Survey on Opinion Mining and Summarization of User Reviews on Web

Vijay B. Raut
P.G. Student of Information Technology,
Pune Institute of Computer Technology,
Pune, India

Prof. D.D. Londhe
Assistant Professor in Dept. of Information Technology,
Pune Institute of Computer Technology,
Pune, India

Abstract— **Large amount of user generated data is present on web as blogs, reviews tweets, comments etc. This data involve user's opinion, view, attitude, sentiment towards particular product, topic, event, news etc. Opinion mining (sentiment analysis) is a process of finding users' opinion from user-generated content. Opinion summarization is useful in feedback analysis, business decision making and recommendation systems. In recent years opinion mining is one of the popular topics in Text mining and Natural Language Processing. This paper presents the methods for opinion extraction, classification, and summarization. This paper also explains different approaches, methods and techniques used in process of opinion mining and summarization, and comparative study of these different methods.**

Keywords— Natural Language Processing, Opinion Mining, Opinion Summarization.

I. INTRODUCTION

Due to growth of web 2.0 and social networks many user access and share data on Internet every day. Social networking sites such as Facebook, Twitter provides their user to send messages to friends, share photos, text information to friends. Users share their views about particular product, movie, hotel, and location on review websites such as www.epinion.com, www.yelp.com, www.tripadviser.com. This creates large source of information i.e. user generated contents in the form of blogs, comments, reviews, wikis, photos etc. This user generated data can act as important source for web mining because it contains user's opinion about particular topic.

Opinion mining also called sentiment analysis is a process of finding users opinion about particular topic. A topic can be a news, event, product, movie, location hotel etc. It involves opinion information retrieval, opinion classification and opinion aggregation. It is one of the booming research topics nowadays [2] [9]. Opinion mining can be used for product feedback analysis, and for decision making to users. Opinion retrieval is a process of collecting opinion text from review websites. Opinion text is subjective information in review, blog, tweet, micro-blog, comment etc.

Opinion classification involves classifying review text into two forms as positive or negative sentiment review. Machine learning is used for text classification using different classification algorithms such as Decision Tree [1], Naive Byes[8], and Support Vector Machines[4] etc. This approach applied for review text classification of data such as movie, product, and news reviews. Sentiment

Analysis is also done by other methods such as dictionaries [5], word lexicons [10], word senses etc.

Opinion summarization is process of generating summary from multiple reviews. It is based on feature selection [4], feature rating [1] and identifying sentence that contain feature. Opinion mining is a topic in Text mining, Natural Language Processing (NLP), and Web mining discipline. Though user generated contents has proven useful in many applications, challenges still exist in process of opinion mining due to unstructured and noisy data on websites [7]. Everyday creation of user-generated content in a large scale and because it involve user attitude there is need of opinion mining.

This paper presents different methods and approaches used in process of opinion mining and opinion summarization. This paper also explains outcomes and comparison of these approaches. Such comparison will be helpful to find methods that best suits for required scenarios in opinion mining.

II. RELATED WORK

Pang and Lee [9] presented survey on sentiment analysis and opinion mining. In that survey they explained opinion oriented information access, challenges, opinion classification and summarization. Mikalai Tsytsarau, Themis Palpanas [2] also have presented Survey on opinion mining. In that survey author explained opinion mining, opinion aggregation and subjectivity analysis. Their study mentioned different work performed on this issue and their comparisons.

Earlier sentiment analysis performed on different domain data such as Movie [11], Products [12] [13], Restaurants [1], and Travel [5] etc. Many authors applied opining mining concept to social network data. In [3] author worked on sentiment analysis of Facebook data from messages written by users. Many researches [15] developed sentiment analysis applications on twitter data. Other issues in opinion mining are emotion recognition, opinion spam detection [2]. In [5] author proposed different methods such as machine learning, machine translation and dictionary for sentiment analysis of text in Hindi language.

Many researchers used machine learning methods for sentiment analysis [3] [4] [8] that involve training of classifier on datasets and use the trained model for new document classification. Some authors suggested another method such as dictionary of word lexicons [6] [7]. The Dictionary approach is based on a prebuilt dictionary that contains opinion polarity values of words. Many resources

are developed in this domain are datasets [9], polarity annoted corpus [5], dictionaries [14]. SentiWordNet [10] is a resource that contains word polarity values which is based on WordNet [14] dictionary.

III. OPINION MINING

Opinion Mining also called sentiment analysis is a process of finding user's opinion towards a topic. Opinion mining concludes whether user's view is positive, negative, or neutral about product, topic, event etc [2]. Opinion mining involves analyzing user's opinion, attitude, and emotion towards particular topic. This consists of first categories text into subjective and objective information, and then finding polarity in subjective text [2]. Opinion mining can be performed word, sentence or document level.

Opinion mining and summarization process involve three main steps, first is Opinion Retrieval, Opinion Classification and Opinion Summarization. Different tasks involved in this process are shown in Fig. 1. Review Text is retrieved from review websites. Opinion text in blog, reviews, comments etc. contains subjective information about topic. Each crawled reviews are treated as a document vector. Reviews classified as positive or negative review. Opinion summary is generated based on features opinion sentences by considering frequent features [1] [6].
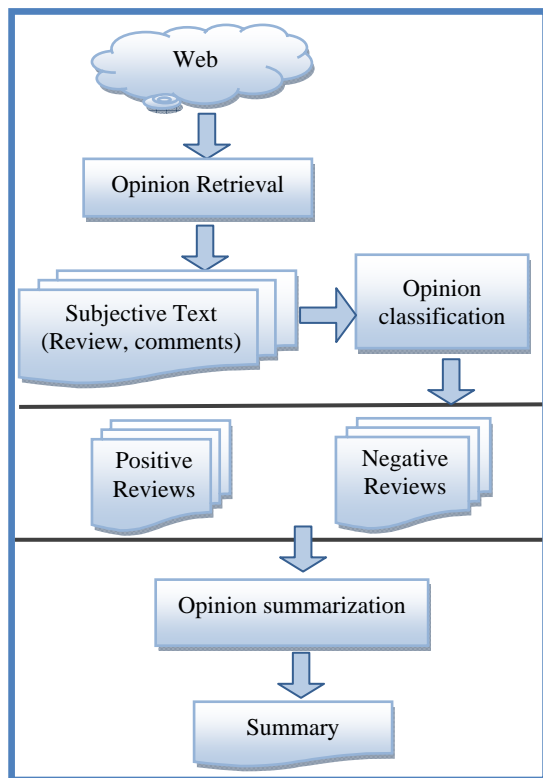

Fig. 1: Architecture of opinion mining and summarization.

A. Opinion Retrieval

Opinion retrieval is a process of collecting reviews text from review websites. Different review websites contain reviews for products, movies, hotels, news etc. Information retrieval techniques such as web crawler [6] can be applied to collect review text data from many sources and store them in database. This step involves retrieval of reviews, micro-blogs, comments etc of user. We should only consider the data which contain subjective data but not the objective data. Reviews are retrieved by query based information retrieval techniques [6].

B. Opinion Classification

Primary step in sentiment analysis is classification of review text. Given a review document $D = \{d_{1...}d_1\}$ and a predefined categories set $C = \{positive, negative\}$, sentiment classification is to classify each d_i in D, with a label expressed in C [2]. There are many approaches for sentiment classification in opinion text. Machine learning and lexicon based approach is more popular.

1) Machine learning approach for opinion classification

The machine learning approach uses supervised learning method for classification of review text. The first step is to train a classifier using sample of reviews with its class (positive/negative). Then the built model of trained classifier is used to predict category of new text reviews. Popular machine learning classifiers for text categorization are Support Vector Machines (SVM) and Naive Bayes (NB).

a) SVM classifier for opinion classification

SVM is a machine learning classifier widely used for text categorization. The review text to be classified is converted into word vectors. SVM constructs a hyper-plane using these vectors which separates data instances of one class from another. SVM finds this hyper-plane using training instances also called support vectors. In the binary categorization of text, the hyper-plane which classifies document d_j as $C_j \in \{1, -1\}$ can be represented by weight vector $\vec{w}$ [11].

$$\vec{w} := \sum \alpha_j\, c_j \overrightarrow{d_j},\ \alpha_j \geq 0, \quad (1)$$

Where α_j is a multiplier and for $\overrightarrow{dj}$ that α_j are greater than zero are support vectors [11]. Test instance is classified by determining which side of $\overrightarrow{w}$'s hyper-plane they fall on.

b) Naïve Bayes classifier

Naïve Bayes classifier categories text based on Bayes theorem of posterior probability [11]. If D is dataset of training data instances $X = (X_1,\ X_2\ ...\ X_n)$ having n attributes and m classes $C_1,\ C_2\ ...\ C_n$. The classifier predicts text X belongs to class having higher probability values for given conditions. This is shown in equation (2)

$$P(C_i/X) > P\big(C_j/X\big)\ for\ 1 \leq j \leq m\ , j \neq i \qquad (2)$$

Where $P(C_i/X)$ calculated using Bayes theorem,

$$P(C_i/X) = \frac{P(X/C_i)P(C_i)}{P(X)} \qquad (3)$$

2) Lexicon based approach for opinion classification

The lexicon Approach predicts sentiment of review text using databases which contain word polarity values e.g. SentiWordNet [10]. Review text is classified by calculating and averaging polarity score of individual words in sentences. Many factors such as word position, word relationships, negation handling should be considered while sentiment classification using lexicon based approach. Dictionary based approach works well in short reviews text. Fig. 2 explains lexicon based approach for classification of reviews into positive or negative class. Review text is classified based on sentences polarity score.

An equation that calculate sentiment polarity of review text using a resource [2] is

$$S(D) = \frac{\sum_{w \in D} S_w \cdot weight(w) \cdot modifier(w)}{\sum weight(w)} \quad (4)$$

In equation 2, S_w is a polarity score value of word w. This value is generated for dictionary using function $weight$ (). Issues such as negation handling, word position in sentence etc are handled by operator $modifier$ ( ) [2].
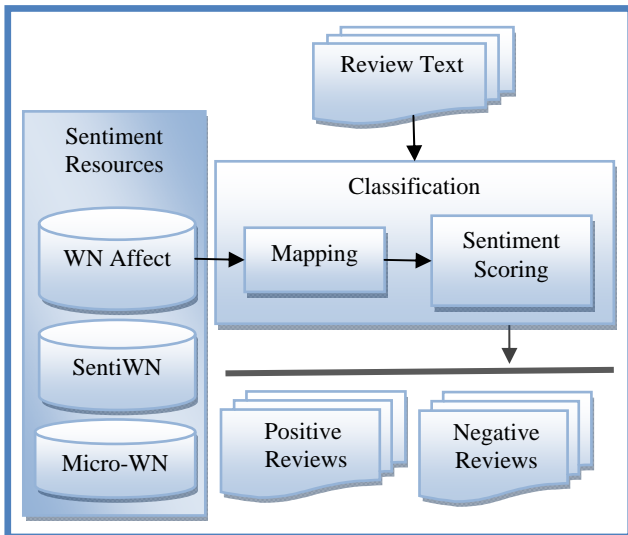

Fig. 2: Lexicon based approach for opinion classification

C. Opinion Summarization

Summarization of opinions is a major part in opinion mining process. Summary of reviews provided should be based on features or subtopics that are mentioned in reviews. Therefore, feature extraction [4] and opinion summarization are key issues. Many researchers worked on summarization product reviews [2]. The opinion summarization process mainly involve following two approaches.

1) Feature based summarization

This type summarization involve finding of frequent terms (features) that are appearing in many reviews. The summary is presented by selecting sentences that contain particular feature information. Features present in review text can be identified using Latent Semantic Analysis (LSA) [4] [7] method. For a short summary of product reviews, product features and opinion words associated

with it can be presented as summary. Sentences in which feature and opinion words are present are displayed in summary of reviews.

a) LSA method for feature identification

This method forms a term-document co-occurrence matrix. In this matrix terms represents rows and column represent documents. This matrix shows term frequency of every term in a document. By applying singular value decomposition (SVD) method [4] to the above matrix as $A = USV^T$ where U and V are matrices with orthonormal columns (i.e. $UU^T = VV^T = I$) and S is a diagonal matrix of A. Rows of resultant matrix represents most important terms(features) [7]. Sentences that contain these terms can be presented in a summary

2) Term Frequency based summarization

Term frequency is count of term occurrences in a document. If a term has higher frequency it means that term is more import for summary presentation. In many product reviews certain product features appear frequently and associated with user opinions about it. In this method sentences are scored by term frequency [6]. The summary is presented by selecting sentences that are relevant and which contain highest frequency terms.

Opinion Summarization process is shown in Fig. 3. It shows review text is preprocessed which involve sentence segmentation and tokenization of sentence in terms. After calculating term frequency of each term, each sentence score and relevance is calculated. As per the compression rate highest scoring and relevant sentences are presented in summary.
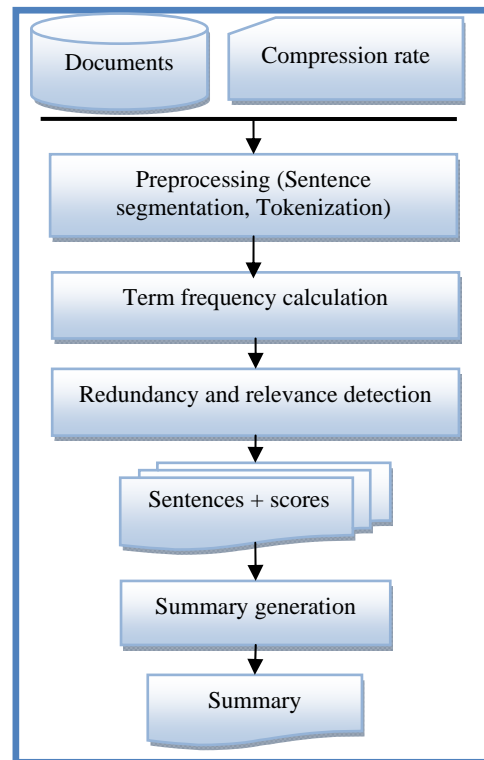

Fig. 3: Opinion summarization process

IV. COMPARISON OF METHODS

Opinion classification and summarization methods are evaluated based on following parameters

A. Evaluation of Opinion Classification

Review text classified as positive or negative by opinion classification method. Effectiveness of method is determined by precision or recall values.

$$Precision = \frac{TP}{TP+FP} \qquad (6)$$

$$Recall = \frac{TP}{TP+FN} \qquad (7)$$

Where,

- True positives (TP) - number of reviews correctly labeled as belonging to particular class (positive/negative).
- False positives (FP) - number of reviews incorrectly labeled as belonging to particular class.
- False negatives (FN) - number of reviews were not labeled as belonging to the particular class but should have been labeled.

B. ROUGE metric for Opinion Summarization

ROUGE (Recall-Oriented Understudy for Gisting Evaluation) metric compares automated summary generated by algorithm with set of reference summaries created by individuals. ROUGE is calculation of recall value between human and automated summary.

$$ROUGE - N = \\ \frac{\sum_{S \in \{Reference\ Summaries\}} \sum_{gram_n \in S} Count_{match}(gram_n)}{\sum_{S \in \{Reference\ Summaries\}} \sum_{gram_n \in S} Count(gram_n)} \qquad (5)$$

Equation (5) shows ROUGE metric, where n is word n-gram length, $gram_n$, and $Count_{match}(gram_n)$ is the maximum number of   n-grams co-occurring in a automated summary and a set of human summaries.

Table 1, gives synopsis of different approaches used for opinion mining and summarization and results obtained by them.

TABLE I- SYNOPSIS OF SOME METHODS USED FOR OPINION MINING AND SUMMARIZATION.

Author	Methods	Dataset	Result (Accuracy %)
Aditya Joshi, Balamurali A.R., Pushpak Bhattacharyya [5](2010)	Machine Learning, Machine Translation (MT), Hindi SentiWordNet dictionary	Travel Reviews	In – language sentiment analysis-78.14%, MT - based sentiment analysis - 65.96%, Resource based - sentiment analysis-60.31%
Chien-Liang Liu, Wen-Hoar Hsaio, Chia-Hoang Lee, Gen-Chi Lu, Emery Jou [4](2012)	SVM classifier, LSA method for feature based summarization	Movie Reviews	SVM classifier - 85.4%

Author	Methods	Dataset	Result (Accuracy %)
Jingjing Liu, Stephanie Seneff, Victor Zue [1](2012)	Word phrase Extraction, word phrase sentiment scoring and classification for sentiment analysis	Hotel Reviews	Decision Tree classifier for word phrase classification - 77.9%
Alexandra Balahur, Mijail Kabadjov, Josef Steinberger, Ralf Steinberger, Andrés Montoyo [7](2012)	Sentiment dictionary Resources, LSA based opinion summarization	Blogs	Sentiment analysis and Summarization ROUGE R1 negative class - 0.268, positive class-0.275
Elena Lloret, Alexandra Balahur, José M. Gómez, Andrés Montoyo, Manuel Palomar [6](2012)	Sentiment Resource, Machine learning, Term Frequency based summarization	Product Reviews	Summary-ROUGE-1 (10% compression) Precision-30.16, Recall-20.54,
Alvaro Ortigosa, José M. Martín, Rosa M. Carro [3](2013)	Machine Learning, Word Lexicons	Facebook Messages	Sentiment analysis-83.27%
Alexandra Trilla, Francesc Alias[8](2013)	Machine Learning	Semeval Dataset, Twitter Dataset	SVM-58.12% - Semeval Dataset, SVM-72.76% - Twitter Dataset

V. CONCLUSION

Due to web and social network, large amount of data are generated on Internet every day. This web data can be mined and useful knowledge information can be fetched through opinion mining process. This paper discussed different opinion classification and summarization approaches, and their outcomes. This study shows that machine learning approach works well for sentiment analysis of data in particular domain such as movie, product, hotel etc., while lexicon based approach is suitable for short text in micro-blogs, tweets, and comments data on web. Due to applications of opinion detection in various domains such as product, travel, movie etc, it is emerged as a popular topic in web mining.

REFERENCES

[1] Jingjing Liu, Stephanie Seneff, and Victor Zue, "Harvesting and Summarizing User-Generated Content for Advanced Speech-Based HCI", IEEE Journal of Selected Topics in Signal Processing, Vol. 6, No. 8, Dec 2012, pp.982-992.

[2] Mikalai Tsytsarau, Themis Palpanas "Survey on mining subjective data on the web", Data Mining Knowledge Discovery, Springer 2012, pp.478-514.

[3] Alvaro Ortigosa, José M. Martín, Rosa M. Carro, "Sentiment analysis in Facebook and its application to e-learning", Computers in Human Behavior Journal Elsevier 2013.

[4] Chien-Liang Liu, Wen-Hoar Hsaio, Chia-Hoang Lee, Gen-Chi Lu, and Emery Jou "Movie Rating and Review Summarization in Mobile Environment", IEEE Transactions on Systems, Man, and Cybernetics-Part C: Applications and Reviews, Vol. 42, No. 3, May 2012, pp.397-406.

[5] Aditya Joshi, Balamurali A. R., Pushpak Bhattacharyya "A Fall-back Strategy for Sentiment Analysis in Hindi a Case Study" Proceedings of ICON 2010: 8th International Conference on Natural Language Processing, Macmillan Publishers, India.

[6] Elena Lloret, Alexandra Balahur, José M. Gómez, Andrés Montoyo, Manuel Palomar, "Towards a unified framework for opinion retrieval, mining and summarization" Journal of Intelligent Information Systems Springer 2012, pp.711-747.

[7] Alexandra Balahur, Mijail Kabadjov, Josef Steinberger, Ralf Steinberger, Andrés Montoyo, "Challenges and solutions in the opinion summarization", Journal of Intelligent Information Systems Springer 2012, pp.375-398.

[8] Alexandra Trilla, Francesc Alias "Sentence-Based Sentiment Analysis for Expressive Text-to-Speech", IEEE Transactions on Audio, Speech, and Language Processing, Vol. 21, No. 2, February 2013, pp.223-233.

[9] Bo Pang, Lillian Lee, "Opinion Mining and Sentiment Analysis", Foundations and Trends in Information Retrieval Vol. 2, Nos. 1–2 (2008).

[10] Esuli, A., & Sebastiani, F. (2006). "SentiWordNet: A publicly available resource for opinion mining". In Proceedings of the 6th international conference on Language Resources and Evaluation (LREC'06), pp.417–422.

[11] Pang B, Lee L, Vaithyanathan S. "Thumbs up? Sentiment classification using machine learning techniques". Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP) 2002.

[12] Dave K, Lawrence S, Pennock D. "Mining the peanut gallery: opinion extraction and semantic classification of product reviews". Proceedings of the 12th international conference on World Wide Web, ACM, New York, NY, USA, WWW'03.

[13] Hu M, Liu B. "Mining and summarizing customer reviews". Proceedings of the 10th ACM SIGKDD international conference on knowledge discovery and data mining. ACM, New York, NY, USA, KDD 2004.

[14] Fellbaum, C. (1998). WordNet: An electronic lexical database. The MIT Press.

[15] Aditya Joshi, Balamurali A R, Pushpak Bhattacharyya, Rajat Mohanty, "C-Feel-It: A Sentiment Analyzer for Micro-blogs", Proceedings of the ACL-HLT 2011, pp.127-132.