[Future Computing and Informatics Journal](https://digitalcommons.aaru.edu.jo/fcij)

Volume 3
Issue 1 *Volume 3, Issue 1, 2018*

Article 8

2018

A survey on opinion summarization technique s for social media

Mohammed Elsaid Moussa
Computer Science Dept., Faculty of Computers and Information, Helwan University, Cairo, Egypt, mohammed.elsaid@fci.helwan.edu.eg

Ensaf Hussein Mohamed
Computer Science Dept., Faculty of Computers and Information, Helwan University, Cairo, Egypt, ensaf_hussein@fci.helwan.edu.eg

Mohamed H. Haggag
Helwan University, mohamed.haggag@fci.helwan.edu.eg

Available online at www.sciencedirect.com

ScienceDirect

Future Computing and Informatics Journal 3 (2018) 82–109
http://www.journals.elsevier.com/future-computing-and-informatics-journal/

A survey on opinion summarization techniques for social media

Mohammed Elsaid Moussa*, Ensaf Hussein Mohamed, Mohamed Hassan Haggag

Computer Science Dept., Faculty of Computers and Information, Helwan University, Cairo, Egypt

Received 13 October 2017; revised 23 November 2017; accepted 19 December 2017
Available online 12 January 2018

Abstract

The volume of data on the social media is huge and even keeps increasing. The need for efficient processing of this extensive information resulted in increasing research interest in knowledge engineering tasks such as Opinion Summarization. This survey shows the current opinion summarization challenges for social media, then the necessary pre-summarization steps like preprocessing, features extraction, noise elimination, and handling of synonym features. Next, it covers the various approaches used in opinion summarization like Visualization, Abstractive, Aspect based, Query-focused, Real Time, Update Summarization, and highlight other Opinion Summarization approaches such as Contrastive, Concept-based, Community Detection, Domain Specific, Bilingual, Social Bookmarking, and Social Media Sampling. It covers the different datasets used in opinion summarization and future work suggested in each technique. Finally, it provides different ways for evaluating opinion summarization.

Keywords: Natural language processing; Sentiment analysis; Opinion summarization; Social media; Opinion mining; Tweet summarization

1. Introduction

Sentiment Analysis is a broad area that includes opinion mining, sentiment classification, and opinion summarization; Opinion Summarization is the process of automatically summarizing many opinions that are related to the same topic [1]. In sentiment analysis, Opinion Summarization involves many preprocessing steps such as tokenization, part of speech, stemming; making it different from traditional summarization [2]. It is one of the most valued and powerful NLP technologies [3]. In Social Media, it is about how to locate the most relevant posts with opinions to a given topic [4]. It will permit understanding hidden events and sentiments on different incidents [5]. Sentiment Summarization is also distinct from the factual data summarization, as sentences that were viewed as instructive from the factual point of view may not contain sentiment at all, making them useless from the sentiment perspective [6].

What makes this survey study important is that lately, there is an increased research interest in Opinion Summarization since it has turned into a pattern among individuals to give their sentiments on different features of products in blogs, review posts, and social networks [7]. As an example, on Amazon, some popular items could get thousands of reviews, making it hard for candidate clients to experience every one of the audits to settle on a choice to buy [8]. This large volume of data puts us in need for an automatic opinion summarization system and causes extraordinary challenges on the summarization system [9]. It would be useful for clients and manufacturers if the user reviews could be automatically processed and presented in a summarized form [10].

Opinion Summarization could be easily integrated into real-life applications, which will save users' time and effort [3]. For example, through Twitter opinions, politicians can review their public image and companies can check their customers'

* Corresponding author.
E-mail addresses: mohammed.elsaid@fci.helwan.edu.eg (M.E. Moussa), ensaf_hussein@fci.helwan.edu.eg (E.H. Mohamed), mohamed.haggag@fci.helwan.edu.eg (M.H. Haggag).
Peer review under responsibility of Faculty of Computers and Information Technology, Future University in Egypt.

https://doi.org/10.1016/j.fcij.2017.12.002

feedbacks [9]. It also plays a significant role in the semantic analysis of Social Media and Social Media Analytics [11].

There are two main approaches to generating textual summaries [12]: 1) **Extractive Summarization** in which the summary is consisting entirely of content extracted from the input [11] and 2) **Abstractive Summarization** in which the summary contains some content not present in the source for example as paraphrased material [11].

Developing opinion-tracking systems is commercially valuable [13]. As an example, Summly[1] (*a mobile app that sorts news by topics, and uses an analytical tool to condense text*) was sold to Yahoo for a reported $30 million US dollars[2] making its author Nick D'Aloisio (born November 1, 1995) one of the youngest self-made millionaires ever.

The structure of this survey is as follows: Section 1 gives an overview of the opinion summarization topic, its importance, and its generic approaches. Section 2 describes the scope and methodology we used while conducting this survey. Section 3 describes the current opinion summarization challenges for social media. Then in Section 4, we cover the steps to be done before the opinion summarization process itself, like preprocessing, Features Extraction, Noise Elimination, and handling of Synonym Features. Section 5 is the core of this survey, and it covers the various techniques used in Opinion Summarization. Section 6 shows the ways for evaluating opinion summarization. Finally, Section 7 concludes this survey along with the suggested future work to be done in opinion summarization over social media.

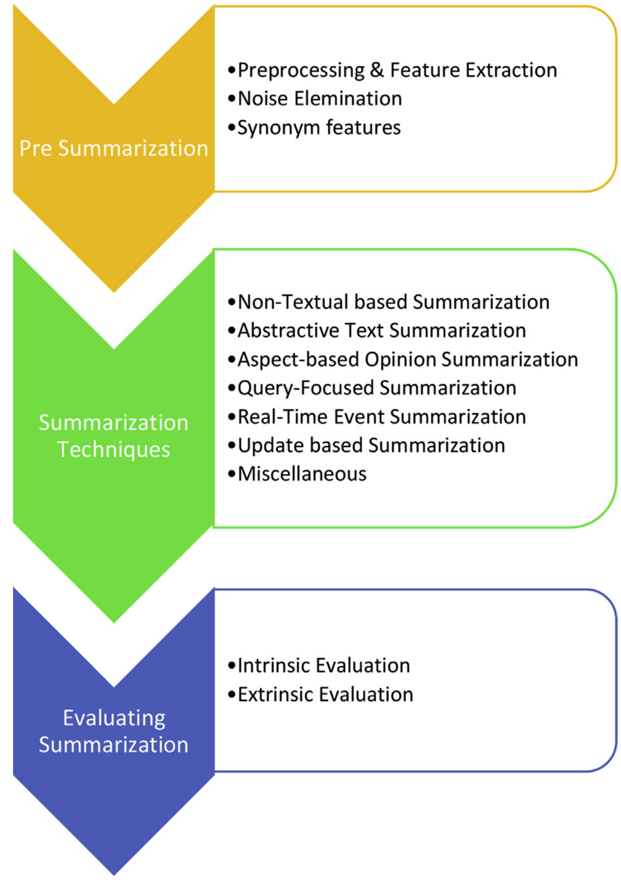


Fig. 1. Scope of the opinion summarization phases and techniques.

2. Scope and methodology

The goal of this paper is to conduct a comprehensive review of opinion summarization techniques for social media. During this survey, we did not get deep into equations and techniques used in summarization. Instead, we tried to cover many approaches used for opinion summarization and showing their techniques, domains, datasets, evaluation criteria, and their suggested future work if applicable, along with challenges and pre-summarization steps needed. So, one can see this survey as a significant horizontal covering of the topic, instead of the vertical covering of its techniques in depth.

We tried to cover most of the recent papers that focus on these keywords (social, sentiment or opinion, summarization) during the last few years. In few cases, some papers that cover only two of these three keywords were also used in this survey especially when the selected topic could be expanded easily to cover the third keyword. In Fig. 1 we see a schematic representation of the structure of the main survey sections.

3. Opinion summarization challenges

Navigating through all daily tweets to review crucial issues is very challenging [4]. Blogs and social media posts have no predefined rules, and they are profoundly unstructured, noisy, with a casual structure of dialect [7]. They often contain emoticons, sarcasm, and non-dictionary-standard words [12]. They are also composed by non-specialists with numerous mistakes in spelling, grammar, punctuations, and capitalization [7], with continuously changing conventions [11]. Which implies that we cannot utilize a lexicon or knowledge base *like Freebase,*[3] *Wikipedia*[4] *to find relevant events* [14] making it more difficult for analysis [12].

Additionally, unsupervised learning methodologies that were utilized for traditional text analysis had a weak performance when applied to social media posts and reviews due to its short texts nature [14].

Lately, new challenges appeared such as integration of social media data into document summarization [11], such as grouping synonym features when summarizing products features for sentiment analysis in products reviews [15]. Also for decision making, human evaluation of the data is so difficult due to its massive volume, and due to particularities of a given medium such as the characters limit in Twitter [11], redundancy, and irrelevant information caused by ambiguity in search keywords [14].

[1] http://summly.com/index.html.
[2] http://www.stuff.co.nz/technology/digital-living/8474218/Teens-multi-million-dollar-Yahoo-payday.
[3] https://www.freebase.com.
[4] https://www.wikipedia.org.

Social media Multi-document summarization also presents new challenges due to redundancy such as retweets, and use of extra features that could alter the significance of every message, such as counts of likes and favorites [16]. There is additionally need to manage the various, short, disparate, and boisterous nature of tweets [17].

Other challenges that are facing opinion summarization especially in commercial services include poor accuracy resulting from difficulties in the understanding language along with the scalability issue which requires deeper and complicated NLP technologies to be handled [18].

4. Pre-opinion summarization

In this section, we will cover the steps to be done before the opinion summarization process itself, like preprocessing, Features Extraction, Noise Elimination, handling of Synonym Features.

4.1. Preprocessing & feature extraction

A crucial phase in opinion summarization is feature extraction phase which simplifies the intricacy of the classification task by reducing the feature space [7]. Samples of the feature sets that are used in summarization as were used by Ref. [17] are **word-based features** such as cue words & phrases, non-cue words, abbreviations & acronyms, opinion words, vulgar words, and emoticons. And **symbol-based features** such as twitter-specific symbols and indicative punctuations. Other features that were used in Ref. [12] are listed in Fig. 2. Moreover, Fig. 3 illustrates general steps for pre-processing and feature generation (feature identification, feature extraction, feature refinement).

In 2015 [7], suggested an extensive preprocessing technique using some transformation and filtering tasks that need to be carried out before the phase of feature extraction. Their preprocessing steps include: Removing unwanted URL's, stop words, tags, special and repeated words, Handling abbreviations, overused and incorrect punctuation words, along with conflating repetitive symbols. Rao and Shah also provided an unsupervised technique for automating feature extraction task from a data set. Their technique includes: applying POS Tagging using Stanford Parser[5] [19], automating rule generation process, filtering extracted features based on frequency count, and finally refining the extracted features. Rao and Shah used user reviews from large websites such as Amazon, CNet,[6] Team-bhp,[7] and Car wale[8] that are belonging to mobile phones and cars. Their results showed a considerable reduction in the number of unnecessary features. A similar approach was used by Ref. [20] in their preprocessing step applied to some individual Arabic tweets to exclude unwanted features and noise, but besides removing of hyperlinks, hash letters, and redundant tweets, they added a text normalization phase following rules in Ref. [21].

In 2013, [12] also suggested a list of steps for preprocessing phase that include: replacing @username, #word, and target (of sentiment) with “ATUSER”, “word”, and “TARGET” respectively, removing URLs, replacing abbreviated slangs with their actual phrase equivalences, along with splitting each tweet into smaller snippets. In the same year, [17] proposed a collocation-based phrase extraction method which is resistant to accidental and un-conventionalized Twitter noise. Their method was based on extracting keywords as frequent nonstop words and extracting key phrases through finding frequent n-gram collocations.

4.2. Noise elimination

Noise includes all the irrelevant or low-quality information for summarization including spam, slang, non-English jokes, and sarcasm [22]. In 2014, [23] proposed a content-based approach to filter spam tweets. They utilized machine learning algorithms and compression-based text classifiers to filter the noisy/spam tweets. They also released their Twitter spam dataset as a public dataset along with a free WEKA[9] library for compression-based text filtering.

In 2012, [24] provided several noise elimination techniques to sift out spam and irrelevant tweets: First, they utilized Apache Nutch project's[10] language detector plus Twitter's provided language information to filter non-English tweets. Next, they filtered spam, useless tweets, and replies using a set of heuristics. Finally, they removed repeated characters using a normalization process [25] took a different approach when handling noisy data in short messages regarding grammar and the syntactic rules. They utilized Character N-Grams for this task which comprises the substrings of length n of the original text. The advantage of their model is its tolerance to noise and spelling mistakes which reduces the possibility of severe spelling mistakes.

4.3. Synonym features

Synonym features are the words or expressions that refer to the same characteristic, like “picture” and “photo” which refer to the same feature of the camera [15]. Synonym features should be grouped together because people often use different words or phrases to describe the same feature [26].

[27] Argued that some form of supervision is needed for handling synonym features the problem, since its solution depends on the user application requirements. In 2011, [15] proposed a technique to cluster and group synonym features, they showed that nor unsupervised learning or using the thesaurus dictionaries were doing great in this procedure. They reformulated the problem as a semi-supervised learning problem and utilized two soft constraints to label initial

[5] http://nlp.stanford.edu/software/lex-parser.shtml.
[6] http://www.cnet.com.
[7] http://www.team-bhp.com.
[8] http://www.carwale.com.
[9] http://www.cs.waikato.ac.nz/~ml/weka.
[10] http://nutch.apache.org.

f1	Document (or tweet) overall sentiment score using the unsupervised polarity detection algorithm
f2	Number of positive words
f3	Number of negative words
f4	Number of negation words
f5	Number of negation words followed by a positive word
f6	Number of negation words followed by a negative word
f7	Inverse sentiment
f8	Number of positive words followed by target
f9	Number of negative words followed by target
f10	Number of negation words followed by target
f11	Number of positive words followed by a negative word
f12	Number of negative words followed by a positive word
f13	Number of target words followed by a positive word
f14	Number of target words followed by a negative word
f15	Number of negation words followed by a positive word which is followed by target
f16	Number of negation words followed by a negative word which is followed by target

Fig. 2. Features used in sentiment analysis and summarization of twitter data [12].

examples automatically: Sharing words ("battery life", "battery", and "battery power"), and Lexical similarity ("movie" and "picture"). They also used positive and negative correlations to differentiate between the various feature expressions. They used the expectation-maximization EM algorithm formulated in Ref. [28], with few adjustments as in Fig. 4. Zhai et al. used five datasets for empirical evaluations (Home theater, Insurance, Mattress, Car, and Vacuum) and used Entropy and Purity for evaluating clustering. They suggested experimenting with other semi-supervised learning methods.

5. Opinion summarization techniques

Opinion mining and summarization procedure (Fig. 5) include three fundamental steps; Opinion Retrieval, Opinion Classification, and Opinion Summarization [30]. The opinion summarization process alone (Fig. 6) incorporates two methodologies (Feature-based summarization and Term Frequency based summarization) in which review text is preprocessed (sentence segmentation and tokenization) then each sentence score and relevance are calculated.

In 2015, [31] proposed a three steps model for opinion summarization system: Performing preprocessing on unstructured reviews, recognizing regular features using weighting association rule, discovering semantics of the sentiment words and lastly summarizing the outcomes.

5.1. Non-textual based summarization

It is worth mentioning that many of the previous works dealt with the word "summarization" differently. Some dealt with it with respect to visualizations or statistics only, and others dealt with it with respect to textual summarization.

For example, in 2014, in Kherwa et al.'s work [32] which provides an approach towards comprehensive sentimental data analysis and opinion mining; they focused on Visualizations by using Google Chart API[11] in Java to visualize the data in a user-friendly manner, as in Fig. 7. In the same year, [33] proposed a Retweeting Structure-aware Approach for Opinion

[11] https://developers.google.com/chart/.

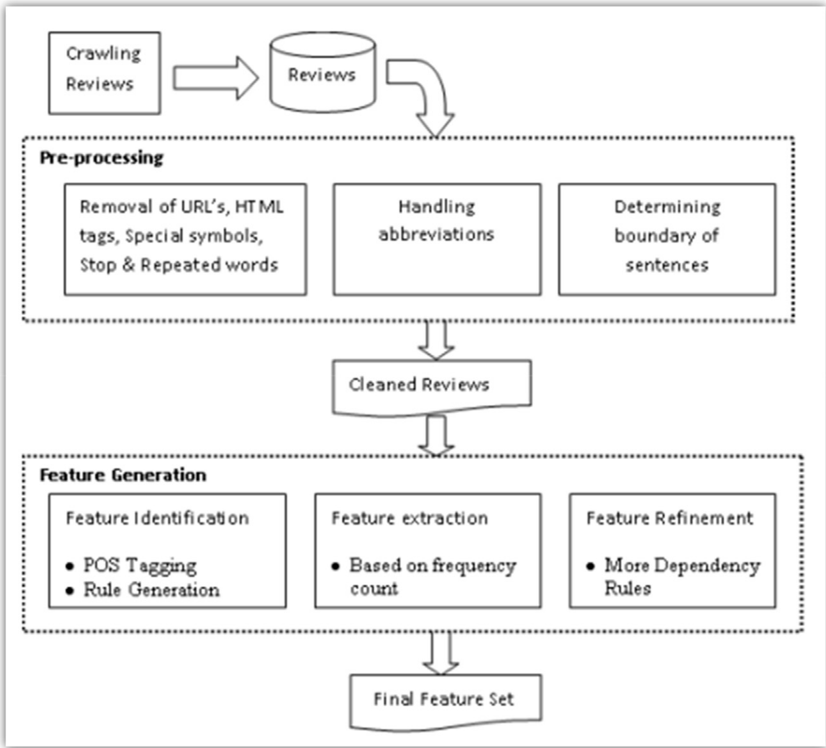


Fig. 3. Architecture for pre-processing and feature extraction steps [7].

```
Input: Labeled examples L; Unlabeled examples U
Output: Class labels of U
Extract SL from U using constraints;
Learn an initial naïve Bayesian classifier f
using L ∪ SL and Equations 1 and 2;
repeat
    // E-Step
    for each example d_i in U (including SL) :
            Compute P(c_j|d_i) using f and Equation 3.
    // M-Step
    Learn a new naïve Bayesian classifier f from L ∪ U
    by computing P(w_t|c_j) and P(c_j) using Equations 1 and 2.
until the classifier parameters stabilize
Classify examples in U to C using the final classifier f.
```

Fig. 4. Soft-constrained expectation maximization algorithm for synonym features [29].

Mining and Sentiment Analysis in Social Networks for website Weibo[12] (*a Twitter-like website in China*). It joins tree-like retweeting structure and examines feeling advancements with a comprehensive perspective. Their system demonstrates the sentiment propagation through opinion summarization chart. Their summarization part was made through an opinion summarization chart. Sample output from their system using real data from 2014 FIFA World Cup is shown in Fig. 8.

Also, the summarization part meant by Hsieh et al. [34] in 2012, which offers a Bilingual Context Mining and Sentiment

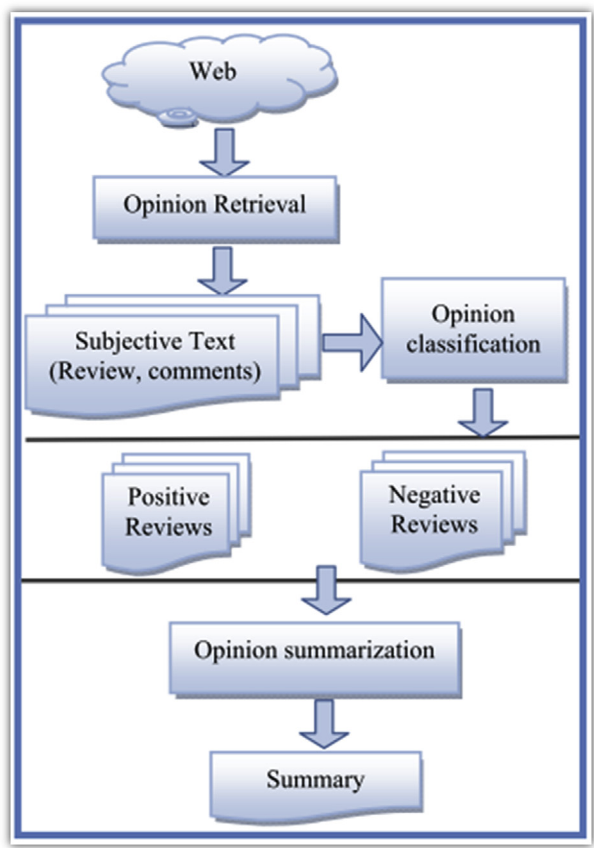


Fig. 5. Architecture of opinion mining and summarization [30].

[12] http://weibo.com.

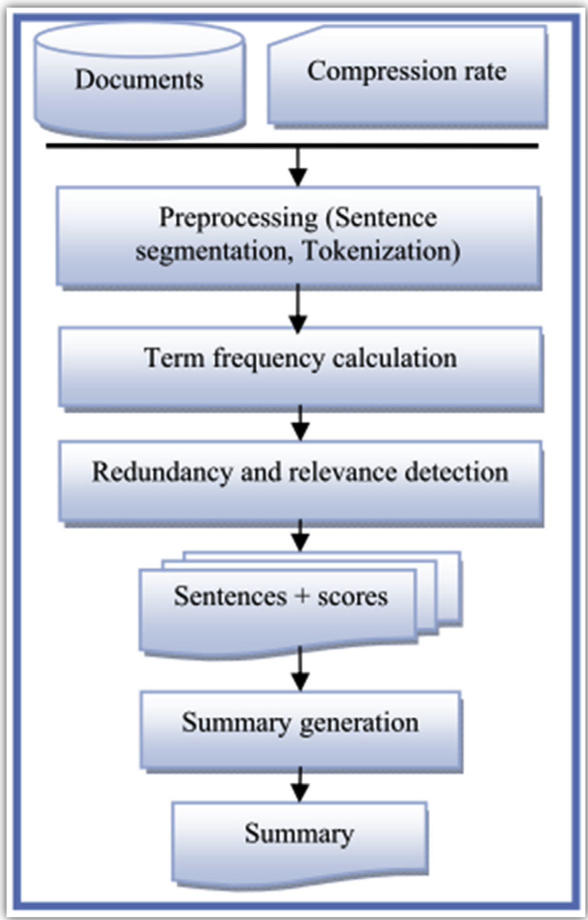


Fig. 6. Opinion summarization process [30].

Analysis Summarization System, is only graphical like in Fig. 9, in which they proposed S.E.R (Social Event Radar) technology. In 2006, [35] proposed graph-based summary along with a time series in their opinion tracking system which gives the trend of opinions from numerous data sources. Sample graph from their system with real data from president selection in Taiwan is represented in Fig. 10. Ku et al. used Data Set collected from TREC[13] corpus (Text REtrieval Conference), NTCIR[14] corpus and articles from web blogs. They suggested considering Opinion holders too in the future.

There is also more work that focuses mainly on statistics such as [10] which proposed an Opinion Summarization and Visualization System (OSVS) to present extracted information components in a graphical structure that encourage clients to have a speedy perspective of product features, and customers' feelings expressed over them. OSVS generates both bar and pie charts using Google Chart API. Fig. 11 shows the main screen of OSVS. In the same year, [36] proposed a methodology that takes into account Topic Modeling and Topic Phrase Mining for Comments Analysis and Visualization. They extracted and summarized Topics within comments as in graphs or clouds. As an example, Sample output with real comments from voyant-tool.org is shown in Fig. 12. Ramamonjisoa et al. suggested further assessment of the must-read comments plus using more criteria like time or semantic relation.

In 2014, [37] proposed an algorithm for calculating collaborated opinion value about the students, their summarization process was done by averaging sentiment scores. They proposed a case study based on the opinions extracted from the remarks given by the educators considering students' execution. Fig. 13 demonstrates a sample of a collaborated opinion score created by their framework.

We can conclude that visualization based summarization is very useful and has many advantages like visually appealing, the quick analysis of data, being easy to read, and concise representation of data. It could also be combined with other categories of summarization.

5.2. Abstractive text summarization

Unlike Extractive Summarization, which utilizes only complete sentences from the original wording, Abstractive Summarization incorporates a reformulation step and uses new terms [38]. As a result of difficulties in text generation, the abstractive summarization is not a common strategy [18]. And up till now the linguistic quality of abstractive summary is still far from satisfactory [39]. There are many techniques in this type of summarizations; these techniques include Template based, Graph-based, Semantic-based, Data Driven, Machine Learning, and Neural Networks.

As an example, for **Template based approach**, [17] proposed a technique for speech act-guided summarization by dealing with Speech acts recognition as a multi-class classification problem utilizing Support Vector Machine SVM for classification. They produced abstractive summaries by utilizing speech acts which catch the basic grounds of tweets from a communicative point of view. They related each tweet with a kind of five speech act types (statement, question, suggestion, comment, or the miscellaneous) as shown in Fig. 14, and generated synopses that incorporate the extracted language materials into speech act-based sentence templates. Fig. 15 showed one of their sample templates. [40] also, followed the same approach when they proposed a speech act-guided summarization approach. They utilized Bagging Ensemble approach with Naive Bayes Classifier; they recognized the speech acts in tweets, separated keywords, and phrases from the tweets, ranked them and embedded them into specific summary templates. Their proposed system architecture is shown in Fig. 16 which has a Content Planner to choose the information to be incorporated in the summary and a Linguistic Generator to pick the right words to express that information.

And for **Graph-based approaches** [41], examined a technique to create abstractive summaries by compressing and merging information from sentences based on word graphs. Their technique is based on extractive summarization initially to help in deciding which of the new sentences are

[13] http://trec.nist.gov.
[14] http://research.nii.ac.jp/ntcir/index-en.html.

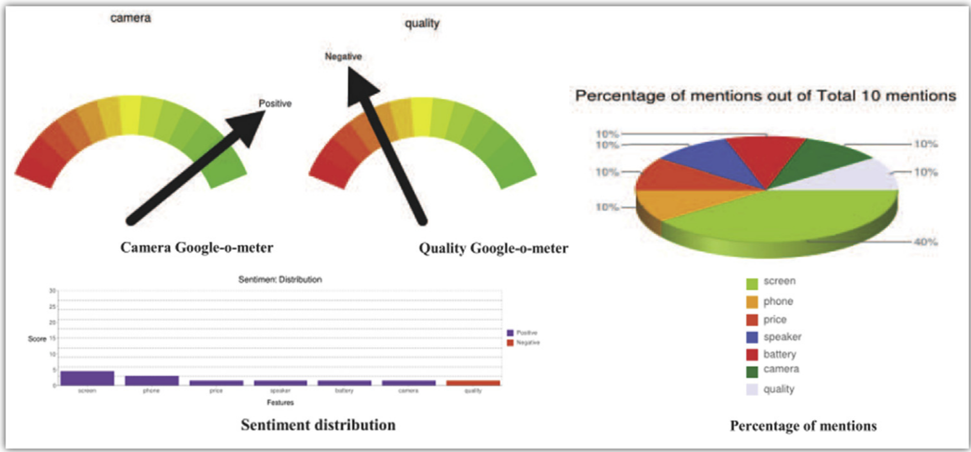
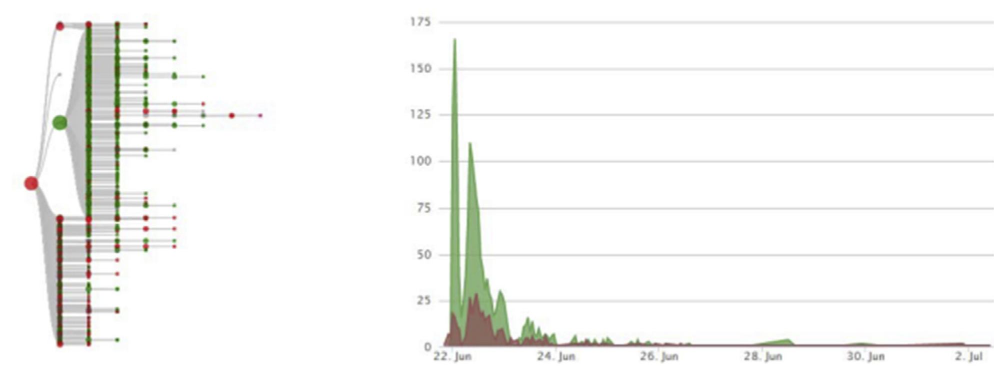


Fig. 7. Sample visualization for summarized features about a phone product [32].



(a) Sentiment propagation graph. The root node is the original tweet, while others are retweeting comments.

(b) Sentiment fluctuation chart. The horizontal axis shows the time line, and the vertical axis shows the number of retweeting comments with a certain sentiment orientation.

Fig. 8. A sample of real-world monitoring [33].

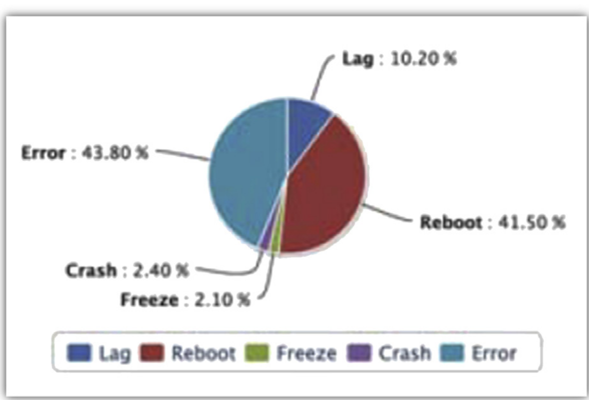


Fig. 9. Quick summary for software relevant issues [34].

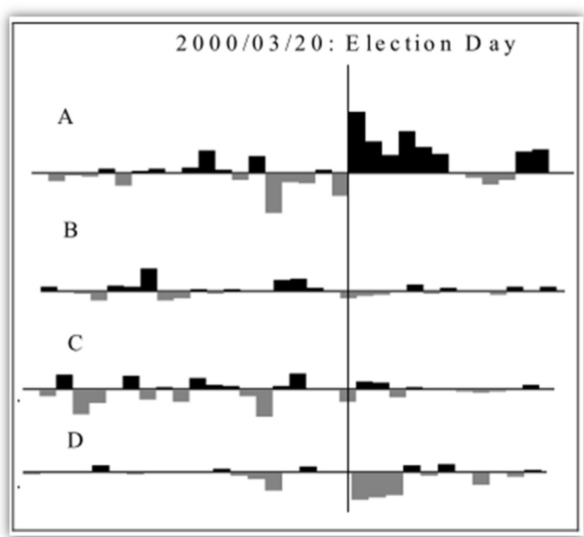


Fig. 10. Graph-based opinion tracking system sample [35].

more appropriate for incorporation in the final abstractive summary. A sample of their word-graph is shown in Fig. 17. [42] proposed a graph-based summarization framework (Opinosis) that creates succinct abstractive summaries of

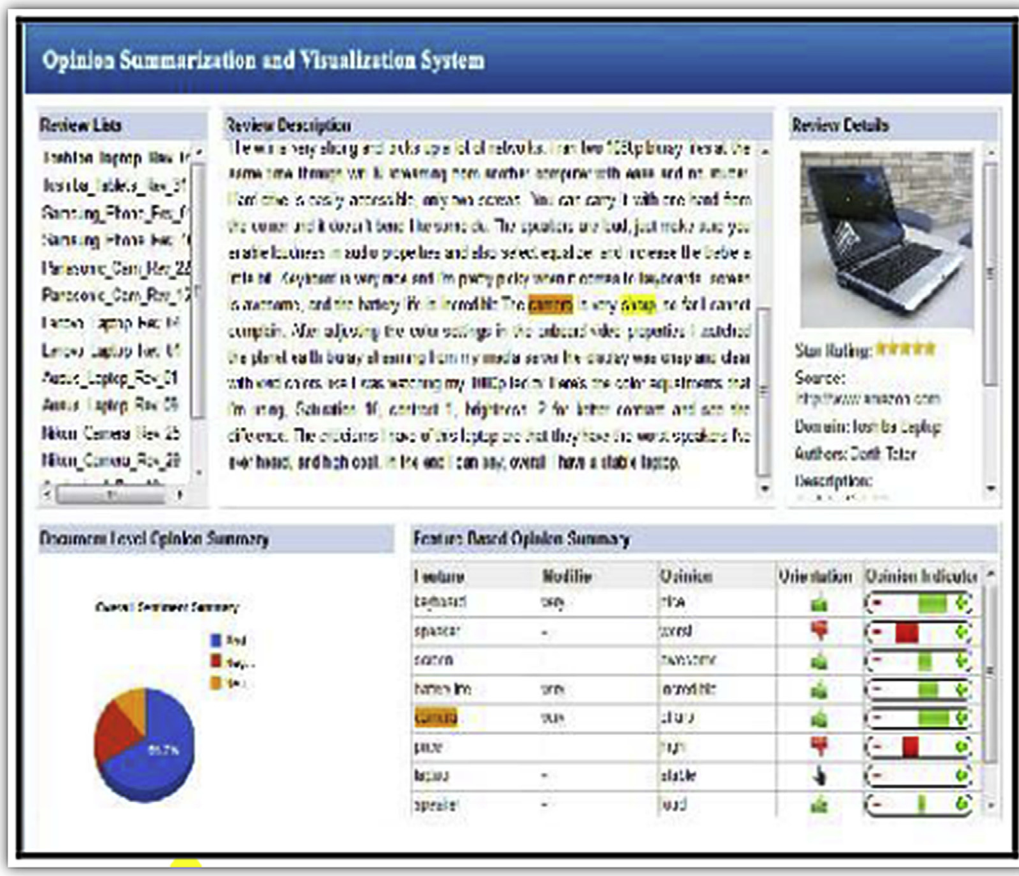


Fig. 11. Opinion summarization and visualization using OSVS [10].

Fig. 12. Summary unigrams example [36].

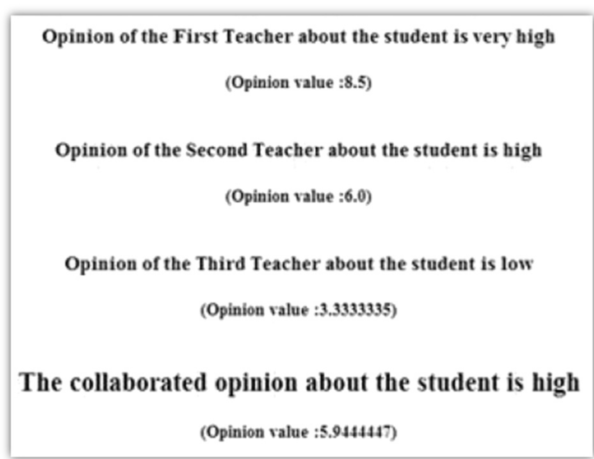


Fig. 13. Sample of collaborated opining [37].

highly redundant opinions, it utilizes shallow NLP and expects no domain knowledge. Their framework consists of four stages (Generation of a valid path, path scoring, collapsed paths, and finally the generation of the summary). Sample Opinosis graph which has three unique properties (Redundancy Capture, Gapped Subsequence Capture, Collapsible Structures) is shown in Fig. 18. The key idea of Opinosis is that it first builds a textual graph of the text to be summarized, then it explores and scores various sub-paths to be used in generating candidate abstractive summaries. One notable thing to mention regarding Opinosis approach is that it can be considered as Extractive summarization but with flavors of abstractive summarization.

Both of the approaches proposed by Refs. [41] and [42] has drawbacks that [43] handled in their graph-based technique that generates summaries of redundant opinions. It utilizes sentiment analysis to combine the statements through compressing and merging information based on word graphs. Since their technique uses sentiment analysis, it overcomes the drawback of [42] that is not being able to fuse sentences that

Types	Example Tweets
Statement	*Libya Releases 4 Times Journalists - http://www.photozz.com/?104k*
Question	*#sincewebeinghonest why u so obsessed with what me n her do?? Don't u got ya own man???? Oh wait.....*
Suggestion	*RT @NaonkaMixon: I will donate 10 $ to the Red Cross Japan Earthquake fund for every person that retweets this! #PRAYFORJAPAN*
Comment	*is enjoying this new season of #CelebrityApprentice.... Nikki Taylor = Yum!!*
Miscellaneous	*65. I want to get married to someone i meet in highschool. #100factsaboutme*

Fig. 14. Speech act types with examples [17].

For "<topic words>**", people** <verb frame> **"**<ngrams>**"**{**,** **(and)** <verb frame> **"**<ngrams>**"**}***.**

Fig. 15. Sample summary template for abstractive summarization [17].

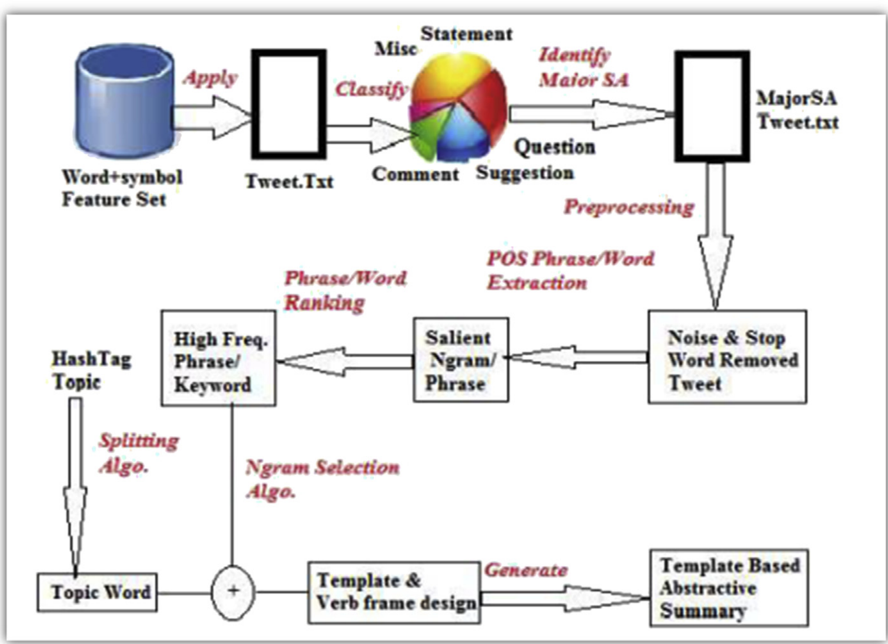


Fig. 16. Architecture of template-based abstractive summarization [40].

could be fused due do the absence a pre-existing connector. It also overcomes the drawback of [41] when a lot of valuable information is missing because of the used policy and grammar obligations, and it overcomes the redundant sentences issues too [43].'s method is based on three steps: building the word graph, confirming the sentence correctness, and getting abstractive summaries (scoring of paths based on the redundancy, fusing sentiments, then summarization and removing duplicate sentences using Jaccard index for similarity measure). Another graph based approach is [44] which

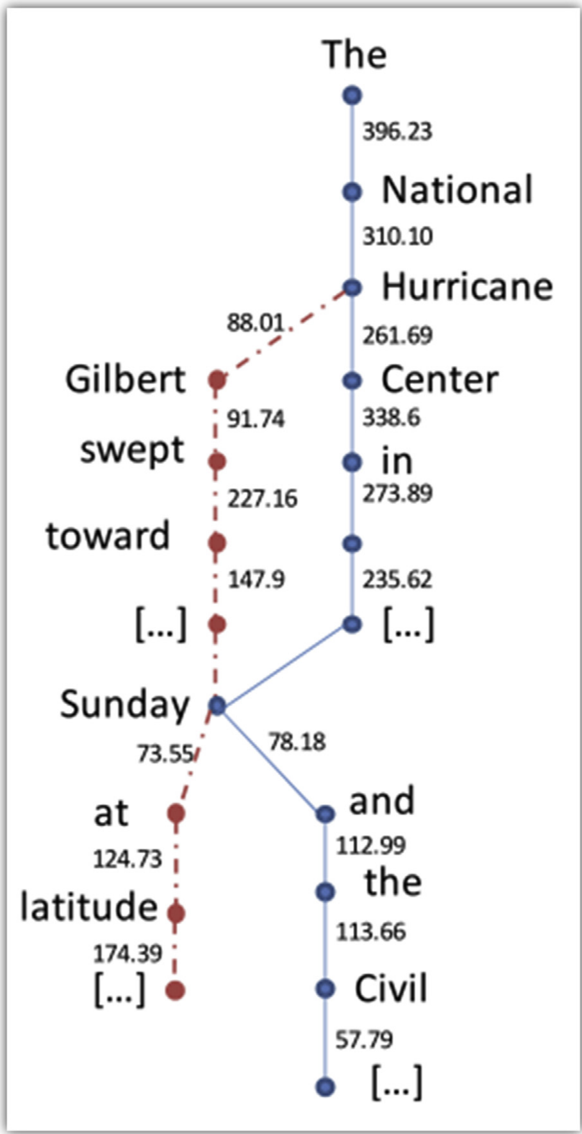


Fig. 17. Sample of word graph representation [41].

proposed an abstractive summarization system for product reviews by exploiting their discourse structure. They used a graph model (Fig. 19) based on the importance and association relations between aspects in the content selection phase of their framework. Their framework relies on the discourse structure and discourse relations of reviews to infer the significance of aspects and the association between them [45]. also, uses word graphs in their Multi-Document Abstractive Summarization approach using Integer Linear Programming ILP Based Multi-Sentence Compression, they initially recognize the most vital document in the multi-document set and align the sentences in the most important document to sentences in other documents. Then they generate K-shortest paths from the sentences using a word-graph structure (Fig. 20). And lastly, they employ integer linear programming (ILP) model with the objective of maximizing information content and readability of the final summary.

And for **Semantic-based or concept level approaches**, [3] proposed and assessed a concept-level methodology for ultra-concise opinion abstractive summarization. The stages of their abstractive summarization pipeline are shown in Fig. 21 in which they integrated sentence text simplification, text analysis, sentence regeneration, internal concept representation, concept analysis and summarization along with surface representation and sentence selection. [46] proposed a framework (Fig. 22) that means to select contents of summary not from the source document sentences but rather from the semantic representation of the source documents. It begins with semantic representation stage by utilizing semantic role labeling, then it groups semantically similar predicate argument structures, and lastly, it uses a genetic algorithm (GA) to rank the predicate argument structures based on weighted features. [47] proposed an ambitious framework for abstractive summarization, which aims at choosing the content of a summary from an abstract representation of the source documents. Their approach requires a semantic analysis of the text to generate the concept of Information Items (INIT) and incorporates many steps such as Semantic Role Labeling (SRL), predicate logic analysis of text, Word-sense disambiguation, coreference resolution, and analysis of word similarity. Their workflow diagram is represented in Fig. 23. [48] proposed a method that first fabricates a concepts pool from the input documents, then new sentences are made by picking and merging informative phrases to maximize the salience of phrases and at the same time satisfying the sentence construction constraints. They utilized integer linear optimization for this.

For **Machine Learning-based approaches**, [49] proposed an approach to generate abstractive ultra-concise summaries of opinions through unsupervised learning. They handled it as an optimization problem, where their aim is to find a succinct and unique set of phrases that are readable and represent key opinions in text. They used a greedy algorithm to solve this optimization problem by systematically exploring the solution space with heuristic pruning.

As an example of **Data-Driven approach**, Facebook AI Research [50] proposed a fully data-driven approach for abstractive sentence summarization (source code available[15]) named NAMAS. Their model is a neural attention-based model that depends on neural machine translation. They focused mainly on news headline-generation and combined a language model based on a standard feed-forward neural network language model (NNLM) [based on that of [51]] to estimate the contextual probability of the next word, with a contextual input encoder [based on that of [52]]. They used many encoders such as Bag-of-Words Encoder, Convolutional Encoder, and Attention-Based Encoder, and used GPUs for training.

A recent follow-up to NAMAS is proposed by IBM Watson [53] which has fascinating additional techniques and improves performance nicely. They cast abstractive text summarization

[15] http://facebook/NAMAS.

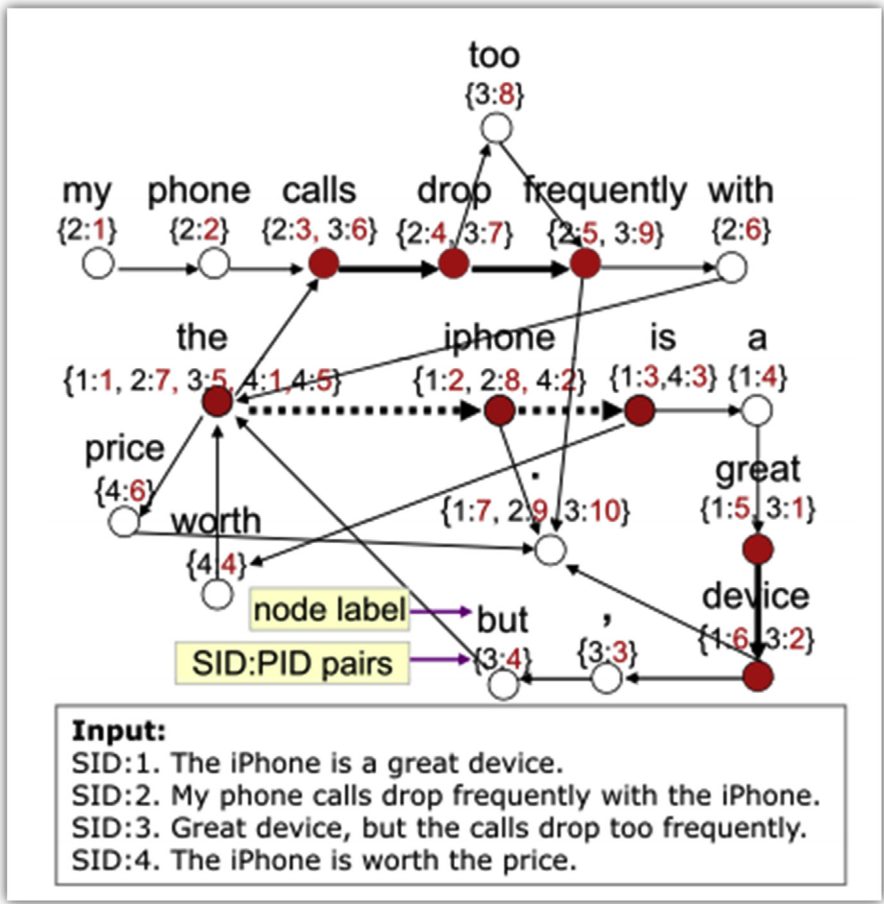


Fig. 18. Sample opinosis graph [42].

as a sequence-to-sequence problem and employed the framework of Attentional Encoder-Decoder **Recurrent Neural Networks** to this problem. They mapped an input sequence of words in a source document to a target sequence of words called summary. Their work is based on many models (1- Encoder-Decoder with Attention, 2- Large Vocabulary Trick, 3- Vocabulary expansion, 4- Feature-rich Encoder, 5- Switching Generator/Pointer, 6- Hierarchical Encoder with Hierarchical Attention). Each of these models addresses a specific problem in abstractive summarization, yielding further improvement in performance.

There are also tries to mix abstractive and extractive techniques together such as [41] who proved that the combination of extractive and abstractive information is a more suitable strategy to adopt towards the generation of abstracts. And [54] who built a **hybrid abstractive/extractive summarizer** combining natural language generation and salient sentence selection techniques. It first selects salient quotes from the input reviews and then embeds them into an automatically produced abstractive summary to either “provide evidence for”, “exemplify”, or “justify” positive or negative opinions.

Examples of datasets used in Abstractive Summarization are: **Opionosis Dataset** [42]; a topic oriented opinion sentences for cars, hotels and products (public dataset[16]), **TGSum Dataset** [55]; a multi-document summarization dataset guided by tweets (public dataset[17]), **Hu and Liu Dataset** [56]; customer reviews of twelve products obtained from Amazon.com and CNet (public dataset[18]), a large corpus of Chinese short text summarization dataset developed from the Chinese microblogging website Sina Weibo (public dataset[19]), Document Understanding Conference DUC datasets[20] [2001–2005] and Text Analysis Conference TAC datasets[21] [2010–2011], Gigaword Dataset,[22] Collected Tweets about latest trending topic on Twitter, Collected reviews from Amazon.com and WhatCar.[23]

Evaluations methods that are used in this category of summarizations are Recall-Oriented Understudy for Gisting Evaluation ROUGE [57], The Pyramid Method [58,59]. And

16 http://kavita-ganesan.com/opinosis-opinion-dataset.
17 http://www4.comp.polyu.edu.hk/~cszqcao/.
18 https://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html#datasets.
19 http://icrc.hitsz.edu.cn/Article/show/139.html.
20 http://duc.nist.gov/data.html.
21 https://tac.nist.gov/data/index.html.
22 https://catalog.ldc.upenn.edu/ldc2012t21.
23 http://www.whatcar.com.

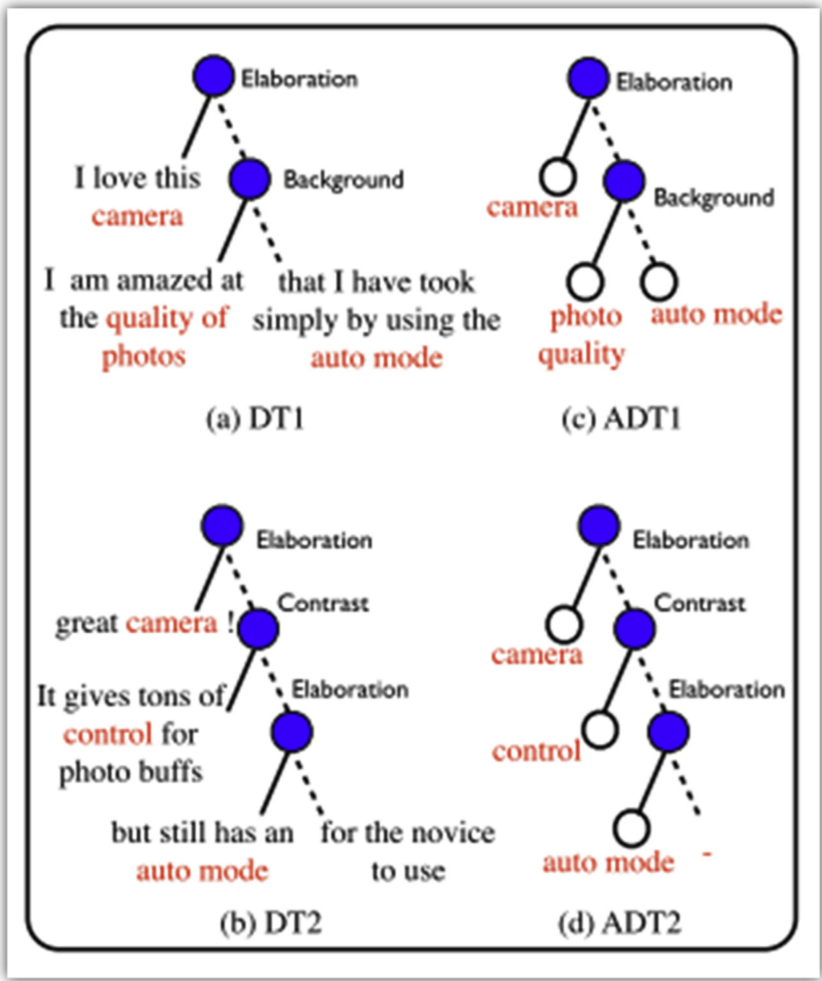


Fig. 19. Graph used in content selection phase in Ref. [44].

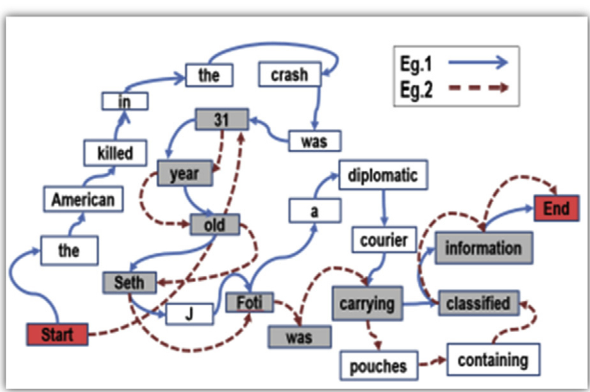


Fig. 20. Word-graph generation from sentences [45].

besides Precession, Recall, and F-Score measures that were used, there is also a source copy rate metrics which is a percentage of tokens in the system summary that occur in the source, to measure the percentage of abstraction (lower is better). Also, the DUC/TAC readability criteria which are Informativeness (the amount of information conveyed) and its readability (linguistic quality, grammaticality, non-redundancy, referential clarity, focus, structure, and coherence) is used. Researchers who work on Abstractive Summarization along with opinions also deal with Polarity (positive/negative/neutral) and Intensity (high/medium/low).

Some manual evaluations are based on pairwise preferences using crowdsourcing services such as crowdflower[24] [humans-in-the-loop] where raters were specifically instructed that their rating should express “overall satisfaction with the information provided by the summary”. Other manual evaluations used the Amazon Mechanical Turk crowd-sourcing system[25] [Human Intelligence Tasks]. And others used Microsoft N-gram service[26] which considered an approximate judge of how readable the system generated phrases are [42,49] also, introduced a readability test to understand if their system summaries are in fact readable, they are mixing sentences from their system summaries with sentences from human gold standard summaries and ask human assessors to pick sentences

[24] http://www.crowdflower.com.
[25] http://www.mturk.com.
[26] http://web-ngram.research.microsoft.com.

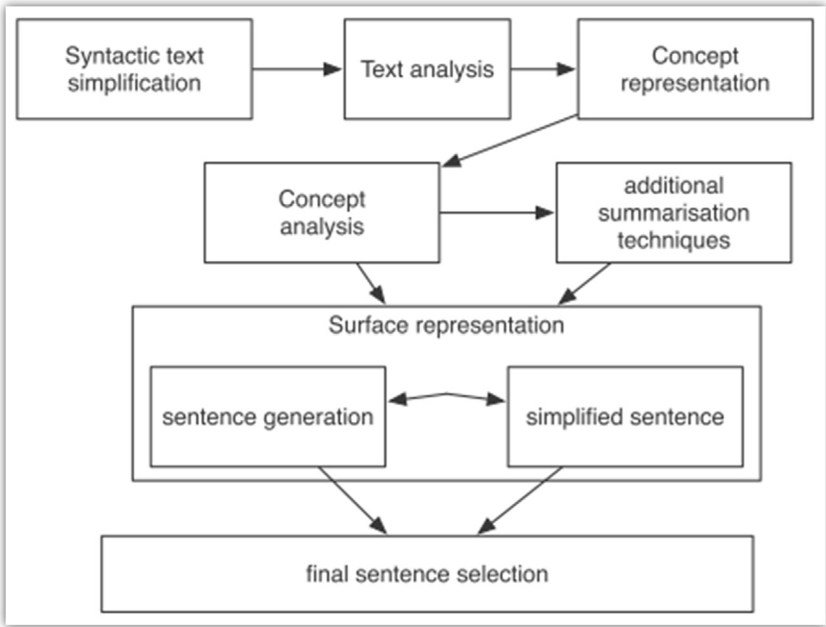


Fig. 21. Stages of abstractive summarization process as in Ref. [3].

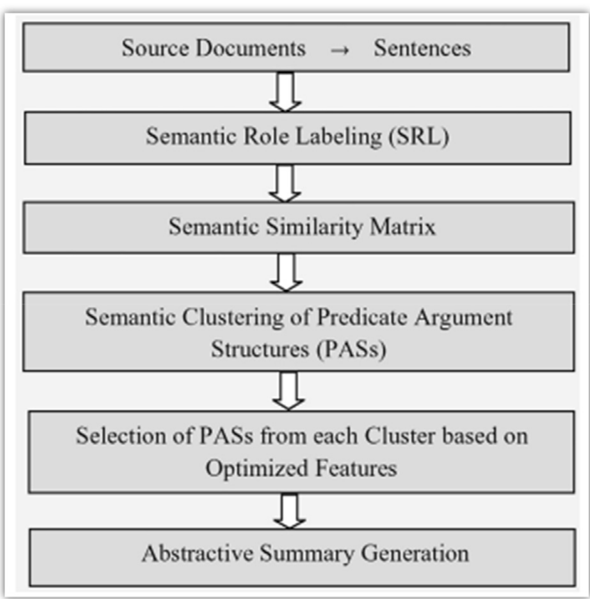


Fig. 22. Multi-document abstractive summarization based on semantic role labeling [46].

that are at least readable, then the system summaries were considered poor in readability if the human assessors pick them most of the time.

Although extensive researches have been done, the linguistic quality of abstractive summary is still far from satisfactory [39]. Future work in this category of summarization include: enhancing summary readability by incorporating the context of keywords and phrases during their extraction, enhancing quality by using paraphrasing techniques, enhancing coherence by addressing phrase level redundancies, building more robust models for summaries consisting of multiple sentences, enhancing grammar quality by improve grammar of the summaries in a data-driven way, enhancing time efficiency, Using of word sense disambiguation, deeper semantics, anaphora resolution, and normalizing the informal language.

Enhancements in Template-based methods include locking down individual nouns and verbs and applying proper verb conjugations such as [suggest "do your homework"] with the double quotes included in the generated summary => [suggest doing your homework], and automatically acquiring pattern matching rules for abstraction schemes by finding a way to gain pattern matching rules automatically, and then possibly reviewing and correcting them manually where needed.

Enhancements in Semantic Graph include integrating graph with Semantic Role Labeling to build a semantic graph for multi-document abstractive summarization, grouping sentences at a deep semantic level by using a similar idea to overlay parse trees, or by using Abstract Meaning Representation (AMR) graphs, plus redefining INITs (Information Items) so that they can be manipulated (compared, grouped, realized as sentences, etc.) more efficiently.

Another potential area of future research concerns: the ability to personalize summaries to the user's needs, Usage of Geographic Neighborhood, Merging Extractive and Abstractive techniques by using the pre-existing extractive techniques to improve recall and abstractive techniques to improve precision of the summary, Merging Abstractive and Update summarization techniques by tracking the same hashtag published on different dates, learning sentence compression techniques from Twitter.

Utilizing deep models in the process of automatic text summarization is still rare, but now researchers are very close

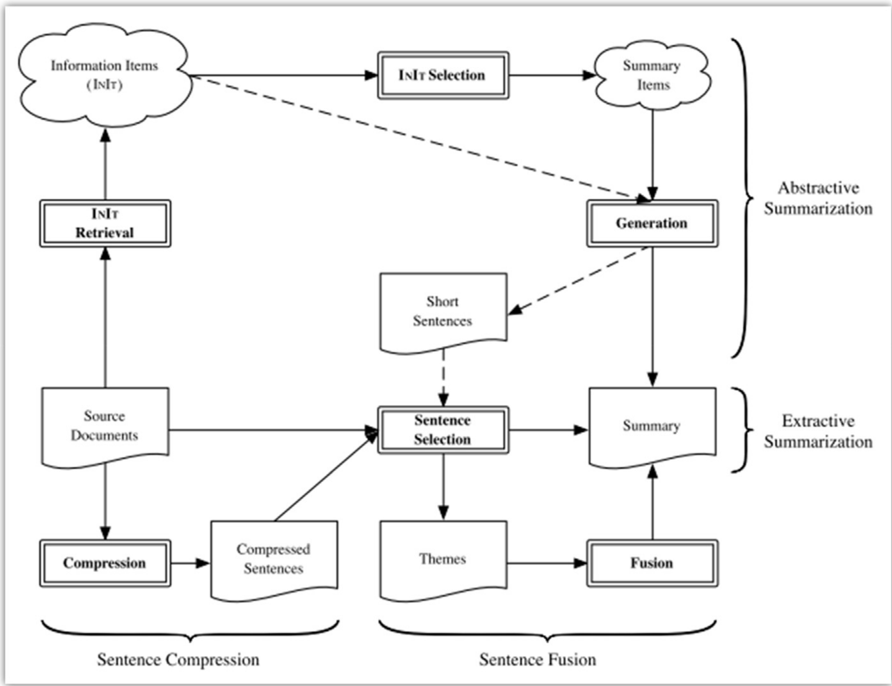


Fig. 23. Workflow diagram of abstractive summarization as in Ref. [47].

to generate abstractive summaries by utilizing the deep learning methods [39]. Enhancements include usage of a hierarchical Recurrent neural network RNN, fixing the rare word problem particularly when working with word-based input, using of neural generative models such as neural translation machine (NTM) [60].

5.3. Aspect-based opinion summarization (Aka) "feature-based opinion summarization."

An aspect-based summarization system could be considered as a multi-document summarization system which summarizes documents depending on different aspects or features of the target item [12]. Fig. 24 shows a brief explanation of the three steps in aspect-based summarization by Kim et al. [18]. Fig. 25 shows a general method for a Summarization of Product Reviews using the client's sentiment, feature occurrences, and the rate of review as proposed by Yang et al. [61] in 2009.

[8] Proposed a framework for opinion summarization based on sentence selection. They evaluated summary's quality regarding aspect coverage and viewpoints preservation. They also predicted review's helpfulness taking into consideration Redundancy and Coverage. They compared their approach with three other baseline approaches on Aspect coverage and Polarity distribution preservation. [12] used a hybrid method between Supervised and Unsupervised Polarity Detection techniques for product-based sentiment summarization of multi-documents which reduced some drawbacks of both technologies. They created many domain dependent lexicons and utilized topic detection algorithms to detect different domains. The advantages of their algorithm include the ability to identify newly added features, with fewer accuracy than manual detection as a disadvantage. [62] presented a feature-based opinion summarization approach which calculates a degree between 0 and 1 for the entire product and its features based on nouns, adjectives, verbs, and adverbs.

[9] Proposed an entity-centric topic-based opinion summarization framework, which remarkably accentuates the insight behind the opinions. They incorporated #hashtags as weakly supervised information into topic modeling algorithms and after that adopted Affinity Propagation algorithm to group #hashtags into coherent topics. [63] proposed a technique for Feature Evaluation for Sentiment Summarization (FESS) which intends to gather user evaluations of sentences on the relevance of these sentences. They find that opinionated content fragments that have supporting arguments are efficient to be used in the summary. [64] exhibited a framework that summarizes restaurants' and hotels' sentiment reviews. A general overview of the system is given in Fig. 26 in which their summarizer extracts a concise sampling of the reviews organized by aspect and sentiment and provides both quantitative and qualitative data at the aspect level.

Examples of datasets used in this category of summarizations include product reviews from Amazon.com, tweets regarding "iPhone", annotated by a group of human annotators, real-life reviews of electronics products with manual annotations, real-life tweets about people and brands, a list of forum posts about companies.

Future work in this area include Detection of mockery, finding more semantic features, building a multi-domain context dependent lexicon, enhancing current approaches to increase the recall, experimenting with other entities than products, giving normal text summarization on a given topic,

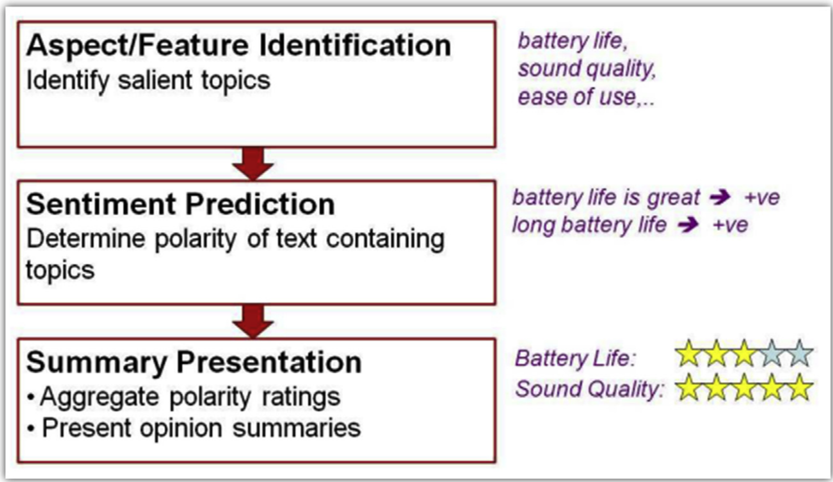


Fig. 24. General three steps of aspect-based opinion summarization [18].

Fig. 25. Summarization of product reviews using the user's opinion [61].

studying semantics underlying #hashtags, trying different corpora and languages, and finding combinations of sentences constituting a decent summary.

5.4. Query-Focused Summarization (Aka) "search based summarization."

Automatic summarization of user-generated online posts such as community QA and blogs present new challenges due to their much wider range of topics than product reviews [65]. These challenges include how to retrieve query relevant sentences, how to cover the main topics in the document, and how to balance these two requests [66].

[65] designed a framework for opinion summarization on community question answering and blog data. They proposed an objective function that considers relevance, coverage, and non-redundancy. [67] proposed a framework for dealing with this type of summarization based on sentence compression,

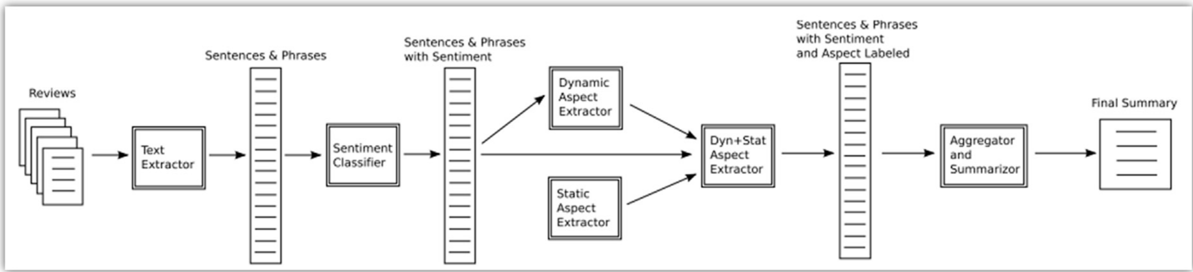


Fig. 26. A sentiment summarizer system for local service reviews [64].

their framework is consisting of three steps: Sentence Ranking for calculating sentences significance, Sentence Compression for producing succinic ranked sentences, and Post-processing for coreference resolution and sentence reordering. The special about their approach is that they join measures of query relevance, content importance, redundancy, and language quality into the compression stage in their tree-based compression function. [68] utilized some matrix factorization method, namely weighted Archetypal Analysis (wAA) to query-focused multi-document summarization (Fig. 27). They displayed synthetic data set as an undirected sentence similarity graph, where nodes represent sentences, edges represent similarity between connected nodes, and each sentence is connected to the given query like in Fig. 28.

[66] studied the query-focused multi-document summarization with focusing on relevance, coverage, and novelty. They proposed a Probabilistic-modeling Relevance, Coverage, and Novelty (PRCN) framework to model topic relevance and coverage; they built a set of features to describe relevance, novelty, and topic balance both from the document and from the query perspective. They also proposed a greedy topic balance algorithm for sentence ranking and extraction. [69] presented a system for exploratory search and topic summarization for Twitter called TweetMotif[27] (open source[28]); they assembled messages by frequent significant terms. TweetMotif extracts a set of subjects to cluster and summarizes these messages. [70] proposed a query-specific opinion summarization system QOS. Their system (shown in Fig. 29) gives back a summary with pertinence to the sentiments and target portrayed by the input question. They utilized an LSI-based strategy to score sentence concerning the query, a lexicon-based method to determine the opinion orientation of a sentence, and diversity penalty for redundancy removal.

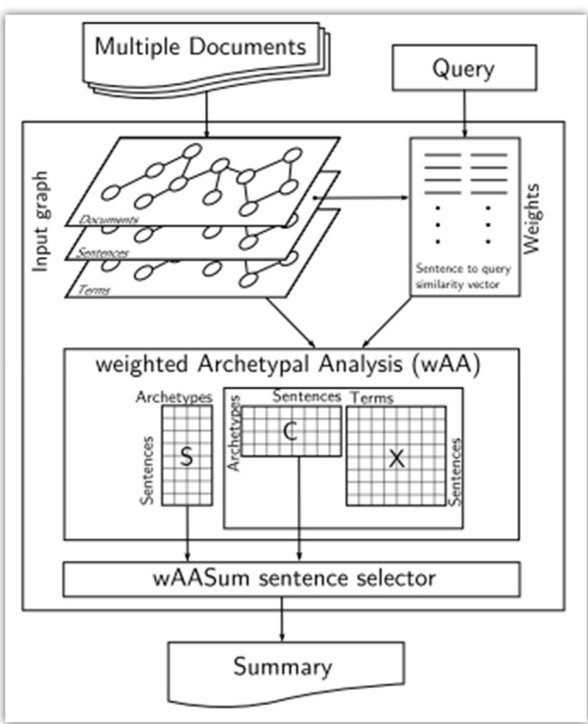


Fig. 27. Query-focused MDS method using wAA [68].

Examples of datasets used in this category of summarizations include Yahoo! Answers dataset from Yahoo! Webscope program, DUC 2005 [71], DUC 2006 [72], DUC 2007 [73], TAC 2008 corpus [74], The dataset from Ref. [75], and miscellaneous topics that fulfill the following criteria (Frequency contrast, Topic diversity, Topic size, and a small number of topics).

Future work in this area include improving the performance of weighted Archetypal Analysis wAA, using sophisticated methods for the query processing/expansion techniques, using WordNet to calculate the semantic similarity between sentences, introducing the multi-layered graph model that emphasizes relations such as n-grams, phrases and semantic role arguments levels, applying query-focused summarization techniques to other summarization tasks such as update and comparative summarization. Developing models based on LDA (Latent Dirichlet Allocation) for document summarization. And studying the opinion ranking and summarization methods, particularly in the opinion search applications.

5.5. Real-time event summarization (Aka) "sub-events based summarization."

An event alludes to any concept of interest that picks up the consideration of the masses. Examples range from worldwide disasters such as earthquakes, political protests or unrest, to

[27] http://tweetmotif.com/about.

[28] https://github.com/brendano/tweetmotif.

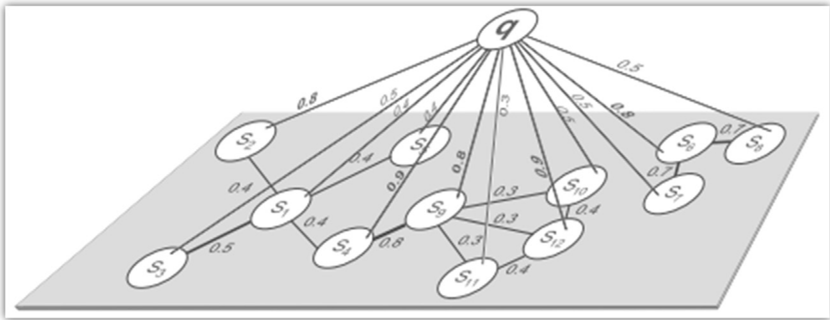


Fig. 28. Sentence similarity graph [68].

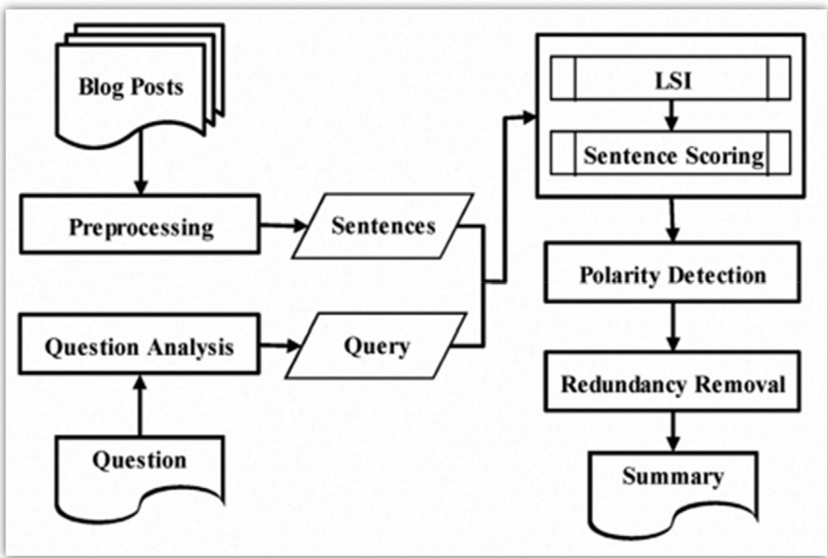


Fig. 29. Query-specific opinion summarization system [70].

dispatches of new customer items [14]. In social media, the real-time event search and the requirement for event recognition raise a critical issue [76]. Real-Time events could be either scheduled events or non-scheduled events.

Amid events like sports games, numerous social media posts are sent describing feelings about these events. Creating sensible summaries of events' critical moments could impart what happened to people who could not track it [24]. Event Summarization intends to produce a representative and succinct textual description of these scheduled events, providing people with other options for viewing the world beyond the conventional journalism [77]. Real-Time summarization includes two operations: **First**, finding vital moments amid event like sub-events, **Second**, discovering few essential tweets that best describes the distinguished sub-event group [78].

Scheduled events summarization gained much attention, and there is much work in this area. [78] performed a Real-Time Summarization of scheduled sub-events for game tournaments like football and cricket from Twitter stream. They presented a New Event Detection (NED) system that operates online with social streams. On the other hand, [16] handled the Rashomon effect [79], by creating subjective summaries for conflicting comments for the same events. Their system summarizes the same event from two or more contradictory viewpoints. They used a human intrinsic summarization evaluation approach. [80] proposed an approach (showed in Fig. 30) that continuously summarizes the real-time tweets stream, by determining first if something new has happened then picking a delegate tweet to depict each sub-event. Their proposed method does not depend on any external knowledge about soccer events making their approach easily applied to other sorts of scheduled events without requiring extra information. They evaluated tweets as Correct, Novel, or Noisy based on manual evaluations.

There are also studies that cover Non-Scheduled events such as [14] which proposed an approach for summarizing non-scheduled events; they introduced two topic models that exploit of data temporal correlation. Their approach depends on topics and modeling n-grams.

Examples of datasets used in this category include tweets from different matches, including FIFA2015 final match, tweets about FA Cup finals of 2012 and 2013 (BBC live text commentaries to be the ground truth), tweets regarding the

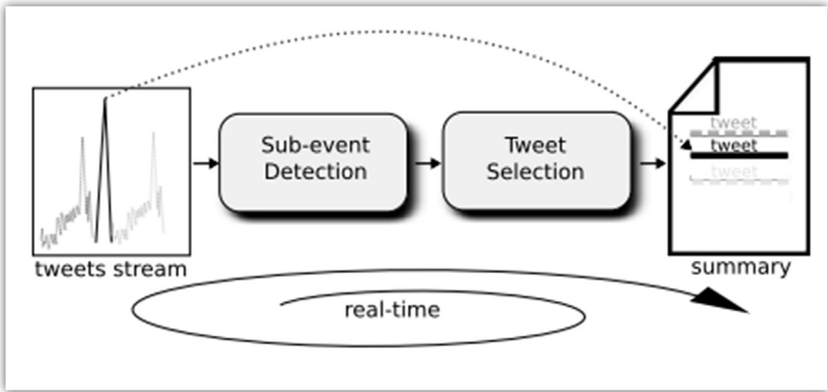


Fig. 30. Two-step process for real-time event summarization [80].

Copa America 2011 championship, and real-world events such as natural disasters, governmental issues, and organization events.

Future work in this area include using paraphrases, handling Out-Of-Vocabulary words, creating a topic-conditioned classifier, identifying feelings regarding events, extending to different areas, for example, governmental issues, alongside with enhancing calculations to recognize which group every fan underpins, assessing technique on different sorts, and working with personalized summaries and factual tweets.

5.6. Update based summarization

Update Summarization is a genuinely recent topic connecting news summarization to online and dynamic settings [81]. As defined at Text Analysis Conference (TAC)2008[29] the update summarization task is to create a 100-word outline of a successive set of some articles for the same point, under the supposition that the reader has already read a given set of earlier documents [74]. Update Summarization related to topic detection and tracking (TDT) which aims to find the latest news report through novelty detection and relevance discrimination [82]. Its goal is to develop systems for efficiently monitoring events' information over time [83].

[84] proposed a generative hierarchical tree model (HTM) based on Hierarchical Latent Dirichlet Allocation (hLDA) for update summarization. They additionally proposed a summary ranking approach that considers various aspects like focus, novelty, and non-redundancy. [85] presented a multi-level summarization framework which extracts sequential update summarization on sudden events and incorporates topic-level and sentence-level summarization technologies.

[86] proposed an Incremental update summarization (IUS) approach that adaptively changed based on the prevalence and novelty of discussions about the event. It joins summarization techniques and supervised regression model. Fig. 31 gives a sample of an IUS system over a time ordered document stream, and Fig. 32 shows how prevalence and novelty map to various conditions of an event. [87] proposed a system for sequential summarization for Twitter trending topics. They proposed a Heuristic (stream-based) approach and an LDA-based (semantic-based) approach for distinguishing subtopics and extracting critical tweets. Their measures for evaluation include coverage, novelty, and correlation. Their evaluations demonstrate that the stream/semantic combination approach perform superior to other approaches.

[82] proposed an extractive content selection framework for update summarization. They built an evolutionary manifold-ranking model which utilizes the iterative feedback mechanism. Likewise, they combined the normalized spectral clustering with evolutionary manifold-ranking to have an ideal sentence determination. Also, they explored a redundancy removal strategy with exponent decay. [88] presented a unified framework for both standard and update summarization, which embraces a topic modeling approach for salience determination and a dynamic modeling approach for redundancy control. Fig. 33 illustrates their approach; they also extended their framework to Chinese multi-document summarization. [89] proposed a co-ranking strategy to address the update summarization task. They integrated two co-ranking processes by including strict constraints. [90] proposed a methodology considering a three-level Hierarchical Dirichlet Process HDP model for update summarization. Their model can detect the birth, splitting, merging, and death of specific aspects and the general foundation data on a given topic.

[91] proposed a Clustering Based Sentence Extractor for Automatic Summarization (CBSEAS) that integrates a specialized method to detect redundancy for enhancing informational diversity. [92] proposed a summarization method based on an incremental hierarchical clustering framework to update summaries as soon as a new document arrives. Their system produces a sentence hierarchical tree to exhibit the complete structure of the documents, and a summary of contents at the current time point is created. Fig. 34 demonstrates their framework. [93] proposed a document

[29] http://www.nist.gov/tac/2008/summarization/update.summ.08.guidelines.html.

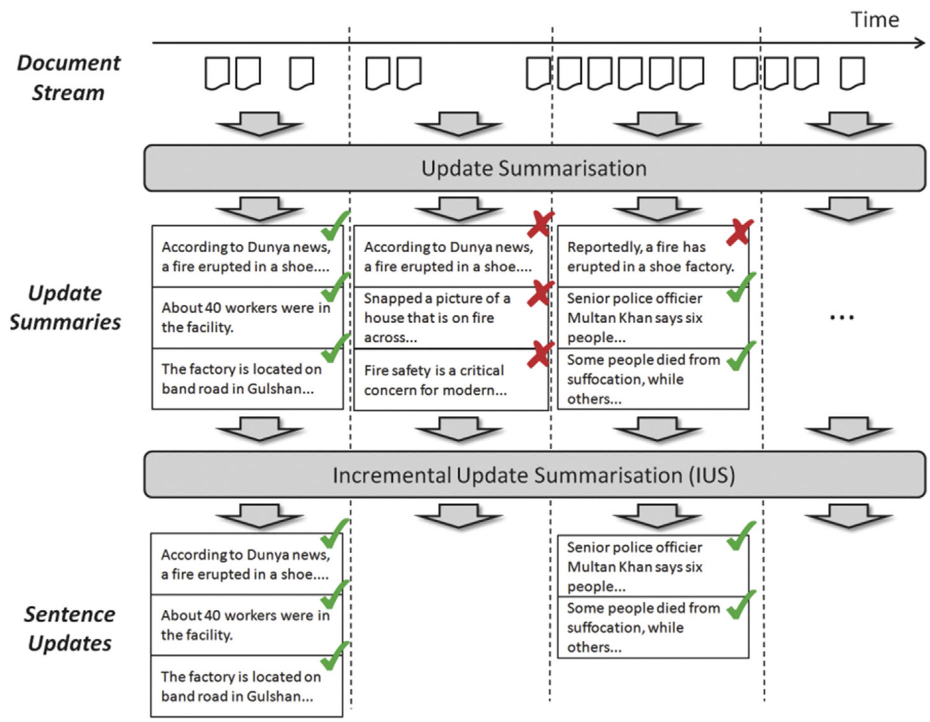
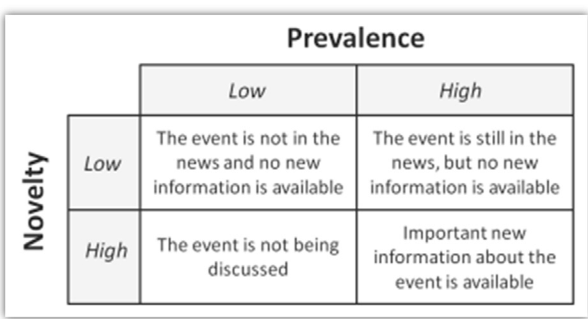


Fig. 31. Illustration of IUS task over a time-ordered document stream [86].



Novelty \ Prevalence	Low	High
Low	The event is not in the news and no new information is available	The event is still in the news, but no new information is available
High	The event is not being discussed	Important new information about the event is available

Fig. 32. Prevalence and novelty relations to different event states [86].

summarization hypothesis based on the theory of information distance. They utilized two approximation methods to gauge data distance (compression, and semantic element extraction).

[94] proposed a summarizer based on the latent semantic analysis (LSA) and proposed the update summarization component which decides the redundancy and novelty of each topic discovered by LSA. The idea of their approach is to use LSA for the creation of a set of subjects contained in the set of recent documents. Then, to indicate their redundancy, novelty, and significance. Finally, the summary is created from the sentences that contain more novel and significant topics.

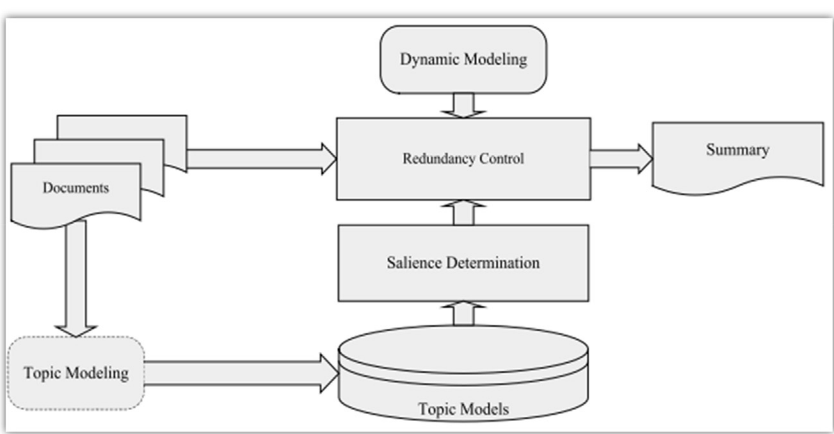


Fig. 33. The unified multi-document summarization framework [88].

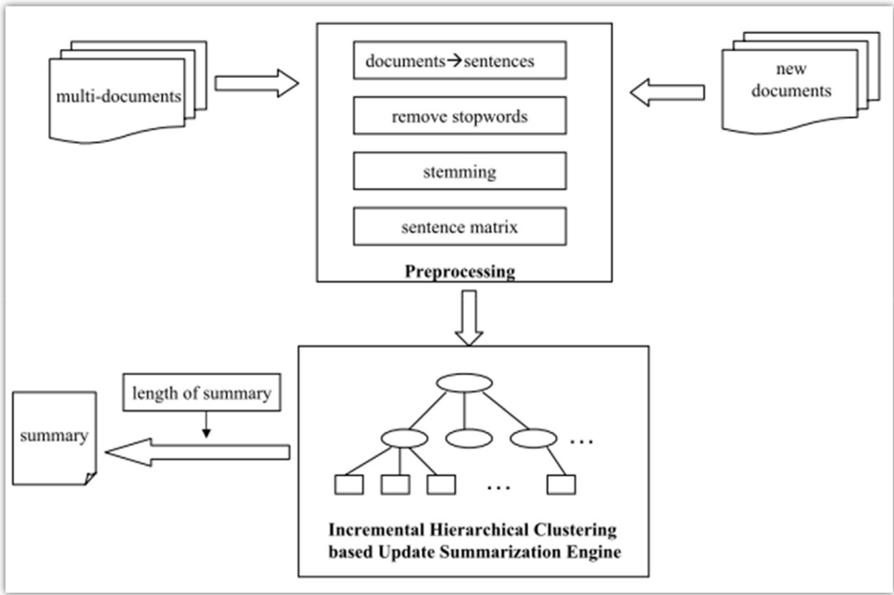


Fig. 34. Update summarization based on incremental hierarchical clustering [92].

Steinberger and Jezek [95] proposed an Update Summarizer, which depends on Iterative Residual Rescaling (IRR). The idea of this approach is the same as in Ref. [94] but with showing the IRR generalization of the basic LSA summarization model that makes the LSA representation of topic/sentence distribution more reliable.

Examples of datasets used in this category of summarizations include DUC2007[30] corpus, TAC (2008–2011) datasets, the dataset in the Sequential Update Summarization (SUS) task of Temporal Summarization (TS) track at Text REtrieval Conference (TREC 2013). Tweets built using Twitter APIs[31] by tracking Twitter official trending, Hurricane Wilma dataset. And for evaluations, the most used evaluation method is ROUGE [57] with human-generated summaries as references, PYRAMID [59], and BE [96], alongside some manual assessments.

Future work in this area includes increasing sensitivity, examining IUS from information filtering perspective, studying how evolutionary manifold-ranking and spectral clustering improve together, and to mine the deep temporal semantic components.

5.7. Miscellaneous

Contrastive Opinion Summarization (COS) aims to catch the representative sentences that have a contrastive meaning which helps in processing diverse assessments and contentions of the same issue [97]. It seeks to compare two documents on sentiment and semantic level [98]. [97] proposed an Expert-Guided Contrastive Opinion Summarization (ECOS) model that could influence a few expert opinions in the mining of enormous volumes of normal opinions. They gathered expert opinions from the websites that provide edited expert opinions for various disputable issues such as procon.org and debate.org. They embraced a heuristic approach with the idea of adding prior expert opinions and did the summarization part by selecting sentences based on the contrastive similitude. Guo et al. created a new dataset about “gay marriage” case and suggested improving clustering results, using of soft clustering strategy, and using more advanced semantic based sentence similarity measure. While not working directly with social media, it is worth to mention the contrastive summarization work that was done by Ref. [98], They provided a three-step algorithm (preprocessing, semantic processing, and summary creation) for evaluating the opinions of Czech senators. Most of preprocessing was done by using NLTK.[32] Topic Comparison part was done using latent topic model algorithms such as Latent Dirichlet Allocation (LDA) [99] and Latent Semantic Analysis (LSA) [100]. Data Set used is about speeches given by senators in the Czech Senate. They evaluated their work using an extrinsic evaluation method [101] proposed by Steinberger and Jezek. Future work suggested by Campr and Jezek is utilizing more sophisticated machine learning algorithms and using other evaluation algorithms.

Concept based summarization (Aka Business related summarization): Organization-related tweets are helpful to organization analysts, internal users, and purchasers [102]. In 2012, Louis and Newman [102] proposed a three-step approach (Concept learning, Tweet clustering, and Cluster ranking and summarization) for arranging company-related tweets into subtopics and creating a descriptive summary for each subtopic. The dataset used by Louis and Newman is

[30] http://www-nlpir.nist.gov/projects/duc/guidelines/2007.html.

[31] http://dev.twitter.com.

[32] http://www.nltk.org.

gathered by using some keywords for each company (using Microsoft crowdsourcing framework [103]). Evaluation is done by creating many summaries using four approaches (concept based, Sentiment only, Frequency only, sentiment + frequency) and taking judges' opinions about their most preferred and informative summary. Their suggestions include customized concepts for various classes of companies and different types of tweets.

Community Detection based Summarization: Subjective event summarization could be combined with more advanced community detection methods [16]. Community detection constitutes a critical instrument for the examination of complex networks by empowering the investigation of mesoscopic structures [104]. Community detection is helpful in numerous social-network analysis applications such as customer segmentation, recommendations, link inference, vertex labeling, and influence analysis [105]. Moreover, can help in many social computing tasks and is already applied in many real-world applications [106]. [107] formulated the informative sentence selection problem in opinion summarization as a community leader detection problem. Their framework evaluated the quality of summary terms of both aspect coverage and viewpoints preservation. They also proposed two algorithms to find the leaders (informative sentences) and communities (sentences with similar aspects and viewpoints). The Data Set[33] used by Zhu et al. is a collection of product reviews crawled from Amazon.com. For evaluation, they used aspect coverage and the polarity distribution preservation. Future work suggested by them is to exploit their sentence extraction method for other tasks, Plus, extending their techniques to various spaces.

Domain Specific Summarization: As in numerous natural language processing applications, methodologies that are specific to a certain domain generally, perform superior to the methodologies that work with generic domains [108,109] did Summarization and Sentiment Analysis on information acquired from the health forum site healthboards.com (health related social networking website), They attempted to discover an association between diseases, drugs, and symptoms. Fig. 35 shows their proposed System Architecture in which the accompanying steps happen: Keyword Extraction, disease−drug−symptoms Association, Sentiment Analysis, Summarization using Lesk based Summarization Algorithm described in Ref. [13].

Bilingual Summarization (Aka Cross-Language Document Summarization): The use of existing summarization techniques to social media with its various languages is faced with extra difficulties, which raises a question of whether to develop language-independent or language-specific approaches [11]. The incorporation of machine translation and summarization open doors for Cross-Language Summarization [11]. While working on their S.E.R (Social Event Radar) technology, Hsieh et al. [34] in 2012, proposed a bilingual sentiment opinion analysis (BSOA) technique. BSOA is implemented alongside lexicon based and domain knowledge. It begins with concept expansion technique for building up a measurable keyword network. They targeted the Chinese language as their second language. Also, while not working specifically in opinion summarization, Xiaojun Wan [110] in 2008, proposed an approach that uses English sentiment resources for Chinese sentiment analysis by utilizing machine translation and ensemble techniques. The framework of their approach is illustrated in Fig. 36. Also, while working on their Opinion Extraction part (bilingual English/Chinese), Ku et al. [35] in 2006 used the Academia Sinica Bilingual Ontological WordNet[34] (abbreviated as BOW). The Sinica BOW is expected as a linguistic infrastructure for knowledge engineering and knowledge representation; it is based upon the relation-based structure of WordNet [111].

[22] defined **Social Bookmarking** as the method of saving bookmarks on Web sites and Web pages and labeling them with keywords for later use. They proposed an approach that exploits users' comments and tags in social bookmarking services like del.icio.us,[35] Digg,[36] YouTube, and Amazon for summarization purposes. They utilized numerous components like term frequency, position score, and sentence length. Moreover, categorized users' feedback into Objective and Subjective Statements through a three-step process (Feature Word Extraction, Scoring sentences, and Summary Generation). Data Set used by Park et al. is arbitrarily inspected bookmarks from del.icio.us. Their evaluation algorithm is ROUGE [57] using manually generated summaries as a gold-standard. Future Work suggested by them is to extract the objective information for the source and to filter out the subjective information of other users.

Social Media Sampling: [112] proposed an optimization-driven method to solve the social message selection problem which they handled as an optimization problem. Their strategy considers numerous contents, social, and user features to deduce the intrinsic level of *informativeness*, *opinionatedness*, *popularity* and *authority* of each message while at the same time guaranteeing the consideration of diverse messages in the final set. They created a manually annotated news-response dataset with a Gold Standard collection generated by human annotators. For Evaluation, Stajner et al. computed both the ROUGE [57] and F1 scores. They suggested Incorporating Extra Message-level or Author-level indicators, moving from extractive sampling to abstractive summarization, and Extending their approach to deal with online sampling. [113] proposed an approach for social media sampling with taking into consideration many factors such as diversity, location, recency, the degree of diffusion effects. They used a greedy iterative clustering technique and developed a weighted dimensional representation of the information units. Next, they proposed a sampling methodology to reduce such large social

[33] Available at http://sites.google.com/site/linhongi2r/data-and-code.

[34] http://bow.ling.sinica.edu.tw.

[35] http://del.icio.us, temporarily http://delicious.com.

[36] http://digg.com.

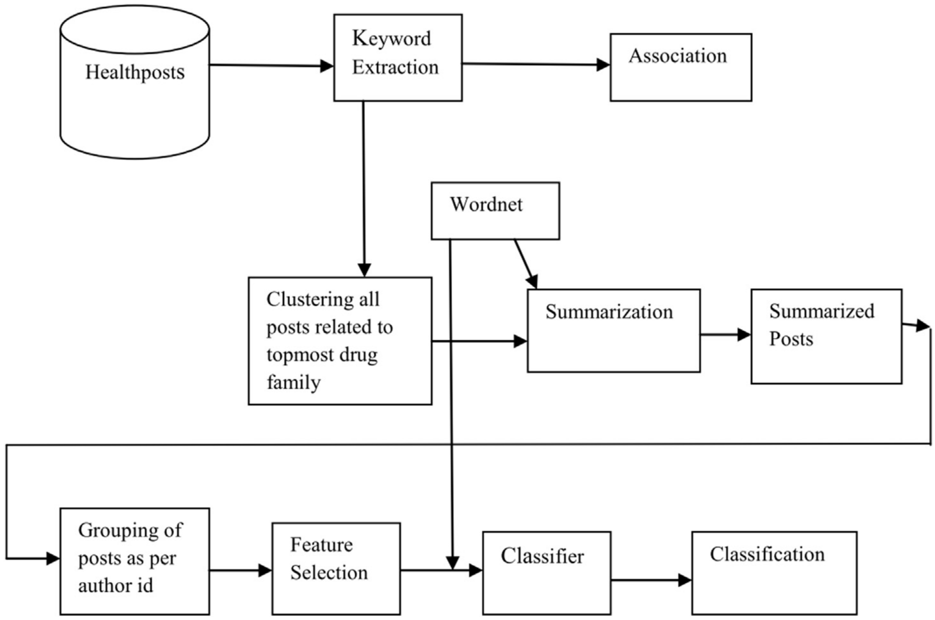


Fig. 35. Summarization and SA from user health posts system architecture [109].

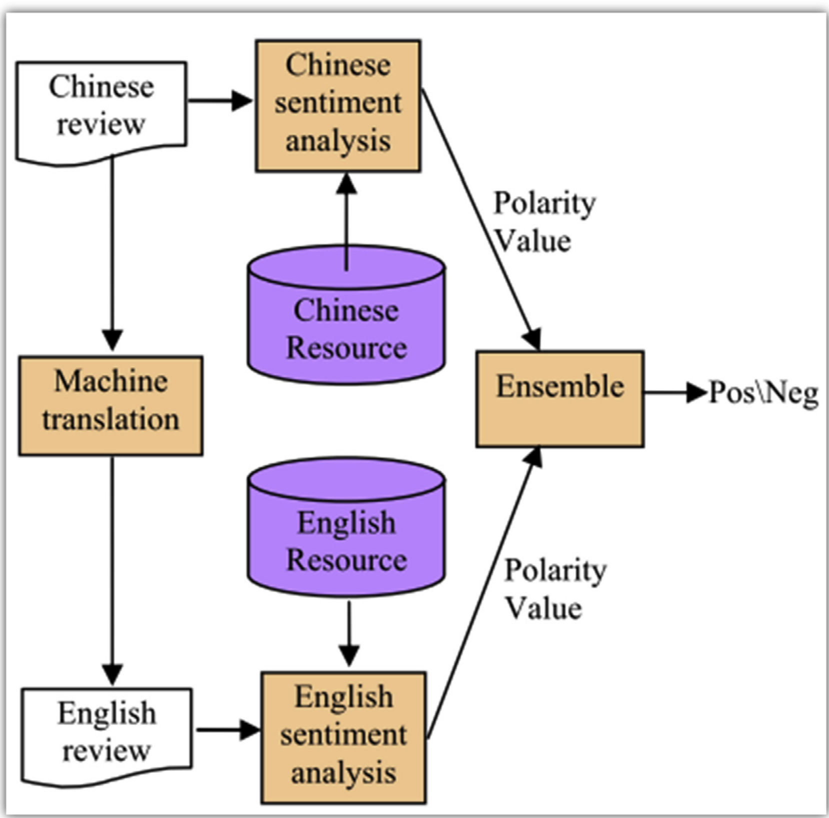


Fig. 36. Bilingual knowledge and ensemble techniques for SA [110].

media spaces. Their three-step methodology depended on sample space reduction, sample generation, and sample ordering. Data Set used by Choudhury et al. is a list of tweets regarding "iPhone" and "Oil Spill", Evaluation is done manually with regular Twitter users.

6. Evaluating opinion summarization

There are many approaches for summarizing text without losing its fundamental parts, making Evaluating summarization task tough [16]. Moreover, the absence of evaluation

$$\text{ROUGE-N} = \frac{\sum_{S \in \{ReferemceSummaries\}} \sum_{gram_n \in S} Count_{match}(gram_n)}{\sum_{S \in \{ReferenceSummaries\}} \sum_{gram_n \in S} Count(gram_n)}$$

Fig. 37. ROUGE-N metric [57].

measures for opinion summarization process is another issue to deal with [18]. The two methods of evaluating summarization are the **Intrinsic Evaluation**, which measures the properties of the nature of the subject and evaluates its objective, and the **Extrinsic Evaluation**, which measures perspectives concerning the effects of its exact function for a human user [114].

6.1. Intrinsic evaluation

In the Intrinsic Evaluation, the machine generated summaries are compared with other summaries written by professionals as a reference, making it a recall-based method [11].

One of the most used Intrinsic Evaluation tools is ROUGE [57], which stands for Recall-Oriented Understudy for Gisting Evaluation. It includes measures to determine summary's quality automatically, such as counting the number of overlapping units such as n-gram, word pairs, and word sequences. The following Equation (Fig. 37) shows ROUGE-N metric, where n is word n-gram length, gram $_n$, and Count $_{match}$ (gram $_n$) is the maximum number of n-grams co-occurring in an automated summary and a set of human summaries. So, ROUGE is basically a calculation of recall value between human and automated summary [30].

[96] described a framework in which summary evaluation measures can be instantiated and compared. They implemented a specific assessment method using very small units of content, called Basic Elements BE, that address some of the shortcomings of n-grams used in ROUGE. BE Package[37]- which is available without restriction - provides three main modules (BE breakers, BE matchers, and BE scorers).

[58] introduced an observationally grounded method for evaluating content determination in summarization called The Pyramid Method, which quantifies the relative significance of facts to be conveyed. It involves semantic matching of content units which lead to more stable, more informative scores, and to a meaningful content evaluation. One of the advantages of the pyramid method is that it permits the investigators to find what critical data is missing from the automated summary, which leads to target improvements of the summarizer. The Pyramid Method is also discussed in more details in Ref. [59], where the authors showed how it could be utilized for investigation of multiple human abstracts into semantic content units, which helps in assigning empirical importance weight to various content units.

Kim et al. [18] suggested attempting to create measures such as Opinion-ROUGE or Opinion-Pyramid method which considers opinion aspects alongside evaluation measures used in general summarization.

6.2. Extrinsic evaluation

Unlike Intrinsic Evaluation, Extrinsic Evaluation measures the quality of a system regarding its ability to solve a particular task such as correctly answering a user query, making it with no need for human summaries to be used as a gold standard [11].

Steinberger and Jezek [101] suggested an extrinsic evaluation method that handles the issue of lacking an accessible dataset, the principle of their method is to be able to classify the summarized document the same way and into the same class as the original (not summarized) full document [98].

Text Analysis Conference (TAC)2009[38] 2010,[39]2011[40] introduced the Automatically Evaluating Summaries of Peers (AESOP) task, which tried to promote research and development of systems that automatically evaluate the quality of summaries regarding their content and readability. For example, one of the participations in TAC 2009 [115] submitted four different metrics for AESOP (amount of content shared between a pair of texts, Use of the IS-A Taxonomy and Addition of Entities, statistical correlations, and discriminative power).

Finally, there is a need for a universal strategy to evaluate summarization systems [116]. Moreover, for fair comparisons between different evaluation techniques, we need good datasets, and evaluation measures [18]. Along with standard human or automatic summary evaluation metrics [11].

7. Summary and concluding remarks

Regardless of numerous research efforts, current opinion summarization studies still have numerous impediments for enhancements [18]. Moreover, there is a real requirement for building a summarization corpus specifically from social media for quickening the advancement in this area [11].

We see that recently there is more research interest in opinion summarization, especially in Abstractive Summarization category. And the current trend among many researchers is the use of deep learning techniques and utilizing the GPUs for training from large-scale data. So, we expect the following three keywords to dominate in the near future for the field of opinion summarization (Abstractive Summarization + Deep Learning + GPUs). And we also expect that more attention will be gained for the deep learning software tools such as TensorFlow,[41] Microsoft Cognitive Toolkit.[42]

In Table 1, we show a summary of Opinion Summarization techniques used in this survey, along with few remarks about them.

[37] http://www.isi.edu/~cyl/BE/.

[38] http://www.nist.gov/tac/2009/Summarization/aesop.09.guidelines.html.

[39] http://www.nist.gov/tac/2010/Summarization/AESOP.2010.guidelines.html.

[40] http://www.nist.gov/tac/2011/Summarization/AESOP.2011.guidelines.html.

[41] https://www.tensorflow.org.

[42] https://www.microsoft.com/en-us/research/product/cognitive-toolkit.

Table 1
Summary of opinion summarization techniques used in the survey.

Ref. No.	Techniques/Approaches/Methods	Data Sources/Domains	Remarks
[32]	Visualization, Aspect based	Products, laws, policies, discussions, forums	
[33]	Visualization	Sports	opinion evolutions with a holistic view
[34]	Visualization, Bilingual, lexicon based, domain knowledge	Brands, Food Safety, Products	constantly crawling all new media, Chinese, Polysemy, Double Negation, Adverb of Degree
[35]	Textual, Visualization, Bilingual	Web Blogs, News, President Elections, TREC & NTCIR corpora	opinion tracking (trend of opinions), Chinese, uses Sinica BOW
[10]	Visualization, Statistical, Aspect based	Products	Interactive Interface
[36]	Visualization, Statistical, Topic Modeling, Topic Phrase Mining	News	Cloud graphs
[37]	Statistical, Score based	Students Remarks by Teachers	collaborated opinion
[17]	Textual, Abstractive, Template based, speech acts, n-grams, SVM	Generic, Trending Topics, News	speech act-based sentence templates
[40]	Textual, Abstractive, Template based, speech acts, Bagging Ensemble, Naïve Bayes Classifier	Tweets about trending topics on Twitter	speech act-based sentence templates
[41]	Textual, Abstractive, Graph-based	DUC 2002	Compressing and merging sentences based on word graphs
[42]	Textual, Abstractive, Graph-based	Opinosis Dataset	Shallow NLP, assumes no domain knowledge
[43]	Textual, Abstractive, Graph-based	DUC 2002 + Opinosis Dataset	uses sentiment analysis, overcomes the redundant sentences issues
[44]	Textual, Abstractive, Graph-based	Hu & Liu Dataset	Discourse Structure
[45]	Textual, Abstractive, Graph-based	DUC 2004, DUC 2005	ILP based multi-sentence compression
[3]	Textual, Abstractive, Concept-level	Products, mobile phones, cars, Amazon	Sentence generation, simplified sentence
[46]	Textual, Abstractive, Semantic based	DUC 2002	Semantic Role Labeling, Genetic Algorithm
[47]	Textual, Abstractive, Semantic-based	TAC 2010	Based on Information Items (INITs)
[48]	Textual, Abstractive, Semantic-based	TAC 2011	Pool of concepts
[49]	Textual, Abstractive, Machine Learning	reviews from CNET	Optimization Problem
[50]	Textual, Abstractive, Data-Driven	DUC 2003, DUC 2004	Neural attention-based model, Headline-generation, Many Encoders
[53]	Textual, Abstractive, Machine Learning	a new dataset consisting of multi-sentence summaries	Sequence to sequence problem, Attentional Encoder-Decoder Recurrent Neural Networks
[41]	Textual, Hybrid Extractive/Abstractive	DUC 2002	
[54]	Textual, Hybrid Extractive/Abstractive	Reviews	Natural language generation and salient sentence selection techniques
[61]	Textual, Aspect based	Product Reviews	
[8]	Textual, Aspect based, Score	Product Reviews from Amazon	Evaluation (aspect coverage & viewpoints preservation)
[12]	Textual, Graphical, Aspect based, hybrid Supervised/Unsupervised Polarity Detection, Topic Detection (Stream-Based, Semantic-Based)	Products	uses median tweet for summarization
[62]	Score, Aspect based	Products, customer reviews, Amazon, C | net	Feature-based summary, Feature Buzz Summary
[9]	Textual, Aspect based	Product reviews, movie review, politicians, celebrities, brands	provide insights behind opinions
[63]	Textual, Aspect based	Forum posts	conversational documents
[64]	Textual, Aspect based, Score	Products, Local Services, Search Query, User Reviews	quantitative and qualitative
[65]	Textual, Query-Specific	General, community question answering and blog data, Yahoo! Answers, TAC 2008 corpus	objective function (Relevance, coverage, non-redundancy)
[67]	Textual, Query-Specific	DUC 2006, DUC 2007	Sentence Ranking, Sentence Compression, and Post-processing for coreference resolution and sentence reordering
[68]	Textual, Query-Specific	DUC 2005, DUC 2006	weighted Archetypal Analysis (wAA)
[66]	Textual, Query-Specific	DUC 2005, DUC 2006	Probabilistic-modeling Relevance, Coverage, and Novelty (PRCN) framework
[69]	Textual, Query-Specific	General	the source code is available, Frequency contrast, Topic diversity, Topic size
[70]	Textual, Query-Specific, Latent Semantic Indexing, lexicon-based, diversity penalty	General	
[78]	Textual, Real Time, Scheduled Events, SVM, Naïve Bayes, Logistic Regression, AFINN scoring	Sports (soccer and cricket)	
[16]	Textual, Real Time, Scheduled Events, intrinsic evaluation	Sports (FA Cup finals)	subjective summary, two (or more) perspectives

Table 1 (*continued*)

[80]	Textual, Real Time, Scheduled Events	Sports (Copa America)	Doesn't require pre-knowledge about the sport sub-events
[14]	Textual, Real Time, Non-Scheduled Events, Search Query, temporal correlation, n-grams	Company events, natural disasters, politics	
[84]	Textual, Update Summarization, hierarchical tree model, Hierarchical LDA	TAC 2008−2011 corpus	
[85]	Textual, Update Summarization, multi-level	TREC 2013 corpus	topic-level, sentence-level
[86]	Textual, Update Summarization, supervised regression model	TREC 2013 corpus	prevalence and novelty
[87]	Textual, Update Summarization, Heuristic (stream-based), LDA-based (semantic-based)	Trending Topics, News, Technology, Sports	Coverage, novelty, correlation
[82]	Textual, Update Summarization, evolutionary manifold-ranking model	DUC 2007, TAC 2008 corpus	Combined the normalized spectral clustering with evolutionary manifold-ranking
[88]	Textual, Update Summarization, topic modeling, dynamic modeling	TAC 2008−2009 corpus	unified framework for both standard and update summarization
[89]	Textual, Update Summarization, co-ranking method	TAC 2011 corpus	
[90]	Textual, Update Summarization, three-level Hierarchical Dirichlet Process HDP	TAC 2008−2011 corpus	
[91]	Textual, Update Summarization, Clustering Based Sentence Extractor	TAC 2008−2009 corpus	
[92]	Textual, Update Summarization, incremental hierarchical clustering framework	TAC 2008 corpus + Hurricane Wilma Releases (Hurricane)	
[93]	Textual, Update Summarization, approximation methods	TAC 2007−2009 corpus	semantic element extraction
[94]	Textual, Update Summarization, LSA	TAC 2008 corpus	
[95]	Textual, Update Summarization, Iterative Residual Rescaling	TAC 2008 corpus	latent semantic space
[97]	Textual, Contrastive Opinion summarization	Gay Marriage, procon.org	select sentence based on contrastive similarity
[98]	Textual, Contrastive Opinion Summarization, LDA, LSA	Political, Czech Senate	semantic and sentiment level
[102]	Textual, Concept-based, Aspect based, score, Concept learning, Tweet Clustering, Cluster ranking, and summarization	Companies	Business related, Sentiment only, Frequency only, Sentiment + Frequency
[107]	Textual, Community Detection, optimization problem	Product Reviews, Amazon	Evaluation (aspect coverage & viewpoints preservation)
[109]	Textual, Domain-Specific, bag-of-word, TD-IDF	Medical, health forums (health boards, patients like me)	Unified Medical Language System
[110]	Textual, Bilingual, machine translation, ensemble techniques	Reviews	Chinese
[22]	Textual, Social Bookmarking	delicious, Digg, YouTube, and Amazon	works with social bookmarks
[112]	Textual, Social Media Sampling, optimization-driven	News and Articles	Considers informativeness, opinionatedness, popularity & authority
[113]	Textual, Social Media Sampling, greedy iterative clustering	General, Products, Trending Topics	The desired level of diversity, considers characteristics (social, content, nodal)

References

[1] Bhattacharjee Saprativa, Ray Anjan Kumar. Sentiment analysis: approaches, applications and challenges. IJIACS Int J Innov Adv Comput Sci March 2015;4:516−27.

[2] Felciah M Lovelin Ponn, Anbuselvi R. A study on sentiment analysis of social media reviews. In: Innovations in information, embedded and communication systems (ICIIECS), 2015 International Conference on; 2015. p. 1−3.

[3] Lloret Elena, Ester Boldrini, Tatiana Vodolazova, Patricio Martínez-Barco, Rafael Muñoz, Manuel Palomar. A novel concept-level approach for ultra-concise opinion summarization. Expert Syst Appl 2015;42(20): 7148−56.

[4] Liao Xiangwen, Huang Yichao, Wei Jingjing, Yu Zhiyong, Chen Guolong. A heterogeneous graph model for social opinion detection. In: Machine learning and cybernetics; 2014. p. 175−85.

[5] Hole Vikrant, Takalikar Mukta. A survey on sentiment analysis and summarization for prediction. Int J Eng Comput Sci (IJECS) December 2014;3(12):9503−6.

[6] Balahur Alexandra, Kabadjov Mijail, Steinberger Josef, Steinberger Ralf, Andr. Challenges and solutions in the opinion summarization of user-generated content. J Intell Inf Syst 2012;39(2):375−98.

[7] Rao Ashwini, Shah Ketan. Model for improving relevant feature extraction for opinion summarization. In: Advance computing conference (IACC), 2015 IEEE international; 2015. p. 1−5.

[8] Zhu Linhong, Sheng Gao, Sinno Jialin Pan, Haizhou Li, Dingxiong Deng, Cyrus Shahabi. Graph-based informative-sentence selection for opinion summarization. In: Advances in social networks analysis and mining (ASONAM), 2013 IEEE/ACM international conference on; 2013. p. 408−12.

[9] Meng Xinfan, Furu Wei, Xiaohua Liu, Ming Zhou, Sujian Li, Houfeng Wan. Entity-centric topic-oriented opinion summarization in twitter. In: Proceedings of the 18th ACM SIGKDD international conference on knowledge discovery and data mining; 2012. p. 379−87.

[10] Kamal Ahmad. Review mining for feature based opinion summarization and visualization. Int J Comput Appl 2015;119(17).

[11] Farzindar Atefeh. Social network integration in document summarization. Innovative document summarization techniques revolutionizing knowledge understanding. IGI-Global; 2014. p. 139−62.

[12] Bahrainian Seyed-Ali, Dengel Andreas. Sentiment analysis and summarization of twitter data. In: Computational science and engineering (CSE), 2013 IEEE 16th international conference on; 2013. p. 227−34.

[13] Cambria Erik, Schuller Bjorn, Xia Yunqing, Havasi Catherine. New avenues in opinion mining and sentiment analysis. IEEE Intell Syst 2013;(2):15−21.

[14] Chua Freddy Chong Tat, Asur Sitaram. Automatic summarization of events from social media. In: Seventh international AAAI conference on weblogs and social media; 2013.

[15] Zhai Zhongwu, Liu Bing, Xu Hua, Jia Peifa. Clustering product features for opinion mining. In: Proceedings of the fourth ACM international conference on web search and data mining; 2011. p. 347−54.

[16] Corney David, Martin Carlos, Göker Ayse. Two sides to every story: subjective event summarization of sports events using Twitter. In: ICMR2014 workshop on social multimedia and storytelling, Glasgow, Scotland; 2014.

[17] Zhang Renxian, Li Wenjie, Gao Dehong, Ouyang You. Automatic twitter topic summarization with speech acts. IEEE Trans Audio Speech Lang Process 2013;21(3):649−58.

[18] Duk Kim Hyun, Ganesan Kavita, Sondhi Parikshit, Zhai ChengXiang. Comprehensive review of opinion summarization. 2011 [Online]. HYPERLINK, http://hdl.handle.net/2142/18702. http://hdl.handle.net/2142/18702.

[19] Manning Christopher D, Mihai Surdeanu, John Bauer, Jenny Finkel, Steven J Bethard, David McClosky. The stanford CoreNLP natural language processing toolkit. In: Association for computational linguistics (ACL) system demonstrations; 2014. p. 55−60 [Online]. HYPERLINK, http://www.aclweb.org/anthology/P/P14/P14-5010. http://www.aclweb.org/anthology/P/P14/P14-5010.

[20] ElSahar Hady, El-Beltagy Samhaa R. A fully automated approach for Arabic slang lexicon extraction from microblogs. In: International conference on intelligent text processing and computational linguistics; 2014. p. 79−91.

[21] Larkey Leah S, Ballesteros Lisa, Connell Margaret E. Light stemming for Arabic information retrieval. In: Arabic computational morphology. Springer; 2007. p. 221−43.

[22] Park Jaehui, Fukuhara Tomohiro, Ohmukai Ikki, Takeda Hideaki, Lee Sang-goo. Web content summarization using social bookmarks: a new approach for social summarization. In: Proceedings of the 10th ACM workshop on Web information and data management; 2008. p. 103−10.

[23] Santos Igor, Igor Miñambres-Marcos, Carlos Laorden, Patxi Galán-García, Aitor Santamaría-Ibirika, Pablo García Bringas. Twitter content-based spam filtering. In: International joint conference SOCOâ€™13-CISISâ€™13-ICEUTEâ€™13; 2014. p. 449−58.

[24] Nichols Jeffrey, Mahmud Jalal, Drews Clemens. Summarizing sporting events using twitter. In: Proceedings of the 2012 ACM international conference on intelligent user interfaces; 2012. p. 189−98.

[25] Aisopos Fotis, Papadakis George, Varvarigou Theodora. Sentiment analysis of social media content using N-Gram graphs. In: Proceedings of the 3rd ACM SIGMM international workshop on social media; 2011. p. 9−14.

[26] Chen Hsinchun, Zimbra David. AI and opinion mining. IEEE Intell Syst 2010;25(3):74−80.

[27] Zhai Zhongwu, Liu Bing, Xu Hua, Jia Peifa. Grouping product features using semi-supervised learning with soft-constraints. In: Proceedings of the 23rd international conference on computational linguistics; 2010. p. 1272−80.

[28] Nigam Kamal, McCallum Andrew Kachites, Thrun Sebastian, Mitchell Tom. Text classification from labeled and unlabeled documents using EM. Mach Learn 2000;39(2−3):103−34.

[29] Zhai Zhongwu, Liu Bing, Wang Jingyuan, Xu Hua, Jia Peifa. Product feature grouping for opinion mining. IEEE Intell Syst 2012;27(4):37−44.

[30] Raut Vijay B, Londhe DD. Survey on opinion mining and summarization of user reviews on web. Int J Comput Sci Inf Technol (IJCSIT) 2014;5(2):1026−30.

[31] More Mugdha, Tidke Bharat. Social media online opinion summarization using ensemble technique. In: Pervasive computing (ICPC), 2015 international conference on; 2015. p. 1−6.

[32] Kherwa Pooja, Sachdeva Anish, Mahajan Dhruv, Pande Nishtha, Singh Praveen Kumar. An approach towards comprehensive sentimental data analysis and opinion mining. In: Advance computing conference (IACC), 2014 IEEE international; 2014. p. 606−12.

[33] Lin Lu, Li Jianxin, Zhang Richong, Yu Weiren, Sun Chenggen. Opinion mining and sentiment analysis in social networks: a retweeting structure-aware approach. In: Utility and cloud computing (UCC), 2014 IEEE/ACM 7th international conference on; 2014. p. 890−5.

[34] Hsieh Wen-Tai, Ku Tsun, Wu Chen-Ming, Chou Seng-cho T. Social event radar: a bilingual context mining and sentiment analysis summarization system. In: Proceedings of the ACL 2012 system demonstrations; 2012. p. 163−8.

[35] Ku Lun-Wei, Liang Yu-Ting, Chen Hsin-Hsi. Opinion extraction, summarization and tracking in news and blog corpora. In: AAAI spring symposium: computational approaches to analyzing weblogs; 2006. 100107.

[36] Ramamonjisoa D, Murakami R, Chakraborty B. Comments analysis and visualization based on topic modeling and topic phrase mining. In: The third international conference on E-technologies and business on the web (EBW2015); 2015. p. 1.

[37] Virmani Deepali, Malhotra Vikrant, Tyagi Ridhi. Sentiment analysis using collaborated opinion mining. Commun Res Rep 2014. vol. abs/1401.2618, https://arxiv.org/abs/1401.2618.

[38] Mirchev Uri, Last Mark. Multi-document summarization by extended graph text representation and importance refinement. Innovat Doc Summ Tech Revolutionizing Knowl Underst 2014:28−53.

[39] Hu Baotian, Chen Qingcai, Zhu Fangze. Lcsts: a large scale Chinese short text summarization dataset. arXiv prepr arXiv 1506.05865 2016;4.

[40] Shaikh Gulab R, Padulkar Digambar M. Template based abstractive summarization of twitter topic with speech act. 2014.

[41] Lloret Elena, Palomar Manuel. 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. p. 61−6.

[42] Ganesan Kavita, Zhai ChengXiang, Han Jiawei. Opinosis: a graph-based approach to abstractive summarization of highly redundant opinions. In: Proceedings of the 23rd international conference on computational linguistics; 2010. p. 340−8.

[43] Bhargava Rupal, Sharma Yashvardhan, Sharma Gargi. ATSSI: abstractive text summarization using sentiment infusion. Procedia Comput Sci 2016;89:404−11.

[44] Gerani Shima, Mehdad Yashar, Carenini Giuseppe, Ng Raymond T, Nejat Bita. Abstractive summarization of product reviews using discourse structure. In: EMNLP; 2014. p. 1602−13.

[45] Banerjee Siddhartha, Mitra Prasenjit, Sugiyama Kazunari. Multi-document abstractive summarization using ilp based multi-sentence compression. In: The twenty-fourth international joint conference on artificial intelligence (IJCAI 2015); 2015. p. 1208−14.

[46] Khan Atif, Salim Naomie, Kumar Yogan Jaya. A framework for multi-document abstractive summarization based on semantic role labelling. Appl Soft Comput 2015;30:737−47.

[47] Genest Pierre-Etienne, Lapalme Guy. Framework for abstractive summarization using text-to-text generation. In: Proceedings of the workshop on monolingual text-to-text generation; 2011. p. 64−73.

[48] Bing Lidong, Piji Li, Yi Liao, Wai Lam, Weiwei Guo, Rebecca J Passonneau. Abstractive multi-document summarization via phrase selection and merging. arXiv prepr arXiv 1506.01597 2015. p. 1587−1597, https://arxiv.org/abs/1506.01597.

[49] Ganesan Kavita, Zhai ChengXiang, Viegas Evelyne. Micropinion generation: an unsupervised approach to generating ultra-concise summaries of opinions. In: Proceedings of the 21st international conference on world wide web; 2012. p. 869−78.

[50] Rush Alexander M, Chopra Sumit, Weston Jason. A neural attention model for abstractive sentence summarization. In: The 2015 conference on empirical methods in natural language processing; 2015. p. 379−89. Lisbon, Portugal.

[51] Yoshua Bengio R, Vincent Pascal, Jauvin Christian. A neural probabilistic language model. J Mach Learn Res 2003;3(Feb):1137−55.

[52] Bahdanau Dzmitry, Cho Kyunghyun, Bengio Yoshua. Neural machine translation by jointly learning to align and translate. arXiv prepr arXiv 1409.0473 2014. p. 1−15, https://arxiv.org/abs/1409.0473v1.

[53] others Nallapati Ramesh, Zhou Bowen, Gulcehre Caglar, Xiang Bing. Abstractive text summarization using sequence-to-sequence rnns and beyond. arXiv prepr arXiv 1602.06023 2016. https://arxiv.org/abs/1602.06023.

[54] Di Fabbrizio Giuseppe, Stent Amanda J, Gaizauskas Robert. A hybrid approach to multi-document summarization of opinions in reviews. In: INLG 2014; 2014. p. 54.

[55] Cao Ziqiang, Chen Chengyao, Li Wenjie, Li Sujian, Wei Furu, Zhou Ming. Tgsum: build tweet guided multi-document summarization dataset. AAAI; 2016. p. 2906−12.

[56] Hu Minqing, Liu Bing. Mining and summarizing customer reviews. In: Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining; 2004. p. 168−77.

[57] Lin Chin-Yew. Rouge: a package for automatic evaluation of summaries. In: Text summarization branches out: proceedings of the ACL-04 workshopvol. 8; 2004.

[58] Nenkova Ani, Passonneau Rebecca. Evaluating content selection in summarization: the pyramid method. Columbia University Academic Commons [distributor]; 2004 [Online]. HYPERLINK, http://hdl.handle.net/10022/AC: P:20539. http://hdl.handle.net/10022/AC: P:20539.

[59] Nenkova Ani, Passonneau Rebecca, McKeown Kathleen. The pyramid method: incorporating human content selection variation in summarization evaluation. ACM Trans Speech Lang Process(TSLP) 2007; 4(2):4.

[60] Luong Minh-Thang, Sutskever Ilya, Le Quoc V, Vinyals Oriol, Zaremba Wojciech. Addressing the rare word problem in neural machine translation. arXiv prepr arXiv 1410.8206 2014. https://arxiv.org/abs/1410.8206.

[61] Yang Jung-Yeon, Myung Jaeseok, Lee Sang-goo. The method for a summarization of product reviews using the user's opinion. In: Information, process, and knowledge management, 2009. eKNOW'09. International conference on; 2009. p. 84−9.

[62] Jmal Jihene, Faiz Rim. Customer review summarization approach using Twitter and SentiWordNet. In: Proceedings of the 3rd international conference on web intelligence, mining and semantics; 2013. p. 33.

[63] Mangnoesing Gino, van Bunningen Arthur, Hogenboom Alexander, Hogenboom Frederik, Frasincar Flavius. An empirical study for determining relevant features for sentiment summarization of online conversational documents. In: Web information systems engineering-WISE 2012. Springer; 2012. p. 567−79.

[64] Sasha Blair-Goldensohn, Kerry Hannan, Ryan McDonald, Tyler Neylon, George A Reis, Jeff Reynar. Building a sentiment summarizer for local service reviews. In: WWW workshop on NLP in the information explosion eravol. 14; 2008. p. 339−48.

[65] Wang Lu, Raghavan Hema, Cardie Claire, Castelli Vittorio. Query-focused opinion summarization for user-generated content. In: COLING 2014, the 25th international conference on computational linguistics: technical papers; August 23-29 2014. p. 1660−9. Dublin, Ireland.

[66] Luo Wenjuan, Zhuang Fuzhen, He Qing, Shi Zhongzhi. Exploiting relevance, coverage, and novelty for query-focused multi-document summarization. Knowl Base Syst 2013;46:33−42.

[67] Wang Lu, Raghavan Hema, Castelli Vittorio, Florian Radu, Cardie Claire. A sentence compression based framework to query-focused multi-document summarization. arXiv prepr arXiv 1606.07548 2016. https://arxiv.org/abs/1606.07548.

[68] Canhasi Ercan, Kononenko Igor. Weighted archetypal analysis of the multi-element graph for query-focused multi-document summarization. Expert Syst Appl 2014;41(2):535−43.

[69] O'Connor Brendan, Krieger Michel, Ahn David. TweetMotif: exploratory search and topic summarization for twitter. In: ICWSM; 2010.

[70] Jin Feng, Huang Minlie, Zhu Xiaoyan. A query-specific opinion summarization system. In: Cognitive informatics, 2009. ICCI'09. 8th IEEE international conference on; 2009. p. 428−33.

[71] Dang Hoa Trang. Overview of DUC 2005. In: Proceedings of the document understanding conference. vol. 2005; 2005.

[72] Dang Hoa Trang. Overview of DUC 2006. In: Proceedings of the document understanding conference. vol. 2006; 2006.

[73] Dang Hoa Trang. Overview of DUC 2007. In: Proceedings of the document understanding conference. vol. 2007; 2007.

[74] Dang Hoa Trang, Owczarzak Karolina. Overview of the tac 2008 opinion question answering and summarization tasks. In: Proc. of the first text analysis conference; 2008.

[75] Clarke James, Lapata Mirella. Global inference for sentence compression: an integer linear programming approach. J Artif Intell Res 2008; 31:399−429.

[76] Atefeh Farzindar, Khreich Wael. A survey of techniques for event detection in twitter. Comput Intell 2015;31(1):132−64.

[77] Shen Chao, Liu Fei, Weng Fuliang, Li Tao. A participant-based approach for event summarization using twitter streams. In: HLT-NAACL; 2013. p. 1152−62.

[78] Hole Vikrant, Takalikar Mukta. Real time tweet summarization and sentiment analysis of game tournament. Int J Sci Res (IJSR) Sep. 2015; 4(9):1774−80.

[79] Heider Karl G. The Rashomon effect: when ethnographers disagree. Am Anthropol 1988;90(1):73−81.

[80] Zubiaga Arkaitz, Spina Damiano, Amig Enrique, Gonzalo Julio. Towards real-time summarization of scheduled events from twitter streams. In: Proceedings of the 23rd ACM conference on Hypertext and social media; 2012. p. 319−20.

[81] Farzindar Atefeh, Inkpen Diana. Natural language processing for social media. Synth Lect Hum Lang Technol 2015;8(2):1−166.

[82] He Ruifang, Qin Bing, Liu Ting. A novel approach to update summarization using evolutionary manifold-ranking and spectral clustering. Expert Syst Appl 2012;39(3):2375−84.

[83] Aslam Javed A, Ekstrand-Abueg Matthew, Pavlu Virgil, Diaz Fernando, Sakai Tetsuya. TREC 2013 temporal summarization. In: The twenty-second text REtrieval conference; 2013.

[84] Li Rumeng, Shindo Hiroyuki. A hierarchical tree model for update summarization. In: Advances in information retrieval. Springer; 2015. p. 660−5.

[85] Zhang Chunyun, Ma Zhanyu, Zhang Jiayue, Xu Weiran, Guo Jun. A multi-level system for sequential update summarization. In: Heterogeneous networking for quality, reliability, security and robustness (QSHINE), 2015 11th international conference on; 2015. p. 144−8.

[86] McCreadie Richard, Macdonald Craig, Ounis Iadh. Incremental update summarization: adaptive sentence selection based on prevalence and novelty. In: Proceedings of the 23rd ACM international conference on conference on information and knowledge management; 2014. p. 301−10.

[87] Gao Dehong, Li Wenjie, Cai Xiaoyan, Zhang Renxian, Ouyang You. Sequential summarization: a full view of twitter trending topics. on Audio Speech Lang Process IEEE/ACM Trans 2014;22(2):293−302.

[88] Wang Hongling, Zhou Guodong. Toward a unified framework for standard and update multi-document summarization. ACM Trans Asian Lang Inf Process (TALIP) 2012;11(2):5.

[89] Wan Xiaojun. Update summarization based on Co-Ranking with constraints. In: COLING (posters); 2012. p. 1291−300.

[90] Li Jiwei, Li Sujian, Wang Xun, Tian Ye, Chang Baobao. Update summarization using a multi-level hierarchical dirichlet process model. Proc COLING 2012: Tech Phy Pap December 2012:1603−18. Mumbai.

[91] Bossard Aurelien, Rodrigues Christophe. Combining a multi-document update summarization system−CBSEAS−with a genetic algorithm. In: Combinations of intelligent methods and applications. Springer; 2011. p. 71−87.

[92] Wang Dingding, Li Tao. Document update summarization using incremental hierarchical clustering. In: Proceedings of the 19th ACM international conference on Information and knowledge management; 2010. p. 279−88.

[93] Long Chong, Huang Min-Lie, Zhu Xiao-Yan, Li Ming. A new approach for multi-document update summarization. J Comput Sci Technol 2010; 25(4):739−49.

[94] Steinberger Josef, Jezek Karel. Update summarization based on latent semantic analysis. In: Text, speech and dialogue; 2009. p. 77−84.

[95] Steinberger Josef, Jezek Karel. Update summarization based on novel topic distribution. In: Proceedings of the 9th ACM symposium on document engineering; 2009. p. 205−13.

[96] Hovy Eduard, Lin Chin-Yew, Zhou Liang, Fukumoto Junichi. Automated summarization evaluation with basic elements. In: Proceedings of the fifth conference on language resources and evaluation (LREC 2006); 2006. p. 604−11.

[97] Guo Jinlong, Lu Yujie, Mori Tatsunori, Blake Catherine. Expert-guided contrastive opinion summarization for controversial issues. In: Proceedings of the 24th international conference on world wide web companion; 2015. p. 1105−10.

[98] Campr Michal, Jezek Karel. Contrastive summarization: comparing opinions of Czech senators. J Theor Appl Inf Technol 2015;77(1).

[99] Campr Michal, Jezek Karel. Comparative summarization via latent dirichlet allocation. In: Dateso; 2013. p. 80−6.

[100] Campr Michal, Jezek Karel. Comparative summarization via latent semantic analysis. In: WSEAS international conference. Proceedings. Recent advances in computer engineering series; 2012.

[101] Steinberger Josef, Jezek Karel. Evaluation measures for text summarization. Comput Inf 2009;28(2):1001−25.

[102] Louis Annie, Newman Todd. Summarization of business-related tweets: a concept-based approach. In: COLING 2015, the 24th international conference on computational linguistics: posters; December 2012. p. 765−74. Mumbai.

[103] DiPalantino Dominic, Vojnovic Milan. Crowdsourcing and all-pay auctions. In: Proceedings of the 10th ACM conference on electronic commerce; 2009. p. 119−28.

[104] Symeon P, Yiannis K, Athena V, Ploutarchos S. Community detection in social media, performance and application considerations. J Data Min Knowl Discov 2012;24(3):515−54.

[105] Qi Guo-Jun, Aggarwal Charu C, Huang Thomas. Community detection with edge content in social media networks. In: Data engineering (ICDE), 2012 IEEE 28th international conference on; 2012. p. 534−45.

[106] Tang Lei, Liu Huan. “Community detection and mining in social media. Synth Lect on Data Min Knowl Discov 2010;2(1):1−137.

[107] Linhong Zhu, Sheng Gao, Sinno Jialin Pan, Haizhou Li, Dingxiong Deng, Cyrus Shahabi. The pareto principle is everywhere: finding informative sentences for opinion summarization through leader detection. In: Recommendation and search in social networks. Springer; 2015. p. 165−87.

[108] Imran Muhammad, Castillo Carlos, Diaz Fernando, Vieweg Sarah. Processing social media messages in mass emergency: a survey. ACM Comput Surv (CSUR) 2015;47(4):67.

[109] Mane Vinod L, Panicker Suja S, Patil Vidya B. Summarization and sentiment analysis from user health posts. In: Pervasive computing (ICPC), 2015 international conference on; 2015. p. 1−4.

[110] Wan Xiaojun. Using bilingual knowledge and ensemble techniques for unsupervised Chinese sentiment analysis. In: Proceedings of the conference on empirical methods in natural language processing; 2008. p. 553−61.

[111] Huang Chu-Ren. Sinica BOW: integrating bilingual WordNet and SUMO ontology. In: Natural language processing and knowledge engineering, 2003. Proceedings. 2003 international conference on; Oct 2003. p. 825−6.

[112] Stajner Tadej, Thomee Bart, Popescu Ana-Maria, Pennacchiotti Marco, Jaimes Alejandro. Automatic selection of social media responses to news. In: Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining; 2013. p. 50−8.

[113] De Choudhury Munmun, Counts Scott, Czerwinski Mary. Find me the right content! Diversity-based sampling of social media spaces for topic-centric search. In: ICWSM, 5th international AAAI conference on weblogs and social media; 2011.

[114] Jones Karen Sparck, Galliers Julia R. Evaluating natural language processing systems: an analysis and reviewvol. 1083. Springer Science ＼& Business Media; 1995.

[115] Steinberger Josef, Kabadjov Mijail, Pouliquen Bruno, Steinberger Ralf, Poesio Massimo. WB-JRC-UT's participation in TAC 2009 Update summarization and AESOP tasks. In: Proceedings of the text analysis conference, Gaithersburg, MD (November 2009); 2009.

[116] Das Dipanjan, Martins Andre FT. A survey on automatic text summarization. Lit Surv Lang Stat II Course CMU 2007;4:192−5.