Mining Context Information from Consumer's Reviews

Silvana Aciar
Instituto de Informática
Universidad Nacional de San Juan, Argentina
Av. Central y Meglioli
Rivadavia, San Juan, Argentina - 5400
saciar@iinfo.unsj.edu.ar

ABSTRACT

Consumer reviews, opinions and shared experiences are popular ways to express preferences and interest of tourists in most of the popular tourism sites inside of the typical valuation of product using any valuation scale. The most critical issue on opinion mining is how to extract information that can be understood and utilized by computers from written text by users/consumers in natural language. Several approaches using artificial intelligence have been used to deal with this problem even so; problem that has been less addressed but not less important is the identification of context information embedded in consumer's opinions due the arduous task of processing natural language in which reviews have been expressed. This paper addresses this problem based on classification text mining techniques to identify review's sentences containing contextual information to be then processing and incorporated in a recommender system. This approach was exemplified by a case study using reviews from www.tripadvisor.com.

Keywords
Contextual Information, Opinion Mining, Text Mining.

1. INTRODUCTION

Review comments are one of the most powerful and expressive source of user preferences. Product review forums and discussion groups are popular ways for consumers to exchange their experiences with a product [5] [6] [7]. There is growing evidence that such forums inform and influence consumers' purchase decisions [5] [6] [10]. These reviews provide valuable information about consumer's behavior that can be used to infer preferences and interests about future products. However, usage of this information is not an easy task due to the difficulties of incorporating unstructured data [3]. The consumer reviews are in free form text and they prefer to use natural language to express their opinion. It is difficult for a program to "understand" the text information and use these data. Several approaches using artificial intelligence techniques and text mining address the problem of identifying consumer's ratings for a product [10] [12]. However, there is a problem that has been less addressed in research until now but it is not less important, this problem is the identification of context information embedded in consumer's reviews.

In this paper we present a technique for detecting review's sentences containing contextual information. In future work we will study the incorporation and combination of such information and user's ratings in a recommendation. The rest of the paper is organized as follows: Section 2 provides a brief review of the literature on opinion mining. Section 3 presents the detail of the classification process and an illustrated exampled of its implementation. Finally, Section 4 concludes the paper and provides directions for future research.

2. Literature review

In the literature there are several researchers that analyze opinion and reviews to obtain consumer's information. In mining opinion, several approaches using artificial intelligence techniques and text mining address the problem of identifying consumer's ratings or opinion (positive or negative) for a product from consumer's reviews [10] [12]. Sentiment analysis is focused on the extraction of the relevance of product's feature based on sentiments of consumer reviews expressed in review sentences. Natural Language Processing (NLP) and supervised, unsupervised learning techniques, association rules have been used in sentiment analysis [12] [11 [18]. The sentiment classification relies with the classification of the reviews based on their polarity (positive or negative). Text mining and mutual information are used in sentiment classification [16][17] [15][1]. All of the previous approaches analyze reviews to extract product's features and classify opinions but they did not capture the context in which the reviews were expressed. In this work, we propose the use of text mining tools to obtain classification rules to identify contextual sentences containing contextual information into a review. The problem of determining whether a review or sentence express consumer's preferences is not easy to solve. It is also about context information. Without such information, any preference is of little practical use. So one should not only talk about preferences extraction from consumer's reviews, but also about the context information that preferences have been expressed upon. In this work contextual information is any condition that allow identify the temporal preference of a consumer for some product feature. For example the date when the review was written, weather condition that conditioning a trip, trip objective, etc. For example a consumer might prefer cheap hotels when he travels with his family for holidays but he prefers expensive hotels when he makes business trips. So, contextual information is the objective of trip: holidays and business. This information is easier for a consumer to express in a review where he can write using free text form. Thus, we need to be able to analyze the natural language text accurately to identify and extract user's preferences and the context on which preferences have been expressed. The proposed review classification process is presented in next section.

3. REVIEW CLASSIFICATION PROCESS

The classification process follows the implementation of text mining process described in [1] to classify review's sentences in digital camera domain into good, bad and quality categories. Once the sentences have been classify into one category, they can defines a set of metrics to obtain the rating that reviewer gives to some feature of a digital camera (positive or negative). But the preferences about some products can be changed according some situations, also can be contradictories. So in this case situational information (contextual information) has to be analised. In this paper we apply the text minig tools applied in [1] to define rules set that allow us to identify sentences that containing contextual information. As is defined in [1] shallow parser and classification algorithms based on term frequencies do not provide good results due the size of the sentences involved in the classification process. So, rule based classification techniques are employed. As described before, two categories have been defined to classify the sentences: "Contextual", and "Preferences". "Contextual" category groups those sentences that contain information about the context in which the review have been expressed. "Preferences" category groups those sentences that contain information about some features that consumer have evaluated. The Text-Miner Software Kit (TMSK) and the Rule Induction Kit for Text (RIKTEXT) have been used to obtain the classification rule sets [14]. Figure 2 shows the inputs and outputs of both miner tools. The best rule set is selected based on a combination of complexity and error-rate considerations. RIKTEXT finds the rule set with the minimum error-rate and then finds a less complex rule set whose error-rate is reasonably close to this minimum error-rate.

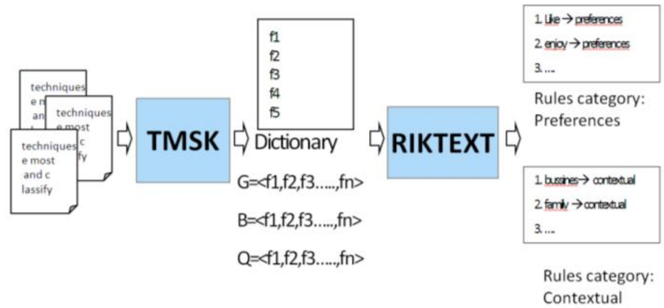

Figure 2. TMSK and RIKTEXT miner tools: input and outputs

3.1 Rules extraction

We analyzed reviews in tourism domain. Many tourism sites such as www.tripadvisor.com, www.virtualtourist.com, www.viajeros.com and www.travelpod.com enable consumers to exchange information, opinions and recommendations about destinations, tourism products and services, with sometimes diaries of travel experiences and ratings of a particular product or hotel. In a study made by TripAdvisor.com 83% of the user write travel reviews [9]. Online travel review writers are mostly motivated by a concern for other consumers, helping a travel service provider and needs for extraversion/positive self-enhancement. In [19] the role and impact of online reviews as useful tourist information providers are investigated. They found that 20% of consumers rely on other user's reviews when planning a trip and looking at other tourists' comments and travel blogs is the most popular online activity [9]. Decision making tools utilized in tourism sites need the automatic discovery, analysis and generalization of tourism consumer opinions, especially via the automatic recognition of tourist preferences and satisfactions when they consume tourism products. An example was conducted in the tourism domain where users write opinions about hotels, restaurants, trips, etc. The objective was getting a set of classification rules of "Contextual" and "Preferences" categories. The data we used are 100 reviews from www.tripadvisor.com arbitrarily selected from available reviews. Since the reviews were not in XML format, a special processing program was necessary to transform the data. Each sentence of each review is treated as a document. Once the data is in XML format, it is ready to be processed by TMSK to generate the dictionary and a set of labeled vectors.

A dictionary of 1250 words was generated and was used to generate vectors. The vectors have been splitted into training and tests portions. Test cases are selected randomly in RIKTEXT and we specified how many cases should be used for testing. We choose two-thirds of the available cases for training and the rest for testing. The results are presented in Table 1. As you can see, it displays a number of rule sets to classify review sentences in “Preference” category.

Each rule set is numbered under the column “RSet”. A “*” delineates the rule set with the minimum error rate. A “**” indicates the best rule set according to the error rate and simplicity. “Rules” is the number of rules in the rule set.

Table 1: Rule Set to Classify Sentences into Preference Category

Table of pruned rule sets
(* = minimum error; ** = within 0-SE of minimum error)

RSet	Rules	Vars	Train Err	Test Err	Test SD	MeanVar	Err/Var
1	33	160	0.1656	0.3181	0.0430	0.0	0.00
2**	25	30	0.1746	0.3181	0.0430	0.0	0.67
3	24	24	0.1785	0.3433	0.0435	0.0	2.00
4	22	22	0.2191	0.3679	0.0434	0.0	3.00
5	18	18	0.2439	0.3896	0.0430	0.0	4.00
6	1	1	0.2318	0.4565	0.0450	0.0	0.14

Random test cases, 412(33.3%) test cases
****

Selected rule set
1. like >=1 --> Preferences
2. prefer --> Preferences
3. interest >=1 --> Preferences
4. don't & like --> Preferences
5. disappointed --> Preferences
6. enjoy --> Preferences
7. favourite >=1 --> Preferences
8. want >=2 --> Preferences
9. sunny --> Preferences
10. not & wanted --> Preferences
11. look & for >=1 --> Preferences
12. find >=1 --> Preferences
13. use --> Preferences
14. feeling --> Preferences
15. interested & in --> Preferences
16. read & about >=1 --> Preferences
17. offer --> Preferences
18. wish --> Preferences
19. was >=1 --> Preferences
20. had --> Preferences
21. expect--> Preferences
22. loved --> Preferences
23. decision --> Preferences
24. ordered --> Preferences
25. passion >=1 --> Preferences
26. [TRUE] --> ~Preferences

Additional Statistics (Training Cases):
precision: 74.9533 recall: 89.1542 f-measure: 79.9880

Additional Statistics (Test Cases):
precision: 68.1159 recall: 75.6032 f-measure: 72.2121

“Vars” indicates the total number of conjuncts in the left-hand-side of the rules. The column “Train Err” gives the error-rate of the rule sets on the training data. “Test Err” is an error-rate estimate and Test SD is the standard deviation of the estimate. “Mean Var” is the average number of variables of the resampled rule set that approximates in size the rule set for the full data. “Err/Var” gives an indication of the quality of the solution. The chosen rules are those that have minimum error rate or are very close to the minimum but may be simpler than the minimum (**). Precision, recall and F-measure obtained from training and test cases are shown at the end of the table.

Table 2 shows the rule set obtained to classify review sentences in “Contextual” category. For each review sentence is performed a set of rules and if any rule can be applied the sentence is classify in this category.

Table 2: Rule Set to Classify Sentences into Contextual Category

Table of pruned rule sets
(* = minimum error; ** = within 0-SE of minimum error)

RSet	Rules	Vars	Train Err	Test Err	Test SD	MeanVar	Err/Var
1	73	160	0.0000	0.0193	0.0024	0.0	0.00
2	72	151	0.0007	0.0193	0.0024	0.0	1.00
3	42	87	0.0080	0.0158	0.0022	0.0	1.44
4	30	58	0.0118	0.0174	0.0023	0.0	1.97
5	29	54	0.0123	0.0168	0.0023	0.0	1.75
6	9	15	0.0201	0.0133	0.0020	0.0	2.54
7	8	13	0.0206	0.0143	0.0021	0.0	4.00
8	7	11	0.0211	0.0139	0.0021	0.0	6.00
9**	11	8	0.0218	0.0120	0.0019	0.0	6.33
10	5	6	0.0227	0.0149	0.0022	0.0	7.50
11	4	4	0.0236	0.0149	0.0022	0.0	11.00
12	3	3	0.0251	0.0139	0.0021	0.0	18.00
13	2	2	0.0301	0.0152	0.0022	0.0	60.00
14	1	1	0.0504	0.0200	0.0025	0.0	243.00

Random test cases, 412(33.3%) test cases
****

Selected rule set
1. wedding>=2 --> Contextual
2. anniversary>=2 --> Contextual
3. business>=4 --> Contextual
4. holiday>=1 & increase>=1 --> Contextual
5. children>=1 & dietary>=1 --> Contextual
6. situated & for --> Contextual
7. birthday --> Contextual
8. take & for --> Contextual
9. planning --> Contextual
10.christmas --> Contextual
11. weekend --> Contextual
12. [TRUE] --> ~Diet

Additional Statistics (Training Cases):
precision: 74.9533 recall: 89.1542 f-measure: 79.9880

Additional Statistics (Test Cases):
precision: 68.1159 recall: 75.6032 f-measure: 72.2121

A dictionary with related words and synonymous have been created to identify into reviews the words involved on rules due that the word involved in a rule can be written by the user in different ways. For example the word "loved" found in rule 22 to classify "Preferences" category can be write by the user in a review as "love".

3.2 Illustrated example

Once we have obtained the rule set to classify review sentences we have performed a controlled experimentation to evaluate the classification rule set. 50 new reviews from www.tripadvisor.com have been used to obtain the sentences that contain contextual information and the sentences with preferences' information. The amount of sentences involved in reviews varies between 1 and 14 sentences with an average of 6.5 sentences. The set of rules obtained in the previous section is applied to each sentence of the new reviews to classify it into one category. For example we applied the set rule in 1 of the 50 reviews, this review is shown in Figure 3.

The first sentence has been classified into the "Contextual" category. The second sentence has been classified into the "Preference" category and the last sentence is irrelevant because none of the rules has been applied as it is illustrated in the following

Sentence 1: I stayed there for a **business** trip and the **weekend** in mid February 2010.

Contextual rules: rule 3, rule 11
Preferences rules: none
Classification: CONTEXTUAL

Sentence 2: While **I've** been to Paris frequently I still struggled to **find** a hotel that is privately run and that **offers** good value with friendly staff.

Contextual rules: none
Preferences rules: rule 12, rule 17, rule 20
Classification: PREFERENCES

Sentence 3: The Apollon **offered** just this with a small but spotless bath room and a comfy bed and nice interior design.
Contextual rules: none
Preferences rules: rule 17
Classification: PREFERENCES

Sentence 4: It's located in the Montparnasse residential area so instead of views of the Champs Elysees you **find** a flower shop over the street and other essentials for Paris neighborhoods **like** brasseries with oysters up the street opposite the metro station

Contextual rules: none
Preferences rules: rule 1, rule 12
Classification: PREFERENCES

Sentence 5: Hope this helps you.

Contextual rules: none
Preferences rules: none
Classification: IRRELEVANT SENTENCE

Figure 3. One of the consumer's reviews from www.tripadvisor.com used in the case study.

Applying the rule sets on the 50 reviews we have obtained that 326 sentences have been classified of which 63 have been classified into the Contextual category, 71 into the Preferences category and 194 are irrelevant because none of the rules has been applied. This classification has been made using the automatic process described in previous sections. In order to evaluate the accuracy of the automatic classification we manually performed a classification process. For the 50 new reviews, we manually have evaluated each one of the sentences in order to identify if the sentences contain contextual information and preferences information. Comparing the result obtained using text mining process with the result obtained manually we can see that sentences of 8 reviews have been bad classified into Contextual category and sentences of 12 reviews have been bad classify into Preferences category. Analyzing these cases we have observed that the rules have been applied, however some rules are not specific enough to determine if the sentence refers to preferences' information. It is the case of the application of rule 1 from Table 1 on the sentence 4 of review show in Figure 3. The word "like" does not refer to a desire or wish, it refers to equal or equivalent. Another reason of the differences of results is that there are some "Contextual" and "Preferences" sentences that are not consider by the rules. Also some sentences have been classified in both categories. Once we have obtained the manual classification we have applied evaluation measures such as MAE (Mean Absolute Error), Precision, Recall and Fmeasure. Table 3 shows the result obtained in the classification of both categories. The bit difference obtained in the automatic classification according the manual classification mentioned before is evidence in the result obtained in Table 3.

Table 3: Resume of the results obtained in the experiment

	MAE	Precision	Recall	Fmeasure
Contextual	0,11	0,87	0,89	0,88
Preferences	0,01	0,91	0,90	0,91

4. CONCLUSIONS AND FUTURE WORK

This paper presents an automatic identification process of reviews containing information about user's preferences and information about the context in which this review was written. The identification of such information is not an easy task. The main problem is dealt with natural language used by reviewers to write their opinion. The process presented in this paper uses classification rules obtained from text mining tools.

The rules have been obtained for tourism domain where 100 reviews from www.tripadvisor.com have been used for training and test in the rule generation process. 50 new reviews from the same site have been used on an experiment to evaluate the accuracy of the rule to classify reviews sentences in these categories. The results obtained are considered good due the high value obtained in Precision, Recall and F-measure and the low value of MAE measure. Based on this result we can say that the automatic identification of contextual and preferences information can be made accurately using the text mining techniques presented in this papers. In further work we will try to refine the rules using stem dictionary in order to improve the classification process in sentences with words with different meaning such as like, have, etc. As the objective of this paper was the study of the identification of sentences containing contextual information and preference's information, the study of deriving further knowledge from these sentences and its incorporation in a recommendation process is under study now.

5. REFERENCES

[1] Aciar, S., Zhang, D., Simoff, S., and Debenham, J. 2006. Recommender System Based on Consumer Product Reviews. In Proceedings of the 2006 IEEE/WIC/ACM international Conference on Web intelligence. 719-723.

[2] Akehurst. G. User generated content: the use of blogs for tourism organisations and tourism consumers.Publisher Springer Berlin / Heidelberg.ISSN 1862-8516 (Print) 1862-8508 (Online)- pp. 51-61. 2007.

[3] Adomavicius, G. "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, vol 17, no. 6, 2005.

[4] Chevalier N. and D. Mayzlin: The effect of word of mouth on sales: Online book reviews, NBER Working Paper Series, National Bureau Of Economic Research, USA. 2003.

[5] Curien, N. E. Fauchart, G. Laffond and F. Moreau: Online consumer communities: escaping the tragedy of the digital commons, In Brousseau E. and N. Curien (eds), Internet and Digital Economics, Cambridge University Press, 2006.

[6] Dellarocas, C.: The digitization of Word-of-Mouth: Promise and challenges of online feedback mechanisms, Management Science, 49 (10): pp. 1407-1424. 2003

[7] Dellarocas, C.: Strategic manipulation of Internet opinion forums: Implications for consumers and firms, Working Paper, Sloan School of Management, MIT, Cambridge. 2004.

[8] Feldman, R. 2006 Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data.

[9] Gretzel, U. Online Travel Review Study. Role & Impact of Online Travel Reviews. Research Report. Laboratory for Intelligent Systems in Tourism. Texas A&M University. 2007.

[10] Liu, Bing. Web Data Mining. Exploring Hyperlinks, Contents, and Usage Data. 532 p. 177. ISBN: 978-3-540-37881-5. 2007.

[11] Liu, B., Hu, M., Cheng, J., 2005. Opinion Observer: Analyzing and Comparing Opinions on the Web. In Proceedings of the 14th International Conference on World Wide Web (WWW'05), Chiba, Japan, pp.342-351

[12] Pang, B. and Lee, L. Opinion Mining and Sentiment Analysis. Foundations and Trends in Information Retrieval. Vol. 2, Nos. 1–2 (2008) 1–135. 2008.

[13] Senecal, S and J. Nantel: The influence of online product recommendations on consumers' online choices. Journal of Retailing, 80, Elsevier, pp. 159-169.2004.

[14] Sholom M. Weiss, Nitin Indurkhya, Tong Zhang, and Fred Damerau. Text Mining: Predictive Methods for Analyzing Unstructured Information, Springer-Verlag, New York, 2004.

[15] Turney, P. Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews, Proceedings of the 40th Conference on Association for Computational Linguistics (ACL), Philadelphia, PA, 2002, pp.417-424.

[16] Wiebe, J. Bruce, R. and O'Hara, T. .Development and Use of a Gold Standard Data Set for Subjectivity Classifications, Proceedings of the 37th Conference on Association for Computational Linguistics (ACL), College Park, MD, 1999, pp.246-253

[17] Wei, C. Yang, C. S. and Huang, C .Turning Online Product Reviews to Customer Knowledge: A Semantic-based Sentiment Classification Approach, Proceedings of the 10th Pacific Asia Conference on Information Systems. Malaysia, 2006.

[18] Yang, C. C., Wong, Y. C., and Wei, C. 2009. Classifying web review opinions for consumer product analysis. In Proceedings of the 11th international Conference on Electronic Commerce ICEC '09. ACM, New York, NY, 57-63

[19] Zhang, L. Pan, B., Smith, W.W., & Li, X. (2009). An exploratory study of travelers' use of online reviews and recommendations: A qualitative approach. Journal of Information Technology and Tourism, 11(2): 157-167.