Query-Based Multi-Document Evaluation

Devansh Trivedi, Maulik Soni, Sahil Misra.
Department Of Computer Engineering
K.J. Somaiya Institute of Engineering and Information Technology,
Ayurvihar Complex, Everard Nagar, Sion, Mumbai-400022
Maharashtra, India
Project Guide: -Prof. Sarita Ambadekar

Abstract- This paper provides mechanisms to present the textual information effectively. Browsing through various reviews over the web searching a specific topic is a tedious task which is time consuming. The goal of the system is, given information needed by user in the form of a query provide it's precise reply from the documents containing review. The reply can be extracted from a single or multiple documents pertaining to that topic provided by the user, tailored to specific facets (i.e., features) and/or sentiments related to the facet. The System returns a summarized response to the query provided by the user containing only relevant information along with its polarity.

Keywords -- Query Based Multi Document Sentiment Analyzer, SVM.

1. INTRODUCTION

To facilitate the task of synthesizing sentiments expressed in reviews on a particular product P specified in a user query/question Q, we use a query-based, multi-document sentiment analyzer. Given Q and a set of reviews R in the form of documents on P extracted in response to Q from documents provided by the user about the topic, system creates an extractive answer for Q, which is tailored towards the information needs expressed in Q. Answer for Q is (i) sentiment-specific if Q asks for positive (negative, respectively) information about P, (ii) facet-specific if Q queries specific facets of P (such as the battery life of a camera). QMSE summarizes multiple reviews in response to specific, user's information needs. In designing QMSE, we have learned how to (i) identify products ,facets, and sentiment keywords to determine the information needs of a user expressed in a query(ii) Form cluster for collecting sentences related to these information needs (iii) find the most-informative sentences in a review that capture the expressed opinions on the topic covered (iv)Assign these sentences into the cluster (iv) Form a sentiment summary denoting the polarity of the views mentioned in document about the facet asked in query.

2. REVIEW OF LITERATURE

There are four different query based-summarization approaches presented in the literature: (i) extraction, which generates summaries by identifying and retaining representative segments of a text, (ii)

abstraction, which creates concise sentences to capture the content of a text, (iii) fusion, which combines phrases of a text coherently to create the corresponding summary, and (iv) compression, which removes sections of a text that are considered unimportant .Multi-document, captures the overall content of a collection of documents in a single summary. In this section, we discuss existing multi-document (sentiment) summarizers, some of which are query-based and focus on mining the opinions in a document collection to be summarized, the same as QMSE. The core aim of any multi-document summarizer system is that of processing multiple sources of information and outputting relatively brief but broad report or summary. Uses of multi-document summarizer systems vary widely, from summarization of closed-domains documents. We have adopted some novel techniques to address this such as

- Sentence Ordering Model
- Heuristic Sentence Filtering
- Paragraph Clustering

Carenini et al. [1] introduced (i) MEAD, which extracts sentences that most accurately describe the features of P, and (ii) Summarizer of Evaluative Arguments, which adopts natural language methods to create a summary that includes statistical information on all the features of P, such as the number of users that mention a given feature, compiled using a set of reviews.

Conrad et al. [2] propose a query-based approach based on Fast-Sum, which is a multi-document summarizer, for summarizing sentiment documents in the legal domain. Given a user's legal question, Conrad et al. retrieve a set of top-10 documents D from Blogsearchengine.com and use FastSum to rank sentences in D.Redundant sentences are removed from the ranked list, and the remaining highest-ranked sentences yield the summary of D.

Zhuang et al. [3] rely on WordNet, statistical analysis, and movie knowledge to generate extractive, feature-based summaries of movie reviews. The authors identify sentences in a set of reviews on the same movie M which describe some of the movie features 2 The sentiment score of a summary (review, respectively) is a numerical value that reflects the rating of a product P in a summary (review, respectively), which is computed based on the positive and negative sentences in a summary (review, respectively).

3. DEVELOPMENT METHODOLOGY

3.1. Selecting Sentences to be Included in a answer

In creating a reply, QMSE includes some sentences in clusters created to determine which sentences are to be extracted from which cluster and included in the summary, QMSE relies on the ranked cluster labels. Using the ranked labels, QMSE includes in the summary of Q one sentence from a cluster at a time, starting from the cluster with the highest-ranked label up till the limit of the summary size of 250 words is reached. If the number of sentences that should be included in a summary exceeds the number of generated clusters after selecting the highest-ranked sentence in each cluster, then in the subsequent iterations QMSE selects the next highest-ranked sentence S In cluster C with the lowest similarity score, denoted LSS, with respect to the sentence(s) S in C that has (have) already been included in the summary for Q.

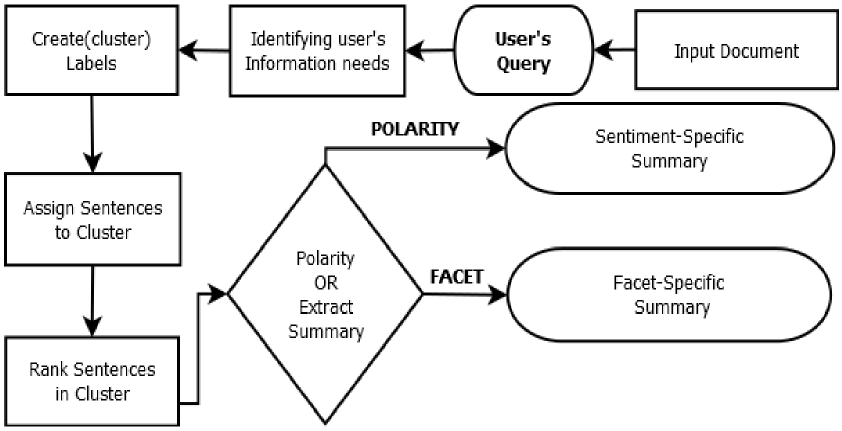

Figure 1: The Overall Summarization Process of QM

3.2. Sentiment-Specific replies

The Sentiment-Specific summary for Q is created by using only the highly-ranked sentences (along with probably the ones with the lowest LSS score) in clusters which satisfy the sentiment (i.e., positive or negative) specified in Q (identified using the keyword tagger introduced in are included in the summary for Q). To determine the positive or negative polarity of a sentence S in a cluster, QMSE calculates the overall sentiment score of S by subtracting the sum of its negative word SentiWordNet scores from the sum of its positive word SentiWordNet scores that reflects the (degree of) sentiment of S such that if its sentiment score is positive (negative,respectively), then S is labeled as positive.

3.3. Facet-Specific replies

To create the Facet-Specific replies for Q, QMSE first identifies the labels that are highly similar to each of the facets F specified in Q. To identify cluster labels highly similar to F, QMSE employs a reduced version of the word-similarity matrix which contains 13% of the most frequently-occurring words (based on their frequencies of occurrence in the Wikipedia documents), and for the remaining 87% of the less-frequently-occurring words, only exact-matched correlation factors, i.e., correlation factors of values 1.0, are used.

3.4. Proposed System

The overall process of QMSE is illustrated below. The design methodology includes (i) identifying products, facets, and sentiment keywords in a user's question using a multi-class SVM on a number of novel features, (ii) finding opinions on various facets of P using a novel sentence clustering algorithm based on word correlation factors, (iii) condensing each individual review to exclude sentences in the review that are redundant, relatively uninformative, or lack of opinions, and (iv) ensuring that each QMSE-generated multi-document summary is non-redundant, coherent, and concise by employing a simple, yet highly effective, sentence selection algorithm. Identifying the User's Information Needs QMSE adopts a one-against-all implementation of a multiclass SVM to identify information needs expressed in a query. To train a multi-class SVM, each training instance is an input vector of a non-numerical, non-stopword3 K in a query Qand is a succession of '1' ('0', respectively), each of which represents the presence (absence, respectively) of an SVM-feature F (defined by us below) if F applies (does not apply, respectively) to K.

• **Is-Singular** is set (to '1') if K is in a singular form. Products, sentiment keywords, and non-essential terms tend to be expressed in singular form.

• **Is-Adjective** is set if K is given an adjective part-of-speech(POS) tag. QMSE employs the Stanford POS tagger (nlp.stanford.edu/software/tagger.shtml), denoted Stanford-POS,for assigning POS tags, such as noun, verb, or adjective, token words (in Q). Sentiment keywords specified in Q are often adjectives which describe different aspects of a productin Q.

• **Is-Sentiment** is set if K is a sentiment keyword, which is determined by using a list of more than 4,000 sentiment keywords provided by the General Inquirer (wjh.harvard.edu/inquirer/homecat.htm).

• **Is-Capitalized** is set if the first letter of K is capitalized. The first character of a product is often capitalized.

• **Is-After-Preposition** is set if K appears immediately after a preposition, which is identified using Stanford-POS. Both products and facets tend to occur after a preposition in Q.

• **Is-After-Apostrophe** is set if K appears immediately after a term in a Saxon genitive form, i.e., a traditional term for the apostrophe-s. Facets often appear after a term in the Saxon-genitive form in Q.

• **Is-Before-Sentiment** is set if K appears immediately before a sentiment keyword inQ. Both products and facets are often followed by a sentiment keyword in Q.

• **Is-Stopword** is set if K is a stopword, which is a non essential term. We compiled our own list of 531 stopwords using multiple stopword lists posted online for this feature.

• **Is-5W1H** is set if K is one of the keywords frequently used ain formulating questions, i.e., "what", "when", "where","who", "why", and "how". 5W1H terms are treated as non essential terms, since "when", "where", "who", and "why"do not appear often in sentiment questions, whereas "how" and "what", which appear more often, do not have a direct impact on the information needs specified in users' questions.
Our claim has been verified by the high accuracy achieved by QMSE in generating summaries in response to a user's query. To verify that each of the chosen SVM-features listed above is accurate in identifying keywords (in users' queries) that are either Products, Facets, Sentiment Keywords, or Non-Essential Terms, we conducted an empirical study using a dataset, denoted for analyzing the performance of QMSE's multiclass SVM. Property-DS consists of 3,000 opinion questions randomly extracted from Yahoo! Answers (answers.yahoo.com)and WikiAnswers(wiki.answers.com). Keywords in each of the questions in Property-DS were identified as products, facets, sentiment keywords, or non-essential terms by independent assessors prior to conducting the evaluation. Table-1 shows the types of keywords that the (previously introduced) SVM-features are supposed to identify.

SVM-Features\ Keyword Types	**Product**	**Facet**	**Sentiment Keyword**	**Non-Essential Term**
Is-Singular	Yes		Yes	Yes
Is-Capitalized	Yes			
Is-Adjective			Yes	
Is-Sentiment			Yes	
After-Preposition	Yes	Yes		
After-Apostrophe		Yes		
Before Sentiment	Yes	Yes		
Is–Stopword				Yes
Is-5W1H				Yes

Table 1:Keyword types that are identified by each of the previously introduced SVM-features

Following example shows a sample user query, denoted Q. Our system divides this user's query with the help of SVM features by using Table 1.

Query Q : What do people like about twilight's story?

Words in query:
What
do
people
like
about
twilight's
story
capitalizedWords:
Tags:
What/WP
do/VBP
people/NNS
like/IN
about/IN
twilight/NN
s/PRP
story/NN
Stopwords:
do

4. CONCLUSION

We have introduced QMSE, a query-based, extractive, multi-document sentiment summarizer, which summarizes reviews retrieved from various documents, in response to a query Q posted by a web user who is interested in (positive and/or negative) feedback compiled by other users on a product P. With the development of QMSE we have introduced a new label extraction approach to identify the facets of a product and a novel clustering algorithm based on word correlation factors to group sentences in different reviews based on the facets of a product addressed in the sentences.

REFERENCES

[1] G. Carenini, R. Ng, and A. Pauls.Multi-document summarization of evaluative text. In EACL, pages 305–312, 2010.
[2]J. Conrad, J. Leidner, F. Schilder, and R. Kondadadi.Query based opinion summarization for legal blog entries. In ICAIL, pages 167–176, 2009.
[3] L. Zhang, Y. Pan, and T. Zhang. Focused named entity recognition using machine learning. In ACM SIGIR, pages 281–288, 2004.
[4] X. Hu and B. Wu. Classification and S2ummarization of pros and cons for customer reviews. In IEEE/WIC/ACM WI-IAT,pages 73–76, 2009.
[5] Sun Park; Dept. of Computer Eng.,Uni. Of Honam, Gwangju : ByungRae Cha, Query-Based Multi-Document Summarization Using Non-Negative Semantic Feature & NMF Clustering ,In IEEE,pages 609-614,2008.
[6] Chao Shen and Tao Li ,Learning to Rank for Query-Focused Multi-Document Summarization In IEEE, pages 1550-4786,2011.