Trends in

Applied Sciences Research

ISSN 1819-3579

Academic Journals Inc.

www.academicjournals.com

Trends in Applied Sciences Research 6 (10): 1141-1157, 2011
ISSN 1819-3579 / DOI: 10.3923/tasr.2011.1141.1157

Sentiment Classification Using Sentence-level Lexical Based Semantic Orientation of Online Reviews

Aurangzeb Khan, Baharum Baharudin and Khairullah Khan
Department of Computer and Information Sciences, Universiti Teknologi PETRONAS Perak, Malaysia

Corresponding Author: Aurangzeb Khan, Department of Computer and Information Sciences, Universiti Teknologi PETRONAS Perak, Malaysia

ABSTRACT

Sentiment analysis is the process of extracting knowledge from the peoples' opinions, appraisals and emotions toward entities, events and their attributes. These opinions greatly impact on customers to make their choices regarding online shopping, choosing events, products and entities. With the rapid growth of online resources, discussion groups, forums and blogs; people communicate through these means of internet on daily basis. As a result, the vast amount of new data in the form of customer reviews and opinions are being generated progressively. So it is desired to develop an efficient and effective sentiment analysis system for online customer reviews and comments. In this study, the rule based domain independent sentiment analysis method is proposed. The proposed method classifies subjective and objective sentences from reviews and blog comments. The semantic score of subjective sentences is extracted from SentiWordNet to calculate their polarity as positive, negative or neutral based on the contextual sentence structure. The results show the effectiveness of the proposed method and it outperforms the word level and machine learning methods. The proposed method achieves an accuracy of 97.8% at the feedback level and 86.6% at the sentence level.

Key words: Sentiment classification, feature extraction, review mining, text mining, sentiment analysis

INTRODUCTION

With the increasing availability of electronic documents and the rapid growth of the World Wide Web, the task of automatic classification of text documents and online customer reviews becomes the key method for information organization and knowledge discovery. Proper classification of e-documents, online news, blogs, e-mails and digital libraries need text mining, machine learning and natural language processing techniques to get meaningful knowledge (Khan *et al.*, 2010). Rapidly growing online resources, online discussion groups and forums and blogs has lead to people commentating via the internet and a vast amount of new data in the form of customer reviews, comments and opinions about a product, events and entities being generated more and more. The reviews about any entity, e.g., banks, hotels, airlines and online shopping items including books, digital cameras, mobile phones, notebooks, etc., are useful in decision making for both the customer and manufacturer. When a customer wants to travel abroad by air, how is the decision made about which airline is feasible to travel on and which hotel or restaurant is more suitable to stay in, and which bank is more appropriate for banking as compared to other bank facilities? For online shopping, which brand is he/she going to buy and why? The sentiments from online reviews have a great influence on others in decision making (Liu, 2010a; Amato *et al.*, 2011).

Before the web, it was relatively very difficult to collect people reviews, and there was almost no computational study on opinions because little opinionated text was available. At that time, for one's decision making, opinions were normally collected from friends, families and neighbors. On the other hand, for an organization, surveys were conducted to aid decision making about their events and products from relevant groups of people. Now, with the rapid growth of social media content on the internet in the last few years, the world has been altered and the web is the best way for people to express their views regarding anything on the various social network sites, discussion forums and blogs. If we want to buy a product, travel abroad or stay at a hotel, we are no longer limited to asking our friends and families because there are many user reviews available on the Web. For a company, there may no longer be a need to conduct surveys from focus groups in order to gather consumer opinions about its products and those of its competitors because there is plenty of such information publicly available on the internet (Liu, 2010b).

So it is desirable to develop an efficient and effective sentiment analysis technique that is able to analyse the customer review and classify it into positive, negative or neutral opinions about any entity. Several researchers have been working on the sentiment analysis using a domain dependent framework for feature and feedback level opinion classification. A few are using machine learning techniques for classification at the document level. In this study, a domain independent rule based method is proposed for semantically classifying sentiment from online customer reviews and comments. The method is effective as it takes a review, checks individual sentences and decides its semantic orientation considering its structure and the contextual dependency of each word.

BACKGROUND AND RELATED RESEARCH

Researchers have taken a keen interest in sentiment analysis for the last few years. It has attracted a great deal of attention because of its challenging research problems and the wide range of applications for both academia and industry (Liu, 2010a). It needs a computational study for extracting knowledge from the people's opinions, appraisals and emotions toward entities, events and their attributes. In today's international global world market and highly growing internet usage, people prefer online shopping, banking, ticket reservation, hotel booking, etc. So sentiment analysis from online customer reviews is becoming a requirement of an organization, customer and also manufacturer. Different researchers have been working on different aspects of this area. The existing work on sentiment analysis can be categorized into document, sentence and word/feature level classification (Liu, 2010b).

Word or feature level sentiment analysis gets much importance by applying the NLP and statistical methods. Several researchers have worked on extraction of features and opinion-oriented words (Popescu and Etzioni, 2004; Hu and Liu, 2004a). The same technique was presented by new mechanism in Popescu and Etzioni (2004) for product features for extraction of customer opinions. Andreevskaia and Bergler (2006) presented Natural Language Processor Linguistic Parser to parse each review, to split text into sentences and to produce part of speech tags for each word like noun, verb, adjective, etc. A few authors have taken term senses into account and assume that a single term can be used in different senses and can present different opinions. They use Synset from WordNet for different senses of the same term. Attardi and Simi (2006) used opinionated words for opinion mining from blogs (Shi *et al.*, 2010).

The machine learning techniques performed better then lexicon and rule based approaches (Baccianella *et al.* 2009). They use bag-of-words (BoW), Part-Of-Speech (POS) information and sentence position as features for analyzing reviews and representing reviews as feature vectors to a learning device usually Naïve Bayes and SVM. But these feature extraction methods are also dependent on tools like POS Tagger and no contextual information is considered. Zhao *et al.* (2008), proposed a method for sentiment classification based on conditional random fields (CRFs) in response to the two special characteristics of contextual dependency and label redundancy in sentence sentiment classification. CRFs capture the contextual constraints on the sentence sentiment. A Hierarchical framework is used for introducing redundant labels and capturing the label redundancy among sentiment classes. The Hierarchical structure is very costly and ineffective in a large scale data set.

Most of the existing work focused on document level sentiment classification (Pang and Lee, 2004; Pang *et al.*, 2002), used a machine learning technique with a minimum cuts algorithm for sentiment classification. Topic oriented classification models normally represent a document as a set of terms in which topic sensitive words are important. In contrast, polar terms such as "excellent" and "worst" are considered essential to sentiment-oriented classification. The sentiment structures in sentence context are more expressive than individual polar term based features (Hu and Li, 2011). 'The full story of how lexical items reflect attitudes is more complex than simply counting the valences of terms' (Polanyi and Zaenen, 2004).

In study of Sarvabhotla *et al.* (2010), the problem of attributing a numerical score (one to five stars) to a review is presented. They use the feature representations of reviews and describe it as a multi-label classification (supervised learning) problem, and present two approaches using Naïve Bayes (NB) and Support Vector Machines (SVM's) (Subramanian and Ramaraj, 2007). Yu and Hatzivassiloglou (2003), presented a system which classifies documents and then checks subjectivity of sentences in it. The machine learning approach with the integration of compositional semantics of sentiment classification is presented by Choi and Cardie (2008). The Support Vector Machine (SVM) algorithm with 'bag of words' (BoW) to classify movie reviews is presented by Whitelaw *et al.* (2005), in which a few types of special features are selected. However, the limitation of this approach is that, it only focuses on adjectives and their modifiers that express appraisal. The method in Turney (2002) extracts the polarity of phrases using the Point-wise Mutual Information (PMI) between the phrases and seed words. Most of the above mentioned techniques use flat feature vector (a bag-of-words) BoW methods used to represent the documents. However, statistical based techniques rely on subject, domain and language style to gather large amounts of significant data with statistics while neglecting contextual information and syntactical structure which in turn affects the accuracy of the sentiment classification at small textual composition levels. So the techniques may not accurately represent the information that can be extracted at sentence level. To measure sentiment on the phrase or sentence level, opinion oriented words were proposed by simple methods for combination of individual sentiments (Kim and Hovy, 2003) and supervised (Alm *et al.*, 1990) statistical techniques. A machine learning method is proposed to using both lexical and syntactic features for sentiment analysis. These methods, however, missed vital contextual information. So, the individual sentence is important for extracting semantic orientation.

Rule based techniques approaching the analysis of word dependency and structure of contextual information for sentiment orientations were proposed by Moilanen and Pulman (2007) and Dey and Haque (2009). Opinion extraction from noisy text data was proposed at multiple levels of granularity using domain knowledge for contextual structure and WordNet for semantic orientation (Moilanen and Pulman, 2007; Umer and Khiyal, 2007).

The limitations of these techniques are manually developed domain-dependent lexicons and inability to deal with long complex sentences. A lexical system for sentiment analysis at various grammatical levels is presented by Dey and Haque (2009). This approach used a wide-coverage lexicon, accurate parsing and sentiment sense disambiguation semantic orientation So, the contextual information of all the parts of speech is essential for the semantic orientation, as was shown by Neviarouskaya *et al.* (2009). All the content, parts of speech and the structure of the sense in the sentence play a vital role in sentiment analysis The main limitations of the existing approaches are the concentration on sentence structure and the contextual valance shifter is low; lexicon based systems suffer from limitations in lexical coverage, Word since disambiguation which is ignored, rule of term weighting and the polarity score is too generalized; moreover, less attention is given to attenuation or the imperial expression or the confidence level of the sentiment orientation in the expression is ignored, and there is no proper rule for handling the noisy text with photonic symbols and special characters.

In this study we proposed a method of sentiment classification at the sentence level applying rules for all parts of speech to score their semantic strength, contextual valence shifter, expression or sentence structure based on dynamic pattern matching and word sense disambiguation is addressed. The system identified opinion type, strength, confidence level and reasons. It deals with the SentiWordNet (Esuli and Sebastiani, 2006), WordNet (http://wordnet.Princeton.edu), as the knowledge base with the additional capability of strengthening the knowledge base with modifiers and contextual valence shifter information and is used for all parts of speech.

PROPOSED FRAMEWORK

In this section, the proposed sentence level sentiment classification method is described in detail. In the first step, sentences are split into subjective and objective ones based on lexical dictionary. Subjective sentences are further processed for extraction to classify as positive, negative or neutral opinions. A rule based lexicon method is used for the classification of subjective and objective sentences. From subjective sentences, the opinion expressions are extracted and their semantic scores are checked using the SentiWordNet directory. The final weight of each individual sentence is calculated after considering the whole sentence structure, contextual information and word sense disambiguation. Figure 1 shows the overall process of the sentiment analysis of the proposed method. The steps are described below.

- Split reviews into sentences and make a Bag of Sentences (BoS)
- Remove noise form sentences using spelling correction, convert special characters and symbols (photonics) to their text expression, use POS for tagging each word of the sentence and store the position of each word in the sentence
- Make a comprehensive dictionary (feature vector) of the important feature with its position in the sentence
- Classify the sentences into objective and subjective sentences using both machine learning and lexical approaches
- Using a lexical dictionary as a knowledge base, check the polarity of the subjective sentence as positive, negative or neutral
- Check and update polarity using the sentence structure and contextual feature of each term in the sentence

Figure 1 shows the overall process of the sentiment analysis of the proposed system.

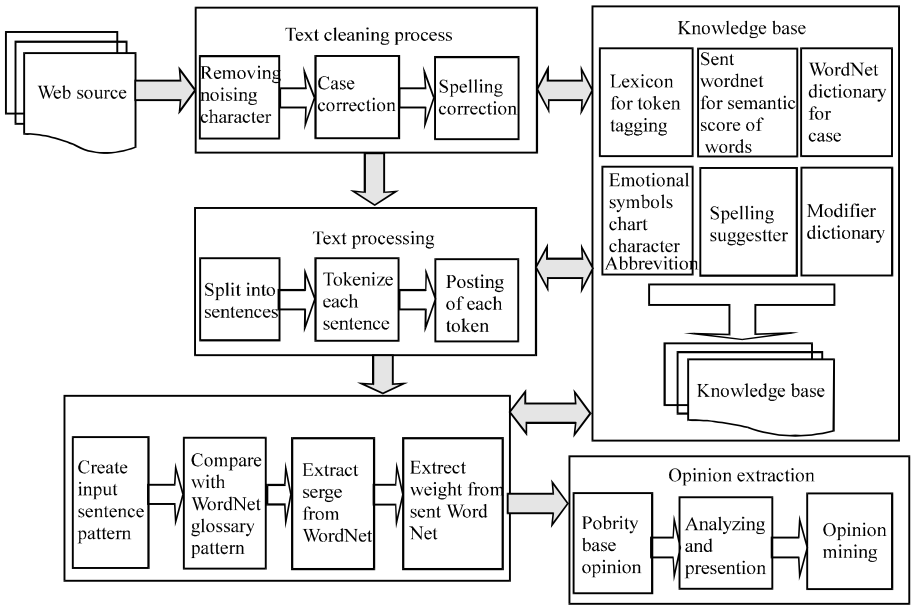

Fig. 1: Proposed Architecture for sentiment analysis

Table 1: POS Tags

Pos-id	POS_name	POS_abbrivation	SentiWordNet_abrv
1	Noun	NN	n
2	Adjective	JJ	a
3	Verb	VB	v
4	Adverb	RB	r
5	Nouns	NNS	n
6	Adjectives	JJS	a
7	Verbs	VBZ	v

Sentence splitter and processing noisy text: Here, reviews/comments are split into sentences to extract the feature level sentiment score from the SentiWordNet. A BoS is made from the split sentences an each sentence is stored with a review-id and sentence-id. After applying the POS, the position of each word in the sentence is also stored for further processing.

Sentence boundary identification: The sentence boundary identification is important in order to split reviews/comments into correct sentences. For this purpose, we have implemented a rule based module, where "." is consider as the sentence boundary, when it is not preceded by predefined word, i.e. Pvt., Ltd., etc.; the "." is also ignored after an abbreviation list (defined in dictionary) and immediately after digits which do not follow a space character.

Part of speech (POS) tagger: For assigning a tag to each word in a sentence, we used the POS tagger by adopting the Stanford trigger lexical database as the knowledge base and connect it with our system with some changes for efficient and effective tagging. The system extracts the review and comments from the web using a crawler and then cleans it and applies the POS for tagging. A tag is assigned to each word, like, JJ, JJS, VB, VBS, RB, NN, NNS, DT, etc., as described in Table 1.

Subjectivity lexicon: To construct the metadata of subjective words, the Point Wise Mutual Information (PMI) method is used. The PMI based subjectivity lexicon is formed to check subjective words in sentences (Baharudin *et al.*, 2010). The proposed rule based module is used to extract those sentences which contain opinions or subjective words referring to the SentiWordNet, WordNet or subjectivity lexicon knowledge base:

$$\begin{aligned} PS(w) &= PMI(w,pos) - PMI(w,neg) \\ &= \log_2 \frac{P(w,pos)/P(pos)}{P(w,neg)/P(neg)} \\ &= \log_2 \frac{P(w,pos)}{P(w.neg)} \end{aligned} \tag{1}$$

where, POS and neg are sets of positive and negative sentences and w is a given word.

The subjectivity score S (w) is described according to the said method.

$$S\ (w) = \log_2 \frac{P\ (w, subjective)}{P\ (w,objective)} \tag{2}$$

The word w is considered as a subjective word, if S (w) is larger than a given threshold.

Feature and opinion word position extraction: The algorithm for sentiment classification uses opinion terms or expressions to determine polarity of sentences based on contextual information and sentence structure. The position of each word in a sentence is important for the semantic orientation and correct pattern extraction for word sense disambiguation. Also, product features and opinion words are extracted from tagged sentences using the word position. Features are selected at run time after suggesting the most frequent feature list extracted from the opinionated sentences. To extract opinion words from sentences, first we focus on finding features that emerge explicitly as nouns or noun phrases in reviews. The following steps are used.

- Use the Part of speech (POS), to tag every word of the sentence and store each word position with its assigned tag
- Collect the nouns, noun phrases and adjectives with their positions
- Noun phrases are observed as product features
- For each sentence in the review, if it contains any feature word, extract any nearby adjective and consider such adjectives as opinion words
- Adjectives and/or adjective preceded by adverbs are observed as opinion words
- Frequent product features are selected from key noun phrases

Sentiment sentence extraction: Here, we apply the subjective sentence extraction method to classify sentences into objective and subjective ones. In previous study (Baharudin *et al.*, 2010), the supervised learning approach is used to extract the subjective sentences. In this study, a rule based module is proposed to extract those sentences which contain opinions, subjective expressions

Table 2: Semantic weight assigned to sentences

SEN_ID	Sentence	Weight
1	KUL-BKK A320 pretty modern cabin crew okay need to polish on their smiles and social skills	-0.30
2	Nice flight cheap price	1.75
3	For the price I paid no complaints	0.625
7	The one on the way in was really dirty	-0.50
8	On the way out of Bali the plane seemed brand new it was clean too	0.875
9	AirAsia service was bad on the way in but great on the way out	-0.525
10	The flight attendants seemed to ignore us on the way in but were kinder on the way out	0.10
11	One thing I've noticed though is the lack of safety cards along with the magazine and Buy-on-board list in every seat are we supposed to share safety cards	0.25
14	AirAsia offers good value for money considering the ticket prices but is definitely not my carrier of choice even for short flights	0.225
15	But their cheap tickets allowed us to stay at a better hotel than we would have if wed flown a full-fare airline KUL-TWU SDK-KKI and KKI-SIN	0.475
16	Overall a good experience	0.325
17	Ouly downside was not receiving the meals we had prepaid for 3 months in advance when booking the tickets	0.125
18	This is a major inconvenience for vegetarians who have nearly no other choice to get a meal on-board becanse the meal selection in general is very poor on Air Asia they are nsually out of stock on most items you ask about	0.625
20	There was no way to reassign your seat using ouline check-in two days before the it flight not even if you are willing to pay for	1.25
23	Check in was fine and boarding not a problem either Seats were more than adequate and the cabin staff were as helpful as they needed to be	-0.625
25	To be honest for a low cost airline this was actually a fantastic flight	0.625

or terms referred to in SentiWordNet, WordNet or the subjectivity lexicon knowledge base. Consideration of each term weight in sentence is important for sentiment classification (Kasam and Hyuk-Chul, 2006).

Word sense disambiguation: One unique aspect of this work is to check the word sense disambiguation. The proposed method extracts the semantic pattern of the desired sentence using the opinion expression position in the sentence. Then, all possible patterns for that opinion expression for all possible senses are extracted based on the WordNet glossaries; the system locates an exact pattern match of the desired sentence and extracts the sense no. from the WordNet synset. The semantic score for that sense no. is extracted from SentiWordNet, which gives very efficient results. If patterns are not exactly matched, then it checks for the nearest pattern and the score of that nearest pattern is extracted from SentiWordNet. The result of proposed process is described in Table 2, 3, 4 and 5. In Table 2 reviews are spitted into sentences and only subjective sentences are selected for semantic orientation.

From Table 2 we take sentence No. 25 with its semantic weight. Table 4 shows the semantic score of each term in the sentence. The matching algorithm is applied on this sentence to extract the sense of the semantic term fantastic from WordNet. Our proposed system extracts the pattern for the sentiment term and matches it with WordNet synset terms; there are four senses of the word fantastic which have both negative and positive scores, but here the sense with the positive score is to be extracted. So, the system exactly extracts the positive score for the term fantastic as 0.375 from SentiWordNet, as shown in Table 4 and 5. The process is described in Fig. 3.

Table 3: Description of terms weight

	To be honest for a low cost airline this was actually a fantastic flight			
Word	POS_ID	POS-score	NEG-score	Position
To	1		1	
Be	3	0.25	0.125	2
Honest	2	0.75	0	3
Low	2	0	0.25	6
Cost	1			7
Airline	1			8
Was	3			10
Fantastic	2	0.375	0.375	13
Flight	1	0.25	0	14

Table 4: WordNet sense patterns

		Pattern
Word	Sense No.	Pattern
Fantastic	2	/IN-/NNS-/DT-WRD-/NN-/NN-/DT
Fantastic	5	/NN-/VBP-/VBD-WRD-/NN-/JJ-/NNS
Fantastic	4	/IN-/JJ-RB-WRD-/NNS-/IN-/NN
Fantastic	3	/JJ-/JJ-/DT-WRD-/NN-/IN-/PRP$

Table 5: SentiWordNet semantic score for term fantastic

Gloss	Synset-terms	Neg-Score	Pos-score	ID	ID1	POS
Extravagantly fanciful in design, construction, appearance;"Gaudi's fantastic architecture"	Fantastic#5	0.375	0.375	1796452	a	9869
Existing in fancy ouly; "fantastic figures with bulbous heads the circumference of a bushel"- Nathaniel Hawthorne	fantastical#1 fantastic#4	0.625	0	1936778	a	10611

The tag sentence is ([To/NN be/VB honest/JJ for/IN a/DT low/JJ cost/NN airline/NN this/DT was/VBD actually/RB a/DT fantastic/JJ flight/NN ./.) and the pattern extracted is VBD//WRD-//NN, which matches the sense no.5 in WordNet, and the semantic score of sense no.5 for the term "fantastic" is 0.375. As shown in Table 4. So, it extracts the positive score of the term "fantastic" which is an accurate semantic score according to the sentence structure.

There are still problems in semantic scores as is seen in Table 4, where the word LOW has a negative score but here it has a positive sense Low-cost, so we tackle the problem by extracting a bigram word in our next step.

Knowledge base for sentence structure and contextual information: Knowledge base conations are SentiWordNet, WordNet and predefined intensifier dictionaries for domain independent polarity classification for positive, negative and neutral opinions. Sentiment words are usually classified into positive and negative categories. For this purpose, we extract the semantic score of each opinion word using the SentiWordNet dictionary containing the semantic score of more than 117662 words. Then, we check the structure and associated words (which affect the weight of the opinion word) in the sentence and update the polarity accordingly. The main aspect of this work is a knowledge base for the contextual information of each word in a sentence which really modifies the strength of the opinion. The knowledge base (calculates semantic strength for each sentence) contains negation words, enhancers, reducers, model nouns, context shifters and other intensifiers with their semantic scores.

Negation: Negation words reverse the polarity of opinion words by checking their position in a sentence. The words are (Not, Never, N't, Doesn't, Can't, Nor, Don't, Wouldn't, No, etc). The result will be the opposite if the system fails to recognize the negation word. So, for the recognition of the semantic expression in the sentence, we use the word sense disambiguation to extract the exact or nearest semantic score of the opinion expression.

Contact shifter: There are a few types of context shifters to populate our knowledge base with semantic scores; they are followed by some specific rules for semantic weight extraction from sentences and are shown below.

- The contact shifter (But, except, however, only, although, though, while, whereas, etc.)
- Contradictory nature contact shifter (Although, Despite, While)
- Mobilizing or modal contact shifter (Would, Should)
- Pre-Supposition contact shifter (Miss, forget, refused, assumed, hard, harder, less, etc.)

If sentences have any such type of word, then the polarity will be recalculated by checking their position in respect to the opinion expression, because these words affect the polarity of the opinion word. The negation words reduce its effect to nothing.

Modifiers (Enhancer and reducer): If there is a modifier word in the sentence (Slightly, somewhat, pretty, really, very, extremely, (the) most), closer to the sentiment terms, then its polarity will be recalculated by referring to its weightage dictionary. The score of the opinion word will be affected in the sentence by checking its position in the sentence. E.g., in the sentence the staff was very nice and cooperative, very is enhancing the weight of the nearest opinion word, i.e., nice.

Modifiers of certain nouns: Certain modifiers like (a (little) bit of, a few, Minor, Some, a lot, Deep, Great, a ton of) effect the sentence polarity, so recalculate the polarity if such types of words occur. Use the dictionary of the weights of the words/terms to assign weights to each sentence accordingly.

CONTEXTUAL SEMANTIC ORIENTATION OF SENTENCES

In this section, we describe the process of assigning a weight to each sentence and deciding whether the review is positive, negative or neutral. We used the rule based method to check the polarity of sentences and the contextual information at the sentence level. The process is used to extract the contextual information from the sentence and calculate its semantic orientation using SentiWordNet, WordNet and predefined intensifier semantic score dictionaries. From the results, it is clear that contextual information and consideration of sentence structure for correct sense extraction is very important for the useful sentiment classification. The main contributions of this work are sentence level semantic pattern extraction for word sense disambiguation, consideration all the parts of speech (POS) of the sentence for semantic orientation and it is a domain independent based polarity classification. However, the limitations of this study include the dependency on a lexical dictionary and limited word sense disambiguation. We evaluate the system on several datasets and online comments that it's outperformed. The following process shows the overall polarity calculation of the proposed method to observe the sentence structure.

- Split the reviews into sentences and a Bag of sentences is created (BoS):

REVIEWS: = split corpus
SENT: = Split Reviews
REW_ID:= Assign ID to each Review
SENT_ID:= Assign ID to each sentence
WORD_LIST:= list of words in sentence
WORD_POSITION: = position of each word in a sentence

- Classify the sentences into subjective and objective
- Applying POS and clean the sentences and take subjective sentences for further processing:

SENT – sentences to be tagged
WORD_LIST:= list of words in sentence
For each WORD in SENT compare with LEXICON and tag it
RETURN TAG_SENT

- Check each sentence and find the required word (WRD), if exist in the sentence, the extract its position in the sentence. X= POS_WRD
- Check the Opinion Word (OW) in the sentence by calculating its position as (X-5) and (X+5) in the sentence. If found then mark is as opinion sentence and assign The word to N ie (N=OW)
- For the correct sense, extract the sense-id from WordNet using semantic pattern of the desired sentence, refer to SentiWordNet the semantic score of WRD is extract on the basis of that sentence structure

SELECT only NN JJ RB VB from TAG_SENT
Place WRD at NN JJ RB VB place
CONCATINATE tags of k+3 and k-3 with WRD
RETURN DES_PATTERN
SLT_PATTERN – extracted pattern from WordNet glossary
SELECT SENSE_NO of SLT_PATTERN from WORDNET

- Now calculate its word semantic orientation and assign a weight to this word from the SentiWordNet dictionary (OW_SEM_SCOR):

SENTIM_WORD_SCORE:= extract positive negative score from the SentiWordNet according to SENSE_NO
IF the POSITIVE_SCORE is greater than NEGATIVE_SCORE THEN
SENTIM_WORD_SCORE:= POSITIVE_SCORE
ELSE the POSITIVE_SCORE is less than NEGATIVE_SCORE THEN
SENTIM_WORD_SCORE:= NEGATIVE_SCORE

- Sentence level polarity is calculated as consider the sentences to calculate the average score in the sentences the following rules are take into consideration:

MODIFIER_WEIGHT:= weight of SENT_SENTIM_WORD in MODIFIER_DICT
MODIFIER_DICT: = list of Modifier which affects the score of positive and negative polarity
IF SENT_SENTIM_WORD is similar JJ OR SENT_SENTIM_WORD is similar RB THEN
CHECK (SENT_SENTIM_WORD + 3) and (SENT_SENTIM_WORD - 3) for Modifier from MODIFIER_DICT
IF WORD found as MODIFIER THEN calculate overall weight.

- If there is negation word (Not, Never, N't, Does'nt, Cannt, Nor, Don't, Would'nt, No) near the N, Check (N+3) and (N-3) then reverse its polarity. e.g. (OW=+0.8 → OM= -0.8)
- If there is any type of context shifter in the sentence then the polarity will be recalculated because these words affect the polarity. The position of the contact shifter were checked in sentences, then check the nearest opinion word may be JJ, JJS, noun NN, NNS or VB, VBS, if its score is negative then it will be change it after recalculating its weights and vice versa. The negation words reduce its effect to nothing
- Check the modifier word in the sentence, if exists then recalculate the polarity referring the weightage dictionary the same process will be repeated that score of which opinion word will be effected. e.g, in the sentence the staff were very nice and cooperative, in this sentence the very is enhance the weight of the nearest opinion word i.e., nice
- Curtains nouns affect the sentence polarity, so recalculate the polarity if such types of word occur. From the dictionary of the weights of words/terms, assign weights to each sentence accordingly. The steps of rule base system for contextual valance shifter is describes as below:

IF the MODIFIER is a negation modifier THEN
SENTIM_WORD_SCORE:= Reverse the polarity of SENT_SENTIM_WORD
IF the MODIFIER is a intensifier THEN
SENTIM_WORD_SCORE:= intensifying MODIFIER_WEIGHT obtained from MODIFIER_DICT
SENTIM_WORD_SCORE:= SENTIM_WORD_SCORE + MODIFIER_WEIGHT
IF the MODIFIER is a decelerator OR IF the MODIFIER is enhancer OR IF the MODIFIER is context shifter THEN
SENTIM_WORD_SCORE:= intensifying MODIFIER_WEIGHT obtained from MODIFIER_DICT
SENTIM_WORD_SCORE:= SENTIM_WORD_SCORE + MODIFIER_WEIGHT

- Calculate the final weights of each sentence and review to decide if it is positive, negative or neutral. So, the opinion strength for both sentence and feedback is calculated by assigning the combined opinion weight to the sentence and review using the Eq. 3 and 4:

$$\text{SentenceScore(Sen)} = \frac{\sum_{i=1}^{n} \text{Score(i)}}{n} \quad (3)$$

Where,

Score (Sen), are the positive or negative score of sentence Sen, Score(i) is the positive, negative score of ith word in sentence S. n is the total no. of words in Sen:

$$\text{ReviewScore(Rew)} = \frac{\sum_{i=1}^{n} \text{Score(Sen)}}{n} \tag{4}$$

Where,

Rew(Score), are the positive or negative score of Review Rew, Score(Sen) are the positive, negative score of ith sentences in review. n is the total no. of sentences in the review.

Machine learning based classification: In this section, we processed the Naïve bayes algorithm for sentence based sentiment classification. We train the Naïve bayes classifier by taking the annotated sentences from our system (Baharudin *et al.*, 2010). This time, we take the annotated sentences extracted from our rule based system as training data and test on different review datasets for the sentiment classification of positive and negative opinions. This word level classification assigns the positive and negative polarity to the new sentences as described in Fig. 2.

- Train word level classifier using n-grams as input
- Use the classifier to mark the sentences as positive or negative

Naïve bayes sentence classification: A supervised classification algorithm is used for classification of label sentences. The classifier is trained on a small labeled dataset and tested on similar examples. Naïve bayes calculates the prior probability frequency of each label in the training set. Each label is given a likelihood estimate from the contribution of all features from the feature set; the label is assigned (nltk.classify.naivebayes) to the highest likelihood estimated sentences. We employ the python-NLTK naïve bayes[2] algorithm for feature based sentence classification.

Naive Bayes algorithms find the probability for a label by first using the Bayes rule to express P (label | features) in terms of P(label) and P(features | label):

$$P(D \mid c_i) = \prod_{j=1}^{n} P(d_j \mid c_i) \tag{5}$$

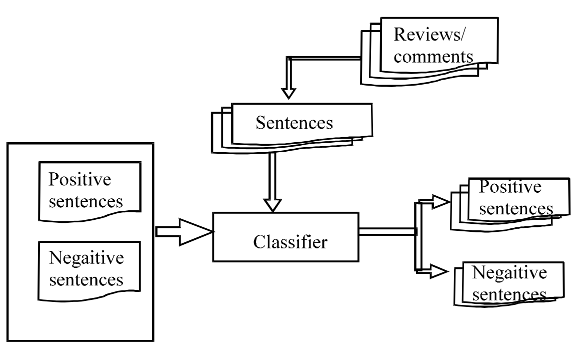

Fig. 2: Automatic classification of sentences

EXPERIMENTS AND RESULTS

Three types of online customer review datasets are collected for our system performance. (1) Popular publicly available corpus from movie-review polarity dataset v2.0 IMDB movie reviews (http.nltk.classify.naivebayes). The data set consists of 1000 positive and 1000 negative reviews in individual text files; also, the sentences polarity dataset (includes 5331 positive and 5331 negative processed sentences/snippets (Pang and Lee, 2005). We take the positive and negative sentences to check the performance of our proposed system. (2) We extract 1000 reviews from Skytrax (http://www.cs.cornell.edu/people/pabo/movie-review-data/), where there are more than 2.5 million independent reviews for over 670 airlines and 700 airports. After splitting the reviews into sentences, an average of 8 sentences per review is found. We extract the subjective lexicons and semantic orientation from all the positive and negative sentences. (3) We perform our experiments on the dataset of about 2600 downloaded hotel reviews, which are collected from TripAdvisor (http://www.airlinequality.com/), one of the popular review sites about hotels and travelling. We extract only text of reviews using text file. Table 6 and 7 show the customer reviews, no. of sentences per review/comment and the objective and subjective sentences in the reviews.

All the datasets were processed to remove the noise and clean up the special characters and symbols and check for spelling mistakes; furthermore, we apply the POS tagger and classify the sentences into subjective and objective as shown in Table 7. The movie reviews data has already been processed for positive and negative sentences. Only the subjective sentences were taken for further processing to find the semantic orientation at the individual sentence level. Figure 3 shows the classification of subjective and objective sentences (taken from our proposed system) for the airline and movie review datasets.

The subjective sentences were processed for semantic orientation by taking the contextual features and using the SentiWordNet for the semantic score. The weight is calculated using the Eq. 3 and 4 for the final opinion orientation. The results were evaluated by using precession and recall. Table 8 shows the overall accuracy of our proposed method.

Figure 4 shows the graphical representation of positive, negative or neutral opinion orientation for comments and review data. The system achieves accuracy of about 93% for the feedback level and about 86% at the sentence level. So, the rule based system with the lexical system performs better then machine learning and word level sentiment analysis.

Table 9 shows the overall performance of the proposed system compared to the machine learning and Hu and Liu (2004b) methods, taking the feature list as seed for the opinion orientation. Our system improves the semantic extraction efficiency up to 2% and the opinion

Table 6: Processed datasets

Datasets	Comments	Sentences	Sentences/comments(average)
Movie reviews	Already processed	10662	10
Airline reviews	1000	7730	8
Hotel reviews	2600	25663	10

Table 7: Sum of Opinion sentences

Dataset	Reviews	Sentences	Subjective	Objective	Percentage (sub/Ogj)
Movie reviews	Already processed	10662	8530	2132	80/20
Airline reviews	1000	7730	5405	2325	70/30
Hotel reviews	2600	25663	17704	7969	68/32

Fig. 3: Percentage of subjective and objective sentences for customer reviews

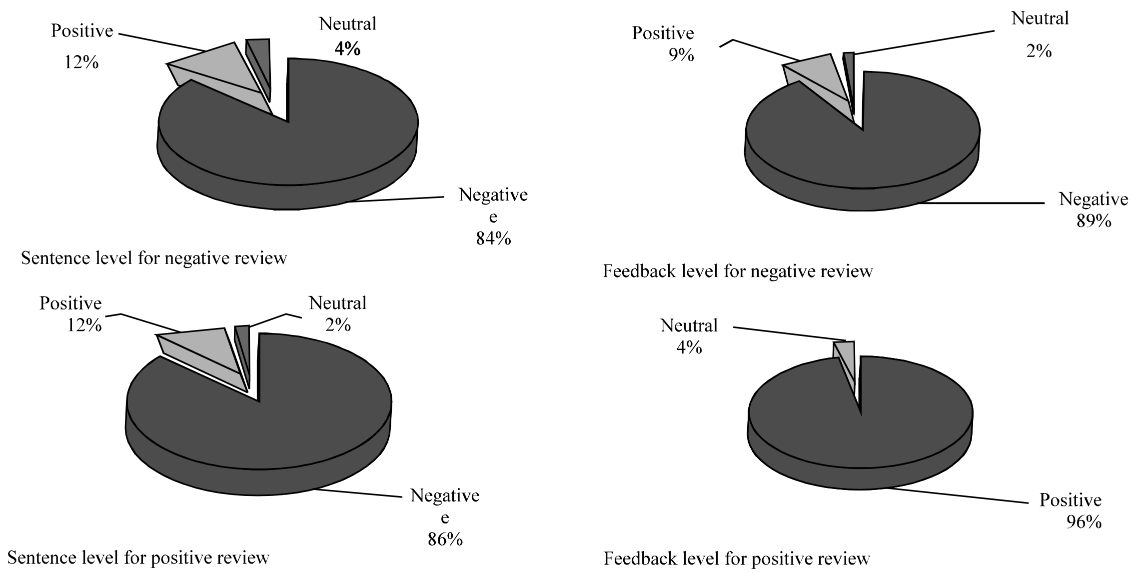

Fig. 4: Accuracy of proposed method using different datasets

Table 8: Accuracy of opinion orientation for positive and negative sentiments for noisy text

		Sentence level accuracy		Feedback level accuracy	
Datasets	Sentiment orientation	Machine learning (%)	Proposed system (%)	Machine learning(%)	Proposed system (%)
Movie reviews	Positive	78	86.50	80	96.80
	Negative	73	84.80	74	95.70
	Weighted average		86.00		97.00
Hotel reviews	Positive	79	81.80	80	83.70
	Negative	77	76.00	75	79.50
	Weighted average		80.00		82.00
Airline reviews	Positive	75	87.80	82	94.80
	Negative	73	84.00	77	89.90
	Weighted average				93.00

Table 9: Evaluation for customer review dataset

	Hu and Liu (2004b)	Machine learning	Proposed method
Opinion oriented sentence extraction	69.3	55	80.0
Sentiment orientation accuracy at sentence Level	84.2	74	86.6
Feedback level	----	79	97.8

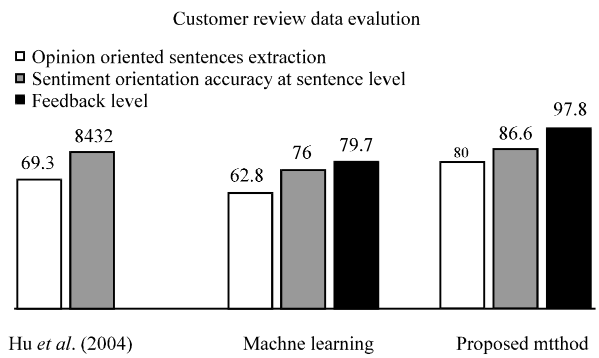

Fig. 5: Percentage of subjective and objective sentences for customer reviews

sentences extraction up to 10%. Our main contribution is the extraction of sentence level semantic orientation taking all parts of speech and sentence contextual structure. However, it depends on the lexicon dictionary which is the main drawback of this work. Figure 5 shows the performance of our proposed system as compared to machine learning (Hu and Liu, 2004b).

CONCLUSION AND FUTURE WORK

In this study, a rule based sentiment analysis approach is proposed for opinion classification. The contextual information and the sense of each individual sentence are extracted according to the pattern structure of the sentence. The semantic score for the extracted sense is assigned to the sentence using SentiWordNet. The final semantic weight is calculated after checking each semantic orientation of each term in the sentence. The decision is then made to check the polarity of positive, negative or neutral opinions. The results show that the sentence structure and contextual information in the review are important for the sentiment orientation and classification. The sentence level sentiment classification performs better than the word level semantic orientation. The limitations include the dependency on lexicons and the lack of term sense disambiguation. The experiments are performed on three types of customer review datasets. From the results, it is clear that the proposed method achieves an average accuracy of 86.6% at the sentence level and 97.8% at the feedback level for customer review datasets without removing noise from the text.

In future, the plan is to improve extraction of the acute sense of sentence and remove noisy text for an efficient semantic orientation. Furthermore, the knowledgebase will be improved for the semantic scores of all parts of speech.

REFERENCES

Alm, C.O., D. Roth and R. Sproat, 1990. Emotions from text: machine learning for text-based emotion prediction. Proceedings of the Human Language Technology Conference on Empirical Methods in Natural Language, October 2005, USA., pp: 579-586.

Amato, F., R. Canonico, A. Mazzeo and A. Picariello, 2011. Statistical and lexical analysis for semi-automatic extraction of relevant information from legal documents. J. Applied Sci., 11: 639-646.

Andreevskaia, A., and S. Bergler, 2006. Mining wordnet for fuzzy sentiment: Sentiment tag extraction from wordnet glosses. Proceedings of the 11th Conference of European Chapter of the Association for Computational Linguistics, (EACL'06), Trento, Italy, pp: 209-216.

Attardi, G. and M. Simi, 2006. Blog mining through opinionated words. Proceedings of the 15th Text Retrieval Conference, Nov. 14-17, National Institute of Standards and Technology, Maryland, pp: 2-7.

Baccianella, S., A. Esuli and F. Sebastiani, 2009. Multi-facet rating of product reviews. Proceedings of the 31th European Conference on IR Research on Advances in Information Retrieval, (ECIR'09), Heidelberg, pp: 461-472.

Baharudin, B., L.H. Lee and K. Khan, 2010. A review of machine learning algorithms for text-documents classification. J. Adv. Inform. Technol., 1: 4-20.

Choi, Y. and C. Cardie, 2008. Learning with compositional semantics as structural inference for subsentential sentiment analysis. Proceedings of the Conference on Empirical Methods in Natural Language Processing, (EMNLP'08), USA., pp: 793-801.

Dey, L. and S.M. Haque, 2009. Opinion mining from noisy text data. Int. J. Document Anal. Recognition, 12: 205-226.

Esuli, A. and F. Sebastiani, 2006. SentiWordNet: A publicly available lexical resource for opinion mining. Proceedings of the 5th Conference on Language Resources and Evaluation, (CLRE'06), Italy, pp: 417-422.

Hu, M. and B. Liu, 2004a. Mining opinion features in customer reviews. Proceedings of the 19th National Conference on Artificial Intelligence, July 25-29, AAAI Press, San Jose, California, pp: 755-760.

Hu, M. and B. Liu, 2004b. Mining and summarizing customer reviews. Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data, Aug. 22-25, ACM Press, Washington, USA., pp: 168-177.

Hu, Y. and W. Li, 2011. Document sentiment classification by exploring description model of topical terms. Comput. Speech Language, 25: 386-403.

Kasam, A. and K. Hyuk-Chul, 2006. Consolidation of diversifying terms weighting impact on IR system performances. Inform. Technol. J., 5: 7-12.

Khan, A., B. Baharudin and K. Khan, 2010. Sentence based sentiment classification from online customer reviews. Proceedings of the Conference on Frontiers of Information Technology, (FIT'10), ACM, pp:1-6.

Kim, S.M. and E. Hovy, 2003. Determining the sentiment of opinions. Proceedings of the 20th International Conference on Computational Linguistics, (ICCL'03), USA., pp: 1367-1373.

Liu, B., 2010a. Sentiment Analysis and Subjectivity. In: To Appear in Handbook of Natural Language Processing, Indurkhya, N. and F.J. Damerau (Eds.). 2nd Edn., University of Illinois, Chicago, USA., pp: 1-38.

Liu, B., 2010b. Sentiment analysis: A multi-faceted problem. IEEE Intell. Syst., 1: 1-5.

Moilanen, K. and S. Pulman, 2007. Sentiment composition. Proceedings of the Recent Advances in Natural Language Processing International Conference, Sept. 27-29, Borovets, Bulgaria, pp: 378-382.

Neviarouskaya, A., H. Prendinger and M. Ishizuka, 2009. Semantically distinct verb classes involved in sentiment analysis. Proceedings of the International Conference on Applied Computing, (ICAC'09), Japan, pp: 27-34.

Pang, B., L. Lee and S. Vaithyanathan, 2002. Thumbs up? Sentiment classification using machine learning techniques. Proceedings of the Conference on Empirical Methods on Natural Language Processing, July 6-7, Association for Computational Linguistics Press, Philadelphia, USA., pp: 79-86.

Pang, B. and A.L. Lee, 2004. A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. Proceedings of the 42nd Annual Meeting of the Association for Computational, Nov. 16-24, Kimberly Patch, pp: 271-278.

Pang, B. and L. Lee, 2005. Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales. Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics, June 17, Association for Computational Linguistics Press, Michigan, pp: 115-124.

Polanyi, L. and A. Zaenen, 2004. Contextual valence shifters. http://www.aaai.org/Papers /Symposia/Spring/2004/SS-04-07/SS04-07-020.pdf.

Popescu, A.M. and O. Etzioni, 2004. Extracting product features and opinions from reviews. http://turing.cs.washington.edu/papers/emnlp05_opine.pdf.

Sarvabhotla, K., P. Pingali and V. Varma, 2010. Supervised learning approaches for rating customer reviews. J. Intell. Syst., 19: 79-94.

Shi, H., G. Zhou and P. Qian, 2010. An attribute-based sentiment analysis system. Inform. Technol. J., 9: 1607-1614.

Subramanian, S.K. and N. Ramaraj, 2007. Automated classification of customer emails via association rule mining. Inform. Technol. J., 6: 567-572.

Turney, P.D., 2002. Thumbs up or thumbs down? Sentiment orientation applied to unsupervised classification of reviews. Proceedings of the 40th Annual Meeting on Association for Computational Linguistics, July 11-12, Philadelphia, USA., pp: 417-424.

Umer, M.F. and M.S.H. Khiyal, 2007. Classification of textual documents using learning vector quantization. Inform. Technol. J., 6: 154-159.

Whitelaw, C., N. Garg and S. Argamon, 2005. Using appraisal groups for sentiment analysis. Proceedings of the 14th ACM International Conference on Information and Knowledge Management, Oct. 31-Nov. 5, ACM Press, pp: 625-631.

Yu, H. and V. Hatzivassiloglou, 2003. Towards answering opinion questions: Separating facts from opinions and identifying the polarity of opinion sentences. Proceedings of the Conference on Empirical Methods in Natural Language Processing, (CEMNLP'03), USA., pp:129-136.

Zhao, J., K. Liu and G. Wang, 2008. Adding redundant features for CRFs-based sentence sentiment classification. Proceedings of the Conference on Empirical Methods in Natural Language Processing, October 2008, USA., pp: 117-126.