Contrastive Opinion Summarization based on Blog Threads

Vidhita Kothari[1], Arindam Chaudhuri[2], Saifee Vohra[3]

[1]*PG Student, Computer Engineering Department, MEFGI, Gujarat, India*
[2]*Associate Professor, Computer Engineering Department, MEFGI, Gujarat, India*
[3]*Assistant Professor, Computer Engineering Department, MEFGI, Gujarat, India*

ABSTRACT

Social media has become a vital medium for people to analyze others opinions while making decision. These opinions consist of both positive as well as negative opinions about that particular product/topic. There are large amount of opinions shared using blogs, forums or other opinion dedicated websites. Thus referring all those opinions is a challenging task for user. Existing opinion summarization work is limited to separating positive and negative opinions on basis of particular aspects or features. This paper gives a solution to this problem by generating a contrastive opinionated summary which highlight both positive as well negative opinion in form of pair so user can view both types of opinions at same time. For the computation of contrastive summary basic level similarity measures are used. Our goal would be to generate a contrasting opinion summary using advanced sentence similarity measures for better results. This facilitates direct comparison among the opinions in form of a contrastive summary which helps in quick decision making.

Keyword: *Contrastive opinions, Contrastive summary, Conflicting opinions.*

1. INTRODUCTION

Opinions are present in all aspects of our daily activities and thus influence our behavior, our observation and affect how we see and assess things around us. This is the reason why we often search for the opinions of others, especially when we are in concern to make a decision. With the encroachment in web technologies, people can easily express their opinions on variety of topic using platforms such as opinion sharing websites. Since big amount of opinions are available about a subject matter, it makes difficult for users to assimilate all the opinions. A stab was made by generating a concise and digestible summary for large number of opinions which is called opinion summarization.

Opinion Summarization can be performed at different levels of granularity like at document level, sentence level or at aspect level. In case of document level mining, a document is considered as a single entity to be observed. Similarly for sentence level mining, a single sentence and for aspect level mining, different aspects of an entity are taken into consideration. Early studies on opinion mining and summarization has focused on classification of all the opinions as either positive or negative and determining the final polarity of the entire document. But still a user needs to go through all the positive and negative opinions to get a clear idea. It has been found that people are more interested in reading a topic when both sides of it are reflected i.e. positive and negative reviews on that topic in single frame. Thus the approach of contrastive opinion summarization is to produce an automatic summary of contrasting pair of sentences on same aspect to make user digest the mixed opinions. A contrastive summary reflects both positive and negative sides of that particular aspect so users have a clear image of that particular aspect. Detection of such contrastive pair of sentences is important step in formation of contrastive opinion summarization. Thus contrastive opinion summary would be really helpful for understanding both sides of the particular aspect for a common user.Traditional opinion summarization techniques resulted separation of positive and negative opinion on some specified topic. Now the question arises is what more can be done after separating positive and negative opinions for more clear understanding of user. A Contrastive summary is an approach where user can view both the side of coin together and decide according to his/her convinces.

In case as an illustration consider, some customer may say affirmative things about the phone X such as “the phone is exceptionally good for a technical person” but others might say “the phone operating system is not easy to understand, not happy to use.” So it can be seen that both the opinions are contradictory but are on same topic i.e. phone X but both the opinions are made under different conditions. When there are many such contrastive pairs of sentences about same aspect user would need to understand how to understand those types of sentences. So instead of referring thousands of opinions COS highlights the most contrastive and representative sentence pairs in form of summary. From the summary for above mentioned example one can conclude that if a user is a technical person phone is best for him otherwise not. From the sets of positive and negative opinions which is generally the result of an opinion summarizer, COS aspire to extract most representative opinions from the set of various types of opinions and compute a summary including collection of contrastive sentence pairs.

In this study, we proposed a system that would generate contrasting opinion pairs which would make the task of undergoing all the opinions for the particular topic/aspect of the user comparatively easier and less time consuming. In this study selection of most representative and contrastive pair is done using cosine similarity measure. As per our knowledge, movie reviews dataset that gives aspect as well its respective opinions is not available publicly. We created a new movie reviews dataset for different aspects with its respective opinions.

The rest of the study is organized as follows: Section 2 gives the current work done on contrastive summarization, Section 3 describes problem definition, Section 4 describes the methodology of proposed framework, Section 5 shows experimental results, and Section 6 concludes the paper with future enhancements.

2. RELATED WORKS

In this section we discuss some related research work to our contrastive opinion summarization research.

This work is about general area of Opinion summarization task. Opinion summarization task aims at classifying opinions based on sentiment prediction and generate a general summary reflecting those extracted opinions. Large amount of research has been conducted in this area[13]. In [13] shows the general opinion summarization task and challenges. In [14] the author gives the overview of opinion summarization classification and introduces contrastive opinion summarization.

In research paper [2] M. G. Ozsoy et al. provided a novel approach called Contrastive Max-Sum Opinion Summarization (CMSOS) which takes into consideration contrastiveness and representativeness together. This method produces a list of pairs of the most representative sentences relevant to given aspect/topic. In CMSOS model for sentence similarity Cosine Similarity measure is used with Term Frequency (TF) and Inverse Term Frequency (TF-IDF). Better result is obtained with the combination of Cosine and TF-IDF methods. The authors have also created a fresh Turkish dataset for the Contrastive Opinion Summarization task.

In paper [3] H. D. Kim et al. introduce a study of a dilemma known as contrastive opinion summarization. The aim of COS is to extract the most comparable sentences containing contrastive pair of opinions. The author proposed two general methods based on similarity measures i.e. content similarity measures and contrastive similarity measures. Content similarity measures the content or sentences in the same group of opinions while contrastive similarity measure the content or sentences lying in two different groups of opinions. The two algorithm proposed in this paper are Representativeness-First (R-F) algorithm and Contrastiveness-First (C-F) algorithm. C-F method gives better result in terms of precision and aspect coverage. Use of advanced semantic based similarity measures can be used for better results.

In research paper [7] R. Sipos et al. presented an approach to select pairs of snippets from reviews in a way that creates a summarizing product comparison. They have proposed a submodular objective function that aligns the snippet into pairs. Here snippets are selected from the product reviews due to their easy availability. Using a supervised learning approach they have achieved simplification across various product pairs by using user feedback on the given pairs. Unlike COS, this model compares snippets of two different products.

In paper [1] J. Guo et al. improved the limitation of traditional contrastive opinion summarization (COS) by integrating expert opinions with the ordinary opinions. The author proposed a technique called Expert Guided COS (ECOS) for controversial issues where semi supervised Probabilistic Latent Semantic Analysis (PLSA) is used to extract topic/arguments from opinions. ECOS model aims to select most contrastive argument pairs for controversial topics. This model integrates expert opinions with the ordinary opinion and generates a summary based on them.

3. PROBLEM DEFINITION

People pay more attention in referring the topics on which contrasting opinions have been expressed, in accepting these conflicting opinions, and in examine their evolution over instant and space. So there is a need of combine varied opinions in the summaries. To make possible direct comparisons between different topics/products, an approach to construct short and comparative summaries is required based on opinions. Contrastive Opinion Summarization is one of the solutions to problem. It provides a summary reflecting contrasting opinions in aspect level. The goal of the task is to identify the contrastive pairs of opinions from positive and negative reviews respectively and generate a contrastive summary for better understandability of users.

4. PROPOSED METHOD

The proposed framework has five major modules. These modules are Input (Movie Reviews), Labeling of movie reviews, Sentiment Prediction , Selection of most Representative Reviews from positive and negative reviews, Selection of most Contrastive Pair of Reviews, Summary Generation and Output (Contrastive Opinion Summary).

The selection of most representative review from each set of reviews i.e. positive and negative set and the selection of contrasting pair of reviews is done using advanced similarity measures.

The figure 1 below shows a diagrammatic view of the proposed framework along with its modules and their flow of interactions.

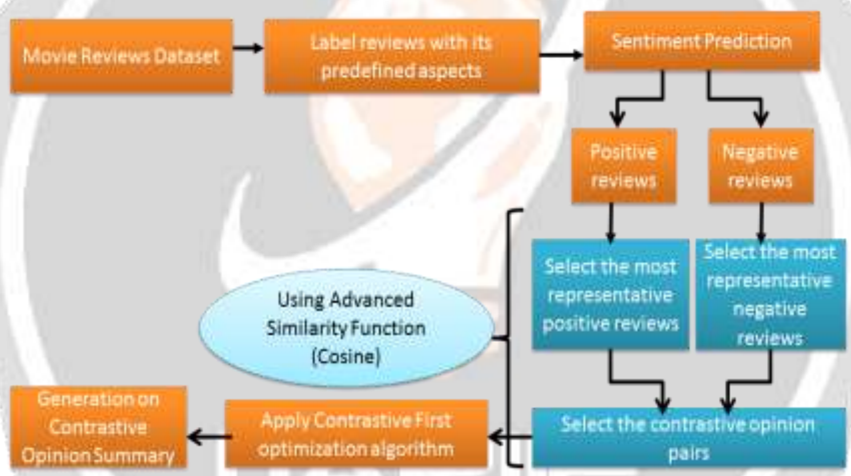

Fig 1: Working of the system

The selection of most representative review from both set positive and negative reviews involve computation of Content Similarity Function while the selection of contrasting pair of reviews involves the computation of Contrasting Similarity Function. These two similarity functions are computed using advance similarity measure namely Cosine Similarity.

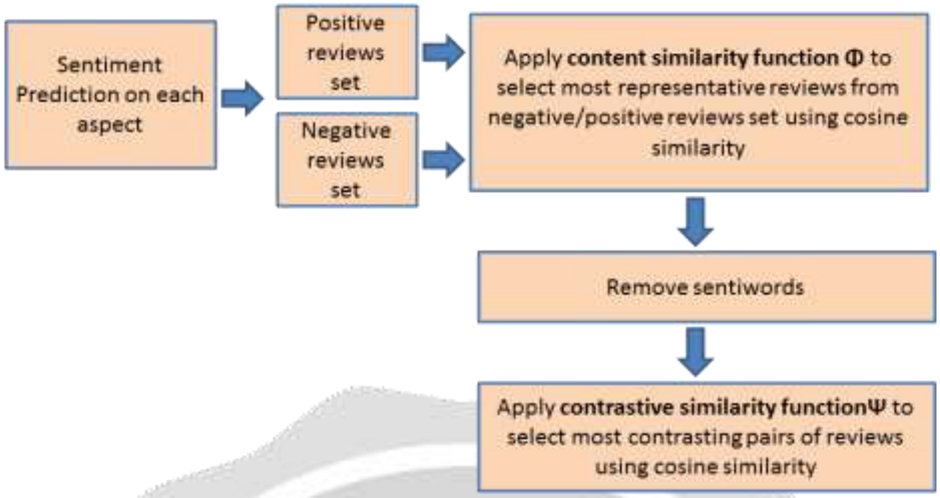

Fig 2 : Steps Performed for Computing Similarity Measures

4.1 Sentiment Prediction

This step predicts the sentiment associated with a sentence. That is, it tries to identify whether the given sentence is positive or negative with respect to a product considered. It defines the score of 1.0 for positive statement and -1.0 for the negative statement.

4.2 Selection of most Representative Opinion

This step is applied to the set positive and negative opinions individually which is output of sentiment prediction. During this step the computation of Content Similarity Function ϕ is done on positive set of opinions using the advance similarity measure i.e. Cosine Similarity. Similarly the same procedure is repeated for the set of negative opinions. The formula for computing Content Similarity Function [3] is as below:

$$\emptyset(s_1, s_2) = \frac{\sum_{u \in s_1} max_{v \in s_2} w(u,v') + \sum_{v \in s_2} max_{u' \in s_1} w(u',v)}{|s_1| + |s_2|} \quad (1)$$

where w(u,v) is term similarity function and $|s_1|$ and $|s_2|$ are the total counts of words in sentence s_1 and s_2 respectively. The strategy based on cosine similarity [18] of two vectors is used for the computation of content similarity and contrastive similarity functions. The cosine similarity between two sentences can be calculated as follows:

$$\cos(a, b) = \frac{a.b}{\|a\|\|b\|} \quad (2)$$

As the frequencies can never be negative, the cosine value will be in the range of [0,1] indicating 1 in the same direction which signifies two strings are same and 0 indicating in the different direction which signifies two strings are not at same. This cosine value is used for the further computation of similarity functions.

4.3 Selection of Contrastive Pair of Opinions

The most representative reviews from each set of positive and negative each are used for computation of Contrastive Similarity Function Ψ. First the sentimental words i.e. adjectives and negations are removed and the similarity is measures using Cosine Similarity. Rest of the computation is same as the formula for ϕ. The Contrastive Similarity function gives the corresponding contrastive pairs that are to be included in the summary.

4.4 Ranking of contrastive opinion pairs

After finding the contrasting pair of sentences now Contrastive First (CF) [3] is used to decide the order of contrasting pair of sentences in the final summary. This algorithm focuses on both representativeness and contrastiveness in the summary.

$$S^* = \arg\max \left(\lambda\, r(S) + (1-\lambda) c(S)\right) \quad (3)$$

$$S^* = \arg\max \left(\frac{\lambda}{|X|} \sum_{x \epsilon X} max_{i \epsilon [1,k]}\ \phi(x, ui) + \frac{\lambda}{|Y|} \sum_{y \epsilon Y} max_{i \epsilon [1,k]}\ \phi(y, vi) + \frac{1-\lambda}{k} \sum_{i=1}^{k} \psi(ui, vi)\right) \quad (4)$$

where ψ (ui, vi) is contrastive pair of opinions and $\phi(x, ui)$ is content similarity among i.e. positive opinions and $\phi(y, vi)$ is content similarity among y i.e. negative opinions. The contrasting pair with maximum content similarity value is considered for this task.

4.5 Summary Generation

Once the order of contrastive pair of sentences is decided then final summary can be presented in tabular for highlighting contrastive opinions on particular topic/aspect. Contrastive Opinionated Summary will be generated positive and negative reviews will be displayed for every aspect taken into consideration.

5. EXPERIMENTAL RESULTS

The following section explains the dataset used in our experiment and the results obtained.

5.1 DATASET DESCRIPTION

The proposed system uses the real time movie review dataset. These reviews contain both positive and negative opinions in form of short or long comments by the audience. Movie reviews are collected from some reputed online movie review site like rottentomatoes.com. Rotten Tomatoes is the most trusted measurement of quality for filmed entertainment. As the web's leading aggregator of movie and TV reviews from professional critics, Rotten Tomatoes offers the most comprehensive guide to what's fresh. As the proposed system requires only the audience reviews so crawler is used here. Real-time dataset is acquired using a tool available known as import.io and the labeling of reviews with particular aspect is done manually. New movie reviews dataset will be in form of aspect with its respective opinion. Reviews are split into individual sentences for the proposed system.

Table -1 : Movie Review Dataset with sentiment prediction

ID	Movies	Positive Reviews	Negative Reviews	Total Reviews
1	Kick	25	19	44
2	Bajirao Mastani	28	27	55
3	Dilwale	29	28	57
4	Tanu weds Manu	24	23	47
5	Fan	25	24	49
6	Kapoor and Sons	23	25	48

7	Prem Ratan Dhan Payo	22	24	46
8	Tamasha	20	25	45
total				391

5.2 PARAMETERS OF EVALUATION

The performance of the proposed system is evaluated using parameters such as precision, recall and F-measure.

Precision is fraction of retrieved instances that are relevant. Recall is the fraction of relevant instances that are retrieved. F-measure is the measure of test's accuracy. F-measure is defined as harmonic mean of precision and recall. Precision, recall and F-measure can be calculated as below:

$$Precision = \frac{|Relevant\ Sentence\ pairs\ \cap\ Retrieved\ Sentence\ pairs|}{|Retrieved\ Sentences\ pairs|} \quad (5)$$

$$\text{Recall} = \frac{|\text{Relevant Sentence pairs} \cap \text{Retreived Sentence pairs}}{|\text{Relevant Sentence pairs}|} \quad (6)$$

$$F - measure = \frac{2 \times Recall \times Precicion}{(Recall + Precicion)} \quad (7)$$

To calculate these measures, relevant values in the reviews are identified manually. The proposed system mines contrastive opinion pairs. Using this precision, recall and F-measure are calculated for movie reviews.

5.3 RESULTS

The proposed system gives accuracy of 78.9%. Precision, Recall and F-measure for the proposed system is shown in the figure 3 as follows

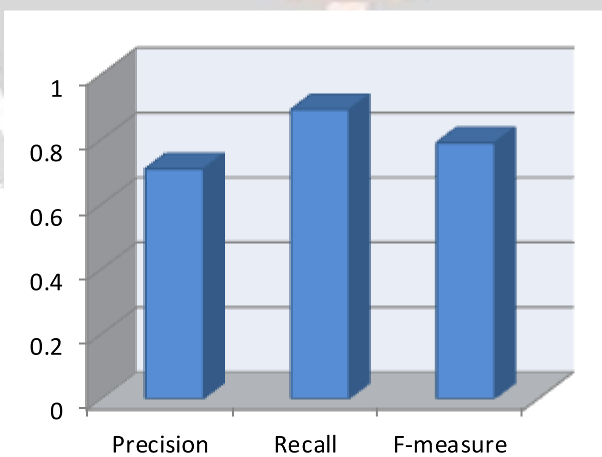

Fig 3: Result analysis of proposed system

We compare the results of our method with the results reported in [1] and [2] in Table 2.

Table 2 : Comparative Analysis of Proposed System

Method	Recall	Precision
Kim et al-RF	0.737	0.503
Kim et al-CF	0.804	0.537
CMSOS	0.889	0.649
Proposed System	**0.891**	**0.708**

The methods reported in [1] labeled as Kim et al-RF where the task of clustering is carried out before selecting the contrasting pairs so the clustering algorithm must be optimal and Kim et al-CF used basic similarity measures which did not lead to a meaningful contrastive summary. The CMSOS method uses both representativeness and contrastiveness in parallel. The proposed system performs well using cosine similarity measure as it give the highest similarity score independent of size of the text. The advance similarity measure gives better result as compared to other methods in selection of contrasting pairs and generates a meaningful summary.

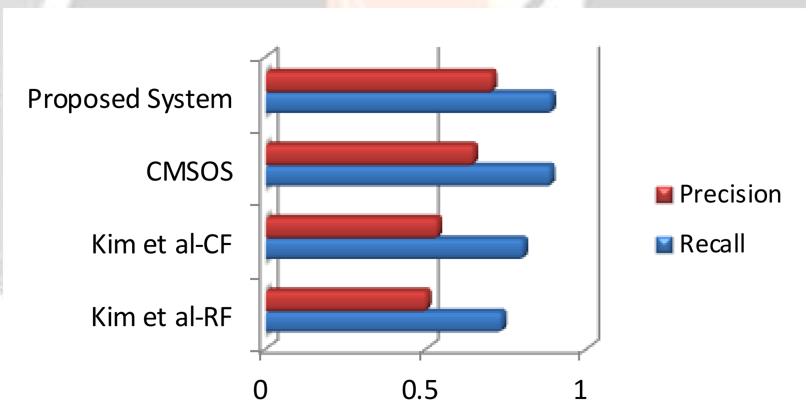

Fig 4: Comparative Analysis with Existing Solutions

6. CONCLUSION AND FUTUREWORK

The proposed system will gives an informative contrastive summary on various movies with their predefined aspects. It would be interesting to further study how to develop algorithms to achieve better approximate solutions to the optimization problem using the proposed framework. The proposed system gives better result using advance similarity measures and it can be used to generate a contrasting summary on other types of data to check the generality of the system. Larger real time dataset can be used for future enhancement of the proposed system. Advanced Semantic based similarity functions can be used.

7. REFERENCES

[1] J. Guo, Y. Lu, T. Mori, and C. Blake, “Expert-Guided Contrastive Opinion Summarization for Controversial Issues,” *Proceedings of the 24th ACM international conference on World Wide Web Companion*, pp. 1105–1110, 2015.

[2] M.G.Ozsoy and R.Cakici,"Contrastive Max-Sum Opinion Summarization," *Information Retrieval Technology*. Springer , International Publishing, vol. 1, pp. 256–267, 2014.

[3] H.D.Kim,and C.Zhai, "Generating Comparative Summaries of Contradictory Opinions in Text," *Proceedings of the 18th ACM conference on Information and knowledge management* ,pp. 385–393, 2009.

[4] M.Campr and K.Jezek, "Comparing Semantic Models for Evaluating Automatic Document Summarization," *Text, Speech, and Dialogue*, Springer,International Publishing, pp. 252–260, 2015.

[5] M. Campr and K. Jezek, "Topic Models for Comparative Summarization," *Text, Speech, and Dialogue*,Springer Berlin Heidelberg, pp. 568–574, 2013.

[6] M. J. Paul and R. Girju, "Summarizing Contrastive Viewpoints in Opinionated Text," *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, Association for Computational Linguistics, pp. 66–76, 2010.

[7] S. Ruben, and T. Joachims,"Generating comparative summaries from reviews," *Proceedings of the 22nd ACM international conference on information and knowledge management*, pp. 1853-1856, 2013.

[8] M. De Marneffe, "Finding contradictions in text," *Association for Computational Linguistics,* vol.8, pp. 1–25, 2005.

[9] S. Harabagiu, A. Hickl, and F. Lacatusu, "Negation , Contrast and Contradiction in Text Processing," *American Association for Artificial Intelligence*,vol.6, pp. 755–762.

[10] P. Achananuparp, X. Hu, and X. Shen, "The Evaluation of Sentence Similarity Measures," *Proceedings of the 10th international conference on Data Warehousing and Knowledge Discovery*, Springer Berlin Heidelberg, pp. 305-316, 2008.

[11] A. Popescu and M. Pennacchiotti, "Detecting Controversial Events from Twitter," *Proceedings of the 19th ACM international conference on Information and knowledge management*, pp.1873-1876, 2010.

[12] H.D.Kim, "General unsupervised explanatory opinion mining from text data," Ph.D. dissertation, University of Illinois at Urbana-Champaign, 2013.

[13] K. Khan, B. Baharudin, A. Khan, and A. Ullah, "Mining opinion components from unstructured reviews: A review," *J. King Saud Univ. Comput. Inf. Sci.*, vol. 26, no. 3, pp. 258–275, 2014.

[14] H. Kim and K. Ganesan, "Comprehensive review of opinion summarization," *Illinois Environ.*, pp. 1–30, 2011.

[15] S. Bhattacharjee, A. Das, U. Bhattacharya, S. K. Parui, and S. Roy, "Sentiment analysis using cosine similarity measure," *Recent Trends Inf. Syst. (ReTIS), IEEE 2nd Int. Conf.*, pp. 27–32, 2015.

[16] A. Jeyapriya and C. S. K. Selvi, "Extracting aspects and mining opinions in product reviews using supervised learning algorithm," *2nd Int. Conf. Electron. Commun. Syst. ICECS*, pp. 548–552, 2015.

[17] J. Jin, P. Ji, and R. Gu, "Engineering Applications of Arti fi cial Intelligence Identifying comparative customer requirements from product online reviews for competitor analysis," *Eng. Appl. Artif. Intell.*, vol. 49, pp. 61–73, 2016.

[18] S. Zhu, "TOP-K Cosine Similarity Interesting Pairs Search," *Fuzzy Systems and Knowledge discovery (FSKD), 7th International Conference,* vol. 3, pp. 1479–1483, 2010.

[19] M.Hu and B.Liu, "Mining and Summarizing Custmer Reivews," *In Proceedings of the 10th ACM SIGKDD International Conference on Knowledge discovery and data mining,* pp. 168-177, 2004.