A Review and Comparative Analysis of Sentiment Analysis Techniques

Article *in* Informatica · July 2022

DOI: 10.31449/inf.v46i6.3991

2 authors:

Shaha Alotaibi
Princess Nora bint Abdul Rahman University

Amal Al-Rasheed
Princess Nuorah bint Abdul Rahman University

A Review and Comparative Analysis of Sentiment Analysis Techniques

Shaha Al-Otaibi and Amal Al-Rasheed*
E-mail: stalotaibi@pnu.edu.sa, aaalrasheed@pnu.edu.sa
Department of Information Systems, College of Computer and Information Sciences, Princess Nourah Bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia

Overview paper

Keywords: sentiment analysis, feature extraction, machine learning, social networking

Received: February 8, 2022

Social networking platforms have become a major source of information, which covers a wide range of topics and has gained a large volume of usage by people around the world. Platforms such as Twitter, Facebook, Instagram, and LinkedIn have attracted huge numbers of users who create public profiles and communicate with other users in the network. They exchange videos, posts, and comments. Social networking requires appropriate techniques to analyze the huge amount of complex, and frequently updated data generated. Sentiment Analysis is one such method of handling this vast volume of data and extracting useful knowledge from it. Social networking contents are analyzed using different techniques to gain insight from this data and use it in decision-making processes. The aim of this work is to study the sentiment analysis concept and present state-of-the-art techniques as well as provide a comparative study of these techniques.

Povzetek: Pregledni članek opisuje metode za analizo mnenj v socialnih omrežij tipa Twitter, Faceook, Instagram in LinkedIn.

1 Introduction

Sentiment analysis has emerged as one of the most active research areas at the intersection of computer science and linguistics, with the goal of determining the emotions embedded in a text. Sentiment can be characterized as a positive, negative, or neutral evaluation that is stated through textual information, such as a review posted online for a movie, book, or product. It has become a common capability of most social networking platforms and is used by business, marketers, and political analysts. Sentiment analysis understands the expressed words and evaluates how language is used to convey human emotions [1].

Recently, social networking platforms have gained a high rate of usage by attracting lots of people, which has resulted in the generation of large volumes of data. However, appropriate techniques are required to analyze such large, complex, and frequently changing social media data. Additionally, to handle this vast volume of data and gain insight from it, sentiment analysis appears as a means to understand human attitudes and feelings. It contributes significantly to the decision-making process and related issues. For example, decision-making may include purchasing a product, making an investment, or exploring new ideas [2], and consumers are constantly interested in learning from others' experiences while performing these activities. Nowadays, there are many reviews on social networks that can be used in sentiment analysis to understand the polarity of reviews, so that a user can quickly determine if a review is positive or negative without having to read the entire context, which helps in decision-making [3].

Sentiment analysis is commonly used in many domains such as, social networking platforms, marketing, customer service, clinical medicine, and healthcare. Among the important areas for sentiment analysis in health is the area of mental health with the use of persuasive technology [4]. Chatbots, for example, use sentiment analysis a lot for mental health analysis [5].

Sentiment analysis is still a growing area of research and has a broad range of applications. Many studies have applied and discussed issues that influence sentiment analysis processing and affect its accuracy [6]. This work will review the sentiment analysis literature to investigate all the related aspects as well as present a comparative analysis of the recent studies in this domain. It introduces the social networking context and its framework and illustrates the position of sentiment analysis in this framework by providing the different processes and the related issues based on previous research. Many of the previous works have focused on aspects related to sentiment analysis while ignoring other aspects. The main

* Corresponding author

objective of this work is to present comprehensive research for sentiment analysis, covering the framework, processes, techniques, and datasets. We can express the overall contributions of this work as follows:

1. It proposes a clear and precise description of the social networking and sentiment analysis contexts, as well as describing the framework and the relationship between them.
2. It explains the different processes that are required to accomplish the sentiment analysis task.
3. It presents the different techniques that have been applied in sentiment analysis, then reviews the state-of-the-art research.
4. It summarizes the results of each work and aggregates the analyzed output.
5. It provides a comparative study based on select criteria to provide a guide for researchers to achieve good results.

The rest of this paper is organized as follows: Section 2 presents an overview of the social networking context and its framework, followed by an explanation of sentiment analysis and the feature extraction process in sections 3 and 4, respectively. Section 5 introduces the different techniques that have been applied in sentiment analysis. Next, section 6 presents an overview of the state-of-the-art works that have been applied to sentiment analysis. The comparison between the different techniques of sentiment analysis is presented in section 7. Finally, the main conclusions drawn from this work are presented in section 8.

2 Social networking platforms

This section introduces the social networking analysis framework. Social networking can be defined as a set of platforms through which online users can communicate with each other. To identify a user's interests and preferences, a user profile is created. Many approaches can be used to construct this profile, but the most used method is to allow the user to answer some questions. However, the user may not be willing to give information where the user's interests change frequently [7]. In the last decade, the rapid change in technology has helped users to generate more and more content on social networking platforms. With the help of social networking, users share content related to influences, feelings, opinions, and sentiments, with the help of texts, pictures, video, and so on [8]. Examples of social networking platforms are Instagram, Snapchat, Facebook, and Twitter, where users generate vast volumes of data every day. We need to extract the useful knowledge from this data, such as understanding human feelings and opinions. Hence, social networking platforms are the more sophisticated source to find a large amount of data and can be handled to extract the required knowledge [6].

Social Networking Analysis (SNA) is a new concept that provides techniques for characterizing and studying people's interactions and the connections between them. Using SNA techniques, we can extract information from social networking platforms that might be useful for sentiment analysis, such as the sequence of influence among users or groups of similar users. It involves a three-step process framework: capture, understand, and present [9]. An identify step must be integrated into this framework in order for postings or tweets to be identified prior to the capture step [10], as appears in Figure 1. This identification is achieved by utilizing the keywords specified by users. These keywords are then used in query requests to social networking analysis platform APIs, such as the Twitter API, to collect posts or tweets with the determined keywords. Therefore, the steps presented in [9], [10] and [11] are modified as follows:

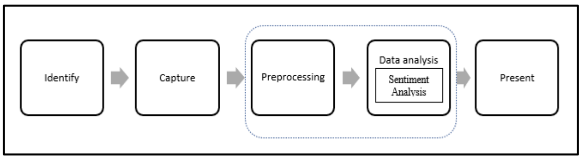

Figure 1: Social networking platforms framework.

- The identify step involves finding the relevant keywords to be used in gathering social networking data.
- The capture step utilizes the keywords to collect the relevant social networking data by checking different social networking sources, archiving related data, and extracting appropriate information.
- The preprocessing step chooses the relevant data and removes noisy and low-quality data. This is followed by the feature extraction process.
- The data analysis step employs different techniques to analyze the data and obtain insights from it.
- The present step displays the findings in a meaningful way.

3 Sentiment analysis

The sentiment analysis process can be shown as a part of the social networking framework displayed in Figure 1. Hence, sentiment analysis is a process of verifying whether a piece of text about a particular entity, such as a product, review, movie, or tweet, is positive, negative, or neutral. It can be used to determine the consumer or social networking user's attitude regarding this particular entity, using a set of variables. Enterprises can use sentiment analysis to discover public opinion of their products or services and to analyze customer satisfaction [3]. Additionally, they can use this analysis to collect crucial feedback about challenges in recently released products.

When discussing sentiment analysis, we should mention opinion mining which has almost the same meaning. However, there is a small difference between them: opinion mining is the process of extracting and analyzing people's opinion on a certain thing, whereas sentiment analysis is the process of looking for sentiment words or expressions and analyzing emotional feelings. The Merriam-Webster dictionary defines sentiment as "an attitude, thought, or judgment encouraged by feeling," while it defines opinion as "a view, judgment established in the mind about a certain subject." The distinction is slight, and they both contain specific characteristics of the other. According to the definitions, an opinion is a person's point of view on a certain topic, whereas a sentiment is about the person's feelings. However, in the majority of cases, opinions reflect positive or negative sentiments, so sentiment analysis and opinion mining are nearly same thing and are used interchangeably in the data analytics literature [12].

Nowadays, most data are available in an unstructured format, hence the vast volume of data created by different channels such as chats, social networking, Reviews, articles, and documents. We need to analyze and understand these volumes of data to extract meaningful knowledge and benefit from it in the decision-making process. Sentiment analysis has emerged as a new technique for understanding people's attitudes in different areas. It supports the processing of huge amounts of data in an efficient and cost-effective way. Moreover, sentiment analysis can identify vital issues in real-time, such as a public opinion crisis on social networking regarding certain products, services, or events. It can support the identification of such situations and discover the human trends and attitudes underlying them [13].

As mentioned above, the sentiment analysis process can be considered as part of the overall process represented in the social networking analysis framework presented in Figure 1. This framework can be extended to explore the sentiment analysis processes presented in Figure 2. Moreover, the sentiment analysis process has been applied at three levels: 1) Document level: at this level, the entire document is classified using sentiment analysis into positive, negative, or neutral results. 2) Sentence level: at this level, each sentence is evaluated to determine whether it reflects a positive, negative, or neutral result. This type of analysis is used for reviews and comments comprising one sentence written by a user in a social network. 3) Entity and Aspect level: sentiment analysis in this level is summarized based on the feature. The classification entails detecting and extracting specific features from the original data. This type is utilized in reviews when we seek sentiments on a specific aspect or feature [13]. For example, assume a buyer purchases a mobile phone and observes that the camera quality is good, but the sound quality is fair. Aspect level analysis can then be used to examine the many aspects of an entity.

4 Feature extraction in sentiment analysis

The feature extraction process is presented in this section. Sentiment analysis includes a sequence of processes such as data preprocessing, feature extraction, and classification, which are applied to find the polarity of data, as shown in Figure 2. Data preprocessing encompasses steps such as removing stop words, stemming, and performing tokenization. Stop words represent common words, such as is, am, are, in, to, which do not reveal any specific opinion. Stemming is a process of returning a word from different forms to its base. Tokenization is a process of dividing a sentence into separate words. Feature extraction is the process of converting the sentence into a bag of words that are represented by the features' vector. Two methods can be applied to find the term occurrence in the corpus: term presence and term frequency. Term presence is the creation of a binary vector representing the extracted features, where 1 indicates that the term is available in the sentence and 0 represents the absence of the term. Term frequency is the creation of a feature vector representing the number of occurrences of the words. Term frequency can be applied using n-gram features, which are widely used in computational linguistics. As a result, the n-gram denotes the number of terms that appear in a text or speech corpus together. When only one term is used as a feature, it is called a unigram; when two terms are used, it is called a bigram, and so on. [13], [14].

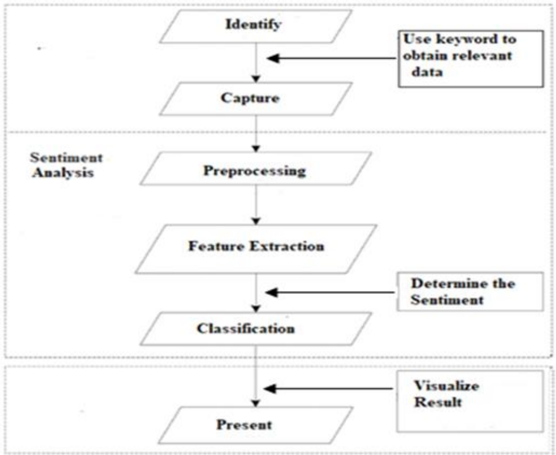

Figure 2. The sentiment analysis process.

5 Review of literature on sentiment analysis techniques

Different types of techniques have been applied in sentiment analysis for social networking. These techniques can be categorized into three main categories: machine learning techniques, lexicon techniques, and hybrid techniques. The following sections present these categories and review the associated literature.

5.1 Machine learning techniques

Machine learning techniques can be further classified into three main classes: supervised, unsupervised, and deep learning techniques. The following subsections present these techniques and the different algorithms that have been applied in sentiment analysis.

5.2 Supervised learning techniques

Supervised learning techniques can be used upon the availability of labeled data for training the model. In this context, the two phases applied include training the model and the prediction of new cases [15]. In the training phase, the labeled data set is passed to the classification algorithm, which produces a model as an output. Then, the test data is passed to the model to predict the class of the new instances. These techniques have been applied to sentiment analysis; for example, several studies used emoticons and hashtags to create the training set, with the emoticons used as class labels to determine the polarity of social networking posts [14]. Pang et al. proposed an algorithm that calculates the probability of every term in the training dataset, which extracts data from movie reviews to identify whether they are positive or negative. The algorithm classifies the new instance of data, then splits the newly classified sentence into separate word features. Next, the constructed model uses the probabilities calculated in the training phase to compute the conditional probabilities of the combined features in order to predict its class [16]. Additionally, the Bayesian Network algorithm was applied to enhance the aspect-based sentiment analysis of reviews from Arabic hotels [17] . The terms' dependency is represented as a directed graph, which is acyclic and contains a set of nodes representing the words as variables, and edges, which represent the dependency between these variables. The hyperplane is used by the Support Vector Machine (SVM) algorithm to offer the most separation between classes with the highest margin of hyperplane, lowering the upper bound of classification errors [16]. It can be used for dimensionality reduction and noise removal with various weighting systems such as term frequency–inverse document frequency (TF-IDF), term occurrence, and Binary Occurrence that uses chi-square as a feature selection [18]. Twitter data has been applied to interpret the insight hidden in public opinion. It was used to classify the sentiments of Twitter posts, in order to determine whether they were positive or negative. The unigram was also applied as a feature extraction method [3]. The Decision Tree algorithm has been applied on many types of datasets, where the training data is divided into smaller parts to identify patterns hidden in the data and used for the classification process [14]. Moreover, the Artificial Neural Network (ANN) is used for text classification using Reuters corpus documents as a training dataset in [19].

5.3 Unsupervised learning techniques

These techniques are used when the available data is unlabeled; such data is easy to collect in some domain areas. The keywords are used to determine the sentences' categories. Clustering algorithms are one of the solutions used to perform sentiment analysis in unlabeled data. These techniques can help to categorize the users' sentiment texts into positive, negative, and neutral. Clustering algorithms were applied to perform clustering and deliver an appropriate number of clusters in an adequate run-time [14], [20]. Additionally, tweets were clustered into positive and negative tweets using the unsupervised approach based on the spectral clustering approach to perform sentiment analysis [21]. The clustering-based approach was developed to perform an aspect-based sentiment analysis that is concerned with gathering and categorizing general opinions about the features of a certain product or service. Many solutions for aspect clustering are monolingual and need to be applied in several languages. This algorithm categorizes the aspects related in semantic manner and performs the sentiment analysis in several languages [22].

5.4 Deep learning techniques

Deep learning techniques refer to the multi-layer methods that adapt hidden layers of the neural network. These techniques differ from other machine learning techniques in how they deal with the feature extraction process. Supervised and unsupervised learning approaches extract features either manually or using a feature extraction method, while deep learning techniques learn and extract features in an automatic way to achieve a high accuracy in the learning process. Deep learning is currently one of the best solutions for many sentiment analysis problems. It can be classified into three main categories: Deep Neural Networks (DNN), Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN). In addition, there are other methods such as the Recursive Neural Network, Belief Neural Network, and hybrid neural networks that are based on integrated techniques [23]. The DNN is a complex mathematical model with a set of layers; some of them are hidden layers that can process data in several ways. The DNN has three layers: an input layer that contains the input data, other hidden layers that contain processing nodes known as neurons, and an output layer that contains one or more neurons whose outputs are the network outputs [23], [24]. An enhancing DNN model has been applied to sentiment analysis for social networking, using a classifier combined with a word embedding model and a linear machine learning algorithm [25]. Additionally, the CNN is a specific type of neural network used in various areas such as Natural Language Processing (NLP), recommender systems, and computer vision. It comprises the convolutional and pooling or subsampling layers to deliver inputs to the classification layer [23]. It has been applied for handling a huge amount of unstructured data, which is a complex process. Feedforward neural networks with many hidden layers have been investigated using CNN [26]. The RNN is a type of neural network that has a directed cycle of connections between neurons. This connection structure produces feedback loops inside the RNN [27]. The RNN algorithm for sentiment analysis was proposed in [28] to recommend the places that are nearby to the user's current position. It analyzed the different reviews extracted from different social networking sites.

5.5 Lexicon techniques

A lexicon can be defined as a set of predefined words, each of which has a polarity score associated with it. It is considered the simplest approach for sentiment analysis. In lexicon approaches, the classifier uses the lexicon directly and, matching each word with its polarity, classifies the entire sentence [20]. Many dictionaries are available online that provide information regarding the pinion words and their polarities. Popular lexicons include Sentiwordnet 3.0 (assigns to each sentiment of WordNet three sentiment scores: positivity, negativity, or objectivity), LIWC (Linguistic Inquiry and Word Count), and General Inquirer. Moreover, many datasets of words are manually created with determined semantic and cognitive categories [29], [14]. These techniques have been applied in sentiment analysis for text messages posted on social networking platforms like Twitter [30]. The sentiment analysis of text is predicted based on the overall polarity of the words that constitute the messages. SentiStrength has been applied to classify the sentiment of textual messages. This algorithm is based on a lexicon that contains words and sentences that are used mostly in social networking [31]. Additionally, the lexicon is built by integrating three online dictionaries, saving only the occurrence words that are repeated. The seed words were selected using TF-IDF methods, then the word set was expanded using synonyms of the seeds to enhance classification performance [32]. The sentiment analysis algorithm was applied in Twitter and studied the patterns repeated in the data contexts. Then, the polarity and scores of words in the lexicons were determined by it [33]. Moreover, semantic and statistical methods have been applied in sentiment analysis. Hence, the semantic orientation determines how positive or negative a word is. Esuli and Sebastiani have proposed a method that starts by expanding the initial seed words using the WordNet dictionary. Their assumption is that words with similar orientations are likely to have similar interpretations [34]. Additionally, a statistical analysis for the collection of documents has used the K-Nearest Neighbors (KNN) algorithm. The model designs a classifier with several features: n-gram feature, pattern feature, punctuation feature, keyword-based feature, and word feature. It used the KNN classifier and calculates the accuracy of all algorithms. It focuses on classifying the tweets into positive and negative sentiments [35].

5.6 Hybrid techniques

In much of the literature, two or more techniques have been combined to enhance the sentiment analysis process. In the following paragraphs, some of these hybrid techniques will be introduced. The algorithm proposed in [36] was designed to examine feedback on the same topic in order to derive useful contextual knowledge. Semantic similarity measurements were used to verify a given opinion's semantic orientation. It took into account context-dependent opinions by employing linguistic rules to carry out semantic orientation of contextually dependent opinions. Then, in order to judge contextually dependent opinions, it collected contextual information from other reviews for the same product feature. Peng and Shih investigated an unsupervised learning method that uses rules of part-of-speech (POS) patterns to extract the sentiment phrases of each review. For each unknown sentiment phrase, they used each phrase as a query to find the top-N relevant snippets. After gathering the sentiment of the lexicon, the prediction sentiments of unknown emotion phrases were computed. The sentiments of adjacent known sentiment terms inside the snippets were then used to identify predictive sentiments. For opinion extraction, they only examined sentences that contained at least one detected sentiment phrase. Finally, the POS pattern was used to finish the opinion extraction [37]. Additionally, utilizing TF-IDF weight on the raw data, a technique based on the k-means clustering algorithm was devised. Here, a voting process was employed to obtain a stable clustering result. After that, the documents were divided into two groups: positive and negative [38]. The semantic orientation approach with other machine learning approaches, such as the n-gram model, has been used to analyze movie reviews [39]. The result gained by the machine learning approach was accurate; however, the model's training process required a significant amount of time. Although less accurate, the semantic orientation technique is more efficient in real-time applications. The performance of the underlying POS tagger determines the performance of semantic orientation. In ref. [40], the authors used a hybrid method that combined a corpus-based algorithm with SVM, resulting in good accuracy for light-stemmed data. Moreover, the three-stage cascade method that focuses on identifying the polarity shift problem in document level sentiment analysis was proposed in [41]. Each document was broken down into a set of subsentences. It created a hybrid model that combines certain principles with a statistical technique to find explicit and implicit polarity shifts. Then, to reduce polarity shift in negations, it presented a polarity shift exclusion approach. Finally, it trained the base classifiers on training subsets that were split by various categories of polarity shifts and used a weighted combination of the component classifiers for sentiment analysis. The hybrid algorithm was suggested in ref. [42] that combined the K-means and DENCLUE (DENsity-based CLUstEring) methods to utilize the precise number of clusters and an adequate clustering performance, applying this to four different Twitter datasets.

6 Recent techniques in sentiment analysis

As presented above, the machine learning, lexicon, and hybrid approaches used to optimize these solutions can solve the sentiment analysis problem either as a classification or clustering problem. In this section, recent studies will be presented to show the new works that have been published in the area of sentiment analysis research. Currently, many approaches have been adapted to deal with sentiment analysis and tackle the large volume of data that is generated by different social networking platforms. In this section, some of the works that have been applied to sentiment analysis problems will be presented to understand and analyze users' behaviors and attitudes regarding different online activities. Usually, social networking users send messages containing different multimodal contents. Such contents are generally brief and lack explicit sentiment words. To analyze the context of messages and understand the sentiment associated with them [31]. a model has been proposed that entails the semantic correlation between various modalities as well as the impacts of tweet context information.

Many research works were applied to customer reviews created during online shopping. The study that looked at product reviews in support of consumers' purchase decisions is presented in [43]. The authors of this study reviewed existing research on information fusion processes and methodologies. In the analysis stage, they summarized the main conclusions that point to future research directions in this area. Additionally, a comparative study was conducted in an educational context, using three classifiers including an evolutionary algorithm. It aimed to mine the public opinions labeled with learning-centered emotions [44]. A new regression-based approach was proposed to learn, analyze, and classify products and shop information based on customer experience. The data gathered from the benchmark of the Unified Computing System (UCS), which is a server for data-based computer products, aligns with a hardware evaluation, visualization support, and software management [45]. Furthermore, a sentiment analysis approach based on deep learning was used on product reviews acquired from Twitter. The TF-IDF weighted glove word embedding approach was integrated with the Convolutional Neural Networks (CNNs) method. This architecture consists of five tiers: a weighted embedding tier, convolution tier, max-pooling tier, followed by the LSTM and dense tiers [46].

Aspect-based sentiment analysis (ABSA) techniques, which anticipate the sentiment of specific aspects in a text, are also an increasing new trend. The recognition of the relevant contexts for different aspects is a challenge, hence the method was proposed for feature extraction networks to eliminate noise and enhance the aspect features for sentiment [47]. Additionally, the algorithm aggregates scores for the high occurrence class and the second class to predict the result score [48]. The predicted score is used to calculate the probability for each sentence to be part of a certain class from the five-star scale using a lexicon-based system. The final probabilities are utilized to determine the overall sentiment at a document-level. Moreover, aspect-based sentiment analysis has been applied to identify both sentiments and aspects. The contextual word representations are obtained from the pre-trained linguistic model named BERT (Bidirectional Encoder Representations from Transformers) [49].

One of the most attractive domains for sentiment analysis is analyzing traveler reviews. Sharing opinions between travelers can help people and decrease the risk of making bad decisions regarding their accommodation. Automatic review summarizations are a promising means of improving information processing of traveler reviews. The study conducted in [50] focused on extracting the relevant text features and performing sentiment analysis to compile review summaries. It proposed a systematic approach that used four classifiers to identify helpful reviews: Decision Tree (DT), Random Forest (RF), Logistic Regression (LR), and SVM. It then classified the sentences into six characteristics, which were subsequently examined to generate review summaries. Additionally, other research has explored the impact of traveler reviews to make direct endorsements in text. It applied a text mining approach to the online reviews to identify the effect of explicit recommendations [1].

The deep learning model mentioned above can be considered a means to solving sentiment analysis problems. For example, in a recent study of Chinese language [31], the failure to fully utilize emotional information was identified as an issue. A single Chinese character might have multiple meanings in different words, and the character embeddings are combined with the word embeddings to obtain a more accurate meaning of the information. The CNN was used to recognize the sentiment classification of sequence semantic features. Thus, the suggested model for sentiment information extraction and analysis is based on a multi-neural network. Additionally, the deep neural networks model has been applied to evaluate the sentiment of text through combining the end-to-end memory neural network and termed Recurrent Memory Neural Network. The goal of this model is to address a weak interaction in the attention mechanism by creating a multielement attention mechanism that creates powerful attention weights and more accurate aspects based on sentiment representation [51]. Moreover, neural network research has considered the tweet-text information and user-connection information in Twitter data. However, the Attentional-graph Neural Network based Twitter Sentiment Analyzer (AGN-TSA) has been proposed to perform sentiment analysis using attentional-graph neural networks. AGN-TSA uses a three-layered neural structure that includes a word layer, a user layer, and an attentional graph network layer to merge Twitter text and user connection information [52].

Sentiment analysis for short text is challenging because of the lack of context. Song et al. addressed this issue by proposing a text representation model for short text sentiment analysis based on probabilistic linguistic terms sets (PLTSs). In this model, every word is represented as a PLTS that entirely describes the possibilities for the word's sentiment polarity. They employed support vector machines and a polarity classification framework, which is a supervised and unsupervised learning technique [53]. In addition, a new study has presented a new approach for Twitter sentiment analysis that merges the results of standard classifiers and Natural Language Processing (NLP) techniques [54]. The proposed technique uses a fuzzy metric to decide the significance of each classifier. Hence, the fuzzy metrics are combined with the Choquet fuzzy to produce the final label. Furthermore, a machine learning technique based on a lexicon is proposed in ref. [55]. The polarity of the reviews is analyzed through negations, intensifiers, acronyms, and punctuation. Next, the words are used for computing the score, and the machine learning algorithms are used.

Two well-known clustering algorithms, K-means and DENCLUE (DENsity-based CLUstEring), were combined in one model in [56] to exploit the accurate number of clusters from K-means and the clustering performance from DENCLUE. This model was applied to Twitter datasets and its effectiveness was tested against state-of-the art methods.

Finally, sentiment analysis has been applied in many different situations to gain insight from people's behavior and related feedback. For example, the research presented by Ruz et al. addressed the problem of sentiment analysis during critical events such as environmental disasters or social activities. They applied Bayesian network classifiers to implement sentiment analysis on Spanish datasets [57]. Moreover, the ambiguity in political, religious, and social matters leads to extremism among people who are represented by their sentiments on social media platforms using English, which is the most commonly used language for sharing viewpoints on social media. Asif et al. studied the sentiment analysis of social media using multilingual textual data to determine the amount of extremist sentiment. They built a multilingual lexicon with different intensity weights. Then, the Multinomial Naive Bayes and Linear Support Vector Classifier algorithms were applied to perform the classification [58]. Authors in [59] , [60] studied COVID-19 and recognized the strange situation of pressure that was considered on each country to establish plans to manage the population and utilize the existing resources in more convenient way. The data was collected from Twitter based on hashtags that including COVID-19, coronavirus, new case, deaths, recovered, and so on. The authors have proposed a hybrid heterogeneous Support Vector Machine approach.

7 Comparative analysis and discussion

As presented in the previous sections, many techniques have been applied by the researchers to solve sentiment analysis problems. The different techniques were analyzed and compared by showing their performance and the type of datasets, as presented in Table 1.

Table 1: Summary of different sentiment analysis approaches.

Ref.	Source	Techniques	Domain	Accuracy%
[16]	EMNLP (2002)	Naïve Bayes	Movie reviews	81
[17]	Information Processing and Management (2019)	Bayesian Network	Hotel reviews	59
[18]	I4CT (2014)	SVM	Movie reviews & SFU educational review	83
[19]	ICCIA (2016)	ANN	Reuters corpus documents	68
[41]	Information Processing and Management (2016)	SVM, logistic regression & Naïve Bayes	Book, DVD, and Electronics reviews (Amazon.com)	83.8, 84.6 & 85.6
[21]	International Journal of Computer Applications (2016)	Spectral clustering approach	Twitter (Tweets about movies)	72
[25]	Expert System with applications (2017)	DNN	SemEval2013, Seval2014, Vader, STS-Gold, IMDB, PL04	82.51, 82.51, 71.92, 85.71, 75.47, 70.82
[26]	InAES (2017)	CNN	Twitter	75
[28]	CITS (2017)	RNN	Place reviews from different social networking sites	90.34
[35]	International Journal of Advanced Computer Science and Applications (2017)	Corpus-based using KNN	Twitter	81
[22]	Knowledge-Based Systems (2020)	K-means & DENCLUE (DENsity-based CLUstEring)	Restaurant reviews (English, Dutch, Russian, Spanish & Turkish)	71
[32]	SERA (2016)	Lexicon-based (3 online dictionaries)	Sanders Twitter Sentiment Corpus by Niek J. Sanders1	58.11
[33]	Information Processing and Management (2016)	Lexicon-based with TF-IDF	Twitter	87.5
[34]	4th ACM International Conference on Information and Knowledge Management (2005)	Corpus-based with Cosine-normalized TF-IDF	HM2, TL3, & KA4 term sets	61.18, 68.53 & 65.49
[37]	IEEE/WIC/ACM (2010)	Corpus-based with Unsupervised snippet-based sentiment	Chinese UGC corpus	80

1 http://www.sananalytics.com/lab/twitter-sentiment/
2 http://www.ldc.upenn.edu/Catalog/
3 http://www.altavista.com/
4 http://www.tasa.com/

		classification		
[38]	ISKE (2010)	k-means & TF-IDF	Movie reviews	78.33
[39]	The 38th Hawaii International Conference on System Sciences (2005)	N-gram Classifiers & Semantic Orientation approach	Movie reviews	85.54
[40]	AEECT (2013)	Corpus-based approach & SVM	Twitter	87.2
[3]	International Journal of Advanced Computer Science and Applications (2018)	SVM	Twitter	87
[50]	Tourism Management (2020)	Decision Tree (DT), Random Forest (RF), Logistic Regression (LR), and SVM	Online hotel reviews (TripAdvisor.com)	70.4, 73.7, 69.5 & 63.7
[44]	Expert Systems with Applications (2020)	EvoMSA Multinomial NB KNN Decision Tree B4MSA Bernoulli NB SVC Linear SVC Random Forest Deep Learning (LSTM CNN_5a Deep Learning CNN_10a CNN_LSTM_7a BERT)	sentiTEXT & eduSERE corpus	93, 84 87, 79 79, 68 85, 72 92, 83 87, 76 90, 79 90, 79 89, 77 89, 79 91, 68 90, 80 88, 74 93, 83
[47]	Complex and Intelligent Systems (2020)	ML based regression model	Product reviews	98
[61]	Tsinghua Science and Technology (2020)	Probability based model	Twitter	73

[45]	Complex and Intelligent Systems (2020)	Regression problem	Product information from UCS	98
[46]	Concurrency and Computation - Practice & Experience (2020)	TF-IDF with CNN	Twitter	87.12
[51]	Neurocomputing (2020)	Recurrent Memory Neural Networks	Twitter	94.09
[1]	Journal of Hospitality and Tourism Management (2020)	Binary logistic regression	Traveler reviews	66.86
[48]	Journal of Information Science (2020)	Lexicon-based approach using Score Aggregation	Hotel datasets from TripAdvisor & restaurant reviews from CitySearch	79
[31]	International Journal of Electrical Engineering & Education (2020)	Multi-neural network	JOYBUY user comment dataset for food	88.13
[57]	Future Generation Computer Systems (2020)	NB, TAN, BF TAN, SVM & RF	Spanish datasets (Catalan independence referendum)	78.1, 80.8, 80.8, 82.9 & 85.8
[58]	Telematics and Informatics (2020)	Linear SVC	Multilingual dataset (acebook)	82.1
[53]	Knowledge-Based Systems (2020)	SVM	Movie reviews Stanford Twitter TripAdvisor reviews	76 72 78
[54]	Journal of Information Science (2020)	Fusion + NLP	Stanford, Stanford Twitter Sentiment, SemEval-2016 Movie Reviews	81, 83.6, 66.6, 83.8
[55]	International Journal of	Negation-intensifier-	Books dataset	90.3

	Technology and Human Interaction (2020)	punctuation-acronyms (NIPA) + SVM NIPA + Naïve Bayes	(Taboada)	91.3
[56]	Knowledge-Based Systems (2020)	K-means & DENCLUE (DENsity-based CLUstEring)	SemEval dataset for English, Dutch, Russian, Spanish, Turkish	69, 61, 63, 78, 51
[52]	Information (2020)	Attentional-graph Neural Network based Twitter Sentiment Analyzer (AGN-TSA)	Twitter	94.62
[49]	22nd Nordic Conference (2019)	Pre-trained linguistic model named BERT	SemEval-2016 (Restaurant, Laptop Reviews)	96.3, 98.4
[59]	Inf Syst Front (2021)	Hybrid heterogeneous approach that applied SVM	Twitter data, COVID19 related data	96

From Table 1, it has been observed that the accuracy of the sentiment analysis process depends upon the dataset, the selected features, and the classification or clustering algorithm that has been applied. Hence, the impact of data quality on the sentiment analysis process is affected by the high readability of the datasets. Additionally, most of studies consider the overall accuracy or F-score as a reliability metric and ignore the processing time. In such situation it is difficult to determine which technique has the best performance. However, when considering some algorithms that have been applied on Twitter data, we observed that the following algorithms gave the highest accuracy arranged from high to low: hybrid heterogeneous SVM, Attentional-graph Neural Network, Recurrent Memory Neural Networks, Corpus-based approach & SVM , TF-IDF & CNN, SVM, Lexicon-based with TF-IDF, Corpus-based using KNN, and CNN.

We also observed that English is the most used language in sentiment analysis, but we found that some researchers have applied sentiment analysis for other languages, such as Arabic, Chinese, Dutch, Russian, Spanish, and Turkish. Additionally, the datasets used in sentiment analysis are mostly extracted from Twitter data, movie reviews, or traveler reviews. Therefore, sentiment analysis has mostly been applied in product and service reviews from business perspectives. Moreover, some new research has applied sentiment analysis to situations related to critical events, such as environmental disasters, social activities in social networking, or extremism. This presents the possibility of new research that applies sentiment analysis to many new situations related to different human activities and behaviors.

8 Conclusion

This paper discussed the different issues related to sentiment analysis. First, the social networking framework was presented, and the processes were determined, with an explanation of how the sentiment analysis process was integrated as a part of this framework. Then, the existing techniques that have been applied in sentiment analysis were explained in more detail. These techniques are generally divided into three main categories: machine learning, lexicon-based, and hybrid approaches. Additionally, sentiment analysis based on machine learning can be further classified into three categories: supervised machine learning, unsupervised machine learning, and deep learning techniques. The social networking content is textual data containing the emotional behavior of users, which can be analyzed and used in decision-making processes. The sentiment analysis of such content is performed using different techniques that classify social networking posts and reviews based on three main processes: preprocessing, feature extraction, and classification or clustering. These three processes have been applied in all techniques of sentiment analysis. Finally, a brief review was presented of the recent studies, and a comparative study of different works was provided based on a set of criteria including the technique used, domain, and accuracy.

Acknowledgement

Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2022R235), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

References

[1] J. Guerreiro and P. Rita, "How to predict explicit recommendations in online reviews using text mining and sentiment analysis," Journal of Hospitality and Tourism Management, vol. 43, p. 269–272, 2020 https://doi.org/10.1016/j.jhtm.2019.07.001

[2] A. J. Sanur Sharma, "Role of sentiment analysis in social media security and analytics," Wiley Interdisciplinary Reviews- Data Mining and Knowledge Discovery, vol. 10, no. 51,, p. 1366, 2020. https://doi.org/10.1002/widm.1366

[3] S. Al-Otaibi, A. Alnassar, A. Alshahrani, A. Al-Mubarak, S. Albugami and N. Almutiri, "Customer satisfaction measurement using sentiment analysis," International Journal of Advanced Computer Science and Applications, vol. 9, no. 2, pp. 106-117,

2018.
https://doi.org/10.14569/ijacsa.2018.090216

[4] T. Kolenik and M. Gams, "Persuasive Technology for Mental Health: One Step Closer to (Mental Health Care) Equality?," IEEE Technology and Society Magazine, vol. 40, no. 1, pp. 80-86, 2021. https://doi.org/10.1109/mts.2021.3056288

[5] T. Kolenik and M. Gams, "Intelligent cognitive assistants for attitude and behavior change support in mental health: state-of-the-art technical review," Electronics, vol. 10, no. 11, p. 1250, 2021.
https://doi.org/10.3390/electronics10111250

[6] D. M. E. D. M. Hussein, "A survey on sentiment analysis challenges," King Saud University - Engineering Sciences, vol. 30, no. 4, p. 330–338, 2018.
https://doi.org/10.1016/j.jksues.2016.04.002

[7] M. Reformat and K. S. Golmohammadi, "Rule- and OWA-based semantic similarity for user profiling," International Journal of Fuzzy Systems, vol. 12, no. 2, pp. 87-102, 2010.

[8] S. Al-Otaibi, A. A. Al-Rasheed, B. AlHazza, H. A. Khan, G. AlShfloot, M. AlFaris, N. AlFari, and N. AlKhalaf. ", "Finding Influential Users in Social Networking using Sentiment Analysis," Informatica, vol. 46, no. 5, 2022.
https://doi.org/10.31449/inf.v46i5.3829

[9] W. Fan and M. D. Gordon, "The power of social media analytics," Communications of the ACM, vol. 57, no. 6, p. 74–81, 2014.

[10] O. Chong, S. Sheila and A. Soliman, "Social media analysis framework: the case of Twitter and super bowl ads," Journal of Information Technology Management, vol. XXVI, no. 1, pp. 1-18, 2015.

[11] E. Younis, "Sentiment analysis and text mining for social media microblogs using open-source tools: an empirical study," International Journal of Computer Applications, vol. 112, no. 5, pp. 44-48, February 2015.

[12] F. A. Pozzi, E. Fersini, E. Messina and B. Liu, Sentiment Analysis in Social Networks, 1st ed., Henderson, NV, USA: Morgan Kaufmann, 2017. https://doi.org/10.1016/b978-0-12-804412-4.00001-2

[13] G. Beigi, X. Hu, R. Maciejewski and H. Liu, "An overview of sentiment analysis in social media and its applications in disaster relief," Sentiment analysis and ontology engineering, vol. 639, p. 313–340, 2016.
https://doi.org/10.1007/978-3-319-30319-2_13

[14] R. Tejwani, "Sentiment Analysis: A Survey," ArXiv, pp. 1-3, 2014.

[15] S. Kaur & R. Mohana, "Prediction of sentiment from macaronic reviews," Informatica , vol. 42, no. 1, pp. 127-137, 2018.

[16] B. Pang, L. Lee and S. Vaithyanathan, "Thumbs up? Sentiment Classification using Machine Learning Techniques," in proc. of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2002), Philadelphia, PA, USA, 2002.
https://doi.org/10.3115/1118693.1118704

[17] M. Al-Smadi, M. Al-Ayyoub, Y. Jararweh and O. Qawasmeh, "Enhancing aspect-based sentment analysis of Arabic hotels' reviews using morphological, syntactic and semantic features," Information Processing and Management, vol. 56, no. 2, pp. 308-319, 2019.
https://doi.org/10.1016/j.ipm.2018.01.006

[18] N. Zainuddin and A. Selamat, "Sentiment analysis using support vector machine," in proc. International Conference on Computer, Communications, and Control Technology (I4CT), Kuching, Malaysia, 2014.
https://doi.org/10.1109/i4ct.2014.6914200

[19] L. Vega and A. Mendez-Vazquez, "Dynamic neural networks for text classification," in proc. International Conference on Computational Intelligence and Applications (ICCIA), Leuven, Belgium, 2016.
https://doi.org/10.1109/iccia.2016.15

[20] D. Sharma, M. Sabharwal, V. Goyal and M., "Sentiment Analysis Techniques for Social Media Data: A Review," in proc. First International Conference on Sustainable Technologies for Computational Intelligence, Singapore, pp. 75-90, 2020.
https://doi.org/10.1007/978-981-15-0029-9_7

[21] M. Unnisa, A. Ameen and S. Raziuddin, "Opinion mining on Twitter data using unsupervised learning technique," International Journal of Computer Applications, vol. 148, no. 12, pp. 12-19, 2016.
https://doi.org/10.5120/ijca2016911317

[22] L. R. C. Pessutto, D. S. Vargas and V. P. Moreira, "Multilingual aspect clustering for sentiment analysis," Knowledge-Based Systems, vol. 192, no. 9, 2020.
https://doi.org/10.1016/j.knosys.2019.105339

[23] N. C. Dang, M. N. Moreno-García and F. D. P. la, "Sentiment analysis based on deep learning: a comparative study," Electronics, vol. 9, no. 3, pp. 483-512, 2020.
https://doi.org/10.3390/electronics9030483

[24] C. C. Aggarwal, Neural Networks and Deep Learning, Berlin, Germany: Springer, 2018.

[25] O. Araque, I. Corcuera-Platas, F. J. anchez-Rada and C. A. Iglesias, "Enhancing deep learning sentiment analysis with ensemble techniques in social

applications," Expert System with Applications, vol. 77, p. 236–246, 2017. https://doi.org/10.1016/j.eswa.2017.02.002

[26] A. M. Ramadhani and H. S. Goo, "Twitter sentiment analysis using deep learning methods," in proc. 7th International Annual Engineering Seminar (InAES), Yogyakarta, Indonesia, 2017.

https://doi.org/10.1109/inaes.2017.8068556

[27] D. Britz, "Recurrent neural networks tutorial, part 1–introduction to RNNs.," WILDML Artificial Intelligence, Deep Learning, and NLP, 17 September 2020. [Online]. Available: wildml.com.

[28] G. Preethi, P. V. Krishna, M. S. Obaidat, V. Saritha and S. Yenduri, "Application of deep learning to sentiment analysis for recommender system on cloud," in proc. 2017 International Conference on Computer, Information and Telecommunication Systems (CITS), Dalian, China, 2017. https://doi.org/10.1109/cits.2017.8035341

[29] P. Patil and P. Yalagi, "Sentiment analysis levels and techniques: a survey," International Journal of Innovations in Engineering and Technology (IJIET), vol. 6, no. 4, pp. 523-528, 2016.

[30] W. Etaiwi, D. Suleiman and A. Awajan, "Deep Learning Based Techniques for Sentiment Analysis: A Survey," Informatica , vol. 45, no. 7, 2021. https://doi.org/10.31449/inf.v45i7.3674

[31] Y. Li, Q. Jin, M. Zuo, H. Li, X. Yang, Q. Zhang and X. Liu, "Multi-neural network- based sentiment analysis of food reviews based on character and word embeddings," International Journal of Electrical Engineering and Education, pp. 1-12, 2020. https://doi.org/10.1177/0020720920928492

[32] S. Park and K. Y, "Building thesaurus lexicon using dictionary-based approach for sentiment classification," in proc. IEEE 14th International Conference on Software Engineering research, Management and Applications (SERA), Towson, MD, USA, 2016. https://doi.org/10.1109/sera.2016.7516126

[33] H. Saif, Y. He and M. Fernandez, "Contextual semantics for sentiment analysis of twitter," Information Processing and Management, vol. 52, no. 1, p. 5–19, 2016. https://doi.org/10.1016/j.ipm.2015.01.005

[34] A. Esuli and F. Sebastiani, "Determining the semantic orientation of terms through gloss classification," in proc. of 4th ACM International Conference on Information and Knowledge Management, Bremen, Germany, 2005. https://doi.org/10.1145/1099554.1099713

[35] M. R. Huq, A. Ali and A. Rahman, "Sentiment analysis on Twitter data using KNN and SVM," (IJACSA) International Journal of Advanced Computer Science and Applications, vol. 8, no. 6, pp. 19-25, 2017. https://doi.org/10.14569/ijacsa.2017.080603

[36] C. Wu and L. Shen, "A new method of using contextual information to infer the semantic orientations of context dependent opinions," in proc. International Conference on Artificial Intelligence and Computational Intelligence, Shanghai, China, pp. 274-278, 2009. https://doi.org/10.1109/aici.2009.406

[37] T. C. Peng and C. C. Shih, "An unsupervised snippet-based sentiment classification method for chinese unknown phrases without using reference word pairs," in proc. IEEE/WIC/ACM International Conference on Web Intelligence and intelligent Agent Technology, Toronto, Canada, 2010. https://doi.org/10.1109/wi-iat.2010.229

[38] G. Li and F. Liu, "A Clustering-based approach on sentiment analysis," in proc. IEEE International Conference on Intelligent Systems and Knowledge Engineering, ISKE, Hangzhou, China, pp. 31-337, 2010. https://doi.org/10.1109/iske.2010.5680859

[39] L. Z. Chaovalit, "Movie review mining: a comparison between supervised and unsupervised classification approaches," in proc. of the 38th Hawaii International Conference on System Sciences, Hawaii, USA, pp. 112, 2005. https://doi.org/10.1109/hicss.2005.445

[40] N. A. Abdulla, N. Ahmed, M. Shehab and M. Al-Ayyoub, "Arabic sentiment analysis: Lexicon-based and corpus-based and corpus-based," in proc. IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT), ., Amman, Jordan, 2013. https://doi.org/10.1109/aeect.2013.6716448

[41] R. Xia, F. Xu, J. Yu, Q. Yong and E. Cambria, "Polarity shift detection, elimination and ensemble a three-stage model for document-level sentiment analysis," Information Processing and Management, vol. 52, no. 1, p. 36–45, 2016. https://doi.org/10.1016/j.ipm.2015.04.003

[42] A. I. Hajar Rehioui, "New clustering algorithms for Twitter sentiment analysis," IEEE Systems Journal, vol. 14, no. 1, pp. 530-537, 2020. https://doi.org/10.1109/jsyst.2019.2912759

[43] Z Fan, G. Li and Y. Liu, "Processes and methods of information fusion for ranking products based on online reviews: an overview," Information Fusion, vol. 60, pp. 87-97, 2020. https://doi.org/10.1016/j.inffus.2020.02.007

[44] M. L. B. Estrada, R. Z. Cabada, R. O. Bustillos and M. Graff, "Opinion mining and emotion recognition applied to learning environments," Expert Systems With Applications, Vols. 150, 113265, 2020 https://doi.org/10.1016/j.eswa.2020.113265

[45] Y. Shanshan and L. Xiaofang, "Machine learning based customer sentiment analysis for recommending shoppers, shops based on customers' review," Complex and Intelligent Systems, vol. 6, no. 5, pp. 1-14, 2020. https://doi.org/10.1007/s40747-020-00155-2

[46] A. Onan, "Sentiment analysis on product reviews based on weighted word embeddings and deep neural networks," Concurrency Computation- Practice and Experience, no. Special Issue, pp. 1-12, 2020. https://doi.org/10.1002/cpe.5909

[47] K. Shuang, Q. Yang, J. Loo, R. Li and M. Gu, "Feature distillation network for aspect-based sentiment analysis," Information Fusion, vol. 61, pp. 13-23, 2020. https://doi.org/10.1016/j.inffus.2020.03.003

[48] P. J. Khiabani, M. E. Basiri and H. Rastegari, "An improved evidence-based aggregation method for sentiment analysis," Journal of Information Science, vol. 46, no. 3, p. 340–360, 2020. https://doi.org/10.1177/0165551519837187

[49] C. Sun, L. Huang and X. Qiu, "Utilizing BERT for aspect-based sentiment analysis via constructing auxiliary sentence," in proc. NAACL HLT-Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Turku, Finland, 2019. https://doi.org/10.18653/v1/2021.naacl-main.146

[50] C. Tsai, K. Chen, Y. Hu and W. Chen, "Improving text summarization of online hotel reviews with review helpfulness and sentiment," Tourism Management, Vols. 80, 104122, 2020. https://doi.org/10.1016/j.tourman.2020.104122

[51] N. Liu and B. Shen, "ReMemNN: A novel memory neural network for powerful interaction in aspect-based sentiment analysis," Neurocomputing, vol. 395, pp. 66-77, 2020. https://doi.org/10.1016/j.neucom.2020.02.018

[52] M. Wang and G. Hu, "A novel method for Twitter sentiment analysis based on attentional-graph neural network," Information, vol. 11, no. 2 , p. 92, 2020. https://doi.org/10.3390/info11020092

[53] C. Song, X. Wang, P. Cheng, J. Wang and L. Li, "SACPC: A framework based on probabilistic linguistic terms for short text sentiment analysis," Knowledge-Based Systems, Vols. 194, 105572, 2020. https://doi.org/10.1016/j.knosys.2020.105572

[54] M. Emadi and M. Rahgozar, "Twitter sentiment analysis using fuzzy integral classifier fusion," Journal of Information Science, vol. 46, no. 2, p. 226–242, 2020. https://doi.org/10.1177/0165551519828627

[55] T. Sahu and S. Khandekar, "A machine learning-based Lexicon approach for sentiment analysis," International Journal of Technology and Human Interaction, vol. 16, no. 2, pp. 8-22, 2020. https://doi.org/10.4018/ijthi.2020040102

[56] L. Rafael, C. Pessutto, D. S. Vargas and V. P. Moreira, "Multilingual aspect clustering for sentiment analysis," Knowledge-Based Systems, Vols. 192, 105339, 2020. https://doi.org/10.1016/j.knosys.2019.105339

[57] G. Ruz, P. Henríquez and A. Mascareño, "Sentiment analysis of Twitter data during critical events through Bayesian networks classifiers," Future Generation Computer Systems, vol. 106, p. 92–104, 2020. https://doi.org/10.1016/j.future.2020.01.005

[58] M. Asif, A. Ishtiaq, H. Ahmad, H. Aljuaid and J. Shah, "Sentiment analysis of extremism in social media from textual information," Telematics and Informatics, Vols. 48, 101345, 2020. https://doi.org/10.1016/j.tele.2020.101345

[59] H. Kaur, S. Ahsaan, B. Alankar and V. Chang, "A Proposed Sentiment Analysis Deep Learning Algorithm for Analyzing COVID-19 Tweets," Information Systems Frontiers, vol. 23, no. 6, p. 1417–1429, 2021. https://doi.org/10.1007/s10796-021-10135-7

[60] S. K. Akpatsa, X. Li, H. Lei and V.-H. K. S. Obeng, "Evaluating Public Sentiment of Covid-19 Vaccine Tweets Using Machine Learning Techniques," Informatica, vol. 46, no. 1, 2022. https://doi.org/10.31449/inf.v46i1.3483

[61] B. Liu, S. Tang, X. Sun, Q. Chen, J. Cao, J. Luo and S. Zhao, "Context-aware social media user sentiment analysis," Tsinghua Science and Technology, vol. 25, no. 4, p. 528–541, 2020. https://doi.org/10.26599/tst.2019.9010021