Fundamentals of Sentiment Analysis: Concepts and Methodology

A.B. Pawar, M.A. Jawale and D.N. Kyatanavar

Abstract Internet has opened the new doors for information exchange and the growth of social media has created unprecedented opportunities for citizens to publicly raise their opinions, but it has serious bottlenecks when it comes to do analysis of these opinions. Even urgency to gain a real time understanding of citizens concerns has grown very rapidly. Since, the viral nature of social media which is fast and distributed one, some issues get rapidly distributed and unpredictably become important through this word of mouth opinions expressed online which in turn has known as sentiments of the users. The decision makers and people do not yet realized to make sense of this mass communication and interact sensibly with thousands of others with the help of sentiment analysis. To understand thoroughly use of sentiment analysis in today's business world, this chapter covers the brief about sentiment analysis including introduction of sentiment analysis, early history of sentiment analysis, problems of sentiment analysis, basic concepts of sentiment analysis with mathematical treatment, sentiment and subjectivity classification comprises of opinion mining and summarization, past scenarios of opinion or sentiment collection and their analysis. Methodologies like Sentiment Analysis as Text Classification Problem, Sentiment analysis as Feature Classification with mathematical treatment are explored. Also, Economic consequences of sentiment analysis on individual, society and organization with the help of social media sentiment analysis are provided as supporting component.

Keywords Feature extraction · Sentiment analysis · Opinion mining

A.B. Pawar (✉)
Department of Computer Engineering, S.R.E.S', College of Engineering,
Kopargaon 423 603, India
e-mail: anil.pawar1983@gmail.com

M.A. Jawale
Department of Information Technology, S.R.E.S', College of Engineering,
Kopargaon 423 603, India
e-mail: jawale.madhu@gmail.com

D.N. Kyatanavar
Savitribai Phule Pune University, Pune, Maharashtra, India
e-mail: kyatanavar@gmail.com

1 Introduction

Opinion mining also termed as sentiment analysis is the mining of opinions of individuals, their appraisals, and feelings in the direction of certain objects, facts and their attributes. As stated by Liu [10, 11] and Jawale et al. [6], in the latest years, opinion mining has attracted great deal of concentration from both the academicians and industry persons because of various challenging research issues and support of sentiment analysis for a broad set of applications. Opinions play a very important role in making a proper decision. As it is wise to get or listen to the opinions from other people while we make a choice. This scenario is not only true in the case of individual choice but today it is useful and right for organizations also. Very little computational study was carried out on opinions prior to the introduction of World Wide Web (WWW) due to limited availability of opinionated text for such analysis. In earlier days, when the person used to go for taking a decision, she usually used to ask for opinion from different sources either friends or relatives.

As stated by Liu [10] and Jawale et al. [7, 8] when an organization is in need of opinions about their product or service from customer or general public, they generally conduct surveys or opinion polls from a group. Because of increase in the contents on the WWW through social media, the world has changed and became wealthy in data through advancement of Web. Currently people can put their reviews about products on respective business organizational sites and can express their opinions on nearly everything in various blogs and discussion forums. If anyone wants to purchase a product, there is no need to ask about the product to someone from friends and family. As many user reviews are easily available on the Web. So for the industries perhaps it is not required to conduct surveys to get customer opinions about their product or service or about their competitors because ample of information about the same is publicly available.

As stated by [8], due to growth of internet and data contents on Web, searching the sites where opinions are available and then monitoring such sites on the internet is quite an intensive job because the opinion contents are available on different sites, and in turn every site may also have large amount of opinions. It is possible that, in most of the cases, opinions or judgments about particular thing are not directly expressed. So it becomes difficult for a user to identify such opinion related sites, read and extract opinion related contents, analyze such contents, get useful summary out of it, and use this summary to form an opinion. 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.

It is observed in Liu [11] that, the opinion contents which are available online on internet as well as off line are containing mostly textual information used by the customer to provide relevant product feedback. The information available in textual format can be generally classified as either facts or opinions. An objective expression about entities, objects or events and their attributes is known as facts. On the other hand, opinion is a subjective expression that describes sentiments of an individual,

assessment of performance or emotions about entities, things or events and their attributes.

Opinion is a broad concept. Majority of the current research study from Liu [9, 11, 14] based on opinion mining concentrated on textual information processing. In addition, it pays attention on opinion mining and retrieval of factual information provided and expressed in opinionated text. It also supports information retrieval, text classification, text mining and natural language processing. Processing of opinions was very little focused and studied concept to get exact opinions of the customer than factual information in recent research of sentiment analysis.

Following section introduces the problem of sentiment analysis and its challenges in bringing automation in opinion mining system.

Problems with Sentiment Analysis

Research in the field of sentiment analysis began with the study of the problem of subjectivity classification and sentiment classification, which mainly considered the sentiment analysis problem as a text classification problem. The subjectivity classification is the field which identifies whether a given text document contains factual information or opinionated information. Then the sentiment classification is responsible for categorizing an opinion into either positive opinion or negative opinion from the set of opinionated documents. In reality, it is required to have in-depth analysis of these opinionated documents because the user is interested to know what opinion have been expressed on certain product, service or on individual. From the survey of a product, one wants to know what features of the product have been praised or criticized by user of the product.

Let us consider an example of opinion expressed by a user on fruit Guava:

"(1) I purchased 2 kg of Guava day before yesterday. (2) They appeared very fresh when I bought them home. (3) The shape of each Guava was very appropriate. (4) The color of some of the pieces was bright. (5) I prepared a tasty salad of it. (6) According to my mother I bought them for higher rate."

In this context, main issue is exactly what we want to extract from this review? We may notice that there are several opinions in this provided input review. In this one, sentences (2), (3) and (4) express three positive opinions on fruit-Guava, while sentences (5) is expressing a fact. Sentence (6) expresses negative opinion about the same. It is also observed that, all the opinions presented by user here have some targets on which the opinions are expressed or given. In this sense, the opinion in sentence (2) is on the appearance of the fruit while the opinions expressed in sentences (3) and (4) are on the features such as 'shape' and 'color' respectively. The opinion given in sentence (6) is expressed on the cost of the fruit, but the opinion expressed in sentence (5) is on consumer herself i.e. on 'me', not on the fruit. Finally, it is also noticed that the source or holder of opinions is also part of this opinionated text. The source or holder of the opinions in sentences (2), (3) (4) and (5) is the author of the expressed review, where as in sentence (6) it is 'my mother'. With the help of this example, one can realize that challenges in sentiment analysis or opinion mining are part of intensive computing and require huge data analysis to get accurately analyzed business opinion.

Before proceeding further, this chapter highlights various basic concepts which are related with opinion mining field, like:

What is an opinion or sentiment?

What is an emotion?

What is an object?

What is a feature?

What are the types of opinions?

What is an opinion orientation?

Without the knowledge of these factors, it would be nearly impossible for any business organization decision maker to understand the trend of customer choice, individual expectations and decisions for improvement measures to be carried out in organization product development procedure. This chapter also highlights the growth of the product contents on Web and their diversified formats of representations, the language contents used by the consumers while expressing the reviews on the product and challenges in sentiment analysis. The need of automated sentiment or opinion mining system has been illustrated. The consideration of sentiment analysis as opinion classification problem rather than text classification is also brought out. This chapter also gives the details about motivation, issues in sentiment analysis as well as scope and objective of the proposed research work, contributions of the proposed research work and detailed outline of the report of research work.

2 Sentiment Analysis: Basic Concepts

In sentiment analysis, the opinionated text is important for decision making based on its analysis. So while collecting the opinionated text and treating it as an input for the opinion mining systems, one has to understand the basic terminologies associated with the opinion mining. The following section highlights the basic concepts related to this opinionated text and its further processing for decision making process.

As stated by Liu [10], information available in textual format can be classified into two main things: Facts and Opinions.

An objective expression made by user regarding certain objects, entities or events and their attributes is known as facts. In the similar way, a subjective expression which describes emotions of a person, her sentiments and performance assessment about objects, entities and events and their characteristics is known as opinion. Generally, opinions can be expressed on anything in this world including a product, person, business industries, or a topic. It indicates that mostly opinions are expressed on target entities which are having their own components and attributes. So in opinion mining, an object can be divided into the hierarchical levels based on the part-of relation. So, exactly object can be defined as:

Object

An object is mainly any entity which can be anything in real world i.e. person, organization, event, product, topic etc. Liu [11]. Let, *Obj* represents object here. So,

it can be represented by a pair, *Obj*(*L*, *A*), where L represents hierarchy of levels of components and different sub levels of components. A indicates attribute set of *Obj*. Similarly, every component in the set can have the sub-component set and their respective attribute set. Following example explained below illustrates the definition of object*Obj*, levels of components *L*, and attributes *A*.

Consider phone as general class. So a particular brand of phone can be considered as object. Then this phone has a set of components like battery, speaker and screen, and also its attributes will be quality of voice, slimness of the phone i.e. size and heaviness of the phone. Then component named battery also has separate set of attributes, e.g., battery time, and battery dimension.

So with the help of this definition, an object forms a tree or hierarchy. So phone object can be represented as tree or hierarchy shown in Fig. 1.

Here, the root of the tree represents the object itself (In Fig. 1. it is 'Phone'). After this, every node which is not a root represents the component (for this example, the components of Phone object are Battery, Speaker, Screen) or sub-component of the object (in this case sub-components are represented in {...} brackets i.e. for Battery component, the sub-components are {Battery Life}, {Battery Size}, for Speaker component, the sub-components are {Type}, {Volume Level}, and for Screen component, the sub-components are {Screen Size}, {Screen Color}) and every connection i.e. link represents the available part-of relation. Additionally, a set of attributes defined for that node are in connection with the node. It is important because an opinion is given for any node and on attribute of any node which is associated with the object.

While expressing the opinion, one can comment on object i.e. the phone, which is the root in Fig. 1. These opinions may be like "I don't like this phone", or one can express opinion on one of its attributes as given in Fig. 1 like "The quality of voice of this phone is poor". Similarly, opinions can be expressed on any component of the Phone or any property of that component as shown in Fig. 1.

In theory context, it is sufficient to define an opinionated text object as described earlier, but in practice to understand this opinion concept clearly, we can simplify this definition. The question can be raised here about this simplification in object definition. Basically, there are two main reasons behind it which are illustrated as follows:

Reason 1: Opinions are generally expressed by user in the natural language and this natural language processing is really difficult task as it requires extensive

Fig. 1 Tree structure of an object in opinion mining
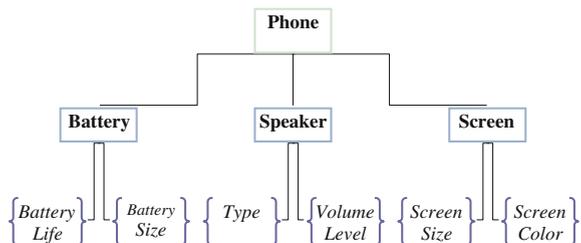
computational text processing. So, to do this text processing, defining an object is extremely challenging.

Reason 2: As defined earlier, object in opinionated text is represented by using hierarchical representation of an object. For normal users; it may become complicated to utilize this form of representation of an object while expressing the opinion about an object. So, instead of this tree kind of hierarchical representation of the object, we will use the concept feature to represent the components and attributes of opinionated text object.

Simplification of Object into Feature Form

Object itself can acts as a feature and becomes a special feature, which is represented as the root in the tree representation.

So, the opinion expressed on this object is then known as general opinion on the object. For this example as shown in Fig. 1 the general opinion statement can be "I like Nokia Phone". When the opinion is expressed on specific feature of the object, then it is known as specific opinion which is given on some feature of that object. For Phone object, specific opinion can be expressed as like [11] "Nokia Phone is having a very good touch screen", here Nokia Phone has a feature 'touch screen'. In real world, people are often habitual with product feature term than using feature term for an object. Even, at the point the objects are considered as topics or events, the feature term may not sound common and simpler one. Additionally, it is observed that researchers also used the terms either topic or aspect to get meaning as feature.

In this research work, we have chosen the term feature along with the term object. Since, both of these terms are required because in majority of the applications, the essential concern of the client or user is a set having interesting objects i.e. a set of products which are related to each other. After this, user is required to know about each feature which are expressing about an opinion document belonging to which object. This feature term is somewhat confusing if one considers the field of machine learning, where a feature indicates data characteristic or attribute. So to evade this conflict, the term which will be used further in this thesis is object feature to indicate feature of an object. To elaborate it in clearer manner, the following mathematical notations will help in this context.

Let, an object O has feature f in an opinionated text. Let d be a document, containing reviews about the products. Now, this opinionated document is collection of expressed opinion by user in statements. So, document d can be described with the help of series of sentences.

$$d = <s_1, s_2, \ldots, s_m>$$

Opinion Passage

An opinion passage is generally expressed on an object O about its feature f in document d which is a collection of sentences. These sentences can be positive opinions or negative opinions on feature f.

From this, it is understood that opinion can be expressed as collection of statements in document d on features f of an object O. Even possibility is that a single sentence may consist of opinions provided on multiple features of an object. The same is given in following sentence.

"The color of this guava fruit is nice but it does not taste that good".

This statement expresses opinion on features i.e. 'color' and 'taste' of an object 'fruit'. In recent studies, it is observed that current research focuses on sentences in opinionated text.

Explicit and Implicit Feature

When user expresses the opinion on any object by considering the features of an object, then one has to carefully categorize these opinions based on whether these are expressed explicitly or implicitly by using features of the object. So based on this expression, such features are called as either explicit or implicit features. The definition of these features is given below:

Explicit feature in sentence s of any document d is defined as follows: The feature f is called explicit feature if the feature or its alternatives is present in a sentence s. For example, in the next statement 'color' is an explicit feature expressed in opinionated text.

"The color of this Guava fruit is pleasant".

Similarly, in the following statement, the opinion is expressed on the 'size' feature of the fruit, but it is not appeared in the sentence directly but it is implied, so it is an implicit feature. "This fruit is big". In this sentence, big is not synonym of feature 'size', but it indicates opinion towards this feature, so such indications are known as a feature indicator. Mainly adjectives and adverbs are used as feature indicators. Some of the adverbs and adjectives are very commonly used like nice, good, bad or worst. Apart from these, other adjectives adverbs may be used for indicating specific feature, like beautiful can be used to indicate appearance, etc. These kinds of feature indicators can be mapped directly with their respective features.

Opinion Holder and or Opinion Sources

Mainly opinions on certain objects are expressed by users. Users may be individual person, group, and organization. It means that these users are authors of the opinions. In the field of sentiment analysis, such users are known as holder of an opinion. These holders of opinion are also known as opinion sources. Whenever feedback or review about particular thing is published or a blog is written, the opinion is given by the author of that review or blog. Additionally opinion holders are considered vital in the news articles as they provide explicit view about a situation, place, organization or person which is published in news articles. To understand it, consider the sentence, "Sarita expressed her disagreement on the purchased car." The opinion holder in this sentence is 'Sarita' since she is the opinion source in this sentence as 'Sarita' is mentioned explicitly in this sentence.

Opinion and Opinion Orientation

With the help of all above concepts related to opinion in sentiment analysis, opinion can be defined as, a positive or negative view, appraisal, emotion or feeling about object O with the help of feature f given by an opinion holder. These views on opinion lead us towards the next concept i.e. opinion orientation. Opinion orientation decides orientation of the expressed opinion on an object O for a feature f into positive, negative or neutral form. The term opinion orientation is alternatively used in many research papers as 'sentiment orientation', 'polarity of opinion', or even 'semantic orientation'.

Based on these concepts, the opinion mining system can be expressed in theory with the help of Object Model for opinion mining. It is described in the next section in depth.

Object Model

The object model can be derived as follows. Let, object O is a set of finite features as given below,

$$F = \{f_1, f_2, \ldots, f_n\}$$

This feature set F may include the object as a special feature. As described earlier, these features are the attributes of object. So, each feature $f_i \in F$ can be represented as a finite set of words or phrases where,

$$W_i = \{W_{i,1}, W_{i,2}, \ldots, W_{im}\}$$

Which are the feature synonyms, and finite set of feature indicators is expressed as of the feature.

Opinionated Document Model

A general document d containing opinion on set of objects as $\{Obj_1, Obj_2, \ldots, Obj_n\}$ from set of opinion holders represented as $\{OH_1, OH_2, \ldots, OH_n\}$. Now, these opinions expressed on each object Obj_1 are a subset F_i of features.

There are mainly two types of opinions, called as direct opinion and comparative opinion which are defined as below.

Direct Opinion

A direct opinion is defined in the form of quintuple $(Obj_i, f_{ik}, Obj_{ijk1}, h_i, t_l)$, where an object is represented as Obj_i, the feature is denoted as f_{ik} of an object obj_i, and the orientation or polarity of the opinion on feature f_{ik} of object obj_i is denoted as obj_{ijkl}, opinion holder is represented as h_i and t_l denotes the time when an opinion is commented by opinion holder h_i. The orientation of an opinion obj_{ijkl} can be identified as positive, negative or neutral. For feature, f_{ik}, the opinion holder h_i selects words or sentences from the corresponding alternative W_{ik}, or a sentence from the feature indicator set as I_{ik} to explain the feature, and then it is articulated as a positive, negative or neutral opinion on the feature in the opinionated text.

More specifically, the direct opinions are divided into two sub-types based on direct opinion expression on an object and on its effect on some other object. The first case considers opinions which are directly expressed on an object or its features; the next example clarifies this concept. "The color of Guava fruit is really fresh." Here, the object is Guava fruit and its directly stated feature in this sentence is 'color'. In the latter case, opinions expressed for an object are commented based on the effect of an object on some other objects. It is very easily adopted and frequently used in the medical field where patients expresses their opinions on drugs or describe drug side effects very frequently. If patient expresses opinion like "After taking this medicine, my leg pain reduced." Then it describes an expected effect of the specified drug on the leg, and so it implies a positive opinion about that medicine. For simplicity, we are considering these both sub types under the direct opinion category in the further discussion.

Comparative Opinion

A comparative opinion as the name itself explores that it is generally expressed on two or more objects based on their similarities or differences, and the object preferences are given by the opinion holder based on some of the shared features or attributes of these objects. Comparative opinion is described with the help of comparative or superlative form of an adverb or adjective, but it is not true for all comparative opinions.

Strength of Opinion

Strength of an opinion is mainly defined by opinion orientation obj_{ijkl} as described in quintuple of opinion. Each opinion comes in different strengths where few are very strong and few are weak. The strong opinion may be like opinion expressed in the following statement "This phone is rubbish" and weak opinion may be expressed like, "As per my thinking this phone is nice". So, it is possible to determine power of opinions. A feeling of contentedness, happiness, joyfulness, or delightedness is expressed with the help of positive opinion. In practice, as per the application, one can select scale to measure the strength of opinion. The more depth of this discussion indirectly leads towards the emotions.

Emotions

Emotions are human beings subjective feelings and thoughts expressed by them. The significant study has been done on the emotions and studied in many fields like biology, social science, psychology, philosophy, etc. Still it is found that researchers are not agreed on basic set of emotions. Mainly, there are six types of emotions, namely joy, love, surprise, anger, fear and sadness. These emotions further can be classified into their own subtypes.

Though emotions are influencing the strength of opinion and are exhibiting close relationship with opinion, there is no equivalence between emotion and opinion. So when the impact of emotions or opinions is discussed, it is important to differentiate between these two things i.e. feelings and language expressions used to describe feelings. As stated earlier, there are only six types of primary emotions; but on other

side there are many expressions of the language that can be used to convey emotions. In the same sense, there are also many opinion expressions used to explain positive sentiments or negative sentiments by users. Therefore, mainly sentiment analysis attempts to conclude individual sentiments based on expressions provided with the help of any language to portray their emotions.

Extensive Purpose of Sentiment Analysis

The main purpose of sentiment analysis, aims to conclude positive sentiments or negative sentiments from given opinionated text. But this objective can be extended to discover the additional information which is significant for use of opinions in the practical decision making process. So the objectives of direct opinion mining can be illustrated as: In given opinionated document d, firstly, identify the quintuple of opinion i.e. (obj_i, f_{ik}, obj_{ijkl}, h_i, t_l) and secondly, find out all the feature indicators I_{ik} and synonyms (W_{ik}) of each feature f_{ik} in d.

Issues present in the Feature based Opinion Mining

Here we will discuss the different issues related to feature based opinion mining [2].

Issue 1

In the feature based opinion mining, each time it is expected that the quintuple containing five pieces of information about an opinion should be related to each other. It indicates that, the opinion have to be expressed on feature of object at specified time by opinion holder. It can be concluded from this condition that, sentiment analysis is really challenging problem since in opinionated text; even discovering each piece of information itself is very complicated. Even it becomes worse when in opinionated text, sentence does not explicitly mention some pieces of information required in quintuple, and may be stated with the help of pronouns, language expressions, and the context of opinionated text.

The following blog example highlights this issue with the help of further illustration. Each statement is identified with the number for simplicity and understanding in blog opinionated text.

"(1) Last Monday, I purchased a 'Samsung' phone and my granny got her gift as an 'Asus' phone. (2) She called me to inform me the same. (3) I could not listen to her as voice of her phone was not clearly audible. (4) The picture taken from camera was fine. (5) I am really happy and comfortable with my phone. (6) I got the phone as per my expectations with good voice. (7) So the phone was really displeasure for us. (8) My granny was not at all satisfied with her gift and sent it back to the retailer day before yesterday."

The identified objects in this given blog contain opinionated text describing 'Samsung phone' and 'Asus phone', which is really difficult to identify in practical way. To get an idea about what does 'my phone' indicate and what does 'her phone' shows in sentence (3) and sentence (5) is actually difficult. Sentence (4) is really conflicting one as it is not having mention of either 'Samsung phone' or 'Asus phone' and it also does not contain a pronoun. Then the question raised here is that 'the camera' belongs to which phone. Sentence (6) directly gives a positive opinion about a phone and quality of its voice. In sentences (7) and (8), it is very difficult to get information

about which 'phone' it is and what 'the gift' is. Here, the opinion holder of all the opinions is the author himself of the opinionated blog text except in sentence (8) the opinion holder is 'my granny' instead of author of the blog.

Issue 2

It is not needed to identify the quintuple with all five pieces of information every time for each application in practice. Since some time some of them may be unknown and unwanted. In case of reviews about a product, the reviewed object i.e. product is assessed in each review. When the review is written, the author of the review always knows that a review site typically records and displays review information. Though this information is not required every time, one needs to take out such review information from the Web pages of review sites because of structured data extraction issue.

The example stated earlier in issue1 additionally revealed another issue, known as, subjectivity. It states that, in any document which is having collection of opinions, some sentences always give opinions where as some does not. As seen in earlier example, sentences (1), (2) and (8) do not comment on any opinions. It gives the clue for sentence subjectivity concept which is defined and described in the next section.

Sentence Subjectivity

Generally, factual information about the world is always represented with the help of objective sentences, while subjective sentences are used to express personal emotions or beliefs. As example given in issue1, sentences (1), (2) and (8) are categorized as objective sentences, whereas all remaining sentences are called as subjective sentences. Many forms of subjective expressions are like desires, opinions, speculations, beliefs, suspicions, and allegations. Therefore, a subjective sentence may or may not contain an opinion every time. For example, the following sentence indicates the sentence subjectivity but it does not state either positive opinion or negative opinion on any precise smart phone object. "I expected a smart phone with good reception signal power." In the same way, not every objective sentence contains any opinion.

Explicit Opinion and Implicit Opinion

As discussed earlier a feature which is specified explicitly and a feature which is implied in the sentence, an explicit opinion is defined as an opinion which is expressed explicitly in a sentence on feature f in a subjective sentence of the opinionated text [2]. Similarly, an implicit opinion is defined as an opinion on feature f which is implied in an objective sentence of the opinionated text. The sentence as presented below is an example of an explicit positive opinion in subjective sentence. "Clarity of this TV is excellent." and in the similar way the next sentence expresses an implicit negative opinion in an objective sentence. "The earphone failed working in two days." It is noted that, this sentence presents a fact with the help of objective sentence; but still it is a negative opinion on an 'earphone' which is indicated implicitly.

Opinionated Statement

An opinionated statement is defined as a statement that explicitly expresses positive or negative opinion or implicitly implies positive or negative opinions. This

opinionated sentence may be either of these two categories viz., subjective sentence or objective sentence. One needs to understand that subjective sentences and opinionated sentences are two different concepts, though these sentences which are opinionated are most of the time the subset of subjective sentences. The ways with which opinionated and subjective sentences are determined are also similar. In this research work to avoid the reader confusion we will use these terms alternative to each other. The function that is used for distinguishing subjective or objective sentences is known as subjectivity classification. The next section describes in depth about this concept.

Sentiment and Subjectivity Classification

The term sentiment classification is used to classify an opinionated document based on whether it expresses the positive opinion or negative opinion. This is also called as the sentiment classification done at document level because it takes the complete document as the basic information unit used as an input for performing opinion mining [10]. The more details about sentiment classification at document level are illustrated in the following section. The most of the existing researcher done the study based on the document is known to be opinionated as the basic information unit. However, this sentiment analysis can be applied to the individual sentences of the document too. But one has to note that, we cannot assume each sentence to be opinionated as part of the opinionated document. The function that classifies a given sentence into either opinionated sentence or non-opinionated sentence is known as subjectivity classification. Even the resultant sentences which are opinionated are also devised into positive sentence or negative sentence, which is further known as sentence-level sentiment classification [1, 13].

Sentiment Classification at Document Level

Document level sentiment classification is used to determine whether the document d is expressing a positive or negative opinion when given a set of opinionated documents D and each document $d \in D$.

More specifically, sentiment classification done at the document level can be illustrated better in the following way. Consider, a document d containing opinionated sentences which provides an opinion on object O, then identify the orientation or polarity of the opinion expressed on an object O, i.e. determine what is the orientation of that opinion on feature f. It can be represented in the quintuple (o, f, s, h, t), where $f = O$ and h, t, O are assumed to be known or not relevant one.

During this study it is observed that the traditional research work done on classification of sentiments is based on the theory that the document d which is opinionated contains usually the product reviews on an object O and these opinions are usually expressed by only one opinion holder h. Such kind of assumption can be useful for user reviews on products and services. But it cannot be applicable for the opinionated text generated by blog posts and a forum and since in these posts, the opinion holder may provide her opinions on multiple targets and may compare those opinions using superlative or comparative sentences.

Additionally, sentiment classification done at document-level can be carried out using supervised learning and unsupervised learning methods. These methods are described below.

Sentiment Classification at Document Level Based on Supervised Learning

Sentiment classification can be done as a supervised learning problem with two class labels which are positive and negative. In existing research studies, the training and testing data used are mostly reviews about a product of certain organization.

Sentiment classification is similar with classic topic-based text classification. But distinction can be made between two as sentiment classification classifies the documents into either positive or negative whereas text classification using classic topic based methods classifies documents into already defined topic classes like movies, sports, food, education, etc. Additionally, in classification using topic-based methods, the main emphasis is given on topic related words. On the other hand, in sentiment classification importance is not given to topic-related words during the classification. Instead, it puts emphasis on identifying expressed sentiment, feeling, emotion or opinion words which expresses either positive opinions or negative opinions like good, bad, outstanding, nice, etc.

It is found that present supervised learning classification techniques can be directly used to perform sentiment classification; such techniques are Naïve Bayesian Classification Algorithms, Nearest Neighbor Algorithm, Decision Tree Classifier, Artificial Neural Networks and Support Vector Machines, etc. It is proved earlier that features in classification are considered using unigrams (also called as a bag of individual words) have shown the expected results with supervised learning techniques either Naïve Bayesian Algorithm or Support Vector Machine. The main thing to be clarified here is about use of features. Here, the features are considered as data attributes as described in machine learning, instead of object features which are mentioned in earlier section. Following are the examples of such features used in existing research work [1]:

Term (T) and Term Frequency (TF)

The feature considered as individual word or word n-grams is called as term and its occurrence count in the document is known as term frequency. During the classification word positions may be another issue. So from the field of information retrieval, the feature Term Frequency—Inverse Document Frequency weighting scheme can be applied. These features are common part of traditional topic-based text classification, but it is found that these are quite effective and useful one in sentiment classification also.

Part of Speech Tags (POS)

It is the method used to assign a Part-of-Speech to every word present in the sentence. Every word in the sentence is assigned a tag like, verb, noun, prepositions and adjective etc. In English language, mainly adjectives are used to identify subjectivity and opinions. So, the earlier researchers used these adjectives as significant indicators of either subjectivities or opinions and counted these adjectives as the special features in the field of opinion mining.

Opinion Words and Opinion Phrases

In the opinionated text positive or negative sentiments which are commonly used to express emotions of the opinion holder are called as opinion words. We can consider the beautiful, good, amazing, etc. as positive opinion words and negative opinion words like, bad, weak, poor, etc. Additionally, as stated earlier, adjectives, adverbs, nouns like rubbish, junk, and crap and even verbs like hate, like, etc. are used as opinion words.

Moreover, instead of such individual opinion words, there are also idioms and phrases which can be used to indicate opinions. Consider an example, "These opera tickets cost us an arm and a leg". Here, 'cost someone an arm and a leg' is a phrase which means having a negative impact of something. Therefore, opinion phrases and opinion words play a vital role in performing sentiment analysis.

Syntactic Dependency

It is defined as words dependency which is based on features generated from dependency trees or parsing. Several researchers can use it in their research in the field of sentiment analysis.

Negation

This term is vital in sentiment analysis since its appearance normally change the opinion orientation results. In the next example, the sentence "I am not happy with this camera" is negative. Handling of negation words requires careful analysis because not all occurrences of such words are used to indicate negation like in the sentence 'not' in 'not only ... but also' does not change the orientation status.

Sentiment Classification at Document Level based on Unsupervised Learning.

In sentiment classification opinion phrases and opinion words are the most important indicators. Hence use of unsupervised learning methods for sentiment analysis which is based on such opinion phrases and opinion words can be obvious in the research of sentiment analysis. Mainly, these methods perform classification by using fixed syntactic phrases that are mostly used by the user to state the opinions. Generally, these algorithms consist of the following steps.

Step 1

Initially, from the given document, it does the extraction of phrases containing adverbs or adjectives. Because as subjectivity and opinions can be indicated from adverbs and adjectives as described earlier. However, it is observed that subjectivity is indicated with the help of adjective, but there may be insufficient relative information to state or decide its opinion orientation results based on such isolated adjectives. So, these algorithms extract two successive words, out of these two one member represents either an adjective or adverb and the other gives context dependent word. If the Part-of-Speech tag of these two successive words conform to any of the patterns as shown in Table 1 then only these words will be extracted. As shown in line 2 of Table 1 two successive words are extracted if there is match for the pattern as like first word is an adverb and the second word is an adjective, but the third word cannot be a noun.

Table 1 POS patterns for two words extraction

First word		Second word	Third word (Not Extracted)
1.	JJ	NN or NN	Anything
2.	RB, RBR, or RBS	JJ	not NN nor NNS
3.	JJ	JJ	not NN nor NNS
4.	NN or NNS	JJ	not NN nor NNS
5.	RB, RBR, or RBS	VB, VBD, VBN, or VBG	Anything

3 Sentiment and Subjectivity Classification

The term sentiment classification is used to classify an opinionated document based on whether it expresses the positive opinion or negative opinion [10]. This is also called as the sentiment classification done at document level because it takes the complete document as the basic information unit used as an input for performing opinion mining. The more details about sentiment classification at document level are illustrated in the following section. The most of the existing researcher done the study based on the document is known to be opinionated as the basic information unit. However, this sentiment analysis can be applied to the individual sentences of the document too. But one has to note that, we cannot assume each sentence to be opinionated as part of the opinionated document. The function that classifies a given sentence into either opinionated sentence or non- opinionated sentence is known as subjectivity classification. Even the resultant sentences which are opinionated are also devised into positive sentence or negative sentence, which is further known as sentence-level sentiment classification [10, 13].

Sentiment Classification at Document Level

Document level sentiment classification is used to determine whether the document d is expressing a positive or negative opinion when given a set of opinionated documents D and each document $d \in D$.

More specifically, sentiment classification done at the document level can be illustrated better in the following way. Consider, a document d containing opinionated sentences which provides an opinion on object O, then identify the orientation or polarity of the opinion expressed on an object O, i.e. determine what is the orientation of that opinion on feature f. It can be represented in the quintuple (o, f, S, h, t), where $f = O$ and h, t, O are assumed to be known or not relevant one.

During this study it is observed that the traditional research work done on classification of sentiments is based on the theory that the document d which is opinionated contains usually the product reviews on an object O and these opinions are usually expressed by only one opinion holder h. Such kind of assumption can be useful for user reviews on products and services. But it cannot be applicable for the opinionated text generated by blog posts and a forum and since in these posts, the opinion holder

may provide her opinions on multiple targets and may compare those opinions using superlative or comparative sentences.

Additionally, sentiment classification done at document-level can be carried out using supervised learning and unsupervised learning methods. These methods are described below.

Sentiment Classification at Document Level Based on Supervised Learning

Sentiment classification can be done as a supervised learning problem with two class labels which are positive and negative. In existing research studies, the training and testing data used are mostly reviews about a product of certain organization.

Sentiment classification is similar with classic topic-based text classification. But distinction can be made between two as sentiment classification classifies the documents into either positive or negative whereas text classification using classic topic based methods classifies documents into already defined topic classes like movies, sports, food, education, etc. Additionally, in classification using topic-based methods, the main emphasis is given on topic related words. On the other hand, in sentiment classification importance is not given to topic-related words during the classification. Instead, it puts emphasis on identifying expressed sentiment, feeling, emotion or opinion words which expresses either positive opinions or negative opinions like good, bad, outstanding, nice, etc.

It is found that present supervised learning classification techniques can be directly used to perform sentiment classification; such techniques are Naïve Bayesian Classification Algorithms, Nearest Neighbor Algorithm, Decision Tree Classifier, Artificial Neural Networks and Support Vector Machines, etc. It is proved earlier that features in classification are considered using unigrams (also called as a bag of individual words) have shown the expected results with supervised learning techniques either Naïve Bayesian Algorithm or Support Vector Machine. The main thing to be clarified here is about use of features. Here, the features are considered as data attributes as described in machine learning, instead of object features which are mentioned in earlier section. Following are the examples of such features used in existing research work:

Term (T) and Term Frequency (TF)

The feature considered as individual word or word n-grams is called as term and its occurrence count in the document is known as term frequency. During the classification word positions may be another issue. So from the field of information retrieval, the feature Term Frequency- Inverse Document Frequency weighting scheme can be applied. These features are common part of traditional topic-based text classification, but it is found that these are quite effective and useful one in sentiment classification also.

Part of Speech Tags (POS)

It is the method used to assign a Part-of-Speech to every word present in the sentence. Every word in the sentence is assigned a tag like, verb, noun, prepositions and adjective etc. In English language, mainly adjectives are used to identify subjectivity

and opinions. So, the earlier researchers used these adjectives as significant indicators of either subjectivities or opinions and counted these adjectives as the special features in the field of opinion mining.

Opinion Words and Opinion Phrases

In the opinionated text positive or negative sentiments which are commonly used to express emotions of the opinion holder are called as opinion words. We can consider the beautiful, good, amazing, etc. as positive opinion words and negative opinion words like, bad, weak, poor, etc. Additionally, as stated earlier, adjectives, adverbs, nouns like rubbish, junk, and crap and even verbs like hate, like, etc. are used as opinion words.

Moreover, instead of such individual opinion words, there are also idioms and phrases which can be used to indicate opinions. Consider an example, "These opera tickets cost us an arm and a leg". Here, 'cost someone an arm and a leg' is a phrase which means having a negative impact of something. Therefore, opinion phrases and opinion words play a vital role in performing sentiment analysis.

Syntactic Dependency

It is defined as words dependency which is based on features generated from dependency trees or parsing. Several researchers can use it in their research in the field of sentiment analysis.

Negation

This term is vital in sentiment analysis since its appearance normally change the opinion orientation results. In the next example, the sentence "I am not happy with this camera" is negative. Handling of negation words requires careful analysis because not all occurrences of such words are used to indicate negation like in the sentence 'not' in 'not only ... but also' does not change the orientation status.

Sentiment Classification at Document Level based on Unsupervised Learning

In sentiment classification opinion phrases and opinion words are the most important indicators. Hence use of unsupervised learning methods for sentiment analysis which is based on such opinion phrases and opinion words can be obvious in the research of sentiment analysis. Mainly, these methods perform classification by using fixed syntactic phrases that are mostly used by the user to state the opinions. Generally, these algorithms consist of the following steps.

Step 1

Initially, from the given document, it does the extraction of phrases containing adverbs or adjectives. Because as subjectivity and opinions can be indicated from adverbs and adjectives as described earlier. However, it is observed that subjectivity is indicated with the help of adjective, but there may be insufficient relative information to state or decide its opinion orientation results based on such isolated adjectives. So, these algorithms extract two successive words, out of these two one member represents either an adjective or adverb and the other gives context dependent word. If the Part-of-Speech tag of these two successive words conform to any of the patterns as shown in Table 1 then only these words will be extracted. As shown in line 2 of

Table 1 two successive words are extracted if there is match for the pattern as like first word is an adverb and the second word is an adjective, but the third word cannot be a noun.

Step 2

The orientation estimation of the phrases which are extracted can be calculated with the measure of Point-wise Mutual Information (PMI) as given in equation:

$$PMI(t1, t2) = log_2 \left(\frac{Pr(t1 \wedge t2)}{\Pr(t1)\Pr(t2)} \right)$$

In this equation, the term $Pr(t1 \wedge t2)$ is the co-occurrence probability of t1 and t2, and another term $\Pr(t1)\Pr(t2)$ represents the probability of the occurrences of these two terms if and only if they are statistically not dependent on each other. Therefore, this degree is a measure of the level of statistical dependence between them and log of this proportion gives the measure of information that we are getting about the occurrence of one of the words when we see its connection with other word.

Now, based on this we can calculate the individual phrase opinion orientation based on its connection with the words like outstanding which is a positive reference and based on its connection with the words like unfortunate which is a negative reference. So the individual phrase opinion orientation can be calculated using the following equation.

$$OO(phrase) = PMI(phrase, \text{``excellent''}) - PMI(phrase, \text{``poor''})$$

The probabilities can be calculated by targeting the queries to any information retrieval tool and based on its response in terms of hits i.e. count. For every targeted request, information retrieval tool provides the response in terms of documents which are relevant to the request made by user, which represents the total count. So, it is easy to calculate the probability of the two terms together as well as separately as per the PMI calculation equation.

Step 3

Based on given collection of reviews, the algorithm can compute the average object orientation value of all the provided phrases and categorize the provided reviews as important and useful if average of this object orientation is positive; otherwise it is not recommended as important and useful.

Sentiment Classification at Sentence Level

Now for the similar task we can compute the sentence-level classification. Suppose the task is given as below. For a sentence s, perform the two important sub-tasks which are given below [1].

Perform Subjectivity Classification

In this one can find out sentence s is either an objective sentence or a subjective sentence as per the opinion expressed.

Perform Sentiment Classification at Sentence Level
In this, if sentence s is subjective sentence, then one can find out sentence s is either a positive opinion or negative opinion.

In this case, it is observed that the quintuple (o, f, OO, h, t) is not utilized for describing the task since the classification at sentence level is a middle step. Then again, the two sub-assignments of the classification done at sentence-level are still very significant because (1) it sort out those sentences which are not having any opinion, and (2) after the identification which are the objects and which are the features of that objects are expressed in these sentences, it provides help to decide that the provided opinions on the underlying objects and also on the features of that objects are either positive or negative. It is found that significant study has been done on these problems, but in separate directions as classification problem, so, we can easily apply the supervised learning methods for their solution. The work reported by earlier researchers was focused on subjectivity classification. Classifier used for doing subjectivity classification is the Naïve Bayesian Classifier.

The main drawback of such supervised learning is, it requires extensive manual efforts to do annotation of a big number of training examples. Another approach to reduce the manual classification attempt is a bootstrapping method which was proposed for automatically labeling the training data. In this case, the algorithm considers the uses of two classifiers with high precision namely, HP-Subject and HP-Object which can categorize automatically subjective sentences and objective sentences. These classifiers list out lexical tokens such as words and n-grams words which give clues for doing better subjectivity classification. The HP-Subject determines a subjective sentence if it indicates more than two well-built subjective clues and HP-Object identifies objective sentence if it does not show strongly subjective clues. The merit of these classifiers is that precision is very high but demerit is that its recall is low. These extracted sentences are then added to the training data in order to learn about the patterns. These patterns are then utilized to perform identification of other sentences which are either subjective sentences or objective sentences. These classified sentences are further added in the training data set, and algorithm continues with its next continues loop. For pattern learning, we can provide syntactic template set to limit the types of patterns to be learned by this classifier. Few examples of syntactic templates and their respective patterns as example are given in Table 2.
Assumptions made for Sentiment Classification at Sentence Level

Table 2 Syntactic template pattern with examples	Syntactic template pattern	Example
	<subject> passive-verb	<subject> was satisfied
	<subject> active-verb	<subject> complained
	Active-verb <dobject>	endorsed <dobject>
	noun aux <dobject>	fact is <dobject>
	passive-verb prep <np>	was worried about <np>

It is most of the time assumed that the sentence is provided by a single opinion holder to express a single opinion. This hypothesis is suitable only when a simple sentence expresses a single opinion. Consider the example, "Clarity of this TV is excellent." But in fact, compound sentences can provide one or more opinions. Consider the same example with opinion expressed as, "Quality and Clarity of this TV is excellent but its SMPS is poor." provides both the types of positive as well as negative opinions in mixed form. This sentence gives positive opinion for 'quality' and 'clarity', but expresses negative opinion for 'SMPS', even this whole compound sentence is positive for the camera. At the last, it should be kept in mind that subjective sentences are subset of opinionated sentences, and additionally, many objective sentences are also used to express the opinions. So, we need to identify opinions from given textual information with the consideration of subjective as well as objective sentences.

4 Opinion Mining and Summarization

The Internet has changed every individual's life to a great extent which was very sophisticated and simpler one previously. Generally, people are using the Web technology for their functionalities like information retrieval, to perform communication, to enjoy their relaxation and to have an entertainment, to done online shopping, etc. Even, Internet forums are also used as medium of information and online resource exchange. The commercial review [14] Web sites are used to express opinions on the products, individuals, and on organizations. These commercial review websites allow users to express their opinions in their own ways, so the collection of such reviews for specific product is available enormously. Hence it becomes challenge for the customers to observe all these reviews to make a decision on certain product choice and its selection.

So, with the help of opinion mining, an extraction technique can be developed to score the reviews and summarize the opinions to end user. From this, user can decide whether to recommend the product or not based on these opinions mined.

Past Scenario of Collecting Opinions and Their Analysis

During the decision making process we always take in account that "What is the opinion or thinking of other persons?" This type of view always had been a vital part of information for most of us during the decision making to get the idea or views of people in the decision. When we decide to make a decision we want to hear other persons for their opinions. This is applicable not only for individual's performance assessment but also for organizations. In recent days, when WWW has grown largely and easily available, many of us used to ask our friends to recommend a certain product or vote poll for local elections, reference for individual, etc. In past when an organization requires finding general public opinions about its products and related services, they mainly go for surveys and their center of attention is groups. As there is a tremendous growth of the data contents on the Web due to social media

in the last few years, the world has been changed completely. Internet and the Web made it very simpler to discover the opinions of users and share their knowledge and their feedback in the massive group of unknown people that are never been our individual good wishers or professional criticizer. Generally, these are people we have never heard of. Today, it is observed that many users are expressing their opinions on Internet without any constraints.

Due to this explosive growth of Web, finding such sites which have opinions and monitoring such sites on the Web became really challenging job. Since a large numbers of such opinion sites are available and each one structured differently, and each site may contain large amount of text which is opinionated. Most of the times, opinions are indirectly expressed by the user in their posts and blogs. It is not possible for the human expert to locate opinionated Web sites, then take out opinionated text, understand it, summarize it, and organize it into usable forms for the decision making. So, it demands the need of automated opinion discovery and summarization system which is the main theme of this research work.

Sentiment Analysis as Text Classification Problem

In sentiment analysis field research is done mostly by academicians. In this study, it is found that, sentiment analysis is treated as a problem of text classification. Additionally, two associate research areas have been comprehensively studied in this study which are: (1) classification of document which is opinionated based on either a document expresses a positive opinion or negative opinion, and (2) classification of a sentence either subjective sentence or objective sentence, and for the sentences which are subjective, again classify it on the basis of either it expresses a positive opinion, negative opinion or neutral opinion. First research area is universally known as sentiment classification at document level, which focuses on finding the broad feeling of the opinion holder in the given opinionated text. With the help of given product reviews; it helps to recognize and understand whether the opinion holder is positive or negative on the subject of that product. The second area concentrates on the sentences individually to identify whether this sentence is responsible for contributing any opinion or not which is known as subjectivity classification, and also it is known as sentence-level sentiment classification.

Sentiment Analysis as Feature Classification

During the study, it is observed that in most of the cases it is important to done the classification of the opinionated texts is either at the sentence level or at the document level. However, such classification is not giving the necessary details required for other different types of applications. In general, a positive opinion on a certain feature of an object in the opinionated document does not indicate that the person has positive opinions about all the features of that object. Similarly, a negative opinion on a certain feature of an object in the opinionated document does not indicate that the person has negative opinions about all the features of that object. Writer may put in writing both aspects of the object in an opinionated text as positive and negative. These types of details are not given by the classification at document-level and classification at sentence-level. So, for such information, there is a need to identify features of an

object that means there is a requirement of extraction of all five pieces of quintuple i.e. full model of an object. As declared previously, at the feature level, the mining task deals with discovery of every quintuple (obj_i, f_{ik}, obj_{ijkl}, h_i, t_l) and determine all the synonyms W_{ik} and feature indicator is I_{ik}.

Here, we basically perform two important mining tasks as given below:

(1) Determine all the features of an object that have been expressed. Consider the example sentence, "The taste of tea is very good." and here, the feature of an object is 'taste'.
(2) Identify the expressed opinions on the features of an object is a positive remark, negative remark or neutral one.

Here in the above example sentence, the opinion expressed on the feature 'taste' of a tea object is positive.

This model identifies the opinion sentences and opinion objects on which opinions are given in those sentences, and subsequently classifies the opinions as positive, negative or neutral kind of. Generally, the targeted objects are real world entity and their features are identified. Here, in opinion mining, an object can be manufactured goods, any types of service, human being, any institute, any firm, occasion, theme, etc. on which opinions are expressed. Consider the case of product, in review sentences of product, it finds features of an object which is product that has been expressed by the user and identifies the comments expressed are either positive type or negative type. In the sentence, for example, "The battery of this Laptop is giving less backup," here, the opinion is expressed on 'battery' of the considered object 'Laptop' and the expresses negative kind of opinion. This level of thorough investigation is the need of real existence applications during the decision making process. Since for the improvement of product one must be familiar with which are the components or which are the features of the product are positively recommended or not recommended by the users. Generally, this type of detailed investigation of information is not acknowledged by sentiment classification and subjectivity classification.

5   Economic Consequences of Sentiment Analysis—Impact on Individual, Society and Organization

Many business analysts of online reviews believe that these reviews considerably control their marketing strategies as well as user purchasing decisions. Today, users are always searching for and dependent upon online advice and recommendations, suggestions rather than asking particular individual about his or her choice. Due to the growth and easy availability of Internet, it is easily possible to discover about the opinions expressed and shared feedbacks of people from all the corners of the world regardless of their professions and expertise in the particular domain. And conversely, many users are putting their opinionated text available to strangers via the Internet and social media sites. The key findings about impact on individual, society and organizations of opinion mining are stated below as per the two surveys

carried out in America on 2000 American people in year 2008; details are reported by Horrigan [3].

Online investigation about a product minimum once is performing by 81 % of Internet users before purchasing it; 20 % Internet users perform such investigation on a usual day the majority of the time; During selecting the various daily services, most of the readers of online reviews of hotels, restaurants, and a variety of services like hospitals, among 73 % and likely 87 % users reported that reviews had a major influence while doing selection of such services and products; Consumers report says that users are easily ready to pay starting from 20 to 99 % additional for a 5-star-rated products than that of a 4-star-rated products based on reviews given by various users.Out of all users, 32 % users have given and commented an evaluation about a person, a product, or a service via an online evaluation system, and 30 % users counting 18 % of online senior general public have posted and expressed an online observation or comment about the service or product during their online purchase.

It is noticed that utilization of services and goods is not the merely inspiration and foundation behind citizens expressing opinions online. Even, there are so many factors involved in this posting of reviews online. Mostly, for political information, it is used extensively.

Following example illustrate this fact more clearly. Survey report of more than 2500 Americans reported by Lee and John [12] pointed that the 31 % of Americans were 2006 election campaign Internet users; 28 % users believed that a most important issue for these online actions was to reach and to know judgments and perspectives from within their society, and 34 % users said that a most important cause was to get judgments and perspectives from outside their society and to see its impact during election; 27 % users had stared online for the approval of outside organizations posted online; 28 % users declare that the majority of the sites they utilize to share their point of observation on certain public issues, but 29 % users said that most of the sites they use challenge their point of view against others views; 8 % users posted their own political expressed comments, opinions online without considering any public issues.

The mentioned above two report findings are just representative example that shows user expectations and dependency based on online recommendations and advice is major cause at the back the surge of attention in innovative systems that concerned in a straight way with opinions as a primary rank object and gives them opinion analysis as early as possible. This suggests an unambiguous necessity to assist customers by constructing improved information analysis systems as compared to those that are presently in way of life.

6   Conclusion and Future Scope

In this chapter, detailed and clear suggestions about findings and conclusions of opinion into the present and current research study are studied for analysis of an opinion. Current research studies major contribution is to distinguish either opinion is positive, negative or neutral on the basis of explicit opinions and its explicit feature recognitions only. Not provided a great deal of attentions on implicit opinion and

its implicit feature extraction and identifications. So the planned research studies center of attention will be on recognizing and identifying explicit as well as implicit opinions with the help of their explicit and implicit features. This sensible and large requirement in major applications and the technological challenges will stay the field lively and active.

References

1. Dalal, M.K., Zave, M.A.: Automatic text classification: a technical review. Int. J. Comput. Appl. (0975–8887) **28**(2), 37–40 (2011)
2. Eirinaki, M., Pisal, S., Singh, J.: Feature-based opinion mining and ranking. J. Comput. Syst. Sci. 1175–1184 (2012)
3. Horrigan, J.A.: Online Shopping (2008) http://www.pewinternet.org/2008/02/13/online-shopping/
4. http://nishithsblog.wordpress.com/2013/09/27/social-media-facts-and-statistics-for-2013/
5. http://www.academia.edu/6723240/Mining_Opinion_Features_in_Customer_Reviews
6. Jawale, M.A., Dr., Kyatanavar, D.N., Pawar, A.B.: Design of automated sentiment or opinion discovery system to enhance its performance. In: Proceedings of International Conference on Advances in Information Technology and Mobile Communication 2013 and In: Journal of ACEEE 2013, pp. 48–53 (2013) http://searchdl.org/public/book_series/LSCS/2/79.pdf
7. Jawale, M.A., Dr., Kyatanavar, D.N., Pawar, A.B.: Development of automated sentiment or opinion discovery system: review. In: Proceedings of ICRTET 2013 (2013)
8. Jawale, M.A., Dr., Kyatanavar, D.N., Pawar, A.B.: Implementation of automated sentiment discovery system. In: Proceedings of IEEE International Conference on Recent Advances and Innovations in Engineering (ICRAIE- 2014), pp. 1–6 (2014). ISBN: 978-1-4799-4041-7
9. Leong, C.K., Lee, Y.H., Mak, W.K.: Mining sentiments in SMS texts for teaching evaluation. In: Expert Systems with Applications, pp. 2584–2589 (2012)
10. Liu, B.: Sentiment analysis and subjectivity. In: Handbook of Natural Language Processing, 2nd edn. pp. 1–38 (2012)
11. Liu, B.: Sentiment analysis: a multi-faceted problem. In: IEEE Intelligent Systems, pp. 1–5 (2010)
12. Rainie, L., Horrigan, J.: Election 2006 online, Pew Internet & American Life Project Report, Jan 2007
13. Tang, H., Tan, S., Cheng, X.: A survey on sentiment detection of reviews. In: Science Direct, Expert Systems with Applications, pp. 10760–10773 (2009)
14. Yin, C., Peng, Q.: Sentiment analysis for product features in Chinese reviews based on semantic association. In: International Conference on Artificial Intelligence and Computational Intelligence, pp. 82–85 (2009)