

OPINION MINING ALGORITHM

Marathe Dagadu Mitharam

*Lecturer, Computer Management, RCPIMRD SHIRPUR, North
Maharashtra University Jalgaon, Thalner, Shirpur, Maharashtra, India*

ABSTRACT

In this research paper creates algorithms for opinion mining. It discovers positive, negative or neutral opinion on a particular product as well as a comparative sentence of product. To capture the opinion and it classifies an evaluative text as being positive or negative. For example, given an object review or blogs, the system determines whether the review expresses a positive or a negative sentiment of the reviewer. To discover the opinion of Internet forums, discussion groups, blogs, etc. Classification and tree structure of opinion done using WEKA Software.

KEYWORDS: *Opinion Mining, Algorithm for Opining Mining, Sentiment Classification, Comparative Sentences and Relation Mining*

Article History

Received: 14 Jul 2018 | Revised: 19 Jul 2018 | Accepted: 24 Jul 2018

INTRODUCTION

Generally product is good or bad shows the opinion for peoples. In this paper opinion mining means to give opinion on content on the Web, e.g. reviews, forum discussions, and blogs etc. The problem is intellectually challenging, but practically useful. Simply, teacher teaches good or bad it is an opinion of students. Opinion mining provides valuable information for placing advertisement in web pages or not. If in a page people express positive opinions for product then the industry place the advertisement on websites. If people express negative opinions for product then cannot place. Product is best it shows some review, blogs, forum etc.

Opinion mining depends also on parts of speech. Business spends a huge amount of money to find consumer sentiments and opinions. The accuracy is usually reasonable (greater than 80%) if the sentences are either positive or negative, but if neutral sentences are included, the accuracy often drops significantly. Sentences containing negations also pose difficulties.

Some opinion examples

This cloth is best.

This cloth is bad. Etc.

Using Sentiment Words and Phrases

In sentiment words and phrases that express positive or negative sentiments (or opinions). In sentiment words that indicate positive or negative opinions are important[1], e.g. great, excellent, amazing, horrible, good, bad, worst, short, long, best etc.

- Manually find a set of seed positive and negative words. Separate seed sets are prepared for adjectives, adverbs, verbs and nouns [2].
- Grow each of the seed set by iteratively searching for their synonyms and antonyms in WordNet until convergence, i.e., until no new words can be added to the set. Antonyms of positive (or negative) words will be added to the negative (or positive) set.

Apart from a set of opinion words, there are also idioms, which can be classified as positive, negative and neutral as well. Many language patterns also indicate positive or negative sentiments. They can be manually compiled and/or discovered using pattern discovery methods.

Note that the opinion orientations of many words are domain and/or sentence context dependent. Such situations are usually hard to deal with. It can be easy in some cases. For example, “small” can be positive or negative. However, if there is a “too” before it, it normally indicates a negative sentiment, e.g., “this camera is too small for me”[2].

- **Using Supervised Learning:** we need to prepare a set of manually labeled positive, negative and neutral sentences as the training data. If sentiment words and phrases, idioms and patterns are used also as attributes, the classification results can be further improved. Opinions are hard to express with a few keywords. Sentences containing negations and clauses starting with “but”, “however”, etc., need special handling since one part of the sentence may be positive and another part may be negative,

e.g., “The pictures of this camera are great, but the camera itself is a bit too heavy.”

The comparative is used to state that one thing has more (bigger, smaller) “value” than the other. The superlative is used to say that one thing has the most (the biggest, the smallest) “value”.

Algorithm for Opinion Mining

(DM)

It is for sentiments words.

Step I

Identifying the object (ob1)

Initialized F=first feature of object

Initialized PCounter=0

Initialized NCounter=0

Setp-II

Loop

Captures/scan opinion on object (ob1)

Clusters (ob1)A_i.... A_j

//Length of opinion sentences is less than 200 characters.

Step-III

If Length <=200 characters (ob1).

Then

If the opinion is positive or negative

Check opinion as positive or negative of object

(Greater, excellent, great, excellent, amazing,

horrible, good, bad, worst, short, long, best etc

Find out Synonyms of object features

F=(f1,f2,f3...fn)

Store in dataset

End if

End if

//because different users use different words or phrases.

Scan next opinion of an object

If Opinion="Positive" (Synonyms of positive)

Then

PCounter=PCounter+1

End if

If Opinion="Negative"(Synonyms of Negative)

Then

NCounter=NCounter+1

End if

Calculate Ratio of Positive and Negative

End loop

EXPERIMENT

This algorithm implement in server-side scripting language PHP. To analyze data using WEKA software. Opinion mining it is most important in this era because when purchasing any kind of data then we review some opinions given in blog, forum etc.

To create 10 Computer as a client, every client passes the sentiments opinion about the product in that blog. Write PHP script for mining the opinion using the above algorithm just as in client-server basis but yet it's not implements in websites.

- Extract statistical information and discover interesting user objects.
- Cluster the objects group according to their navigational behavior.
- Discover potential correlations between opinion and user groups

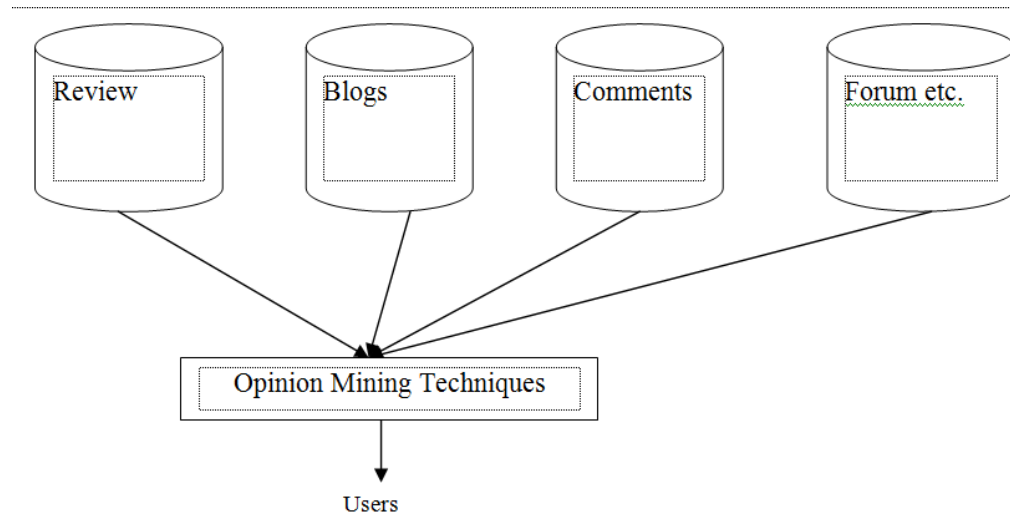


Figure 1

CLASSIFICATION

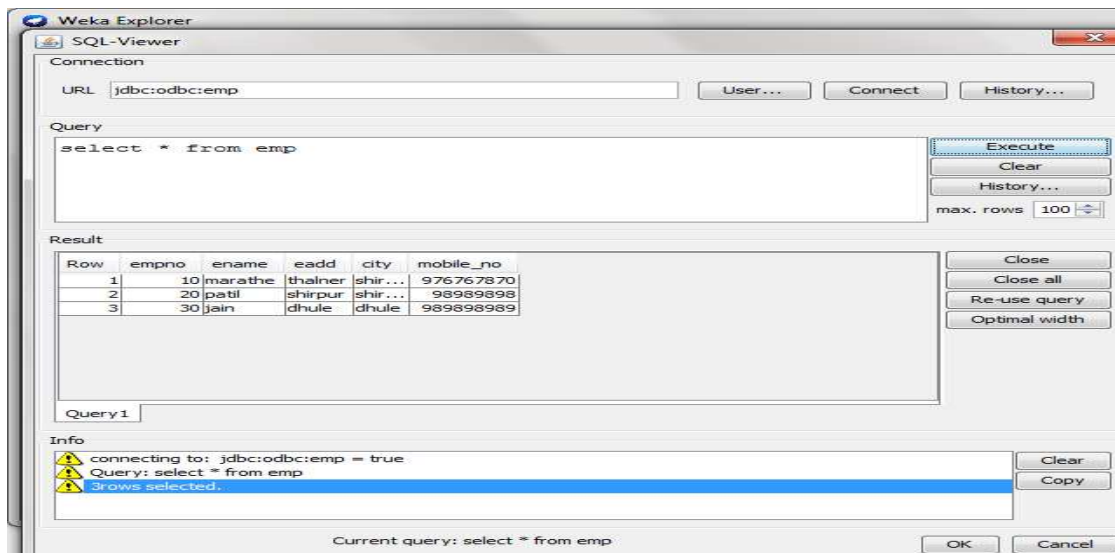


Figure 2

RESULTS

Using the proposed algorithm client getting opinion should be defined in a server on positive or negative of particular products. Directly expressing positive or negative opinions on an object or its features is only one form of evaluation.

e.g. battery life of the camera is good

Good: 46

Bad: 22

Goodness: $46 / (46+22) = 0.68$

Badness: $22 / (46+22) = 0.32$

CONCLUSIONS

Opinions can be expressed on anything, e.g. a product, an individual, an organization, an event, a topic, etc. e.g. the battery life of this camera is too short. This paper has attempted to for the purpose of web opinion mining. The proposed methods were successfully tested on some opinion. If we want to check opinion algorithm refer to this paper. The results which were obtained after the analysis were satisfactory and contained valuable information about the opinion.

REFERENCES

1. *Web data mining – Bing Liu*
2. *PPT for Web usage mining- Bing Liu*
3. *Srivastava, J., Cooley, R., Deshpande, M., Tan, P.N. (2000). Web Usage Mining: Discovery and Applications of Usage Patterns from Web Data. ACM SIGKDD, Jan 2000.*
4. *Jaideep Srivastava Paper*
5. *WCA. Web characterization terminology & definitions.*
6. *Http://www.w3.org/1999/05/WCA-terms/. Vigente al.*
7. *Weka user manual*
8. *Paper for Sentiment Analysis.*
9. *Video for Machine Learning lecture -2 Sentiment Analysis (text classification)*
10. *www.cs.cornell.edu/home/llee/omsa/omsa.pdf*
11. *MISHRA, NIDHI, and CK JHA. "OPINION MINING FROM TEXT IN MOVIE DOMAIN." International Journal of Computer Science Engineering and Information Technology Research (IJCSEITR) ISSN: 2249-6831.*
12. *Cambria, Erik; Hussain, Amir (2012). Sentic Computing: Techniques, Tools, and Applications Springer.*
13. *Vasilis Vryniotis (2013). The importance of Neutral Class in Sentiment Analysis*
14. *Moshe Koppel; Jonathan Schler (2006). "The Importance of Neutral Examples for Learning Sentiment" Computational Intelligence 22. pp. 100–109.*

