Sentimental Analysis – A Survey of Some Existing Studies

Prabakaran Thangavel, Ravi Lourduswamy

Sacred Heart College (Autonomous), Tirupattur(dt), Tamilnadu – 635 601, India.

Abstract

Sentimental Analysis is a process of computing and categorizing the expressed opinions of people about certain event, subject or product as positive, negative or neutral. The major objective of Sentimental Analysis is to help data-driven decisions using insights from replies in social media, surveys and product reviews. Sentiment Analysis can be done with words, sentences, documents, features or aspects, concepts, phrases, links, clauses and implications. Recently, there has been a lot of attention on sentiment analysis especially from researchers in the fields of text mining and natural language processing. But due to extreme absence of annotated datasets which are used to train models in various domains, the accuracy of sentiment analysis has been hindered. Many types of research have been done to confront the challenge and enhance sentiment analysis classification. Sentiment analysis is important as it helps in identifying the emotional and attitude states of people. Positive or negative feelings of people can be expressed in different ways. This research article talks about, in subtle terms, the different ways to deal with sentiment analysis mostly in Machine Learning, Lexicon-based, Hybrid and Ontology-based approaches. This research article gives point by point perspective of the distinctive applications and challenges of Sentiment Analysis.

Keywords 1

Sentiment Analysis, Machine Learning, Lexicon-based, Corpus-based, Hybrid and Ontology

1. Introduction

Sentimental Analysis is a process of determining whether the text, document and media content as positive, negative or neutral. The foremost objective of Sentimental Analysis is to help data-driven decisions using perceptions from responses in social media, surveys and product reviews. Sentiment Analysis can be done with words, sentences, documents, features or aspects, concepts, phrases, links, clauses and implications. The bag of words recovered from word cloud is used for word-level Sentiment Analysis. At the sentence level Sentiment Analysis is done with sentences in reviews and comments written by users. A document-level analysis is done by classifying the opinions expressed in an entire document into different sentiments. Features can be used likewise for opinion mining and Sentiment Analysis. Feature level classification is done by identifying and extracting the different product features from raw source data. The feature analysis is done when a desired sentimental aspect or feature is to be got from a review. This article is structured as follows: section-2.presents background information related to the survey. Section-3.presents the related works conducted on various aspects of sentimental analysis. Section-4.presents the methodology of the survey conducted. Detailed discussion on open issues and challenges of sentiment analysis is presented in section-5. Finally, this article is concluded in section-6.

ISIC'21: International Semantic Conference, February 25-27, 2021, New Delhi, India
EMAIL: prabagaran@shctpt.edu (Prabakaran Thangavel)
ravi@shctpt.edu (Ravi Lourduswamy)

CEUR Workshop Proceedings (CEUR-WS.org)

2. Background

2.1 Machine Learning Approaches

Machine Learning (ML) techniques have a variety of features that help in the construction of classifiers for the expressed sentiments in texts. ML approaches make use of ML algorithms for Sentimental Analysis to extract the linguistic and syntactic features and form standard classifications of those features. Another popular method used for sentimental analysis is Deep-Learning which makes use of neural network techniques.

Supervised learning: Supervised learning is a technique that makes use of labelled datasets to create a model. A variety of supervised models can be created depending on the type of procedure that is employed to create the model.

Decision tree classifiers: Decision tree classifiers prepare the information space in different levels of decreasing order that distinguishes data based on the estimation of their credit. A recursive procedure is followed while creating the data division space that will result in the leaf hubs having the base quantities of the records used as the end goal of characterization.

Linear classification: Linear classifiers are of different types. The Sustenance Vector machines are one kind of linear classifiers that make use of distinct direct separators for classification. Support Vector Machines (SVM) which is used in neural networks makes use of supervised learning method to create decision planes that provide the decision boundaries specification. Decision planes consist of sets of objects that have different class membership. A linear classifier is a line that separates the sets of objects classified according to their corresponding spheres. If the partitioning is more than two dimensions, the classifier is a curve that is known as the hyperplane classifier. The SVM produces 0 or 1 which is positive or negative from the organized input in a vector space using its portrayed information. The content that is archived in a unique space cannot be used for learning. It has to be configured for calculation using ML algorithms. The pre-preparation of the content archive involves the transformation of every word to measurement and indistinguishable words will relate to the same measurement. For content classification, SVM has been proved to be useful for viable learning calculations.

Rule-based Technique When an 'if-then' rule is applied to a relationship that consists of an antecedent and its corresponding consequent, then it is a Rule-based Technique.

Antecedent ->consequent

An antecedent pronounces a state and can be symbolized as a single token or several tokens that are connected by the "$\wedge$" operator. A token can be either "?" which symbolizes a proper noun or a word or the token can be "#" that denotes the consequence of the state described by the antecedent.

$$\{token1 \wedge token2 \wedge ... \wedge token\} \Rightarrow \{+|-\}$$

The three simple rules A, B and C relates to words describing three sentiments, each denoting an antecedent.

{Good} ⇒ {Positive}… … (A)
{Bad} ⇒ {Negative}…..… (B)
{Ok} ⇒ {Neutral}……..… (C)

Probabilistic classifiers: Probabilistic classifiers also known as the generative classifier makes use of a mix of models that generatively examines each section for a particular term using the mix and classifies the sections accordingly. The three probabilistic classifiers are Naïve Byes, Maximum Entropy and Bayesian Network.

Naive Bayes Classifier (NB): NB is a probabilistic classifier that relies on upon Bayes hypothesis with solid and innocent freedom suppositions. It is a champion among the most principal content order procedures with different applications such as archive classification, email spam discovery, individual email sorting, and slant recognition and dialect identification. It performs better in numerous perplexing genuine issues. The Naive Bayes Variations are Bernoulli Naive Bayes, Binarized Multinomial Naive Bayes and the Multinomial Naive Bayes. Each framework passes on a very surprising result since they use a unique model. When the different events of the words matter a great deal in the characterization issue multinomial naive bayes is utilized. Binarized multinomial naive bayes is used when the regularities of the words don't accept a key part in the arrangement.

Bayesian Network (BN): The Naïve Bayes classifier is the independence of the features. Assumption of Naïve Bayes is to expect that every one of the components is completely dependent. This prompts to the BN which shows a coordinated acyclic graph and whose nodes correspond random variables, and edges that represent conditional dependencies. BN is viewed as an entire factor with their association. In this way, an entire joint probability distribution over each one of the elements is resolved for a model. The computation complexity of the BN is exceptionally costly in text mining. So, BN is used very little. BN is utilized to consider a true issue.

Maximum Entropy (MaxEnt) Classifier: MaxEnt classifiers are feature-based models that are most preferred when it is required to uniformly satisfy a given constraint. Unlike the Naïve Bayes model, MaxEnt does not make any autonomous assumption of the features. So MaxEnt allows the addition of features like bigrams and phrases without the problem of feature overlapping. The MaxEnt principle is explicitly important in the case of information that needs to be examined to determine if a given distribution is consistent with the testable information. MaxEnt or its variations are useful because of their accuracy. Consistency results of the algorithm are another major advantage of it. The ability of the classifier is to work with a huge amount of data describes its performance or its efficiency. When it comes to handling different types of data in a single platform and classify them according to it, the flexibility of the classifier is almost perfect.

Unsupervised Learning: Unsupervised learning technique applies a comparison formula between sentiment values in a lexicon with the components of the text at hand. The words in the lexicon have predefined values and they are applied to similar words in the text. Two most commonly used unsupervised techniques are hierarchical clustering and partial clustering.

Semi-Supervised Learning: Semi-supervised learning involves the use of both the unsupervised and supervised techniques in its classification models. The semi-supervised learning makes use of both the labelled and some unlabelled data in its training sets. So the goal of semi-supervised learning could be to either to predict the values of the unlabelled data in the training set in which case it would be called transductive semi-supervised learning or to find the values of the test sets in which case it would be called the inductive semi-supervised learning.

2.2 Lexicon based Approaches

This approach depends on sentiment Lexicons. Lexicon is considered as an important indicator for sentiment, which is called opinion word. Lexicon can be separated into the dictionary approach and corpus-based approach

Dictionary-based approach: An arrangement of sentiment word is gathered manually with known instructions. The conclusion set is created by looking in the prominent repository WordNet for their proportionate word and antonyms. The next iteration starts when the words are added into the seed list. In the absence of finding any new words, the iterative process stops. After the system completes, a manual appraisal can be done to evacuate or amend errors. Dictionary approach has an inconvenience, that is, the inability to discover feeling words with space and setting specific introductions.

Corpus-based approach: Corpus-based method deals with the issue of finding sensitivity words with setting specific presentation. Corpus construct techniques depend on syntactic illustrations or examples that are found together with a semblance of assessment words to find other inferencing words in a broad corpus. The objectives are for conjunctions like AND, OR, BUT, EITHER E-XOR. The conjunction in AND case that are conjoined descriptive words, for the most part, have the comparative introduction. This contemplation is called notion consistency, which is for the most part not solid for all predictable practicality. There are also adversative expressions on account of BUT they are shown as opinion changes to figure out whether two conjoined descriptors are of the same or distinctive introductions. The corpus-based approach uses factual or semantic techniques to discover assumption extremity.

2.3 Hybrid Approaches

A hybrid approach is a combination of ML and Lexicon approaches used in the Sentimental Analysis. Although the hybrid approach may not be used frequently, they are known to produce results better than the approaches mentioned earlier. [1]

2.4 Ontology-based Approaches

Ontology applies the hierarchy of concepts to represent domain knowledge. An ontology can be defined as an unambiguous, machine-readable representation of a common formulation of concepts. An Ontology is used to represent knowledge formally by modelling the terms of a particular domain and by capturing the semantic relationship between the terms. The relationship between terms is particularly important for aspect-level sentiment analysis more specifically in product reviews since product reviews are usually qualified by their aspects. Ontological approaches are generally used to capture the relation between the product and their properties in a hierarchical order. Such ordering is done using definitions of fundamental Entities and interrelationship between the entities that can be interpreted by machines. An ontology model incorporates the entities, concepts, classes or objects, their properties and the relation between them. For the sharing of knowledge among researchers about a particular domain, the ontology provides common vocabulary definitions. Such ontological definitions not only help to share a common knowledge understanding among people and software agents but also allows the reusability of domain knowledge. Storage of knowledge in an ontology model is done using either OWL/XML or RDF/XML format. The merging of two ontology models which have the same domain knowledge is possible. Information retrieval from an ontology model is done by querying it. [2]

In the concept of sentiment analysis portraits different levels, various approaches, many classifications and evaluation methods are displayed in figure 2.1.

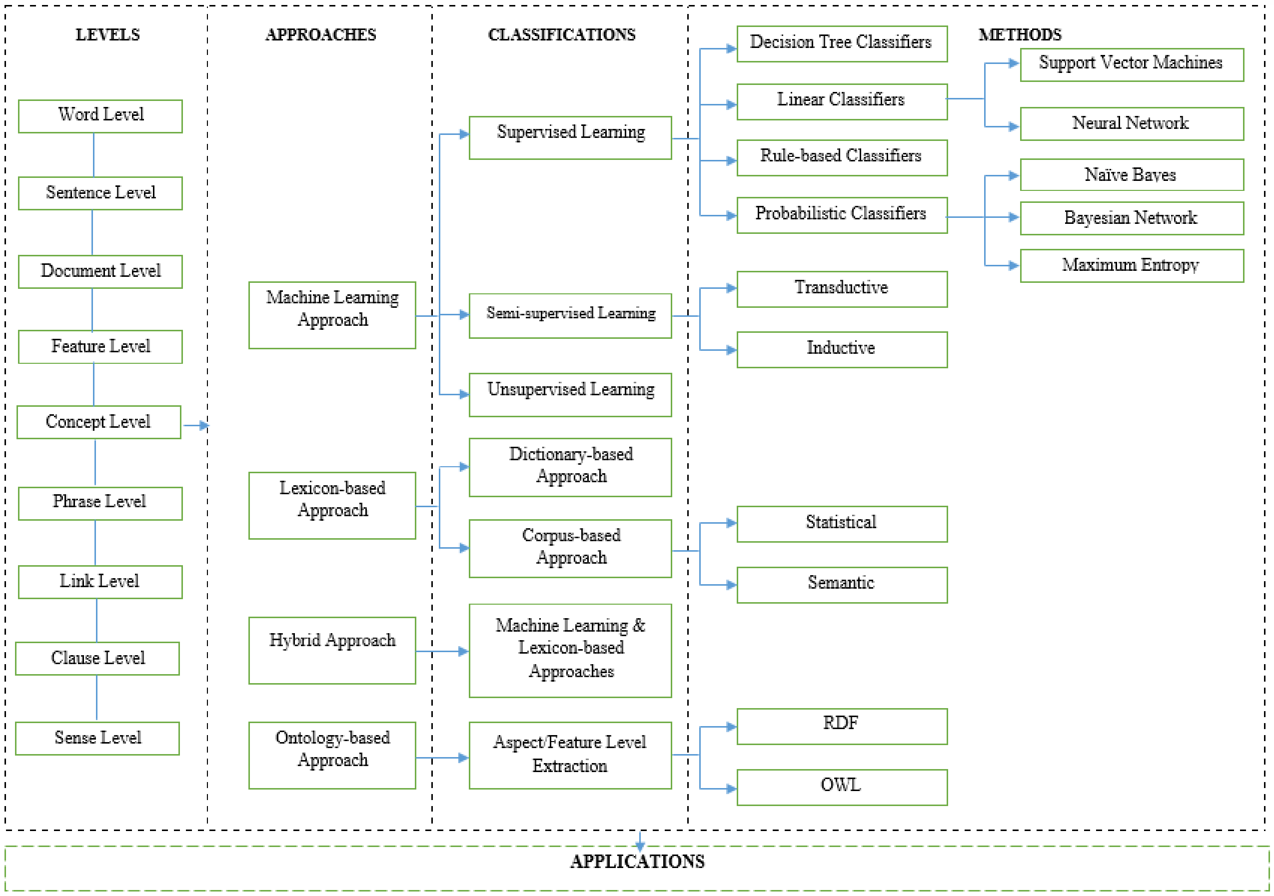

Figure 2.1 Levels of sentiment analysis, approaches, classifications and method

3. Related Works

Various sentimental approaches are surveyed [3-8] and these approaches are compared based on the criteria such as issues addressed, techniques, datasets used, accuracy and limitation [3]. Sentiment analysis tools are also surveyed [6].

The tools for Sentiment Analysis that are used in different domains were also studied analyzing its features and performance. SentiWordNet, LIWC, EMOTIONS, SenticNet, Happiness Index, AFINN, PANAS-t, Sentiment140 Dataset, NRC, EWGA, FRN [6], Mahout and Weka [9] are the sentiment analysis tools reported in the literature.

Twenty four articles are compared in the survey to study sentiment analysis for its importance, effects and the challenges in sentiment evaluation. The sentiment review structure and the challenges in sentiment analysis are initially compared. The domain-dependence of the sentiment challenges is an essential feature that is revealed by the comparison. The popularity of the negation challenges in the structure of all the reviews types differed in its implicit or explicit meaning. The result of such comparison provided a capacity to measure the effects of each sentiment challenge based on the structure of the review types. The topic, nature and the structure of the review regulate the appropriate challenges for the assessment of reviews on sentiment analysis. The next comparison that is done between the relevance of the challenges in sentiment analysis relevant with the accuracy rate. From this comparison, two things were evident, namely, the status of the sentiment challenges for evaluation of the sentiments and the method of finding the fitting challenge to improve accuracy. The relation between the amount of theoretical and technical use of sentiment techniques to resolve sentiment challenges is also found. It was also established and explained that the theoretical type of sentiment challenges is the growing area of research. By proving that with the growth in research in a sentiment analysis the average accuracy has reduced, the inference from the average of accuracy based on the number of researches in each challenge has been established. The compassion circle could be extended to new research literature for further work. [12].

4. Methodology

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) are a method of reporting a minimum set of items grounded on shreds of evidence in systematic reviews and meta-analyses. The PRISMA *flow diagram* in figure 4.2 exhibits how information moves along the different phases in *systematic reviews*. The number of records identified, included and excluded, and the reasons for exclusions are charted out in the PRISMA *flow diagram*. [13].

The following procedures of the systematic literature review were carried out to survey the current study and research in and around the subject of sentiment analysis.

4.1 Search Process

The search process included the search for pertinent research literature on the internet through websites, online libraries and databases like Springer, IEEE Xplore, Web of Science, Scopus, Science Direct, ACM Digital Library, Elsevier and Google Scholar, etc.. (Listed in Table 4.1). The vital terms or strings that were used for search are ‘‘survey on sentiment analysis”, ‘‘survey on opinion mining”, and “Art of survey on sentimental analysis”. It was found that searching the references of the research articles of the various studies with the three key terms yielded more articles.

Table 4.1

List of research articles searched in different databases.

Databases	Springer	IEEE Xplore	Web of Science	Scopus	Science Direct	Google Scholar /Proceedings	Elsevier	ACM Digital Library
No. of Research Articles	8	16	4	12	8	7	2	9

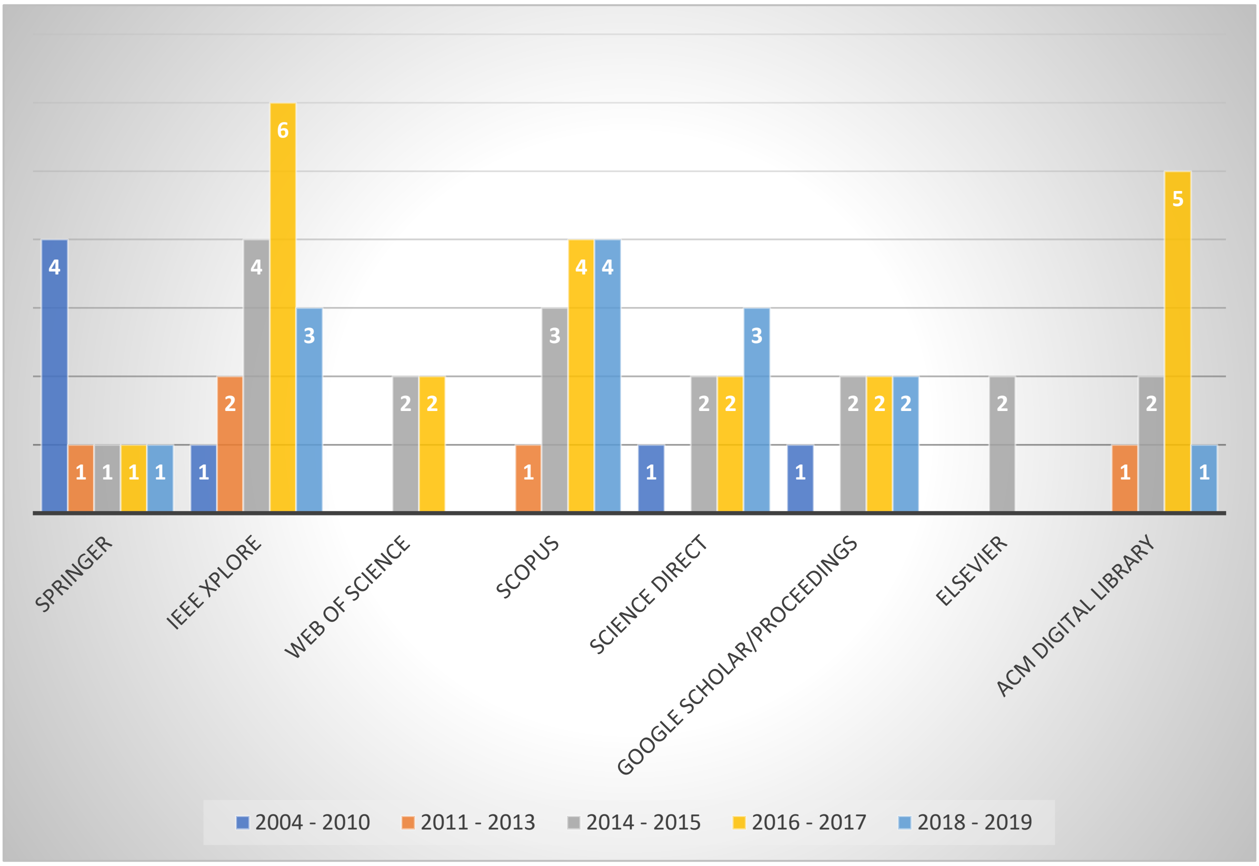

Figure 4.2: List of research articles searched in different databases – Year wise.

4.2 The criterion for including or excluding articles

The number of literature available on sentimental analysis is quite vast. To set the number of literature to that which can be managed or those with a focus, certain criteria were devised in the selection of the literature for review.

i) Criterion for Including – Only those literature on sentiment analysis or on tools and techniques for sentiment analysis that was published during the period 2005–2019 (mentioned in figure 4.2) are considered. If the same or similar content is published in more than one journal or conference proceeding, the most complete version of the literature was chosen.

ii) Criteria for Excluding – Works published in unknown conferences or journals, research articles that were not relevant to above the keyword search or not relevant to the survey and 'white' articles as mentioned in table 4.2.

4.3 Document Retrieval and Bibliography Management

The search process described above was used to identify the relevant literature and these were then checked by their title and abstract using the criteria of including or excluding literature. Once all the relevant research articles that were identified for the review process, they were downloaded extracting their data and to analyse them further. In Figure 4.2 is a PRISMA flowchart to illustrate the process of the search that was done for the selection for literature for review [13].

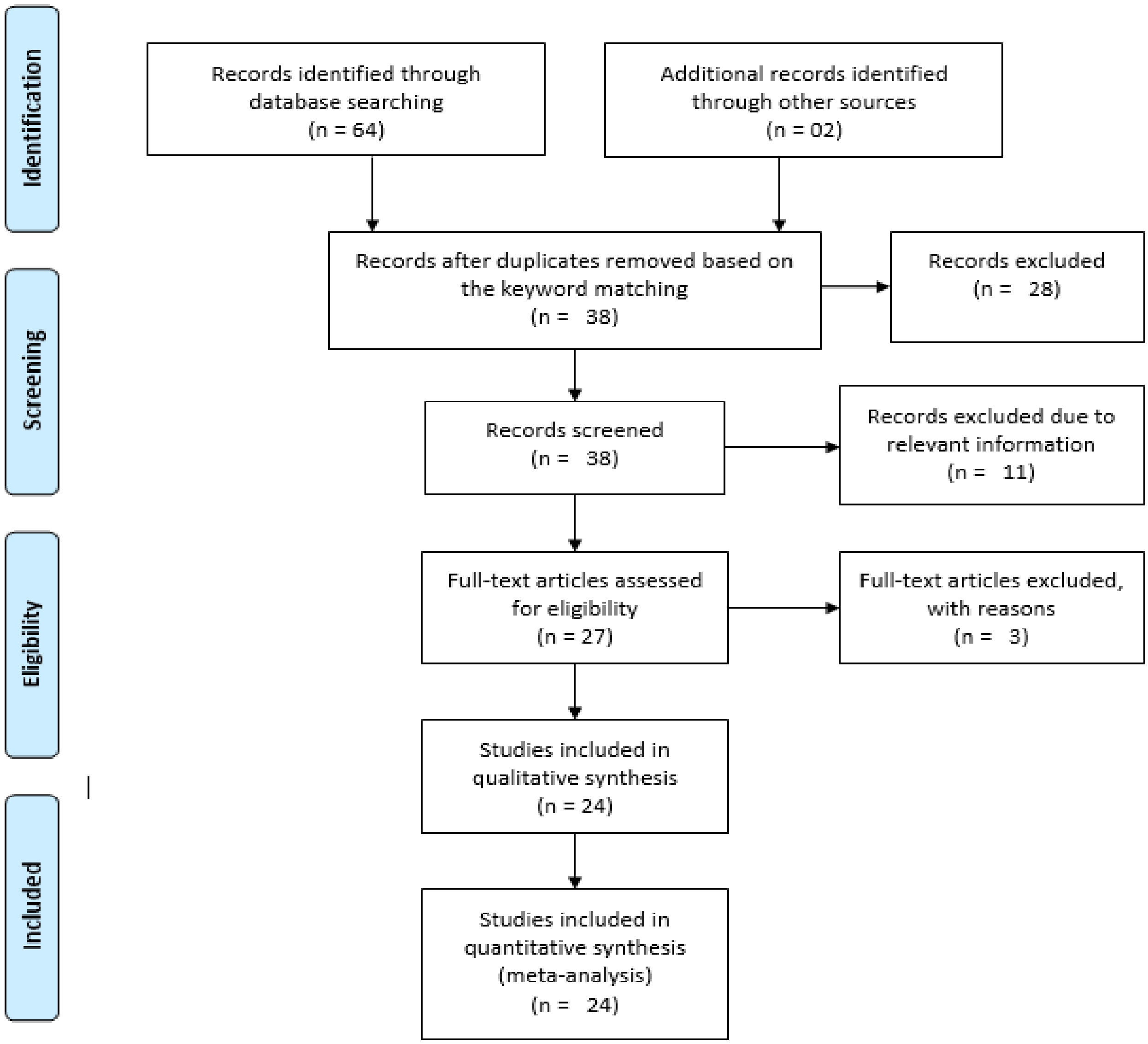

Figure 4.2 PRISMA flowchart for sentiment review of literature

Inclusion and exclusion criteria based on PRISMA Model as listed in the table 4.2.

Table 4.2
Findings and methodology of selected research literature

Reference No	Data Content	Modality	Sentiment Measurement	Findings (Research Issues)	Methodology /Approach
[1]	User Generated Data	Text	No	- Detection of Fake Reviews and Spams - Language Dependency	- Lexicon
[2]	Social Media Content	Text/Image	No	- Detection of Fake Reviews and Spams - Language Dependency	- Lexicon
[3]	User Generated Data	Text	No	- Detection of Fake Reviews and Spams - Language Dependency	- Lexicon

[4]	User Generated Data	Text	No	- Detection of Fake Reviews and Spams - Language Dependency	- Lexicon
[5]	User Generated Data	Text	No	- Detection of Fake Reviews and Spams - Language Dependency	- Lexicon
[6]	User Generated Data	Text	No	- Detection of Fake Reviews and Spams - Limitations in Classification Filtering - Domain Dependency	- Lexicon
[7]	User Generated Data	Text	No	- Domain Dependency - Comparisons - Negations - Sarcasm	- Lexicon
[8]	User Generated Data	Text	No	- Detection of Fake Reviews and Spams - Language Dependency	- Lexicon
[9]	User Generated Data	Text	No	- Detection of Fake Reviews and Spams - Language Dependency	- Lexicon
[10]	User Generated Data	Text	No	- Detection of Fake Reviews and Spams - Domain Dependency - Sarcasm - Interrogative Sentences - Sentiment without sentiment words - Conditional sentences	- Lexicon
[11]	User Generated Data	Text	No	- Detection of Fake Reviews and Spams - Language Dependency	- Lexicon
[12]	User Generated Data	Text	No	- Detection of Fake Reviews and Spams - Language Dependency	- Lexicon
[14]	User Generated Data	Text	No	- Detection of Fake Reviews and Spams - Language Dependency	- Lexicon
[15]	Facial Extraction (biometrics)	Image	No	Excluded in the PRISMA Model	- Lexicon
[16]	Online News Galleries	Text and Image	Yes		- Lexicon
[17]	Micro-Blogging Contents. Like Twitter and Messaging Systems	Text and Image	Yes		- Lexicon
[18]	User Review on Social Content and Geo Located Map	Text	Yes		- Lexicon
[19]	Audio and Video signals: Emotion Categorization	Audio, Video and Text	No		- Lexicon

[20]	Multimedia Content (Product Reviews)	Text	Yes		- Lexicon
[21]	User Posts (Reviews on Social Media Content)	Text	Yes	- Detection of Fake Reviews and Spams - Language Dependency	- Lexicon
[22]	Reviews, forum discussions, blogs and social networks	Text	Yes	- Detection of Fake Reviews and Spams - Language Dependency	- Lexicon
[23]	Video Streaming (Content-Based Retrieval)	Video	No	Excluded in the PRISMA Model	- Lexicon
[24]	User generated data	Text	No	- Detection of Fake Reviews and Spams - Language Dependency - Sentiment without sentiment words	- Lexicon
[25]	Social Media Contents	Emoji Symbols on User Reviews	No	Excluded in the PRISMA Model	- Lexicon
[26]	User-generated data	Text	Yes		- Lexicon
[27]	Research Articles	Text	Yes	- Detection of Fake Reviews and Spams - Language Dependency	- Lexicon
[28]	User Reviews on the Internet	Text	Yes	Excluded in the PRISMA Model	- Lexicon
[29]	User Reviews on Social Media	Text	Yes		- Lexicon
[30]	Extracting the Text on Image on Social Media	Text	No		- Lexicon
[31]	User-generated data	Text	No		- Lexicon
[32]	Multimedia Contents	Text, Audio and Video	No		- Lexicon
[33]	Twitter Data	Text	Yes	- Detection of Fake Reviews and Spams - Language Dependency	- Lexicon
[34]	Healthcare Review Information	Text and Image	No	Excluded in PRISMA Model	- Lexicon
[35]	Multimedia Signals	Image and Video	No		- Lexicon
[36]	User opinions on Social Media	Text	Yes		- Lexicon

[37]	User communication (Speech information)	Audio	No		- Lexicon
[38]	User Reviews (Based on Text, Speech and Visual)	Text	Yes		- Lexicon
[39]	User generated data	Text, Image	No		- Lexicon
[40]	User generated content	Text, Audio	No		- Lexicon
[41]	Audio Fusion Information	Audio	No		- Lexicon
[42]	User generated content	Text	No		- Lexicon
[43]	Multimedia content	Text, Image	No		- Lexicon
[44]	User generated content	Image and Video	No		- Lexicon
[45]	Image Dataset	Image	No		- Lexicon
[46]	User generated content	Image	No		- Lexicon
[47]	User generated data	Text	No		- Lexicon
[48]	Multimedia Contents	Text and Image	No		- Lexicon
[49]	User generated content	Text	No	- Detection of Fake Reviews and Spams - Language Dependency	- Lexicon
[50]	Twitter Data	Text	Yes	- Detection of Fake Reviews and Spams - Language Dependency - Sentiment without sentiment words	- Lexicon
[51]	YouTube - User generated Reviews	Text	Yes	- Detection of Fake Reviews and Spams - Language Dependency	- Lexicon
[52]	User generated content	Text	No	Excluded in the PRISMA Model	- Lexicon
[53]	Textual and Visual Data in Social Media	Text and Image	No		- Lexicon
[54]	User generated content	Text	No		- Lexicon
[55]	User generated content	Text	No	- Detection of Fake Reviews and Spams - Language Dependency	- Lexicon
[56]	User generated content on Web	Text and Visual Data	No	Excluded in the PRISMA Model	- Lexicon
[57]	Image content in social media	Image	No		- Lexicon
[58]	Microblog Image content	Image	No		- Lexicon
[59]	User generated content	Text, Image	No		- Lexicon

[60]	User generated content	Text	Yes		- Lexicon
[61]	User generated Image	Image	No		- Lexicon
[62]	Twitter Data	Emoji	Yes		- Lexicon
[63]	Text-based Survey	Text	No	- Detection of Fake Reviews and Spams - Language Dependency - Sentiment without sentiment words	- Lexicon
[64]	User generated Images	Image	No	Excluded in the PRISMA Model	- Lexicon
[65]	User generated Review Text	Text	No		- Lexicon
[66]	Pre – Processing the text-based of tokenization				

5. Issues, Challenges and Research Gap

Having studied the contents of the literature on Sentimental Analysis, the following issues and challenges were evident:

Finding Fake Reviews and Spams: Internet contents especially in social media comprises both authentic and spam contents. Therefore, it is necessary to remove the spam and fake contents before pre-processing. [4] [6] [10]

Limitations in Classification Filtering: Unrelated opinions are removed to establish the most popular opinion. The classification filtering techniques do not provide the expected results. [6]

Language Dependency: Maximum of the work focused only on the English text based content and thus, most of the resources are available in English. [4]

Domain Dependency: Opinion mining depends on the domain text used. [6] [10]

Word Sense Disambiguation: Exact meaning of a word based on the context needs to be extracted as words can have different meanings for different fields. [7]

Comparisons: To decide the polarity for relative sentences can be a challenge. It is challenging to find the highly positive or highly negative rating based on the intensity of opinion that is given. It is called a degree of polarity. [7]

Negations: If the negations are not handled properly can give completely wrong results. [7]

Sarcasm: Identify and analyze emotions voiced in the text at a more fine-grained level. [7] [10]

Interrogative Sentences: When dealing with question type sentences, the sentence itself may not contain positive or negative sentiments but the keywords used in such a sentence might express positive or negative sentiments. [10]

Sentiment without sentiment words: At times there could be sentences without any of the keywords that express sentiments such good, better, best, worst, bad and so on but the sentences itself might be used to express positive or negative feedback about some particular product, service or policy. [10]

Conditional sentences: Conditional sentences create problems similar to interrogative sentences and therefore are a challenge in sentimental analysis. [10]

The following two important issues were not tackled in the studies that were done so far:

Multimedia Content – so far the above survey pointed out the 99% of content taken for the sentimental analysis based on text. Few articles only tell about the multimedia content like audio, image and video-based sentiment analysis based on lexicon approach.

Machine Learning Approach – different approaches proposed for sentiment analysis (like Lexicon based, Machine learning based and Ontology-based), but no implementation based on multimedia content using machine learning.

6. Conclusion

There has been a lot of attention recently on sentiment analysis especially from researchers from the fields of text mining and natural language processing. But due to extreme absence of annotated datasets which is used to train models in various domains, the accuracy of sentiment analysis has been hindered. Several kinds of research have been done out to confront the challenge and enhance sentiment analysis classification. Sentiment analysis is important as it helps in identifying the emotional and attitude states of people. Positive or negative feelings of people can be expressed in different ways. This research article talks about, in subtle terms, the different ways to deal with sentiment analysis mostly in Machine Learning, Lexicon-based, Hybrid and Ontology-based approaches. This research article gives a point by point perspective of the distinctive applications and challenges of Sentiment Analysis.

References

[1] Megha Joshi, Purvi Prajapati, Ayesha Shaikh and Vishwa Vala. “A Survey on Sentiment Analysis”. International Journal of Computer Applications - 163(6):34-38, April 2017.

[2] Thakor, Pratik and S. Sasi. “Ontology-based Sentiment Analysis Process for Social Media Content.” INNS Conference on Big Data (2015).

[3] Abirami, A. and V. Gayathri. “A survey on sentiment analysis methods and approach.” 2016 Eighth International Conference on Advanced Computing (ICoAC) (2017): 72-76.

[4] Jain, S. and P. Singh. “Systematic Survey on Sentiment Analysis.” 2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC) (2018): 561-565.

[5] Kaur, Harpreet & Mangat, Veenu & Nidhi,. (2017). A survey of sentiment analysis techniques. 921-925. 10.1109/I-SMAC.2017.8058315.

[6] S. Rajalakshmi, S. Asha and N. Pazhaniraja, "A comprehensive survey on sentiment analysis," 2017 Fourth International Conference on Signal Processing, Communication and Networking (ICSCN)*, Chennai, 2017,* pp.1-5, doi: 10.1109/ICSCN.2017.8085673.

[7] Kolkur, Seema et al. “Study of Different Levels for Sentiment Analysis.” (2015).

[8] K.H. Rizwana, Dr. B. Kalpana, “A Survey on Sentiment Analysis and Opinion Mining”. International Journal of Pure and Applied Mathematics, Volume 118, No.18, 2681-2688, ISSN: 1311-8080 (printed version), ISSN: 1314-3395 (online version), 2018.

[9] Jyoti and Seema Rao. 2016. A Survey on Sentiment Analysis and Opinion Mining. Article 53, 1–5. DOI:https://doi.org/10.1145/2979779. 2979832

[10] M., Vidisha & Vala, Jay & Balani, Prem. (2016). A Survey on Sentiment Analysis Algorithms for Opinion Mining. International Journal of Computer Applications. 133. 7-11. 10.5120/ijca2016907977.

[11] S. Vasanthraj, A. Martin, A. Meiappane, J. Madhusudnan, “A Survey on Sentiment Analysis Applied in Opinion Mining”, Journal of Network Communication and Emerging Technologies (JNCET), Volume 1, Issue 1, ISSN: 2395-5317, 2015.

[12] Doaa Mohey El-Din Mohamed Hussein,A survey on sentiment

analysis challenges, Journal of King Saud University - Engineering Sciences, Volume 30, Issue 4, 2018, Pages 330-338, ISSN 1018-3639, https://doi.org/10.1016/j.jksues.2016.04.002.

[13] PRISMA Model - http://www.prisma-statement.org/PRISMAStatement/FlowDiagram.

[14] T. Al-Moslmi, N. Omar, S. Abdullah and M. Albared, "Approaches to Cross-Domain Sentiment Analysis: A Systematic Literature Review," in *IEEE Access*, vol. 5, pp. 16173-16192, 2017, doi: 10.1109/ACCESS.2017.2690342.

[15] Zheng, Yufeng & Elmaghraby, Adel. (2011). A brief survey on multispectral face recognition and multimodal score fusion. 543-550. 10.1109/ISSPIT.2011.6151622.

[16] Helen Caple & John S. Knox (2015) A framework for the multimodal analysis of online news galleries: What makes a “good” picture gallery?, Social Semiotics, 25:3, 292321, DOI: 10.1080/10350330.2014.1002174

[17] Baecchi, C., Uricchio, T., Bertini, M. *et al.* A multimodal feature learning approach for sentiment analysis of social network multimedia. *Multimed Tools Appl* **75,** 2507–2525 (2016). https://doi.org/10.1007/s11042-015-2646-x

[18] Evangelia Polymerou, Despoina Chatzakou, and Athena Vakali. 2014. EmoTube: A Sentiment Analysis Integrated Environment for Social Web Content. Article 20, 1–6. DOI:https://doi.org/10.1145/2611040.2611081

[19] Sikandar, M.. “A Survey for Multimodal Sentiment Analysis Methods.” (2014).

[20] Mohammad Soleymani, David Garcia, Brendan Jou, Björn Schuller, Shih-Fu Chang, Maja Pantic, A survey of multimodal sentiment analysis, Image and Vision Computing, Volume 65, 2017, Pages 3-14, ISSN 0262-8856, https://doi.org/10.1016/j.imavis.2017.08.003.

[21] Lin Yue, Weitong Chen, Xue Li, Wanli Zuo, and Minghao Yin. 2019. A survey of sentiment analysis in social media. *Knowl. Inf. Syst.* 60, 2 (August 2019), 617–663. DOI:https://doi.org/10.1007/s10115-018-1236-4

[22] Joshi, N. and Suhasini A. Itkat. “A Survey on Feature Level Sentiment Analysis.” (2014).

[23] Calic, Janko & Campbell, N. & Dasiopoulou, Stamatia & Kompatsiaris, Ioannis. (2005). A Survey on Multimodal Video Representation for Semantic Retrieval. 1. 135 - 138. 10.1109/EURCON.2005.1629877.

[24] Hussein, D.-M.E.D.M. “A survey on sentiment analysis challenges”. Journal of King Saud University – Engineering Sciences (2016), http://dx.doi.org/10.1016/j.jksues.2016.04.002.

[25] G. Varni, G. Volpe, and A. Camurri. 2010. A System for Real-Time Multimodal Analysis of Nonverbal Affective Social Interaction in User-Centric Media. *Trans. Multi.* 12, 6 (October 2010), 576–590. DOI:https://doi.org/10.1109/TMM.2010.2052592.

[26] D'Andrea, Alessia & Ferri, Fernando & Grifoni, Patrizia & Guzzo, Tiziana. (2015). Approaches, Tools and Applications for Sentiment Analysis Implementation. International Journal of Computer Applications. 125. 26-33. 10.5120/ijca2015905866.

[27] Toqir A. Rana and Yu-N Cheah. 2016. Aspect extraction in sentiment analysis: comparative analysis and survey. *Artif. Intell. Rev.* 46, 4 (December 2016), 459–483. DOI:https://doi.org/10.1007/s10462-016-9472-z.

[28] M., Vidisha & Vala, Jay & Balani, Prem. (2016). A Survey on Sentiment Analysis Algorithms for Opinion

Mining. International Journal of Computer Applications. 133. 7-11. 10.5120/ijca2016907977.

[29] Siti Salwa Hasbullah, Diana Maynard, Rita Zaharah Wan Chik, Farahwahida Mohd, and Myzan Noor. 2016. Automated Content Analysis: A Sentiment Analysis on Malaysian Government Social Media. Article 30, 1–6. DOI:https://doi.org/10.1145/2857546.2857577

[30] R. Ji, D. Cao and D. Lin, "Cross-Modality Sentiment Analysis for Social Multimedia," in 2015 IEEE International Conference on Multimedia Big Data (BigMM), Beijing, China, 2015 pp. 28-31. doi: 10.1109/BigMM.2015.85.

[31] Zhang, L, Wang, S, Liu, B. Deep learning for sentiment analysis**:** A survey. *WIREs Data Mining Knowl Discov.* 2018; 8:e1253. https://doi.org/10.1002/widm.1253

[32] Poria S, Cambria E, Howard N, Huang G & Hussain A (2016) Fusing audio, visual and textual clues for sentiment analysis from multimodal content, Neurocomputing, 174 (A), pp. 50-59.

[33] Giachanou, Anastasia & Crestani, Fabio. (2016). Like It or Not: A Survey of Twitter Sentiment Analysis Methods. ACM Computing Surveys. 49. 1-41. 10.1145/2938640.

[34] R. Pang, A. Baretto, H. Kautz and J. Luo, "Monitoring adolescent alcohol use via multimodal analysis in social multimedia," 2015 IEEE International Conference on Big Data (Big Data), Santa Clara, CA, 2015, pp. 1509-1518, doi: 10.1109/BigData.2015.7363914.

[35] Camurri A., Mazzarino B., Ricchetti M., Timmers R., Volpe G. (2004) Multimodal Analysis of Expressive Gesture in Music and Dance Performances. In: Camurri A., Volpe G. (eds) Gesture-Based Communication in Human-Computer Interaction. GW 2003. Lecture Notes in Computer Science, vol 2915. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-24598-8_3.

[36] Shah, Rajiv Ratn. (2016). Multimodal Analysis of User-Generated Content in Support of Social Media Applications. 423-426. 10.1145/2911996.2912032.

[37] Nicu Sebe, Ira Cohen, Theo Gevers , and Thomas S. Huang, "Multimodal Approaches for Emotion Recognition: A Survey", SPIE 5670, Internet Imaging VI, (17 January 2005); doi: 10.1117/12.600746

[38] Cummins, Nicholas & Amiriparian, Shahin & Ottl, Sandra & Gerczuk, Maurice & Schmitt, Maximilian & Schuller, Björn. (2018). Multimodal Bag-of-Words for Cross Domains Sentiment Analysis. 10.1109/ICASSP.2018.8462660.

[39] Atrey, P.K., Hossain, M.A., El Saddik, A. *et al.* Multimodal fusion for multimedia analysis: a survey. *Multimedia Systems* **16,** 345–379 (2010).https://doi.org/10.1007/s00530-010-0182-0.

[40] Jaimes A., Sebe N. (2005) Multimodal Human Computer Interaction: A Survey. In: Sebe N., Lew M., Huang T.S. (eds) Computer Vision in Human-Computer Interaction. HCI 2005. Lecture Notes in Computer Science, vol 3766. Springer, Berlin, Heidelberg.https://doi.org/10.1007/11573425_1.

[41] Multimodal information fusion of audio emotion recognition based on kernel entropy component analysis zhibing xie and ling guan International Journal of Semantic Computing 2013 07:01, 25-42.

[42] Dumas B., Lalanne D., Oviatt S. (2009) Multimodal Interfaces: A Survey of Principles, Models and Frameworks. In: Lalanne D., Kohlas J. (eds) Human Machine Interaction. Lecture Notes in Computer Science, vol 5440. Springer, Berlin, Heidelberg.https://doi.org/10.1007/978-3-642-00437-7_1.

[43] Tadas Baltrusaitis, Chaitanya Ahuja, and Louis-Philippe Morency. 2019. Multimodal Machine Learning: A Survey and Taxonomy. *IEEE Trans. Pattern Anal. Mach. Intell.* 41, 2 (February 2019), 423–443.

DOI:https://doi.org/10.1109/TPAMI.2018.2798607.

[44] Shah, Rajiv Ratn. (2016). Multimodal-based Multimedia Analysis, Retrieval, and Services in Support of Social Media Applications. 10.1145/2964284.2971471.

[45] Tao Chen, Felix X. Yu, Jiawei Chen, Yin Cui, Yan-Ying Chen, and Shih-Fu Chang. 2014. Object-Based Visual Sentiment Concept Analysis and Application. In Proceedings of the 22nd ACM international conference on Multimedia (MM '14). Association for Computing Machinery, New York, NY, USA, 367–376. DOI:https://doi.org/10.1145/2647868.2654935

[46] Yan-Ying Chen, Tao Chen, Winston H. Hsu, Hong-Yuan Mark Liao, and Shih-Fu Chang. 2014. Predicting Viewer Affective Comments Based on Image Content in Social Media. In Proceedings of International Conference on Multimedia Retrieval (ICMR '14). Association for Computing Machinery, New York, NY, USA, 233–240. DOI:https://doi.org/10.1145/2578726.2578756.

[47] Swagatam Das, Sayan Maity, Bo-Yang Qu, P.N. Suganthan, Real-parameter evolutionary multimodal optimization — A survey of the state-of-the-art, Swarm and Evolutionary Computation, Volume 1, Issue 2, 2011, Pages 71-88, ISSN 2210-6502, https://doi.org/10.1016/j.swevo.2011.05.005.

[48] H. Wang, A. Meghawat, L. Morency and E. P. Xing, "Select-additive learning: Improving generalization in multimodal sentiment analysis," 2017 IEEE International Conference on Multimedia and Expo (ICME), Hong Kong, 2017, pp. 949-954, doi: 10.1109/ICME.2017.8019301.

[49] Walaa Medhat, Ahmed Hassan, Hoda Korashy, Sentiment analysis algorithms and applications: A survey, Ain Shams Engineering Journal,Volume 5, Issue 4, 2014, Pages 1093-1113, ISSN 2090-4479, https://doi.org/10.1016/j.asej.2014.04.011.

[50] Kharde, Vishal & Sonawane, Sheetal. (2016). Sentiment Analysis of Twitter Data: A Survey of Techniques. International Journal of Computer Applications. 139. 5-15. 10.5120/ijca2016908625.

[51] Asghar, Dr. Muhammad & Ahmad, Shakeel & Marwat, Afsana & Kundi, Fazal. (2015). Sentiment Analysis on YouTube: A Brief Survey. Brist. 3. 1250-1257.

[52] Yuan J., You Q., Luo J. (2015) Sentiment Analysis Using Social Multimedia. In: Baughman A., Gao J., Pan JY., Petrushin V. (eds) Multimedia Data Mining and Analytics. Springer, Cham. https://doi.org/10.1007/978-3-319-14998-1_2

[53] Quanzeng You. 2016. Sentiment and Emotion Analysis for Social Multimedia: Methodologies and Applications. In Proceedings of the 24th ACM international conference on Multimedia (MM '16). Association for Computing Machinery, New York, NY, USA, 1445–1449. DOI:https://doi.org/10.1145/2964284.2971475

[54] K. Schouten and F. Frasincar, "Survey on Aspect-Level Sentiment Analysis," in IEEE Transactions on Knowledge and Data Engineering, vol. 28, no. 3, pp. 813-830, 1 March 2016, doi: 10.1109/TKDE.2015.2485209.

[55] M. Desai and M. A. Mehta, "Techniques for sentiment analysis of Twitter data: A comprehensive survey," 2016 International Conference on Computing, Communication and Automation *(ICCCA)*, Noida, 2016, pp. 149-154, doi: 10.1109/CCAA.2016.7813707.

[56] Louis-Philippe Morency, Rada Mihalcea, and Payal Doshi. 2011. “towards multimodal sentiment analysis: harvesting opinions from the web”. Association for Computing Machinery, New York, NY, USA, 169–176. DOI:https://doi.org/10.1145/2070481.2070509.

[57] J. Islam and Y. Zhang, "Visual sentiment analysis for social images using transfer learning approach", pp. 124-130,doi:10.1109/BDCloud-SocialCom-SustainCom.2016.29.

[58] Cao, D., Ji, R., Lin, D. *et al.* "Visual sentiment topic model based microblog image sentiment analysis". Multimed Tools Appl 75**,** 8955–8968 (2016).https://doi.org/10.1007/s11042-014-2337-z.

[59] J.Kehrer and H.Hauser, "Visualization and Visual Analysis of Multifaceted Scientific Data: A Survey," in IEEE Transactions on Visualization and Computer Graphics, vol. 19, no. 3, pp. 495-513, March2013,doi:10.1109/TVCG.2012.110.

[60] Q. Fang, C. Xu, J. Sang, M. S. Hossain and G. Muhammad, "Word-of-Mouth Understanding: Entity-Centric Multimodal Aspect-Opinion Mining in Social Media," in IEEE Transactions on Multimedia, vol. 17, no. 12, pp. 2281-2296, Dec. 2015, doi: 10.1109/TMM.2015.2491019.

[61] Christian Otto, Matthias Springstein, Avishek Anand, and Ralph Ewerth. 2019. Understanding, Categorizing and Predicting Semantic Image-Text Relations. In Proceedings of the 2019 on International Conference on Multimedia Retrieval (ICMR '19). Association for Computing Machinery, New York, NY, USA, 168–176. DOI:https://doi.org/10.1145/3323873.3325049.

[62] Ziad Al-Halah, Andrew Aitken, Wenzhe Shi, Jose Caballero, "Smile, be Happy: Emoji Embedding for Visual Sentiment Analysis", Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2019.

[63] Saxena Diksha, Gupta Shubham, Jos eph Joy, Prof. Mehra Richa, "Sentiment Analysis", International Journal of Engineering, Science and Mathematics Vol. 8 Issue 3, March 2019,ISSN:2320-0294 Impact Factor: 6.765.

[64] Eric C. Yang, "Development of an integrated multimodal optical imaging system with real-time image analysis for the evaluation of oral premalignant lesions, J. Biomed. Opt. 24(2) 025003 (21 February 2019) https://doi.org/10.1117/1.JBO.24.2.025003.

[65] Sankar, H, Subramaniyaswamy, V, Vijayakumar, V, Arun Kumar, S, Logesh, R, Umamakeswari, A. Intelligent sentiment analysis approach using edge computing-based deep learning technique. Softw: Pract Exper. 2020; 50: 645–657.https://doi.org/10.1002/spe.2687

[66] Vijayarani, S., Ms J. Ilamathi, and Ms Nithya. "Preprocessing techniques for text mining-an overview." International Journal of Computer Science & Communication Networks 5.1 (2015): 7-16.