

A Novel Approach for Sentiment Analysis of Customer Reviews

Pankaj Vijay Ghule, Prof. Rugraj R.

Department of Computer Engineering, Alard College of Engineering and Management, Pune, India

ABSTRACT: with the rapid growth of the Internet the number of online reviews and recommendations is increasing. Both users and organizations use this data for their needs. Users check the reviews before purchasing any item so that they can compare between two or more items. Organizations use these reviews to understand the positive points and issues about their product and hence can make decisions accordingly. However, the reviews are generally disorganized and not ordered, causing difficulties in knowledge acquisition and information navigation. We propose a product aspect ranking framework, which will identify the important product aspects, in order to improve the usability of the various reviews. In particular, given the consumer reviews of a product, we will first identify product aspects and determine consumer opinions on these aspects via a sentiment classifier. We then develop a aspect ranking algorithm to understand the importance of aspects. We then weight these aspects and then decide the overall rating of the product.

KEYWORDS: Consumer reviews, Aspect identification, Sentiment classification, Aspect ranking, Product aspects.

I. INTRODUCTION

In recent years, the explosion of social networking sites, blogs and review sites provide a lot of information. Millions of people express uninhibited opinions about various product features and their nuances. This forms an active feedback which is of importance not only to the companies developing the products, but also to their rivals and several other potential customers. Sentiment Analysis is the task of tapping this goldmine of information. It retrieves opinions about certain products or features and classifies them as recommended or not recommended, that is positive or negative. The sentiment regarding a particular product in a review is seldom explicitly positive or negative; rather people tend to have a mixed opinion about various features, some positive and some negative. Thus the feature specific opinion matters more than the overall opinion. Consider a review “I like Micromax’s multimedia features but the battery life sucks.” This sentence has a mixed emotion. The emotion regarding multimedia is positive whereas that regarding battery life is negative. Hence, it is of utmost importance to extract only those opinions relevant to a particular feature (like battery life or multimedia) and classify them, instead of taking the complete sentence and the overall sentiment. In this work, we propose a method that represents the features and corresponding opinions in the form of a graph where we use dependency parsing to capture the relations between the features and their associated opinions. The idea is to capture the association between any specific feature and the expressions of opinion that come together to describe that feature. This is done by capturing the short range and long range dependencies between the words using dependency parsing. Clustering is done on the graph to retrieve only those opinion expressions that are most closely related to the target feature (user specified feature) and the rest are pruned. We apply merging in the final phase of our algorithm to merge the opinions about any 2 features that cannot be described independent of each other. We apply our method to domain specific reviews to test the efficacy of the system. We achieved a high accuracy across all domains over the baseline. We compare our approach with state-of-the-art systems [4] where we achieve a comparable accuracy despite data limitations. The system performance improved greatly not only over the naïve baseline but also over the chosen improved baseline

II. LITERATURE SURVEY

Chen et. al [1] use dependency parsing and shallow semantic analysis for Chinese opinion related expression extraction. They categorize relations as, topic and sentiment located in the same sub-sentence and quite close to each other (like the rule “an adjective plus a noun” is mostly a potential opinion-element relation), topic and sentiment

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 7, July 2016

located in adjacent sub-sentences and the two sub-sentences are parallel in structure (that is to say, the two adjacent sub-sentences are connected by some coherent word, like although/but, and etc), topic and sentiment located in different sub-sentences, either being adjacent or not, but the different sub sentences are independent of each other, no parallel structures any more.

Wu et. al [2] use phrase dependency parsing for opinion mining. In dependency grammar, structure is determined by the relation between a head and its dependents. The dependent is a modifier or complement and the head plays a more important role in determining the behaviors of the pair. The authors want to compromise between the information loss of the word level dependency in dependency parsing as it does not explicitly provide local structures and syntactic categories of phrases and the information gain in extracting long distance relations. Hence they extend the dependency tree node with phrases." Hu et. al [3] used frequent item sets to extract the most relevant features from a domain and pruned it to obtain a subset of features. They extract the nearby adjectives to a feature as an opinion word regarding that feature. Using a seed set of labeled Adjectives, which they manually develop for each domain, they further expand it using WordNet and use them to classify the extracted opinion words as positive or negative.

Lakkaraju et. al [4] propose a joint sentiment topic model to probabilistically model the set of features and sentiment topics using HMM-LDA. It is an unsupervised system which models the distribution of features and opinions in a review and is thus a generative model. Most of the works mentioned above require labeled datasets for training their models for each of the domains. If there is a new domain about which no prior information is available or if there are mixed reviews from multiple domains intermixed (as in Twitter), where the domain for any specific product cannot be identified, then it would be difficult to train the models. The works do not exploit the fact that majority of the reviews have a lot of domain independent components. If those domain independent parameters are used to capture the associations between features and their associated opinion expressions, the models would capture majority of the feature specific sentiments with minimal data requirement.

III. PROPOSED SYSTEM

We have proposed a product aspect ranking framework to find the vital aspects of products from various customer reviews. We add to a probabilistic viewpoint positioning calculation to construct the significance of different aspects by all the consumers. The modules can be classified as

- Pre-processing
- Product Aspect Identification
- Sentiment Classification
- Aspect Ranking.
- Negation Handling

A. Preprocessing :

The pre-processing module involves Tokenization, Stop-word Removal and Stemming.

Tokenization and Stop-word Removal: Tokenizing (i.e. breaking a string in its desired constituent parts) is fundamental to all NLP tasks. In lexical examination, tokenization is the procedure of separating a flood of content into words, expressions, images, or other important components called tokens. Multiple tokens will be the input for ahead or future processing like parsing or text mining. Stop words are words which are filtered out in pre or post phase of processing data.

Stemming:

Stemming is the term utilized as a part of data recovery to portray the procedure for decreasing curved (or in some cases derived) words to their oath stem, base or root shape for most part of composed word structure. Stemming programs are commonly known as stemming algorithms or stemmers. A straightforward stemmer turns upward the arched structure in a lookup table. The advantages of this approach is that it is efficient and fast.

Synonym Removal:

A synonym is a word that would mean exactly or nearly the same as another word in the same language.

Synonym may be present like headphone and earphone represent the same aspect. So these should be grouped as one aspect.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 7, July 2016

B. Product Aspects Identification:

Generally, a product can have several aspects. For instance, iPhone has aspects like appearance, applications, 3G network. Recognizing vital item aspects will enhance the ease of use of various reviews and is advantageous to both customers and firms. Clients can profitably make better decisions by taking into consideration important aspects, while firms can focus on improving the way of these perspectives and thusly redesign things reasonably. From the Pros and Cons reviews, we first identify the unique aspects because consumer uses different words for same aspect. So this will reduce the accuracy of ranking algorithm. So here we use synonym clustering to obtain unique aspects. We collect synonym terms of aspect as a feature. The isodata (Iterative Self Organizing Data Analysis technique) clustering algorithm is used to do synonym clustering.

C. Sentiment Classification

After the identification of important aspects the next step is sentiment classification. In this step the sentiments expressed on each aspect is identified. The sentiment is classified as a positive or a negative sentiment for that particular aspect. Thus we obtain aspects and sentiments related to those aspects. Dependency Extraction Algorithm is used for sentiment classification.

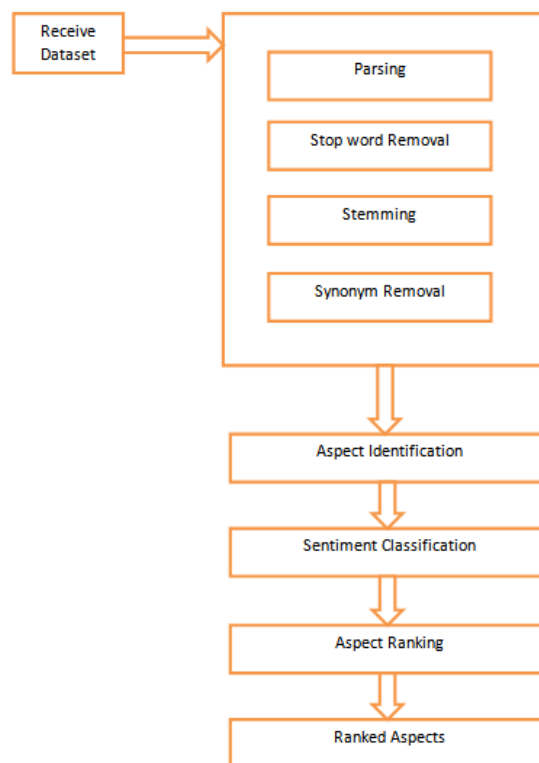


Fig 1: Proposed System Architecture

D. Aspect Ranking

Subsequent to performing Sentiment grouping we have a set of aspects alongside sentiments connected with them. Now we need to find weight of each of the aspects. TFIDF is a numerical measurement that is expected to reflect how vital a word is to a report, in an accumulation or corpus. The IDF is a measure of the amount of data the word gives, that is, whether the term is regular or uncommon over all docs.

TF: This indicates Term Frequency, which measures how frequently a term comes in a document. Since each data set is of variable length, there is a probability that a term would occur multiple times i.e. a larger number of times in long documents than shorter ones. In this manner, the term frequency is divided by the report length (otherwise known as the total number of terms in the report) as a method for normalisation.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 7, July 2016

IDF: Inverse Document Frequency measures how critical a term is. While figuring TF, all terms are viewed pretty much as of same importance. Regardless it is understood that particular terms, eg. "of", "is", "that", may show up a huge amount of times however have little worthiness. Thus we need to reduce weight of frequently occurring terms and scale upwards the rare terms.

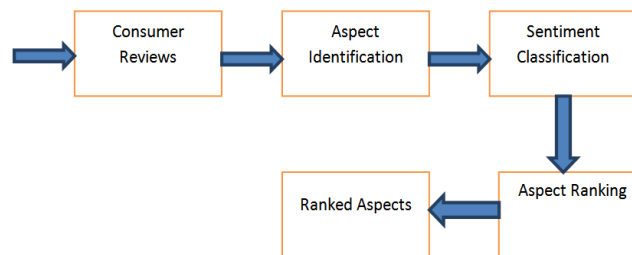


Fig 2: System Flow

E. Negation Handling & Linguistic Parsing

Negation simply means to reverse the polarity of lexical element besides an negator, changing good (+2) to not good (-2). We term this as switch negation. There are various subtitles that are related to negation and need to be considered. One is the consideration that there are negators, including not, none, nobody, nothing and other words, such as without or lack, which do have a similar effect, some of these could occur at more distance from the lexical item which they have effect, a reverse search is required to identify these negators, one which is related to that part-of-speech involved.

Also the user may enter his reviews in native languages. It is necessary to capture these sentiments as well. These reviews in native languages are passed to a linguistic parser which will output the English version of the sentiment and this is used for sentiment analysis.

Algorithm

Dependency Retrieval Algorithm for sentiment classification:

Consider there are 'n' features where the dimension of F is indicated by 'n'. The algorithm for obtaining the set of words from comment, that show the opinion related to the target feature f_t will be as below:

- i : Initialize all nodes as cluster C_i , as $1 \dots n$
- ii. Each C_i having a cluster head Ch which is denoted by the given cluster information based on the cluster features.
- iii. For each ($C_i \neq \text{null}$)
 - Extract the all features from Ch
 - If(C_i index is 0)
 - Create a new cluster otherwise
 - Calculate the weight of current cluster C_i and classified clusters.
- iv. Collect all weight list from all clusters.
- v. find best maximum weight cluster group.
- vi. Assign current cluster to highest weight group.
- vii. end for

Preprocessing Algorithm:

- Step 1: Review every comment C from dataset D.
- Step 2: Apply tokenization to input review
- Step 3: Perform stop-word removal functionality on C
- Step 4: Use stemmer algorithm to obtain the root words.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 7, July 2016

F. Dataset

Numerous experiments are being done on product comments, which we obtain from various web applications. First we created one web portal like e-commerce application where user can buy various products. For the users discussion purpose we made there one forum where user can enter or update his/her own comments about the specific products. The same data we have used for processing purposes, it is multidimensional data and we store it into mysql as well as .csv file format.

The most utilized datasets is from web applications, which contains survey of five specific electronic items (e.g., Nikon Quickpix 4500). Every sentence is physically commented on with viewpoint terms and also annotator agreement has been accounted for. Every sentence is chosen to express clear positive or negative suppositions. There are some sentences having clashing suppositions about viewpoint terms (e.g., "The screen is clear however little, its a 4.8-inch screen").

I. Proposed Results

For the proposed system performance evaluation, we calculate matrices for accuracy. We implement the system on java 3-tier MVC architecture framework with INTEL 3.0 GHz i7 processor and 8 GB RAM. Some user comments are positive or negative, and the data contains around 80,000 user's comments. The system finally classifies all the comments as positive, negative as well as neutral. Negation handling also works at the time of aspect classification. Here table 1 shows the estimated system performance with different existing systems. So, proposed results are around on satisfactory level.

Approach	Feature selection	Data Source	Accuracy
Lexical Resource	POS Apriori	Amazons customers Reviews	87.07%
Lexical Approach	Graph Distance Measurement	Users Blog Posts	82.85%
Hybrid	n-gram	Movie based review	90.05%
Naive Bayes	Information Gain	Canteen services reviews	91.75%
Naive Bayes and SVM	Based on minimum cuts	Movie reviews from users	85.90%
Proposed Approach	NLP and ML	Specific Product based Review	95.90% (Estimated)

Table 1: Performance Analysis of Proposed System

Table 1 shows the estimated result after completion of the system implementation.

IV. CONCLUSION

We have proposed sentiment classification approach based on product aspects and then ranking these aspects. The modules for this include Preprocessing Module, Product Aspect Identification, Sentiment Classification, Aspect

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 7, July 2016

Ranking. The preprocessing module contains the sub modules for tokenization, Stop word removal and Stemming. Synonym handling is done for aspects with same meaning but different name to avoid ambiguity. Sentiment classification is done based on aspects and ranking of these aspects is decided based on their calculated weight. We have also considered the Negation handling feature for obtaining better accuracy. We will also consider the intensity of the reviews to minimize the errors in classification. This framework will allow the user not only to judge between the different products but also the user can judge the individual aspects of the products.

V. ACKNOWLEDGMENT

I am thankful to Prof. Rugraj R. for guiding and suggesting valuable comments and helping me regarding project information. He not only helped for solving a number of issues but also gave me direction and guided me. I would also thank Prof. Archana, M.E. Coordinator Department of Computer Engineering, for encouraging me and giving me continuous guidance. I also want to thank Prof. Priyadarshini Doke for all her valuable guidance and assistance.

REFERENCES

- [1] Zheng-JunZha, Jianxing Yu, Jinhui Tang, Meng Wang, and Tat-Seng Chua. Product Aspect Ranking and Its Applications. IEEE transaction on data and knowledge engineering ,vol.26,no.5, Oct 2014.
- [2] N. D. Valakunde and Dr. M. S. Patwardhan 2013 "Multi-Aspect and Multi-Class Based Document Sentiment Analysis of Educational Data Catering Authorization Process". Book By Han and Kamber. Data Mining.
- [3]. Jiaxiong Wang and Andy Dong 2010 "A Comparison of Two Text Representations for Sentiment Analysis". "Centimeters-Br: a New Social Web Analysis Metric to Discover Customers Sentiment"
- [4]. Renate Lopes Rosa, Demstenes Zegarra Rodriguez., 2013 IEEE 17th International Conference on Department of Computer Engineering, MIT AOE
- [5]. "Sentiment Analysis on Tweets for Social Events" Xujuan Zhou and Xiaohui Tao, Jianming Yong., Proceedings of the 2013 IEEE 17th International Conference on Computer Supported Cooperative Work in Design.
- [6] "Sentiment Analysis in Twitter using Machine Learning Techniques" Neethu M S and Rajasree R., IEEE – 31661
- [7]. "Twitter Sentiment Analysis: A Bootstrap Ensemble Framework" Ammar Hassan* and Ahmed Abbasi+ and Daniel Zing., SocialCom/PASSAT/Big Data/EconCom/BioMedCom 2013
- [8]. "Text Feeling Analysis Algorithm Optimization Platform Development in Social network" Yiming Zhao, Kai Niue, Zhejiang He, Jiaru Lin, and Xinyu Wang., 2013 Sixth
- [9]. International Symposium on Computational Intelligence and Design. Sentiment Analysis: A Combined Approach Rudy Prabowo, Mike Thelwall.
- [10]. Osimo David and Mureddu Francesco, "Research Challenge on Opinion Mining and Feeling Analysis", ICT Solutions for power and policy modeling
- [11]. McDonald R., Hannan K., Neylon T., Wells M., and Reynar J., "Structured models for fine-to-coarse sentiment analysis," in Proceedings of the Association for Computational Syntax (ACL), pp. 432–439, Prague, Czech Republic: Association for Computational Linguistics, June 2007.
- [12]. Kamp G, Benamara F., Cesarano C., Picariello A., Reforgiato D. and Subramanian VS, "Sentiment Analysis: Adjectives and Adverbs are better than Adjectives Alone". ICWSM '2006 Boulder, CO USA
- [13]. Wilson T., Wiebe J. and Hoffmann P., "Recognizing Background Split in Phrase-Level Sentiment Analysis", Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing (HLT/EMNLP), pages 347–354, Vancouver, October 2005. c 2005 Association for Computational Syntax
- [14]. Liu B., "Sentiment Analysis and Subjectivity", Department of Computer Science, University of Illinois at Chicago, 2010.
- [15]. Frank E., Xia L, Bouckaert R. R., Zhen N: Bayes Naive for Text Classification with Unbalanced Classes, 2007
- [16]. Turney, Peter D. Thumbs up or thumbs down?: semantic orientation applied to unverified classification of reviews. in Proceedings of Annual Meeting of the Association for Computational Linguistics (ACL-2002). 2002