

Context Based Classification of Reviews Using Association Rule Mining, Fuzzy Logics and Ontology

Razia Sulthana A^{1*}, Subburaj Ramasamy²

Department of Information Technology, SRM University, Kattkulathur-603203, TamilNadu, India

Coressponding author, e-mail: raziasulthana.a@ktr.srmuniv.ac.in^{1}, subburaj.r@ktr.srmuniv.ac.in²

Abstract

The Internet has facilitated the growth of recommendation system owing to the ease of sharing customer experiences online. It is a challenging task to summarize and streamline the online textual reviews. In this paper, we propose a new framework called Fuzzy based contextual recommendation system. For classification of customer reviews we extract the information from the reviews based on the context given by users. We use text mining techniques to tag the review and extract context. Then we find out the relationship between the contexts from the ontological database. We incorporate fuzzy based semantic analyzer to find the relationship between the review and the context when they are not found therein. The sentence based classification predicts the relevant reviews, whereas the fuzzy based context method predicts the relevant instances among the relevant reviews. Textual analysis is carried out with the combination of association rules and ontology mining. The relationship between review and their context is compared using the semantic analyzer which is based on the fuzzy rules.

Keywords: context, reviews, recommendation, ontology, classification

1. Introduction

Reviewing commercial products is pervasive in online purchases. The Opinion of the reviewer is not only conveyed by the review text specified by the customer or user but also depend on the context of the text. A method for mining and summarizing the review of the user is proposed in [1]. It highlights run downs of client surveys of items sold on the web. Here, highlights extensively mean item elements (or traits) and capacities. Given an arrangement of client surveys of a specific item, the assignment includes recognizing elements of the item that clients have communicated their sentiments on and for every component, recognizing audit sentences that give positive or negative sentiments. With this, they produce a summary of the information retrieved from the product reviews.

The problem of extracting the opinion of the customers review by the computers automatically done in [2] where artificial intelligence approaches are given to solve these kinds of issues and still the user's need efficient systems. In this work, researcher provides the classification based on the text mining techniques to review the contextual information of the review's sentences. But, the author proposes to extend the work further to extract the knowledge representation of the contextual information.

A recommendation system [3] extracts the contextual information of the user's current query and compares the query with the user's previous history and other users history for the same kind of query and make recommendation. A system [4] predicts the user's interests based on the relevant feedback system and a recommendation system is designed based on the relevant feedback of the user's reviews. Here, the user's interests are ranked based on the recommendations history they provide previously.

An approach for mining the context is proposed in [5]. In this approach, semantic relations between the concepts are not clearly extracted and recommendations do not provide efficient results. It uses opinion mining methodology based on the semantic of the web to enhance the results obtained with traditional natural language processing techniques and sentiment analysis processes. This work shows how to improve feature-based opinion mining by using ontologies at the feature selection stage and analysis-based method for sentiment analysis. An overview of association rule mining in [6] demonstrates that review comparison on distributed data over the data taxonomies can be efficiently carried out with association rules

like Apriori. It discusses about the different types of association rule mining which is used to increase the efficiency of the algorithms. It refers the category of databases on which way the association rules are assigned.

In [7] a system for product recommendation based on the domain ontology is proposed with sentiment analysis of the tagged word. In this work, researchers present the conceptual model which relates the semantic relation between context features and opinion words. The analysis is taken down at the feature level of the label provided for the product which helps to extract the exact opinion of the sentiment word mentioned by the users.

Text mining techniques [8] in contextual recommendations for products based on the user recommendations extracts the contextual information. Pre-filtering is used to filter out the irrelevant items. Modeling consists of using the context within the recommendation models. In post filtering, contextual information is used to build a traditional recommendation model to reorder. The work is carried out for building the recommendation system using rule mining technique with some machine learning specifications.

Personalized summarization [10] is proposed to customize the user's review history. It works based on the customer's previous browsing history or by user's credentials. Here, click through approach is used to extract the common features of the users and find the desired products for the current user. We also find the reviews that are more related to the products searched and user needs and classify the positive/negative opinions.

The user reviews are compared with mining the context through the competitive intelligence in [11]. They develop a system to extract the relative relations from client conclusion information and construct similar connection maps which helps supervisors in distinguishing the potential operation dangers and supporting procedure choices. This work uses correlation assessment as data sources for business hazard administration.

A context-aware personalized recommendations system in [12] analyzes the user's context while they access the multimedia information. This method shows an experimental use of latent preferences for ranking items. Performance analysis of this paper shows that there is an improvement in the mean average precision.

The recommendation system proposed in [12] provides a user based contextual recommendation system based on the user behavior. In this approach the author constructs item clustered trees to extract the user behavior based on their preference cart. The system also takes the users past context for providing a better recommendation for their items.

In [13] the author describes about decision making process on consumers reviews. This method analyzes the e-commerce websites and provides the recommendation to users based on the comparative results of the products. It shows that this type of context aware system mostly provide better recommendation than the available benchmark system.

Ontology mapping is applied in [14] to categorize the learning resources. This approach calculates the similarity and weights of the features, thereby the accuracy and effectiveness of system is increased through ontology mapping. A service based framework is proposed in [15] which distribute's the concepts of services hierarchically using a hierarchical ontology. The ontology can thus be used to increase the effectiveness of the system.

The paper is organized as follows. In section 2 the research method is proposed which includes context extraction, ontology in recommender system, fuzzy logics and association rules and finally rules generation using fuzzy logics and association rule. In section 3 the results are discussed and finally the paper is concluded in section 4.

2. Research Method

In this paper, we propose an approach to develop recommendation systems for products based on the opinion of users. The recommendation system process is shown in Figure 1. The reviews are crawled from the corpus and tagged by using Stanford parser. Then we tag the reviews which gives the features of the context given. The analysis done in two different manners where in the first step we extract the features of context and in the second stage we extract the frequent words from the review. Once the frequent words are retrieved then it is sent to the semantic analyzer to extract the opinion of the given context.

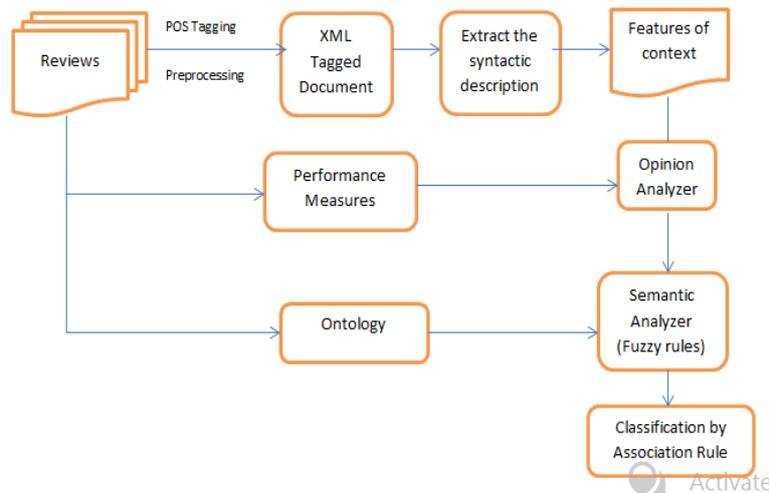


Figure 1. Framework of context based review classification

2.1. Context Extraction

In the first step of our recommendation process we extract the user context instead of going by the meaning of the specified word whereas the traditional text based recommendation extracts the keywords. Here, the term context refers the text that is related with the target word which is also used to extract the opinion of the user review. We extract the syntactic descriptions of reviews into noun phrases and verb phrases. We find the features (words) that are used to express the opinions rather than tag all the words in the context. Given the input as crawled review database we find the frequent words that are available in the review databases and then we extract the opinion words which provide the semantic orientation of the context. We do this with the use the natural language processor for Part of Speech (POS) tagging. It develops the XML tagged output of the review which differs from the usual Part Of Speech tagging. The tagged sentence is saved in the database with the tag specifications.

2.2. Ontology in Recommender System

During the second step we use ontology as a way to increase the performance of the recommender system. Ontology helps in predicting the implicit knowledge of users and makes it explicit for computers to predict the same. Ontology help extend recommender system to a multi-class environment.

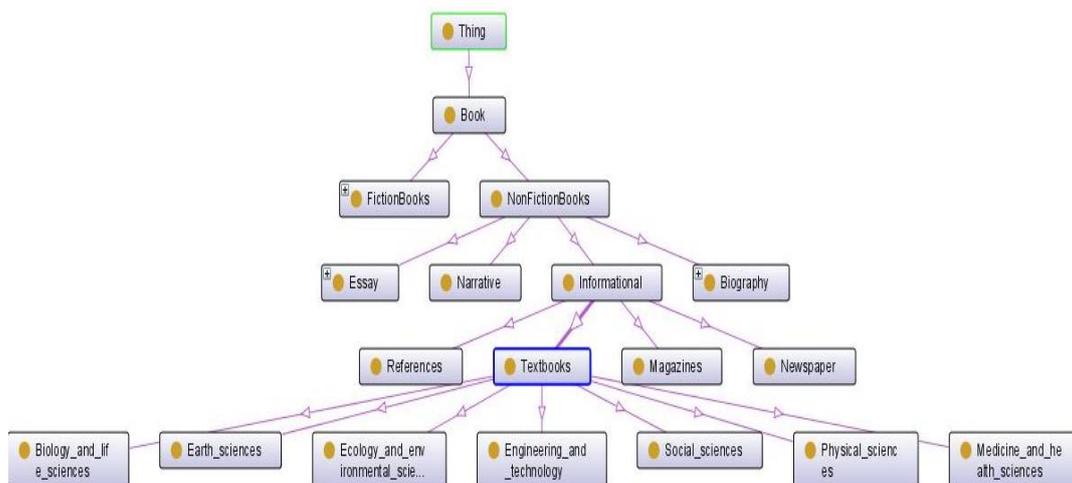


Figure 2. Ontology for book domain

Figure 2 shows the generalized view of the ontology for book domain. The reviews are provided by the users for the particular books and the sentences are separated based on the parsing. In this, the main objective is to find out the text book name by using the features which is available in the reviews that provided in the ontology. It is also important to generate the well-defined training set to extract the proper domain based on the analysis of features.

2.3. Fuzzy Logics and Association Rules

In the final step we develop the association rules using fuzzy logics. While generating the association rules for recommendation system we identify the target product. This helps the system to recommend easily the specific product rather than to generate the rules for several targets. To obtain associations among the reviews, we identify the reviews correspond to the product. While analyzing the context, we employ 4-fold cross validation approach, wherein three groups are retained as training data and the other group is used for testing. We then frame the association rules for the training data and test the same with collaborative ontology mining.

Out context based recommendation system mines the context from the user's reviews given for the products and finds the associations between the user reviews. We discover the level of associations between the reviews and link them. Initially, we predict whether the rating is positive or negative and then we chose the rule based on the prediction label. If user review is positive, then corresponding transaction is tagged as [review: positive] and if user review is negative, then the corresponding transaction is tagged as [review: negative]. If the user doesn't give the review for the product then, the transaction is simply mentioned as [review: null].

The results of opinion analyzer and features context are given to the semantic analyzer which extracts the context opinion semantically by applying fuzzy rules. After we retrieve the semantic view of the context the classification is done by using association rules. Rules are classified based on the semantic view of the context and relation between the tagged objects in the context.

Opinions of the frequent words are identified from the previous learning context or the previous recommendations using ontology.

2.4. Rule Generation Using Fuzzy Logics and Association Rules

We have framed an algorithm (Algorithm 1) to identify the new words that we encounter in the reviews and they possible ways of updating the feature set using fuzzy logics and association rules.

Algorithm 1: To identify the rare word features based on the fuzzy logics

Input: Feature set $F = (f_1, f_2, f_3 \dots f_y)$ and opinion words for each feature set $O = (o_1, o_2, o_3 \dots o_z)$ obtained from training set

Output: Feature set

For each context $c_i = (c_1, c_2, c_3 \dots c_n)$ in the review $r_j = (r_1, r_2, r_3 \dots m)$

If $((c_i \in F) \parallel (o_i \notin O))$

$O \leftarrow o_i$

Else if $(c_i \notin F)$

$F \leftarrow c_i$

Find an noun $(n_i) \parallel$ noun phrase (n_p) such that $(n_i \parallel n_p \in O)$

$O \leftarrow o_i$

Store the c_i in Featureset F of the word

Fuzzy Rule Example

An example review sentence is listed below

Sentence 1: The cost of the mobile is very low.

Sentence 2: The battery backup of the mobile is low.

The sentence based recommendation system evaluates both the sentence and identifies that they have the same noun as "mobile" and opinion as "low. But, if we review the context using the fuzzy rules, one is positive and the other is negative. Because the former is about the cost and the latter one is about battery backup. In this work, fuzzy rules extract the semantic knowledge of the sentence.

3. Results and Discussions

Amazon dataset is taken for implementing the proposed system. The dataset has information that clearly defines the social relationships between the products they buy in the sites. The reviews were collected for the period between November 2010 and April 2015. One fourth of dataset which totally has 15,000 reviews is taken as testing dataset. We have evaluated the obtained results with the benchmark performance measures, precision and recall. Precision signifies the probability of true positive reviews against the total positive reviews and recall signifies the probability of true positive rate. The performance measures are described below:

$$\text{Precision} = \frac{\text{true positive}}{\text{true positive} + \text{false positive}}$$

$$\text{Recall} = \frac{\text{true positive}}{\text{true positive} + \text{false negative}}$$

Table 1. shows the comparison of our previous sentence based review classification [14] with the proposed context based review classification. And also, the results in Figure 3 clearly shows the precision and recall is improved better when compared to the sentence based review classification. True positive is improved and false positive is reduced by using the context based review classification.

Table 1. Comparison of Context Based and Sentence Review Classification

	Class	True Positive	False Positive	Precision	Recall
Sentence based	Positive	0.479	0.663	0.419	0.479
	Negative	0.337	0.521	0.392	0.337
Context based	Positive	0.534	0.474	0.554	0.618
	Negative	0.526	0.466	0.523	0.474

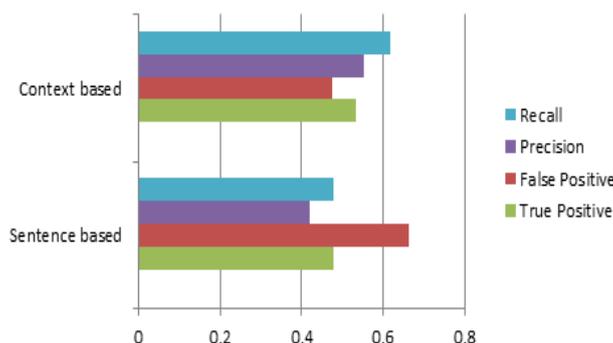


Figure 3. The performance measures against context based and sentence based

4. Conclusion

In this paper, we have proposed the context-aware recommendation classification approach based on user reviews using fuzzy logics. The text mining techniques are used to tag the given context and to provide the connectivity between the words in the review context. It also helped to provide the semantic similarity between the contexts. Textual analysis is done with the combination of association rules and ontology mining. The relation between review and their context is compared using the semantic analyzer which is based on the fuzzy rules. By the experiments, we show that the association rules are quite appropriate for review recommendation systems to provide the efficient classification.

References

- [1] Hu M, Liu B. *Mining and Summarizing Customer Reviews*. In Proceedings of the tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM 2004 Aug 22: 168-177.
- [2] Aciar S. *Mining Context Information from Consumers Reviews*. In Proceedings of Workshop on Context-Aware Recommender System. ACM 2010 Sep 30; 2010.

- [3] Hariri N, Mobasher B, Burke R, Zheng Y. *Context-aware Recommendation Based on Review Mining*. In Proceedings of the 9th Workshop on Intelligent Techniques for Web Personalization and Recommender Systems (ITWP 2011) 2011 May: 30.
- [4] Ramesh LS, Ganapathy S, Bhuvaneshwari R, Kulothungan K, Pandiyaraju V, Kannan A. Prediction of User Interests for Providing Relevant Information Using Relevance Feedback and Re-ranking. *International Journal of Intelligent Information Technologies (IJIT)*. 2015 Oct 1; 11(4): 55-71.
- [5] Penalver-Martinez I, Garcia-Sanchez F, Valencia-Garcia R, Rodriguez-Garcia MA, Moreno V, Fraga A, Sanchez-Cervantes JL. Feature-based Opinion Mining through Ontologies. *Expert Systems with Applications*. 2014 Oct 1; 41(13): 5995-6008.
- [6] Kotsiantis S, Kanellopoulos D. Association Rules Mining: A Recent Overview. *GESTS International Transactions on Computer Science and Engineering*. 2006 Jan; 32(1):71-82.
- [7] Nie X, Liu L, Wang H, Song W, Lu J. The Opinion Mining Based on Fuzzy Domain Sentiment Ontology Tree for Product Reviews. *JSW*. 2013 Nov 1; 8(11): 2682-7.
- [8] Domingues MA, Sundermann CV, Manzato MG, Marcacini RM, Rezende SO. *Exploiting Text Mining Techniques for Contextual Recommendations*. In Web Intelligence (WI) and Intelligent Agent Technologies (IAT), 2014 IEEE/WIC/ACM International Joint Conferences on IEEE 2014 Aug 11; 2: 210-217.
- [9] Kavasoglu Z, Gunduz Ogdücü S. Personalized summarization of customer reviews based on user's browsing history. *International Journal on Computer Science and Information Systems*. 2013 Jul 1; 8: 147-58.
- [10] Xu K, Liao SS, Li J, Song Y. Mining Comparative Opinions from Customer Reviews for Competitive Intelligence. *Decision Support Systems*. 2011 Mar 31; 50(4): 743-54.
- [11] Alhamid MF, Rawashdeh M, Dong H, Hossain MA, El Saddik A. Exploring Latent Preferences for Context-aware Personalized Recommendation Systems. *IEEE Transactions on Human-Machine Systems*. 2016 Aug; 46(4): 615-23.
- [12] Song L, Tekin C, van der Schaar M. Online Learning in Large-scale Contextual Recommender Systems. *IEEE Transactions on Services Computing*. 2016 May 1; 9(3): 433-45.
- [13] Shi F, Ghedira C, Marini JL. Context Adaptation for Smart Recommender Systems. *IT Professional*. 2015 Nov; 17(6): 18-26.
- [14] Gang LV, Sheng-bing C. An Improved Entity Similarity Measurement Method. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*. 2014 Dec 1; 12(4): 1017-22.