Study on Sentiment & Opinion for "Review Text" Corpus

Supratim Bhattacharya, Saberi Goswami and Jayanta Poray
Department of Computer Science & Engineering
Techno India University, West Bengal
EM /1 - Salt Lake, Sector - V, Kolkata - 91, INDIA
Email::bhattacharya.supratim@gmail.com, tuki.saberi@gmail.com, jayanta.poray@gmail.com

***Abstract*—Now a days, Microscopic blogs became a very distinguished and famed communicating tool among different internet users. Millions of users share their opinions on different aspects everyday. Thus microblogging websites play a significant role as a group of mixed resources of data for analysing opinion and sentiment. The main aim of Sentiment Analysis is to determine the sensibility of a speaker or a writer with respect to different topic for the overall contingent polarity of the document. Sentiment Analysis is also one of the major task in Natural Language Processing(NLP). The major challenges are focused on the edge of 4 V's(Velocity, Variety, Volume & Veracity)) where the data are purely unstructured. In this paper, a particular product reviews for different companies and categories are scanned for opinion mining purposes which may further be extended to different product also. Using the corpus, we build a classifier, used to determine positive, negative & neutral sentiments for a document[1]. A conventional process for illustration of sentiment classification is proposed with comprehensive procedure.**

***Index Terms*—Sentiment Analysis, Opining Mining, big-data, natural language processing, Text Mining.**

I. Introduction

The mental activity or reaction of acquiring knowledge and perceive thought, experience and the senses is called cognition. This obtained knowledge transforms to express Sentiments. As a whole sentiment is an attitude, thought, or judgment apt towards feeling. Sentiment analysis, which is also known as opinion mining, studies peoples' sentiments towards certain entities. Now a days, internet is a bucket of resources with respect to the information about sentiment analysis. From a user's perspective, people are able to post their own content through various social media, such as forums, blogs, or different online social networking sites[23].

Most of the time, these contents are purely unstructured. If it can be transformed into a structured form, then it can be analyzed. If it can be analyzed, then it can be interpreted to provide certain information. But the major challenge is to identify the type of count or interpretation that can be made from that type of contents. Brute force manual procedure may be an option but not a good one. In general the unstructured as well as semi-structured data is huge in comparison to the structured data. In a normal observation, generally 80% data is in unstructured form, whereas structured data represents only 20% of the information available to the organization. Which signifies that 80% of the data is purely in unstructured format. If conclusion need to be drawn from analyzing only 20% of their data, then there is a massive potential is hidden in analysing unstructured data. Exploring this potential directs towards the Big Data Analysis challenge. Text Analytics is a tool to address this issue [2].

Websites related to social media and blogs have emerged as an originator of different kind of information. Now-a-days people used to post real time messages about their opinion on different topics, perform discussion on current topics and express their views. In fact they also convey their sentiments towards different product they use in daily life. The respective manufacturing company of this type fo products have started to analyze these microblogs to get a sense of generalised sentiments for their different category of products. Sentiment analysis has become an area of interest with the rapid increase of available text data containing opinions, judgement and recommendations. The diversity of the data and various applications using sentiment analysis give rise to different issues that have yet to be fully disclosed in the existing systems. The challenge is to build technology and to built a opinion detection methods to detect and summarize the overall sentiment. Opinion detection systems using sentiment analysis have been developed to target customers and evaluate the success of marketing campaigns, to know the user experience with certain products on their image of brands[3].

Our aim is to determine user's emotions and attitudes and to adapt its behavior accordingly. Our target is to make a user detected socio-emotional behaviours and sentiment.

to detect users'

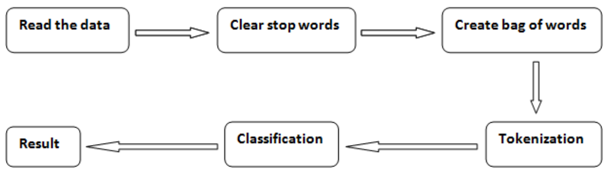

Fig. 1. Types of Information System

number of

As the bloggers and users of social mnedia increases everyday, data from these sources can be utilised in opinion mining and sentiment analysis.We use a dataset which is a combination of different blogs or reviews. It contains a huge number of short messages collected from different micro-blogging platform. The message content vary from individual thoughts to general statements. For example, any retail organization may be interested in the following questions[21]:

In which

1) Which type of product customers are interested most?
2) What type of brand they are looking for?
3) Any seasonal effect for a particular product?
4) Attraction towards different combination of products?

Interesting outcome may be observed where any political party studied about the popularity of any program which they have organized. All this information can be obtained from micro-blogging or any social media services.

In this paper, we study how microblogging can be used for sentiment analysis purposes. We show how to use reviews as a corpus for sentiment analysis and opinion mining. We use reviews for the following reasons:

1) Reviews are used by different people to express their opinion about different topics, thus it is a valuable source of people's opinions.[4]
2) Reviews contains an enormous number of text posts and it grows every day. The collected corpus can be arbitrarily large.[4]
3) Reviews audience varies from regular users to celebrities, company representatives and politicians. Therefore, it is possible to collect text posts of users from different social[4] and interests groups.
   Reviews'
4) Also reviews audience may represented by users from many countries.[4]
   reviews'

We collected a corpus of review text posts from different reviews split automatically between three sets of text:

In the context of text-categorisation, author et al [20] describes the followings:

1) Texts containing positive emotions[20].
2) Texts containing negative emotions.[20]
3) Objects texts that only state a fact or do not express any emotions.[20]
   Objective

We perform a linguistic analysis from our data corpus and we show how to build a sentiment classifier that uses the collected corpus as the training data.

II. Literature Review

The main aim of Sentiment Analysis is to highlight on the aspect of positive or negative sentiment. Feelings or judgement is the nutshell of human activity which influence our behaviour.

Different Levels of Analysis[19] As per <author et al>...

1) Document Level:- Whether a entire document express a positive or negative sentiment.[19]
2) Sentence Level:- Differentiates sentences that provided circumstantial information from sentences that expresses subjective views and opinions.[19]
3) Entity & Aspect Level:- The concept is based on an opinion which consists of a sentiment and to perform comprehensive and impenetrable analysis[19].

Also <author et al> described that there are 2 types...

According to [18] there are two types of opinions

1) Regular Opinion :- Expresses a sentiment only on an specific entity or an entity characteristics.
   the characteristics of an entity.
2) Comparative Opinion :- Compares multiple entities based on some of their shared aspects.

Sentiment words or opinion words used to express positive and negative sentiments. A part from individual words there are also phases and idoms. A list of such words or phases is called a **sentiment lexicon**[18].

are

In [18] we find that

1) In different domain different sentiment word(negative or positive)or phrases may have opposite orientation in acceptance of meaning.
2) A sentence comprises of sentiment words sometimes may not convey any sentiment or opinion.
   comprising
3) Sarcastic sentences containing sentiment words sometimes causes difficult scenario in performing any judgement.
4) It may happen with many sentences where opinion may be implied without sentiment words.

The main objective is to extract a data slice from an huge set of unstructured data to make a inter / intra relationship among various subproblems.

According to [24], A opinion consists of two key components: a target 't' and a sentiment 'p' on the target, i.e. (t,p)
where
't' can be any entity or aspect of the entity for expression the opinion expressing
'p' is a expressed as positive, negative or neutral sentiment, or a numeric rating score expressing the intensity or strength of the sentiment (e.g. 1 to 5 stars).

In general, +ve, -ve and neutral are called sentiment orientation... [24]

[24]. Told that positive, negative and neutral are called sentiment (or opinion) orientation (or polarisation).

From different study we came to know that an opinion is a quadruple,

(g, s, h, k) <- italics

where

g is the opinion (or sentiment) target,
s is the sentiment about the target,
h is the opinion,
k is the time used for opinion expression.

Entity:- An entity 'e' is described as a product, service, topic, issue, person, organization or event. It is described with a pair,

e:(S,W)

where

S is denoted by various ranking of parts and sub-parts and

W is a set of attributes of e.

Maintaining the hierarchy of Entity whether it is problem-attribute or sub-problem attribute is hard. To simplify, the hierarchy is made in two levels and use the terms aspects to denote both parts and attributes.

So, opinion can be redefined as quintuple,

$(e_i, a_{ij}, s_{ijkl}, h_k, x_l)$,

where

e_i is the name of an entity,
a_{ij} is an aspect of e_i,
s_{ijkl} is the sentiment on aspect,
a_{ij} of entitye_i,
h_k is the opinion holder,
x_l is the time when the sentiment is conveyed by h_k.

The opinion s_{ijkl} is positive, negative or neutral which may be expressed with different intensity and strength levels, e.g. 1 to 5. When the sentiment varies on the entity itself, the distinct aspect GENERAL is used to indicate it. Here e_i and a_{ij} together represent the opinion target[25]. line gap

Several important aspects of opinion mining

Important aspect:-

1) Five pieces of information in the quintuple must correspond to one another. That is, the sentiment s_{ijkl} must be specified by opinion holder h_k about aspect a_{ij} of entity l_i at time t_l. Any mismatch is an error.
2) All the five component is essential. components are
3) Does not covers all possible facts.
4) Provides a framework to transform[17] unstructured text to structured data. This quintuple above is a database schema.

Definition:-

Entity category and entity expression:- An entity category represents a unique entity, while an entity expression is an definite word or expression that appears in the text indicating an entity category.

The process of grouping entity expressions into entity categories is called entity categorization.

Aspect category and aspect expression:- An aspect class of an entity focuses on a unique aspect of the entity, while an aspect produces distinctive and actual word or phrase within the text indicating an aspect class.

Explicit aspect expression:- Aspect interpretation that are nouns and noun phrases named explicit aspect expressions.

Implicit aspect expression:- Aspect expressions containing nouns or not noun phrases are called implicit aspect expressions. other (parts of speech)

Sentence subjectivity:- An *objective sentence* presents some factual information about the world, while a *subjective sentence*expresses some personal feelings, views, or beliefs.

Model of entity:

An entity e_i is represented by a finite set of aspects $A_i = \{a_{i1}, a_{i2}, ... a_{in}\}$. This entity e_i can be expressed with any one of a finite set of its entity expressions $\{ee_{i1}, ee_{i2}, ..., ee_{is}\}$. Each aspect $a_{ij} \in A_i$ of entity e_i can be explained with any one of its finite set of aspect mentioned as $\{ ae_{ij1}, ae_{ij2}, ..., ae_{ijm} \}$.

Model of opinion document:-

An opinion script d contains opinions on a bunch of entities $\{e_1, e_2, ... e_r\}$ and a subset of their aspects from a set of opinion holders $\{h_1, h_2, ..., h_p\}$ at some particular time point.

Sentiment analysis consists of the following 6 main tasks:-

Task 1**(entity extraction and categorization):** Extract all entity statements in D, and categorize or group synonymous entity expressions into entity clusters (or categories). Each entity expression cluster indicates a unique entity e_i.

Task 2**(aspect extraction and categorization):** Extract all aspect expressions of the entities, and categorize these aspect expressions into clusters. Each aspect statement is a group of entity e_i represents a exclusive aspect a_{ij}. , representing an ...

Task 3 **extraction and categorization):** Extract opinion holders for opinions from text or structured data and categorize them. The task is analogous to the above two tasks.

Task 4 **(time extraction and standardization):** Extract the times when opinions are given and standardize different time formats. The task is also analogous to the above tasks. This

Task 5 **(aspect sentiment classification):** Determine whether the is sentiment on an aspect a_ij is positive, negative or neutral, or accredit a numeric sentiment rating to the aspect.

Task 6 **(opinion quintuple generation):** Generate all sentiment quintuples (e_i, a_{ij}, s_{ijkl}, h_k, t_l) expressed in record d based on the credentials of the above tasks.

Sentiment analysis (or opinion mining) based on this framework is often called aspect-based sentiment analysis (or opinion mining), or feature-based[15] sentiment analysis (or opinion mining).

Different Types of Opinions

Regular opinion: A *regular opinion* is often referred to simply as an *opinion* in the literature and it has two main sub-types:-

Direct opinion: A *direct opinion* signifies an opinion that expressed directly on an entity or an entity aspect.

Indirect opinion: An *indirect opinion* is a *sentiment* that is evaluated indirectly on an entity or aspect of an entity depending on its effects on some other entities. This sub-type often occurs in the medical domain.

Comparative opinion: A comparative opinion expresses a relation of similarities or dissimilarities between two or more entities and/or a choice of the opinion holder depending on the aspects of the choice of the entities.

Explicit and Implicit Opinions:

Explicit opinion: An explicit opinion is a subjective statement that gives a regular or comparative opinion.

Implicit (or implied) opinion: An *implicit opinion* is an unbiased statement that implies a regular or relative opinion. Such an equitable statement usually encompases a enticing or undesirable fact.

Document Sentiment Classification

From various studies like [20]we classify Opinion document as the document level sentiment classification because it determines the overall document as an essential information unit.

It states that, given an opinion document *d* concluding an entity, determine the whole sentiment *s* of the opinion holder bearing the entity, i.e., determine s expressed on aspect GENERAL in the quintuple –

(–, GENERAL, s, –, –),

where the entity *e*, opinion holder *h*, and time of opinion *t* are assumed known or irrelevant (do not care).

There are two formulations based on the type of value that s takes. If s takes categorical values, e.g., positive and negative, then it is a classification problem. If s takes numeric values or ordinal scores within a given range, e.g., 1 to 5, the problem becomes regression.

Sentiment classification or regression assumes that the opinion document *d* (e.g., a product review) evaluates sentiment on an individual entity *e* and contains opinions from a single opinion holder *h*. But if an opinion document estimate more than one entity, then the opinion on the entities can be contradicting.[20]

Supervised Learning for Sentiment Classification

From the study of [14] we came to know that sentiment classification is usually formulated as a two-class classification problem, positive and negative. Training and testing data used are normally product reviews. Since online analysis have appraisal scores assigned by their reviewers, e.g., 1 to 5 stars, the positive and negative classes are determined using the ratings.

Sentiment classification is essentially a text classification problem. Traditional text classification mainly classifies documents of different topics, e.g., politics, sciences, and sports. In such classifications, topic related words are the key features. However, in sentiment classification, sentiment or opinion words that indicate positive or negative opinions are more important[14].

In their descriptions about the terms and frequency, author et al [13] states the following features:

[13] states the following features:

Terms and their frequency: These features are individual words (uni-gram)and their n-grams with associated frequency counts. They are also the most common features used in traditional topic-based text classification. In some cases, word positions may also be considered.[13]

Part of speech:The part-of-speech (POS) of each word can be important too. Words of different parts of speech (POS) may be treated differently. Adjectives are important indicators of opinions. Thus, some researchers treated adjectives as special features. However, one can also use all POS tags and their n-grams as features.[13]

Sentiment words and phrases: Sentiment words are words within a language that are used to express positive or negative sentiments. Mostly adjectives and adverbs are used as sentiment words, but nouns (e.g., rubbish, junk, and crap) etc and verbs (e.g., hate and love) etc can also be used to denote sentiments. Apart from individual words, there are also sentiment phrases and idioms.[13]

Rules of opinions: Apart from Along with sentiment words and idioms, other language compositions and expression that can be used to express or imply sentiments and opinions.[13]

Sentiment shifters: These are expressions that are used to change the sentiment orientations, e.g., from positive to negative or vice versa. Negation words are the most important class of sentiment shifters.[13]

Syntactic dependency: Words dependency-based features generated from parsing or dependency trees are also tried by researchers.[13].

Penn Treebank Part-Of-Speech (POS) tags

textbfSentiment Classification Using Unsupervised Learning

From this study From the study by author et al [6]... [6] we came to know that sentiment words are often the dominating factor for sentiment classification. In unsupervised manner, sentiment can be classified and distribute in words and phrases. Here the Classification is observed based on some fixed syntactic patterns which are likely to be used to express opinions.

The syntactic patterns are based on part-of-speech (POS) tags. The algorithm defined by [6] is composed of three steps:

In the first step, two consecutive words are extracted if their POS tags conform to any of the patterns mentioned in the following table.

Bold

Tag	Description	Tag	Description
CC	Coordinating conjunction	PR$	Possessive pronoun
CD	Cardinal number	RB	Adverb
DT	Determiner	RBR	Adverb, comparative
EX	Existential there	RBS	Adverb, superlative
FW	Foreign word	RP	Particle
IN	Preposition or subordinating conjunction	SYM	Symbol
JJ	Adjective	TO	to
JJR	Adjective, comparative	UH	Interjection
JJS	Adjective, superlative	VB	Verb, base form
LS	List item marker	VBD	Verb, past tense
MD	Modal	VBG	Verb, gerund or present participle
NN	Noun, singular or mass	VBN	Verb, past participle
NNS	Noun, plural	VBP	Verb, non-3rd person singular present
NNP	Proper noun, singular	VBZ	Verb, 3rd person singular present
NNPS	Proper noun, plural	WDT	Wh-determiner
PDT	Predeterminer	WP	Wh-pronoun
POS	Possessive ending	WP$	Possessive wh-pronoun
PRP	Personal pronoun	WRB	Wh-adverb

Table: Tags and their descriptions of POS.

In step 2, Sentiment Orientation (SO)is estimated of the extracted phrases using the Pointwise Mutual Information (PMI)(that measures the degree of statistical dependence between two terms). measure which is defied by [6] as:

$$PMI(term_1, term_2) = log_2\{\frac{Pr(term_1) \wedge Pr(term_2)}{Pr(term_1)Pr(term_2)}\} \quad (1)$$

Here, $Pr(term_1 \wedge term_2)$ is the actual co-occurrence probability of $term_1$ and $term_2$, and $Pr(term_1)Pr(term_2)$ is the co-occurrence probability of the two circumstances if they are statistically self-sufficient. The sentiment orientation (SO) of a phrase is computed based on its association with the decisive quotation reference word "excellent" and the negative reference word "poor".

$$SO(phrase) = PMI(phrase, \text{"excellent"}) - PMI(phrase, \text{"poor"}) \quad (2)$$

In Step 3, the average review is calculated and classification of positive or negative is done based on the value of SO.

Another unsupervised approach used by [12] is the lexicon-based method, which uses a dictionary of sentiment words and phrases with their associated orientation and strength, and computes computer sentiment score based on the incorporated incorporates intensification and negation.

Author et al [12] ... [12] experimented following four strategies in order to transfer sentiment classification:-

(1) Exercise on a combination of defined reviews from other domains where data are available for testing on the objective domain;

(2) Training a classifier, but reducing the set of appearences to those only observed in the objective domain;

(3)Using ensembles of classifiers from domains with available labeled data and testing on the objective domain;

(4) Combining small amounts of labeled data with huge amounts of unspecified data in the objective domain (this is the traditional semi-supervised learning setting).

line gap

Here SVM was used by [12] for the first three strategies and EM was used in the last strategy.

III. Contributions

Human preference are practically unpredictable. Psychology and Sociology are the stream of science that helps us study this things. Emotions and behaviours displays a characteristic at a certain moment. Some testing strategies is based on Hypothetical Testing. We can treat social media as a huge psychological database. We can analyse these text with the power of Machine Learning.

The contributions are as follows:

1) Objective reviews are collected by just collecting the slice of the huge data.
2) Next method allows to collect negative and positive sentiments such that no human effort is needed for classifying the documents.
3) We perform statistical analysis for of the generated slice.
4) A review system for sentiment classification is created.
5) An evaluation is made based on the set of real review posts.

Author et al [11] states that Overall sentiment Analysis is being done in respect to Extract Aspect from different reviews.

In first Phase two task need to be executed.

Task 1: Aspect term Extraction Given a group of sentences with pre-described entities, aspect terms are being classified in the sentence and returns a list containing all the recognizable aspect terms. Aspect term identified a particular aspect of the target entity.

Task 2: Aspect term polarity detection From the aspect terms within a sentence need to determine the polarity(positive, negative, neutral or conflict).

In second Phase another two task need to be executed.

Task 3: Aspect term categorization Given a predefined set of aspect categories, identify the aspect categories discussed in a given sentences.

Task 4: Aspect category polarity detection Given a set of pre-identified aspect categories, determine the polarity of each aspect category.

IV. PROPOSED MODEL

The flowchart is as follows:-

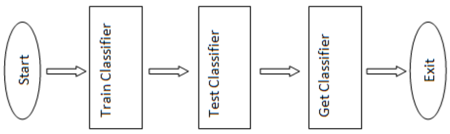

Fig. 2. Flow Chart

step

First is to train the classifier with train data and level i.e. positive or negative depending upon the review. Second is the test classifier and finally get sentiment of a given sentence of a word.

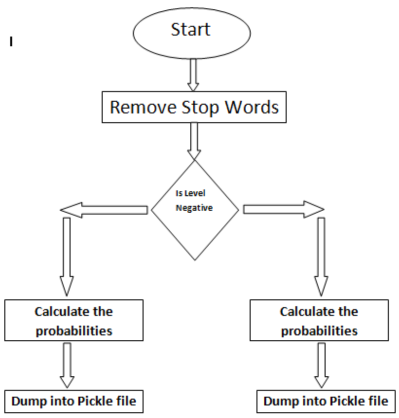

Fig. 3. Train Classifier

In fig 3, firstly we need..

First job is to remove the stop words i.e. neutral words consisting of Noun, pronoun and adverbs. Now for the remaining words based on the levels we divide it into positive and negatives and calculate the probabilities and each word is dump into pickle file.

and special characters

classes

of

and

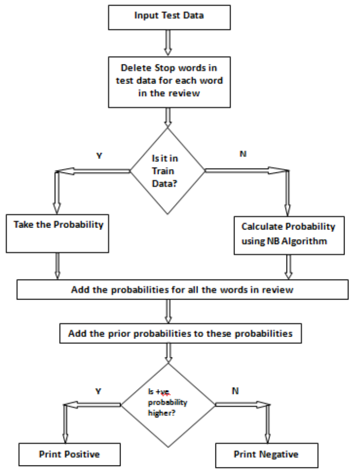

Fig. 4. Test Classifier and Get Sentiment

in fig 4 then we again

Here, the input is test data, again delete the stop words. Next Then for each word in the review, check is made if the word is within the training data or not. If it is yes then we take the probability associated with it which could be positive or negative. If not then we calculate the probability based on Naive Bayes algorithm. After repeating for all these words, we add all these probabilities individually for both positive and negative along with prior probabilities. If dimension of positive probabilities is higher then it is positive or else it is negative.

V. THEORETICAL RESULTS

overall

In order to implement the process we used **Algorithm used: - Naive Bayes Algorithm** is used.

the

Classification Procedure - Supervised Learning

following

Let's consider Taken the example as 3 positive sentiment & 2 negative sentiments,

Doc	Text	Class
1	I loved the movie	+
2	I hated the movie	-
3	a great movie, good movie	+
4	poor acting	-
5	a good movie, great acting	+

So a total of 10 unique words are identified and listed below:-

I, loved, the movie, a, great, hated, good, acting, poor

The unique words and their corresponding frequencies are shown in he below fig. figure below.

Doc	I	loved	the	movie	hated	a	great	poor	acting	good	class
1	1	1	1	1	-	-	-	-	-	-	+
2	1	-	1	1	1	-	-	-	-	-	-
3	-	-	-	2	-	1	1	-	-	1	+
4	-	-	-	-	-	-	-	1	1	-	-
5	-	-	-	1	-	1	1	-	1	1	+

Positive outcome are separated which is shown in the below figure.

Doc	I	loved	the	movie	hated	a	great	poor	acting	good	class
1	1	1	1	1	-	-	-	-	-	-	+
3	-	-	-	2	-	1	1	-	-	1	+
5	-	-	-	1	-	1	1	-	1	1	+

The probability of positive outcome is

P(+) = $3/5$ = 0.6 [Total document = 5, out of which 3 is positive]

Calculating each word being positive i.e. need to calculate below probabilities:-

P(I/+); P(loved/+); P(the/+); P(movie/+); P(a/+); P(great/+); P(good/+); P(acting/+)

To calculate that below mentioned formula is used:-

P(w_k/+) = (n_k+1)/n+—vocabulary—

where

n_k : number of times work k occurs in these cases (+)

n : number of words in (+) case. Here it is 14

there are 14 words in the vocabulary

vocabulary : Total Unique words.

While testing if any unknown word is found, then use nk = 0 and find its probability being +ve or -ve.

This is how we calculate probabilities using Naive Bayes.

Describe the result

VI. Discussion

1) We used rule based method is used, and some of the aspect term weren't extracted correctly by the rules.
2) If quantity of the aspect term is high, then it is difficult to assign polarity.
3) Unstructured data compared to structured data is difficult to process & analyse.
4) Classifier may have some limitations in respect to changing emotions.
5) A component need to be incorporated in terms of human analysis in sentiment analysis, as automated systems are not able to analyze historical tendencies of the individual.[8]

The accuracy of a sentiment analysis system depends on how well it agrees with human judgments. This is usually measured by alternative measures based on precision and recall based on the two objective categories of negative and positive texts. However, according to research human raters. It is observed that about 80%[40] of the time. Thus, a program which accomplish 70% efficiency in classifying sentiment is doing nearly as well as humans, even though such accuracy may not be so impressive. If a program were "correct" 100% of the time, humans would still contradict with it about 20% of the time, since they conflict with the answer. On the other hand, computer systems may experience errors compared to human intervension, and thus the figures are not entirely comparable. In general, the utility for practical financial application of sentiment analysis as it is defined in academic aspect has been called into question, mostly since the simple one-dimensional model of sentiment from negative to positive yields rather little actionable information for a client worrying about the effect[9] of public discourse on e.g. brand or corporate reputation.

VII. Conclusion

Now a days as the dataset is pure unstructured, it is a huge challenge to analyse & make a framework on Sentiment Analysis as well as Opinion Mining. We put up the challenge and ensures a conversion of unstructured data to semi-structured data but It is quite difficult to make it fully structured.

A process is designed to analyse Sentiments and Opinion. Accuracy & efficiency is not compared with other system but it is good enough to be used in regular micro-blogs and reviews in order to analyse Sentiments.

Considering the many benefits discussed in the article, companies see sentiment analysis as a major benefit computation for sales and improving their marketing strategies as well. To accomplish this, some companies develop their own tools and others rely on outsourcing companies that specialize in sentiment analysis.

References

[1] Abbasi, Ahmed, Hsinchun Chen, and Arab Salem. Sentiment analysis in multiple languages: Feature selection for opinion classification in web forums. ACM Transactions on Information Systems (TOIS). 26(3),(1972)

[2] Abdul-Mageed, Muhammad, Mona T. Diab, and Mohammed Korayem. Subjectivity and sentiment analysis of modern standard Arabic. in Proceedings of the 49 th Annual Meeting of the Association for Computational Linguistics:shortpapers, 2011.

[3] Akkaya, Cem, Janyce Wiebe, and Rada Mihalcea. Subjectivity word sense disambiguation. in Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing (EMNLP-2009). 2009.

[4] Andrzejewski, David, Xiaojin Zhu, and Mark Craven. Incorporating domain knowledge into topic modeling via Dirichlet forest priors. in Proceedings of ICML. 2009.

[5] Andrzejewski, David, Xiaojin Zhu, and Mark Craven. Incorporating domain knowledge into topic modeling via Dirichlet forest priors. in Proceedings of ICML. 2009.

[6] Asur, Sitaram and Bernardo A. Huberman. Predicting the future with social media. Arxiv preprint arXiv:1003.5699, 2010.

[7] Aue, Anthony and Michael Gamon. Customizing sentiment classifiers to new domains: a case study. in Proceedings of Recent Advances in Natural Language Processing (RANLP-2005). 2005.

[8] Banea, Carmen, Rada Mihalcea, Janyce Wiebe, and Samer Hassan. Multilingual subjectivity analysis using machine translation. in Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP-2008). 2008.

[9] Benamara, Farah, Baptiste Chardon, Yannick Mathieu, and Vladimir Popescu. Towards Context-Based Subjectivity Analysis. in Proceedings of the 5th International Joint Conference on Natural Language Processing (IJCNLP-2011). 2011.

This is usually measured by alternative quantifiers;
which is based on precision and recall with respect to the two objective categories
of negative and positive texts.

[10] Bilgic, Mustafa, Galileo Mark Namata, and Lise Getoor. Combining collective classification and link prediction. in Proceedings of Workshop on Mining Graphs and Complex Structures. 2007.

[11] Boiy, Erik and Marie-Francine Moens. A machine learning approach to sentiment analysis in multilingual Web texts. Information retrieval, 2009.

[12] Chaudhuri, Arjun. Emotion and reason in consumer behavior2006: Elsevier Butterworth-Heinemann.

[13] Dave, Kushal, Steve Lawrence, and David M. Pennock. Mining the peanut gallery: Opinion extraction and semantic classification of product reviews. in Proceedings of International Conference on World Wide Web (WWW-2003). 2003.

[14] Esuli, Andrea and Fabrizio Sebastiani. Determining the semantic orientation of terms through gloss classification. in Proceedings of ACM International Conference on Information and Knowledge Management (CIKM-2005). 2005.

[15] Ganesan, Kavita, ChengXiang Zhai, and Jiawei Han. Opinosis: a graphbased approach to abstractive summarization of highly redundant opinions. in Proceedings of the 23rd International Conference on Computational Linguistics (COLING-2010). 2010.

[16] Ghose, Anindya, Panagiotis G. Ipeirotis, and Arun Sundararajan. Opinion mining using econometrics: A case study on reputation systems. in Proceedings of the Association for Computational Linguistics (ACL-2007). 2007.

[17] Goldberg, Andrew B. and Xiaojin Zhu. Seeing stars when there aren't many stars: graph-based semi-supervised learning for sentiment categorization. in Proceedings of HLT-NAACL 2006 Workshop on Textgraphs: Graph-based Algorithms for Natural Language Processing. 2006.

[18] Hu, Minqing and Bing Liu. Mining and summarizing customer reviews. in Proceedings of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-2004). 2004.

[19] Kim, Soo-Min and Eduard Hovy. Extracting opinions, opinion holders, and topics expressed in online news media text. in Proceedings of the Conference on Empirical Methods in Natural Language Processing(EMNLP-2006). 2006.

[20] Liu, Bing, Wee Sun Lee, Philip S. Yu, and Xiao-Li Li. Partially supervised classification of text documents. in Proceedings of International Conference on Machine Learning (ICML-2002). 2002.

[21] Medlock, Ben and Ted Briscoe. Weakly supervised learning for hedge classification in scientific literature. in Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics. 2007.

[22] Mukherjee, Arjun, Bing Liu, Junhui Wang, Natalie Glance, and Nitin Jindal. Detecting Group Review Spam. in Proceedings of International Conference on World Wide Web (WWW-2011, poster paper). 2011.

[23] Riloff, Ellen, Siddharth Patwardhan, and Janyce Wiebe. Feature subsumption for opinion analysis. in Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP-2006).2006.

[24] Wang, Dong and Yang Liu. A pilot study of opinion summarization in conversations. in Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics (ACL-2011). 2011.

[25] Wu, Yuanbin, Qi Zhang, Xuanjing Huang, and Lide Wu. Phrase dependency parsing for opinion mining. in Proceedings of Conference on Empirical Methods in Natural Language Processing (EMNLP-2009). 2009.