Reviews Sentiment analysis for collaborative recommender system

Alia Karim Abdul Hassan
Computer science Dept.
University Of Technology
Baghdad, Iraq
110018@uotechnology.edu.iq

Ahmed Bahaa aldeen abdulwahhab
Informatics Dept.
Middle Technical University
Baghdad, Iraq
ahmed80.ab@gmail.com

Abstract: *recommender system nowadays is used to deliver services and information to users. A recommender system is suffering from problems of data sparsity and cold start because of insufficient user rating or absence of data about users or items. This research proposed a sentiment analysis system work on user reviews as an additional source of information to tackle data sparsity problems. Sentiment analysis system implemented using NLP techniques with machine learning to predict user rating form his review; this model is evaluated using Yelp restaurant data set, IMDB reviews data set, and Arabic qaym.com restaurant reviews data set under various classification model, the system was efficient in predicting rating from reviews.*

Keywords: recommender systems, sentiment analysis, opinion mining, natural language processing, text classification.

1. INTRODUCTION

Recommender systems collect information on user's preferences for particular items and try to predict user needs in the field of products, information services [1]. Recommender systems are constructed using two paradigms: content based recommender systems and collaborative filtering recommender systems. The content based recommender systems are depending on item features. Content based recommender systems. Recommend items that similar in the feature with that chosen by the user before. Collaborative filtering (CF) recommender systems depending on the similarity between users, collaborative filtering. Filtering considers two users are similar if they chose same items before, and recommended the same item to each other [2].

So recommender system needs user or objects data to work fluently, but most RS. Suffering from cold start problems (new user cold start and new item cold start) that are caused by lack of information about the new user or new items added to the recommender system database, and the second problem of data sparsity which is caused by insufficient user rating during purchasing process in the recommender system. This yield many works to tackle these two problems [3][4]]. Beyond the traditional methods to build a recommender system, there is a growing paradigm to exploit additional unstructured text reviews as an added source of information about products and services [3].

User review and opinions could be useful in recommender system especially in the area of services like restaurants, movies, hospitals, doctors, more rather in measurable products [5][6].

The contribution of this paradigm is, firstly, it decreases the effect of data sparsity problem in collaborative filtering systems by using the user review as for additional information about an item. Second, it provides a way to extend recommender system to measure services where numerical rating could be impossible or insufficient, third this research handle both English and Arabic languages for sentiment analysis

2. LITERATURE REVIEW

Sentiment analysis (SA.) or opinion mining is the study of emotions toward an item or entity [7]. Sentiment analysis Can be treated as a classification problem; Sentiment analysis will determine whether the sentence expresses positive or negative opinion [7][8]. The most important application of Sentiment analysis is product reviews; these reviews are important for the business holders as they can decide according to user's opinion, and also for users because they Can be recommended for products according to opinions of other users [8].

PMI-IR (point wise mutual information and information retrieval) was proposed by Turing, who computed sentiment orientation is (SO) to two adjectives (poor, and excellent) and measure the synonymy of the word according to these two terms. PMI-IR used to build a synonymy lexicon [9]. Qumsiyeh. And Ng. is Urged a recommender system. That provides recommendations about multimedia items, by using data about features extracted from users reviews about a specific movie or picture [10].

Kim proposed a movie search engine depends on previous user ratings, reviews. When a user types a query which is expanded by selecting an item appeared which is an issue from the last review, but this approach leads to the over- fitting problem and decreasing in diversity in recommender results [11]. Content based recommender system also exploits user reviews in building review based product profile [5]. Ganu et al. proposed restaurant recommender system, that cluster users based on their reviews and the items which liked by them [12].

In recent years after the revolution of the social network, directions appeared to apply sentiment analysis on social network text posts, Molla. Made a sentiment analysis by using a post about Samsung products using a

different account for Samsung company [6]. Dmah. H. and Xiao Zheng develop this direction and deploy sentiment analysis on twitter post for a trusted friends relate to a particular person account[13].

This work proposes a technique to alleviate the problem of data sparsity (poor user ratings), by training a naïve bays classifier to specify user preferences from the review (like/dislike) depending on review features and Word Net corpus.

3. Sentiment classification techniques

Sentiment classification methods can be classified into:-
3.1 machine learning techniques
This paradigm for sentiment analysis uses well-known machine learning algorithms to classify text according to linguistic features [7][14].
3.1.1 Probabilistic classifiers
Probabilistic classifiers are based on hypothesizing a generative model which comes from product distribution over original features space, two Probabilistic important classifiers are:-

- Naïve bays classifier: - this classifier is founded on Bayes theorem, this model computes the positive probability of the class, based on word distribution in a document; this model needs a bag of words feature extraction and ignore the position of the word in document [7][14].
- Maximum entropy classifier: - also known as a conditional exponential classifier, it uses labeled feature set as a vector, then calculates the possibility of each feature that could be computed to specify the most likely class for the feature set.

3.1.1.2 Linear classifier
Linear classifier trains a model to classify objects to their class with a line. Assume $\bar{X} = \{x1, x2, ..., xn\}$ is a vector of word frequencies, $\bar{A} = \{a1, a2, an\}$ is a vector of features, so linear classifier would predict a line that is
P= $\bar{A}\bar{X} + b$, where b is a scalar, and P is a line spreading plan between classes. Linear classifier is divided into:-

- Support vector machine: - support vector machine is a supervised classifier used for classification and regression analysis. The idea behind the SVM is to determine a line that is separating in a search space, which separates different classes. SVM works well with text data, because of sparse nature of the text data. Text data are irrelevant, but somehow are correlated and could be organized in linearly separable categories [7][14][15].
- Neural networks: -this paradigm is based on the neuron, NN consists of many neurons, multilayer neural networks can be used for nonlinear boundaries. The input to the neuron is X which is a vector of word frequencies in the document and set of weights paired with each neuron to calculate class prediction $p = A.\overline{Xi}$.[14]

3.1.1.3 Decision tree classifier: - decision tree uses the division of training data, in which the data is partite based on the condition of feature value, here we can use division by appearance or absence of a word to classify a document [14][15].
3.1.1.4 Rule based classifier: - it depends on modeling the data with a set of rules. If a word holds positive emotion text, it considered as positive, while if the word holds negative emotion, it classified as negative [14].
3.1.2 Unsupervised learning
This machine learning approach has no target to learn from, it is learning by observation. The most used approach in sentiment analysis and recommender system are clustering, which operates by gathering the similar items into one cluster based on Euclidian distance or other distance measurements. K-means, hierarchical clustering, Gaussian mixture model, hidden Markov are known clustering methods [14].
3.2 Lexicon based approach
This method uses opinion words by determining positive words to specify the desired state, and negative words to define unacceptable state [7].lexicon based method is divided into two categories:-

- Dictionary based approach: - this mechanism a collected opinion word with known orientation. This detection could be expanded by searching Word Net corpora for synonyms and antonyms.
- Corpus based approaches: - this method tries to measure the sentiment orientation by finding co-occurrence pattern words to determine their sentiment, and get another opinion words which have similar context [7][14][15].

4. The Proposed system

The proposed system consists of three components. The first one is web scraper which is used to scrap user's reviews from web sites and other social networks, while the second component is responsible for analyzing user reviews and specifying positive and negative sentiments from collected review data set, the third one is a collaborative filter that provides recommendations. This paper focus on a study the second part which is the sentiment analyzer for reviews see figure below:-
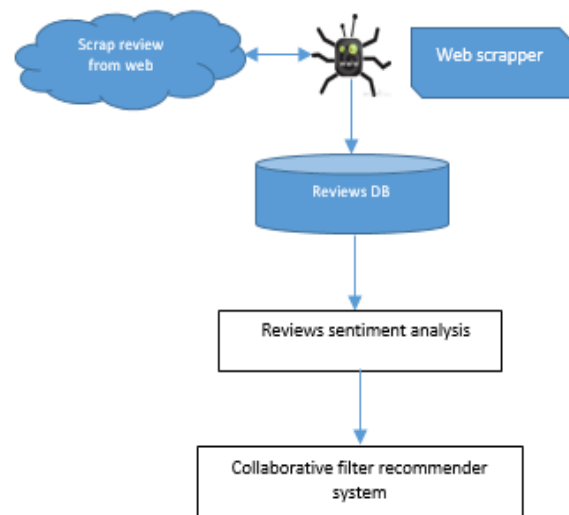

Figure 1. Overview of proposed system

The sentiment analysis component system begins by scraping reviews from the web and preprocess the reviews data set for sentiment analysis operation. The preprocessing process aims to refine the reviews text from parts that decrease the efficiency of machine learning for sentiment analysis operation like stop words and finding the stems (root) of words in each review text and saving the result refined text in corpus list, preprocessing algorithm below illustrate this procedure:-

---

Algorithm1: reviews preprocessing algorithm

Input: n reviews dataset

Output: Corpus of n normalized reviews

I=0;   n=number of reviews in dataset;

1. **While** (i< n) do
2. input review (i)
3. keep English charecters [A-Z,a-z] in review(i)
4. convert review(i) to lower case
5. **for** all words in review(i) do :

    Stem (word) in review (i)

    Remove stop word in review (i)

6. add review to list corpus[ ]
7. i=i+1
8. end

The resulted corpus of reviews is ready for next step which is building the bag of words, dividing the dataset of a bag of words into a training set, and testing set. After that machine learning naïve bays system or any other machine learning model would be trained to predict the positive sentiment and negative one from reviews as algorithm below:-

---

Algorithm 2: Reviews sentiment analysis ML model

Input: corpus of normalized reviews [ ]

Output: trained machine learning model to predict sentiment from reviews (positive negative)

1. Build a bag of word (BOW) from corpus list
2. Divide BOW into training set and testing set
3. Train a machine learning classifier model
4. Calculate confusion matrix (CM)
5. End

This research also handles the sentiment analysis for the Arabic language, which has a specialty in the process because it's an oriental language. The word in the Arabic language contains the word's character and vowel diacritical marks that specify the pronounce of the character that formulates the word. The existence of vowel diacritical marks may cause under fitting for the machine learning system, so this vowel diacritical marks have to be removed in the normalization preprocessing, we can notice algorithm 3 , its begin with read the review and unify the characters than contains vowel diacritical marks like أ , آ, إ to be ا , و to ؤ, ي to ى , and ئ to ى and removing the noise (vowel diacritical marks) like ّ Tashdid, َ Fatha,ُ |Damma, ّ Tanwin Damm, ِ Kasra, ٍ Tanwin Kasr, ْ Sukun. Then complete the preprocess as in English preprocessing by removing stop words and stemming for all words in the normalized review.

Algorithm 3:- Arabic reviews preprocessing

Input: n Arabic reviews data set
Output: corpus of normalized reviews
I=0, n= number of reviews in dataset
1. While (i<n) do :
2. Input review(i)
3. Normalize Arabic characters cases:-
    a. إأآا   to ا
    b. ى   to ي
    c. ؤ   to ء
    d. ئ   to ء

4. Delete noise in Arabic text reviews:
    ّ | # Tashdid
    َ | # Fatha
    ً | # Tanwin Fath
    ُ | # Damma
    ٌ | # Tanwin Damm
    ِ | # Kasra
    ٍ | # Tanwin Kasr
    ْ | # Sukun
    -   # Tatwil/Kashida
5. for all words in reviews (i) do
    Remove Arabic stop words
    Stem (word) in review (i)
6. Add a normalized review to list corpus [ ]
7. End.

5. RESULTS

In the experiment of the sentiment analysis system, we use three datasets, Yelp data set for restaurant reviews which contain 1000 reviews and IMDB movie reviews data set that contains 1582000 reviews, and Arabic restaurant reviews data set scraped from qaym.com contains 8364 reviews. The reviews text file consists of three headers, the user_id; the review is one line text written by a visitor to a restaurant in yelp restaurants data set or movie audience in IMDB or qaym.com and rating that takes either 1 for like or -1 for dislike. The sentiment analysis system is implemented using Anaconda Python 3.6 with the NLTK. Library for natural language processing. The system read the data

from reviews file, execute preprocessing operations (splitting, stemming, and remove stop words) then we divide the data set into 80% training set and 20% test set. Then we implement three machine learning models to compare for the better one and to observe their issues in this area (sentiment analysis), naïve bays, logistic regression, and decision tree.These models are trained using training data; then the model predicts set of values to be compared with the original test set to compute the confusion matrix (CM). In order to calculate accuracy and recall the system use CM matrix which is used to understand the accuracy of our system and how data gets misclassified before tunning our model. The result for experiments illustrated table1.

Table 1: Results of sentiment analysis system

Dataset	Machine learning	Precision	Recall	Accuracy
Yelp	Naïve bays	0.88	0.73	**0.73**
	Logistic regression	0.46	0.75	0.71
	Decision tree	0.47	0.47	0.71
IMDB	Naïve bays	0.78	0.73	**0.82**
	Logistic regression	0.91	0.81	**0.89**
	Decision tree	0.65	0.53	0.67
Arabic Qaym.com	Naïve bays	0.75	0.79	0.80
	Logistic regression	0.82	0.83	**0.82**
	Decision tree	0.80	0.78	0.80

We compute the accuracy, precision, and recall using following laws from confusion matrix

$$accuracy = \sum true\ prediction / \sum test\ set \qquad (1)$$

$$precision = \frac{true\ positive}{true\ positive + false\ positive} \qquad (2)$$

$$recall = \frac{true\ positive}{true\ positive + false\ negative} \qquad (3)$$

6. DISCUSSION

From testing sentiment analysis in three datasets, the smaller one is Yelp restaurant, the big IMDB, and the Arabic qaym.com restaurant reviews. We can observe the accuracy and efficiency of the prediction of the user sentiment reviews to an either positive or negative. The reader can notice that the naïve bays approach work efficiently and accurately with Yelp, IMDB, and qaym.com data sets in adequate training time. While the logistic regression was not efficient with *a* small dataset (yelp) in compare with naïve bays, its accuracy is 0.71 while in IMDB it registered 0.89 and 0.82 in qaym.com datasets as logistic regression build an efficient classifier with large training data sets. Lastly, the decision tree yielded lower accuracy in two datasets (Yelp and IMDB) while it does well with Arabic qaym.com data set. Decision tree needs a lot of features and consuming

training time about 4 minutes to be trained for IMDB and qaym.com, and its accuracy according to the required training set feature size cannot be as good as Naïve bays or logistic regression.

7. CONCLUSION

We propose a sentiment analysis system that can feed ratings extracted from user reviews to a recommender system. Sentiment analysis success in predicting user satisfaction or dissatisfaction by classifying reviews into either positive or negative. This approach could compensate the deficiency in user rating about an item in recommender system (data sparsity).

User reviews could be more suitable to rate things that are not appropriate to be evaluated numerically, and this could yield to explain for a specific recommendation. For future directions, the researchers can build recommender systems depend on attribute about items taken from reviews. In more than one languages like Arabic and English to reduce new item cold start, also this would open the way to treat the recommender system as information retrieval problem by retrieving relevant recommendation from relevant user reviews

8. REFERENCE

[1] J. Babodilla, F.Ortega, A.Hernando, A.Gutiersz, "recommender system survey," knowledge-based systems Journal, vol 46, p 109-132,2013.

[2] Francesco Ricci, Lior Rokach, Bracha Shapira, Paul B.Kantor, Recommender systems handbook. Springer, 2011.

[3] R. Martinas, Garcia," A collaborative filtering approach based on user's reviews," in the proceeding of Brazilian conference on intelligent systems, pp: 204-209, 2014.

[4] F.O Isinkaye, Y.O. Folajimi, B.A. Ojokoh, 2015, "Review: Recommendation systems: principles, methods, and evaluation," Egyptian Informatics Journal, Issue 16, pp: 261-273.

[5] L. Lu, M. Medo, C. Ho, Yeung, Y. Zhang, Z. Zhang, Tao, Zhou,"recommender systems," Physics Reports, Vol 519, pp 1-49,2012.

[6] C.weng, S.chi-Fai, F.Chung, "Integrating collaborative filtering and sentiment analysis: A rating inference approach," in the proceeding of ECAI workshop on recommender systems, pp: 62-66, 2006.

[7] W.Medhat, A. Hassan, H. Korashy, " sentiment analysis algorithms and applications a survey," Ain- shams engineering journal, vol5, pp: 1003-1113, 2014.

[8] M.Issa, V. Piek "A lexicon model for deep sentiment analysis and opinion mining applications," decision support system journal, vol 2, pp: 680-688, 2012.

[9] P.Turneg, "Thumbs up or thumbs down? Sentiment orientation applied to unsupervised classification of reviews", in the proceeding of the 40th annual

meeting of the association for computational linguistics, pp: 417-424, 2002.

[10] R. Qumsiyeh, Y.Ng. ,"predicting the rating of multimedia items for making personalized recommendations," in the proceeding of the 35th international ACU SIGIR conference on research and development information retrieval, pp" 475-484, 2012.

[11] H. Kim, K. Han, J. Cho, J.Hong, "movie mine: personalized movie content search by utilizing user comments," IEEE transactions and consumer electronics, volume 58, pp: 1416-1426, 2012.

[12] G. Ganu, Y.Kakadkar, A. Marian, "Improving the quality of prediction using textual information in online user reviews," information system journal, vol 38, issue 1, pp: 1-15, 2013.

[13] D.Alahmadi, X.Zeng, " ISITS: implicit social trust and sentiment based approach to recommender systems," expert systems with applications journal, vol. 42, pp: 8840-8849, 2015.

[14] A. Bhardwaj, Y. Narayan, V. Anarj, Pawnan, M. Dutta, " sentiment analysis for Indian stock market predictor using Sensex and nifty," in proceeding 4th international conference on eco-friendly computing and communication systems, pp: 85-91, 2015.

[15] M. Giatsoglou, M.Vozalis, K.Dimantras, A. Vakali, G.Sarigiannidis, K. Chatzisavvas, " sentiment analysis leveraging emotions and word embedding's," expert systems and applications journal, volume 69, pp: 214-224, 2017.

Biography

Assist. Prof. Dr. Alia K. Abdul Hassan, Computer Science Dep. /University of Technology. B.Sc. Computer Science/ University of Technology/1993, MSc.from Computer Science/ University of Technology/ 1999, Ph.D. Computer Science/ University of Technology /2004. Assistant professor since 20/3/2008, Position Deputy Dean of Computer Science Department since Feb 2015 till now. Supervised on 22 MSc. &PhD thesis in Computer Science since 2007. Publications Published more than (45) papers in International Conferences and Journals. Current Research Interests Soft computing, Green computing, AI, Data Mining, Software engineering, Document Recognition, Electronic Management, Computer security.

Ahmed B. Abdul wahhab. Ph.D. student in Dept. Of computer science/university of technology. B.Sc. in computer science and education UOB. 2001, MSc computer science from Iraqi commission for computer and informatics 2005, lecturer in the Middle technical university/ technical college of management (Dept. of Informatics) since 2005, teaching data structure, database, and data compression, and programming.