Lexicon based Approach for Sentiment Classification of User Reviews

Muhammad Zubair Asghar[1], Rahman ullah[1], Shakeel Ahmad[1], Fazal Masud Kundi[1], Irfan ullah Nawaz[1]

Institute of Computing and Information Technology Gomal University Dera Ismail Khan
Email: zubair@gu.edu.pk

Abstract: With the advent of web, online user reviews are getting more and more attention of the researchers because valuable information about products and services are available on social media like twitter1. These reviews are very helpful for organizations as well as for new customers showing interest in these products or services. But this data is generated in tremendous amount which is out of control of manual mining methods. These reviews need a model that has the ability to gauge these shared reviews according to predefined categories. This work introduces a rule based approach to find the opinion classification of reviews. The system can automatically crawl reviews from social media sites, classify these reviews as subjective and objective and then calculate polarity score for subjective reviews at word level. This method shows impressive results and out-performs the baseline method by achieving 86% and 82% accuracy at feedback and sentence level respectively for comments and 96% at feedback and 85 % at sentences for reviews.

[Muhammad Zubair Asghar, Rahman ullah, Shakeel Ahmad, Fazal Masud Kundi, Irfan ullah Nawaz. **Lexicon based Approach for Sentiment Classification of User Reviews**. *Life Sci J* 2014;11(10):468-473] (ISSN:1097-8135). http://www.lifesciencesite.com. 64

Keywords: Product review, Sentiment Analysis, Opinion Mining, Classification. Blog Mining. 1www.twitter.com

1. Introduction

Information technology and its rapid advancement make sure for the manufacturing firms to collect user views about their products and services to get help in designing new strategies and developmental tasks towards better change. Change includes design, development, marketing, and customer initiatives and applications (Buttle et al. 2003; Kumar 2005; MZ Asghar, 2009). It includes the collection of customer details, customer interactions, customer behavior, and customer preferences. The system also makes predictions for marketing and sales and identifying strategic aims (Berry et al. 1997; Ganapathy et al. 2004; Tseng & Huang, 2004)). In general Statistical survey is used for collecting customer details and for observing customer behavior (Van Bennekom, Frederick C (2002), Fowler, Floyd J (1995), Vavra, G Terry (1997)). Research conducted previously take into account the numerical as well as categorical data for recommending and personalizing specific product and analyzing criteria to make customer more loyal (Lee et al. 2007; Lin, Wen-Bao (2007); Lin et al. 2008).

Although the online analytical processing (OLAP) is used to analyze user reviews, yet data mining techniques used for this purpose are more efficient (Berry et al. 1997; Berson et al, 1997; Fayyad et al, 1996; Han et al, 2001; E Thomsen, 2002, MZ Asghar et al, 2014). User views shared in natural language are unstructured and semi-structured and it needs much knowledge from other related areas for handling these texts (B Lent et al, 1997; D Merkl et al, 1998, A Visa, 2001).

Textual data, just like numerical data, highlight numerous issues involved in natural language processing as well as promotes business intelligence and competitive intelligence. These techniques needs to be enhanced as the data size is getting large and large for online data e.g. memos, webpages and even short text messages. Customers share their ideas about products and services and make comparison with different products and services. Companies also get hundreds of emails from users commenting on the services and products. The designers are unable to get valuable information until they get better understanding of these reviews.

These reviews are helpful for buyers and they can use these reviews for decision making process. Understanding these reviews is not crucial but these reviews are present in massive amount so it is not feasible to process these reviews manually. So getting these reviews inside the circle we need an automated model that can handle this massive amount of data. Some researchers have worked to deal with these reviews and opinion mining is their specific focus (K Dave et al, 2003; Gamon et al, 2005; Huang, Chun-Che (2007); B Liu et al, 2005; A. M. Popescu, & O Etzioni, (2005); P.D Turney, 2002; MZ Asghar et al, 2014).

This study provides a rule based approach that can crawl user reviews from web platform and then classify these reviews into subjective and objective reviews. The subjective reviews are then categorized into positive and negative reviews at the end.

2. Previous Work

This section presents a related work conducted so for on sentiment analysis of twitter messages regarding product reviews.

Research on user reviews, posted on micro blogging networking sites, associated to products and services is very recent. In sentiment classification word, sentence, document semantic orientation is determined. A lot of work is done taking word as a basic unit of processing. [Vasileios Hatzivassiloglou and Kathleen R. McKeown (1997)] extracted opinions from adjective bearing sentences keeping in view the constraints of linguistics.. In 2002, WordNet approach emerges which used semantic distance in between words. [J. Kamps and M. Marx. 2002] proposed PMI (point wise mutual information) for measuring sematic distance between two words to facilitate the assessment of sentiment strength and [Turney et al. 2003] used a cosine method for calculating distance in LSA, it produce better results.

Many approaches have been adopted for performing sentiment analysis on social media sites. Knowledge based approaches classify the sentiments through lexicons in which sentiment polarity of words and linguistic patterns are defined (Turney et al, 2003). Specifically, for twitter sentiment analysis not a single approach has been used by researchers. In this regard hybrid approach has been adopted by combining methods based on lexicon with those based on ML and NLP techniques, in order to get advantage of both content as well as connectivity patterns among users (MZ Asghar, 2013).

Movie reviews are discussed and several machine learning approaches are used with common text features and sentiment classification at document level for the classification. An information retrieval classifier is introduced which is capable of feature extraction as well as scoring reviews. A combined approach of PMI, semantic orientation factors and structural relationship into the features of SVM are introduced. Another machine learning approach is introduced which highlights the opinion detection and minimum cut in graph.(Pang et al, 2002,2004; Dave et al, 2003; Tony Mullen and Nigel Collier (2004); CE Osgood et al, 1957;). [Bo Pang et al, 2005] further enhanced their idea by introducing multipoint scale. A comparative study of ML approaches and semantic orientation is carried out by [Pimwadee Chaovalit and Lina Zhou (2005)].

However, our approach to mine the user reviews corresponding with products and services and it provides a better and effective way for Lexicon based opinion mining.

3. Methodology

This work will find sentiment orientation of opinionated words present in user review; it is a combination of corpus based as well as dictionary based techniques. Features like emoticons and capitalization of words are also considered as they play role of intensifiers in user reviews and largely appear in the informal social media language. The overall data flow of the proposed system is given in figure 1.

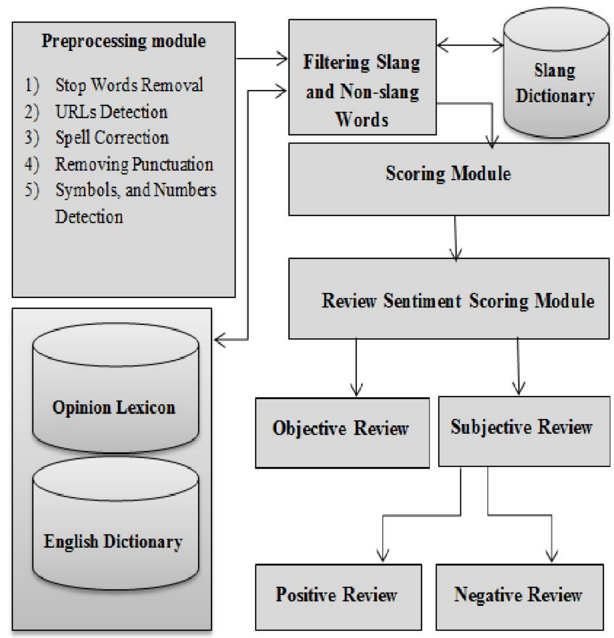

Fig.1: Model Diagram of the Proposed Methodology

3.1 Preprocessing Module

In this phase all URLs (WWW.example.com), and hash tags (#topic), are removed. Preprocessing module computes the fraction of the capitalized words. Spell correction module tries for spell correction and repeated character is tagged by a predefined weight. Emoticons are annotated manually and their scores are available in the annotated table. Exclamation marks are counted and remaining punctuation marks are removed. POS tagger[1] is used to tag verbs, adjectives and adverbs

Algorithm1. explains all these steps in detail.

Algorithm 1: Computing Polarity Score and Categorization of Subjective and Objective Reviews

[1] [1]http://www.infogistics.com/textanalysis.html

Input SentiWordNet, Slang_Dict, Lexicon
<W, pos, pol_score>:
R review
Thr threshold
Negation List = {not, never…}
Context-Shift-List = {but, however…}
Enhancer-Reducer-List = {slightly, very…}
Output: Word sentiment score,
Review sentiment score,
Objective reviews,
Subjective (positive, negative) reviews
Begin:
1. Get (W, POS, largest_sent_score) from SWN,slang lexicon;
2. For review R calculate (Fc, Ar, Ec);
3. Compute sentiment Sint(T) for intensifiers

$$S(R) = \frac{(1+\frac{Fc+ Ar+Ec}{3})}{N(op)}$$

4. For each review R compute (opinion groups) do
5. Get adjective groups (AGi)
6. Get verb groups (VGi)
7. Count emoticons (Ne)
8. Count slang words
9. Calculate sentiment score of opinion groups

$$S(T) = \sum_{i=1}^{N(op)} (\text{opinion groups})$$

10. End for
11. Calculate overall sentiment score of tweet

$$S(R) = \frac{(1+\frac{Fc+ Ar+Ec}{3})}{N(op)} * \sum_{i=1}^{OG(E)} \text{Score}(AGi + VGi + Nei * S(Ei) + S(SW) \qquad (3)$$

12. Return Sentiment of review,
13. if Abs (Score (R)) > Thr then
14. Return: R is subjective
15. Get (W, pol_score) from SWN;
16. for (i=1; i<=n; i++)
17. pos_score(W) ← pos_scorep(i) / n
18. neg_score(W) ← neg_scorep(i) / n
19. obj_score(W) ← obj_scorep(i) / n
20. End for
21. If pos_score(W) > neg_score(W)
22. max_pol_score(W) ← [pos_score(W)]
23. Else if neg_score(W) > pos_score(W)
24 max_pol_score(W) ← -[neg_score(W)]
25 Else
26. max_pol_score(W) ← [obj_score(W)p]
27. End if
28. if W preceded by NL then
29 max_pol_score (W) ← pol_score(W) * -1;
30. if W preceded by ERL then
31. max_ pol_score (W) ← pol_score(W) + getERL(enhancer_reducer_Word, score);
32. if W preceded by CSL then
33. max_pol_score (W) ← pol_score(W)+ getCSL(context_shifter_Word, score);
34. Return: word sentiment score
35 End if
36. Else
37. Return: T is objective
38. End if
39. Return: R.S (Sentiment score).
40. End For
41 End begin

3.2 Adjectives, Adverbs and Verbs Sentiment Orientation

Dictionary based methods handle semantic orientation of verbs and adverbs while semantic orientation of adjectives is carried out through corpus based method. Since semantic orientation of adjective is domain dependent, therefore we apply corpus based methods to manipulate it in product reviews domain.

In this work following equation 1 adopted from the work of [Vasileios Hatzivassiloglou and Kathleen R. McKeown (1997)] is applied to predict opinions.

$$\eta = W\,T\,X \qquad (1)$$

In equation 1, "X" is the vector of observed counts in the various conjunction categories for a specific adjective pair and "W" represents a weight vector learnt during training. "Y" is the response which is non-linearly related to η through the inverse logit function.

$$Y = \frac{e^{\eta}}{1 + e^{\eta}} \qquad (2)$$

In equation "2", "y" presents correlation between words. Initially, seed list of adjectives with assigned values and similarity values are used to compute the conjoined semantic score "y" of adjectives.

Verbs are also sentiment carriers (like, dislike). Lexicon based methods are used to compute semantic orientation of verbs and adverbs as they are domain free. Initial seed list containing positive and negative sentiment score of verbs and adverbs, is extended by using WordNet (J. Kamps and M. Marx. 2002). Other commonly used verbs and adverbs are manually annotated and values domain is from -1 to +1. For example; one user says, "This is beautiful pen" and; other says, "This is so beautiful pen".

3.3 Handling Slang Words

The informal nature of social media is a great diversion in the field of mining. Slang words are also lies under the factor discussed and needs to handle in some specific way because slang words also convey the author ideas. Slang a language of generating non-standard and irregular words and phrases

(Wikipedia[2]) such as hahaha, b4. A module is introduced which will identify slang and non-slang regular words. A slang dictionary is applied for handling slangs, while the WordNet and SentiWordNet are applied for regular words.

3.4 Overall Tweet Sentiment

Adjectives and nouns are grouped together and named as adjective group, while verb group combines verbs and adverbs. Score of adjective group is determined by multiplying adjective score and noun score obtained from corpus and lexical resources, similarly verb group score is obtained by multiplying verb score and adverb score. Slang word lexicon is introduced in which slang definitions are provided so for handling slangs the system will contact with slang lexicons. In absence of terms its default score is inserted 0.5. The average of all the opinion intensifiers (capitalization, word emphasis, adjectives groups, verb groups, emoticons, exclamation mark) is calculated according to equation 3 given below:

$$\mathrm{S(R)} = \frac{\left(1+\frac{\mathrm{Fc+\,Ar+Ec}}{3}\right)}{\mathrm{N(op)}} * \sum_{\mathrm{i=1}}^{\mathrm{OG(E)}} \mathrm{Score(AGi +} \quad \mathrm{VGi+ \quad Nei*SEi+S(SW)} \qquad (3)$$

In equation 3, N (op) represents total number of opinion groups present as well as emotion icons in the tweet. "F_c" fractions of words capitalized "Ar" shows the count of repeated alphabets. "Ec" count of exclamation marks. "W(AGi)" ith adjective group, "W(VGi)" ith verb group, "W(Ei)" ith emoticon, "Nei" ith emoticon and "S(SW)" is used for score of slang word[2].

Fc, Ar, Ec are named as sentiment intensifiers. Reviews score is re-arranged to 1 and -1 if they crossed the mentioned limits.

4. Result and Discussion

To analyze the mechanism and results of the proposed methodology we conduct an experiment. For evaluation we extract 625 product reviews from social media. The system executes the dataset and categorized reviews into 495 opinionative and 130 non-opinionative reviews, opinionative reviews are further classified into 466 positive and 29 negative reviews as shown in Table 1. All the execution is carried out in accordance with the proposed methodology mentioned in section 3.

Table 1: Categorization of opinionative and non-opinionative Reviews

Dataset	**Reviews**	**Non-Opinionative Reviews**	**Opinative Reviews**	
Twitter	**625**	**130**	495	
			Positive	**Negative**
			466	29

4.1 Evaluation Techniques

The evaluation techniques are similar to those discussed in earlier research work by (M.Z Asghar et al., 2014) (FM Kundi et al., 2014)

A manually annotated data set is used for classifying opinion classes under the instructions of the techniques discussed in previous research of (Bo Pang, and Lee Lillian (2008), Turney Peter, 2002). The evaluating techniques (precision, accuracy, recall and F-measure) for the above research are given in table 2 and 3.

Table 2: Accuracy of Positive and Negative Sentences

Dataset	**Sentences**	**Total no. sentences**	**Positive sentences**	**Negative sentences**	**Accuracy**
Twitter	**Positive**	310	249	61	0.803
	Negative	185	35	150	0.810

Table 3: Computation of Precision (P), Recall (R), and F-Measure (F)

	Positive			**Negative**		
	P	**R**	**F**	**P**	**R**	**F**
Harry Potter	.456	.418	.436	.822	.631	.714
R. Swaminathan et al	.770	.780	.774	.510	.660	.574
Political Miner:	.780	.790	.794	.540	.680	.601
Proposed approach	**.876**	**.803**	**.837**	**.710**	**.810**	**.756**

5. Conclusion and Future Work

Opinions are a special type of information users shared on the twitter about any product or service currently introduced and it is totally different from facts. Content classification methods are not effective enough as there is large gap between the opinions of different persons. During the evaluation of product reviews we come to know that it is feasible to develop a mechanism that will classify the user

[2] [2]http://en.wikipedia.org/wiki/Slang/ <Accessed:01/2014.

opinion about any product or service, but one thing is clear that categorizing a large number of opinions is a tedious work due to variation in opinions as well as inflections in language. This approach can be combined with other approaches to get higher degree of accuracy.

This work can be further extended by applying lexical rules for the natural language by using appropriate language patterns. Opinionated text can be enhanced and made easy by incorporating semantic module using intelligent framework.

References

1. Buttle F. Customer relationship management: concepts and technologies. 2009
2. Berry MJ, Linoff G. Data mining techniques: for marketing, sales, and customer support. John Wiley & Sons, Inc. 1997.
3. Ganapathy S, Ranganathan C, Sankaranarayanan B. Visualization strategies and tools for enhancing customer relationship management. Communications of the ACM. 2004; 47.11: 92-99.
4. Huang CC. "Rough set-based approach to feature selection in customer relationship management." Omega. 2007; 35.4: 365-383.
5. Van B, Frederick C. Customer surveying: a guidebook for service managers. Customer Service Press. 2002.
6. Fowler FJ. Improving survey questions: Design and evaluation. 1995; 38.
7. Lee S, Lee S, Park Y. "A prediction model for success of services in e-commerce using decision tree: E-customer's attitude towards online service." Expert Systems with Applications. 2007; 33.3: 572-581.
8. Lin, W. "The exploration of customer satisfaction model from a comprehensive perspective." Expert Systems with Applications. 2007; 33.1: 110-121.
9. Lin C, Hong C. Using customer knowledge in designing electronic catalog. Expert systems with Applications. 2008; 34.1: 119-127.
10. Gamon M, Aue A, Corston-Oliver S, Ringger E. Pulse: Mining customer opinions from free text. In Proceedings of advances in intelligent data analysis VI, 6th international symposium on intelligent data analysis IDA. Madrid, Spain. 2005: 121–132.
11. Zhan J, Han TL, Ying L. Gather customer concerns from online product reviews–A text summarization approach. Expert Systems with Applications. 2009; 36.2: 2107-2115.
12. Kundi FM, Ahmad S, Khan A, Asghar MZ. Detection and Scoring of Internet Slangs for Sentiment Analysis Using SentiWordNet. Life Science Journal. 2014; 11(9).
13. Zhan, Jiaming, Han Tong Loh, and Ying Liu. "Gather customer concerns from online product reviews–A text summarization approach." Expert Systems with Applications. 2009; 36.2: 2107-2115.
14. Asghar MZ, Qasim M, Ahmad B, Ahmad S, Khan A, Khan IA. Health miner: opinion extraction from user generated health reviews. international Journal of academic research part a. 2013; 5(6): 279-284.
15. Berson A, Smith SJ. Data warehousing, data mining, and OLAP. McGraw-Hill, Inc.1997.
16. Fayyad UM, Piatetsky-Shapiro G, Smyth P. From data mining to knowledge discovery: An overview. In U. M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, & R. Uthurusamy (Eds.), Advances in knowledge discovery and data mining. Menlo Park, CA, USA: American Association for Artificial Intelligence. 1996: 1–34.
17. Han J, Kamber M, and Pei J. Data mining: concepts and techniques. 2006.
18. Thomsen E. OLAP solutions: Building multidimensional information systems (2nd ed.). Wiley. 2002.
19. Asghar MZ, Khan A, Ahmad S, Kundi FM. A Review of Feature Extraction in Sentiment Analysis. Journal of Basic and Applied Scientific Research. 2014: 4(3):181-186
20. Lent, B., Agrawal, R., & Srikant, R. (1997). Discovering trends in text databases. In Proceedings of the third international conference on knowledge discovery and data mining (pp. 227–230).
21. Merkl D. Text Data Mining. Dale R, Moisl H, Somers H.(eds.). A handbook of natural language processing: techniques and applications for the processing of language as text. 1998.
22. Visa A. Technology of text mining. In Proceedings of machine learning and data mining in pattern recognition, second international workshop, MLDM. Leipzig, Germany. 2001: 1–11.
23. Dave K, Lawrence S, Pennock DM. Mining the peanut gallery: Opinion extraction and semantic classification of product reviews. In Proceedings of the 12th international conference on World Wide Web. Budapest, Hungary. 2003: 519–528
24. Liu B, Hu M, Cheng J. Opinion observer: Analyzing and comparing opinions on the Web. In Proceedings of the 14th international conference on World Wide Web. Chiba, Japan. 2005: 342–351.
25. Popescu AM, Etzioni O. Extracting product features and opinions from reviews. In

Proceedings of human language technology conference and conference on empirical methods in natural language processing, HLT/EMNLP. Vancouver, BC, Canada. 2005: 339–346.
26. Asghar MZ, Khan A, Kundi FM, Qasim M, Khan F, Ullah R, Nawaz IU. Medical opinion lexicon: an incremental model for mining health reviews. International Journal of Academic Research Part A. 2014; 6(1): 295-302.
27. Turney PD. Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews. In Proceedings of the 40th annual meeting on association for computational linguistics. 2002: 417-424
28. Asghar MZ, RahmanUllah, Ahmad B, Khan A, Ahmad S, Nawaz IU. Political miner: opinion extraction from user generated political reviews. Sci.Int(Lahore). 2014; 26(1): 385-389.
29. Hatzivassiloglou V, McKeown KR. Predicting the semantic orientation of adjectives. In Proceedings of ACL. 1997: 174-181.
30. Kamps J, Marx M. Words with attitude. In Proc. of the First International Conference on Global WordNet. 2002: 332-341.
31. Turney PD, Littman ML. Measuring praise and criticism: Inference of semantic orientation from association. ACM Transactions on Information Systems (TOIS). 2003; 21.4: 315-346.
32. Asghar MZ, Khan A, Ahmad S, Kundi FM. Preprocessing in natural language processing. Journal of Basic and Applied Scientific Research.; 2014; 4(3):181-186.
33. Pang B, Lee L, Vaithyanathan S. Thumbs up?: sentiment classification using machine learning techniques. In Proceedings of the ACL-02 conference on Empirical methods in natural language processing. Association for Computational Linguistics. 2002; Volume 10: 79-86.
34. Mullen T, Collier N. Sentiment analysis using support vector machines with diverse information sources. In Proceedings of EMNLP. 2004 : 412-418.
35. Osgood CE. The Measurement of Meaning. University of Illinois. Press 1957.
36. Pang B, Lee L. A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. In Proceedings of ACL. 2004: 271-278.
37. Pang B, Lee L. Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales. In Proceedings of ACL. 2005: 115-124.
38. Chaovalit P, Zhou L. Movie review mining: A comparison between supervised and unsupervised classification approaches. In Proceedings of HICSS. 2004; 4.
39. Pang B, Lee L. Opinion mining and sentiment analysis. Foundations and trends in information retrieval. 2008; 2(1-2): 1-135.
40. Asghar MZ, Khan AR, Asghar MJ. Computer assisted diagnoses for Red Eye (CADRE). International Journal of Computer Science and Engineering. 2009; 1(3): 163-170.