Sentiment Ranking for Opinion Extraction by Weighted Feature Scheme

Mukundaram Prem Kumar
Department of Information Technology,
Easwari Engineering College,
Chennai, India.
mukund.ram96@gmail.com

Sridhar Amirneni
Department of Information Technology,
Easwari Engineering College,
Chennai, India.
sridharamirneni@gmail.com

Sujith Prabhu
Department of Infomation Technology,
Easwari Engineering College,
Chennai, India.
sujith.prabhu78@gmail.com

Abstract— **Sentiment Analysis is the analysis of emotions and opinions in a given text using Natural Language Processing. It helps in understanding the user attitude and feelings by analyzing reviews/comments posted by people through social media, blogs etc. Internet availability has given us an easier way of accessing the information and expressing our opinions on various topics or products. These reviews are of immense importance to the targeted organization since they largely affect the outlook of the organization. People tend to consider these reviews when purchasing or using a product. This paper aims to analyze the user's sentiment in the form of free text by extracting the opinionated or sentiment words from the sentences using various techniques. They are further classified into positive or negative sentiment words and are grouped under various categories using a sentiment dictionary based on which every word will be assigned a weight. The major refinement in the proposed system is the estimation of different feature weights for evaluation of overall sentiment score and this increases the accuracy by a huge margin. The average of all the values is computed and the review is rated on a scale of 1 to 5 indicating the sentiment of the user. The sentiment score in addition to providing the user feedback also gives us an insight into the various features and helps an enterprise to focus on improving the below average features to boost the product on the whole.**

Keywords: **sentiment analysis, opinion mining, feature based, star rating; weighing scheme**

I. INTRODUCTION

Internet usage is developing at a very swift rate and is experiencing an increase in the number of user's day by day. Social media is gaining more and more popularity and is used by people express the views about any product, movie, restaurant etc. These reviews can be of huge importance to businesses as they provide an insight as to how they are performing in the society. This analysis helps in improving their business and also their customer relationship. The emotional understanding of users enables us to take decisions to improve the business.

The reviews in general can be collected through polls or by collecting the opinions in the form of text. Nowadays, most of the text reviews are unstructured in nature. Many users express their opinion through blogs or product reviews. It is practically infeasible for humans to analyze each and every review manually and finally analyze the user opinions as a whole.

Also, the same reviews when analyzed by different individuals might yield a different result. So, an automated system which can analyze the sentiment of the users will solve this problem. Since this system evaluates all the reviews on the same scale, the analysis will be more accurate and also it can be performed quickly when compared to manual analysis. This is known as sentiment analysis or opinion mining.

Sentiment analysis or opinion mining is used for natural language processing to deduce subjective information. The objective is to evaluate the attitude of a writer with respect to a document. Sentiment analysis was initially used to classify a review as positive, negative or neutral. This is done by estimating the polarity of sentiment words in a given text. E.g.: Good, bad, unhappy, brilliant etc. This method sometimes produces errors as the context under which the words were used was not considered. E.g.: Consider a mobile phone; if the review is "The processor of this mobile is slow", the opinion is of the processor and not of the phone. Hence this leads to the development of feature based sentiment analysis in which the sentence is parsed to find out the sentiment of the user and find whether it is based on the entire product or a feature. There are many other challenges pertaining to sentiment analysis and some of the problems have been addressed in the following sections.

The different sentiment analysis approaches fall into one of the following three categories:

1. Knowledge Based Techniques
2. Statistical Methods
3. Hybrid Approaches

Knowledge-based techniques work based on the identification of emotion words like happy, angry, sad or bored. Statistical methods are based on machine learning techniques. Some of the methods which use this are Latent semantic analysis, SVM, bag of words and Semantic Orientation. Some methods employ deeper analytics and try to find out the person associated with the sentiment and also the entity due to which he is affected/benefitted. In order to mine the opinion context and feature, methods make use of grammatical relationships which are obtained using deep parsing. Hybrid approaches use a combination of machine learning techniques and knowledge based techniques like semantic networks.

IJSMS

However, when the analysis of human reviews is compared with the systems for performing automated analysis, a huge discrepancy is found. Since humans often disagree among themselves i.e. each and every individuals' review varies, this error is acceptable and efforts are being taken to minimize this error.

II. RELATED WORKS

Voice of customers has been investigated by David[1] with respect to technical and functional quality of services in the marketing field. The impact of social media analytics to the economic contribution of industry and thereby to the country was suggested and demonstrated by Benxiang [2]. Cheng [3] analyzed Chinese microblogs on travel sites and provided an insight of consumers so as to enable the marketing strategy for Tourism in China. Padmini, Tannirua [4] developed a new customer satisfaction model by acquiring insights from the experience of patients so as to provide a significant improvement in the healthcare services.

Pang [5] applied machine learning techniques for extracting subjective portions of the document for sentiment classification. Ding [6] built an opinion observer, a linguistic rule based system to detect the polarity of opinions based on the context using opinion aggregation function. Harb[7] proposed an automatic extraction of positive and negative adjectives from reviews for learning the data set using the association rules and achieved an accuracy of 71% for positive adjectives and 62% for negative adjectives. Damien[8] developed a sentiment analysis model for film reviews using machine language and NLP techniques. He proposed that linguistic methods are suitable for smaller data set while machine learning techniques are suitable for a larger data set. Jine-Cheon [9] proposed a method that compares different online genres for movie documents using NLP techniques that identify the frequently used positive and negative terms.

Duk [10] proposed a multi-perspective classification to summarize the reviews of a product features. Chamlertwat [11] proposed a feature based customer feedback analysis on smart phones and summarized the best product feature among various phone brands. Somprasertsri [12] developed an ontology-based opinion model for extracting features of product from customer reviews. Keefe [13] proposed a method that used a range of feature selectors like proportional difference (PD), sentiwordnet subjectivity scores (SWNSS), sentiwordnet proportional difference (SWNPD) along with Naïve Bayes and SVM classifiers and confirmed that proportional differential (PD) is the best feature selector for the classification of larger data set.

Li [14] proposed a method for clustering documents into positive and negative groups by using TF-IDF weighing scheme. Ghag[15] proposed a sentiment TF-IDF model that finds the sentiment-stop-word by using the proportional frequency of the words and by comparing the distribution of words among positive and negative documents in a movie dataset.

The above systems did research which focused on sentiment analysis at the sentence levels and only a few works have been done on feature based sentiment analysis. The drawbacks of using NLP based techniques is that full parsing is not done due to which the accuracy of the system will be affected. The drawback of feature based analysis is that although the features were identified, their subjective importance was not considered. The weights added using TF-IDF analysis did not provide much accuracy.

III. PROPOSED MODEL

In this section, we are going to present a conceptual model of our system and explain various aspects of it in a systematic manner. The proposed system makes use of two modules namely:

1) Initial Data Modeling
2) Sentiment Estimation

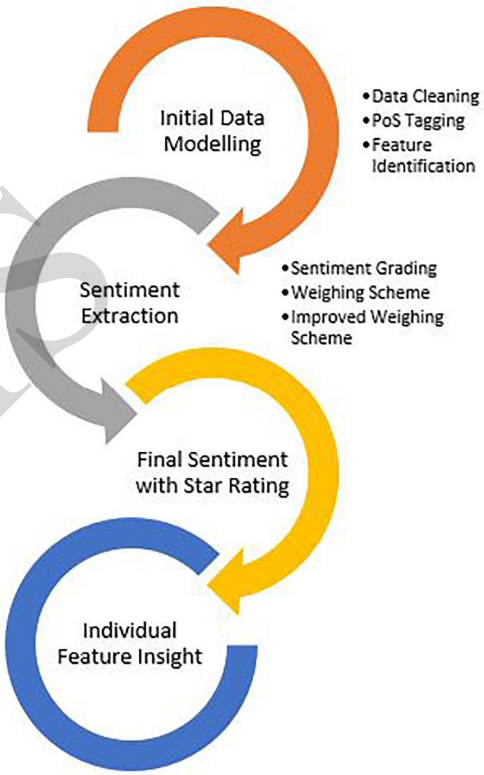


Fig 1. Overall Flow Model

A. Initial Data Modeling

Before analyzing the data for sentiments, the data needs to be checked for imperfections in order to remove irrelevant data. Once this is done, the data must be converted into a form suitable for analysis by sorting the data and grouping them into various categories. This is done by PoS tagging (Parts of Speech). Once data is categorized, the various features pertaining to the domain are identified.

1) Data Cleaning

The real-world data is incomplete or error filled lacking attribute values, lacking useful values, contains too many outliers, a disparity in values, etc. This data is not useful in

IJSMS

doing analysis or using as a corpus for sentiment analysis. The various tasks which are used to extract meaningful data include data cleaning, integration, transformation, reduction.

Data Cleaning has various steps like filling the missing values, identifying the outliers and correcting conflicting data. The missing values can be corrected using various methods like:

- Removing the record.
- Using the element mean
- Using the element mean for all the records displaying the same characteristics.
- Forecast the missing value by using a learning algorithm.

The outliers are identified and the noisy data is smoothened out by binning, clustering or regression. The conflicting data is corrected by using domain knowledge or expert decision.

Combining various databases, data cubes or files is called Data Integration. Data Transformation involves normalization, aggregation, generalization and attribute construction. Decreasing the number of attributes, decreasing the number of attributes values or reducing the number of tuples form the foundation of Data Reduction.

2) *PoS Tagging*

This is used to assign a tag to all the words present in the sentence. This is a form of supervised learning which is used to assign the parts of the sentence to the word. The various tags present are:

Table I. Parts of Speech Tagset

Code	Type	Example
CC	Coordinating conjunction	And, but, or
DT	Determiner	A, the
FW	Foreign word	Mea culpa
IN	Preposition/subordinating conjunction	Of, in, by
JJ	Adjective	yellow
JJR	Adjective, comparative	bigger
JJS	Adjective, superlative	wildest
MD	Modal	Can, should
NN	Noun, singular or mass	llama
NNS	Noun, plural	llamas
NNP	Proper noun, singular	IBM
NNPS	Proper noun, plural	Carolinas
PRP	Personal pronoun	I, you, he
RB	Adverb	Quickly
RBR	Adverb, comparative	faster
RBS	Adverb, superlative	fastest
TO	To	to
UH	Interjection	Ah, oops
VB	Verb, base form	eat
VBD	Verb, past tense	ate
VBG	Verb, gerund/present participle	eating

The main problem in this is for the same words depending on the instance or the type of usage the tag can change.

E.g.: For the word ‘top’

- X stood at the top of the stairs (Noun)
- The top button of his shirt (Adjective)
- Some civil servant earning twenty-eight thousand a year tops. (Adverb)
- Toast topped with baked beans (Verb)

The automatic method used for tagging is Maximum Entropy Markov Model (MEMM) Decoding. It is a greedy algorithm which is based on maximizing the value of P(tag | nearby words, nearby tags). This has a relatively high accuracy than the other methods used for PoS Tagging.

conjunction: though

<rule LC=positive, RC=negative, solution=positive>
<rule LC=negative, RC=positive, solution=negative>

conjunction: but

<rule LC=positive, RC=negative, solution=negative>
<rule LC=negative, RC=positive, solution=positive>

where LC is left clause, RC is right clause, w is word, nw is next word. Consider a sentence with conjunction ‘but’. “The phone is good but has a bad display”. Although the sentence may seem positive, the part after ‘but’ changes the sentiment of the sentence. So, a positive score is given for the feature ‘Phone’ and a negative score is given for the feature ‘Display’.

double negation: not <neg adjective>

<rule w=negative, nw=positive, solution=negative>
<rule w=negative, nw=negative, solution=positive>

If there are two negation terms occur in a sentence either consecutively or in different parts of the sentence, the sentiment of the sentence is evaluated as positive. Consider an example for double negation. “Sound Quality is not bad”. When a negative word ‘bad’ is prefixed with another negative word ‘not’, it becomes positive. Hence feature ‘Sound Quality’ is given positive score.

3) *Feature Identification*

Generally, reviews are not given as a whole instead they are based on any feature of the product. These features have different weights which have to be determined. These features vary from review to review. There are various methods present to identify the feature methods. The method with the highest accuracy is Entropy Weighted Genetic Algorithm (EWGA).

This method uses the information-gain (IG) factor to give weights to various features. These weights are then used to find the selected features. The algorithm has the following steps:

a) Estimate weight using IG.
b) Add IG selected features as part of the initial solution.
c) Assess and select the solutions based on the fitness function.
d) Crossover pairs the attributes at point that maximizes total IG difference between two solutions
e) Mutate the solutions based on feature IG weights.

IJSMS

B. Sentiment Estimation

Generally, verbs, adverbs, adjectives express the sentiment of a user in a sentence. A frequency matrix is developed for every feature that is identified. A sentence is considered as an

$$\text{Sentiment Score} = \frac{\sum_{i=1}^{i=n} f(i)_w}{T_n}$$

opinion only if the sentence contains both the feature word and the sentiment word. The sentiment score is calculated for the respective feature using weights.

Consider a sample review: 'The phone is very good and the sound quality is excellent. The front camera is not good.' In this review, the reviewer has specified about the phone, sound quality, and camera. The various sentiment words are 'very good', 'excellent' and 'not good'.

1) *Sentiment Grading*

A sentiment dictionary is used to categorize the sentiment words and assigns them their corresponding values. The value for any word is between -1 and +1. The negative words are assigned a value between -1 and 0 whereas a positive word obtains a value between 0 and 1. Positive words can be categorized into:

- Grade 1 - E.g.: Brilliant (Indicates the highest level of appreciation), Weight = 1
- Grade 2 - E.g.: Very good (Indicates a high level of appreciation), Weight = 0.80
- Grade 3 - E.g.: Good (Indicates appreciation), Weight = 0.60
- Grade 4 - E.g.: Not bad (Indicates acceptance), Weight = 0.20

Negative words can be categorized into:

- Grade 1 - E.g.: Worst (Indicates the highest level of disappointment), Weight = -1
- Grade 2 - E.g.: Very Bad (Indicates a high level of disappointment), Weight = -0.80
- Grade 3 - E.g.: Bad (Indicates a moderate level of disappointment), Weight = -0.60
- Grade 4 - E.g.: Not good (Indicates non-acceptance), Weight = -0.20

2) *Weighing Schemes*

The sentiment score can be calculated using the following steps:

- Identify the feature words and the associated sentiment words.
- Consider the weight of the sentiment for every feature word.
  Eg: The Songs were brilliant in the movie. The weight of sentiment word is 1. The feature is Song.
  Hence, the total sentiment score for the sentence is 1.
  Repeat the above step for all the feature word.
- Add all the values and divide the value by the total number of feature words.
  Where $f(i)_w$ is the weighted sentiment score for a particular word,
  T_n is the total number of feature words.

The resulting sentiment score lies between -1 and 1. If the value is <0, then the review is a negative review. If the value is >0, then the review is a positive review.

The star ratings are assigned based on the following rules:
1 Star – Sentiment score <= -0.5
2 Star – Sentiment score between -0.50 and -0.20
3 Star - Sentiment score between - 0.20 and 0.20
4 Star - Sentiment score between 0.20 and 0.50
5 Star - Sentiment score between >= 0.50

For the values lying in the middle, we used the following solution:
E.g.: Consider the total sentiment score to be 0.20.

Rule 1: If the number of positive sentiment words > numbers of negative sentiment words, then assign the higher star rating. In this case, the value 0.20 will be given a 3-Star rating.

Rule 2: If the number of positive sentiment words < numbers of negative sentiment words, then assign the lower star rating. In this case, the value 0.20 will be given a 2-Star rating.

Some of the reviews from Amazon.com and their evaluation using the above method is shown below:

Review 1: Awesome phone...!!! got shareable invite… anyone interested let me know....

Review 2: Very clear sounding voice. Noise cancellation works great the other way.

Display is super sharp, pixel-packed, with adjustable auto brightness for every level of ambient light! Good color reproduction, warm hue helps prevent eye fatigue.

Amazing Battery! I've been consistently getting 2 full days with light use, 1.5 with medium use, 1 full day with really heavy use, about 7 hours full-on non-stop blast.

Review 3: I'm not satisfied with this cellphone. I'm seriously disappointed with this cell and tap does not respond. I guess I'll keep using it till the time I can afford a new mobile.

Review 4: This phone has problem with its Mic. The person on the other side is not able to hear my voice at all. Even with some earphones, the sound quality is bad. The battery charger cord is not working properly.

Review 5: The phone charger is very disappointing and I had to change my charger within a week of purchase.

IJSMS

Consider Review 3. The method will work as follows:

- It consists of three sentences.
- I'm not satisfied with this cellphone.
- I'm seriously disappointed with this cell and tap does not respond.
- I guess I'll keep using it till the time I can afford a new mobile.

Each sentence is evaluated individually.

Sentence 1: The feature word 'Cellphone' is identified. The system then looks for the associated sentiment word. The associated Sentiment word is 'not satisfied'. According to negation rule, not satisfied indicates a negative sentiment and its sentiment score is identified from its corresponding category. The score for this sentence is -0.6.

Sentence 2: The feature word 'Cell' is identified and its associated sentiment word 'seriously disappointed' is identified (negative sentiment). The sentiment score is -1. Also, this sentence contains another part. It has the feature word tap and the sentiment expressed is 'does not respond' (negative). The sentiment score is -0.8. Hence the total sentiment score for this sentence is -1 + (-0.8) = -1.8.

Sentence 3: Sentence 3 consists of a feature word "Mobile" but there is no associated sentiment word. Hence, this sentence will be treated as an opinion.

The total score for this review is -0.6 + (-1.8) = -2.4

The number of feature words are 3.

Hence, the sentiment score is -2.4 / 3 = -0.80

This value of -0.80 indicates a negative review since value is less than 0 and it also indicates that it is a strongly negative review and hence gets a 1-star since the value is < -0.50.

The below method generated many differences in ratings and when analyzed, the reason was found to be the impact factor of feature words. For better understanding, consider a review, "The phone is amazing". The sentiment score according to the above method for this sentence is 1. Consider another review," the camera of this phone is awesome". The sentiment score for this sentence according to the above methods is also 1. But when the reviews were compared, the first one received a 5-Star rating and the second one received a 4-Star rating. This is because of the difference in features mentioned in the above two reviews. The first one was about the phone and the second one was only about the camera which is a feature of the phone.

Table II. Sentiment Estimation Using Weighing Scheme

R. No.	Feature word	Associated Sentiment Word	Sentiment Score	Total Score	User Rating	System Rating
1	Phone	Awesome	1	1	5	5
2	Call Quality	Very Clear	0.80	0.85	4	5
	Display	Super sharp	1			
	Colour	Good	0.60			
	Battery Life	Amazing	1			
3	Cell phone	Not Satisfied	-0.60	-0.80	1	1
	Tap	Does Not respond	-0.80			
	Cell	Seriously Disappointed	-1			
4	Mic	Problem	-0.40	-0.60	2	1
	Sound Quality	Bad	-0.60			
5	Phone Charger	Disappointing	-1	-1	3	1

3) *Improved Weighing Scheme:*

To minimize the errors and increase the accuracy, the major improvement made to the above method was assigning weights to the feature words based on its customer's value. Consider a mobile phone, both battery and color represent features of a mobile phone. But the feature 'battery' generally has more significance to a mobile than color. Hence, to estimate the weight of each feature, we devised a method which is stated below:

Step 1: Identify the total number of occurrences of every feature word for the dataset. Also, include the synonyms of the features.

Step 2: Compute the total number of occurrences of all feature words.

$$\text{Feature Weight} = \frac{\sum f(k)}{T_r}$$

Step 3: The weight for a particular feature word can be found using the following formula:

where f(k) = Total number of occurrences for a feature,
T_r = Total number of reviews.

Table III. Weights of the Features

Feature Word	Occurrences	Synonym occurrences	Total Occurrences	Feature Weight
Battery	450	253	653	0.67
Touch	502	170	672	0.69
Call	387	22	409	0.42
Camera	400	97	497	0.51
Sound Quality	282	137	419	0.43
Processor	352	252	604	0.62
Colour	198	06	204	0.21
Charger	140	35	175	0.18
Size	204	108	312	0.32
Display	513	120	633	0.65
Platform	380	114	594	0.61
Memory	400	197	497	0.51
Mic	287	122	409	0.42
Total Number of Reviews			975	

Note: If the review talks about the phone on the whole, the feature weight will always be 1.

Consider Review 3 again for which all the steps are same except for computation of sentiment score.

Sentence 1: The feature word 'Cellphone' is identified. The feature weight is 1. The score for this sentence is -0.6.

Sentence 2: The feature word 'cell' is identified and its associated sentiment word 'seriously disappointed' is identified (negative sentiment). The sentiment score is -1.

This sentence also contains another part. It has the feature word 'tap' and the sentiment expressed is 'does not respond' (negative). The feature weight for tap is 0.69 (refers to the Touch feature weight since 'tap' refers to 'touch'. The sentiment score is 0.69 x -0.8 = -0.55. Hence the total sentiment score for this sentence is -1 + (-0.55) = -1.55.

Sentence 3: Sentence 3 consists of a feature word 'Mobile' but there is no associated sentiment word. Hence, this sentence will be treated as an opinion.

The total score for this review is -0.6 + (-1.55) = -2.15. The number of feature words is 3. Hence, the sentiment score is -2.15 / 3 = -0.71. The value -0.71 indicates a negative review and hence given a 1-Star since the value is < -0.50.

Table IV. Sentiment Estimation Using Improved Weighing Scheme

R. No.	Feature word	Associated Sentiment Word	Sentiment Score	Total Score	User Rating	System Rating
1	Phone	Awesome	1	1	5	5
2	Call Quality	Very Clear	0.33	0.44	4	4
	Display	Super sharp	0.65			
	Colour	Good	0.12			
	Battery Life	Amazing	0.67			
3	Cell phone	Not Satisfied	-0.6	-0.71	1	1
	Tap	Does Not respond	-0.55			
	Cell	Disappointed	-1			
4	Mic	Problem	-0.25	-0.25	2	2
	Sound Quality	Bad	-0.25			
5	Phone Charger	Disappointing	-0.18	-0.18	3	3

Thus, when the improved weighing scheme is used, the computed results are in accordance with the user results. This is because the feature weights have been computed with respect to user's importance. When the user's priority for a particular feature change, the same will be reflected in its feature weight. Hence, this method is highly accurate to estimate the sentiment and its star rating. Moreover, since the calculation involves only basic operations, the computation time is very low when compared to the existing methods.

IV. INDIVIDUAL FEATURE INSIGHT

Any organization needs to know the areas of improvement to increase the success rate of its business. Hence, the sentiment analysis alone is not the solution. The proposed system in addition to performing sentiment analysis can be used to extract feature based opinions and analyze various features that are responsible for success/failure of a product. Using the data from Table III, we get a simple feature analysis as follows:

- For each feature, note the corresponding number of positive and negative reviews.
- Using visualization tools, construct any chart to find out the impact of every individual feature.

IJSMS

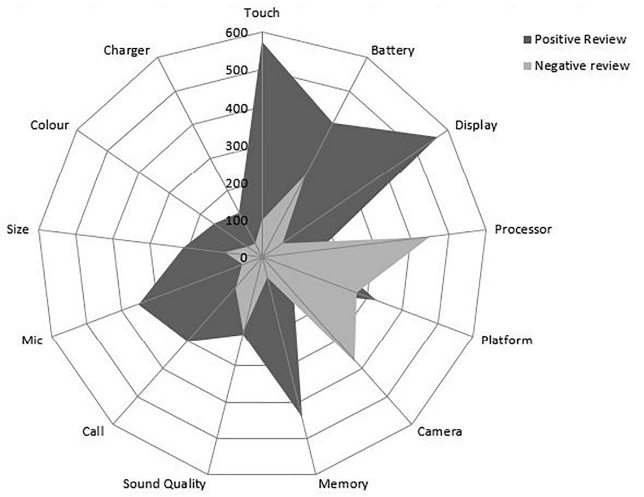


Fig 2. Individual Feature Analysis

From the radar map, it can be easily understood that users were not satisfied with the processor, camera, and sound quality. Even though other features did receive negative reviews, a majority of users found those features to be good and hence only the highly negative features are significant contributors to the failure of this product.

V. CONCLUSION

The proposed system provides a framework for sentiment analysis by comparing features of a particular subject based on the online reviews generated by the users and thereby predict the sentiment of the user. The system can produce a two-fold output i.e. Apart from indicating the sentiment score and star rating for a review it also provides an insight for businesses to identify the areas of improvement by highlighting the individual features that affect the overall product. It helps in reducing the manual effort spent in analyzing these reviews and it is very effective since it uses the customer's' views directly. The system is highly accurate as every feature is given its priority based on the customer's view and helps to identify which feature induces a positive sentiment on the customer and those that don't. The only challenge lies in identifying sarcasm statements which are very complex to compute but even this can be incorporated by using a sentiment dictionary which is context aware.

REFERENCES

[1] David, Dagger, "Generating positive word-of-mouth in the service experience", Managing Service Quality, 2011, 133-151.
[2] Benxiang, Gerritsen, "What do we know about social media in tourism? A review", Tourism Management Perspectives,2014, 27-36.
[3] Cheng, Edwards, "Social Media in Tourism - a visual analytic approach", Current Issues in Tourism, 2015, 1080-1087.
[4] Padmini, Tannirua, "Seeking intelligence from Patient experience using text mining: analysis of emergency data", Information Systems Management, 2015, 220-228.
[5] Pang, Lee, "Thumps up? Sentiment Classification using Machine Learning techniques", Proceedings of Empirical Methods in Natural Language Processing 2002, 79-86.
[6] Ding, Liu, "The Utility of Linguistic Rules in Opinion Mining", Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development on Informations Retrieval, SIGIR'07, Amsterdam, 2007
[7] Harb, Plantie, Dray, Roche, Trousset, Poncelet, "Web Opinion Mining: How to extract opinions from blogs?", CSTST '08 International Conference on Soft Computing as Transdisciplinary Science and Technology, 2008, 211-217.
[8] Damien, Bothorel, Guimier De Neef, Boull, "Automating Opinion analysis in Film reviews: the case of statistic versus Linguistic approach", Language Resources and Evaluation Conference, Morocco, 2008, 94-101.
[9] Jine-Cheon, Thet, Khoo, "Comparing sentiment expression in movie reviews from four online genres", Online Information Review, 2010, 317-338.
[10] Duk, Kavita, Parikshit, Xiang, "Comprehensive Review of Opinion Summarization", International Journal of Computer Engineering and Applications, 2014
[11] Chamlertwat, Bhattarakosol, Rungkasiri, Haruechaiyasak, "Discovering Consumer Insight from Twitter via Sentiment Analysis", Journal of Universal Computer Science, 2012, 973-992.
[12] Somprasertsri, Lalitrojwong, "Mining Feature-Opinion in Online Customer Reviews for Opinion Summarization", Journal of Universal Computer Science, 2010, 938-955.
[13] Keefe, Koprinska, "Feature Selection and Weighting Methods in Sentiment Analysis", Proceedings of the 14th Australasian Document Computing Symposium, Sydney, Australia, 2009
[14] Li, Liu, "Application of clustering method on sentiment analysis", Journal of Information Science, 2010, 127-139.
[15] Ghag, Shah, "SentiTFIDF – Sentiment Classification using Relative Term Frequency Inverse Document Frequency", International Journal of Advanced Computer Science and Applications, 2014