Automated Sentiment or Opinion Discovery System

M.A. Jawale
I T Dept.S.R.E.S COE,
Kopargaon, SPPU, Pune (M.S)
JJTU, Rajasthan, India

D.N. Kyatanavar, Ph.D.
S.R.E.S COE, Kopargaon,
SPPU, Pune (M.S)
JJTU, Rajasthan, India

A.B. Pawar
Comp.Dept, S.R.E.S COE,
Kopargaon, SPPU, Pune (M.S.)
JJTU, Rajasthan, India

ABSTRACT

Today's world is relied on computer technology's advancement to get the best whatever they want or select. Since the possibility of sharing and exchanging information on internet, it is really easiest task than ever before and same technology aids are providing us ample amount of data, information while selecting best of services, best of products available as well as best of individual based on quality features they possesses. Even due to emerge of social media like blogs, forums, communities, twits, etc., now it is far superior to give feedback on any organization, services provided, product qualities, and on individual skills very easily. Additionally, like individual internet user, all kind of organization experts, management teams, analysts, government agencies are focusing on such data and its analysis for their business growths and trends in today's competitive world. In the same sense, this research paper focuses on development of automated opinion mining system to help, to analyze, to evaluate user's reviews, to provide on click solution of reviews mining for business decision making process. After comparing the results of proposed system with existing opinion mining system, it is found that first it combines the opinion mining system development approaches used earlier i.e. dictionary based and corpus based together which is rarely found. Also, it provides more accuracy in obtained results to make this system more trustworthy and efficient.

Keywords

Explicit Features, Feature Extraction, Implicit Features, Opinion Mining, Sentiment Analysis.

1. INTRODUCTION

Sentiment analysis or opinion mining is the significant study about individual opinions, appraisals, and emotions toward certain objects, facts and their attributes. In recent years, it has been attracted great attentions from both academia and industry due to many challenging research issues and support of sentiment analysis for a wide range of applications. Opinions are vital since whenever want to make proper decision it is wise to listen others opinions. This is not only true for individuals' selection, but it is useful and true for organizations too. However, in recent days, there was very little computational study did on opinions before the introduction of World Wide Web (WWW) because of limited availability of opinionated text for such analysis. In the past, when an individual go for certain decision, he/she typically used to ask for opinions from friends, families and others too. When an organization wanted to find opinions of the general public about its products and services, they were conducting surveys. With the very high growth of the social media content on the WWW, the world has been transformed and became rich in data through advancement of web. People can now post reviews of products on respective business organizational sites and can express their reviews on almost anything in discussion forums and blogs, and even at social network sites. Now if one wants to buy a product, no need to ask about the one's friends and family's opinion. For the industries, it may no need to conduct surveys or focus groups in order to gather consumer opinions about its products and those of its competitors because there is ample information publicly available [1].

As stated earlier, due to growth of internet and data contents on webs, finding opinion sites and monitoring them on the web is quite intensive task because the opinion contents are available on different sites, and each site may also have a huge volume of opinionated text. In many cases, opinions are not directly expressed and are hidden in long forum posts and blogs. So, it becomes difficult for a human reader to find relevant sites, extract related contents with opinions, analyses them, summarize them, and organize them into compatible and usable forms. Thus, to do all these tasks, there is need of automated opinion discovery and summarization system in the field of business decision making process to enhance the business strategies and business profit in the competitive world [2].

It is observed that, opinion contents which are available online as well as offline are containing mostly textual information used by the customer to provide relevant product feedback. So, this textual information in the world can be broadly categorized into two main types either facts or opinions. Facts are known as objective expressions about entities, objects, events and their properties. On the other hand, Opinions are usually subjective expressions that describe individual sentiments, assessment of performance or emotions toward entities, objects, events and their properties [3].

Most of the recent research study based on opinion mining was concentrated on textual information processing [11], [12]. Also, it has been focused on opinion mining and retrieval of factual information provided and expressed in opinionated text. It also supports information retrieval, text classification, etc., and many other text mining and natural language processing tasks. Processing of opinions was very little focused and studied concept to get exact opinions of customer than factual information.

This proposed research work intends to do opinion mining from opinionated text rather than traditional factual text processing [11], [12].

2. MOTIVATION AND RELATED WORK

The research community has studied individual sub problems of the sentiment analysis in the significant way. It is found, the most considered and well-studied sub problem is opinion orientation classification problem at the document level by many researchers as mentioned by [4] , followed by sentence level and feature level opinion classification reported by [5].

Based on a set of language patterns, [6] defined certain rule, to perform the extraction of two noun phrases from collected customer reviews and found good accuracy for recall with conducted experiments for the opinion mining. However, it lead to low precision for the two noun phrases extraction because of wrong Part of Speech (POS) tags generation in the said experimental work. So, it results in the need of improvement in the text sentiment orientation systems

The existing reported solutions are still far from required and expected one. The main issue is that the current studies are still coarse and not with finer details. For example, on opinion classification can be done on many conceptual rules that govern opinions as described and evaluated by [7] and there are even more expressions that can be used to share these concepts. Still there is need to do depth study on these concepts of the research field since the very little details available on said concepts in the existing research works.

With the explosive growth of internet and proliferation of social networking and e-commerce in the real life of users, the information available in the opinions or reviews expressed by the people has grown by leaps and bounds i.e. with Word of Mouth (WOM) [12]. In this research direction, [8] presented an opinion search engine system that incorporates two novel opinion mining algorithms to perform the opinion mining. The opinions are considered based on features and the orientation of these opinions is also largely based on the features rather than a product as a whole in this research work. But, it is found, people seem to like or dislike a specific product because of some feature associated and provided with the product. It gives the requirement of the research work that there is not only classifiers required to classify the input reviews as positive or negative one, but also there is need to extract the most potential features of each reviewed item, and assigns opinion scores on them. This is another reason to research on automated sentiment or opinion discovery system to utilize the feature extraction properly for further opinion mining.

Feature extraction and synonym grouping got attention of the research communities in the research field of opinion mining in recent years, but they remain to be very challenging as the reasons are identified and illustrated by [9].In opinion mining, Object extraction is the simplest task since many existing information extraction algorithms can be applied to do this task. However, integrating and matching of all 5 pieces of information in the quintuple of opinion mining is found still lacking [9], which is the main outcome of the research communities who focused on individual sub-problems of sentiment analysis rather than consideration of the same as multi-faceted one. This leads the issue of sentiment analysis accuracy, i.e., what is the accuracy of currently researched sentiment analysis algorithms and techniques? This question is not easy to answer because there are so many sub problems with individual issues. So it intends to the motivation to provide a completely automated solution for opinion mining. The key concept of study is to fully understand the whole range of issues and pitfalls in the sentiment analysis, cleverly manage them based on design and implementation of the automated system, and determine what portions can be done automatically and what portions need human assistance [9].

Existing sentiment analysis is focused on feature extraction techniques to improve the opinion mining. But most of the research work taken in account only explicit features of the expressed opinion or reviews by the customer and very little or no focus is given to identify potential of implicit features of expressed opinions. To find such implicit features, semantic understanding is needed, which requires more sophisticated techniques and in depth analysis of reviews. However, implicit features occur much less frequent than explicit ones. Thus, mostly existing study on opinion analysis focus on finding features that appear explicitly as nouns or noun phrases in the reviews [12].

Here, it gives clear indication about findings of opinion into the existing research work involved for opinion analysis. Their main contribution is to identify either opinion is positive, negative or neutral based on explicit opinion and its features identification only [12]. A very little focus is given on implicit opinion and its implicit features too. So, the proposed research work will focus on identifying explicit as well as implicit opinions through their explicit and implicit features analysis.

At the end, one has to understand that opinion mining requires dealing with several different challenges, such as identification of segment of text which is opinionated, identification of the opinion holder, determination of the positive or negative polarities of opinion. Even it is concerned for human reviews, emotions and sentimental discussion too. Everyone has his or her own perception and concern about a particular problem, issue, or topic on which he or she would like to express certain opinion. Surprisingly, opinionated text may be fake, irrelevant and or based on ambiguous information and very challenging to describe in the text form.

During this research work literature review, it sense that there is huge need of automated sentiment analysis or opinion discovery system with the enhanced performance and trusting results of the opinion mining for the business decision making and to trace the consumer trends for the organization products. Additionally, the usage of sentiment analysis field is not limited with the business organizations; they are extensively used by all kinds of organizations to perform individual assessment, to analyze organization growth and down trends, as well as to get the idea about the public trends on certain issues within short period of time. Even, as the social media emerged in everyone's life with the simplified and easy to access nature, it influenced the need of sentiment analysis at each possible field related to users of the social media.

3. IMPACT OF SOCIAL MEDIA IN SENTIMENT ANALYSIS
3.1 Introduction
The "social media" term specifies the wide range of Internet based and mobile based services that allow users to participate in online exchanges, contribute user authorship content, or join online communities i.e. information exchange and resource sharing [11]. Today, Social media is a powerful resource for people to exchange information. It is explosive growing source of texts ranging from everyday observations to involved discussions by the users of the social media. The increasing of new social networking sites has produced new opportunities as well as new challenges to individuals, organizations, and the society at large extend. Mainly, Individuals are using this social media for diverse purposes like, communication, coordination, co-creation, research, entertainment, mentoring, networking, friendship, career advancement, training and learning. The business organizations are also involved and connected with these social platforms to enhance performance and business through group sourcing, online collaboration and communication with the product users, online customer and business communities, mobile computing services and social media services too.

3.2 Research Challenges and Opportunities

These new advancements raise new challenges and changes with novel and important research issues, which gain the attention of researchers in this field. Today, it covers the research areas likes qualitative and quantitative approaches, social network analysis and evaluation, data mining, sentiment analysis or opinion mining. Moreover, social media facilitate users to create, share, and coordinate content. Even, it is useful to apply novel methods like decision making models, design framework patterns, and opinion or sentiment mining analysis, business intelligence can be obtained based on social media content analysis to do actionable decisions [10].

The social media elements are increasing and they are gaining popularity in people. However, these social networks were in past and still today are famous in young generation. But, in the following Fig.1, it shows that social media is not just for the teens and young people. The Fig.1 shows that increase in usage by the 55 to 64 year old user is greater than 100% for Facebook, Twitter and Google+, So it is no surprise that social media covers all the kinds of consumers about whom business organizations are looking after and it gives them platform to enhance their business decision strategies based on this huge information source expressed by these user on their certain products.

These social media open great marketing opportunities for businesses from all corners. Before, enlisting the benefits of using social media contents for the business purpose, one should highlight on the key findings of social media and then he or she will realize the requirements of using social media contents in business decision making process.
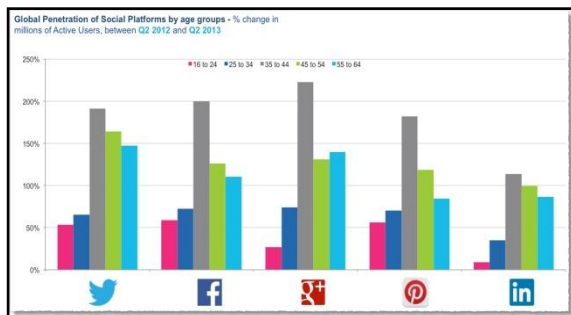

Fig. 1: Statistics Summary of Year 2013 of Increase in Social Media Usage by Age Groups of World Population[10]

The Global Web Index study report [10] published in year 2013 found these to be the major insights into social media usage and can influence the business organizations as shown in Fig. 1. Active usage of the major, global social platforms is growing worldwide very rapidly with mobile being the key driver for all the users.

3.3 Benefits of Social Media Content Analysis

Now, this gives the following benefits to the business organizations based on social media (i) to promote and enhance the name of your brand and business across the globe; (ii) to publish and tell the details to customers about your goods and services; (iii) to identify customer trends and find out what customers think of your business; (iv) to advertise your products and attract new customers to increase

business market; (v)to provide regular follow ups and build stronger relationships with existing customers.

In addition, many businesses are using social media into their marketing strategies to improve the business. Even for governments and parliamentarians, social media offer new ways to interact with citizens and to collect their views regarding various schemes launched for citizens. It allows companies, organizations, governments to reach millions of people.

4. SYSTEM MODEL

4.1 Objective

Based on the issues addressed for sentiment analysis or opinion mining systems main objective of the research work was set as: "To build an automated opinion discovery system to enhance performance of sentiment or opinion analysis based on feature extraction sentiment analysis sub technique, natural language processing and data mining techniques in an integrated way."

4.2 Proposed Model

As shortcomings observed in earlier work emphasizes and motivates the need of the system which can consider the sentiment analysis as multi-faceted problem, can determines user's sentiments based on implicit as well as on explicit features, can handle large amount of data & finally will give more trustworthy solution for business decision making process to any organization by having single- click solution to enhance their business strategies. The proposed system is step towards the same integrated system which will give the automation of entire sentiment analysis process.

Fig. 2 gives an architectural model for automated sentiment or opinion analysis system [3]. The system performs the opinion analysis in three main steps: Data collection and its pre-processing, opinion mining and generate opinion status.

The input to the system is in the form of review datasets where dataset will include product name and an entry page for all the reviews of the products. The output is the opinion status of the reviews. *i.e.* positive, negative or neutral opinion about the product. Based on given input, the system first downloads all the reviews, and puts them in the dataset for further processing. Following section outline the main components of the system as shown in Fig. 2. The details about this model are discussed in depth in [2].

Data pre-processing
The data from the dataset is pre-processed to set the data in the format which is acceptable to next computation of system. The outcome will be formatted dataset which will be the input for the opinion miming engine. As studied in literature review, most of the existing dataset or database files available for opinion mining are in the form of web pages formats, either in HTML or XML tags. So before opinion mining, it is necessary to filter out such tags to get opinion dataset only.

Opinion Mining Engine
This is the main component of the system model. It consists of two major steps for its computation namely, *Feature Extraction and Opinion Direction Identification.* The illustration of these steps is given in next section.
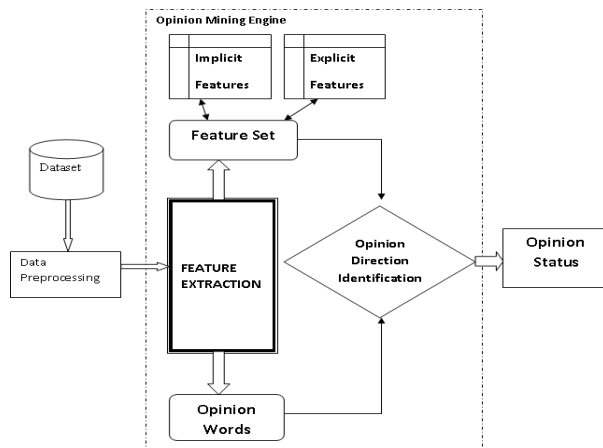
Fig. 2: Automated Sentiment Discovery System Model

The *feature extraction*, which is the first and foremost task of the work, it extracts "*frequent*" features that a lot of people have expressed their opinions on in their reviews, and then finds those *infrequent* ones. The *opinion direction identification* takes the generated features and states the opinions about the feature through the opinion status component of the system.

To perform part of speech tagging, use of the Natural Language Processor (NLP) like linguistic parser is used, which parses each sentence and yields the part-of-speech tag of each word (whether the word is a noun, verb, adjective, etc.) and identifies simple noun and verb groups. It will help to identify the expressed opinion is implicit or explicit type. The next step is to find features that people are most interested in. In order to do this, system will use association rule mining to find all frequent item sets. Not all frequent features generated by association mining are useful or are genuine features. There are also some uninteresting and redundant ones. Feature pruning aims to remove these incorrect features. Opinion words are words that people use to express a positive or negative opinion. Observing that people often express their opinions of a product feature using opinion words that are located around the feature in the sentence, we can extract opinion words from the review dataset using all the remaining frequent features (after pruning). After opinion features have been identified, one can determine the semantic orientation (i.e., positive or negative) of each opinion sentence.

Generate Opinion Status
This consists of two steps: (1) for each opinion word in the opinion word list, there is need to identify its semantic orientation and (2) then decide the opinion orientation of each sentence based on the dominant orientation of the opinion words in the sentence. The final output is status of the reviews as positive, negative or neutral opinion. Based on obtained scores of positive, negative and neutral opinions, scoring summary for each feature of the product is visualized for better understanding of the system users.

4.3 Mathematical Modeling
This section describes the details about the mathematical modeling used for the system computation and steps carried out to implement the system. For this proposed automated opinion discovery system, it is able to represent using set theory. Before describing whole system in mathematical form, it is important to analyze its problem solving and solution delivery capabilities.

As in computer computation, the problem definitions are mainly categorized as P-class, NP-complete and NP-hard problems. Thus, the defined and developed system is on the basis of producing the orientations of opinions positive, negative or neutral on the provided input dataset of user reviews about certain product features. It implies that, the developed system is result oriented and based on it further decision making process will work; so, it sense the selected problem definition leads towards P-class problem; where P-class problem is polynomial, time-stipulated and depends on fixed algorithm execution and solution-oriented. This section focuses on set theory based mathematical modeling of proposed system which include representation of system *input, process, rules and output* sets of the system. Further the relationship between these elements is mapped in Venn diagram representation in Fig. 3. Lastly, this modeling shows process state diagram to represent system in computational form as given in Fig. 4.

Input
This system deals with number of reviews given by users on certain products, which will be given as input for proposed system. For experimental analysis, such review dataset can be obtained online which are collections of reviews collected from various blogs, twits, forums, etc. as shown in Table 1. Here, the input for automated discovery system is considered for single product features having user's reviews from twitter dataset. Even, one can take directly reviews online via the web page of such product organization. Here, consider the input dataset given to this system for further analysis as set

$I=\{i_1,i_2,i_3..i_n\}$..(1)

Where, *I* is input dataset having reviews about certain product as i_1, i_2, i_3... i_n. For example, for 'Mobile Phone' product, the input dataset samples are shown in Table 1. The same input dataset '*I*' can be represented in the form of set 'S'.

$S=\{s_1,s_2,s_3...s_n\}$...(2)

Where, s_1, s_2, $s_3...s_n$ are the review sentences of users containing opinions or sentiments about the product.

Table 1. Input Dataset

User 1	Mobile has a good camera; it is of 5 Mega pixel.
User 2	The voice quality of this phone is great and so is the reception, but the battery life is short.
User 3	Speaker of Mobile is amazing.
User 4	The picture quality is amazing.
User 5	The battery life is long.

Processes

The processes are main elements of any system computation. These are the functions of the system to represent the complete flow of execution comprising of initial process to final step. Each process does certain task as part of system computation. For this proposed system, the main processes are described below: Let, *P* represent process, then

P1: Data preprocessing to make data compatible for next processing; *P2*: Classify the comments as fact and sentiments; *P3*: Classify the sentiments as implicit opinion and explicit opinion; *P4*: Feature extraction to extract implicit as well as explicit features of the product; *P5*: Opinion orientation to identify and to classify opinions; *P6*: Status indicator of oriented opinion i.e. either the input user review is positive, negative or neutral; *P7*: Visualization of each opinion status to decision making process

So set '*P*' is represented as

$$P= \{P1, P2, P3, P4, P5, P6, P7\}.. (3)$$

Rules

These rules are mainly used to decide the status of opinion orientation *i.e.* positive, negative or neutral. So, based on used algorithm in this system, the following main defined rules are

$$R= \{R1, R2, R3\}... (4)$$

Where, *R1*= If the final score is positive; then opinion on the feature in sentence S is considered as *Positive.*

R2 = if the final score is negative, then opinion on the feature in sentence S is considered as *Negative.*

R3 = otherwise opinion on the feature in sentence S is *Neutral.*

Output(s)

Output of the proposed system is alternative representation of equation (4),

$$O= \{O1, O2, O3\}.. (5)$$

Where, O1: *Positive* opinion; O2: *Negative* opinion; O3: Neutral opinion.

Now, the following Fig. 3 shows the relationship between the elements of System Set *S i.e. I, P, and O*
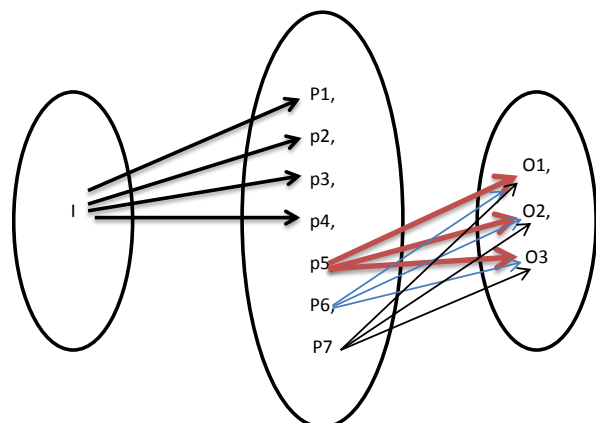

Fig. 3: Venn diagram to represent Mapping between Input, Process and Output of System

To support the functional mapping of the proposed system, the process state diagram is given in Fig. 4.
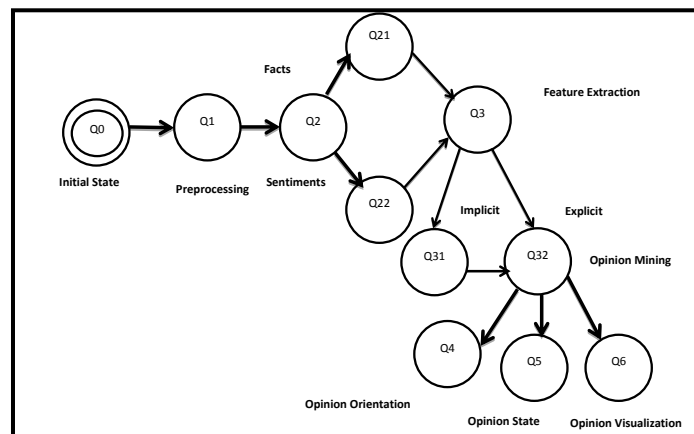

Fig. 4: Process State Diagram to Represent Function Mapping of System

5. EXPERIMENTAL ANALYSIS
5.1 Proposed System Evaluation
The developed system is implemented by using the following software & hardware environment. The main software includes: *Word Net is* a large lexical database of English language words. Nouns, verbs, adjectives and adverbs are grouped into sets of cognitive synonyms (synsets), each expressing a distinct concept. Synsets are interlinked by means of conceptual-semantic and lexical relations. Second is Natural Language Toolkit, or more commonly NLTK, is a suite of libraries and programs for symbolic and statistical Natural Language Processing (NLP) is used for the Python programming language and third is *Python* which is a remarkably powerful dynamic programming language that is used in wide applications. The proposed system is tested on the following minimum hardware configuration: RAM: 256MB; Processor: Intel Core series @ 2.27 GHz; Disk Space: 2 GB.

Implementation Steps carried out for Development
Initially, collected text files (i.e. user opinions) as an input extracted from standard dataset is given to opinion mining system. It classifies the comments into *facts* and *sentiments*. Further, from the obtained sentiments, the proposed system classifies sentiments as *Implicit* and *Explicit features.* For explicit features comments, system finds the "*feature*" and "*opinion-word*" and creates the association database of these features. Also, it applies association database on implicit features comments too. Finally, after this processing, system displays the orientations of each feature (implicit as well as explicit) as positive, negative and neutral one.

5.2 Obtained Results
This section highlights the initial experimental analysis results of the developed system. The GUI as shown in Fig. 5 contains several navigation links for the given input and gives the step by step results from the data preprocessing up to opinion trend visualization during system computation.

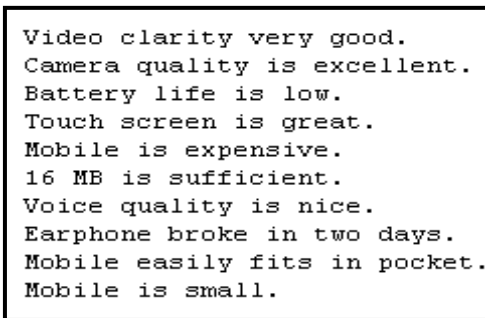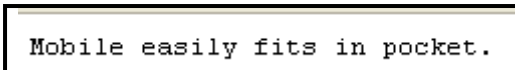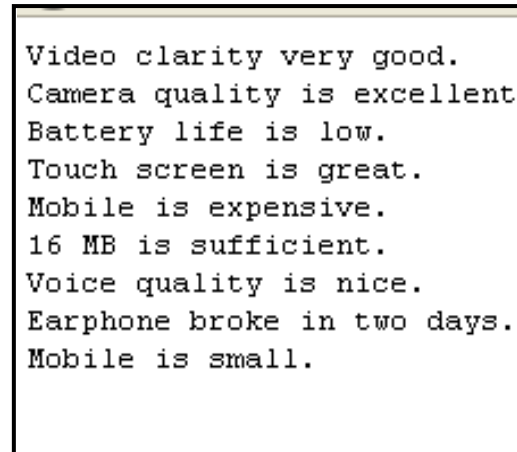**Fig. 5: Main GUI Console of Proposed System**

Fig. 6: Preprocessed File details for the Mobile Phone Review File

Fig. 6 shows the preprocessed review data for the next step of fact and sentiment comments separation as shown in the Fig.7.

(a) Fact Separation

(b) Comments Separation

Fig. 7: Fact & Sentiment Comment Separation

After this, feature extraction technique is applied to separate implicit and explicit features from the product review as given in the Fig. 8.
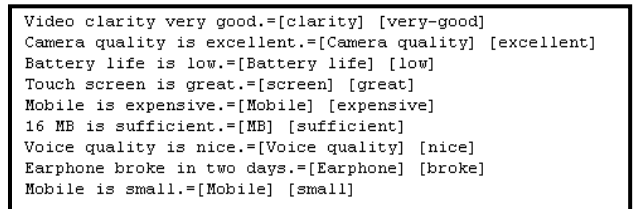

Fig. 8: Output(s) of Used Feature Extraction Technique

The algorithms were used to implement this process, namely, Fact Separation Algorithm, Implicit Feature Extraction Algorithm, and Explicit Feature Extraction Algorithm.
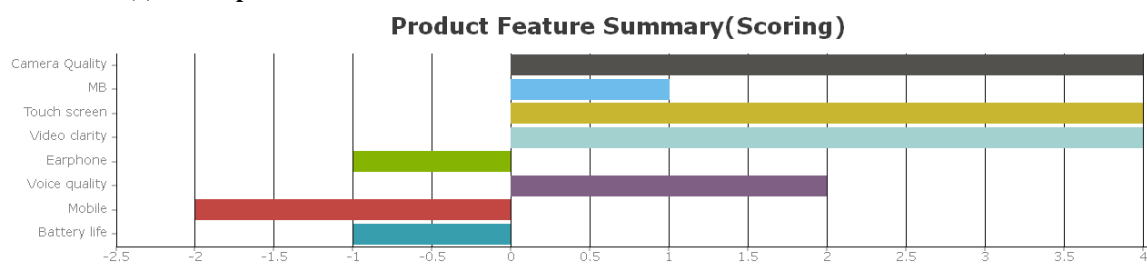

Fig. 9: Final Review Analysis based on Extracted Features for Mobile Phone Reviews

Based on the extracted features for the given review data, the developed system gives the finally the output of the system with the scoring value for each extracted feature of product is shown as like in Fig. 9.

5.3 Comparison with Other Opinion Mining Systems (OMSs)

Existing OMSs also provides the evaluation based on such complementary collection of review data set. To take advantage of such evaluation, the following two OMS are selected: Feature-Based Summarization (FBS), which was implemented using mainly on three perspectives of OMS i.e. effectiveness of feature extraction, effectiveness of opinion sentence extraction, and accuracy of orientation prediction of opinion sentences. Opinion Observer (OB), which was

intended to predict the semantic orientations of opinions on product features.

The main shortcomings observed during these OMSs performance evaluation against the proposed system on same review data set as input are enlisted in Table 2.

Table 2. Comparison of Proposed Work with Other OMSs

System	Human Tagger Used	Features Extracted	Automated
FBS	Yes	Explicit	No
OB	Yes	Explicit & Implicit	Semi-Automated
Proposed System	**No**	**Explicit & Implicit**	**Automated**

As shown in Table 2, use of human tagger slows down the sentiment analysis system where user demands quick analysis of opinions online as well as it is not possible for human being to perform same analysis for such huge volume of review data. Therefore, in proposed system, this task is implemented as part of system. User needs more accurate and trust worthy opinions while selecting their products, individuals or interested objects and hence, as stated earlier, to get more exact review outcomes, implicit and explicit review analysis is important. Therefore, proposed system has taken this issue in consideration and provided review results based on these features. Finally, in proposed system, whole computation is automatic and found very rare in existing opinion analysis systems for their complete execution. Hence, it speeds up the review analysis process and provides more trust worthy results in comparison with FBS and OB system.

6. CONCLUSION AND FUTURE SCOPE

In this paper, main shortcomings observed during existing OMS performance evaluation on same review data set as input of proposed system are identified. Use of human tagger in FBS and OB slows down the sentiment analysis system where user demands quick analysis of opinions online as well as it is not possible for human being to do it for huge volume of review data. Therefore, in proposed system, this task is implemented as part of system. User needs more accurate and trust worthy opinions while selecting their products, individuals or interested objects and hence, as stated earlier, to get more exact review outcomes, implicit and explicit reviews are important. Therefore, proposed system has taken this issue in consideration and provided review results based on these features which is found very rare in existing FBS and OB like opinion mining systems. Finally, in proposed system, whole computation is automatic and found very rare in existing opinion analysis systems for their complete execution. Hence, it speeds up the review analysis process and provides more trust worthy results in comparison with FBS and OB system. The proposed system automates the sentiment analysis by providing the single click solution to the business analysts with the computation power of computer machine. Also, it reveals the benefits of having such online opinion mining system which overcomes the drawback of existing opinion mining systems which are developed based on individual research directions as sentiment classification and feature-based opinion mining. It combines the advantages of these research directions and results produced by proposed system as integrated one are found promising & trustworthy. Since, it takes in account both types of sentiment features which are specified by the user as implicit as well as explicit which is observed very rare in earlier research works. This system can provide faster and efficient results, if it would be executed in parallel computing environment for review data processing and its analysis as part of future work.

7. REFERENCES

[1] M.A.Jawale, Dr.D.N.Kyatanavar, A.B.Pawar 2013 Development of Automated Sentiment or Opinion Discovery System: Review In Proc. of ICRTET 2013.

[2] M.A.Jawale, Dr.D.N.Kyatanavar, A.B.Pawar 2013 Design of Automated Sentiment or Opinion Discovery System to Enhance Its Performance In Proc. of Int. Conf. on Advances in Information Technology and Mobile Communication 2013and In Journal of ACEEE 2013.

[3] M.A.Jawale, Dr.D.N.Kyatanavar, A.B.Pawar 2014 Implementation of Automated Sentiment Discovery System In Proc. of IEEE ICRAIE'14.

[4] Chee Kian Leong, Yew Haur Lee, Wai Keong Mak 2012 Mining sentiments in SMS texts for teaching evaluation In Expert Systems with Applications pp. 2584–2589.

[5] Chunxia Yin, Qinke Peng 2009 Sentiment Analysis for Product Features in Chinese Reviews Based on Semantic Association In International Conference on Artificial Intelligence and Computational Intelligence,pp.82-85.

[6] Huifeng Tang, Songbo Tan, Xueqi Cheng 2009 A survey on sentiment detection of reviews In Science Direct Expert Systems with Applications, pp. 10760–10773

[7] Bing Liu 2010 Sentiment Analysis: A Multi-Faceted Problem In IEEE Intelligent Systems, pp.1-5.

[8] Magdalini Eirinaki, Shamita Pisal , Japinder Singh 2012 Feature-based opinion mining and ranking In Journal of Computer and System Sciences,pp.1175–1184

[9] Mita K. Dalal,Mukesh A. Zave 2011 Automatic Text Classification: A Technical Review In International Journal of Computer Applications (0975 – 8887) Volume 28– No.2, pp.37-40.

[10] http://nishithsblog.wordpress.com/2013/09/27/social-media-facts-and-statistics-for-2013/

[11] http://searchdl.org/public/book_series/LSCS/2/79.pdf

[12] http://www.academia.edu/6723240/Mining_Opinion_Features_in_Customer_Reviews.