Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 44 No. 5 (2023)

A Perspective Study on Mining Techniques for Sentiment Analysis

[1]Dr.C.Victoria Priscilla, [2]Mrs .P.Shareefa Parveen

[1]Department Of Computer Science
SDNB Vaishnav College For Women, University Of Madras, Chennai, India.
[2]Department Of Computer Science
JBAS College for Women, University Of Madras, Chennai, India.

Abstract: "In the past 20 years, the landscape of social media has undergone substantial growth, emerging as a prominent platform that attracts a substantial and diverse audience. Within this digital realm, a wealth of opinionated content is disseminated, comprising various forms of data, including emails, user feedback, tweets, posts, pins, web content, and textual information such as product sentiments. The influx of this unstructured data across multiple domains has spurred the demand for data mining. Data mining, in turn, allows for the identification of valuable patterns within this vast and unstructured dataset. An area of considerable interest in the field of Natural Language Processing (NLP) is sentiment analysis, wherein the sentiments expressed in text are analyzed.

In this context, this study not only provides a concise overview of social networking platforms but also conducts an in-depth examination of the methodologies and tools employed in sentiment analysis. Furthermore, potential limitations are scrutinized, paving the way for opportunities to advance research in the future."

Keywords: Data Mining Techniques, Social Media, Social Network, Sentiment Analysis, Opinion Mining

1. Introduction

Social media is a web-based technology that makes it easier for many individuals to communicate socially through a network. Due to the most recent technological revolution, social media is expanding quickly and becoming an integral part of daily life. The astounding growth is a result of more people using smartphones, such as BlackBerry, Android, and iPhones. With these smartphones, nearly anyone can easily access any social networking platform. These social networking platforms' smartphone versions are incredibly user-friendly because they are so simple to use. Additionally, the use of Map services on mobile devices for finding directions and locations was impressive.

Six Degrees, the first well-known social media platform, was established in 1997. Users can upload a profile and friend other users using it. The first blogging platforms rose to prominence in 1999, sparking a social media phenomenon that endures today. [1]

2. Social Networking Sites:

According to Statista estimates, using social media will rank among the top internet activities on May 29, 2022. Around the world, there were 3.6 billion users of social media in 2020, and by 2025, that number is expected to reach approximately 4.41 billion. [2]

Fig 1: Social Networking Sites

Facebook: The dominant brand The most widely used social network globally is Facebook, which was the first to cross one billion registered accounts and has about 2.5 billion monthly active members [2].

Instagram: A photo-sharing app with around 1 billion active accounts per month[2].

WhatsApp: This app has made sharing and communicating instantly possible. Users per month: 1 billion approx[2].

The other newest social media platforms are Triller, WT, SocialValence, Flip, Popbase, Elpha,Yubo, Peanut, HouseParty, Caffeine, Steemit, Goodreads, Twitch, CaringBridge, WattPad, Crunchyroll, Soundcloud, Mocospace, CouchSurfing, italki, Medium, Ello, Vimeo, Giphy, Tribe, Kuaishou, Imgur, Influenster, FilmAffinity, Open Diary, Bubbly[2][Figure 1].

3. Data Mining And Its Techniques

3.1 Data Mining

Due to the vast amount of data that is kept in files, databases, and other repositories, it is becoming more and more crucial, if not absolutely necessary, to create effective tools for data analysis, interpretation, and the extraction of knowledge that may be useful for decision-making. Data mining, commonly referred to as Knowledge Discovery in Databases (KDD), "is the process of extracting implicit, undiscovered, and potentially relevant information from databases in a non-trivial way". [3]

3.2. Techniques:

The following are some of widely used data mining techniques in social media [Figure 2].

- Apriori algorithm
- KNN
- Decision Tree
- Association Rule
- Genetic Algorithms
- ANN
- K-Means
- Support Vector Machine
- Naive Bayes Classifier
- Page Rank Algorithm
- AdaBoost

Fig 2: Data Mining Techniques

4. Sentiment Analysis (SA)

As an alternative, opinion mining is a field of study that focuses on analysing people's feelings or attitudes on various subjects, events, people, issues, services, products, organisations, and their attributes [4]. Sentiment analysis is viewed as a subfield of computational linguistics, natural language processing, data mining, machine learning, and which also incorporates sociology and psychology. Although the history of natural language processing (NLP) dates back to the 1950s, sentiment analysis and people's opinions received little attention until the 2005s. The popularity of social media has fueled sentiment analysis's growth over the past few years.

We will primarily address three key questions in this section: the significance of sentiment analysis, the need for this survey, additionally, this survey's contributions.

4.1 Significance of Sentiment Analysis

Sentiment analysis has become more important as the amount of information available on social media has grown. From a business standpoint, sentiment analysis can offer online recommendations and guidance to both customers and businesses. On the one hand, e-commerce platforms can leverage customer preferences shown by the data to examine their goods and services. On the other hand, due to the virtual aspect of online buying, it can be challenging to fully and objectively grasp a product and determine whether a customer is open to hearing what other customers have to say.

Another significant element from a political standpoint is the enormous demand for political information. People don't just use the internet to express or seek their opinions for commercial purposes. Online social media sites like Twitter and Facebook were seen to be major factors in the onset and spread of the events. Authorities can find this kind of sensitive information in advance with the aid of sentiment analysis. Shutting down Internet communication routes, for example, would deny advocates of terrorism access to such services.

In conclusion, sentiment analysis and opinion mining tasks are crucial for a variety of reasons, including the traditional consumer and businesses conducting surveys to learn what customers think of corresponding goods or services. They are also crucial for national security and the analysis of public opinion.

4.2 Need Of this Survey?

Current studies on SA examine technical specifics or concentrate on particular sentiment analysis facets. Those are now obsolete due to the field's quick progress. Additionally, there is no survey that contains all information regarding algorithms, common tools, approaches etc.,

The work of Pang et al. [5], which sought to address the issues raised by sentiment-aware applications and included information on issues regarding manipulation, privacy, and the economic impact that is produced by opinion-oriented information-access services, has directly enabled opinion-oriented information-access applications. Medhat et al. [6] has categorized articles according to the techniques used, which can help the researchers who are familiar with certain techniques to use them in the sentiment analysis and choose the appropriate technique for a certain application. Wiegand et al.[7] survey is on the role of negation in sentiment analysis, in which various computational approaches modeling negation are introduced. In particular, this work focuses on aspects such as negation word detection, scope of negation, and limitation and challenges of negation modeling.

Ravi et al.[8] is organized on the basis of sub-tasks to be performed, that is to say, machine learning, natural language processing techniques, and applications of sentiment analysis. Saif et al. [9] presented an overview of eight publicly available manually annotated evaluation datasets for Twitter sentiment analysis and a common limitation of most of these datasets. Vinodhini et al. [10] presented a short survey that covers the techniques and challenges appeared in the field of sentiment analysis. Tsytsarau et al[11] reviewed the development of sentiment analysis and also discussed the gradual progress of a research direction, namely contradiction analysis.

Schouten et al.[12] focused on aspect-level sentiment analysis, to find and aggregate sentiment on entities mentioned within documents or aspects of these entities. Tang et al.[13] discussed related issues and main approaches to word sentiment classification, subjectivity classification, opinion extraction, and document sentiment classification. Giachanou et al. [14]investigated and briefly described the algorithms of sentiment analysis in Twitter, in which researcher discussed tasks related to Twitter opinion retrieval, tracking sentiments over time, irony detection, emotion detection, and tweet sentiment quantification. Of all those surveys proposed , Liu [15] is regarded as an encyclopedia on sentiment analysis and opinion mining .This article has summed up all important research topics in the field of sentiment analysis,. With more than 400 bibliographic references the author has summed up all details.

4.3 Contributions of this survey

Giving a thorough introduction and presenting fresh perspectives on this topic are the objectives of this paper. In conclusion, the article's contributions are as follows:

We examined the literature in the field of sentiment analysis from a variety of angles and listed the advantages and disadvantages of different methodologies. With brief descriptions of the algorithms and their originating sources, various sentiment analysis methodologies are grouped. Beginners interested in sentiment analysis will find this work useful for giving them a broad overview of the entire research area.

5. SENTIMENT ANALYSIS TASKS

5.1 Granularity-Oriented Sentiment Analysis:

Sentiment analysis tasks can be divided into three categories based on granularity:

- Word level
- Sentence level
- Document level

5.1.1 Sentiment analysis at the word level

Words are the fundamental building blocks of language, and the subjectivity of a word's corresponding sentence or document is directly tied to its polarity. A sentence with an adjective in it has a very high likelihood of being subjective. Additionally, the term a person chooses to express themselves with not only reflects their demographic characteristics, such as gender and age, but also their motivations, personalities, social standing, and other psychological or social features. Word therefore serves as the foundation for text sentiment analysis. The two most popular techniques at the moment are those based on machine learning and natural language processing technology**.**

5.1.2 Sentence-level sentiment analysis:

Sentence-level sentiment analysis is preferred for complicated jobs like handling conditional sentences or sarcastic sentences. 5.5.2.2 Sentence-level sentiment analysis The author of [32] offers a BERT (Bidirectional Encoder Representation from Transformers) + BiGRU (Bidirectional Gated Recurrent Unit) model that first converts words into vectors using the BERT model, obtains contextualized embedding's, and then uses the BiGRU to analyze sentiment. A more complex sort of speech act, sarcasm involves saying the exact opposite of what the speaker means. In order to identify sarcastic statements in product evaluations, Tsur et al. [33] introduced a brand-new semi-supervised technique named SASI. Semi-supervised pattern acquisition and sarcastic categorization are the two stages of SASI. In an effort to address the issue of automatically detecting.

5.1.3 Document-level Sentiment Analysis

Cross-domain and cross-language sentiment analysis are the two biggest obstacles to document-level sentiment analysis.

By including document-level sentiment labels in the context vectors used as the foundation for calculating the distributional similarity between words by Bollegala et al., sentiment sensitivity is achieved in the thesaurus. [31] The author has created a method for performing cross-domain sentiment analysis that makes use of sentiment-sensitive thesaurus (SST). They used labelled data from numerous source domains and unlabeled data from target domains to address the feature mismatch in cross-domain sentiment categorization. Then, during the training and testing phases of a binary classifier, the produced thesaurus is used to extend feature vectors.

5.2 Task-oriented sentiment analysis

Tasks of SA

- Polarity classification,
- Feature/aspect-based sentiment analysis.
- Polarity at specific scale,
- Beyond polarity,
- Subjectivity/Objectivity identification,

5.2.1 Polarity classification:

5.1.1 Polarity classification: A fundamental study in sentiment analysis, polarity classification examines the positive, negative, or neutral nature of the expressed opinions in a document or a sentence about a certain feature or attribute of a target. In 2002, Pang et al [16] investigated the efficacy of using naive Bayes and SVM to the sentiment classification job on movie reviews for detecting the polarity of product evaluations and movie reviews with different approaches. Based on updated SentiWordNet (SWN) sentiment scores, Khan et al[17] recommended vocabulary.

5.2.2 Feature /Aspect-based sentiment analysis

Penalver-Martinez et al. [30] developed a novel approach to improve the outcomes of conventional sentiment analysis methods by utilizing new Semantic Web-guided solutions. By incorporating ontologies into the feature selection process, this technique enhances feature-based opinion mining and introduces a fresh vector analysis-based approach for sentiment analysis. Identifying the thoughts stated on various elements or aspects of an entity is known as feature or aspect-based sentiment analysis [29], which is a more granular analysis activity. Aspect extraction differs from entity extraction in the context of sentiment analysis. Aspect-based analysis basically seeks out explicit aspect expressions, which are typically nouns and noun phrases from the text of the supplied domain.

5.2.3 Polarity at specific scale

There are various scales on which to categorize document polarity. In order to anticipate whether star ratings on a 3 or 4 star scale will be positive or negative, Snyder et al. [18] and Pang et al. [19] enlarged the basic task of two-class classification. To increase prediction accuracy in recommenders, the sentiment-based rating prediction system (RPS) was suggested in [20]. The recommender system has accurately predicted the rating by taking into account these parameters. For incorporating human emotion into intelligent computer systems, Karyotis et al. [21] introduced a novel framework for emotion modelling. The fuzzy technique is assessed in terms of its capacity to represent affective states in comparison to other existing machine learning approaches.

5.2.4 Entities beyond polarity (entity, opinion holder, spatial information, and temporal information analysis)

The goal of fine-grained opinion analysis is to locate subjective expressions in text and to identify the sources and targets of those expressions. Using fine-grained analysis, it is possible to recognise different sorts of opinion entities, including opinion sources [22], opinion holders [23], opinion expressions [24], opinion targets [25], and features of a target [26].

5.2.5 Subjectivity or objectivity identification

Subjectivity or objectivity identification represents a different study avenue. In a case study using resources that were sense-aligned, Banea et al. [27] sought to evaluate the transfer of subjectivity between languages. By employing either cross-lingual or multilingual training reinforced with bootstrapping, the framework in this model is able to predict subjectivity labelling for unseen senses.

By providing a difference-based scoring formula, Karimi et al[28] developed a language model-based structure that can help with subjectivity identification by decreasing the impact of common topic relevant words in the process of differentiating subjective papers from objective ones.

5.3 Methodology Oriented Techniques:

5.3.1 Supervised Learning Methods:

As the name suggests, supervised learning involves a supervisor serving as an instructor. In essence, supervised learning refers to the process of teaching or training the computer utilizing labelled data. Which indicates that the right answer has already been assigned to certain data. In order for the supervised learning algorithm to analyses the training data (set of training examples) and create a proper result from labelled data, the machine is then given a fresh set of examples (data). [35].

Fangzhao Wu et al [36] has proposed a new approach to extract heterogeneous sentiment knowledge from massive unlabeled micro blog messages and incorporate them into a unified framework to train sentiment classifiers for micro blog sentiment classification. The techniques used were SVM, Naive Bayes and Logistic Regression. Using a combination of semi-supervised and supervised learning techniques, Siaw Ling LO et al. [37] have proposed a methodology to identify and rank the High Value Social Audience of a Twitter account owner with minimal annotation work. A model to perform the analysis of the whole tweet texts was developed by Ward van Zoonen et al. and published in [38]. In this paper, the coding performance of three classifiers—Linear Support Vector Machine, Naive Bayes, and Logistic Regression—was compared using data from tweets linked to work from Dutch employees. When the randomly chosen training set has at least 4000 tweets, the linear support vector machine performs satisfactorily. Using the "SENNA" deep learning framework, Wang et al. [39] implemented deep learning for "entity recognition." SVM, KNN, Logistic Regression, and Naive Bayes were utilised as classifiers. When compared to other methods, SVM produced the most promising outcomes.

SVM outperformed other classifiers with accuracy and recall of 89.8% and 89.0%, respectively.

Author Celli et al in [40] has analyzed using Random Forest, the style of personality and communication in the diffusion of news articles. As classification algorithm Logistic Regression is used, with 66% training and 33% test split. The results show that it is possible to predict correctly about 60% of positive and negative mood sharers in Twitter using personality types and communication styles. In particular, positive mood sharers can be detected with more recall and negative mood sharers with more precision. A model by Igawa et al[41] to analyze the behavior of the content produced by bots for fraud evaluation, improving the detection of spamming activities in Online Social Network using tweets of 2014 FIFA World Cup. It obtained accuracy of 88.7% with Random Forest and Neural Network. In [42] Perikos and Hatzilygeroudis (2016) has extensively evaluated text like news, articles. An ensemble classifier system which is based on three main classifiers, a naïve Bayes learner, a maximum entropy and a knowledge based tool is used and concluded that sentimental analysis of tweets is better to be conducted by statistical approaches. Bouazizi and Ohtsuki[43]obtained best results to tweets collected from 2014 to 2015 for detecting sarcastic comments using pattern based features with an approach using Random Forest, Support Vector Machine, k Nearest Neighbours (k-NN) and Maximum Entropy. The overall accuracy obtained reaches 83.1% using the classifier Random Forest for an F1-score equal to 81.3%.

In[44] Nair et al developed a model which processed Cleveland data from heart disease dataset and all parameters yielded higher results and implementation was carried with decision tree and Spark's MLlib, the machine learning library. The user tweets the health data which is filtered by the application in near real time and apply the machine learning model on the extracted health data to predict the health status.

Cui et al in [45] presented a model which improved results using distant based supervised algorithm along with SVM and LibSVMtool. Author Perez-Gallego et al in [46] showed improved results for tweets with emoticons using ensemble based algorithms .Naive Bayes where the implementation was carried on with techniques like Naive Bayes, Logistic Regression and SVM. We have investigated the behaviour of ensembles in a scenario where it is assumed that the data distribution would change between the training and testing stages.

Author Alsinet et al[47]formulated a model for automatically labeling the relationship between the sentiments which he implemented with SVM and obtained 60% accuracy. By using the techniques like Naive Bayes, SVM, Logistic Regression, and Random Forest author Jianqiang and Xiaolin[48] with dataset from Stanford Twitter Sentiment SemEval 2014,STS Gold, SS-Twitter, SE-Twitter has proposed a model which obtained F-Score of 0 .37 for SemEval 2014.

In [49] Jain and Kumar formulated a model for health domain with tweets collected from Sep 2016 to Nov 2016 and implemented the model with classifiers like SVM, Naive Bayes and Logistic Regression where SVM out performed Naive Bayes. And parameters F-Score, Precision, Recall and Accuracy were used as metrics for evaluation.

In [50] researcher Keshavarz and Abadeh used datasets like Sanders, Presidential debate corpus, Healthcare Reform (HCR), SemEval 2013 and Stanford and demonstrated a model which was implemented with genetic algorithm . And improved results were obtained.

Geo-tagged tweets were gathered by author Anna et al[51] from Hurricane Sandy Collection and analyzed sentiments of the tweets belonging to the environmental crisis. The tweets with geo-location using SentiStrength were collected and SVM and Naive Bayes were applied. SVM produced accuracy of around 76%. Author Singh et al [52] applied SVM and Naive Bayes to analyze sentiment polarity of the tweets belonging to food(health)domain. SVM performed better than Naive Bayes.

Another work was carried by Xiaomei et al[53] with health, tea party and politics dataset. Implementation was carried with SVM and Naive Bayes where SVM yielded improved accuracy. In [54] Khan et al applied Naive Bayes for a model by collecting political and non-political tweets as dataset. The algorithm achieved accuracy of 85%.

In 2017 author Bouazizi and Ohtsuki[55] proposed a model with Random Forest and gathered tweets from SENTA dataset . It gave a accuracy of 60.2% for multiclass SA. In 2017 Li et al[56]collected 196,370 tweets and classified them into multiclass related to stock market .And implemented this model with Naive Bayes and decision tree . Naive Bayes yielded accuracy of 72%.

Ghiassi and Lee[57] used Neural Networks and SVM to mine sentiment from 40,000 tweets about Starbucks, Governor Christie, Southwest Airlines, and Verizon. and obtained improved outcomes. A model with SVM for sentiment was proposed by Zhao et al. in 2019[58] and it explored the link between image and text. For all the metrics, this model performed well.

In [59], Park *et al* have developed a model using SVM with Amazon reviews and Yelp reviews .The proposed model uses semi-supervised distributed representation. And has obtained high accuracy. In 2019, Vashishtha and Susan[60] proposed a model using fuzzy rule and has applied twitter data. This model showed good performance for all the four metrics. The model was applied using nine publicly available twitter datasets, four pre-existing models, and sentiment lexicons.

In 2019, Yousif *et al*.[61] have implemented a multi-task learning method on the basis of CNN and RNN. The citation sentiment and citation purpose dataset was used to analyze the technique. The performance was measured based on F-Score, Precision and Recall values. In 2020, author Hassonah et al.[62] applied SVM to a model using Twitter social data and suggested a hybrid machine learning approach for enhancing the SA, while the MVO and Relief models combined their feature selection techniques.Author Xu*et al*.[63] have proposed a model with Naive bayes classifier. The model was implemented with Amazon product and Movie review data and tested with many parameters and high *accuracy* was obtained through continuous testing. For an e-commerce platform, it was suggested. For aspect-based SA for Arabic hotels, author Smadiet al. built a model using deep recurrent neural networks instead of SVM in [64].

The application made use of the review dataset for Arabic hotels. Neural networks were outperformed by SVM. In 2020, the author [65] has suggested a model to study the impact of 2012-2016 stock market events. Here, Twitter dataset were used to evaluate the events. A deep-learning method for categorising the opinion of the user mentioned in reviews was proposed by Abdi et al. in 2019 [66]. Additionally, the suggested model used RNN that included LSTM to address the drawbacks of traditional algorithms and take into account the advantages of sequential processing. All the four parameters were used to access the performance and movie review was the dataset fed for implementation.

A deep learning strategy for performance improvement has been proposed by Park et al.(2020) [67] . The model utilized laptop and restaurant reviews from SemEval 2014 with SVM classifier and obtained accuracy and high F-Score. In 2019, the author Bardhan*et al*.[68] have defined a quasi-qualitative model for understanding the effects of gender mainstreaming in SRH management. Verbal narratives from semi-structured interviews and group discussions were the data's used for implementation. The emotions of the stakeholders were analyzed using SA with machine learning algorithm. Text data from corpus was the dataset. Authors in [69] have proposed a fusion based hybrid model for corona tweets with five classifiers and the accuracy is between 79 to 85 percent for the models used. Eight countries dataset are compared for evaluation. In [70] the author uses classifiers based on the sentiment contradiction formula (Sent-C) and the entropy contradiction formula (Ent-C) with statistical aggregates as input data .This dataset contained approximately 7 million tweets, which we assigned with sentiment labels by SentiStrength, as more appropriate for short messages, and applied CTree algorithm on them to detect contradictions.

Authors Duyu Tang etal [71] proposed a model where sentiment embedding's are helpful in capturing discriminative features to predict the positive/negative sentiment of text.In [72] reputation polarity of a tweet using feature selection algorithm is performed and approached in an entity-dependent way. To classify reputation polarity, by incorporating strong, word list-based, sentiment classifiers for social media with social media features such as authority and recursive use of discourse structure in Twitter is used.

Authors in [73] analyzed how figurative language is used on Twitter. Explicit tags in messages by users as #irony, #sarcasm and #not were analyzed in order. The author has taken into account emotional and affective lexical resources, in addition to structural features, with the aim of exploring the relationship between figuratively, sentiment and emotions at a finer level of granularity.

5.3.2 Unsupervised Learning Methods :

In [74] He, Lee, Ng, and Dahlmeier (2017) proposed an attention-based model for unsupervised aspect extraction. The fundamental idea is to use the attention mechanism to emphasize aspect-related words more during the learning of aspect embedding's while underplaying aspect-irrelevant terms. Brody and Elhadad [75] suggested that aspects be first identified using topic models, and then aspect-specific opinion words be identified by taking adjectives into consideration. [76] Classification is done using well-known opinion words. [Figure 3].

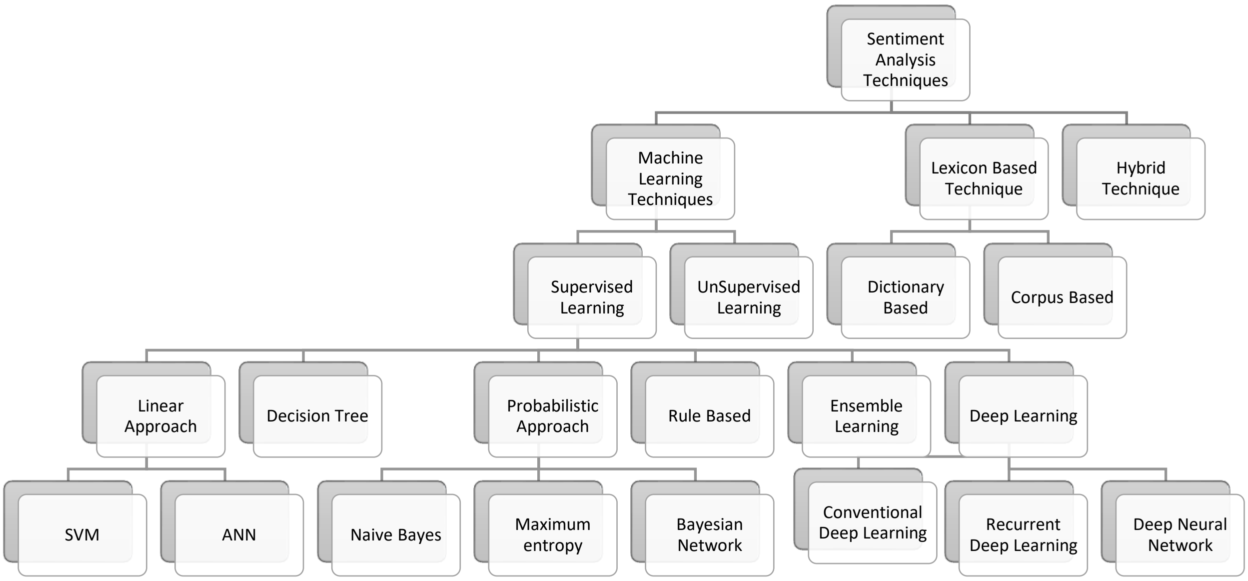

Fig 3: Sentiment Analysis Techniques

Table 2: SENTIMENT CLASSIFICATION ACCURACY OBTAINED

SNO	AUTHOR	REFERENCES	TECHNIQUES/ALGORITHMS	RESULTS BASED ON PERFORMANCE METRICS
1	V. Patti , E. Sulis, D. Irazú Hernández Farías, P. Rosso, , and G. Ruffo	[73]	Decision Tree, Random Forest, SVM, Naive Bayes,Logistic Regression	F-Score of RF for irony versus sarcasm is 69.8%, for irony versus not, 75.2%, and for sarcasm versus not, 68.4%
2	F. Wu, Y. Song, and Y. Huang	[36]	SVM, Naive Bayes, Logistic Regression	F Score > 60%
3	S. Ling, R. Chiong, and D. Cornforth	[37]	SVM and Bootstrap(HYBRID 1), SVM and Bagging(HYBRID 2), Fuzzy Logic, Linear Regression,	Hybrid Obtained F-Score = 98%, 97%, 98% For Three Datasets
4	W. Van Zoonen and T. G. Van Der Meer	[38]	SVM, Naive Bayes, Logistic Regression	Accuracy = 81%
5	Z. Wang, S. Zhang X. Cui, L. Gao, Q. Yin,and L. Ke,	[39]	SVM,KNN, Logistic Regression , Naive Bayes	Accuracy = 89.8%
6	F. Celli, A. Ghosh, F. Alam, and G. Riccardi	[40]	Logistic Regression, Random Forest	F-Score=61.7%
7	R. A. Igawa *et al.*	[41]	Random Forest And Neural Network	Accuracy = 88.7% .

Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 44 No. 5 (2023)

8	I. Perikos and I. Hatzilygeroudis	[42]	Naive Bayes	Accuracy = 85%
10	M. Bouazizi and T. Otsuki	[43]	SVM, KNN, Random Forest	Precision = 98%
11	L. R. Nair, S. D. Shetty, and S. D. Shetty	[44]	Decision Tree	More Accuracy In Less Time
12	L. Cui, X. Zhang, A. K. Qin, T. Sellis, and L. Wu	[45]	SVM	Improved Results
13	P. Pérez-Gállego, J. R. Quevedo, and J. J. del Coz	[46]	Naïve Bayes, Logistic Regression, SVM	Improved Results With Naive Bayes
14	T. Alsinet, J. Argelich, R. Béjar, C. Fernández, C. Mateu, and J. Planes	[47]	SVM	Accuracy 60%
15	Z. Jianqiang and G. Xiaolin	[48]	SVM, Naive Bayes, Logistic Regression, Random Forest	F- Score = 0.37
16	V. K. Jain and S. Kumar	[49]	SVM,Naive Bayes,Logistic Regression	SVM Performed Better Than Naive Bayes
17	H. Keshavarz and M. S. Abadeh	[50]	Genetic Algorithm	Accuracy= 85%
18	A. Squicciarini, A. Tapia, and S. Stehle	[51]	Naive Bayes,SVM	Accuracy =76%
19	A. Singh, N. Shukla, and N. Mishra	[52]	SVM,Naive Bayes	SVM Performance Better Than Naive Bayes
20	Z. Xiaomei, Y. Jing, Z. Jianpei, and H. Hongyu	[53]	SVM,Naive Bayes	SVM Performed Better Than Naive Bayes
21	I. Khan, S. K. Naqvi, M. Alam, and S. N. A. Rizvi	[54]	Naive Bayes	Accuracy= 85%
22	M. Bouazizi and T. Ohtsuki	[55]	Random Forest	Accuracy= 60.2%
23	B. Li, K. C. C. Chan, C. Ou, and S. Ruifeng	[56]	Naive Bayes, Decision Tree	Accuracy= 72%
24	M. Ghiassi and S. Lee	[57]	SVM,Neural Networks	Enhanced Results
25	Z. Zhao *et al.*	[58]	SVM	Enhanced Precision, F-Score, Recall And Accuracy
26	S. Park, J. Lee, and K. Kim	[59]	SVM	Enhanced Accuracy
27	S. Vashishtha and S. Susan	[60]	Fuzzy Rule	More Accuracy
28	A. Yousif, Z. Niu, J. Chambua, and Z. Y. Khan	[61]	Conventional Neural Networks	Enhanced F-Score, ,Precision And Recall
29	M. A. Hassonah, R. Al-Sayyed, A. Rodan, A. M. Al-Zoubi, I. Aljarah, and H. Faris	[62]	SVM	F-Score, Precision, Recall And Accuracy
30	F. Xu, Z. Pan, and R. Xia	[63]	Naive Bayes	Enhanced Accuracy
31	M. Al-Smadi, O. Qawasmeh, M. Al-Ayyoub, Y. Jararweh, and	[64]	SVM	Faster Execution

	B. Gupta			
32	H. Maqsood *et al*	[65]	Support Vector and Regression	MAE AND RMSE
33	A. Abdi, S. M. Shamsuddin, S. Hasan, and J. Piran	[66]	RNN Composed By LSTM	F-Score, Precision, Recall And Accuracy
34	H. jung Park, M. Song, and K. S. Shin	[67]	SVM	Enhanced Accuracy And F-Score
35	R. Bardhan, M. Sunikka-Blank, and A. N. Haque	[68]	Quasi Qualitative Model	Accuracy Obatined
36	M. E. Basiri, S. Nemati, M. Abdar, S. Asadi, and U. R. Acharrya	[69]	CNN, Bigru, NB, SVM, AND Fast Text	Accuracy = 79% To 85 %
38	M. Tsytsarau and T. Palpanas	[70]	Sentiment Contradiction Formula (Sent-C)	Accuracy= 82 %,
39	D. Tang, F. Wei, B. Qin, N. Yang, T. Liu, and M. Zhou	[71]	Knn, SVM, Nn	F-Score Of SVM Is 72.1%
40	M. H. Peetz, M. De Rijke, and R. Kaptein	[72]	Decision Tree	F Score For 2013 And 2012 is 0.55% , 0.49% respectively

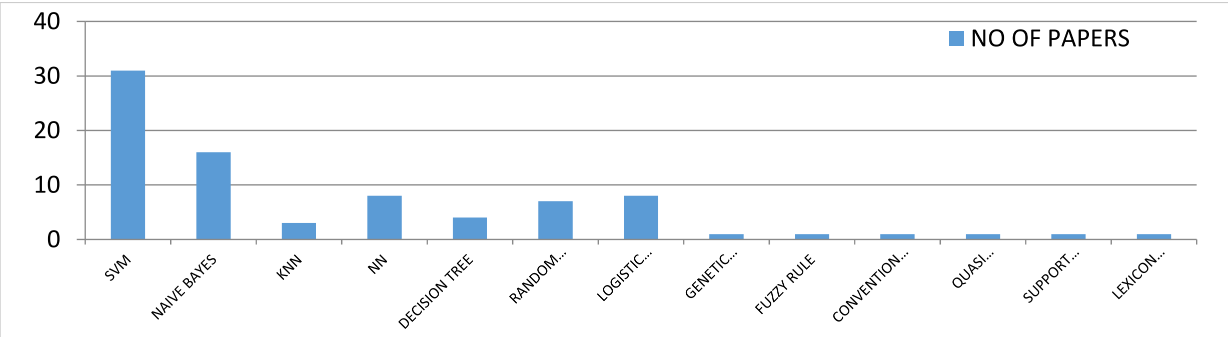

Figure 4

The analysis of the above SA papers [Table 2] have shown that SVM is the most frequently used technique as it always yields better performance [Figure 4]. Naive Bayes is the next technique often used for implementation by authors.

6. Conclusion:

From the review conducted on Sentiment Analysis (SA) the following can be researched further.

- The SA for irony analysis, sarcasm detection, rumor detection, have been studied less.
- Most of the SA work has focused on Twitter data for business decision making.
- Some of the Soft Computing techniques can enhance intelligent analytics . And unstructured data which is available in social media can be experimented with these Soft Computing techniques. Few of the approaches are not analyzed properly. Methods like Fuzzy logic, Genetic Algorithm, Neural Networks and Evolutionary Computing are applied less.
- Most of the study has focused only on the accuracy, other metrics are either not measured or result of those metrics given less importance.

- Since SA is applied mostly on social media and the content in social media are emojis, smileys and unstructured textual data .But so far the focus has been mostly on textual data. A new approach with multi dimensional view to all kinds of data including multimedia like videos has to be developed which can check polarity for all informal data.

References

[1] Hendricks Drew, "The Complete History of Social Media: Then And Now - Small Business Trends," *Small Business TRENDS*, 2021. [Online]. Available: https://smallbiztrends.com/2013/05/the-complete-history-of-social-media-infographic.html. [Accessed: 29-May-2022].

[2] H. Tankovska, "Number of social media users 2025 | Statista," *Jan 28,2021*, 2021. [Online]. Available: https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/. [Accessed: 29-May-2022].

[3] O. Zaïane, "Chapter I: Introduction to Data Mining," in *Principles of Knowledge Discovery in Databases*, 1999, pp. 1–15.

[4] L. Zhang, B. Liu, and S. Francisco, "Encyclopedia of Machine Learning and Data Science," *Encycl. Mach. Learn. Data Sci.*, 2020, doi: 10.1007/978-1-4899-7502-7.

[5] P. Bo and L. Lee, *Opinion mining dan sentiment analysis*, vol. 2, no. 1–2. 2008.

[6] W. Medhat, A. Hassan, and H. Korashy, "Sentiment analysis algorithms and applications: A survey," *Ain Shams Eng. J.*, vol. 5, no. 4, pp. 1093–1113, Dec. 2014, doi: 10.1016/j.asej.2014.04.011.

[7] M. Wiegand, A. Balahur, B. Roth, D. Klakow, and A. Montoyo, "A Survey on the Role of Negation in Sentiment Analysis," *Proc. Work. Negation Specul. Nat. Lang. Process.*, no. July, pp. 60–68, 2010.

[8] K. Ravi and V. Ravi, "A survey on opinion mining and sentiment analysis: Tasks, approaches and applications," *Knowledge-Based Syst.*, vol. 89, pp. 14–46, 2015, doi: 10.1016/j.knosys.2015.06.015.

[9] H. Saif, M. Fernandez, Y. He, and H. Alani, "Evaluation datasets for Twitter sentiment analysis a survey and a new dataset, the STS-Gold," *CEUR Workshop Proc.*, vol. 1096, no. December, pp. 9–21, 2013.

[10] R. M. Chandrasekaran and G. Vinodhini, "Sentiment Analysis and Opinion Mining: A Survey," *Int. J. Adv. Res. Comput. Sci. Softw. Eng.*, vol. 2, no. 6, p. 2277, 2012.

[11] M. Tsytsarau and T. Palpanas, "Survey on mining subjective data on the web," *Data Min. Knowl. Discov.*, vol. 24, no. 3, pp. 478–514, 2012, doi: 10.1007/S10618-011-0238-6.

[12] K. Schouten and F. Frasincar, "Survey on Aspect-Level Sentiment Analysis," *IEEE Trans. Knowl. Data Eng.*, vol. 28, no. 3, pp. 813–830, 2016, doi: 10.1109/TKDE.2015.2485209.

[13] H. Tang, S. Tan, and X. Cheng, "A survey on sentiment detection of reviews," *Expert Syst. Appl.*, vol. 36, no. 7, pp. 10760–10773, Sep. 2009, doi: 10.1016/J.ESWA.2009.02.063.

[14] A. Giachanou and F. Crestani, "Like it or not: A survey of Twitter sentiment analysis methods," *ACM Comput. Surv.*, vol. 49, no. 2, 2016, doi: 10.1145/2938640.

[15] Y. Lin, X. Wang, and A. Zhou, "Opinion spam detection," *Opin. Anal. Online Rev.*, no. May, pp. 79–94, 2016, doi: 10.1142/9789813100459_0007.

[16] V. S. (Pang B, Lee L, "Thumbs up?: sentiment classification using machine learning techniques.," *Proc. ACL-02 Conf. Empir. methods Nat. Lang. Process.*, vol. 10, no. 9, pp. 79–86, 2002, doi: 10.1016/0096-6347(45)90048-2.

[17] F. H. Khan, U. Qamar, and S. Bashir, "eSAP: A decision support framework for enhanced sentiment analysis and polarity classification," *Inf. Sci. (Ny).*, vol. 367–368, pp. 862–873, 2016, doi: 10.1016/j.ins.2016.07.028.

[18] B. Snyder and R. Barzilay, "Multiple aspect ranking using the good grief algorithm," *NAACL HLT 2007 - Hum. Lang. Technol. 2007 Conf. North Am. Chapter Assoc. Comput. Linguist. Proc. Main Conf.*, pp. 300–307, 2007.

[19] B. Pang and L. Lee, "Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales," *ACL-05 - 43rd Annu. Meet. Assoc. Comput. Linguist. Proc. Conf.*, no. June, pp. 115–124, 2005.

[20] X. Lei, X. Qian, and G. Zhao, “Rating Prediction Based on Social Sentiment from Textual Reviews,” *IEEE Trans. Multimed.*, vol. 18, no. 9, pp. 1910–1921, 2016, doi: 10.1109/TMM.2016.2575738.

[21] C. Karyotis, F. Doctor, R. Iqbal, A. James, and V. Chang, “A fuzzy computational model of emotion for cloud based sentiment analysis,” *Inf. Sci. (Ny).*, vol. 433–434, pp. 448–463, 2018, doi: 10.1016/j.ins.2017.02.004.

[22] D. Lingjia and J. Wiebe, “Recognizing opinion sources based on a new categorization of opinion types,” *IJCAI Int. Jt. Conf. Artif. Intell.*, vol. 2016-Janua, pp. 2775–2781, 2016.

[23] A. Katiyar and C. Cardie, “Investigating LSTMs for joint extraction of opinion entities and relations,” *54th Annu. Meet. Assoc. Comput. Linguist. ACL 2016 - Long Pap.*, vol. 2, pp. 919–929, 2016, doi: 10.18653/v1/p16-1087.

[24] L. X. Liao J, Wang S, Li D, “FREERL: fusion relation embedded representation learning framework for aspect extraction. Knowl Based Syst 135:9–17,” pp. 9–17, 2017.

[25] M. Wiegand, C. Bocionek, and J. Ruppenhofer, “Opinion holder and target extraction on opinion compounds - A linguistic approach,” *2016 Conf. North Am. Chapter Assoc. Comput. Linguist. Hum. Lang. Technol. NAACL HLT 2016 - Proc. Conf.*, pp. 800–810, 2016, doi: 10.18653/v1/n16-1094.

[26] S. Poria, E. Cambria, and A. Gelbukh, “Aspect extraction for opinion mining with a deep convolutional neural network,” *Knowledge-Based Syst.*, vol. 108, pp. 42–49, 2016, doi: 10.1016/j.knosys.2016.06.009.

[27] C. Banea, R. Mihalcea, and J. Wiebe, “Sense-level subjectivity in a multilingual setting,” *Comput. Speech Lang.*, vol. 28, no. 1, pp. 7–19, 2014, doi: 10.1016/j.csl.2013.03.002.

[28] S. Karimi and A. Shakery, “A language-model-based approach for subjectivity detection,” *J. Inf. Sci.*, vol. 43, no. 3, pp. 356–377, 2017, doi: 10.1177/0165551516641818.

[29] M. Hu and B. Liu, “Mining and summarizing customer reviews,” *KDD-2004 - Proc. Tenth ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, pp. 168–177, 2004, doi: 10.1145/1014052.1014073.

[30] I. Peñalver-Martinez *et al.*, “Feature-based opinion mining through ontologies,” *Expert Syst. Appl.*, vol. 41, no. 13, pp. 5995–6008, 2014, doi: 10.1016/j.eswa.2014.03.022.

[31] D. Bollegala, D. Weir, and J. Carroll, “Cross-domain sentiment classification using a sentiment sensitive thesaurus,” *IEEE Trans. Knowl. Data Eng.*, vol. 25, no. 8, pp. 1719–1731, 2013, doi: 10.1109/TKDE.2012.103.

[32] J. Shen, X. Liao, and Z. Tao, “Sentence-level sentiment analysis via BERT and BiGRU,” in *Proc.SPIE*, 2019, vol. 11321, doi: 10.1117/12.2550215.

[33] O. Tsur, D. Davidov, and A. Rappoport, “ICWSM - A great catchy name: Semi-supervised recognition of sarcastic sentences in online product reviews,” *ICWSM 2010 - Proc. 4th Int. AAAI Conf. Weblogs Soc. Media*, no. 9, pp. 162–169, 2010.

[34] R. González-Ibáñez, S. Muresan, and N. Wacholder, “Identifying sarcasm in Twitter: A closer look,” *ACL-HLT 2011 - Proc. 49th Annu. Meet. Assoc. Comput. Linguist. Hum. Lang. Technol.*, vol. 2, no. 2010, pp. 581–586, 2011.

[35] B. Rajoub, “Supervised and unsupervised learning,” in *Biomedical Signal Processing and Artificial Intelligence in Healthcare*, 2020, pp. 51–89.

[36] F. Wu, Y. Song, and Y. Huang, “Microblog sentiment classification with heterogeneous sentiment knowledge,” *Inf. Sci. (Ny).*, vol. 373, pp. 149–164, Dec. 2016, doi: 10.1016/j.ins.2016.09.002.

[37] S. Ling, R. Chiong, and D. Cornforth, “Ranking of high-value social audiences on Twitter Ranking of high-value social audiences on Twitter Citation Citation CORE View metadata, citation and similar papers at core.ac.uk provided by Institutional Knowledge at Singapore Management University Rank,” *Decis. Support Syst.*, vol. 85, no. C, pp. 34–48, Feb. 2016.

[38] W. Van Zoonen and T. G. L. A. Van Der Meer, “Social media research: The application of supervised machine learning in organizational communication research.,” *Comput. Human Behav.*, vol. 63, pp. 132–141, Oct. 2016, doi: 10.1016/j.chb.2016.05.028.

[39] Z. Wang, X. Cui, L. Gao, Q. Yin, L. Ke, and S. Zhang, “A hybrid model of sentimental entity recognition on mobile social media,” *Eurasip J. Wirel. Commun. Netw.*, vol. 2016, no. 1, Dec. 2016, doi: 10.1186/s13638-016-0745-7.

[40] F. Celli, A. Ghosh, F. Alam, and G. Riccardi, “In the mood for sharing contents: Emotions, personality and interaction styles in the diffusion of news,” *Inf. Process. Manag.*, vol. 52, no. 1, pp. 93–98, Jan. 2016, doi: 10.1016/j.ipm.2015.08.002.

[41] R. A. Igawa *et al.*, “Account classification in online social networks with LBCA and wavelets,” *Inf. Sci. (Ny).*, vol. 332, pp. 72–83, Mar. 2016, doi: 10.1016/j.ins.2015.10.039.

[42] I. Perikos and I. Hatzilygeroudis, “Recognizing emotions in text using ensemble of classifiers,” *Eng. Appl. Artif. Intell.*, vol. 51, pp. 191–201, May 2016, doi: 10.1016/j.engappai.2016.01.012.

[43] M. Bouazizi and T. Otsuki, “A Pattern-Based Approach for Sarcasm Detection on Twitter,” *IEEE Access*, vol. 4, pp. 5477–5488, 2016, doi: 10.1109/ACCESS.2016.2594194.

[44] L. R. Nair, S. D. Shetty, and S. D. Shetty, “Applying spark based machine learning model on streaming big data for health status prediction,” *Comput. Electr. Eng.*, vol. 65, pp. 393–399, Jan. 2018, doi: 10.1016/j.compeleceng.2017.03.009.

[45] L. Cui, X. Zhang, A. K. Qin, T. Sellis, and L. Wu, “CDS: Collaborative distant supervision for Twitter account classification,” *Expert Syst. Appl.*, vol. 83, pp. 94–103, Oct. 2017, doi: 10.1016/j.eswa.2017.03.075.

[46] P. Pérez-Gállego, J. R. Quevedo, and J. J. del Coz, “Using ensembles for problems with characterizable changes in data distribution: A case study on quantification,” *Inf. Fusion*, vol. 34, pp. 87–100, Mar. 2017, doi: 10.1016/j.inffus.2016.07.001.

[47] T. Alsinet, J. Argelich, R. Béjar, C. Fernández, C. Mateu, and J. Planes, “Weighted argumentation for analysis of discussions in Twitter,” *Int. J. Approx. Reason.*, vol. 85, pp. 21–35, Jun. 2017, doi: 10.1016/j.ijar.2017.02.004.

[48] Z. Jianqiang and G. Xiaolin, “Comparison research on text pre-processing methods on twitter sentiment analysis,” *IEEE Access*, vol. 5, pp. 2870–2879, 2017, doi: 10.1109/ACCESS.2017.2672677.

[49] V. K. Jain and S. Kumar, “Effective surveillance and predictive mapping of mosquito-borne diseases using social media,” *J. Comput. Sci.*, vol. 25, pp. 406–415, Mar. 2018, doi: 10.1016/j.jocs.2017.07.003.

[50] H. Keshavarz and M. S. Abadeh, “ALGA: Adaptive lexicon learning using genetic algorithm for sentiment analysis of microblogs,” *Knowledge-Based Syst.*, vol. 122, pp. 1–16, Apr. 2017, doi: 10.1016/j.knosys.2017.01.028.

[51] A. Squicciarini, A. Tapia, and S. Stehle, “Sentiment analysis during Hurricane Sandy in emergency response,” *Int. J. Disaster Risk Reduct.*, vol. 21, pp. 213–222, Mar. 2017, doi: 10.1016/j.ijdrr.2016.12.011.

[52] A. Singh, N. Shukla, and N. Mishra, “Social media data analytics to improve supply chain management in food industries,” *Transp. Res. Part E Logist. Transp. Rev.*, vol. 114, pp. 398–415, Jun. 2018, doi: 10.1016/j.tre.2017.05.008.

[53] Z. Xiaomei, Y. Jing, Z. Jianpei, and H. Hongyu, “Microblog sentiment analysis with weak dependency connections,” *Knowledge-Based Syst.*, vol. 142, pp. 170–180, Feb. 2018, doi: 10.1016/j.knosys.2017.11.035.

[54] I. Khan, S. K. Naqvi, M. Alam, and S. N. A. Rizvi, “An efficient framework for real-time tweet classification,” *Int. J. Inf. Technol.*, vol. 9, no. 2, pp. 215–221, Jun. 2017, doi: 10.1007/s41870-017-0015-x.

[55] M. Bouazizi and T. Ohtsuki, “A Pattern-Based Approach for Multi-Class Sentiment Analysis in Twitter,” *IEEE Access*, vol. 5, pp. 20617–20639, Aug. 2017, doi: 10.1109/ACCESS.2017.2740982.

[56] B. Li, K. C. C. Chan, C. Ou, and S. Ruifeng, “Discovering public sentiment in social media for predicting stock movement of publicly listed companies,” *Inf. Syst.*, vol. 69, pp. 81–92, Sep. 2017, doi: 10.1016/j.is.2016.10.001.

[57] M. Ghiassi and S. Lee, “A domain transferable lexicon set for Twitter sentiment analysis using a supervised machine learning approach,” *Expert Syst. Appl.*, vol. 106, pp. 197–216, Sep. 2018, doi: 10.1016/j.eswa.2018.04.006.

[58] Z. Zhao *et al.*, “An image-text consistency driven multimodal sentiment analysis approach for social media,” *Inf. Process. Manag.*, vol. 56, no. 6, Nov. 2019, doi: 10.1016/j.ipm.2019.102097.

[59] S. Park, J. Lee, and K. Kim, “Semi-supervised distributed representations of documents for sentiment

analysis," *Neural Networks*, vol. 119, pp. 139–150, Nov. 2019, doi: 10.1016/j.neunet.2019.08.001.

[60] S. Vashishtha and S. Susan, "Fuzzy rule based unsupervised sentiment analysis from social media posts," *Expert Syst. Appl.*, vol. 138, Dec. 2019, doi: 10.1016/J.ESWA.2019.112834.

[61] A. Yousif, Z. Niu, J. Chambua, and Z. Y. Khan, "Multi-task learning model based on recurrent convolutional neural networks for citation sentiment and purpose classification," *Neurocomputing*, vol. 335, pp. 195–205, 2019, doi: 10.1016/j.neucom.2019.01.021.

[62] M. A. Hassonah, R. Al-Sayyed, A. Rodan, A. M. Al-Zoubi, I. Aljarah, and H. Faris, "An efficient hybrid filter and evolutionary wrapper approach for sentiment analysis of various topics on Twitter," *Knowledge-Based Syst.*, vol. 192, p. 105353, 2020, doi: 10.1016/j.knosys.2019.105353.

[63] F. Xu, Z. Pan, and R. Xia, "E-commerce product review sentiment classification based on a naïve Bayes continuous learning framework," *Inf. Process. Manag.*, no. February, p. 102221, 2020, doi: 10.1016/j.ipm.2020.102221.

[64] M. Al-Smadi, O. Qawasmeh, M. Al-Ayyoub, Y. Jararweh, and B. Gupta, "Deep Recurrent neural network vs. support vector machine for aspect-based sentiment analysis of Arabic hotels' reviews," *J. Comput. Sci.*, vol. 27, pp. 386–393, Jul. 2018, doi: 10.1016/j.jocs.2017.11.006.

[65] H. Maqsood *et al.*, "A local and global event sentiment based efficient stock exchange forecasting using deep learning," *Int. J. Inf. Manage.*, vol. 50, pp. 432–451, Feb. 2020, doi: 10.1016/j.ijinfomgt.2019.07.011.

[66] A. Abdi, S. M. Shamsuddin, S. Hasan, and J. Piran, "Deep learning-based sentiment classification of evaluative text based on Multi-feature fusion," *Inf. Process. Manag.*, vol. 56, no. 4, pp. 1245–1259, Jul. 2019, doi: 10.1016/j.ipm.2019.02.018.

[67] H. jung Park, M. Song, and K. S. Shin, "Deep learning models and datasets for aspect term sentiment classification: Implementing holistic recurrent attention on target-dependent memories," *Knowledge-Based Syst.*, vol. 187, Jan. 2020, doi: 10.1016/j.knosys.2019.06.033.

[68] R. Bardhan, M. Sunikka-Blank, and A. N. Haque, "Sentiment analysis as tool for gender mainstreaming in slum rehabilitation housing management in Mumbai, India," *Habitat Int.*, vol. 92, Oct. 2019, doi: 10.1016/j.habitatint.2019.102040.

[69] M. E. Basiri, S. Nemati, M. Abdar, S. Asadi, and U. R. Acharrya, "A novel fusion-based deep learning model for sentiment analysis of COVID-19 tweets," *Knowledge-Based Syst.*, vol. 228, Sep. 2021, doi: 10.1016/j.knosys.2021.107242.

[70] M. Tsytsarau and T. Palpanas, "Managing Diverse Sentiments at Large Scale," *IEEE Trans. Knowl. Data Eng.*, vol. 28, no. 11, pp. 3028–3040, 2016, doi: 10.1109/TKDE.2016.2597848.

[71] D. Tang, F. Wei, B. Qin, N. Yang, T. Liu, and M. Zhou, "Sentiment Embeddings with Applications to Sentiment Analysis," in *IEEE Transactions on Knowledge and Data Engineering*, 2016, vol. 28, no. 2, pp. 496–509, doi: 10.1109/TKDE.2015.2489653.

[72] M. H. Peetz, M. De Rijke, and R. Kaptein, "Estimating Reputation Polarity on Microblog Posts," *Inf. Process. Manag.*, vol. 52, no. 2, pp. 193–216, Mar. 2016, doi: 10.1016/j.ipm.2015.07.003.

[73] E. Sulis, D. Irazú Hernández Farías, P. Rosso, V. Patti, and G. Ruffo, "Figurative messages and affect in Twitter: Differences between #irony, #sarcasm and #not," *Knowledge-Based Syst.*, vol. 108, pp. 132–143, Sep. 2016, doi: 10.1016/j.knosys.2016.05.035.

[74] R. He, W. S. Lee, H. T. Ng, and D. Dahlmeier, "An unsupervised neural attention model for aspect extraction," *ACL 2017 - 55th Annu. Meet. Assoc. Comput. Linguist. Proc. Conf. (Long Pap.*, vol. 1, pp. 388–397, 2017, doi: 10.18653/v1/P17-1036.

[75] S. Brody, "An Unsupervised Aspect-Sentiment Model for Online Reviews_Slide," *Comput. Linguist.*, no. June, pp. 804–812, 2010.

[76] M. Taboada, J. Brooke, M. Tofiloski, K. Voll, and M. Stede, "Lexicon-basedmethods for sentiment analysis," *Comput. Linguist.*, vol. 37, no. 2, pp. 267–307, 2011, doi: 10.1162/COLI_A_00049.