Open Access
Full Length Article

REVIEW ARTICLE

Opinion Mining and Summarization: A Comprehensive Review

Asim Ullah Jan[1], Mohammad Abid Khan[2], Neelam Mukhtar[3]

[1,2,3] Department of Computer Science, University of Peshawar, Pakistan

ABSTRACT

Opinion Mining is concerned with the skillful extraction of vital information from opinionated text. Due to the rapid growth of social media sites, discussion forums, and online stores in the recent past, thousands of opinions are generated on hourly basis. Examining all these reviews from several sources is a dangling task. To grapple with this problem, opinion summarization is a way, where summary is generated from a set of opinionated data. Nevertheless, making an optimal opinion summarization system is a challenging task. This paper presents an overview of the approaches experimented and practiced so far in the field of Opinion Mining and Summarization and a survey of those techniques/approaches. These include: 1) Natural Language Processing and data mining techniques 2) Machine learning, deep learning and lexicon-based methods for sentiment prediction 3) Methods used for summarization. In a nutshell, an innovative framework is presented, which is an amalgam of different types of opinionated summarization techniques.

Keywords: Opinion Mining, Opinion Summarization, Text Summarization, Natural Language Processing, Feature Extraction, Sentiment Classification.

Author`s Contribution	Address of Correspondence	Article info.
[1,2,3]*Manuscript writing, Data analysis, interpretation, Conception, synthesis, planning of research, Interpretation and discussion, Data Collection*	*Neelam Mukhtar* *Email: sameen_gul@yahoo.com*	*Received: Dec 16, 2019* *Accepted: June 14, 2020* *Published: June 30,2020*

Cite this article: *Jan AU, Khan MA, Mukhtar N. Opinion mining and Summarization: A comprehensive Review. J. Inf. commun. technol. robot. appl.2020; 11(1):76-96*

Funding Source: *Nil*
Conflict of Interest: *Nil*

INTRODUCTION

Opinion Mining (OM) or Sentiment Analysis (SA) is a field of computational linguistics and text mining, which aims to analyze people's opinions, attitudes or sentiments towards an entity. Opinion means the views of someone about an entity. The person who gives opinions is called opinion holder and target can be e.g. a product or a service about which the opinion is expressed. In other words, the target about which the opinions are expressed is known as topic. The fundamental terms used in OM model mainly include; aspect, opinion holder, target (object) and time [1]. Time stands for the particular time when the opinion is expressed.

The opinions may be about different marketing products and services, or these may be about different issues, events and individuals. These opinions play an important role in different application domains. As an example, in marketing, it is helpful to know the customers' likes or dislikes about any product or service, which help business executives in formulating a marketing strategy and improve product or service quality. At the other end, the new customers may also want to know the feedback about a product or a service from existing consumers before purchasing it or adopting it. OM systems can be used to predict sales performance and ranking of products and merchants. Opinionated data from Twitter, movie-review sites and blogs were used to predict box-office revenue for film industries. These systems are helpful in advertisement placement. When one posts a negative comment e.g. about a product or a service; the competitors place their advertisements timely and appropriately based on the recommendation of the system. In politics, the field of SA is used to know what people say about a political leader before deciding to vote him/her. One can also predict the results of elections from these opinions [2]. SA systems enable policy makers to analyze public opinions with respect to policies, public services or political issues [3]. These systems can be used to find the influence factors of something on the basis of which people like or dislike it [4]. Tourism is another hot area of OM. Hundreds of traveling sites are available on the web, where people share their experiences and sentiments about different places. Analyzing these opinions can help the tourists in where to travel, in which hotel to stay there and many more related decisions [5, 6]. Customers' positive reviews play a crucial role in recommending products/services of a company. Several recommendation systems endorse items based on customers' opinions [7-10]. Software quality prediction and software recommendation can be enhanced by analyzing existing end-user comments. The literature of the subjected area states many systems which utilize software reviews for product evaluation in order to maintain quality [11, 12].

An example of OM model is presented in Figure 1, where an opinion holder expressed his views about the specific aspect of an object.

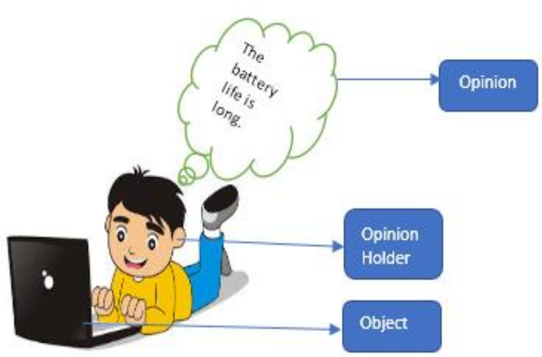

Figure 1. Opinion Mining Model

An opinion is classified as positive, negative or neutral. This classification is known as opinion orientation [13]. Opinion classification or sentiment classification is approached in three different ways; document-level classification, sentence-level classification, and aspect level classification [14]. Different approaches are practiced for classification of the evaluative document to be negative or positive [15, 16]. Sentence level sentiment classification has been explored deeply by different researchers. A sentence is considered to be classified into positive, negative or neutral [17-20]. Aspect-based opinion mining has attracted researchers which has been explored by [21-23]. As most of the researchers in the subject area are more interested in feature-based approach, so different researchers have used their models and perspectives targeting this approach [24]. In state-of-art literature, aspects are also known as features. To avoid the confusion arising due to the term feature that is used in machine learning for a data attribute, the use of the term aspect is becoming more popular than a feature in recent years [1]. Therefore, aspect-based sentiment analysis is also called feature-based opinion mining by Hu and Liu (2004). In this review paper, the terms aspect-based and feature-based will be used interchangeably. In a typical product review, a reviewer may appreciate some of the features while ignoring others due to certain reasons. Thus it is important to find opinions of the reviewers about each individual feature of the product instead of considering the overall opinions in the reviews [25].

Mining opinions and sentiments of people from human languages is an extremely complicated task because it requires deep knowledge of natural language understanding. Most of the opinionated text involves regular and irregular, explicit and implicit, syntactic and semantic rules of language that should be understood. Several existing techniques mainly focus on the syntax of text, in which opinions are expressed explicitly, such as opinions include polarity words, affect terms and their co-occurrence frequencies. However, most of the opinions are often conveyed indirectly using latent semantics, which makes the syntactic approaches ineffective [26]. To overcome these natural language issues, semantic computing is the way, which is a holistic approach to understand natural language by coping different subproblems of extracting meaning and polarity from text. Sentiment analysis task can be divided into a series of subproblems called a suitcase of natural language processing research problems. Several Natural Language Processing (NLP) tasks are required to solve these problems to achieve human-like performance in analyzing sentiments [27, 28]. Opinion mining also has several other domains such as opinion spam detection [29], opinion based entity ranking [30] and opinion based recommendation. Due to the huge volume of data, SA systems need highly computational power and storage capacity. For this purpose, several researches used Hadoop, MapReduce and other parallel approaches [29, 31-33].

With the advancement of Web 2.0, hundreds of social media sites, blogs and forums are emerged, which provide facilities to discuss different entities such as a product, a service, an individual and a specific issue or a topic. With this increase in discussion forums and social sites, opinions are generated exponentially. Several popular products, services and issues acquire hundreds of reviews that cause information overload problem. Reading all these reviews create cognitive stress hence it is a challenging task to read, understand and analyze many opinions. To cope with this problem, summarization is a wise solution. Several types of opinion summarization systems are presented in literature but developing an optimal summarization system is still a challenging task.

In this paper, the focus is on opinion summarization though feature extraction and sentiment classification are also discussed. Existing reviews and survey articles have classified feature extraction, sentiment classification and summarization approaches from different perspectives as shown in **Error! Reference source not found.**

Table 1. Types of Opinion Mining and Summarization Techniques

Reference of Review Articles	Feature Extraction	Opinion Classification	Opinion Summarization
[34]	NLP and Statistical Techniques	Supervised and unsupervised	Single and Multi-Document Summary, Textual Summary and non-textual Summary
[35]	NLP and Mining Techniques	Learning-based and Lexicon-based methods	Aspect-based and non-aspect-based summary
[1]	NLP Techniques, Supervised learning techniques and topic modeling techniques	Supervised learning and unsupervised learning approaches	Aspect and non-aspect based, Single and multi-document Summary, Traditional Summary types
[36]	Lexicon-based and Statistical-based Techniques	Machine Learning and Lexicon-based Approach	Not Given

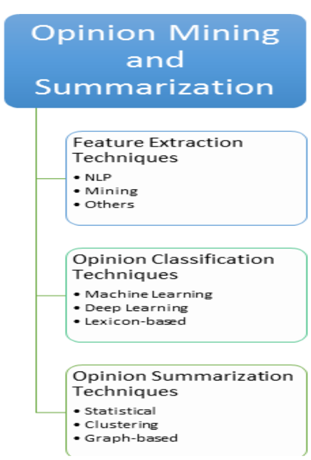

Figure 2. Opinion mining and summarization techniques

In this study, these techniques are categorized in an intuitive manner as shown in **Error! Reference source not found.**An overview of challenges and limitations is also given that remain to be solved in this area. Finally, the article is concluded, and some future research directions are given in the field of sentiment analysis and summarization.

Rest of the paper is organized as follows. The detailed literature review is presented in Section 2. OM and summarization techniques are discussed in Section 3. Section 4 presents an overview of different types of opinion summarization. Section 5 identifies the limitations and Section 6 concludes the paper and highlights future research directions.

LITERATURE REVIEW

Due to rapid growth of social media sites, discussion forums, and online stores, thousands of opinions are generated on an hourly basis about entities including products, services, issues and politicians. Products and services of reputed companies, and issues get thousands of reviews. Examining all these reviews from different sources is a challenging task. To tackle this problem, opinion summarization is a way, in which opinions are extracted and a summary is generated [35]. Opinion summarization is the compact form of original opinionated text while preserving the important content. The term summarization can be formally defined as "summary produced from one or multiple texts, which express key information. Summary size is usually half or less than half of the original text or less than that" [37]. The characteristics of a good summary are preservation of important content, short and concise size and its readability. Opinion summarization is different from conventional text summarization. Opinion mining and summarization firstly extract features from the set of reviews; secondly, it classifies these individual reviews into positive or negative review and finally summarizes them. Opinion summarization uses NLP, machine learning, information extraction, graph theory, visualization and statistical techniques. Different summarization approaches are used by researchers, such as extractive summarization, abstractive summarization, visualized and rate-base summary [38].

Currently, several review papers and books covered opinion summarization study [1, 34-36, 38-56].

Chapter 11 of [46] covers several opinion mining and summarization techniques. In this book, the author first introduces some basic concepts related to definition of opinion mining. Then opinion mining techniques are discussed, covering sentiment classification, opinion spam detection and opinion summarization. The book emphasizes on opinion classification techniques and only a little portion discusses the summary generation. As the book was published in 2007, so recent techniques of sentiment classification and opinion summarization are not covered. Another chapter of [42] is purely focused on sentiment classification techniques, not covering the state-of-the-art summarization methods. Another book of [1] i.e. "Sentiment analysis and opinion mining", covers different opinion summarization approaches. Liu (2012) has divided opinion summarization into aspect-based and structured-based summarization approaches. The book has several topics on aspect-based summarization. Unstructured summarization and non-aspect-based summarization are not reported even briefly in Liu (2012).

Kim et al. [35] covered several articles of opinion mining and summarization. The authors have summarized the work into aspect-based and non-aspect-based summarization. The techniques used in opinion mining and summarization systems are also highlighted in the research of Kim et al. (2011). In this work, opinion summarization systems are divided into textual and non-textual summarization.

Breck & Cardie [39] addressed several research issues in OM systems including opinion lexicon construction, feature identification, identification of opinion holder and sentiment classification. Opinion summarization is also reported along with paragraph-based summarization, single document, and multi-document summarization. However, aspect-based summary, structured and unstructured summaries are not addressed.

Medhat et al. [36] carried out a systematic survey on SA techniques. Feature selection techniques and sentiment classification techniques are briefly discussed in their work. These authors divided SA algorithms into machine learning and lexicon-based approaches. This study focuses on sentiment summarization techniques which are not covered in the survey of [36].

Sun, Luo, & Chen [57] have presented a review on NLP techniques used in OM systems. The authors discussed general NLP techniques required for preprocessing, opinion classification and comparative OM. In this work, opinion summarization is not explored in depth.

TECHNIQUE USED

Opinion mining and summarization borrow algorithms from several neighboring disciplines such as machine learning, deep learning, graph theory and data mining [21, 58-63].

The task of opinion mining and summarization can be divided into the following three phases.

A. Feature Extraction
B. Opinion Classification
C. Opinion Summary Generation

Each of the above phases uses several types of techniques. The detailed description of each technique is given in the sub-sections. For feature extraction and sentiment classification, most researchers used data mining and machine learning algorithms, such as association rule mining, support vector machine and Naive Bayes. WordNet, semantic orientation, collocation extraction, and lexicon-based approach are NLP techniques used in summarization. The detailed overview of opinion mining and summarization techniques is already illustrated in Figure 2.

A. Feature Extraction Techniques

Finding a set of data that is representative of a large volume of data is called feature extraction or topic modeling [47]. In OM, extraction, and selection of feature words is an important phase. Words that frequently occur in the text are considered as feature words [64-66]. Feature words may include a single word or a combination of words. The statistical approach, data mining approach and lexicon-based approach are used for feature extraction. In this study, feature extraction techniques are mainly classified into NLP techniques and data mining techniques. The subsequent sections give brief overview of these techniques.

3.1. Natural Language Processing Techniques

NLP techniques such as part-of-speech (POS) tagging, syntax tree parsing, WordNet and n-gram model are used to identify feature words in the opinionated text. POS tagging is used to select nouns and noun phrases as features [67, 68]. Dependency grammar graph method is adopted by distinguishing the relationship between feature words related to specific opinion words [69]. Shallow parsing is used for aspect identification in short comments [70]. POS tagging, and parsing are though effective techniques for feature extraction, but still some problems exist. Firstly, in the case of large-scale processing, the speed of tagging or parsing is still not efficient. Secondly, such shallow NLP based techniques are not efficient enough to identify every feature, as features do not necessarily always occur as nouns or sometimes they are not explicitly specified in the opinionated text [35].

Several NLP toolkits including NLTK [71], OpenNLP and CoreNLP [72] are used to improve the accuracy of feature selection. These toolkits perform different tasks i.e. POS tagging, word tokenization, name entity recognition and sentence parsing [57].

3.2. Mining Techniques

Data mining approach is another approach for features extraction. It is also called Association rule mining (ARM). ARM finds frequent patterns in data [73, 74]. In review mining, ARM is used for frequent feature extraction of an entity [21, 58, 75-80]. ARM uses support, confidence, and lift measures for extraction of rules. Let X and Y be two words in a corpus whose support and confidence are calculated as follows [73].

$$Support(\mathrm{X} \Rightarrow \mathrm{Y}) = \mathrm{P}(\mathrm{X} \cup \mathrm{Y}) \quad \ldots\ldots\ldots\ldots \quad (1)$$

$$Confidence(\mathrm{X} \Rightarrow \mathrm{Y}) = \frac{\mathrm{Support}(\mathrm{X} \cup \mathrm{Y})}{\mathrm{Support}(\mathrm{X})} = P(\mathrm{X}/\,\mathrm{Y}) \quad \ldots (2)$$

ARM-based approach is used by several researchers for performing the task of feature extraction [21, 81]. Association rule mining is performed after segmentation and tagging of data to learn rules for feature words prediction. These trained rules are then used to extract features from the newly input dataset.

3.2. Other Techniques

Information retrieval and information theory techniques are also used for feature extraction i.e. Term Frequency-Inverse Document Frequency (TF-IDF) and Pointwise Mutual Information (PMI). The detailed description of these techniques is in the following sub-sections.

- **Term Frequency-Inverse Document Frequency (TF-IDF)**

TD-IDF indicates how significant a term is in a text document/set of documents or a corpus. TF-IDF is an Information retrieval technique, which aims to find important words in a corpus. It is also used for stop-words removal in a preprocessing step. In opinion mining and summarization, it is used for feature-words extraction and entity identification [58, 62, 82]. TF-IDF weightage is calculated by the following equation:

$$tf - idf_{t,d} = tf_{t,d} \times idf_t \qquad \text{..........} (3)$$

Where Tf t, d indicates that how many times a term "t" is occurring in document "d" while idf t shows its frequency in the whole corpus. The multiplication of tf t, d and idf t concludes the weightage of these terms.

- **Point wise Mutual Information**

Pointwise Mutual Information (PMI) is an information theory approach used for finding collocations. It is based on conditional probability. In OM and summarization, PMI is used for feature extraction [83, 84] and opinions classification into positive and negative [67]. PMI shows the mutual information between feature and the classes. Equation 4 is used for the calculation of PMI.

$$PMI(x;y) = \log \frac{P(x,y)}{P(x)P(y)} \qquad \text{............} (4)$$

This equation calculates the dependency between terms X and Y by using probability function. If the variables are highly associated, then the PMI value will be maximized.

B. Opinion Classification Techniques

Opinion classification is the identification phase of sentiment polarity detection where the positivity or negativity about an entity is identified. As different people may have different views about similar aspects, this phase helps in discovering the general sentiments (positive, negative or neutral) about these aspects. There are different methods that are used by researchers for sentiment prediction, which are categorized into machine learning based, deep learning-based and lexicon based methods [15, 47, 85, 86].

a. Machine Learning-based Methods

Machine learning techniques are used for identifying the appropriate class for a review. Machine learning algorithms solve the sentiment classification problem as normal text classification using syntactic and linguistic features. Naïve Bayes Classifier, Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Maximum Entropy, Decision Tree and Clustering techniques are examples of classification techniques [36, 70, 75, 87-93].

Naive Bayes classifier is one of the most frequently used supervised classification technique. It computes posterior probability based on word distribution in a review document. Bayes Theorem is used to calculate the probability that a given review belongs to a positive or to a negative class. The following equation calculate the probability.

$$P(Positive|Good) = \frac{P(Good) * P(Good|Positive)}{P(Good)} \qquad \text{.........} (5)$$

P (Positive|Good) is the posterior probability, which we want to find whether the review is positive or negative. P (Good) is the prior probability of a word "Good", how many times an opinionated word "Good" occurs in the review sentences. P (Good|Postive) is the known probability that the word "Good" has the probability to be positive in the trained data.

Support Vector Machine is another supervised linear classification algorithm. SVMs separate the search space into different classes by determining linear separators. Figure 3 shows that there are two classes i.e. O, ■ and three hyper planes. A hyperplane separates the data points into different classes. An optimal hyperplane classifies data so that the distance from its nearest data point on each side is maximized. In Figure 3, the middle hyper plane provides maximum margin separations [36].

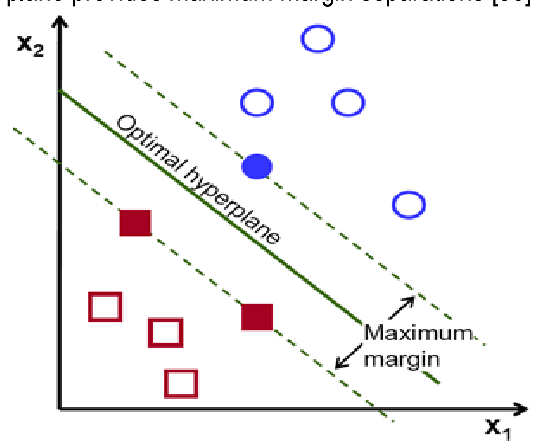

Figure 3. SVMs for classification problem [36]

In opinion mining, SVMs are used to classify a set of opinions based on the polarity into positive and negative. SVMs have been applied in several sorts of text classification tasks and achieved promising success [94, 95].

b. Deep Learning and Artificial Neural Networks

Artificial Neural Networks (ANNs) are machine learning models inspired by biological neural networks, whose functionality is dependent on enormous number of inputs. ANNs are a set of interconnected neurons that receive input, perform progressively complex calculations and then use the output for solving the problems. Such systems learn to perform some tasks by considering training examples regardless of task-specific programming. A single layer ANN has been shown in Figure 4. These networks are used for different purposes including Natural Language Processing (NLP), image processing, speech recognition and social network filtering [96-101].

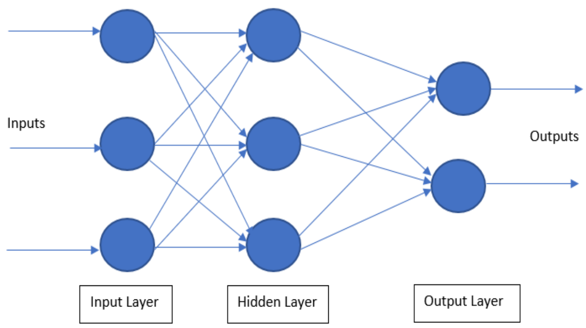

Figure 4. A single layer Artificial Neural Network

Deep learning is the extended form of ANNs that are composed of multiple hidden layers between input and output layers. Deep learning is also called multiple layer neural networks. In Figure 5, an example of multiple layer neural network architecture is shown, two hidden layers H1 and H2 are used to perform classification task. Accordingly, it learns features in the multiple hidden layers from a given input and produces classification result in the output layer. It learns features in a supervised or unsupervised manner within a hierarchy. The upper layers in the hierarchy have more abstract representations than the lower layers. The upper layers evolve in training to solve complex problems in the lower layers [102].Various deep learning architectures such as Deep Neural Networks (DNNs), Deep Recursive Neural Networks (DRNNs), Recurrent Neural Networks (RNNs), deep belief networks and Deep Convolutional Neural Networks (DCNNs) have been applied to various fields i.e. NLP, automatic speech recognition, computer vision and bioinformatics [103, 104]. Deep learning produces promising results in NLP tasks i.e. Part-Of-Speech tagging, chunking, semantic labeling, cross lingual problems and name entity recognition [105, 106]. In state-of-the art systems, deep learning has also been practiced in many sentiment mining tasks including aspect extraction, sentiment extraction, sentiment classification and sentiment lexicon learning [44, 103, 107].

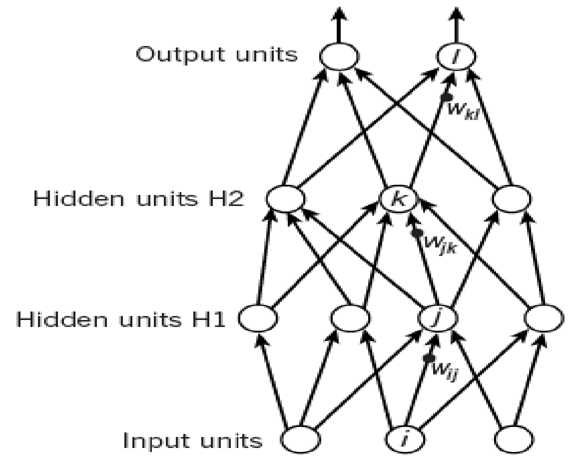

Figure 5. Multilayer Neural Networks [103]

Sentiment classification of different domain data, using same system, is a challenging task. Stacked Denoising Auto-Encoders with sparse recite [108] are used to train data for different domains classification. They successfully perform domain adaptation on large dataset of 22 domains. Dos Santos & Gatti [109] used DCNNs to perform sentiment analysis by exploiting character-to-sentence level information of twitter data. The authors used two convolutional layers. The authors extracted features from individual words by exploiting character-level analysis as well as classifying sentences of any length. The authors compared their results with existing state-of-the art techniques i.e. SVM, Naïve Bayes and Maximum entropy. Hence, satisfactory results are achieved. Dos Santos[110] used Character-to-Sentence Convolutional Neural Networks (CharSCNNs) presented by Dos Santos & Gatti [109]. The authors performed message polarity classification and contextual polarity disambiguation by using these networks. DCNNs are also used for multi-label classification problem by Chen [111]. Poria et.al. used deep convolutional neural networks for multimodal sentiment analysis [112, 113]. The authors extracted features from short text, short video clips and audio clips and combined them to train the classifier. They compared their system with parallelizable decision label data fusion method hence, obtained 14% better performance. Various types of deep learning algorithms are also used to evaluate the performance of different sentiment classification algorithms [114].

c. Lexicon Based Methods

A lexicon is made up of a list of positive and negative words for matching the words in the opinionated text. A set of rules is often used with this word list or the list is combined with the results of parsing or tagging. Dictionary approach, SentiWordNet and Parsing are very popular in opinion classification [21, 22, 76-78, 81, 115, 116]. The term sentiment word dictionary is also used for lexicon interchangeably.

An effective method was proposed by Hu and Liu that was based on WordNet for identifying the opinions about particular features and the orientation by [21, 81]. Initially, these authors started with 30 seed adjectives for each predefined orientation (positive and negative). Then the antonym and the similarity relation in WordNet were used for assigning the negative or the positive orientation to a huge set of adjectives. Thus, by the orientation of adjectives, the orientation of opinion about a feature was decided.

A set of positive and negative words for sentiment prediction was used by [115]. Two sets of sentiment words were used followed by two thesauri to enlarge the seed vocabulary. Based on the orientation of words, the orientation of an opinionated sentence was decided. Sentiment scores were assigned to sentences, representing the polarity and sentiment degree. Zero opinion score was assigned to neutral opinions in addition to positive and negative opinions. To identify the opinions with feature words, dependency relationships were used [69]. These authors adopted a strategy that was similar to Hu and Liu (2004a; 2004b). Firstly, top 100 positive and negative opinionated words were identified from a training set and then WordNet was used for assigning an orientation to the words. The orientation of a word was reversed in the presence of negation words (e.g. not). Lexicon based methods are quite popular. These approaches work well in product review domains where people express their opinions explicitly. In other domains, for example movie reviews, the performance of these methods may be poor due to ignorance of the context or the sarcasm expressed by the people.

Some other techniques are also practiced for opinion classification. In several studies, fuzzy logic and ontology-based methods are extensively used [5, 6, 90, 117].

C. Opinion Summarization Techniques

Opinion summarization is representation of large volume of reviews in compact and concise way, which helps the reader to make quick and efficient decisions. Techniques such as statistical, graph-based, ranking and clustering-based are practiced for summarization of opinionated text. Statistical techniques are used for the non-textual summarization. Similarly, graph-based technique, ranking, and clustering are used for textual summarization. Statistical methods are also utilized for textual summarization but most of the times it is used for producing a structured summary.

Statistical Techniques

The most commonly adopted techniques for generating a summary are the statistical techniques that are adopted by a number of researchers [21, 22, 69]. This form of summary utilizes the results obtained from aspects extraction and the results obtained from the sentiment prediction phase. The general sentiment can easily be understood by observing the number of positive and negative opinions about each aspect. Opinion Observer is developed by Liu et al (2005), that enables the user not only to compare the different aspects of the same product but of different competing products as well by using statistical graphs.

Clustering and Ranking Algorithms

People normally express their opinions about different aspects of an entity. These aspects are normally broken explicitly or implicitly into different sentences or sections. In summary, all of those aspects should be covered, which are incorporated through clustering. Aspect-based clustering is widely used by researchers, where reviews about an aspect are placed in an aspect cluster[75-77, 118]. To mention all aspects in the summary, most informative sentences are extracted from aspect-based clusters. These informative sentences are selected using various ranking algorithms e.g. ClusterRank [62], PageRank [119] and LexRank [120].

- **Graph-based Methods**

A graph is an integration of nodes and edges where nodes represent words, phrases or sentences and edges represent relations between the nodes.The graph-based method is used to rank most informative sentences or phrases for extractive summarization. In Figure 6, weighted graph is illustrated that select informative sentences based on similar words. In this graph, the nodes represent sentences i.e. S1 upto Sn and edges represent number of similar words between sentences.

Several researchers adopted graph-based summarization techniques [38,59,61,80,121-123].

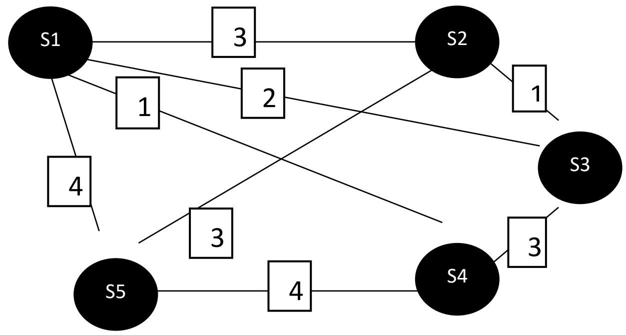

Figure 6. Weighted graph for Informative sentences selection (alZahir et al., 2015)

Ganesan et al. [61] proposed graph-based approach named "Opinosis" that constructs an abstractive summary of redundant opinions. Textual-graph is constructed where nodes represent individual words and edges show the sentence structure. Zhu et al. [121] selected informative sentences using a graph-based method to construct an extractive summary. These authors frame the informative sentence selection problem as a community leader detection problem. Sentences that have most similar words are considered as leader sentences. alZahir et al. [59] utilized multi-edge graph, where nodes represent sentences and edges represent similar words between sentences. The symmetric matrix is used to rank sentences and select most informative sentences for the extractive summary.

TYPES OF SUMMARIZATION

In this review, many types of summaries have been identified including indicative summary, informative summary, extractive summary and abstractive summary. Indicative summary provides information about the text without having internal information within the text, while informative summary produces the compressed version of the text from the original content [37]. Extractive summary identifies key material in the text while abstraction is regenerating the important portions of text and combining these portions of extracted text. Abstractive summarization used compression techniques for pruning of unimportant portions from text [124].

Types of opinion summarization can be divided from different perspectives i.e. aspect-based and non-aspect-based opinion summary, single document and multi-document summary, and textual and unstructured summary. In this study, summaries are divided into non-textual and textual. Non-textual summaries are in the form of rating and statistical graphs that express only the sentiment prediction into positive and negative, while textual summaries provide detailed information of the reviewers in plain text.

A. Aspect-based Summary vs. Non-aspect-Based Summary

In the case of aspect-based opinion summarization, the text is first divided into aspects and then a summary is generated [21, 69, 70], while in the case of non-aspect-based opinion summarization, a generalized summary is produced without considering the aspects [61]. Aspect-based opinion summarization systems follow three steps [125]. Figure 7 shows these steps. Firstly, identify product features from a set of reviews e.g. battery life, sound quality, and ease of use. Secondly, predict the orientation e.g. positive or negative and finally develop an aspect-based summary i.e. rating or textual summary.

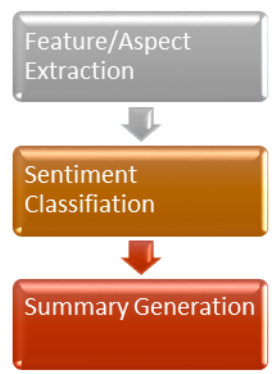

Figure 7. Steps of Aspect-based Opinion Summarization

Aspect-based opinion summarization has two main characteristics. First, it identifies the opinion targets and the sentiments about them. Secondly, it provides the percentage of people who are having positive or negative opinions about the entities and their features. A resulting summary is a form of a structured summary [35]. Typical feature based summary is presented by Hu & Liu [21].This structured summary can also be represented visually [22]. Bar charts are used for visualizing the summary, where the bar above X-axis shows the positive opinions on a feature and the corresponding bar below the X-axis presents the negative opinions that are commented on the same feature. An example of a visual summary is shown in figure 8. Here, different features and their orientations of Cannon Digital Camera are visually represented.

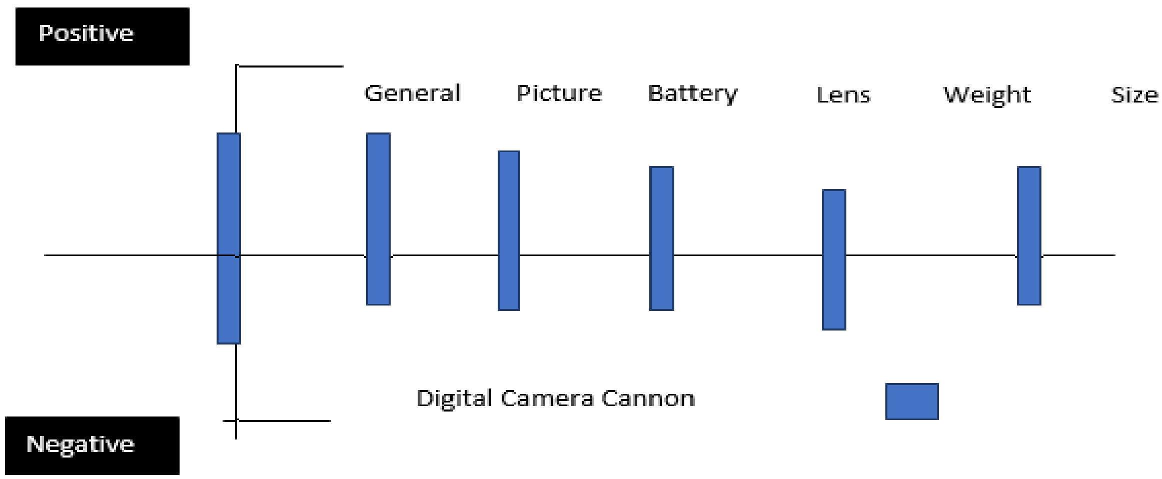

Figure 8. Aspect-based Visual Summary of Digital Camera Cannon (Liu et al., 2005)

Visual comparison of more than one entity is also possible in this summary. Other researchers have adopted structured approach for the summarization of movie reviews [69], summarization of Chinese opinionated text [115] and service reviews summarization [126]. The feature-based summary is not restricted to be in structured form only; it can also be in the form of an extractive and an abstractive-form on the same idea [77, 121].

B. Single document and Multi-Document Summary

Single and multi-document summarizations are traditional approaches, which are used for factual text summarization while opinion summarization is somewhat different from traditional single or multi-document summarization. Opinion-summary is about entities, their features, and appraisal about them. It can be in a quantitative manner i.e. structured summary and feature-based [1].On the other hand, traditional summarization approaches appear in emails summary, educational articles summary, business reports summary and news articles summarization [124]. In single document summarization approach, informative sentences are extracted from the original text and short summary is produced. Multi-document summarization finds and extracts dissimilar information from multi-documents and avoid redundant information to form summary.

Single-document extraction summary utilizes features such as term frequency, the location of the text, identifying key phrases in the text, and sentence length. These features are based on machine learning and NLP techniques. Identifying relations between words, discourse structure identification, recognizing the lexical connection between terms, anaphora resolution and synonyms finding are some major tasks required for single-document extractive summarization. Single-document abstractive summarization approach uses information fusion, language generation, text compression, information extraction, tree-based and ontology-based methods. Multi-document summarization recognizes redundant and important information and produces a coherent summary. For finding redundant information, similarity functions like cosine similarity and Maximum Marginal Relevance (MMR) are used [37].

C. Textual Summary and Non-Textual Summary

A textual summary provides critical information conveying the key opinions while a non-textual summary gives a general overview of an entity [83]. Non-textual summarization uses statistical methods and visualization techniques to form non-textual summary while textual summarization adopts graph theory, statistical model, language generation and compression tools to build a summary in plain text. Types of textual and non-textual summaries are demonstrated in Figure 9.

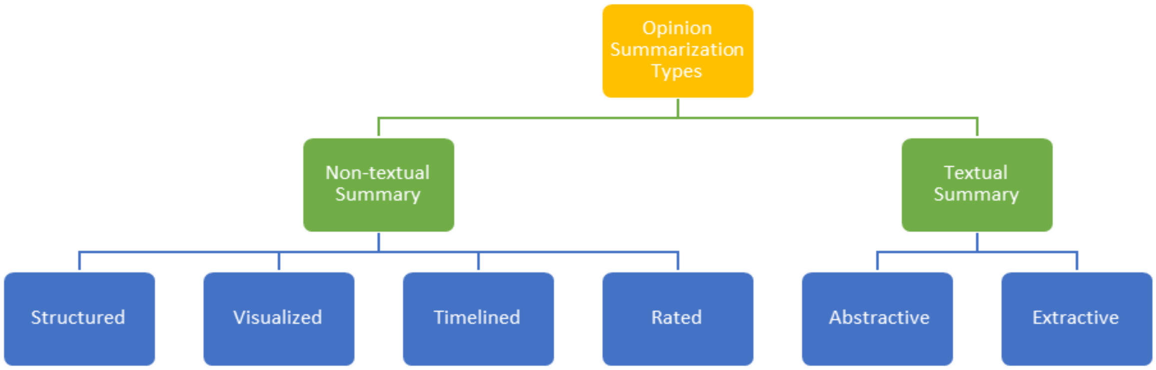

Figure 9. Textual and Non-Textual Summarization

- **Types of Non-Textual Summarization**

Non-textual summarization approaches generate simple summary in the form of rating and other statistical measures. In non-textual summarization, opinion orientations are determined, and the total number of positive and negative opinions is reported. The non-textual summary is useful for conveying general opinions about an entity such as product or service, person or organization but it does not provide the detail of sentiment. Structured, visualized, timeline and rate-based summaries are different types of non-textual summarization.

- **Structured Summary**

This approach produces a summary in a structured format, which is easy to understand. Structured summarization approach is normally used in opinion summarization systems [21, 58, 69, 78]. Figure 10 is an example of the structured summary. Hu & Liu [21] developed structured summarization approach that used Association Rule Mining (ARM) and WordNet for feature extraction and opinion sentence classification. The system builds feature-based summary. Each feature is ranked per its frequency in the reviews. Count number of positive and negative comments about the features and list all the sentences. Zhuang et al. [69] introduced multi-knowledge based, structured summarization approach for movie reviews. Positive and negative opinions about each feature are summed up and a structured summary is formed [21]. More & Tidke [58] proposed a model which used TF-IDF and ARM for extracting frequently opinionated words and their polarity. The summary is subsequently constructed based on polarity into the number of negative and positive opinions. Fan & WU [78] introduced a system to produce structured summary from Chinese text. ARM produced frequent features from the comments. Adjectives dictionary is used to classify sentimental sentences. For semantic orientation of adjectives, the Chinese version of WordNet is used. The system counts the number of positive and negative opinion words in the review and forms a summary.

Galaxy_S4
Aspect: Camera
    Positive: 150
        <individual review sentences>
    Negative: 6
        <Individual review sentences>
Aspect: Size
    Positive: 239
        <individual review sentences>
    Negative: 41
        <Individual review sentences>

Figure 10. Structured Summary

Figure 10 illustrates a structured summary of Galaxy S4 which is a type of a cell phone. Two features i.e. camera and size along with the number of positive and negative comments are shown. In this survey, several articles related to opinion summarization are studied. Several techniques have been used by researchers in the literature for making a structured summary. Table shows these techniques in detail. For feature extraction, association rule mining is widely used. Dictionary approach and other similar NLP techniques are used for opinion classification. In structured summarization, statistical techniques are extensively practiced.

Table 2. Structure Summarization Techniques and Datasets

References	Feature Extraction techniques	Sentiment Classification Techniques	Summarization Techniques	Type of Techniques used	Datasets/Sources
[58]	TF-IDF, ARM	ARM	Polarity Score	IR[1],DM[2]	Not given
[78]	ARM	Dictionary Approach, SO[3]	Structure Summary	DM	Amazon.cn
[69]	Multi-Knowledgebase (WordNet, movie casts)	Trained label Data, statistics, Regular expression	Feature-based Structure summary	Statistics	IMDB (Movies Review)
[21]	ARM	Lexicon Based, WordNet, SO	Structure	DM	Amazon.com, C|net.com

[1]Information Retrieval, [2]Data Mining, [3]Sentiment Orientation

- **Visualized Summary**

Visualization based summarization systems also adopted the structured format that uses statistical methods with graphical presentations. Potthast & Becker [84] proposed opinionated WordsCloud approach for summarization. The increased size of opinionated words in the cloud depends on its frequency. Figure 11 shows WordsCloud visualization summary.

Figure 11. Visualized Summary (Potthast & Becker, 2010)

Liu et al [22] illustrated an approach to compare different product reviews and summarized these reviews using a visualization method called "Opinion Observer". Several other visual summarization systems e.g. "Opinion blocks" and "Opinion Seer" exist in the literature [127, 128]. Several visual summarization techniques have been identified in this study, which are listed in table 3.

Table 3. Visualized Summary Techniques and Datasets

References	Feature Extraction techniques	Sentiment Classification Techniques	Summarization Techniques	Type of Techniques used	Datasets/Sources
[67]	Parsing (nouns are selected as feature)	Label data of existing domain, PMI	Graphical Summary	NLP	Amazon.com
[68]	POS, WordNet, Parsing (N & NP as feature)	SentiWordNet, Feature orientation table	Statistical Feature-based summary, Graph	Statistics, NLP	Several Reviews Sites (Not specifically mentioned)
[84]	Pointwise Mutual Information (PMI)	Dictionary Approach, SO	Visualization Technique	Visualization	YouTube Video Comments, Flickr image Comments
[22]	ARM	Dictionary based Approach	Visualization Technique, statistics	DM, Statistics	Epinions.com, Amazon.com

For feature extraction and sentiment classification tasks, NLP methods such as parsing, and dictionary-based approach are practiced by the researchers while for summarization purpose statistical and graphical techniques are used.

- **Timeline summary**

Opinions change with respect to time. Thus, the summary with a timeline helps to identify different trends of opinions about the entity with respect to time. The summary also helps in developing ideas for further analysis. Summaries with the opinions trends over timeline are developed [23, 115]. Figure 12 presents a timeline summary.

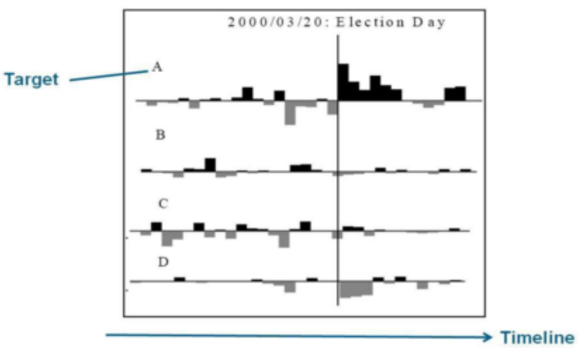

Figure 12. Timeline Summary (Ku et al., 2006)

In Figure 12, the alphabets A, B, C and D demonstrate politicians for whom the public opinions are measured. Public opinions are changed from time to time as shown. The upper bar shows positive opinions and the lower bar shows negative opinions.

- **Aggregated Ratings**

Statistical techniques with text selection are combined to produce better results [70]. Aspects are identified by using clustering and topic modeling. For each aspect, the sentiment identification result of phrases is averaged as the final rating of the sentiment for that aspect. Rating of the aspects is shown by using the phrases that are the most representative.

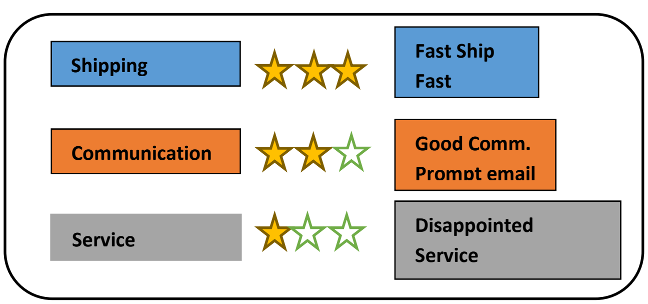

Figure 13. Rating-based Summary (Lu et al., 2009)

Figure 13 illustrates a rating-based summary. Different features of an entity e.g. shipping, communication, and service are discussed. Features that sound good to the customers are rated high while others are rated medium and low. Customers' keywords for describing a feature are also mentioned in the summary i.e. fast ship, good communication and bad service. Techniques used for rating-based summary are presented in table 4.

Table 4. Rating-based Summarization Techniques and Datasets

References	Feature Extraction techniques	Sentiment Classification Techniques	Summarization Techniques	Type of Techniques used	Datasets
[76]	ARM (N and NP), Probability model	Lexicon-based, dictionary-based, Linguistics rules	Aspect-based Clustering	DM, NLP	Amazon.com
[70]	PLSA[1], K-Means Clustering	Naïve Bayes Classifier, Local and Global Prediction	DSR (Detail Seller Rating), Phrases Frequency	DM	Ebay.com
[126]	Syntactic Pattern (Senti words, relative words frequency)	Lexicon (WordNet), Maximum Entropy	Aspect-based rated Summary	NLP, ML[2]	Tripadvisory, zagats.com

Data mining techniques i.e. clustering, classification, and ARM are applied for feature extraction and opinion classification purpose. Rating approach and aspect identification methods are used for summarization.

- **Types of Textual Summarization**

Textual summarization is categorized into two types i.e. abstractive and extractive. These are traditional summarization types used for e.g. news summarization, email summarization and annual report summarization [129]. Textual summaries are also used in review summarization but due to certain limitations e.g. text generations, and grammatical rules violation, these techniques are not widely used for review summarization.

- **Abstractive Summarization**

The abstractive summary is a type of summary that contains the main idea of a source text in different words. In this type of summarization, the system first understands the input text using NLP techniques and then rephrase the input text to generate a concise and generalized summary. Abstractive summarization makes use of advanced language generation, information fusion, extraction and compression techniques to minimize the summary length and maximize the informative content [38, 124, 130, 131]. The abstractive summary is more comprehensive and concise than extractive summary. Many researchers, as evident from the literature, work on abstractive summarization [61, 83]. Ganesan et al [83] proposed a system to generate concise, brief, representative of key information and readable summary from the opinionated text as shown in figure 14.

Mobile Phone Y	Restaurant X
Battery life is short. Big and Clear screen. (8 words)	Good Service. Delicious soup dishes. Very noisy at nights. (9 words)

Figure 14. Abstractive Summary Example (Ganesan et al., 2012)

Important aspects of entity X and Y are summarized in a few words. PMI function is used to generate frequent patterns in the text, which represents key information. For readability issue, the authors used n-gram model. To produce a set of concise, representative and readable summary, the authors framed this problem as an optimization problem. Heuristics algorithms (greedy algorithms) were proposed to solve the optimization problem. Ganesan et al [61] used graph-based approach and produced highly redundant opinions summary. The method extracted key opinions and produced a concise and grammatically correct summary. Table 5 lists approaches that are used for the development of abstractive summary.

Table 5. Abstractive Summarization Techniques

References	Feature Extraction techniques	Sentiment Classification Techniques	Summarization Techniques	Type of Techniques used	Datasets/Sources
[83]	PMI	No	Greedy Algorithm	Machine Learning	CNET
[132]	Not Given	Not given	Graph,Dijkstra's algorithm, TF-IDF	Graph theory, Ranking, Information Retrieval	DUC 2002 newswire corpus
[61]	Graph	No	Graph	Graph Theory	TripAdvisor, Amazon, Edmunds

This study investigated various abstractive summarization articles. Graph theory, ranking and information retrieval methods are highly used by the researchers.

- **Extractive Summarization**

In this type of summarization, important phrases and sentences are extracted from the source text and they are combined to form a summary. Important/central/silence sentences are ranked and selected using PageRank, SentenceRank, ClusterRank and LexRank algorithms [62, 120, 121, 133]. Extractive summaries use statistical analysis of words/phrases, the frequency of phrases or words and their location to extract important content for a summary. Extractive summarization is easier than abstractive one because abstractive summarization needs advanced compression techniques. Extractive summarization has its own limitations, such as large summary size to cover key information, the existence of less relevant information, important information mostly spread out, and co-referencing problems [60]. Researchers have proposed several types of extractive summarization systems [59, 87, 88, 121, 134].

alZahir et al.[59] introduced a multi-edge graph method, where nodes represent sentences and edges represent similar words in the sentences. A symmetric matrix is constructed from the graph where the sum of all the rows produces a ranked vector. Highest ranked sentences are combined to build the summary. Raut & Londhe [87] proposed a model that extracts hotel reviews. Machine learning algorithm (naive Bayes) and SentiWordNet are used for classification of reviews into positive and negative. The relevance-scoring method is used to generate an extractive summary. Zhu et al [121] presented a sentence-based extractive summarization method. Most informative sentences were selected from a set of reviews, which cover several aspects of an entity. Informative sentences were selected using community leader detection algorithm. Ly et al.[88] introduced a cluster-based extractive summarization technique. Frequent features are clustered according to their polarity into positive or negative. The highest polarity score sentences are extracted from both positive and negative clusters to form the summary. Table 6 includes extractive summarization articles and gives detail of the description of different techniques used for summarization.

Table 6. Extractive Summarization Techniques

References	Feature Extraction techniques	Sentiment Classification Techniques	Summarization Techniques	Type of Techniques used	Datasets/Sources
[134]	Recognizing Textual Entailment (RTE)	NG	Sentence Scoring and Ranking	Statistical, Semantics, Ranking	USA Today, CNN, Twitter
[135]	Stanford Parser, Dependency Grammar	NG	Hyperlink-Induced Topic Search (HITS), Logistic Regression (LR)	Graph Theory, Statistics, NLP	NG
[59]	NG	NG	Graph	Graph Theory	Lu.lv (A pdf file on this site)
[87]	Web Crawler	Naïve Bayes, SVM SentiWordNet	Term Frequency, Relevance Score	DM, ML, IR	Tripadvisory.com
[121]	Clustering	Cosine Similarity, SO	Graph, Community leader detection	DM, Graph Theory	Amazon.com
[77]	ARM, Probabilistic Model (characteristic power equation)	Dictionary based approach (Opinion Lexicon)	Feature-Opinion Excerpt, Clustering	DM, NLP	Amazon.com
[79]	LDA[6]	SentiWordNet	Sentence Weightage (Statistical Model)	IR, Ranking	Amazon.com, C|net.com
[88]	ARM	Clustering	Polarity Score	DM	Amazon, CNET
[62]	TF-IDF, Clustering	Not Given	Graph, Cluster Rank, Cosine Similarity, Greedy algorithm	DM, Graph theory, IR, Ranking	AMI Meeting Corpus

Table 6 has listed nine recent articles of extractive summarization. This study concludes that clustering, graph theory, and ranking algorithms are extensively used by the researchers for extractive summarization.

CHALLENGES AND LIMITATIONS

There are several gaps in the field of sentiment analysis and summarization indicating the room for improvement. The sections below briefly discuss these gaps.

- **Limited Words Consideration as an Aspect and Opinion**

Currently, a lot of research is done by considering explicit aspects, but little attention is paid to implicit aspects. Adjectives are mostly considered as main candidates for opinion words although other Parts-Of-Speech such as nouns or verbs can also act as opinion words i.e. opinions can be expressed by opinion bearing verbs and nouns.

- **Pre-processing and NLP Problems**

Mining opinions and sentiments depend on high-level NLP task such as POS tagging, sentence and word tokenization and parsing. For some languages, e.g. Chinese and Japanese, word segmentation is crucial. Parsing and POS tagging produce key information for opinion classification and feature extraction tasks. Though these tasks achieved good results for traditional document classification and summarization, they are unsatisfactory for reviews that are often grammatically incorrect. To deal with this problem, there is a need for the development of algorithms that will correct spellings of words from the context and will assign correct part of speech to the corrected word. Standard dictionaries can be utilized to avoid the spelling mistakes and the context of a word needs to be considered before the assignment of parts-of-speech. The incomplete or insufficient lexicon may also be a source of inaccuracy. Lexicon coverage may be verified by using enough and appropriate data. Problems in the identification of idioms, sarcasm, and irony are noticed and need the attention of researchers. Researchers are also focusing on several other unsolved problems in NLP i.e. co-reference resolution, named-entity recognition, anaphora resolution, word-sense disambiguation and negation handling.

- **Lack of Datasets and Evaluation Techniques**

Usually, researchers crawl the data from the web according to their requirements. Although there is some published data but there are no standardized datasets (that are widely used) for sentiment analysis and summarization. The issue can be resolved by using data from authentic sources. There are several datasets available for developed languages such as English and Chinese but limited datasets are available for less developed languages such as Arabic, Turkish, Urdu and Hindi.

Use of evaluation measures, for each step of opinion mining and summarization, is another issue that needs to be resolved. Conventional automatic text summarization is evaluated using different methods e.g. Rough [136], Document Understanding Conference (DUC) [137] and Pyramid [138]. However, there is no standard evaluation method that is specifically developed for sentiment summarization. Better and strong evaluation techniques are thus required for evaluating automatic opinion summarization. Evaluation of sentiment classifiers is another issue faced by the researchers who are selecting suitable classification algorithm for specific language data [139].

- **Limitation of Learning Algorithms in Summarization**

Though learning based approaches are studied a lot in sentiment prediction but they are not commonly used in sentiment summarization due to the following limitations:

- A huge amount of annotated data is needed which is a big challenge
- Another challenge is to find data in a general domain
- A model that is trained in one domain may not necessarily work well in another domain

- **Limitations of Textual Summarization**

In state-of-the-art literature, extractive summarization is widely studied because of its simplicity. Limited work has been done in abstractive summarization because this type of summary requires advance language generation, information fusion, and compression tools. Fuzzy logic, ontologies-based methods, and graph theory techniques can be used to overcome the problems of an abstractive summary. Extractive summarization system also faces several challenges i.e. large enough to include all key information, contains irrelevant information, important information mostly spread out, and also have co-referencing problems.

CONCLUSION

Opinion mining is an active area of research due to its usefulness and usage in several application domains. With the exponential growth of opinionated data, its analysis and summarization are becoming an essential task. To fulfil these needs, many approaches have been proposed. In this study, a survey of such approaches is presented. Techniques used is each step of opinion summarization are elaborated. Feature extraction techniques, sentiment polarity identification and summarization techniques are discussed in depth. This study indicates that for feature extraction, association rule mining is used widely. For sentiment classification, Naïve Bayes and lexicon-based approaches are extensively practiced. Mostly, the researchers used dataset of Amazon.com1, tripadvisor.com2 and Cnet.com3.

- **Future Research Direction**

Further research is required in the discovery of comparative opinions. In case of discussion forums, dealing with "noisy" input text is an issue. Powerful methods for extracting opinion phrases are required. The utilization of opinion mining in different applications i.e. recommendation systems, smart cities, transportation and traveling also needs attention. Due to several issues in language understanding and generation, the current abstractive opinion summarization systems cannot

[1]www.amazon.com

[2] www.tripadvisor.com

[3] www.cnet.com

provide the accuracy up to the level of customers' satisfaction as there are many complex sentences and expressions that are to be analyzed. Hybrid techniques should be developed to eliminate abstractive and extractive summarization problems. As the volume of data on web is very large, and it keeps on increasing day-by-day, scalability is a big challenge. Strong scalable techniques need to be developed.

REFERENCES

1. B. Liu, "Sentiment analysis and opinion mining," Synthesis lectures on human language technologies, vol. 5, pp. 1-167, 2012.
2. F. A. Pozzi, E. Fersini, E. Messina, and B. Liu, Sentiment Analysis in Social Networks: Morgan Kaufmann, 2016.
3. R. Prabowo and M. Thelwall, "Sentiment analysis: A combined approach," Journal of Informetrics, vol. 3, pp. 143-157, 2009.
4. E. Bilici and Y. Saygın, "Why do people (not) like me?: Mining opinion influencing factors from reviews," Expert Systems with Applications, vol. 68, pp. 185-195, 2017.
5. M. Afzaal, M. Usman, A. Fong, S. Fong, and Y. Zhuang, "Fuzzy Aspect Based Opinion Classification System for Mining Tourist Reviews," Advances in Fuzzy Systems, vol. 2016, p. 2, 2016.
6. F. Ali, D. Kwak, P. Khan, S. R. Islam, K. H. Kim, and K. Kwak, "Fuzzy ontology-based sentiment analysis of transportation and city feature reviews for safe traveling," Transportation Research Part C: Emerging Technologies, vol. 77, pp. 33-48, 2017.
7. K. Bauman, B. Liu, and A. Tuzhilin, "Aspect Based Recommendations: Recommending Items with the Most Valuable Aspects Based on User Reviews," in Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2017, pp. 717-725.
8. K. Bauman, B. Liu, and A. Tuzhilin, "Recommending Items with Conditions Enhancing User Experiences Based on Sentiment Analysis of Reviews," in CBRecSys@ RecSys, 2016, pp. 19-22.
9. Y.-D. Seo, Y.-G. Kim, E. Lee, and D.-K. Baik, "Personalized recommender system based on friendship strength in social network services," Expert Systems with Applications, vol. 69, pp. 135-148, 2017.
10. X. He, T. Chen, M.-Y. Kan, and X. Chen, "Trirank: Review-aware explainable recommendation by modeling aspects," in Proceedings of the 24th ACM International on Conference on Information and Knowledge Management, 2015, pp. 1661-1670.
11. I. Atoum and C. H. Bong, "A framework to predict software “quality in use” from software reviews," in Proceedings of the First International Conference on Advanced Data and Information Engineering (DaEng-2013), 2014, pp. 429-436.
12. W. Leopairote, A. Surarerks, and N. Prompoon, "Software quality in use characteristic mining from customer reviews," in Digital Information and Communication Technology and it's Applications (DICTAP), 2012 Second International Conference on, 2012, pp. 434-439.
13. R. Narayanan, B. Liu, and A. Choudhary, "Sentiment analysis of conditional sentences," in Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 1-Volume 1, 2009, pp. 180-189.
14. M. Ganapathibhotla and B. Liu, "Mining opinions in comparative sentences," in Proceedings of the 22nd International Conference on Computational Linguistics-Volume 1, 2008, pp. 241-248.
15. B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up?: sentiment classification using machine learning techniques," in Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10, 2002, pp. 79-86.
16. P. D. Turney, "Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews," in Proceedings of the 40th annual meeting on association for computational linguistics, 2002, pp. 417-424.
17. M. Gamon, A. Aue, S. Corston-Oliver, and E. Ringger, "Pulse: Mining customer opinions from free text," in Advances in Intelligent Data Analysis VI, ed: Springer, 2005, pp. 121-132.
18. E. Riloff and J. Wiebe, "Learning extraction patterns for subjective expressions," in Proceedings of the 2003 conference on Empirical methods in natural language processing, 2003, pp. 105-112.
19. V. Stoyanov and C. Cardie, "Toward opinion summarization: Linking the sources," in Proceedings of the Workshop on Sentiment and Subjectivity in Text, 2006, pp. 9-14.
20. T. Wilson, J. Wiebe, and R. Hwa, "Just how mad are you? Finding strong and weak opinion clauses," in aaai, 2004, pp. 761-769.
21. M. Hu and B. Liu, "Mining and summarizing customer reviews," in Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining, 2004, pp. 168-177.
22. B. Liu, M. Hu, and J. Cheng, "Opinion observer: analyzing and comparing opinions on the web," in Proceedings of the 14th international conference on World Wide Web, 2005, pp. 342-351.
23. Q. Mei, X. Ling, M. Wondra, H. Su, and C. Zhai, "Topic sentiment mixture: modeling facets and opinions in weblogs," in Proceedings of the 16th international conference on World Wide Web, 2007, pp. 171-180.
24. E. Marrese-Taylor, J. D. Velásquez, F. Bravo-Marquez, and Y. Matsuo, "Identifying customer preferences about tourism products using an aspect-based opinion mining approach," Procedia Computer Science, vol. 22, pp. 182-191, 2013.

25. Q. Su, X. Xu, H. Guo, Z. Guo, X. Wu, X. Zhang, et al., "Hidden sentiment association in chinese web opinion mining," in Proceedings of the 17th international conference on World Wide Web, 2008, pp. 959-968.
26. E. Cambria, S. Poria, R. Bajpai, and B. W. Schuller, "SenticNet 4: A Semantic Resource for Sentiment Analysis Based on Conceptual Primitives," in COLING, 2016, pp. 2666-2677.
27. E. Cambria and A. Hussain, Sentic computing: a common-sense-based framework for concept-level sentiment analysis vol. 1: Springer, 2015.
28. E. Cambria, S. Poria, A. Gelbukh, and M. Thelwall, "Sentiment analysis is a big suitcase," IEEE Intelligent Systems, vol. 32, 2017.
29. D. Sharma, "Study of Sentiment Analysis Using Hadoop," in Big Data Analytics, ed: Springer, 2018, pp. 363-376.
30. S. Bashir, W. Afzal, and A. R. Baig, "Opinion-Based Entity Ranking using learning to rank," Applied Soft Computing, vol. 38, pp. 151-163, 2016.
31. K. Krishnakumari and E. Sivasankar, "Scalable Aspect-Based Summarization in the Hadoop Environment," in Big Data Analytics, ed: Springer, 2018, pp. 439-449.
32. S. B. Mane, Y. Sawant, S. Kazi, and V. Shinde, "Real time sentiment analysis of twitter data using hadoop," IJCSIT) International Journal of Computer Science and Information Technologies, vol. 5, pp. 3098-3100, 2014.
33. Y. Madani, M. Erritali, and J. Bengourram, "Sentiment analysis using semantic similarity and Hadoop MapReduce," Knowledge and Information Systems, vol. 59, pp. 413-436, 2019.
34. B. Pang and L. Lee, "Opinion mining and sentiment analysis," Foundations and trends in information retrieval, vol. 2, pp. 1-135, 2008.
35. H. D. Kim, K. Ganesan, P. Sondhi, and C. Zhai, "Comprehensive review of opinion summarization," 2011.
36. W. Medhat, A. Hassan, and H. Korashy, "Sentiment analysis algorithms and applications: A survey," Ain Shams Engineering Journal, vol. 5, pp. 1093-1113, 2014.
37. D. R. Radev, E. Hovy, and K. McKeown, "Introduction to the special issue on summarization," Computational linguistics, vol. 28, pp. 399-408, 2002.
38. M. E. Moussa, E. H. Mohamed, and M. H. Haggag, "A survey on Opinion Summarization Techniques for Social Media," Future Computing and Informatics Journal, 2018.
39. E. Breck and C. Cardie, "39. Opinion mining and sentiment analysis," in Oxford Handbook of Computational Linguistics, 2nd ed: Oxford Publishing, 2012.
40. W. Maharani, D. H. Widyantoro, and M. L. Khodra, "ASPECT-BASED OPINION SUMMARIZATION: A SURVEY," Journal of Theoretical and Applied Information Technology, vol. 95, p. 448, 2017.
41. R. E. L. Condori and T. A. S. Pardo, "Opinion summarization methods: Comparing and extending extractive and abstractive approaches," Expert Systems with Applications, vol. 78, pp. 124-134, 2017.
42. B. Liu, "Sentiment Analysis and Subjectivity," Handbook of natural language processing, vol. 2, pp. 627-666, 2010.
43. A. Rashid, N. Anwer, M. Iqbal, and M. Sher, "A survey paper: areas, techniques and challenges of opinion mining," IJCSI International Journal of Computer Science Issues, vol. 10, pp. 18-31, 2013.
44. L. Zhang, S. Wang, and B. Liu, "Deep Learning for Sentiment Analysis: A Survey," arXiv preprint arXiv:1801.07883, 2018.
45. B. Seerat and F. Azam, "Opinion mining: Issues and challenges (a survey)," International Journal of Computer Applications, vol. 49, 2012.
46. B. Liu, Web data mining: exploring hyperlinks, contents, and usage data: Springer Science & Business Media, 2007.
47. K. Ravi and V. Ravi, "A survey on opinion mining and sentiment analysis: tasks, approaches and applications," Knowledge-Based Systems, vol. 89, pp. 14-46, 2015.
48. G. Vinodhini and R. Chandrasekaran, "Sentiment analysis and opinion mining: a survey," International Journal, vol. 2, pp. 282-292, 2012.
49. M. Z. Asghar, A. Khan, S. Ahmad, and F. M. Kundi, "A review of feature extraction in sentiment analysis," Journal of Basic and Applied Scientific Research, vol. 4, pp. 181-186, 2014.
50. M. Z. Asghar, S. Ahmad, A. Marwat, and F. M. Kundi, "Sentiment analysis on YouTube: A brief survey," arXiv preprint arXiv:1511.09142, 2015.
51. Y.-Y. Zhao, B. Qin, and T. Liu, "Sentiment analysis," Journal of Software, vol. 21, pp. 1834-1848, 2010.
52. S. ChandraKala and C. Sindhu, "Opinion Mining and sentiment classification a survey," ICTACT journal on soft computing, vol. 3, pp. 420-425, 2012.
53. J. A. Balazs and J. D. Velásquez, "Opinion mining and information fusion: a survey," Information Fusion, vol. 27, pp. 95-110, 2016.
54. M. Adedoyin-Olowe, M. M. Gaber, and F. Stahl, "A survey of data mining techniques for social media analysis," arXiv preprint arXiv:1312.4617, 2013.
55. I. Guellil and K. Boukhalfa, "Social big data mining: A survey focused on opinion mining and sentiments analysis," in Programming and Systems (ISPS), 2015 12th International Symposium on, 2015, pp. 1-10.
56. V. B. Raut and D. Londhe, "Survey on opinion mining and summarization of user reviews on web," International Journal of Computer Science and Information Technologies, vol. 5, pp. 1026-1030, 2014.
57. S. Sun, C. Luo, and J. Chen, "A review of natural language processing techniques for opinion mining systems," Information Fusion, vol. 36, pp. 10-25, 2017.
58. M. More and B. Tidke, "Social media online opinion summarization using ensemble technique," in International Conference on Pervasive Computing (ICPC), 2015., 2015, pp. 1-6.
59. S. alZahir, Q. Fatima, and M. Cenek, "New graph-based text summarization method," in Communications, Computers and Signal Processing (PACRIM), 2015 IEEE Pacific Rim Conference on, 2015, pp. 396-401.

60. V. Gupta and G. S. Lehal, "A survey of text summarization extractive techniques," Journal of Emerging Technologies in Web Intelligence, vol. 2, pp. 258-268, 2010.
61. K. Ganesan, C. Zhai, and J. Han, "Opinosis: a graph-based approach to abstractive summarization of highly redundant opinions," in Proceedings of the 23rd international conference on computational linguistics, 2010, pp. 340-348.
62. N. Garg, B. Favre, K. Reidhammer, and D. Hakkani Tür, "ClusterRank: a graph based method for meeting summarization," Idiap2009.
63. D. Alessia, F. Ferri, P. Grifoni, and T. Guzzo, "Approaches, tools and applications for sentiment analysis implementation," International Journal of Computer Applications, vol. 125, 2015.
64. I. Guyon and A. Elisseeff, "An introduction to feature extraction," in Feature extraction, ed: Springer, 2006, pp. 1-25.
65. I. Guyon, S. Gunn, M. Nikravesh, and L. A. Zadeh, Feature extraction: foundations and applications vol. 207: Springer, 2008.
66. T. A. Rana and Y.-N. Cheah, "Aspect extraction in sentiment analysis: comparative analysis and survey," Artificial Intelligence Review, vol. 46, pp. 459-483, 2016.
67. N. Kurian and S. Asokan, "Summarizing User Opinions: A Method for Labeled-data Scarce Product Domains," Procedia Computer Science, vol. 46, pp. 93-100, 2015.
68. M. K. Dalal and M. A. Zaveri, "Semisupervised learning based opinion summarization and classification for online product reviews," Applied Computational Intelligence and Soft Computing, vol. 2013, p. 10, 2013.
69. L. Zhuang, F. Jing, and X.-Y. Zhu, "Movie review mining and summarization," in Proceedings of the 15th ACM international conference on Information and knowledge management, 2006, pp. 43-50.
70. Y. Lu, C. Zhai, and N. Sundaresan, "Rated aspect summarization of short comments," in Proceedings of the 18th international conference on World wide web, 2009, pp. 131-140.
71. S. Bird, E. Klein, and E. Loper, Natural language processing with Python: analyzing text with the natural language toolkit: " O'Reilly Media, Inc.", 2009.
72. C. D. Manning, M. Surdeanu, J. Bauer, J. R. Finkel, S. Bethard, and D. McClosky, "The stanford corenlp natural language processing toolkit," in ACL (System Demonstrations), 2014, pp. 55-60.
73. J. Han, M. Kamber, and J. Pei, Data mining: concepts and techniques: Elsevier, 2011.
74. B. L. W. H. Y. Ma, "Integrating classification and association rule mining," in Proceedings of the fourth international conference on knowledge discovery and data mining, 1998.
75. S. A. Ahmed, P. Ravi, K. Paramesha, and K. Ravishankar, "AN EFFICIENT APPROACH FOR SENTIMENT ANALYSIS OF HEALTH POSTS," International Journal for Technological Research in Engineering vol. 3, 2016.
76. H. Kansal and D. Toshniwal, "Aspect based summarization of context dependent opinion words," Procedia Computer Science, vol. 35, pp. 166-175, 2014.
77. K. Bafna and D. Toshniwal, "Feature based summarization of Customers' Reviews of Online products," Procedia Computer Science, vol. 22, pp. 142-151, 2013.
78. M. Fan and G. WU, "Opinion Summarization of Customer comments," in International conference on Applied Physics and Industrial Engineering in, 2012.
79. D. Zhang, H. a. Dong, and J. Yi, "Opinion summarization of customer reviews," in International Conference on Automatic Control and Artificial Intelligence (ACAI 2012). 2012, pp. 1476-1479.
80. A. Madasu and S. Elango, "Efficient feature selection techniques for sentiment analysis," Multimedia Tools and Applications, vol. 79, pp. 6313-6335, 2020.
81. M. Hu and B. Liu, "Mining opinion features in customer reviews," in AAAI, 2004, pp. 755-760.
82. G. Li, Z. Lin, H. Wang, and X. Wei, "A Discriminative Approach to Sentiment Classification," Neural Processing Letters, vol. 51, pp. 749-758, 2020.
83. K. Ganesan, C. Zhai, and E. Viegas, "Micropinion generation: an unsupervised approach to generating ultra-concise summaries of opinions," in Proceedings of the 21st international conference on World Wide Web, 2012, pp. 869-878.
84. M. Potthast and S. Becker, "Opinion summarization of web comments," in Advances in Information Retrieval, ed: Springer, 2010, pp. 668-669.
85. P. Manivannan and C. K. Selvi, "Pairwise relative ranking technique for efficient opinion mining using sentiment analysis," Cluster Computing, vol. 22, pp. 13487-13497, 2019.
86. S. Kumar and M. Zymbler, "A machine learning approach to analyze customer satisfaction from airline tweets," Journal of Big Data, vol. 6, p. 62, 2019.
87. V. B. Raut and D. Londhe, "Opinion mining and summarization of hotel reviews," in International Conference on Computational Intelligence and Communication Networks (CICN), 2014., 2014, pp. 556-559.
88. D. K. Ly, K. Sugiyama, Z. Lin, and M.-Y. Kan, "Product review summarization from a deeper perspective," in Proceedings of the 11th annual international ACM/IEEE joint conference on Digital libraries, 2011, pp. 311-314.
89. K. A. Hasan, M. S. Sabuj, and Z. Afrin, "Opinion mining using Naïve Bayes," in IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE), 2015., 2015, pp. 511-514.
90. F. Ali, K.-S. Kwak, and Y.-G. Kim, "Opinion mining based on fuzzy domain ontology and Support Vector Machine: A proposal to automate online review classification," Applied Soft Computing, vol. 47, pp. 235-250, 2016.
91. N. Mukhtar and M. A. Khan, "Urdu Sentiment Analysis using supervised machine learning approach," International Journal of Pattern Recognition and Artificial Intelligence, p. 1851001, 2017.

92. A. Khan, U. Younis, A. S. Kundi, M. Z. Asghar, I. Ullah, N. Aslam, et al., "Sentiment classification of user reviews using supervised learning techniques with comparative opinion mining perspective," in Science and Information Conference, 2019, pp. 23-29.
93. A. Jain and V. Jain, "Sentiment classification of twitter data belonging to renewable energy using machine learning," Journal of Information and Optimization Sciences, vol. 40, pp. 521-533, 2019.
94. T. O'Keefe and I. Koprinska, "Feature selection and weighting methods in sentiment analysis," in Proceedings of the 14th Australasian document computing symposium, Sydney, 2009, pp. 67-74.
95. M. Devika, C. Sunitha, and A. Ganesh, "Sentiment Analysis: A Comparative Study on Different Approaches," Procedia Computer Science, vol. 87, pp. 44-49, 2016.
96. L. M. Rojas-Barahona, "Deep learning for sentiment analysis," Language and Linguistics Compass, vol. 10, pp. 701-719, 2016.
97. M. Zhang, Y. Zhang, and D.-T. Vo, "Neural Networks for Open Domain Targeted Sentiment," in EMNLP, 2015, pp. 612-621.
98. P. Kaladevi and K. Thyagarajah, "Integrated CNN-and LSTM-DNN-based sentiment analysis over big social data for opinion mining," Behaviour & Information Technology, pp. 1-9, 2019.
99. P. Wu, X. Li, S. Shen, and D. He, "Social media opinion summarization using emotion cognition and convolutional neural networks," International Journal of Information Management, vol. 51, p. 101978, 2020.
100. A. Abdi, S. M. Shamsuddin, S. Hasan, and J. Piran, "Deep learning-based sentiment classification of evaluative text based on Multi-feature fusion," Information Processing & Management, vol. 56, pp. 1245-1259, 2019.
101. A. Feizollah, S. Ainin, N. B. Anuar, N. A. B. Abdullah, and M. Hazim, "Halal Products on Twitter: Data Extraction and Sentiment Analysis Using Stack of Deep Learning Algorithms," IEEE Access, vol. 7, pp. 83354-83362, 2019.
102. L. Deng and D. Yu, "Deep learning: methods and applications," Foundations and Trends® in Signal Processing, vol. 7, pp. 197-387, 2014.
103. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, pp. 436-444, 2015.
104. T. Young, D. Hazarika, S. Poria, and E. Cambria, "Recent trends in deep learning based natural language processing," arXiv preprint arXiv:1708.02709, 2017.
105. R. Collobert, "Deep learning for efficient discriminative parsing," in Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics, 2011, pp. 224-232.
106. Q. T. Ain, M. Ali, A. Riaz, A. Noureen, M. Kamran, B. Hayat, et al., "Sentiment Analysis Using Deep Learning Techniques: A Review," International Journal of Advanced Computer Science and Applications (IJACSA), vol. 8, pp. 424-433, 2017.
107. S. Poria, E. Cambria, and A. Gelbukh, "Aspect extraction for opinion mining with a deep convolutional neural network," Knowledge-Based Systems, vol. 108, pp. 42-49, 2016.
108. X. Glorot, A. Bordes, and Y. Bengio, "Domain adaptation for large-scale sentiment classification: A deep learning approach," in Proceedings of the 28th international conference on machine learning (ICML-11), 2011, pp. 513-520.
109. C. N. Dos Santos and M. Gatti, "Deep Convolutional Neural Networks for Sentiment Analysis of Short Texts," in COLING, 2014, pp. 69-78.
110. C. N. dos Santos, "Think Positive: Towards Twitter Sentiment Analysis from Scratch," in SemEval@ COLING, 2014, pp. 647-651.
111. Y. Chen, "Convolutional neural network for sentence classification," University of Waterloo, 2015.
112. S. Poria, E. Cambria, and A. F. Gelbukh, "Deep Convolutional Neural Network Textual Features and Multiple Kernel Learning for Utterance-level Multimodal Sentiment Analysis," in EMNLP, 2015, pp. 2539-2544.
113. E. Cambria, D. Hazarika, S. Poria, A. Hussain, and R. Subramaanyam, "Benchmarking Multimodal Sentiment Analysis," CICLing, 2017.
114. O. Araque, I. Corcuera-Platas, J. F. Sánchez-Rada, and C. A. Iglesias, "Enhancing deep learning sentiment analysis with ensemble techniques in social applications," Expert Systems with Applications, vol. 77, pp. 236-246, 2017.
115. L.-W. Ku, Y.-T. Liang, and H.-H. Chen, "Opinion Extraction, Summarization and Tracking in News and Blog Corpora," in AAAI spring symposium: Computational approaches to analyzing weblogs, 2006.
116. M. Taboada, J. Brooke, M. Tofiloski, K. Voll, and M. Stede, "Lexicon-based methods for sentiment analysis," Computational linguistics, vol. 37, pp. 267-307, 2011.
117. F. Ali, E. K. Kim, and Y.-G. Kim, "Type-2 fuzzy ontology-based opinion mining and information extraction: A proposal to automate the hotel reservation system," Applied Intelligence, vol. 42, pp. 481-500, 2015.
118. S. Riaz, M. Fatima, M. Kamran, and M. W. Nisar, "Opinion mining on large scale data using sentiment analysis and k-means clustering," Cluster Computing, vol. 22, pp. 7149-7164, 2019.
119. H. Huang and C. Li, "A unified graph model for Chinese product review summarization using richer information," in Proceedings of the First International Workshop on Issues of Sentiment Discovery and Opinion Mining, 2012, p. 2.
120. G. Erkan and D. R. Radev, "LexRank: Graph-based lexical centrality as salience in text summarization," Journal of Artificial Intelligence Research vol. 22, pp. 457-479, 2004.
121. L. Zhu, S. Gao, S. J. Pan, H. Li, D. Deng, and C. Shahabi, "Graph-based informative-sentence selection for opinion summarization," in IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), 2013., 2013, pp. 408-412.
122. N. Franciscus, X. Ren, and B. Stantic, "Dependency graph for short text extraction and summarization," Journal of Information and Telecommunication, vol. 3, pp. 413-429, 2019.

123. M. Alfarra, A. M. Alfarra, and A. Salahedden, "Graph-based Growing self-organizing map for Single Document Summarization (GGSDS)," in 2019 IEEE 7th Palestinian International Conference on Electrical and Computer Engineering (PICECE), 2019, pp. 1-6.
124. A. Khan and N. Salim, "A review on abstractive summarization methods," Journal of Theoretical and Applied Information Technology, vol. 59, pp. 64-72, 2014.
125. M. Hu and B. Liu, "Opinion extraction and summarization on the web," in AAAI, 2006, pp. 1621-1624.
126. S. Blair-Goldensohn, K. Hannan, R. McDonald, T. Neylon, G. A. Reis, and J. Reynar, "Building a sentiment summarizer for local service reviews," in WWW Workshop on NLP in the Information Explosion Era, 2008, pp. 339-348.
127. Y. Wu, F. Wei, S. Liu, N. Au, W. Cui, H. Zhou, et al., "OpinionSeer: interactive visualization of hotel customer feedback," IEEE Transactions on Visualization and Computer Graphics., vol. 16, pp. 1109-1118, 2010.
128. B. Alper, H. Yang, E. Haber, and E. Kandogan, "Opinionblocks: Visualizing consumer reviews," in IEEE VisWeek 2011 Workshop on Interactive Visual Text Analytics for Decision Making, 2011.
129. H. Saggion and T. Poibeau, "Automatic text summarization: Past, present and future," in Multi-source, multilingual information extraction and summarization, ed: Springer, 2013, pp. 3-21.
130. P. Mehta, "From Extractive to Abstractive Summarization: A Journey," ACL 2016, p. 100, 2016.
131. S. Gupta and S. Gupta, "Abstractive summarization: an overview of the state of the art," Expert Systems with Applications, vol. 121, pp. 49-65, 2019.
132. E. Lloret and M. Palomar, "Analyzing the use of word graphs for abstractive text summarization," in Proceedings of the First International Conference on Advances in Information Mining and Management, Barcelona, Spain, 2011, pp. 61-6.
133. K. S. Thakkar, R. V. Dharaskar, and M. Chandak, "Graph-based algorithms for text summarization," in 3rd International Conference on Emerging Trends in Engineering and Technology (ICETET), 2010., 2010, pp. 516-519.
134. M.-T. Nguyen and M.-L. Nguyen, "Intra-relation or inter-relation?: Exploiting social information for Web document summarization," Expert Systems with Applications, vol. 76, pp. 71-84, 2017.
135. J. Zhao, S. Liu, and Y. Liu, "Towards Event Detection and Summarization on Microblogging Platforms," Advanced Science and Technology Letters, vol. 123, 2016.
136. C.-Y. Lin, "Rouge: A package for automatic evaluation of summaries," in Text summarization branches out: Proceedings of the ACL-04 workshop, 2004.
137. C.-Y. Lin and E. Hovy, "Automatic evaluation of summaries using n-gram co-occurrence statistics," in Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology-Volume 1, 2003, pp. 71-78.
138. A. Nenkova and R. J. Passonneau, "Evaluating Content Selection in Summarization: The Pyramid Method," in HLT-NAACL, 2004, pp. 145-152.
139. N. Mukhtar, M. A. Khan, and N. Chiragh, "Effective Use of Evaluation Measures for the Validation of Best Classifier in Urdu Sentiment Analysis," Cognitive Computation, pp. 1-11, 2017.