Recommender System Through Sentiment Analysis

Amel Ziani, Nabiha Azizi, Didier Schwab, Monther Aldwairi, Nassira Chekkai, Djamel Zenakhra, Soraya Cheriguene

Recommender System Through Sentiment Analysis

Amel ZIANI[1], Nabiha AZIZI[2], Didier SCHWAB[3], Monther ALDWAIRI[4], Nassira CHEKKAI[5], Djamel ZENAKHRA[2], Soraya CHERIGUENE[2]

[1] Computer science department, Lri laboratory: Information Research Laboratory, Annaba, Algeria
[2] Computer science department, Labged laboratory: Electronic documents control laboratory, Annaba, Algeria
[3]LIG-GETALP Univ. Grenoble Alpes, France
[4]College of Technological Innovation, Zayed University, P.O. Box 144534 Abu Dhabi, United Arab Emirates
[5] Faculty of New Technologies of information and Communication, University of Abdel hamid Mehri, Constantine 2, Algeria.
z_amel1911@live.fr, azizi@labged.net, Didier.Schwab@imag.fr, monther.aldwairi@zu.ac.ae, n.chekkai@gmail.com, dzenakhra@gmail.com, cheriguene@labged.net

***Abstract*—Customer product reviews play an important role in the customer's decision to purchase a product or use a service. Customer preferences and opinions are affected by other customers' reviews online, on blogs or over social networking platforms. We propose a multilingual recommender system based on sentiment analysis to help Algerian users decide on products, restaurants, movies and other services using online product reviews. The main goal of this work is to combine both recommendation system and sentiment analysis in order to generate the most accurate recommendations for users. Because both domains suffer from the lack of labeled data, to overcome that, this paper detects the opinions polarity score using the semi-supervised SVM. The experimental results suggested very high precision and a recall of 100%. The results analysis evaluation provides interesting findings on the impact of integrating sentiment analysis into a recommendation technique based on collaborative filtering.**

Keywords— Recommender systems; Colaborative filtering; Sentiment analysis; Semi-supervised SVM (S3VM);

I. INTRODUCTION

The goal of a Recommender System (RS) is to generate meaningful recommendations to users about items or products that might be of interest to them. This new area of research is gaining more importance mainly due to the effects of widespread use of social media. Most of the existing systems and resources are tailored towards English or other European languages. Despite the fact that Arabic is currently among the top ten most used languages on the Internet, there are very few resources for Arabic recommender systems.

This work is primarily concerned with the task of recommending different products to Algerian customers based on other customers' opinions. We focus on the top four languages used in Algeria: Arabic, Algerian dialect, French and English. Opinions analysis is concerned with the feelings and emotions expressed in a text. It is developing rapidly today because of the widespread usage of web and social media to express very large number of opinions. Therefore, we built a polarity detection system that proven it efficiency in previous experimental works [1-4]. This system transforms reviews texts into a numerical evaluation and feeds them into a recommendation system to implement a collaborative filtering.

Users' emotions are stated explicitly with a vote or implicitly with comments. Those comments are written in natural language with specific vocabulary. They have a polarity score to foresee the vote associated with this comment and make recommendations about items that may interest other customers.

The novelty of this work stems from combining the fields of sentiment analysis and recommendation using collaborative filtering to produce a unique and functioning recommender system. We propose to integrate a semi-supervised classification-based opinions analysis system into a multilingual recommendation system. Firstly, the process of opinions classification extracts the statistical features set such as: number of words, emotionalism, addressing, reflexivity...etc. Secondly, the resulting features vector will be the numerical representation of the review's text in the classification phase by the semi-supervised SVM. Finally, a polarity score is generated to compute the vote for the collaborative filtering of the recommendation phase.

The rest of this paper is organized as follows. Section 2 defines the recommendation systems techniques. Section 3 describes the semi-supervised support vector machine. Section 4 explains in details the description of all the phases of the proposed hybrid system. Section 5 presents the results of the experimental analysis and evaluation.

II. RECOMMENDATION SYSTEMS TECHNIQUES

A recommender system provides suggestions to users, in multiple contexts. For example, when choosing between multiple items or providing the customer with suggested products. Recommender systems are used in most e-commerce websites, where the system displays a list of recommended items to the end user. The core function of a recommender system is to identify potentially useful items for users [5]. In order to predict these, a RS has to be able to predict the utility of these items. Then, based on the results, the system decides which items to recommend.

Recommender systems are commonly classified into three types according to how recommendations are made, namely: content-based filtering (CBF), collaborative filtering (CF), and social filtering (SF) systems. A CBF system suggests user items similar to those he preferred or liked in the past. A CF system suggests user items that people with similar preferences liked in the past while a SF system suggests items according to the preferences of the user's social contacts on social media network. Each of these types of recommendations has its own strengths and weaknesses. In order to address particular short comings and compensate for weaknesses, hybrid filtering (HF) systems combine different recommendation approaches.

III. SEMI-SUPERVISED SVM

Semi-supervised support vector machine (S3VM), proposed by Bennett and Demiriz [6], is a learning method based on cluster assumption. The optimal goal of S3VM is to build a classifier by using labeled and unlabeled data. Similar to SVM, S3VM requires the maximum margin to separate the labeled and unlabeled data. The new optimal classification boundary must satisfy that the classification on original unlabeled data has the smallest generalization error.

Ding et al. [7] explained that the S3VM can be used on all two-class problems in semi-supervised learning. Assuming that there is a dataset, A, with a total of one thousand sets, where one hundred sets are labeled, and the others are not. Only one hundred sets of data may be used with SVM. On the contrary, semi-supervised SVM can make full use of the one thousand sets. Additionally, the performance of semi-supervised SVM is better than SVM. In conclusion, semi-supervised SVM is the best choice, if there are small amounts of labeled data and large amounts of unlabeled data.

Semi-supervised support vector machines have been widely used in many classification problems [8-20]. Because of the lack of labeled Arabic or Algerian datasets we took advantage of the S3VM superior performance with unlabeled datasets.

IV. SYSTEM ARCHITECTURE

As explained earlier the system addresses two problems. First, constructing recommendation system for Algerian users based on opinions analysis. Second, dealing with the four languages, of Algerian users, including: Arabic, Algerian dialect, French and English. Fig. 1 summarizes the different phases of proposed solution.

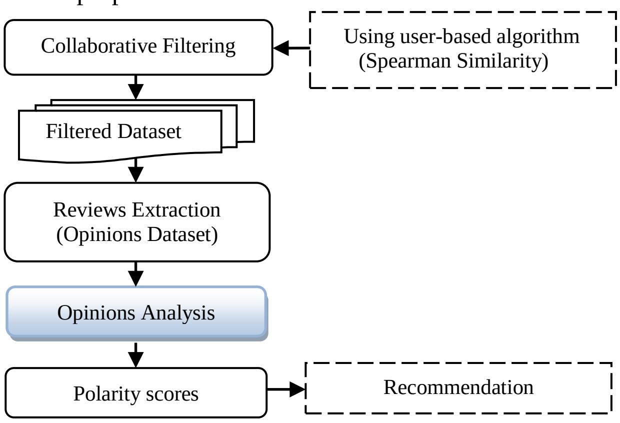

Fig. 1. The proposed process for recommendation system based on opinions analysis.

1. *Collaborative filtering*

Breese et al. [21] divided collaborative filtering algorithms into two categories: memory-based algorithms [22-26] and model-based algorithms [27-39]. Memory-based algorithms compute a prediction by combining ratings of selected users or items that are judged to be relevant. Model-based algorithms use all available ratings to learn a model, which can then be used to predict the rating of any given item by any given user. Memory-based CF algorithms can be further divided into user-based CF algorithms and item-based CF algorithms.

In this work, we used the user-based CF algorithms [23, 25], where a set of *k* nearest neighbors of the target user is identified first by calculating the correlations or similarities between the users' ratings. We experimentally found that spearman similarity was the best similarity measure for the proposed system. Spearman similarity consists of finding a correlation coefficient, not between the values taken by the two variables, but between the ranks of these values. It is defined by:

$$r_s = \frac{\sum_{i=1}^{n}((rang(x_i - \overline{rang(x)})(rang(y_i - \overline{rang(y)}))}{\sqrt{\sum_{i=1}^{n}((rang(x_i - \overline{rang(x)})^2 \sum_{i=1}^{n}(rang(y_i) - \overline{rang(y)})^2}} \quad (1)$$

Where, (x_i) and (y_i) are the observation's ranks in the sample.

The similarities between the users range from -1 to 1. We chose the value 0 as a threshold to identify the closest neighbors of a user. The collaborative filtering with spearman similarity selects the items chosen by similar user's profiles. These last post reviews associated to their preference items.

2. *Opinions analysis*

Fig .2. explains the architecture of the proposed sentiment analysis system. We start first by explaining the features extraction task. Then provide a short overview of the selected features and later we explain the classification phase using the semi-supervised SVM.

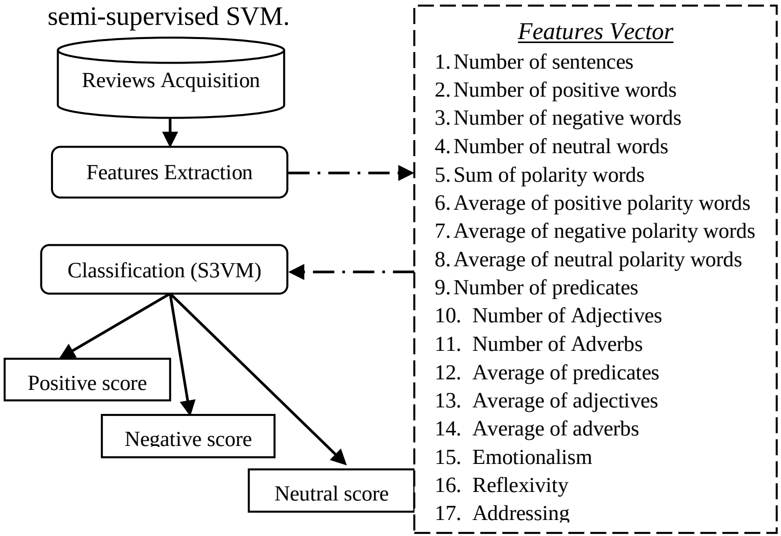

Fig. 2. The proposed process for opinions analysis

1. Reviews acquisition

Because the proposed recommender system is tailored towards Algerian users, we have collected a labeled corpus that comprises almost one thousand reviews in Arabic (MSA), one hundred in Algerian dialect and two thousand in French and English. We also have managed to collect one million unlabeled reviews by using a web crawler. The crawler is a web robot that systematically browses the World Wide Web, and is usually used for the purpose of web indexing.

The main motivation behind using the semi-supervised learning is to employ the large collection of unlabeled data jointly with a few labeled examples for improving generalization performance. Therefore, we have selected two thousand structured reviews for the training of the S3VM.

2. Features extraction

Feature extraction is a rudimentary and essential phase in Sentiment detection process. Therefore, it is important to convert the Arabic review text into a feature vector, so as to process text in a much efficient manner. The selected features used in our system for polarity detection can be summarized as follow.

$$\text{Number of sentences} = \sum sentences \tag{2}$$

$$\text{Number of positive words} = \sum PositiveWord \tag{3}$$

$$\text{Number of negative words} = \sum \text{NegativeWord} \tag{4}$$

$$\text{Number of neutral words} = \sum \text{NeutralWord} \tag{5}$$

$$\text{Sum of polarity words} = \sum PositiveWord + \sum NegativeWord + \sum NeutralWord \tag{6}$$

$$\text{Average of positive polarity words} = \frac{\sum PositiveWord}{\text{Sum of polarity words}} \tag{7}$$

$$\text{Average of negative polarity words} = \frac{\sum NegativeWord}{\text{Sum of polarity words}} \tag{8}$$

$$\text{Average of neutral polarity words} = \frac{\sum NeutralWord}{\text{Sum of polarity words}} \tag{9}$$

$$\text{Number of predicates} = \sum Predicates \tag{10}$$

$$\text{Number of Adjectives} = \sum Adjectives \tag{11}$$

$$\text{Number of Adverbs} = \sum Adverbs \tag{12}$$

$$\text{Average of predicates} = \frac{\sum Predicates}{\sum Predicates + \sum Adjectives + \sum Adverbs} \tag{13}$$

$$\text{Average of adjectives} = \frac{\sum Adjectives}{\sum Predicates + \sum Adjectives + \sum Adverbs} \tag{14}$$

$$\text{Average of adverbs} = \frac{\sum Adverbs}{\sum Predicates + \sum Adjectives + \sum Adverbs} \tag{15}$$

- Emotionalism: The researchers exploited the presence of the adverbs and adjectives in a document as an indicator permitting to determine the opinions. We calculated the emotionalism of a document by counting the number of the adverbs, adjectives and predicates.

$$\text{Emotionalism} = \frac{\sum Predicates + \sum Adjectives + \sum Adverbs}{\sum noun + \sum verbs} \tag{16}$$

- Reflexivity: The reviewers used a lot of reflexivity pronouns such as: «I/me أنا», «I am personally أنا شخصيا », the use of «ي» in « رأيي, I think that », « من وجهة نظري , my point of view »…etc, All of these sentences make reference to an opinion of review, and therefore, we include the measure of the reflexivity. All documents contain a large number of these words will be more subjective. This measure is expressed by Ref (d):

$$Ref(d) = \frac{|\{w \cap w' \setminus w \in d, w' \in R\}|}{|R| + |A|} \tag{17}$$

Where,

d: document

R: reflexivity list

|R|: the number of reflexivity pronouns in *d* from *R*

|A|: the number of addressing pronouns in *d* from *A*

- Addressing: Most reviews contain some addressing words such as: « you أنت, you انتم, yourself نفسك, yourselves أنفسكم, he هو, she هي, them هم, himself نفسه, herself نفسها, themselves أنفسهم», because the reviewers write their opinions addressed to others.

$$Add(d) = \frac{|\{w \cap w' \setminus w \in d, w' \in A\}|}{|R| + |A|} \tag{18}$$

Where,

d: document

A: Addressing list

|R|: the number of reflexivity pronouns in *d* from *R*

|A|: the number of addressing pronouns in *d* from *A*

3. Classification

The Sentiment Polarity Classification is a binary classification task where an opinionated document is labeled with an overall positive or negative sentiment. Sentiment Polarity Classification can also be termed as a binary decision task. The input to the Sentiment Classifier can be opinionated or sometimes not. When given a review is as an input, analyzing and classifying that review, as a good or bad news, is considered to be a text categorization task. Furthermore, this piece of information can be good or bad news, but not necessarily subjective (i.e., without expressing the view of the author). What means this task is a multiclass categorization, where the review can be positive, negative or neutral.

S3VM was chosen as the classification techniques for this phase, in order to benefit from all the collected dataset: labeled and unlabeled in the training.

4. Recommendation

Once the neighbors are obtained and their associated reviews are classified, a weighted average is used to combine the neighbor's item ratings to produce a prediction value for the target user. As soon as all the similarities of the target user *A* with respect to the other users are calculated using Spearman similarity (equation 1) and the *n* most similar users that constitute the vicinity of this target user are defined, the prediction (P_{Aj}) of the value of an item *j* evaluated by the user *A* is calculated using the weighted sum of the estimates of the nearest neighbors who have already estimated the item *j* as follows:

$$P_{Aj} = \overline{v_A} + \frac{\sum_{i=1}^{n} sim(A,i) * (v_{i,j} - \overline{v_i})}{\sum_{i=1}^{n} |sim(A,i)|} \tag{19}$$

Where,

n: number of users present in the neighborhood of *A*, having already voted on item *j*

$v_{j,k}$: Vote of user *i* for object *j*

v_i: Average user rating *i*

$|sim(A, i)|$: Average Similarity

The latter value allows knowing if an item is relevant or not for the target user. This helps the system to generate efficient recommendations for that user. The global range of votes of a recommender system is represented by explicit votes that vary between 0 and 10, if the prediction value ranges between 0 and 5, the item will not be interesting for the user. While the values ranging between 5 and 10 correspond to relevant recommendations for this user.

IV. ANALYSIS OF RESULTS

The system performance has been evaluated with three measures: MeanAbsoluteError (MAE), precision and recall.

- *The MeanAbsoluteError (MAE)* is the most widely used measure in recommendation systems. It estimates the mean of the absolute difference between the estimates and the predictions. The collaborative recommendation system is considered to be performing well when the MAE value is small. For our system, the smaller the MAE, the more efficient the analysis of opinions. This measure is given by the following equation.

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |f_i - y_i| \tag{20}$$

Where,

n is the predictions number,

f_i is the prediction of *i*,

y_i is the evaluation (vote)

- *The precision* which corresponds to the number of comments ranked well in relation to the total number of comments contained in the corpus. The collaborative recommendation system is considered to be performing when the value of the precision is high.
- *The recall*, known as sensitivity, is the number of related comments retrieved over the total number of relevant comments contained in the corpus.

In our system, the higher the precision, the more effective the opinions analysis system. Precision and recall are measured for each class as follows:

$$Precision = \frac{TP}{TP + FP} \tag{21}$$

$$Recall = \frac{TP}{TP + FN} \tag{22}$$

To highlight the importance, and to clarify the effect of the combination between Arabic opinion classification and recommendation techniques, we carried out multiple experiments to calculate the previous measures. We have tested our approach using different datasets in multiple languages to confirm the efficiency of the proposed system.

English dataset: is the *Restaurant_TijuanaRestaurant* dataset [40], which contains 2000 reviews from 50 guests in 40 restaurants. We chose this benchmark because it contains all the necessary information on which our contribution is based (comments and votes).

French dataset: collected from *ldlc.com*[1], "High-Tech Experience" online computer hardware sales website. Our dataset contains 10 users, 5 smart phones and 50 evaluations.

Arabic and dialect dataset: collected from *dz.jumia.com*[2], "JumiaMarket", which is an Algerian-based online shopping website. Our dataset consists of 10 users, 5 oriental clothing for women and 50 evaluations.

TABLE I. THE EXPERIMENTAL RESULTS

	MAE	Precision	Recall
English	0.52	0.96	1.0
French	0.50	1.0	1.0
Arabic and dialect	0.60	0.90	1.0

Table I. shows the results of our experiments. The table shows MAE, precision and recall for the proposed system for the three datasets. We can safe say that our system has high precision and recall. That validates the use of S3VM and confirms that applying sentiment analysis techniques to recommendation systems significantly improves the quality of the recommendations. The system sensitivity in terms of recall was measured at 100% for the datasets. That maximizes the chances for a specific user to access the items he would like to have.

Our experiments validated our assumption of great advantages from combining: the opinions analysis and the recommender systems. Therefore, the new proposed hybrid approach significantly improved the recommendation system performance for Algerian users.

[1] http://www.ldlc.com/ (17.06.2017)

[2] https://dz.jumia.com/ (17.06.2017)

V. CONCLUSIONS AND FUTURE WORK

This study presented a basic tool which can be used to analyze Algerian reviews and comments and detect their polarity, in order to generate meaningful recommendations for users. To achieve this goal, we have tried to use semi-supervised SVM for the opinions classification task to avoid the lack of annotated data problem. The obtained scores from S3VM were used as votes for the recommendation task. To prove the proposed combination efficiency, we have tested and evaluated our system using three datasets (Arabic, Dialect, French and English,). The results were very promising, which encouraged us to continue working along this line. Therefore, we intend to enrich our feature vector with another set of morphological primitives using natural language processing. We plan also to study the choice of using other recommendation techniques.

ACKNOWLEDGEMENT

This work was supported in part by Zayed University Research Office, Research Cluster Award # R17079.

REFERENCES

[1] A. Ziani, N. Azizi, Y. G. Tlili. Detection de polarité d'opinions dans les forums en langue arabe par fusion de plusieurs SVM. In : Proceedings of 20 ème conférence sur le Traitement Automatique des Langues Naturelles RECITAL TALN, 83-95, Sables d'Olonne, France, 2013.

[2] A. Ziani, N. Azizi, Y. G. Tlili. Combining Random Sub Space Algorithm and Support Vector Machines Classifier for Arabic Opinions Analysis. In: International Conference on Computer Science, Applied Mathematics and Applications ICCSAMA: 175-184, Metz, France, 2015.

[3] A. Ziani, C A K A. Kounta, N. Azizi, D. Zenakhra, S. Cheriguene. A New hybrid classification system for subjectivity detection. In: the 2nd Conference on Computing Systems and Applications, Algiers, Algeria, 2016.

[4] A. Ziani, N. Azizi, Y. G. Tlili. Recommender system for sports articles based on Arabic opinions polarity detection with a hybrid approach RSS-SVM. In: 3rd International Conference on Control, Engineering & Information Technology (CEIT), Tlemcen, Algeria, 2015.

[5] G. Adomavicius, A. Tuzhilin. Toward the next generation of recommender systems: A survey of the state of the art and possible extensions. IEEE Transactions on Knowledge and Data Engineering, 2005, pp. 734–749. 9, 14, 21.

[6] K. Bennett, A. Demiriz. Semi-supervised support vector machines. In: Advances in neural information processing systems, vol 11, 1999, pp 368–374.

[7] S. Ding, Z. Zhu, X. Zhang. An overview on semi-supervised support vector machine, In Neural Comput & Applic 28: 2017, pp. 969–978.

[8] B. Xiang, L. Zhou. Improving twitter sentiment analysis with topic-based mixture modeling and semi-supervised training, in: Proceedings of the Fifty Second Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), Association for Computational Linguistics Baltimore, Maryland, 2014, pp.434–439.

[9] L. Becker, G. Erhart, D. Skiba, V. Matula, Avaya. Sentiment analysis on twitter with self-training and polarity lexicon expansion, in: Proceedings of the Second Joint Conference on Lexical and Computational Semantics (SEM), Volume 2 : Proceedings of the Seventh International Workshop on Semantic Evaluation (SemEval 2013), Association for Computational Linguistics, Atlanta, Georgia, USA, 2013, pp.333–340.

[10] W. Baugh, bwbaugh: Hierarchical sentiment analysis with partial self-training, in: Proceedings of the Second Joint Conferenceon Lexical and Computational Semantics (SEM), Volume 2: Proceedings of the Seventh International Workshop on Semantic Evaluation (Sem Eval 2013), Association for Computational Linguistics, Atlanta, Georgia, USA, 2013, pp.539–542.

[11] S. Liu, W. Zhu, N. Xu, F. Li, X. -q. Cheng, Y. Liu, Y. Wang. Co-training and visualizing sentiment evolvement for tweet events, in: Proceedings of the Twenty Second International Conference on World Wide Web Companion, WWW'13 Companion, International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland, 2013, pp.105–106.

[12] C. Tan, L. Lee, J. Tang, L. Jiang, M. Zhou, P. Li. User-level sentiment analysis in corporating social networks, in: Proceedings of the Seventeenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD'11, ACM, NewYork, NY, USA, 2011, pp.1397–1405.

[13] F. Pozzi, D. Maccagnola, E. Fersini, E. Messina, Enhance user-level sentiment analysis on microblogs with approval relations, in: M. Baldoni, C. Baroglio, G. Boella, R. Micalizio (Eds.), AI∗IA 2013: Advances in Artificial Intelligence, Lecture Notes in Computer Science, vol. 8249, Springer International Publishing, 2013, pp.133–144.

[14] C. Johnson, P. Shukla, S. Shukla, On classifying the political sentiment of tweets, http://www.cs.utexas.edu/∼cjohnson/TwitterSentimentAnalysis.pdf, 2011.

[15] S. Hong, J. Lee, J. -H. Lee. Competitive self-training technique for sentiment analysis in mass social media, in: Proceedings of the 2014 Joint Seventh International Conference on and Advanced Intelligent Systems (ISIS) and Fifteenth International Symposium on Soft Computing and Intelligent Systems (SCIS), 2014, pp.9–12.

[16] B. Drury, L. Torgo, J.J. Almedi. Guided self training for sentiment classification, in: Robust Unsupervised and Semi-Supervised Methods in Natural Language Processing, 2011, pp.9–16.

[17] X. Wan. Co-training for cross-lingual sentiment classification, in: Proceedings of the Joint Conference of the Forty Seventh Annual Meeting of the ACL and the Fourth International Joint Conference on Natural Language Processing of the AFNLP: Volume1- Volume1, ACL'09, Association for Computational Linguistics, Stroudsburg, PA, USA, 2009, pp.235–243.

[18] S. Li, Z. Wang, G. Zhou, S.Y.M. Lee. Semi-supervised learning for imbalanced sentiment classification, in: Proceedings of the International Joint Conference on Artificial Intelligence, IJCAI, 2011, pp.1826–1831.

[19] N. Yu. Exploring co-training strategies for opinion detection, J. Assoc. Inf. Sci. Technol.65(10) (2014) 2098–2110.

[20] S. Li, C.-R. Huang, G. Zhou, S.Y.M. Lee, Employing personal /impersonal views in supervised and semi-supervised sentiment classification, in: Proceedings of the Forty Eighth Annual Meeting of the Association for Computational Linguistics, ACL'10, Association for Computational Linguistics, Stroudsburg, PA, USA, 2010, pp.414–423.

[21] J. S. Breese, D. Heckerman, C. Kadie. Empirical analysis of predictive algorithms for collaborative filtering. In Pro c. of the 14th Conf. on Uncertainty in Artificial Intelligence, 1998, pp. 43-52.

[22] C. C. Aggarwal, J.L. Wolf, K-L. Wu, P S. Yu. Horting hatches an egg: A new graph-theoretic approach to collaborative filtering. In Pro c. of the 5th ACM SIGKDD, 1999, pp. 201-212.

[23] J.L. Herlocker, J.A. Konstan, A. Borchers, J. Riedl. An algorithmic framework for performing collaborative filtering. In Pro c. of the 22nd ACM SIGIR, 1999, pp. 230-237.

[24] A. Nakamura, N. Abe. Collaborative filtering using weighted majority prediction algorithms. In Pro c. of the 15th Int. Conf. on Machine Learning, 1998, pp. 395-403.

[25] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, J. Riedl. GroupLens: An Open Architecture for Collaborative Filtering of Netnews. In Pro c. of the ACM Conf. on Computer Supporte d Cooperative Work, 1994, pp. 175-186.

[26] B M. Sarwar, G. Karypis, J A. Konstan, J. Reidl. Item based collaborative filtering recommendation algorithms. In Pro c. of the 10th WWW, 2001, pp. 285-295.

[27] D. Billsus, M J. Pazzani. Learning collaborative information filters. In Pro c. of the 15th Int. Conf. on Machine Learning, 1998, pp. 46-54.

[28] J. Canny . Collaborative filtering with privacy via factor analysis. In Proc. of the 25th ACM SIGIR, 2002, pp. 238-245.

[29] Y-H. Chien, E.I. George. A Bayesian model for collaborative filtering. In Proc. of the 7th Int. Workshop on Artificial Intelligence and Statistics, 1999.

[30] L. Getoor, M. Sahami. Using probabilistic relational models for collaborative filtering. In WebKDD Workshop, 1999.

[31] K. Goldberg, T. Roeder, D. Gupta, C P. Eigentaste: A constant time collaborative filtering algorithm. Information Retrieval, 4(2), 2001, pp. 133-151.

[32] T. Hofmann. Latent semantic models for collaborative filtering. ACM Tran. on Information Systems, 22(1), 2004, pp. 89-115.

[33] B. Marlin. Modeling user rating profiles for collaborative filtering. In Proc. of the 17th Annual Conf. on Neural Information Processing Systems, 2003.

[34] D. Pavlov, E. Manavoglu, D M. Penncok, C L. Giles. Collaborative filtering with maximum entropy . IEEE Intel ligent Systems, 19(6), 2004, pp. 40-48.

[35] D M. Pennock, E. Horvitz, S. Lawrence, C L. Giles. Collaborative filtering by personality diagnosis: A hybrid memory and model-based approach. In Proc. of the 16th Conf. on Uncertainty in Artificial Intel ligenc e, 2000, pp. 473-480.

[36] B M. Sarwar, G. Karypis, J A. Konstan, J. Riedl. Application of dimensionality reduction in recommender systems{a case study . In Proc. of the ACM WebKDD Workshop, 2000.

[37] L. Si, R. Jin. Flexible mixture model for collaborative filtering. In Proc. of the 20th Int. Conf. on Machine Learning, 2003, pp. 704-711.

[38] N. Srebro, T. Jaakkola. Weighted low-rank approximation. In Proc. of the 20th Int. Conf. on Machine Learning, 2003, pp. 720-727.

[39] L H. Ungar, D P. Foster. Clustering methods for collaborative filtering. In Workshop on Rec ommendation Systems at the 15th National Conf. on Artificial Intelligence, 1998.

[40] Y. Zheng. USA, 2015. https://github.com/irecsys (16.06.2017)