

Review on Aspect level Information Retrieval System for Micro Blogging Sites

Manisha Punn, Nidhi Bhalla

Abstract - In this paper we have done systematic review of the Aspect Level Retrieval Systems and have also traced its theory and practice since year 2003 (LDA, ME-LDA, SLDA, ASUM, TSM, JST, JAS and LingPipe Model). It has also been found that today Aspect based, Perspective based systems are gaining significance, therefore, there is an urgent need to improve existing algorithm working in this area, for this we have also given typical flow of such retrieval systems.

Key Terms – Aspect, Information retrieval, LDA, Opinion mining, Perspective, Reviews, Web Objects,

1. Introduction

Information retrieval deals with acquiring relevant information from a collection of information resources. The structured information retrieval method takes advantage from structured information by handling blocks with regular patterns. Conversely, the unstructured retrieval method benefits from handling blocks with irregular patterns ignoring the structured information. The hybrid model uses both models, unstructured retrieval model as well as structured retrieval model. The data units about whom the information is retrieved are referred as web objects [6]. After retrieving the information about web objects, the information is indexed and then ranked.

Aspect [2] is a related object or member or part of a topic of interest on which the evaluation is made. Aspect level information retrieval system retrieves the information based on his/her aspect, perspective etc.

Manisha Punn is currently pursuing masters degree program in Punjab Technical University, Punjab, India.
E-mail: manisha1879@yahoo.co.in

Nidhi Bhalla is currently working as Assistant Professor in swami Vivekanand Institute of Engineering and Technology, Punjab, India.
E-mail: er.nidhi1984@yahoo.com

These web resources may be blogs, forum discussions, social networking sites etc. Aspects are nouns, verbs, adjectives and adverbs of relevance. After extracting the aspects, opinion mining is performed. Opinion mining is performed at document level, sentence level as well as aspect level.

Aspect level opinion mining was earlier known as feature level opinion mining [2]. It aims at retrieving aspects from user reviews and then evaluating the ranks of these aspects and also determining whether aspects are positive, negative or neutral. Users are now more interested in particular aspects rather than complete entity [5]. This lead to the increase in aspect level opinion mining in recent years. Aspect level or perspective opinion mining provides more clear opinions of users. Analyzing these opinions proves to be beneficial as it helps in decision making. According to [5], aspect level opinion mining is still a task with less attention.

2. A Review of Related work

In a study [7], an aspect based summary of aspect based user reviews on products sold online was developed. The proposed method dealt with adjective synonym set and antonym set. The precision and recall values were evaluated which were compared with Fast Term Recognizer FASTR [13]. It was concluded that both precision and recall

values of the proposed system were better than FASTR. For opinion sentence extraction the average precision and recall values were 0.693, 0.642 respectively and sentence orientation accuracy was 0.842..

Most of the work on customer opinion extraction collected information from customer review sites [7]. A study was conducted which considered user reviews from weblog posts [1]. The focus was on two tasks, that is, extracting aspect evaluation relation and extracting aspect of relation using methods combining contextual and statistical clues. The recall and precision values were calculated. For aspect evaluation relation, contextual clues showed 10% improvement in both precision and recall. For aspect of relation, contextual clues showed 10% improvement in precision and 20% improvement in recall. Result concluded that statistical clues needed to be calculated more precisely.

A model was proposed by [8] which identified aspect specific opinion words in an unsupervised manner. Using the database used for this model, another model MaxEnt-LDA [4] was proposed. It was represented as ME-LDA. The ME-LDA model was an extension of LDA [9] but included both opinion words and aspect words. It was concluded that this model could identify meaningful words related to other aspects and could perform good for small training set.

Discovering aspects automatically from reviews is a problem. This problem was tackled by [3]. It proposed a sentence LDA model (SLDA) which assumed that all words in a sentence came from a single aspect. SLDA was further extended to Aspect and Sentiment Unification Model (ASUM). The performance of ASUM was compared with other models like TSM [11], JST [10] and LingPipe [12]. The performance of ASUM was better than all the models.

Recently, a Join Aspect/Sentiment Model (JAS) [5] was proposed to improve aspect level opinion mining for online customer reviews. The performance was compared with many other models one of which was Support Vector Machine (SVM). It was concluded that performance of SVM was the best following which was the performance of JAS.

Many models have been used for aspect level opinion mining and most of them have performed well with the aspect words.

Models	Year
Latent Dirichlet Allocation Model (LDA)	2003
MaxEnt-LDA Model (ME-LDA)	2010
Sentence LDA Model (SLDA)	2011
Aspect and Sentiment Unification Model (ASUM)	2011
TSM	2007
Joint Sentiment/Topic Model (JST)	2009
Joint Aspect/Sentiment Model (JAS)	2013
LingPipe Model	2008

Table 1. List of Models used for Aspect Level Opinion Mining

3. A Review of Methodology

The process of information retrieval using aspect level analysis incorporates following steps:

3.1 Information Retrieval and Corpus Development

The initial step in aspect based retrieval analysis is to retrieve the aspects related to a topic from the web. The information is collected from the reviews posted by the users.

3.2 Indexing and Information Storage

After the information is extracted, the database is created to carry forward the

process of opinion mining. According to [14] opinions are expressed using adjectives, verbs and adverbs.

3.3 Classification of Opinions into Aspects and Polarities

This step deals with determining whether the user opinions are positive, negative or neutral. The intensity of their opinions is also calculated using some metric which tells whether the positive or negative opinions are mild or strong in their degree.

3.4 Evaluation of Results

After the classification phase is over, the results are obtained and the performance of the system is evaluated in terms of its validity, recall and precision.

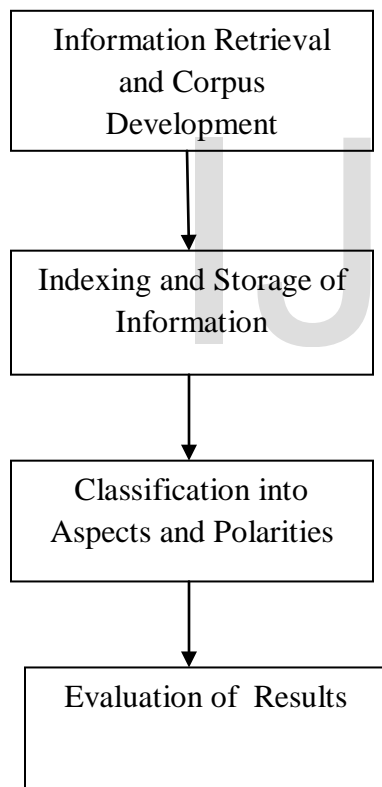


Fig 1. Process of Aspect level opinion mining

4. Conclusion

Extraction of aspects from weblogs is more challenging as compared to extraction from review articles. This is so because there is a more possibility that weblog posts divert

from the topics, vocabulary, goals etc. evaluating these posts have brought many changes in Arab countries in 2011. Election results were also predicted using the posts related to elections.

Joint Sentiment/Topic Model was the first model based on LDA which simultaneously considered topics and sentiments. MaxEnt-LDA was the first model to identify aspect and aspect specific opinion words using supervised maximum entropy (MaxEnt) component in order to separate factual words from opinion words. SVM approach used for aspect level opinion mining showed the best performance but it suffered from intensive labeling labor.

References

- [1] Nozomi Kobayashi, Kentari Inui, Yuji Matsomoto, 2007. "Extracting Aspect Evaluation and Aspect-of Relations in opinion mining."
- [2] Bing Liu, 2012. "Sentiment analysis and Opinion Mining."
- [3] Yohan Jo, Alice Oh, 2011. "Aspect and Sentiment Unification Model for Online Review Analysis."
- [4] Wayne Xin Zhao, Jing Jiang, Hongfei Yan, Xiaoming Li, 2010. "Jointly Modeling Aspects and Opinions with a MaxEnt-LDA Hybrid."
- [5] XU Xueke, CHENG Xueqi, TAN Songbo, LIU Yue, SHAN Huawei, 2013. "Aspect Level Opinion Mining of Online Customer Reviews."
- [6] Zaiqing Nie, Yunxiao Ma, Ji-Rong Wen, Wei-Ying Ma, 2005. "Object-level Web Information Retrieval."
- [7] Minqing Hu, Bing Liu, 2004. "Mining and Summarizing Customer Reviews."
- [8] Samuel Brody, Noemie Elhadad, 2010. "an Unsupervised Aspect Sentiment Model for Online Reviews."
- [9] David M. Blei, Andrew Y. Ng, Michael I. Jordan, 2003. "Latent Dirichlet Association."

[10] C. Lin, Y. He, 2009. "Joint sentiment/topic model for sentiment analysis."

[11] Q. Mei, X. Ling, M. Wondra, H. Su, C. Zhai, 2007. "Topic sentiment mixture: modeling facets and opinions in weblogs."

[12] Alias-i. Lingpipe 4.0.1. <http://alias-i.com/lingpipe>, 2008.

[13] FASTR.
<http://www.limsi.fr/Individu/jacquemi/FAS>
TR/

[14] Yi Fang, Luo Si, Naveen Somasundaram, Zhengtao Yu, 2012. "Mining Contrastive Opinions on Political Texts using Cross Perspective Topic Model."

IJSER