customer express multiple positive and negative opinions of a product within a single review. For example, a customer reviewing a digital camera may use a couple of sentences to recommend the quality of picture but use other sentences to belittle the weight and color of the camera. It is very hard to determine the opinion of such a review as a whole. To simplify the problem, we split reviews into sentences in which case, where it is easier to assign positive or negative sentiments. We do not consider sentences which explicit both positive and negative opinions. We use Maximum Entropy method to split reviews into sentences.

4. DEFINATION

The mathematical model is illustrated properly that defines the problem statement of project, input data, output data and all the necessary constraints regarding project. Solution perspective for proposed scheme,

S= {s, e, i, o, F, DD, NDD, Success, Failure}
s - Initial State: Initially system will be in a state where users
are login.
e - End state: Products are displaying in highly rated manner.
i - Input: I={u, p ,r1 ,r2}.
Where,
u is user.
p is password.
r1 is review.
r2 is rating.
o - Output: display products according to user's reviews, rat-
ings and percentage.
F - Function: {register(), getreview(), mining(), display()}
Class= {user, registration, login, process, product, review, ad-
min}
DD - Deterministic Data function are {review, rating}.
NDD - Non Deterministic data function are {registration, lo-
gin, process, product}.
Success - Desired Product Review, Rating and Percentage are
generated.
Failure - Desired Product Review, Rating and Percentage are not generated.
Function Dependency And Identified Morphisms
1) As user can registration website , we are having many to one relationships between number of user and one website,

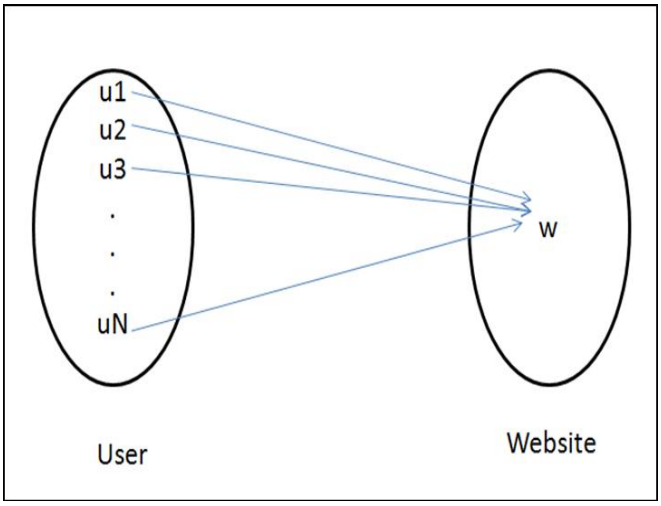

Figure 1: Venn Diagram

2) Whole system have many to many relationship, as system contains number of product and review, and number of product can simultaneously collect number of reviews.

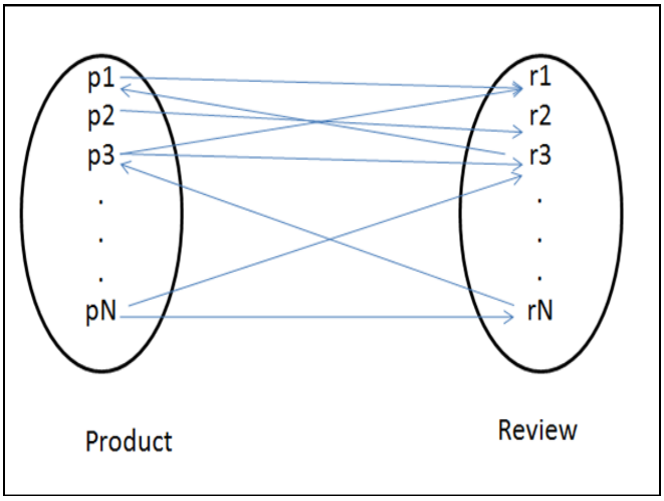

Figure 2: Venn Diagram

5. MAXIMUM ENTROPY ALGORITHM

The aim of the Maximum Entropy model is to use a set of user-specified features and learn appropriate weights. We built a system with Maximum Entropy Classifier that aimed to select feature parameter values to maximize the log of the tweet test data we generated. High weights given to features mean that these are strongly indicative of a certain class.The main idea behind maximum entropy is that one should prefer the most uniform models that also satisfy any given rules.

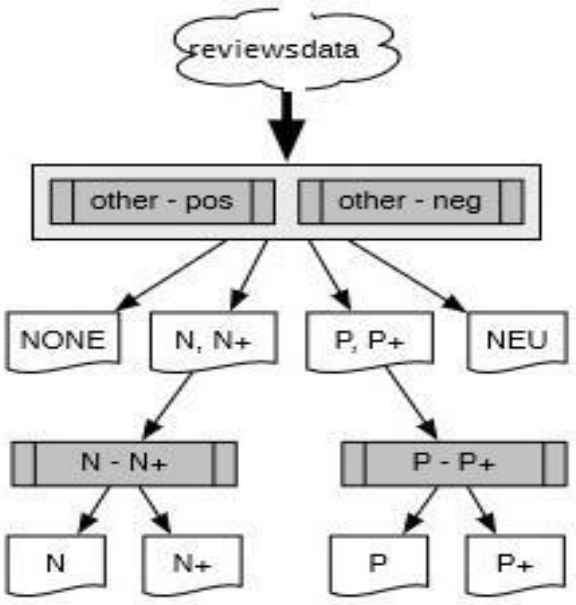

Figure 3: Classification Approach

The sentiment classification considers 6 possible classes: N, N+ → negative polarity; P, P+ → positive polarity; NEU →contains both positive and negative sentiments; NONE →without polarity information. The plus sign (+) signals the sentiment intensity. The first interesting results were achieved by merging five different binary classifiers, one for each class. A first classifier <NONE,other> was used to discriminate between NONE and all the other classes. Two other classifiers <other,neg> and <other,pos> were ap- plied after the first classifier for detecting nega- tive and positive sentiments, respectively. These two latest classifiers make it possible to distinguish between three classes: Positive, Negative, and Neutral. These three classifiers, one can now discriminate between four classes: NONE, Negative, Positive and Neutral. Finally, two other classifiers: <N,N+> and <P,P+>, allow perceiving the sentiment intensity. Only tweets annotated as N and N+ were used for training the <N,N+> classifier, and only tweets marked as P or P+ were used for training the second. That is different from the first three classifiers, which have used all the available data for training.

6. MATHEMATICAL CONCEPTS

To apply maximum entropy method to a domain, we need to select a set of features to use for setting the constraints. For text classification with maximum entropy algorithm, we use word counts as our features. In this paper for each word-class combination we express a feature as:

$$f_{w,c'}(d,c) = \begin{cases} 0 & \text{if } c \neq c' \\ \frac{N(d,w)}{N(d)} & \text{Otherwise,} \end{cases}$$

...(I)

where,
N(d,w) is the number of times word w occurs in document d,
N(d) is the number of words in d.
If a word occurs in one class, we would expect the load for that word-class pair to be higher than for the word paired with other classes.
With this representation, if a word occurs often in one class, we would expect the load for that word-class pair to be higher than for the word paired with other classes. In text classification, we assume that features accounting for the number of times a word occurs should improve classification When constraints are estimated in this fashion, it is sure that a unique distribution exists that has maximum entropy. Moreover, it can be shown that the classification is always of the exponential form:

$$P(c|d) = \frac{1}{Z(d)} \exp(\sum_i \lambda_i f_i(d,c)),$$

...(II)

where each
fi(d,c) is a feature,
ʎ i is a parameter to be estimated,
Z(d) is simply the normalizing factor to ensure a proper probability:

$$Z(d) = \sum_c \exp(\sum_i \lambda_i f_i(d,c)).$$

...(III)

7. CONCLUSIONS

In this system we have proposed the new techniques which gives more accurate rating of product, Because the system analyze feedback which is collected from different websites or given by users and according to that rating is calculated and displayed. We developed a system that collect information from different websites and then use the stored data in the Maximum Entropy algorithm for text classification providing more accurate results. This system is useful for users who are buying products from online websites. Our experimental results indicate that the proposed techniques are very effective in performing their tasks. A recommender system architecture that makes rating predictions using the proposed maximum entropy algorithm and offers personalized recommendation lists to users seeking recommendations is also presented.

REFERENCES

[1] JingnanXu, Xiaolin Zheng, WeifengDing Department of Computer Science "Personalized Recommendation Based on Reviews and Ratings Alleviating the Sparsity.

[2] Problem of Collaborative Filtering " 2012 Ninth IEEE International Conference on e-Business Engineering.

[3] Mohamed Amine Chatti, Simona Dakova, Hendrik Thu s, and Ulrik Schroeder "Tag-Based Collaborative Filtering Recommendation in Personal Learning Environments," IEEE TRANSACTIONS ON LEARNING TECHNOLOGIES, VOL. 6, NO. 4, OCTOBER- DECEMBER 2013.

[4] Prajyoti Lopes, Bidisha Roy, "RecommendationSystem using Web Usage Mining for users of E-commerce site" International Journal of Engineering Research & Technology (IJERT)(2014).

[5] Kamal Nigam, John Lafferty, Andrew McCallum, "Using Maximum Entropy for Text Classifacation" School of Computer Science Carnegie Mellon University Pittsburgh, PA 15213.

[6] Understanding collaborative filtering parameters for personalized recommendations in e- commerce Hong Joo Lee · JongWoo Kim · Sung Joo Park Published online: 24 October 2007 © Springer Science+Business Media, LLC 2007.

[7] Greg Linden, Brent Smith, and Amazon.com Recommendations Collaborative Filtering,"

[8] Fernando Batista ,Ricardo Ribeiro "SentimtntAnalysis and Topic Classification based on Binary".

[9] Maximum Entropy Classifiers" Procesamiento delLenguaje Natural Revista no 50 marzo de 2013, pp77-84 recibido 01-12-12 revisado 14- 02 13 accept do 19-02-13.

[10] Kunpeng Zhang, Ramanathan Narayanan, AlokChoudhary " Mining Online Customer Reviews for Ranking Products".

[11] J. Ben Schafer, Joseph A. Konstan, John Riedl,"E-Commerce Recommendation Application" University of MinnesotaMinneapolis, MN 55455 1-612-625-4002.

[12] RVVSV Prasad, V Valli Kumari, "A C ATEGORICAL REVIEW OF RECOMMENDERSYSTEMS" International Journal of Distribute an Parallel Systems (IJDPS) Vol.3, No.5, September 2012.

[13] D. Pavlov and D. Pennock, "A Maximum Entropy Approach to Collaborative Filteringin Dynamic, Sparse, Hig DimensionalDomains,"Proc. Neural Information Processing Systems (NIPS 2002), BradfordBooks, 2002.

[14] I. Titov and R. McDonald, "A joint model of text and aspect ratings for sentiment summarization," Proceedings of 46th Annual Meeting of the Association for Computational inguistics (ACL'08), 2008.

[15] B. Liu, M. Hu and J. Cheng, "Opinion observer: analyzing and comparing opinions on the Web," Proceedings of the 14 th international conference on World Wide Web, ACM Press, 2005, pp. 342-351, doi:1-59593-046-9.