IEEE Access
Multidisciplinary ⁞ Rapid Review ⁞ Open Access Journal

Received 15 October 2024, accepted 3 December 2024, date of publication 9 December 2024, date of current version 27 December 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3513321

SURVEY

A Review on Text Sentiment Analysis With Machine Learning and Deep Learning Techniques

YONATAN MAMANI-COAQUIRA AND EDWIN VILLANUEVA
Engineering Department, Pontificia Universidad Católica del Perú, San Miguel, Lima 15088, Peru
Corresponding author: Yonatan Mamani-Coaquira (ymamanic@pucp.edu.pe)

This work was supported by Fondo Nacional de Desarrollo Científico, Tecnológico y de Innovación Tecnológica (Fondecyt) of Concytec, which funded the doctoral studies.

ABSTRACT Automating sentiment analysis in texts has become an important task in recent years due to the exponential growth of user-generated content, including comments and opinions on products and services. This represents a valuable opportunity for businesses to glean insights into customer sentiment and, in turn, to refine their offerings. Motivated by this, the machine learning field has witnessed a surge of innovation, with an introduction of models and tools being introduced to streamline sentiment analysis. This paper offers a thorough review of the recent advancements in machine learning and deep learning approaches for text sentiment analysis. We propose a novel framework for studying these models, distinguishing them by their structural intricacies. Additionally, we delve into the challenges, prospects, and emerging directions in research, as illuminated by our framework. Consequently, this paper equips researchers with a detailed panorama of the cutting-edge machine learning methodologies for dissecting text sentiment, easing the way for future explorations in this vibrant field.

INDEX TERMS Machine learning, deep learning, word embedding, text encoding, sentiment analysis, text classification.

I. INTRODUCTION

In recent years we have seen a dizzying growth of user-generated content on different platforms on the internet, such as social networks, e-comerce sites, e-learning platforms, blogs, forums, etc.. These contents have valuable information that can help companies and government make better decisions and offer better products and services [1]. Much of this data contains information about opinions, feelings, emotions and appreciations about products and services consumed by users and that end up influencing the behavior of other users.

Motivated by the need to automate the extraction of this information, the academic community has been actively proposing a variety of methods and models for the identification of emotions and sentiments from a variety of data types, such as texts, audios, images, videos, emoticons, among others. For example, we can find a plethora of methods, referred to as Sentiment Analysis (SA) methods, proposed to identify the polarity of sentiments (positive, negative or neutral) expressed in texts [2], [3], [4]. We can also find methods focused on detecting emotions, such as joy, surprise, anger, disgust, fear and sadness [3], [5], [6].

Most methods for sentiment analysis proposed in recent years are based on Machine Learning (ML) and Deep Learning (DL) models. A key step in the development of such models was the identification of relevant and predictive features, task known as feature extraction (FE). The first features proposed for SA tasks were based on Bag of Word (BoW) [7], Term Frequency-Inverse Document Frequency (TF-IDF) [8], and N-gram vectorization [9], [10]. These features were used in conjunction with classical ML algorithm for the sentiment classification, such as Naive Bayes (NB), Support Vector Machine (SVM), Logistic Regression (LR), Decision Tree (DT), Random Forest (RF), among others [11], [12], [13], [14]. However, due to the limited representativeness of the aforementioned features, several scholars focused their attention on proposing ways to learn more representative features for SA and other textual analisys tasks. By instance, new ways of finding semantic

The associate editor coordinating the review of this manuscript and approving it for publication was Ikramullah Lali.

representations of words (referred to as word embeddings) were proposed, such as Word2Vec [15], Glove [16], and FastText [17]. Along with these new word representations, new models based on deep learning architectures were tried for SA tasks, such as Recurrent Neural Networks (RNN) [18], Convolutional Neural Networks (CNN) [19], and neural networks architectures with Attention Mechanisms (AM) [20], [21], [22], [23]. This new wave of text representations and models has shown promise in SA and related tasks.

The field of sentiment analysis has seen rapid growth, with a large number of methods introduced recently, not only to improve accuracy but also to address emerging forms of textual communication, such as jargon, hyper-abbreviated texts and short words, emoticons, etc. Due to this vibrant wave of innovations, it is difficult for newcomer scholars in the area to have a clear panorama of the methodologies, challenges and research opportunities in the field. There are some reviews of the SA literature that may help new researchers in the area. Jain et al. [11] presented a systematic review on ML techniques for consumer Sentiment Analysis, focusing on sentiment classification, predictive recommendation, and fake hotel and tourism reviews. Minaee et al. [21] conducted a comprehensive review on DL models for text classification, describing different hybrid and independent models used for Sentiment Analysis and listing the 40 most popular datasets in this area. Zhang et al. [1] published a survey of DL models for SA, showing different architectures of hybridized or independent deep neural networks and sharing information from research spanning document, sentence, and aspect levels for opinion processing. Zarandi and Mirzaei [24] reviewed different DL, ML, and Graph Neural Network (GNN) models for sentiment classification, including GAN models.

While these reviews can help to better understand the progress in the area, they focus on limited aspects of the sentiment analysis process, such as FE techniques or classification models. Notwithstanding, several steps are involved in the pipeline of SA model development, such as the selection of the training dataset, the data cleaning and preprocesing steps, the FE method, the selection of performance metrics, the training and evaluation of ML/DL models and the model deployment and application.

Additionaly, on the last few years have appeared several novel SA methods that have not been covered by previous reviews. During our literature review on sentiment analysis studies we found that only a few studies covers novel deep learning models for SA tasks [1], [11], [25], [26], [27]. Furthermore, we observed that the scope of the reviewed SA applications focuses on a few application domains, like hospitals, tourism [11], and e-commerce [24]. On the other hand, with respect to the process of text representation and attribute extraction, most review studies analyze classical methods, like TF-IDF [25], [26], Word2vec, and Glove [1], [11]. We believe that there is a need to delve deeper into most recent text representation methods, such as BERT, Roberta, FastText, etc., in applications with ML/DL models to improve FE.

In the present study we present a comprehensive review of SA models and techniques. For this purpose, we introduce a framework to study all these methods. This framework facilitates the categorization and organization of different processes, techniques, and algorithms from recent research. The main contributions of this study are the following:

- A framework that provides a useful categorization of machine learning and deep learning sentiment analysis models.
- A description and discussion of the combination of recent models in deep learning for sentiment polarity classification.
- Discussion and suggestions for future directions in Sentiment Analysis.

II. CONCEPTUAL BACKGROUND OF SENTIMENT ANALYSIS

Sentiment Analysis is primarily focused on determining the polarity or emotion that the person expresses in different data sources [21], [23], [28], [29].

A. MODALITIES OF DATA FOR SENTIMENT ANALYSIS

Different modalities of data have been used to determine or classify the sentiments or emotions expressed by individuals [22], [29], [30], [31], such as:

- Text
- Image (body gesture or facial expression)
- Video (body gesture or facial expression)
- Audio (voice)
- Emoticon

B. SENTIMENT ANALYSIS TASKS

Tasks for SA can be classified into two general categories: opinion and emotion analysis [12], [23], [32]. However, many expressed opinions do not manifest emotion; therefore, emotions and opinions are not equivalent [3].

1) OPINION MINING

This field of study aims to analyze expressions manifested by people or organizations. The opinion can be about an object, product, service, event, organization, person, etc. The polarity of opinions is usually classified into three classes: positive, negative, or neutral [3], [11], [12], [32]. According to Yadollahi et al. [32], the opinion mining field studies the following problems:

- Subjectivity detection
- Opinion polarity classification
- Opinion spam detection
- Opinion summary
- Detection of expressed argument

2) EMOTION MINING

Emotion mining aims to identify the emotions expressed in texts or other modalities [32]. Emotions are thoughts or feelings that can be subjective [3]. People have six primary emotions such as joy, surprise, anger, disgust, sadness, and fear; these emotions can also be divided into secondary and tertiary and even have emotions according to the degree of intensity [3], [5], [6]. According to Yadollahi et al. [32], the emotion mining field addresses the following problems:

- Emotion detection
- Emotion polarity detection
- Emotion classification
- Emotion cause detection

C. LEVELS OF SENTIMENT ANALYSIS

The analysis of sentiments or opinions can be performed at different levels, such as: document-level, sentence-level, and aspect-level [12], [22], [30], [33], [34].

1) DOCUMENT-LEVEL

This level allows for determining the polarity of the opinion expressed in entire documents, which can be positive, negative or neutral. The analyzed entity is the whole content of the text. It does not allow for comparison or evaluate the polarity of multiple entities.

2) SENTENCE-LEVEL

At this level the aim is to determine the polarity of the sentiment expressed in the sentences of a text (which usually are positive, negative, or neutral). Neutral opinion means no opinion because many sentences contain information that do not represent any sentiment. By analyzing the sentences separately instead of the entire document, a better understanding of the feelings expressed in the text can be inferred. This level of study has two main subtasks, which are 1) subjectivity classification, which determines whether the text is subjective or objective 2) sentence-level sentiment classification, if it is subjective, it determines the positive, negative, or neutral polarity [35].

3) ASPECT-LEVEL

The objective of this level is to determine the sentiment or polarity about aspects contained in a text. A text can have several aspects about which feelings are expressed. For example, in the following text: "the new iPhone 5's screen size is amazing, but its battery life is short", we can see that there are opinions expressed on two aspects, for the iPhone 5 screen it is a positive opinion, but for the battery it is a negative opinion [29]. Therefore, this level has two main subtasks: 1) Aspect-extraction, which aims to identify the aspects mentioned in the text, and 2) Aspect-level Sentiment Analysis (ALSA), which aims to determine the polarity of the sentiment associated with the aspect [33], [36].

III. REVIEW METHODOLOGY

We performed a systematic review (SR) of the scientific literature to identify relevant research works related to Sentiment Analysis. The results of this review led us to propose a Framework to Study Sentiment Analysis methods, described later. We followed the review workflow shown in Figure 1, based on the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) guidelines [37].

TABLE 1. The SR research questions and contributions.

ID	Research questions	Goal
RQ01	What machine learning models that do not involve neural networks are frequently used for sentiment classification?	To identify non-neural ML models and word embedding techniques for Sentiment Analysis tasks.
RQ02	What deep learning and word embedding models are frequently used for sentiment classification?	To identify DL models and word embedding techniques based on neural networks for Sentiment Analysis Tasks.
RQ03	What hybrid deep learning and word embedding models are frequently used for sentiment classification?	To identify hybrid models and word embedding techniques that combine classical and DL techniques for Sentiment Analysis Tasks.
RQ04	What graph-based models are used to extract text features?	To identify techniques based on graphs to encode words and texts.

TABLE 2. Keywords used in the search strings.

Primary concept	Keywords	Description
Sentiment analysis classification	sentiment analysis, opinion mining, sentiment classification & sentiment polarity	These terms allow us to identify research related to sentiment analysis for polarity classification.
Classification algorithms	machine learning	These terms allow for identifying Machine Learning models without neural networks for sentiment classification.
Classification models with neural network	deep learning, neural network & hybrid	These terms help identify Deep Learning and Deep Hybrid Learning models for sentiment classification.
Classification models for feature extraction with graph or node embedding	node embedding & graph embedding	These terms help identify studies that are used for feature extraction with word embedding related to graph.

A. QUESTIONS AND GOALS

The present review aims to find studies based on ML, DL and word embedding models used or proposed for sentiment classification tasks. For this review we pose four research questions (RQ), presented in Table 1.

B. REVIEW PROTOCOL

The search of articles was carried out using Web of Science (WoS) database. This indexes high quality journals and conferences. We follow the PRISMA guidelines to perform the review. This involved four steps: identification, screening, eligibility and Inclusion steps (Figure 2).

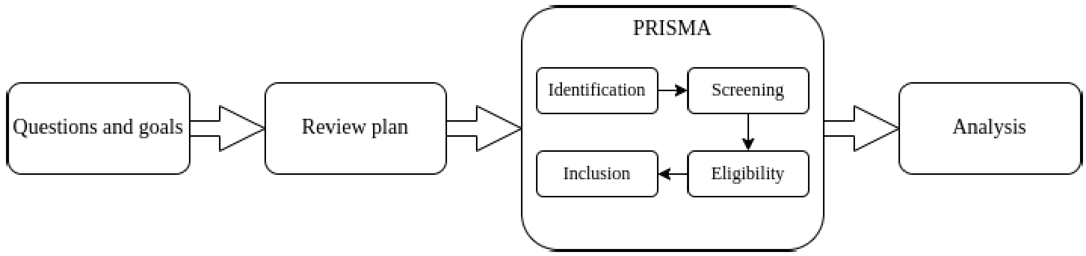

FIGURE 1. Workflow for systematic review.

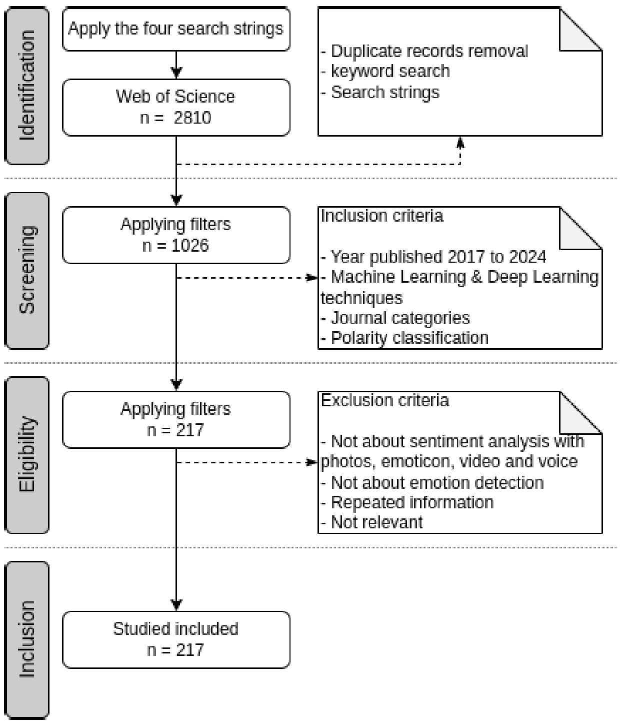

FIGURE 2. PRISMA methodology.

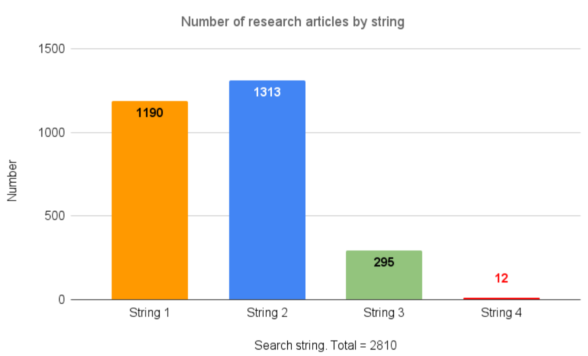

FIGURE 3. Number of research articles found.

For the Identification step we employed four search strings (Table 3) in the WoS database, one for each research question. Table 2 presents the keywords used. In this stage we identified a total of 2,810 articles. Figure 3 shows the distribution of the articles for each search string without using filters.

TABLE 3. Search strings.

Order	Search string	Context description
String 1	TS=("sentiment analysis" OR "opinion mining" OR "sentiment classification" OR "sentiment polarity") AND TS=("machine learning") AND AK=("machine learning") NOT AK = ("neural network*" or "deep learning")	Articles on non-neural Machine Learning models
String 2	TS=("sentiment analysis" OR "opinion mining" OR "sentiment classification" OR "sentiment polarity") AND TS=("deep learning" OR "neural network") AND AK=("deep learning")	Articles on Deep Learning models
String 3	TS=("sentiment analysis" OR "opinion mining" OR "sentiment classification" OR "sentiment polarity") AND TS=("hybrid" AND "deep learning")	Articles of Deep Learning models that include the term hybrid.
String 4	TS=("sentiment analysis" OR "opinion mining" OR "sentiment classification" OR "sentiment polarity") AND TS=("node embedding" OR "graph embedding")	Articles for feature extraction based on graphs or node embeddings.

In the Screening step we applied the inclusion and exclusion criteria of Table 4. These criteria are related to relevance, recentness (year of publication) and knowledge area (computer science, technology or linguistics). With this filtering, it was reduced to 1026 articles.

In the eligibility phase we read the summary of the articles to identify those that are truly relevant to the present study and that allow us to meet our review objectives. We identified a

TABLE 4. Inclusion & exclusion criteria.

Order	Inclusion	Exclusion
String 1	Categories: computer science artificial intelligence, linguistics, language linguistics, and computer science theory methods. Year published: 2024, 2023, 2022, 2021, 2020, 2019, 2018, and 2017. Machine learning techniques	Publications before January 2017 . Publications that do not relate to areas of technology, computer science, or linguistics. Deep learning and neural network.
String 2	Categories: computer science artificial intelligence, linguistics, language linguistics, and computer science theory methods. Year published: 2024, 2023, 2022, 2021, and 2020. Deep learning techniques and neural network.	Publications before January 2020. Publications that don't relate to areas of technology, computer science, or linguistics.
String 3	Categories: computer science artificial intelligence, computer science theory methods, Engineering Electrical Electronic, Multidisciplinary Sciences, and Computer science Information Systems. Year published: 2024, 2023, 2022, 2021, and 2020. Deep learning techniques with hybrid term.	Publications before January 2020. Publications that don't relate to areas of technology, computer science, or linguistics.
String 4	Year published: 2024, 2023, 2022, 2021, and 2020, 2019, 2018. Deep learning, machine learning, graph or node embedding.	Not relevant, and not used word embedding with graph o node knowledge.

total of 217 relevant articles. Figure 4 shows the distribution of the selected articles along the search strings. String 2 resulted in 48.8% of the total relevant articles, followed by String 1 with 30.0%. This suggests that most recent articles on sentiments analysis are based on deep learning models. In addition, the articles we found also demonstrate the use of pre-trained word embedding models to enhance sentiment polarity classification.

IV. RESULTS

This section elaborates on our findings and framework proposal using the 217 studies found in section III. Below, we outline the seven phases of the proposal (see Figure 5) which are based on the results of the systematic review and address the research questions.

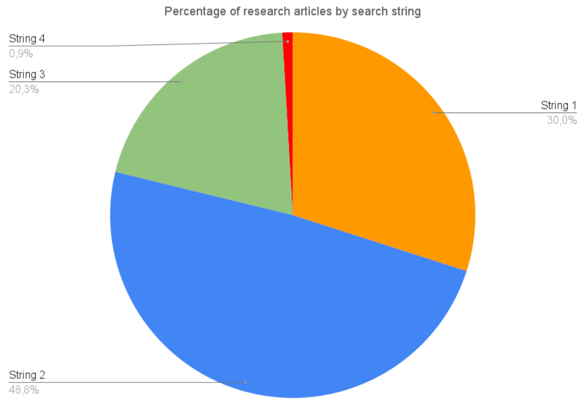

FIGURE 4. The percentage of research articles by the search string.

TABLE 5. Most used dataset for sentiment analysis.

Dataset	Polarity	Registers	Number of studies
Tweets extraction [a]	-	-	66
IMDB movie review	positive and negative	50k	32
Amazon product review	positive and negative	568k	28
Yelp sentiment	positve and negative	560k	11
Sentiment140	positve and negative	1.6M	17
Twitter US Airline Sentiment	positive, neutral and negative	14.6k	16
SemEval 2014	positive, neutral and negative	6k	17
SemEval 2015	positive, neutral and negative	3k	8
SemEval 2016	positive, neutral and negative	4.5k	12
SemEval 2017	positive, neutral and negative	50k	7

[a] Tweets from Twitter are extracted for different domains, and researchers determine the polarity labels. Likewise, other datasets are extracted from blogs, YouTube, Weibo, and Facebook, but they are not used frequently.

A. DATA COLLECTION

In the initial phase of the Framework's data collection, several datasets were identified for sentiment classification, as detailed in Table 5. The majority of research articles utilized datasets from tweet extraction, along with IMDB and Amazon datasets, to determine positive and negative polarity. Figure 6 shows the representativeness of languages in the datasets used in various sentiment classification models. It is evident that English is the most representative language compared to other languages for training and validating different models.

B. DATA CLEANING

During this phase, various tasks are carried out to clean the data. However, the specific cleaning tasks to be used depend on the type of analysis and the dataset used for model development. We identified 16 common cleaning techniques during the review, listed in Table 6.

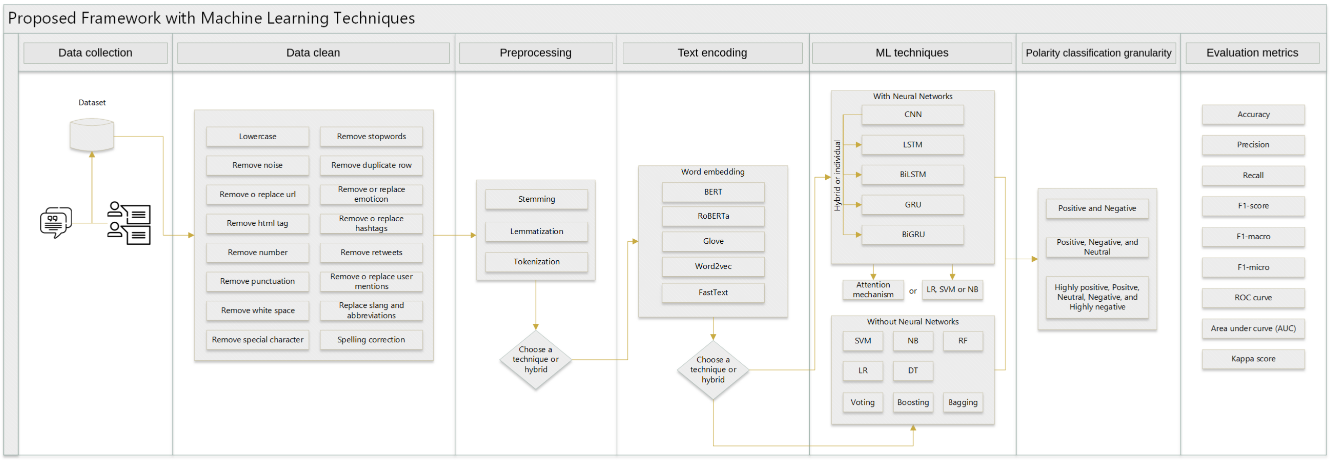

FIGURE 5. Framework proposal for sentiment analysis.

TABLE 6. Data clean techniques for sentiment classification.

S. No.	Technique
1	Lowercase
2	Remove unicode string and noise
3	Remove o replace url
4	Remove html tag
5	Remove number
6	Remove punctuation
7	Remove white space
8	Remove special character
9	Remove stopwords
10	Remove duplicated row
11	Remove emoticon
12	Remove o replace hashtags
13	Remove retweets
14	Remove o replace user mentions
15	Replace slang and abbreviations
16	Spelling correction

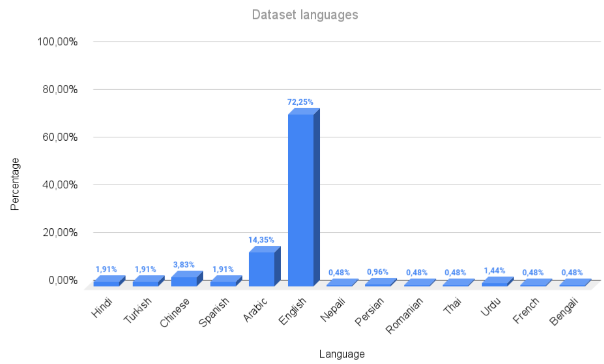

FIGURE 6. Representativeness of languages in common datasets for sentiment analysis.

TABLE 7. Preprocessing techniques.

Technique	Number	Research reference
Tokenization	127	[38], [38]–[97], [97]–[162]
Lemmatization	32	[41], [45], [49], [54], [67], [79], [80], [88], [92], [96]–[99], [105], [108], [110], [111], [129], [133], [135]–[137], [139], [142], [155], [159], [163]–[168]
Stemming	65	[43], [45], [46], [48], [49], [52], [57], [59]–[62], [66], [68], [70], [73], [79], [80], [82], [87], [89]–[92], [100], [101], [104], [109]–[111], [113]–[116], [118], [124], [125], [127], [128], [130], [133], [135], [137], [139], [142], [146]–[149], [152], [155]–[158], [164], [166], [169]–[175]

C. PREPROCESSING

This phase aims to prepare the data for efficient feature extraction. We identified three frequent pre-procesing techniques: stemming, lemmatization, and tokenization (see Table 7). The choice of the technique depend on the objectives defined for developing the model.

D. TEXT ENCODING

Text encoding, is an important step in sentiment classification. It aims to find numerical representations of words and texts. It can be considered as a form of feature extraction for the subsequent phase of model construction. Table 8 shows the most common text encoding techniques we found in papers proposing non-neural models for sentiment classification. It can be seen that TF-IDF and BoW are the most frequently used text encoding techniques in such approaches. In papers that propose neural-network based models, the text encoding techniques are usually different. We list such techniques in Table 9. It can be seen that Word2vec is the most common text-encoding technique in neural-network based models. It is also commonly used Glove, FastText, and BERT techniques as word embedding techniques. Most authors indicate that such encodings contribute to improve

TABLE 8. Common text encoding techniques found in non-neural models for sentiment classification.

S. No.	Technique	Number	Research reference
1	Term Frequency-Inverse Document Frequency (TF-IDF)	29	[41], [43], [47], [49], [50], [52]–[55], [58], [87], [89]–[92], [108], [109], [111], [113], [114], [116], [165], [166], [170], [176]–[180]
2	Bag of Word (BoW)	19	[38], [43], [44], [46], [53], [58], [92], [109], [114], [115], [164], [166], [170], [176], [177], [181]–[184]
3	N-Gram	11	[39], [40], [49], [52], [55], [56], [165], [179], [183], [185], [186]
4	Word2vec	8	[53], [87], [113], [163], [164], [166], [181], [184]
5	Word vector construction	5	[42], [45], [47], [111], [112]
6	Glove	3	[43], [87], [164]
7	Doc2vec	2	[53], [182]
8	StringToWordVector	1	[177]
9	BERT	1	[110]

the performance of sentiment polarity classifiers due to their rich numerical representations of words, that are obtained by pre-training the encoders with vast amounts of data.

E. MACHINE LEARNING TECHNIQUES

Machine learning techniques are at the heart of the process of building sentiment detection models. In the present review we group these techniques into those that are not based on neural networks and those that are. Table 10 lists the non-neural network-based techniques, with the most popular techniques being NB, SVM, LR, and ensemble mmodels (RF, Bagging, Voting, Stacking). Table 11 shows neural-networks based techniques. The most common neural network architectures are CNN, Long Short Term Memory (LSTM) networks, Gated Recurrent Unit (GRU) networks, and AM networks. These architectures are used independently or hybridized with other neural or non-neural models. The techniques with the greatest performance gains are usually hybrid models.

F. POLARITY CLASSIFICATION GRANULARITY

The target categories in sentiment polarity classification is normally determined in the definition of the model development project, since it defines the data labeling requirements for the training and validation phase. Most polarity classification studies target two clases: positive and negative sentiments. It is also common to find three-polarity classification models: positive, negative, and neutral sentiments. Figure 7 shows the relative proportion of the studies for each classification granularity. We can observe that there are a few works targeting five polarities: higly positive, positive, neutral, negative, and higly negative classes. This small number of studies is probably due to the small number of available datasets that have more than 3 labeled classes.

TABLE 9. Common text encoding techniques found in neural models for sentiment classification.

S. No.	Technique	Number	Research reference
1	Word2Vec	51	[62], [65], [67], [71], [73], [74], [77]–[79], [81], [84], [96], [98], [102], [117], [121]–[123], [128], [130], [131], [136], [139], [140], [143], [146]–[149], [151], [159], [160], [172], [187]–[204]
2	Glove	40	[60], [63], [67], [69], [79], [83], [86], [97], [97], [99], [102], [121]–[123], [129], [132], [133], [137], [140], [148], [151], [168], [187], [190], [193], [195], [196], [198], [202], [203], [205]–[214]
3	BERT	26	[59], [61], [94], [117]–[119], [126], [134], [140]–[143], [145], [155], [157], [162], [188], [200], [205], [215]–[221]
4	FastText	18	[61], [79], [97], [97], [100], [121]–[123], [129], [130], [152], [198], [199], [202], [216], [222]–[224]
5	TF-IDF	16	[66], [67], [97], [97], [98], [101], [107], [136], [143], [149], [151], [171], [174], [175], [190], [208]
6	Embedding function or hybrid	10	[120], [124], [125], [127], [135], [144], [150], [153], [156], [161]
7	RoBERTa	5	[85], [103], [138], [157], [158]
8	BoW	5	[66], [171], [174], [175], [195]
9	DistilBERT	3	[155], [157], [158]
10	N-gram	3	[172], [195], [225]
11	Graph Construction	3	[162], [226], [227]
12	Word2Sent	1	[196]
13	Sent2seq	1	[70]
14	ALBERT	1	[154]

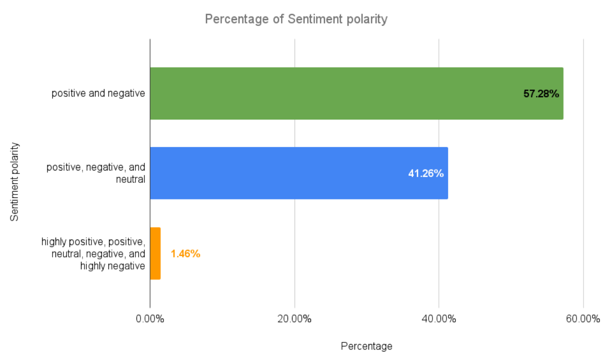

FIGURE 7. Percentage of sentiment polarity most used for sentiment classification.

G. EVALUATION METRICS

Table 12 shows the evaluation metrics we found in the reviewed studies. The most commonly used metrics are

TABLE 10. Machine learning techniques not based on neural networks for sentiment classifier induction.

S. No.	Technique	Number	Research reference
1	Support Vector Machine	43	[43], [45]–[51], [53]–[56], [58], [64], [87], [89]–[92], [108]–[116], [164]–[167], [169], [178]–[181], [183], [184], [228]–[231]
2	Naive Bayes	41	[40]–[42], [44]–[47], [49]–[51], [53]–[58], [64], [87], [90], [91], [110], [111], [115], [116], [164]–[166], [169], [170], [176]–[180], [183], [185], [186], [229]–[232]
3	Random Forest	30	[38], [38], [40], [41], [43], [47]–[49], [52], [53], [57], [87], [89]–[92], [108], [110]–[114], [116], [116], [164], [182], [185], [186], [229], [231]
4	Logistic Regression	28	[38], [38], [40], [41], [43], [44], [46], [47], [49], [50], [55]–[57], [87], [90], [108]–[112], [114], [115], [164], [166], [177], [178], [181], [229]
5	Decision Tree	20	[38], [41], [47], [49], [51], [53], [64], [87], [89], [108], [109], [111], [115], [116], [164], [178], [180], [182], [228], [231]
6	K-nearest Neighbors	15	[40], [45]–[49], [53], [109], [111], [113], [115], [116], [164], [228], [229]
7	Boosting (Adaboost, Gradient Boosting, XGBoost, LogitBoost)	12	[43], [47], [49], [51], [55], [87], [92], [110], [111], [164], [164], [182]
8	Stochastic Gradient Descent	6	[41], [43], [55], [108], [163], [164]
9	Voting	3	[115], [183], [185]
10	Bagging	2	[52], [183]
11	Sequential Minimal Optimization (SMO)	2	[57], [228]
12	Stacking	1	[45]

TABLE 11. Machine learning techniques based on neural networks for sentiment classifier induction.

S. No.	Technique	Number	Research reference
1	LSTM	48	[60], [62], [64], [66], [68], [70], [85], [97], [98], [100], [102], [103], [118], [119], [121], [124], [126], [128]–[131], [134]–[136], [138], [139], [141], [142], [149], [150], [153], [157], [161], [163], [171], [175], [189], [197], [199], [201], [210], [217], [220], [223], [225], [231], [233], [234]
2	CNN	39	[60], [62], [64], [70], [97], [98], [100], [102], [117], [119], [121], [123], [125]–[129], [134], [135], [138], [141], [143], [144], [150], [157], [159], [161], [163], [166], [171], [189], [196], [201], [203], [204], [223], [224], [231], [235]
3	BiLSTM	28	[61], [62], [70], [93], [97], [99], [100], [103], [117], [119], [120], [123], [125], [129], [134], [139], [143], [152], [153], [157]–[159], [161], [171], [172], [197], [199], [217]
4	CNN & LSTM	25	[62], [64], [73], [77], [81], [84], [95], [97], [127], [128], [133], [141], [144], [161], [171], [174], [189], [191], [193], [195], [208], [213], [217], [231], [236]
5	CNN & BiLSTM	17	[59], [73], [76], [86], [97], [107], [127], [129], [132], [148], [155], [159]–[161], [202], [211], [237]
6	GRU	14	[103], [119], [127], [134], [135], [141], [153], [157], [159], [163], [171], [172], [198], [199]
7	LSTM & AM	9	[67], [75], [125], [138], [173], [194], [214], [238]
8	CNN & LSTM or BiLSTM or GRU or BiGRU & SVM or NB or LR	8	[105], [124], [144], [153], [160], [204], [216], [218]
9	BiGRU	7	[68], [96], [120], [124], [125], [152], [158]
10	CNN & BiLSTM & AM	7	[63], [83], [137], [145], [160], [187], [221]
11	LSTM & CNN & AM	3	[65], [191], [192]
12	CNN & GRU	3	[59], [68], [141]
13	CNN & AM	3	[135], [173], [188]
14	BiLSTM & AM	3	[74], [140], [206]
15	CNN & BiGRU	2	[78], [106]
16	BiGRU & BiLSTM & CNN & AM	2	[69], [152]
17	BiGRU & CNN & AM	2	[122], [209]
18	CNN & BiLSTM & GRU & AM	1	[168]
19	BiLSTM & AM & LR	1	[72]
20	Graph Convolutional Network & AM	1	[205]
21	Knowledge Graph & LSTM or BiLSTM	1	[226]
22	BiLSTM & GRU & AM	1	[212]

Accuracy, Precision, Recall, and F1-score. However, other evaluation metrics have been used for model evaluation, like AUC, Kappaca score, etc. Many studies use more than one metric. This is because each metric evaluates or emphasizes some aspect of the results, so using multiple metrics generates a broader view of the model's capabilities.

V. DISCUSSION

In this study we describe a framework for studying articles that propose sentiment classification models. We discuss here some relevant aspects that we found throughout our review.

1) **Sentiment Analysis:** The primary tasks in Sentiment Analysis involves opinion and emotion mining. In our literature review, the most commonly used levels of Sentiment Analysis are sentence and document. However, a sentence can contain multiple sentiments, so it is important to identify and analyze the sentence's final

TABLE 12. Evaluation metrics most commonly used.

S. No.	Metrics	Number
1	Accuracy	206
2	Precision	159
3	Recall	159
4	F1-score	151
5	F1-macro	11
6	AUC	8
7	Kappa score	4
8	F1-micro	3

classification; in any case, it's very complex to identify polarity because the text could have multiple opinions. Therefore, recent deep learning models focuses on the aspect level analysis (see [34], [119], [147], [187], [191], [207], [211], [215], [239]). In those works, the sentiment of each entity is determined. This approach may be more beneficial than document-level analysis since since it can show how the product or service is appreciated at different aspects. Most reviewed studies are typically evaluated in the SemEval dataset (see Table 5).

2) **Preprocessing:** In the present review we have identified preprocessing techniques that have shown their relevance in recent studies. Tokenization is essential for extracting the best features, while Lemmatization and Stemming techniques contribute to improving classification results. It's important to note that these techniques are not universally implemented in all languages and datasets. It is important to analyze and identify the appropriate library to use based on the dataset characteristics. While some studies use these techniques in the cleaning phase, we have observed that Stemming is more commonly used compared to Lemmatization (see Table 7). We suggest using both techniques in the preprocessing phase to maintain an organized workflow.

3) **Text encoding:** The performance of a sentiment classification model depends depends heavily on the choice of text encoding. For ML models not based on neural networks, the TF-IDF and BoW techniques have been the most popularly used coding techniques (see Table 8). However, recent studies have shown that Word2Vec and Glove techniques have demonstrated promising results in non neural-networks models. On the other hand, in recent years there has been an explosion of methods based on artificial neural networks. In the field of sentiment classification, these methods have frequently shown superior predictive performances than traditional ML methods. This is probably due to their non-linear characteristics and ability to digest large amounts of data through the fitting of millions of parameters. It is common that recent sentiment classification models also use neural network-based text encoders in their pipeline, like Glove, Word2Vec, FastText, and BERT (Table 9). This demonstrates that learning good text encodings is the key to improving sentiment identification. They are some interesting studies that use knowledge graphs in the process of text encoding [162], [226] showing promising results. Probably, the incorporation of semantic knowledge of words through knowledge graphs helps to have better numerical representations of texts.

4) **Machine and Deep Learning Models:** Recent studies have proposed several hybrid methods, both neural network-based and non-neural network-based. In Table 10 we list non-neural network-based. It is evident that NB, SVM, RF, LR, and DT are the most commonly used. However, studies utilizing hybrid models such as Bagging, Voting, and Stacking have shown promising results for sentiment classification tasks in some cases. On the other hand, models based on neural networks have demonstrated better results than traditional algorithms like NB, SVM, and RF. Table 11 shows the use of Recurrent Neural Networks (RNN) models such as LSTM, CNN, Bidirectional Long short-term Memory (BiLSTM), GRU, and Bidirectional Gated Recurrent Unit (BiGRU) in numerous applications for Sentiment Analysis due to their ability to identify complex patterns within text more effectively. Furthermore, recent studies have shown that hybrid methods using CNN, RNN, AM, and traditional ML algorithms have produced better results than standalone models.

VI. CHALLENGES AND OPPORTUNITIES

Sentiment Analysis has recently made significant progress in psychology, politics, marketing, economics, and education. From the reviewed studies, we identified several novel ideas. However, there are some challenges and gaps that need to be addressed in future researches. Below, we summarize some challenges that could that could help the progress of the Sentiment Analysis field.

A. REDUCTION OF COMPUTATIONAL COST

Most models require high memory capacities for training and inference. Some techniques, like model distillation and model quantization can help reducing the number of trainable parameters, which in turn would help to accelerate the model training and improve model performance.

B. NEW DATASETS

In our review, we found that most studies use a few datasets to train and evaluate their proposals (see Table 5). In 66 studies we found that the dataset are opinions extracted from Twitter social network; 32 studies use IMDB movie reviews dataset, 28 studies use Amazon reviews dataset, and 17 studies use Sentiment140 dataset. This shows the need to collect new datasets with more records, more diversity of fields, more polarity/sentiment classes and more languages. There is also a need to build datasets oriented toward specific domains that are not represented in existing datasets.

C. ASPECT-LEVEL SENTIMENT CLASSIFICATION

Most research has focused on the sentence and document level due to the availability of datasets such as IMDB, Amazon reviews, Twitter US airline, Sentiment140, and Yelp Sentiment dataset. However, the aspect level is a relatively new and understudied area. For instance, Table 5 shows that SemEval is the only dataset related to aspect labels. Therefore, there is a need for new studies using datasets and models related to aspect-based and domain-specific sentiment analysis.

D. MODELS BASED ON CONCEPT UNDERSTANDING

Sophisticated text encoders, such as Word2Vec (see Table 9) can help develop better sentiment classification models. These techniques consider the semantics, syntax, morphology, and context of the text. However, in order to better comprehend the entire meaning of a verbal language sentence, it is necessary to address the conceptual level. This can allow us to handle words that may not explicitly convey an emotion but are associated with concepts that do.

Deep learning models with CNN, RNN, and AM have shown promising results for Sentiment Analysis. However, there is a need to develop novel architectures with more advanced linguistic, lexical, semantic, and morphological features and concept understanding capabilities. Some promising studies in this line are those that try to incorporate knowledge graphs [162], [226]. Using word concepts is a promising idea because a word may have significant conceptual relations with others through edges, with the nodes being words.

E. NEW DEEP LEARNING ARCHITECTURES

DL models have demonstrated promising results for Sentiment Analysis. However, these models require assistance in comprehending the symbols and words within the text. There is a necessity to design architectures based on deep neural networks or apply graph knowledge with linguistic, lexical, semantic, and morphological features or knowledge in order to fully understand the text.

VII. CONCLUSION

This article identified more than a thousand scientific articles on machine learning (ML) models for sentiment classification. After thoroughly reviewing and analyzing the articles, 217 studies were selected for this review based on their relevance and contribution to the field. The paper proposes a framework to study and disecting Sentiment Analysis methods. The framework comprises the following aspects: data collection, data cleaning, pre-processing, text-encoding, ML techniques, polarity classification, and evaluation metrics. In each aspect, we consider algorithms and techniques that can be used according to the objective of the study. Recent studies have demonstrated that methods based on neural-network word embeddings and deep learning classification methods yield state of the art results. This review can help new researchers in this area to have a panoramic view of the area, as well as to identify challenges, current gaps and new research opportunities.

ACKNOWLEDGMENT

The authors would like to express their gratitude to the Artificial Intelligence Laboratory, Pontifical Catholic University of Peru, for providing them with access to their facilities and servers.

REFERENCES

[1] L. Zhang, S. Wang, and B. Liu, "Deep learning for sentiment analysis: A survey," *WIREs Data Mining Knowl. Discovery*, vol. 8, no. 4, p. e1253, Mar. 2018.

[2] T. Nasukawa and J. Yi, "Sentiment analysis: Capturing favorability using natural language processing," in *Proc. 2nd Int. Conf. Knowl. Capture*, New York, NY, USA, 2003, pp. 70–77.

[3] B. Liu, *Sentiment Analysis and Opinion Mining*, 1st ed. USA: Morgan & Claypool, 2012, doi: 10.1007/978-3-031-02145-9.

[4] R. Tokuhisa, K. Inui, and Y. Matsumoto, "Emotion classification using massive examples extracted from the web," in *Proc. 22nd Int. Conf. Comput. Linguistics (COLING)*, vol. 1, Manchester, U.K., 2008, pp. 881–888.

[5] P. Shaver, J. Schwartz, D. Kirson, and C. O'Connor, "Emotion knowledge: Further exploration of a prototype approach," *J. Personality Social Psychol.*, vol. 52, pp. 1061–1086, Jun. 1987.

[6] W. G. Parrott, *Emotions in Social Psychology: Essential Readings*. Philadelphia, PA, USA: Psychology Press, 2001.

[7] Z. S. Harris, "Distributional structure," *WORD*, vol. 10, nos. 2–3, pp. 146–162, Aug. 1954.

[8] K. S. Jones, "A statistical interpretation of term specificity and its application in retrieval," *J. Documentation*, vol. 28, pp. 11–21, Jan. 1972.

[9] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, Jul. 1948.

[10] A. A. Markov, "An example of statistical investigation of the text Eugene Onegin concerning the connection of samples in chains," *Sci. Context*, vol. 19, no. 4, pp. 591–600, Dec. 2006.

[11] P. K. Jain, R. Pamula, and G. Srivastava, "A systematic literature review on machine learning applications for consumer sentiment analysis using online reviews," *Comput. Sci. Rev.*, vol. 41, Aug. 2021, Art. no. 100413.

[12] K. Cortis and B. Davis, "Over a decade of social opinion mining: A systematic review," *Artif. Intell. Rev.*, vol. 54, no. 7, pp. 4873–4965, Oct. 2021.

[13] F. Hemmatian and M. K. Sohrabi, "A survey on classification techniques for opinion mining and sentiment analysis," *Artif. Intell. Rev.*, vol. 52, no. 3, pp. 1495–1545, Oct. 2019.

[14] S. Osorio Angel, A. Peña Pérez Negrón, and A. Espinoza-Valdez, "Systematic literature review of sentiment analysis in the Spanish language," *Data Technol. Appl.*, vol. 55, no. 4, pp. 461–479, Aug. 2021.

[15] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," 2013, *arXiv:1301.3781*.

[16] J. Pennington, R. Socher, and C. Manning, "Glove: Global vectors for word representation," in *Proc. Conf. Empirical Methods Natural Lang. Process. (EMNLP)*, Doha, Qatar, 2014, pp. 1532–1543.

[17] P. Bojanowski, E. Grave, A. Joulin, and T. Mikolov, "Enriching word vectors with subword information," *Trans. Assoc. Comput. Linguistics*, vol. 5, pp. 135–146, Jun. 2017. [Online]. Available: https://direct.mit.edu/tacl/article-pdf/doi/10.1162/tacl_a_00051/1567442/tacl_a_00051.pdf

[18] J. Elman, "Finding structure in time," *Cognit. Sci.*, vol. 14, no. 2, pp. 179–211, Jun. 1990.

[19] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.

[20] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 1–11.

[21] S. Minaee, N. Kalchbrenner, E. Cambria, N. Nikzad, M. Chenaghlu, and J. Gao, "Deep learning–based text classification: A comprehensive review," *ACM Comput. Surv.*, vol. 54, p. 62, Apr. 2021.

[22] A. Yadav and D. K. Vishwakarma, "Sentiment analysis using deep learning architectures: A review," *Artif. Intell. Rev.*, vol. 53, no. 6, pp. 4335–4385, Aug. 2020.

[23] Z. Kastrati, F. Dalipi, A. S. Imran, K. Pireva Nuci, and M. A. Wani, "Sentiment analysis of Students' feedback with NLP and deep learning: A systematic mapping study," *Appl. Sci.*, vol. 11, no. 9, p. 3986, Apr. 2021.

[24] A. K. Zarandi and S. Mirzaei, "A survey of aspect-based sentiment analysis classification with a focus on graph neural network methods," *Multimedia Tools Appl.*, vol. 83, no. 19, pp. 56619–56695, Dec. 2023.

[25] P. Nandwani and R. Verma, "A review on sentiment analysis and emotion detection from text," *Social Netw. Anal. Mining*, vol. 11, no. 1, p. 81, Aug. 2021.

[26] D. Tiwari, B. Nagpal, B. S. Bhati, A. Mishra, and M. Kumar, "A systematic review of social network sentiment analysis with comparative study of ensemble-based techniques," *Artif. Intell. Rev.*, vol. 56, no. 11, pp. 13407–13461, Nov. 2023.

[27] A. Ullah, S. N. Khan, and N. M. Nawi, "Review on sentiment analysis for text classification techniques from 2010 to 2021," *Multimedia Tools Appl.*, vol. 82, no. 6, pp. 8137–8193, Mar. 2023.

[28] B. Liu, *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*. Cambridge, U.K.: Cambridge Univ. Press, 2015.

[29] E. Cambria, B. Schuller, Y. Xia, and C. Havasi, "New avenues in opinion mining and sentiment analysis," *IEEE Intell. Syst.*, vol. 28, no. 2, pp. 15–21, Mar. 2013.

[30] M. Birjali, M. Kasri, and A. Beni-Hssane, "A comprehensive survey on sentiment analysis: Approaches, challenges and trends," *Knowl.-Based Syst.*, vol. 226, Aug. 2021, Art. no. 107134.

[31] Y. Wang, W. Song, W. Tao, A. Liotta, D. Yang, X. Li, S. Gao, Y. Sun, W. Ge, W. Zhang, and W. Zhang, "A systematic review on affective computing: Emotion models, databases, and recent advances," *Inf. Fusion*, vols. 83–84, pp. 19–52, Jul. 2022.

[32] A. Yadollahi, A. G. Shahraki, and O. R. Zaiane, "Current state of text sentiment analysis from opinion to emotion mining," *ACM Comput. Surv.*, vol. 50, no. 2, pp. 1–33, May 2017.

[33] B. Liu, *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions* (Studies in Natural Language Processing), 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2020.

[34] K. W. Trisna and H. J. Jie, "Deep learning approach for aspect-based sentiment classification: A comparative review," *Appl. Artif. Intell.*, vol. 36, no. 1, Dec. 2022, Art. no. 2014186.

[35] B. Liu and L. Zhang, *A Survey of Opinion Mining and Sentiment Analysis*. Boston, MA, USA: Springer, 2012, pp. 415–463.

[36] S. Poria, E. Cambria, G. Winterstein, and G.-B. Huang, "Sentic patterns: Dependency-based rules for concept-level sentiment analysis," *Knowl.-Based Syst.*, vol. 69, pp. 45–63, Oct. 2014.

[37] M. J. Page et al., "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews," *BMJ*, vol. 372, pp. 790–799, Mar. 2021.

[38] T. Nijhawan, G. Attigeri, and T. Ananthakrishna, "Stress detection using natural language processing and machine learning over social interactions," *J. Big Data*, vol. 9, no. 1, p. 33, Mar. 2022.

[39] A. Chinnalagu and A. K. Durairaj, "Context-based sentiment analysis on customer reviews using machine learning linear models," *PeerJ Comput. Sci.*, vol. 7, p. e813, Dec. 2021.

[40] A. Khattak, M. Z. Asghar, Z. Ishaq, W. H. Bangyal, and I. A. Hameed, "Enhanced concept-level sentiment analysis system with expanded ontological relations for efficient classification of user reviews," *Egyptian Informat. J.*, vol. 22, no. 4, pp. 455–471, Dec. 2021.

[41] A. H. Hussein, I. F. Moawad, and R. M. Badry, "Arabic sentiment analysis for multi-dialect text using machine learning techniques," *Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 12, pp. 693–700, 2021.

[42] H. Ali, H. Farman, H. Yar, Z. Khan, S. Habib, and A. Ammar, "Deep learning-based election results prediction using Twitter activity," *Soft Comput.*, vol. 26, no. 16, pp. 7535–7543, Aug. 2022.

[43] B. Gaye, D. Zhang, and A. Wulamu, "Sentiment classification for employees reviews using regression vector- stochastic gradient descent classifier (RV-SGDC)," *PeerJ Comput. Sci.*, vol. 7, p. e712, Sep. 2021.

[44] S. Chirgaiya, D. Sukheja, N. Shrivastava, and R. Rawat, "Analysis of sentiment based movie reviews using machine learning techniques," *J. Intell. Fuzzy Syst.*, vol. 41, no. 5, pp. 5449–5456, Nov. 2021.

[45] M. Azam, T. Ahmed, R. Ahmad, A. Ur Rehman, F. Sabah, and R. Muhammad Asif, "A two-step approach for improving sentiment classification accuracy," *Intell. Autom. Soft Comput.*, vol. 30, no. 3, pp. 853–867, 2021.

[46] M. M. Ali, "Arabic sentiment analysis about online learning to mitigate COVID-19," *J. Intell. Syst.*, vol. 30, no. 1, pp. 524–540, Jan. 2021.

[47] U. Younis, M. Z. Asghar, A. Khan, A. Khan, J. Iqbal, and N. Jillani, "Applying machine learning techniques for performing comparative opinion mining," *Open Comput. Sci.*, vol. 10, no. 1, pp. 461–477, Dec. 2020.

[48] Z. M. Zain, A. A. Alhajji, N. S. Alotaibi, N. B. Almutairi, A. D. Aldawas, M. M. Almazrua, and A. M. Alhammad, "Ranking beauty clinics in Riyadh using lexicon-based sentiment analysis and multiattribute-utility theory," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 10, pp. 1–10, 2020.

[49] S. Mihi, B. Ait, I. El, S. Arezki, and N. Laachfoubi, "MSTD: Moroccan sentiment Twitter dataset," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 10, pp. 363–372, 2020.

[50] F. M. Plaza-Del-Arco, M. T. Martín-Valdivia, L. A. Ureña-López, and R. Mitkov, "Improved emotion recognition in Spanish social media through incorporation of lexical knowledge," *Future Gener. Comput. Syst.*, vol. 110, pp. 1000–1008, Sep. 2020.

[51] M. Y. Khan and K. N. Junejo, "Exerting 2D-space of sentiment Lexicons with machine learning techniques: A hybrid approach for sentiment analysis," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 6, pp. 1–10, 2020.

[52] L. B. Shyamasundar and P. Jhansi Rani, "A multiple-layer machine learning architecture for improved accuracy in sentiment analysis," *Comput. J.*, vol. 63, no. 1, pp. 395–409, Jan. 2020.

[53] M. Bilgin and H. Köktas, "Sentiment analysis with term weighting and word vectors," *Int. Arab J. Inf. Technol.*, vol. 16, pp. 953–959, Jan. 2019.

[54] M. B. Alvi, N. A. Mahoto, M. A. Unar, and M. A. Shaikh, "An effective framework for tweet level sentiment classification using recursive text pre-processing approach," *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 6, p. 572, 2019.

[55] D. Gamal, M. Alfonse, E.-S. M. El-Horbaty, and A.-B. M. Salem, "Analysis of machine learning algorithms for opinion mining in different domains," *Mach. Learn. Knowl. Extraction*, vol. 1, no. 1, pp. 224–234, Dec. 2018.

[56] M. Ghosh and G. Sanyal, "Analysing sentiments based on multi feature combination with supervised learning," *Int. J. Data Mining, Model. Manage.*, vol. 11, no. 4, pp. 391–416, 2019.

[57] M. Ghosh and G. Sanyal, "An ensemble approach to stabilize the features for multi-domain sentiment analysis using supervised machine learning," *J. Big Data*, vol. 5, no. 1, p. 44, Nov. 2018.

[58] S. Alowaidi, M. Saleh, and O. Abulnaja, "Semantic sentiment analysis of Arabic texts," *Int. J. Adv. Comput. Sci. Appl.*, vol. 8, no. 2, pp. 1–7, 2017.

[59] A. B. Nassif, A. M. Darya, and A. Elnagar, "Empirical evaluation of shallow and deep learning classifiers for Arabic sentiment analysis," *ACM Trans. Asian Low-Resource Lang. Inf. Process.*, vol. 21, no. 1, pp. 1–25, Nov. 2021.

[60] S. Ranjan and S. Mishra, "Perceiving university students' opinions from Google app reviews," *Concurrency Comput., Pract. Exper.*, vol. 34, no. 10, p. e6800, May 2022.

[61] A. K. Durairaj and A. Chinnalagu, "Transformer based contextual model for sentiment analysis of customer reviews: A fine-tuned BERT," *Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 11, pp. 1–7, 2021.

[62] S. M. Mohamed, E. H. Mohamed, and M. A. Belal, "Polarity detection of dialectal Arabic using deep learning models," *Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 11, pp. 212–218, 2021.

[63] T. E. Trueman and E. Cambria, "A convolutional stacked bidirectional LSTM with a multiplicative attention mechanism for aspect category and sentiment detection," *Cognit. Comput.*, vol. 13, no. 6, pp. 1423–1432, Nov. 2021.

[64] P. K. Jain, V. Saravanan, and R. Pamula, "A hybrid CNN-LSTM: A deep learning approach for consumer sentiment analysis using qualitative user-generated contents," *ACM Trans. Asian Low-Resource Lang. Inf. Process.*, vol. 20, no. 5, pp. 1–15, Jul. 2021.

[65] S. Sagnika, B. S. P. Mishra, and S. K. Meher, "An attention-based CNN-LSTM model for subjectivity detection in opinion-mining," *Neural Comput. Appl.*, vol. 33, no. 24, pp. 17425–17438, Dec. 2021.

[66] N. Jamal, C. Xianqiao, F. Al-Turjman, and F. Ullah, "A deep learning–based approach for emotions classification in big corpus of imbalanced tweets," *ACM Trans. Asian Low-Resource Lang. Inf. Process.*, vol. 20, no. 3, pp. 1–16, Mar. 2021.

[67] A. R. Pathak, M. Pandey, and S. Rautaray, "Topic-level sentiment analysis of social media data using deep learning," *Appl. Soft Comput.*, vol. 108, Sep. 2021, Art. no. 107440.

[68] Ö. Çoban, S. A. Özel, and A. Inan, "Deep learning-based sentiment analysis of Facebook data: The case of Turkish users," *Comput. J.*, vol. 64, no. 1, pp. 473–499, Nov. 2019.

[69] M. E. Basiri, S. Nemati, M. Abdar, E. Cambria, and U. R. Acharya, "ABCDM: An attention-based bidirectional CNN-RNN deep model for sentiment analysis," *Future Gener. Comput. Syst.*, vol. 115, pp. 279–294, Feb. 2021.

[70] H. Elfaik and E. H. Nfaoui, "Deep bidirectional LSTM network learning-based sentiment analysis for Arabic text," *J. Intell. Syst.*, vol. 30, no. 1, pp. 395–412, Dec. 2020.

[71] R. Ranjan and A. K. Daniel, "Proposed hybrid model for sentiment classification using CovNet-DualLSTM techniques," *ADCAIJ: Adv. Distrib. Comput. Artif. Intell. J.*, vol. 10, no. 4, pp. 401–418, Feb. 2022.

[72] G. Singh Chauhan, Y. Kumar Meena, D. Gopalani, and R. Nahta, "A two-step hybrid unsupervised model with attention mechanism for aspect extraction," *Expert Syst. Appl.*, vol. 161, Dec. 2020, Art. no. 113673.

[73] M. Umer, I. Ashraf, A. Mehmood, S. Kumari, S. Ullah, and G. Sang Choi, "Sentiment analysis of tweets using a unified convolutional neural network-long short-term memory network model," *Comput. Intell.*, vol. 37, no. 1, pp. 409–434, Feb. 2021.

[74] M. Atzeni and D. Reforgiato Recupero, "Multi-domain sentiment analysis with mimicked and polarized word embeddings for human–robot interaction," *Future Gener. Comput. Syst.*, vol. 110, pp. 984–999, Sep. 2020.

[75] H.-T. Nguyen and L.-M. Nguyen, "ILWAANet: An interactive lexicon-aware word-aspect attention network for aspect-level sentiment classification on social networking," *Expert Syst. Appl.*, vol. 146, May 2020, Art. no. 113065.

[76] S.-T. Park, D.-Y. Kim, and G. Li, "An analysis of environmental big data through the establishment of emotional classification system model based on machine learning: Focus on multimedia contents for portal applications," *Multimedia Tools Appl.*, vol. 80, nos. 26–27, pp. 34459–34477, Nov. 2021.

[77] R. Piryani, B. Piryani, V. K. Singh, and D. Pinto, "Sentiment analysis in nepali: Exploring machine learning and lexicon-based approaches," *J. Intell. Fuzzy Syst.*, vol. 39, no. 2, pp. 2201–2212, Aug. 2020.

[78] S. Soubraylu and R. Rajalakshmi, "Hybrid convolutional bidirectional recurrent neural network based sentiment analysis on movie reviews," *Comput. Intell.*, vol. 37, no. 2, pp. 735–757, May 2021.

[79] F. Es-Sabery, A. Hair, J. Qadir, B. Sainz-De-Abajo, B. García-Zapirain, and I. D. L. Torre-Díez, "Sentence-level classification using parallel fuzzy deep learning classifier," *IEEE Access*, vol. 9, pp. 17943–17985, 2021.

[80] M. Divyapushpalakshmi and R. Ramalakshmi, "An efficient sentimental analysis using hybrid deep learning and optimization technique for Twitter using parts of speech (POS) tagging," *Int. J. Speech Technol.*, vol. 24, no. 2, pp. 329–339, Jun. 2021.

[81] R. K. Behera, M. Jena, S. K. Rath, and S. Misra, "Co-LSTM: Convolutional LSTM model for sentiment analysis in social big data," *Inf. Process. Manage.*, vol. 58, no. 1, Jan. 2021, Art. no. 102435.

[82] S. H. Janjua, G. F. Siddiqui, M. A. Sindhu, and U. Rashid, "Multi-level aspect based sentiment classification of Twitter data: Using hybrid approach in deep learning," *PeerJ Comput. Sci.*, vol. 7, p. e433, Apr. 2021.

[83] H. S. Sharaf Al-Deen, Z. Zeng, R. Al-Sabri, and A. Hekmat, "An improved model for analyzing textual sentiment based on a deep neural network using multi-head attention mechanism," *Appl. Syst. Innov.*, vol. 4, no. 4, p. 85, Oct. 2021.

[84] C. Singla, F. N. Al-Wesabi, Y. Singh Pathania, B. Sulaiman Alfurhood, A. Mustafa Hilal, M. Rizwanullah, M. Ahmed Hamza, and M. Mahzari, "An optimized deep learning model for emotion classification in tweets," *Comput., Mater. Continua*, vol. 70, no. 3, pp. 6365–6380, 2022.

[85] K. L. Tan, C. P. Lee, K. S. M. Anbananthen, and K. M. Lim, "RoBERTa-LSTM: A hybrid model for sentiment analysis with transformer and recurrent neural network," *IEEE Access*, vol. 10, pp. 21517–21525, 2022.

[86] R. A. Saleh, M. Driss, and I. Almomani, "CBiLSTM: A hybrid deep learning model for efficient reputation assessment of cloud services," *IEEE Access*, vol. 10, pp. 35321–35335, 2022.

[87] Y. Didi, A. Walha, and A. Wali, "COVID-19 tweets classification based on a hybrid word embedding method," *Big Data Cognit. Comput.*, vol. 6, no. 2, p. 58, May 2022.

[88] O. Iparraguirre-Villanueva, V. Guevara-Ponce, F. Sierra-Linan, S. Beltozar-Clemente, and M. Cabanillas-Carbonell, "Sentiment analysis of tweets using unsupervised learning techniques and the k-means algorithm," *Int. J. Adv. Comput. Sci. Appl.*, vol. 13, no. 6, pp. 571–578, 2022.

[89] I. Steinke, J. Wier, L. Simon, and R. Seetan, "Sentiment analysis of online movie reviews using machine learning," *Int. J. Adv. Comput. Sci. Appl.*, vol. 13, no. 9, pp. 1–7, 2022.

[90] M. Purba, E. Ermatita, A. Abdiansah, H. Noprisson, V. Ayumi, H. Setiawan, U. Salamah, and Y. Yadi, "Effect of random splitting and cross validation for Indonesian opinion mining using machine learning approach," *Int. J. Adv. Comput. Sci. Appl.*, vol. 13, no. 9, pp. 1–7, 2022.

[91] A. Alsanad, "An improved Arabic sentiment analysis approach using optimized multinomial Naïve Bayes classifier," *Int. J. Adv. Comput. Sci. Appl.*, vol. 13, no. 8, pp. 1–10, 2022.

[92] W. Aljedaani, F. Rustam, M. W. Mkaouer, A. Ghallab, V. Rupapara, P. B. Washington, E. Lee, and I. Ashraf, "Sentiment analysis on Twitter data integrating TextBlob and deep learning models: The case of U.S. airline industry," *Knowl.-Based Syst.*, vol. 255, Nov. 2022, Art. no. 109780.

[93] M. Arbane, R. Benlamri, Y. Brik, and A. D. Alahmar, "Social media-based COVID-19 sentiment classification model using bi-LSTM," *Expert Syst. Appl.*, vol. 212, Feb. 2023, Art. no. 118710.

[94] M. Venugopalan and D. Gupta, "A reinforced active learning approach for optimal sampling in aspect term extraction for sentiment analysis," *Expert Syst. Appl.*, vol. 209, Dec. 2022, Art. no. 118228.

[95] V. Jain and K. L. Kashyap, "Ensemble hybrid model for Hindi COVID-19 text classification with Metaheuristic optimization algorithm," *Multimedia Tools Appl.*, vol. 82, no. 11, pp. 16839–16859, Oct. 2022.

[96] R. Dutta, N. Das, M. Majumder, and B. Jana, "Aspect based sentiment analysis using multi-criteria decision-making and deep learning under COVID-19 pandemic in India," *CAAI Trans. Intell. Technol.*, vol. 8, no. 1, pp. 219–234, Mar. 2023.

[97] S. Aslan, "A novel TCNN–Bi-LSTM deep learning model for predicting sentiments of tweets about COVID-19 vaccines," *Concurrency Comput., Pract. Exper.*, vol. 34, no. 28, p. e7387, Dec. 2022.

[98] D. C. Neagu, A. B. Rus, M. Grec, M. A. Boroianu, N. Bogdan, and A. Gal, "Towards sentiment analysis for Romanian Twitter content," *Algorithms*, vol. 15, no. 10, p. 357, Sep. 2022.

[99] A. Pimpalkar and J. R. Raj R, "MBiLSTMGloVe: Embedding GloVe knowledge into the corpus using multi-layer BiLSTM deep learning model for social media sentiment analysis," *Expert Syst. Appl.*, vol. 203, Oct. 2022, Art. no. 117581.

[100] A. Diwali, K. Dashtipour, K. Saeedi, M. Gogate, E. Cambria, and A. Hussain, "Arabic sentiment analysis using dependency-based rules and deep neural networks," *Appl. Soft Comput.*, vol. 127, Sep. 2022, Art. no. 109377.

[101] R. E. Saudy, A. E. D. M. El-Ghazaly, E. S. Nasr, and M. H. Gheith, "A novel hybrid sentiment analysis classification approach for mobile applications Arabic slang reviews," *Int. J. Adv. Comput. Sci. Appl.*, vol. 13, no. 8, pp. 1–11, 2022.

[102] J.-W. Lin, T. D. Thanh, and R.-G. Chang, "Multi-channel word embeddings for sentiment analysis," *Soft Comput.*, vol. 26, no. 22, pp. 12703–12715, Nov. 2022.

[103] K. L. Tan, C. P. Lee, K. M. Lim, and K. S. M. Anbananthen, "Sentiment analysis with ensemble hybrid deep learning model," *IEEE Access*, vol. 10, pp. 103694–103704, 2022.

[104] V. Kumar, "Spatiotemporal sentiment variation analysis of geotagged COVID-19 tweets from India using a hybrid deep learning model," *Sci. Rep.*, vol. 12, no. 1, p. 1849, Feb. 2022.

[105] A. Londhe and P. P. Rao, "Incremental learning based optimized sentiment classification using hybrid two-stage LSTM-SVM classifier," *Int. J. Adv. Comput. Sci. Appl.*, vol. 13, no. 6, pp. 611–619, 2022.

[106] E. Omara, M. Mousa, and N. Ismail, "Character gated recurrent neural networks for Arabic sentiment analysis," *Sci. Rep.*, vol. 12, no. 1, p. 9779, Jun. 2022.

[107] K. Shanmugavadivel, S. H. Sampath, P. Nandhakumar, P. Mahalingam, M. Subramanian, P. K. Kumaresan, and R. Priyadharshini, "An analysis of machine learning models for sentiment analysis of Tamil code-mixed data," *Comput. Speech Lang.*, vol. 76, Nov. 2022, Art. no. 101407.

[108] M. Karim, M. M. Saad Missen, M. Umer, A. Fida, A. A. Eshmawi, A. Mohamed, and I. Ashraf, "Comprehension of polarity of articles by citation sentiment analysis using TF-IDF and ML classifiers," *PeerJ Comput. Sci.*, vol. 8, p. e1107, Dec. 2022.

[109] N. Aslam, K. Xia, F. Rustam, E. Lee, and I. Ashraf, "Self voting classification model for online meeting app review sentiment analysis and topic modeling," *PeerJ Comput. Sci.*, vol. 8, p. e1141, Dec. 2022.

[110] A. Karanikola, G. Davrazos, C. M. Liapis, and S. Kotsiantis, "Financial sentiment analysis: Classic methods vs. Deep learning models," *Intell. Decis. Technol.*, vol. 17, no. 4, pp. 893–915, Nov. 2023.

[111] S. Ounacer, D. Mhamdi, S. Ardchir, A. Daif, and M. Azzouazi, "Customer sentiment analysis in hotel reviews through natural language processing techniques," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 1, pp. 1–11, 2023.

[112] M. Sha, "A cloud based sentiment analysis through logistic regression in AWS platform," *Comput. Syst. Sci. Eng.*, vol. 45, no. 1, pp. 857–868, 2023.

[113] N. E. Aoumeur, Z. Li, and E. M. Alshari, "Improving the polarity of text through word2vec embedding for primary classical Arabic sentiment analysis," *Neural Process. Lett.*, vol. 55, no. 3, pp. 2249–2264, Jun. 2023.

[114] M. Errami, M. A. Ouassil, R. Rachidi, B. Cherradi, S. Hamida, and A. Raihani, "Sentiment analysis on Moroccan dialect based on ML and social media content detection," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 3, pp. 415–425, 2023.

[115] M. A. Virgananda, I. Budi, and R. S. Ryan, "Purchase intention and sentiment analysis on Twitter related to social commerce," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 7, pp. 1–8, 2023.

[116] D. A. Musleh, I. Alkhwaja, A. Alkhwaja, M. Alghamdi, H. Abahussain, F. Alfawaz, N. Min-Allah, and M. M. Abdulqader, "Arabic sentiment analysis of Youtube comments: NLP-based machine learning approaches for content evaluation," *Big Data Cognit. Comput.*, vol. 7, no. 3, p. 127, Jul. 2023.

[117] C. Yan, J. Liu, W. Liu, and X. Liu, "Sentiment analysis and topic mining using a novel deep attention-based parallel dual-channel model for online course reviews," *Cognit. Comput.*, vol. 15, no. 1, pp. 304–322, Jan. 2023.

[118] M. M. Abdelgwad, T. H. A. Soliman, and A. I. Taloba, "Arabic aspect sentiment polarity classification using BERT," *J. Big Data*, vol. 9, no. 1, p. 115, Dec. 2022.

[119] M. M. Samia, A. Rajee, M. R. Hasan, M. O. Faruq, and P. C. Paul, "Aspect-based sentiment analysis for Bengali text using bidirectional encoder representations from transformers (BERT)," *Int. J. Adv. Comput. Sci. Appl.*, vol. 13, no. 12, pp. 1–9, 2022.

[120] M. Al-Smadi, M. M. Hammad, S. A. Al-Zboon, S. Al-Tawalbeh, and E. Cambria, "Gated recurrent unit with multilingual universal sentence encoder for Arabic aspect-based sentiment analysis," *Knowl.-Based Syst.*, vol. 261, Feb. 2023, Art. no. 107540.

[121] M. Qorich and R. El Ouazzani, "Text sentiment classification of Amazon reviews using word embeddings and convolutional neural networks," *J. Supercomput.*, vol. 79, no. 10, pp. 11029–11054, Jul. 2023.

[122] A. Siagh, F. Z. Laallam, O. Kazar, and H. Salem, "An improved sentiment classification model based on data quality and word embeddings," *J. Supercomput.*, vol. 79, no. 11, pp. 11871–11894, Jul. 2023.

[123] A. Alqarni and A. Rahman, "Arabic tweets-based sentiment analysis to investigate the impact of COVID-19 in KSA: A deep learning approach," *Big Data Cognit. Comput.*, vol. 7, no. 1, p. 16, Jan. 2023.

[124] H. Said, B. S. Tawfik, and M. A. Makhlouf, "A deep learning approach for sentiment classification of COVID-19 vaccination tweets," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 4, pp. 1–10, 2023.

[125] N. Parveen, P. Chakrabarti, B. T. Hung, and A. Shaik, "Twitter sentiment analysis using hybrid gated attention recurrent network," *J. Big Data*, vol. 10, no. 1, p. 50, Apr. 2023.

[126] C. V. Loc, T. X. Viet, T. H. Viet, L. H. Thao, and N. H. Viet, "Pre-trained language model-based deep learning for sentiment classification of Vietnamese feedback," *Int. J. Comput. Intell. Appl.*, vol. 22, no. 3, Sep. 2023, Art. no. 2350016.

[127] D. A. J. Daniel and M. J. Meena, "Deep learning-based hybrid sentiment analysis with feature selection using optimization algorithm," *Multimedia Tools Appl.*, vol. 82, no. 28, pp. 43273–43296, Nov. 2023.

[128] A. Baqach and A. Battou, "CLAS: A new deep learning approach for sentiment analysis from Twitter data," *Multimedia Tools Appl.*, vol. 82, no. 30, pp. 47457–47475, Dec. 2023.

[129] S. Aslan, "A deep learning-based sentiment analysis approach (MF-CNN-BILSTM) and topic modeling of tweets related to the Ukraine–Russia conflict," *Appl. Soft Comput.*, vol. 143, Aug. 2023, Art. no. 110404.

[130] Y. Zahidi, Y. Al-Amrani, and Y. El Younoussi, "Improving Arabic sentiment analysis using LSTM based on word embedding models," *Vietnam J. Comput. Sci.*, vol. 10, no. 3, pp. 391–407, Aug. 2023.

[131] J. Cao, J. Li, M. Yin, and Y. Wang, "Online reviews sentiment analysis and product feature improvement with deep learning," *ACM Trans. Asian Low-Resource Lang. Inf. Process.*, vol. 22, no. 8, pp. 1–17, Aug. 2023.

[132] S. Angamuthu and P. Trojovský, "Integrating multi-criteria decision-making with hybrid deep learning for sentiment analysis in recommender systems," *PeerJ Comput. Sci.*, vol. 9, p. e1497, Aug. 2023.

[133] K. K. Mohbey, G. Meena, S. Kumar, and K. Lokesh, "A CNN-LSTM-based hybrid deep learning approach for sentiment analysis on monkeypox tweets," *New Gener. Comput.*, vol. 42, no. 1, pp. 89–107, Aug. 2023.

[134] N. Habbat, H. Nouri, H. Anoun, and L. Hassouni, "Sentiment analysis of imbalanced datasets using BERT and ensemble stacking for deep learning," *Eng. Appl. Artif. Intell.*, vol. 126, Nov. 2023, Art. no. 106999.

[135] L. Dang, C. Wang, M.-H. Tsou, Y.-E. Hou, and H. Han, "Sentiment analysis of COVID-19 related social distancing using Twitter data based on deep learning," *Multimedia Tools Appl.*, vol. 83, no. 11, pp. 32587–32612, Sep. 2023.

[136] M. S. Başarslan and F. Kayaalp, "Sentiment analysis using a deep ensemble learning model," *Multimedia Tools Appl.*, vol. 83, no. 14, pp. 42207–42231, Oct. 2023.

[137] M. Wankhade, C. S. R. Annavarapu, and A. Abraham, "CBMAFM: CNN-BiLSTM multi-attention fusion mechanism for sentiment classification," *Multimedia Tools Appl.*, vol. 83, no. 17, pp. 51755–51786, Nov. 2023.

[138] R. Ghasemi and S. Momtazi, "How a deep contextualized representation and attention mechanism justifies explainable cross-lingual sentiment analysis," *ACM Trans. Asian Low-Resource Lang. Inf. Process.*, vol. 22, no. 11, pp. 1–15, Nov. 2023.

[139] I. N. Yulita, A. F. Al-Auza'i, A. S. Prabuwono, A. Sholahuddin, F. Ardiansyah, I. Sarathan, and Y. Djuyandi, "Bidirectional long short-term memory for analysis of public opinion sentiment on government policy during the COVID-19 pandemic," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 11, pp. 877–885, 2023.

[140] W. Chang and M. Zhu, "Sentiment analysis method of consumer comment text based on BERT and hierarchical attention in e-commerce big data environment," *J. Intell. Syst.*, vol. 32, no. 1, Dec. 2023, Art. no. 20230025.

[141] H. Murfi, T. Gowandi, G. Ardaneswari, and S. Nurrohmah, "BERT-based combination of convolutional and recurrent neural network for Indonesian sentiment analysis," *Appl. Soft Comput.*, vol. 151, Jan. 2024, Art. no. 111112.

[142] I. Kanwal, F. Wahid, S. Ali, A.-U. Rehman, A. Alkhayyat, and A. Al-Radaei, "Sentiment analysis using hybrid model of stacked auto-encoder-based feature extraction and long short term memory-based classification approach," *IEEE Access*, vol. 11, pp. 124181–124197, 2023.

[143] P. Thiengburanathum and P. Charoenkwan, "SETAR: Stacking ensemble learning for Thai sentiment analysis using RoBERTa and hybrid feature representation," *IEEE Access*, vol. 11, pp. 92822–92837, 2023.

[144] C. Ahle Touate, R. El Ayachi, and M. Biniz, "Classification of sentiment using optimized hybrid deep learning model," *Comput. Informat.*, vol. 42, no. 3, pp. 651–666, 2023.

[145] J. Khan, N. Ahmad, S. Khalid, F. Ali, and Y. Lee, "Sentiment and context-aware hybrid DNN with attention for text sentiment classification," *IEEE Access*, vol. 11, pp. 28162–28179, 2023.

[146] B. B. Al-Onazi, A. Q. A. Hassan, M. K. Nour, M. Al Duhayyim, A. Mohamed, A. Atta Abdelmageed, I. Yaseen, and G. Pasha Mohammed, "Quantum particle swarm optimization with deep learning-based Arabic tweets sentiment analysis," *Comput., Mater. Continua*, vol. 75, no. 2, pp. 2575–2591, 2023.

[147] W. Ahmad, H. U. Khan, T. Iqbal, and S. Iqbal, "Attention-based multi-channel gated recurrent neural networks: A novel feature-centric approach for aspect-based sentiment classification," *IEEE Access*, vol. 11, pp. 54408–54427, 2023.

[148] R. Vatambeti, S. V. Mantena, K. V. D. Kiran, M. Manohar, and C. Manjunath, "Twitter sentiment analysis on online food services based on elephant herd optimization with hybrid deep learning technique," *Cluster Comput.*, vol. 27, no. 1, pp. 655–671, Feb. 2024.

[149] G. Kaur and A. Sharma, "A deep learning-based model using hybrid feature extraction approach for consumer sentiment analysis," *J. Big Data*, vol. 10, no. 1, p. 5, Jan. 2023.

[150] R. K. Dey and A. K. Das, "Modified term frequency-inverse document frequency based deep hybrid framework for sentiment analysis," *Multimedia Tools Appl.*, vol. 82, no. 21, pp. 32967–32990, Sep. 2023.

[151] M. Irfan, N. Ayub, Q. A. Ahmed, S. Rahman, M. S. Bashir, G. Nowakowski, S. M. Alqhtani, and M. Sieja, "AQSA: Aspect-based quality sentiment analysis for multi-labeling with improved ResNet hybrid algorithm," *Electronics*, vol. 12, no. 6, p. 1298, Mar. 2023.

[152] M. Berrimi, M. Oussalah, A. Moussaoui, and M. Saidi, "Attention mechanism architecture for Arabic sentiment analysis," *ACM Trans. Asian Low-Resource Lang. Inf. Process.*, vol. 22, no. 4, pp. 1–26, Mar. 2023.

[153] S. Soumya and K. V. Pramod, "Hybrid deep learning approach for sentiment classification of Malayalam tweets," *Int. J. Adv. Comput. Sci. Appl.*, vol. 13, no. 4, pp. 1–9, 2022.

[154] X. Zhang and Y. Ma, "An ALBERT-based TextCNN-hatt hybrid model enhanced with topic knowledge for sentiment analysis of sudden-onset disasters," *Eng. Appl. Artif. Intell.*, vol. 123, Aug. 2023, Art. no. 106136.

[155] C.-H. Lin and U. Nuha, "Sentiment analysis of Indonesian datasets based on a hybrid deep-learning strategy," *J. Big Data*, vol. 10, no. 1, p. 88, May 2023.

[156] K. B. Muhammad and S. M. A. Burney, "Innovations in Urdu sentiment analysis using machine and deep learning techniques for two-class classification of symmetric datasets," *Symmetry*, vol. 15, no. 5, p. 1027, May 2023.

[157] M. Mujahid, K. Kanwal, F. Rustam, W. Aljedaani, and I. Ashraf, "Arabic ChatGPT tweets classification using RoBERTa and BERT ensemble model," *ACM Trans. Asian Low-Resource Lang. Inf. Process.*, vol. 22, no. 8, pp. 1–23, Aug. 2023.

[158] A. S. Talaat, "Sentiment analysis classification system using hybrid BERT models," *J. Big Data*, vol. 10, no. 1, p. 110, Jun. 2023.

[159] F. Alemayehu, M. Meshesha, and J. Abate, "Amharic political sentiment analysis using deep learning approaches," *Sci. Rep.*, vol. 13, no. 1, p. 17982, Oct. 2023.

[160] H. Tang and R. Zhang, "Research on Chinese microblog sentiment classification based on TextCNN-BiLSTM model," *J. Inf. Process. Syst.*, vol. 19, no. 6, pp. 842–857, Dec. 2023.

[161] N. Singh and U. C. Jaiswal, "Sentiment analysis based on Urdu reviews using hybrid deep learning models," *Appl. Comput. Syst.*, vol. 28, no. 2, pp. 258–265, Dec. 2023.

[162] Y. Jin and A. Zhao, "Bert-based graph unlinked embedding for sentiment analysis," *Complex Intell. Syst.*, vol. 10, no. 2, pp. 2627–2638, Apr. 2024.

[163] M. Wankhade, C. S. R. Annavarapu, and M. K. Verma, "CBVoSD: Context based vectors over sentiment domain ensemble model for review classification," *J. Supercomput.*, vol. 78, no. 5, pp. 6411–6447, Apr. 2022.

[164] E. H. Mohamed, M. E. Moussa, and M. H. Haggag, "An enhanced sentiment analysis framework based on pre-trained word embedding," *Int. J. Comput. Intell. Appl.*, vol. 19, no. 4, Dec. 2020, Art. no. 2050031.

[165] S. Siddiqui, M. A. Rehman, S. M. Daudpota, and A. Waqas, "Opinion mining: An approach to feature engineering," *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 3, pp. 1–7, 2019.

[166] S. Symeonidis, D. Effrosynidis, and A. Arampatzis, "A comparative evaluation of pre-processing techniques and their interactions for Twitter sentiment analysis," *Expert Syst. Appl.*, vol. 110, pp. 298–310, Nov. 2018.

[167] F. H. Khan, U. Qamar, and S. Bashir, "Lexicon based semantic detection of sentiments using expected likelihood estimate smoothed odds ratio," *Artif. Intell. Rev.*, vol. 48, no. 1, pp. 113–138, Jun. 2017.

[168] M. Kamyab, G. Liu, A. Rasool, and M. Adjeisah, "ACR-SA: Attention-based deep model through two-channel CNN and bi-RNN for sentiment analysis," *PeerJ Comput. Sci.*, vol. 8, p. e877, Mar. 2022.

[169] H. Suresh, "An innovative and efficient method for Twitter sentiment analysis," *Int. J. Data Mining, Model. Manage.*, vol. 11, no. 1, pp. 1–18, 2019.

[170] H. Liu and M. Cocea, "Fuzzy information granulation towards interpretable sentiment analysis," *Granular Comput.*, vol. 2, no. 4, pp. 289–302, Dec. 2017.

[171] V. Rupapara, F. Rustam, A. Amaar, P. B. Washington, E. Lee, and I. Ashraf, "Deepfake tweets classification using stacked bi-LSTM and words embedding," *PeerJ Comput. Sci.*, vol. 7, p. e745, Oct. 2021.

[172] A. Suciati and I. Budi, "Aspect-based sentiment analysis and emotion detection for code-mixed review," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 9, pp. 179–186, 2020.

[173] S. Sharmin and D. Chakma, "Attention-based convolutional neural network for Bangla sentiment analysis," *AI Soc.*, vol. 36, no. 1, pp. 381–396, Mar. 2021.

[174] M. L. Barrón Estrada, R. Zatarain Cabada, R. Oramas Bustillos, and M. Graff, "Opinion mining and emotion recognition applied to learning environments," *Expert Syst. Appl.*, vol. 150, Jul. 2020, Art. no. 113265.

[175] B. Gaye, D. Zhang, and A. Wulamu, "A tweet sentiment classification approach using a hybrid stacked ensemble technique," *Information*, vol. 12, no. 9, p. 374, Sep. 2021.

[176] N. Capuano, L. Greco, P. Ritrovato, and M. Vento, "Sentiment analysis for customer relationship management: An incremental learning approach," *Appl. Intell.*, vol. 51, pp. 3339–3352, Jun. 2021.

[177] O. Gokalp, E. Tasci, and A. Ugur, "A novel wrapper feature selection algorithm based on iterated greedy Metaheuristic for sentiment classification," *Expert Syst. Appl.*, vol. 146, May 2020, Art. no. 113176.

[178] A. Madasu and S. Elango, "Efficient feature selection techniques for sentiment analysis," *Multimedia Tools Appl.*, vol. 79, nos. 9–10, pp. 6313–6335, Mar. 2020.

[179] L. Zheng, H. Wang, and S. Gao, "Sentimental feature selection for sentiment analysis of Chinese online reviews," *Int. J. Mach. Learn. Cybern.*, vol. 9, pp. 75–84, Jan. 2018.

[180] M. I. A. Latiffi, M. R. Yaakub, and I. S. Ahmad, "Flower pollination algorithm for feature selection in tweets sentiment analysis," *Int. J. Adv. Comput. Sci. Appl.*, vol. 13, no. 5, pp. 1–8, 2022.

[181] A. Nasser and H. Sever, "A concept-based sentiment analysis approach for Arabic," *Int. Arab J. Inf. Technol.*, vol. 17, no. 5, pp. 778–788, Sep. 2020.

[182] P. Q. Tran, N. T. Trieu, N. V. Dao, H. T. Nguyen, and H. X. Huynh, "Effective opinion words extraction for food reviews classification," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 7, pp. 421–426, 2020.

[183] C. Catal and M. Nangir, "A sentiment classification model based on multiple classifiers," *Appl. Soft Comput.*, vol. 50, pp. 135–141, Jan. 2017.

[184] M. Giatsoglou, M. G. Vozalis, K. Diamantaras, A. Vakali, G. Sarigiannidis, and K. C. Chatzisavvas, "Sentiment analysis leveraging emotions and word embeddings," *Expert Syst. Appl.*, vol. 69, pp. 214–224, Mar. 2017.

[185] B. Al sari, R. Alkhaldi, D. Alsaffar, T. Alkhaldi, H. Almaymuni, N. Alnaim, N. Alghamdi, and S. O. latunji, "Sentiment analysis for cruises in Saudi Arabia on social media platforms using machine learning algorithms," *J. Big Data*, vol. 9, p. 21, Feb. 2022.

[186] A. Mohasseb, M. Bader-El-Den, and M. Cocea, "A customised grammar framework for query classification," *Expert Syst. Appl.*, vol. 135, pp. 164–180, Nov. 2019.

[187] R. Chiha, M. B. Ayed, and C. d. C. Pereira, "A complete framework for aspect-level and sentence-level sentiment analysis," *Appl. Intell.*, vol. 52, pp. 17845–17863, Apr. 2022.

[188] H. Sadr and M. Nazari Soleimandarabi, "ACNN-TL: Attention-based convolutional neural network coupling with transfer learning and contextualized word representation for enhancing the performance of sentiment classification," *J. Supercomput.*, vol. 78, no. 7, pp. 10149–10175, May 2022.

[189] R. Yan, Y. Yu, and D. Qiu, "Emotion-enhanced classification based on fuzzy reasoning," *Int. J. Mach. Learn. Cybern.*, vol. 13, no. 3, pp. 839–850, Mar. 2022.

[190] N. Gozuacik, C. O. Sakar, and S. Ozcan, "Social media-based opinion retrieval for product analysis using multi-task deep neural networks," *Expert Syst. Appl.*, vol. 183, Nov. 2021, Art. no. 115388.

[191] S. Al-Dabet, S. Tedmori, and M. Al-Smadi, "Enhancing Arabic aspect-based sentiment analysis using deep learning models," *Comput. Speech Lang.*, vol. 69, Sep. 2021, Art. no. 101224.

[192] F. Huang, X. Li, C. Yuan, S. Zhang, J. Zhang, and S. Qiao, "Attention-emotion-enhanced convolutional LSTM for sentiment analysis," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 9, pp. 4332–4345, Sep. 2022.

[193] K. Chen and W. Ke, "A hierarchical neural model for target-based sentiment analysis," *Concurrency Comput., Pract. Exper.*, vol. 33, no. 10, p. e6184, May 2021.

[194] J. Yang, X. Zou, W. Zhang, and H. Han, "Microblog sentiment analysis via embedding social contexts into an attentive LSTM," *Eng. Appl. Artif. Intell.*, vol. 97, Jan. 2021, Art. no. 104048.

[195] K. Ayyub, S. Iqbal, M. W. Nisar, S. G. Ahmad, and E. U. Munir, "Stance detection using diverse feature sets based on machine learning techniques," *J. Intell. Fuzzy Syst.*, vol. 40, no. 5, pp. 9721–9740, Apr. 2021.

[196] M. Kasri, M. Birjali, and A. Beni-Hssane, "Word2Sent: A new learning sentiment-embedding model with low dimension for sentence level sentiment classification," *Concurrency Comput., Pract. Exper.*, vol. 33, no. 9, p. e6149, May 2021.

[197] L.-C. Chen, C.-M. Lee, and M.-Y. Chen, "Exploration of social media for sentiment analysis using deep learning," *Soft Comput.*, vol. 24, no. 11, pp. 8187–8197, Jun. 2020.

[198] F. Mehmood, M. U. Ghani, M. A. Ibrahim, R. Shahzadi, W. Mahmood, and M. N. Asim, "A precisely xtreme-multi channel hybrid approach for Roman Urdu sentiment analysis," *IEEE Access*, vol. 8, pp. 192740–192759, 2020.

[199] M. U. Salur and I. Aydin, "A novel hybrid deep learning model for sentiment classification," *IEEE Access*, vol. 8, pp. 58080–58093, 2020.

[200] C. N. Dang, M. N. Moreno-García, and F. De la Prieta, "Hybrid deep learning models for sentiment analysis," *Complexity*, vol. 2021, no. 1, Aug. 2021, Art. no. 9986920.

[201] T. Li, H. Xu, Z. Liu, Z. Dong, Q. Liu, J. Li, S. Fan, and X. Sun, "A spatiotemporal multi-feature extraction framework for opinion mining," *Neurocomputing*, vol. 490, pp. 337–346, Jun. 2022.

[202] F. Alotaibi and V. Gupta, "Sentiment analysis system using hybrid word embeddings with convolutional recurrent neural network," *Int. Arab J. Inf. Technol.*, 5 vol. 19, no. 3, pp. 1–6, 2022.

[203] H. Yang, A. Alsadoon, P. W. C. Prasad, T. Al-Dala'In, T. A. Rashid, A. Maag, and O. H. Alsadoon, "Deep learning neural networks for emotion classification from text: Enhanced leaky rectified linear unit activation and weighted loss," *Multimedia Tools Appl.*, vol. 81, no. 11, pp. 15439–15468, May 2022.

[204] O. Alharbi, "A deep learning approach combining CNN and Bi-LSTM with SVM classifier for Arabic sentiment analysis," *Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 6, pp. 1–8, 2021.

[205] X. Li, R. Lu, P. Liu, and Z. Zhu, "Graph convolutional networks with hierarchical multi-head attention for aspect-level sentiment classification," *J. Supercomput.*, vol. 78, no. 13, pp. 14846–14865, Apr. 2022.

[206] J. Zhou, J. X. Huang, Q. V. Hu, and L. He, "Is position important? Deep multi-task learning for aspect-based sentiment analysis," *Appl. Intell.*, vol. 50, pp. 3367–3378, Jun. 2020.

[207] 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, Jun. 2020.

[208] A. Onan, "Sentiment analysis on product reviews based on weighted word embeddings and deep neural networks," *Concurrency Comput., Pract. Exper.*, vol. 33, no. 23, p. e5909, Dec. 2021.

[209] G. Lu, J. Gan, J. Yin, Z. Luo, B. Li, and X. Zhao, "Multi-task learning using a hybrid representation for text classification," *Neural Comput. Appl.*, vol. 32, no. 11, pp. 6467–6480, Jun. 2020.

[210] G. Lu, X. Zhao, J. Yin, W. Yang, and B. Li, "Multi-task learning using variational auto-encoder for sentiment classification," *Pattern Recognit. Lett.*, vol. 132, pp. 115–122, Apr. 2020.

[211] G. Liu, X. Huang, X. Liu, and A. Yang, "A novel aspect-based sentiment analysis network model based on multilingual hierarchy in online social network," *Comput. J.*, vol. 63, no. 1, pp. 410–424, Jan. 2020.

[212] H.-J. 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," *Knowl.-Based Syst.*, vol. 187, Jan. 2020, Art. no. 104825.

[213] S. L. Ramaswamy and J. Chinnappan, "RecogNet-LSTM+CNN: A hybrid network with attention mechanism for aspect categorization and sentiment classification," *J. Intell. Inf. Syst.*, vol. 58, pp. 379–404, Jan. 2022.

[214] L. Bao, P. Lambert, and T. Badia, "Improving aspect-based neural sentiment classification with lexicon enhancement, attention regularization and sentiment induction," *Natural Lang. Eng.*, vol. 30, no. 1, pp. 1–30, 2022.

[215] L.-C. Cheng, Y.-L. Chen, and Y.-Y. Liao, "Aspect-based sentiment analysis with component focusing multi-head co-attention networks," *Neurocomputing*, vol. 489, pp. 9–17, Jun. 2022.

[216] 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," *Knowl.-Based Syst.*, vol. 228, Sep. 2021, Art. no. 107242.

[217] J. Tao and X. Fang, "Toward multi-label sentiment analysis: A transfer learning based approach," *J. Big Data*, vol. 7, p. 1, Jan. 2020.

[218] C. N. Dang, M. N. Moreno-García, and F. De la Prieta, "Using hybrid deep learning models of sentiment analysis and item genres in recommender systems for streaming services," *Electronics*, vol. 10, no. 20, p. 2459, Oct. 2021.

[219] N. Liu and B. Shen, "Aspect term extraction via information-augmented neural network," *Complex Intell. Syst.*, vol. 9, no. 1, pp. 537–563, Jul. 2022.

[220] S. Vernikou, A. Lyras, and A. Kanavos, "Multiclass sentiment analysis on COVID-19-related tweets using deep learning models," *Neural Comput. Appl.*, vol. 34, no. 22, pp. 19615–19627, Nov. 2022.

[221] C. Yan, J. Liu, W. Liu, and X. Liu, "Research on public opinion sentiment classification based on attention parallel dual-channel deep learning hybrid model," *Eng. Appl. Artif. Intell.*, vol. 116, Nov. 2022, Art. no. 105448.

[222] K. Shrivastava and S. Kumar, "A sentiment analysis system for the Hindi language by integrating gated recurrent unit with genetic algorithm," *Int. Arab J. Inf. Technol.*, vol. 17, no. 6, pp. 954–964, 2020.

[223] K. Dashtipour, M. Gogate, J. Li, F. Jiang, B. Kong, and A. Hussain, "A hybrid Persian sentiment analysis framework: Integrating dependency grammar based rules and deep neural networks," *Neurocomputing*, vol. 380, pp. 1–10, Mar. 2020.

[224] A. Vohra and R. Garg, "Deep learning based sentiment analysis of public perception of working from home through tweets," *J. Intell. Inf. Syst.*, vol. 60, no. 1, pp. 255–274, Aug. 2022.

[225] A. Muslim, A. B. Mutiara, R. Refianti, C. M. Karyati, and G. Setiawan, "Comparison of accuracy between long short-term memory-deep learning and multinomial logistic regression-machine learning in sentiment analysis on Twitter," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 2, pp. 1–8, 2020.

[226] F. A. Lovera, Y. C. Cardinale, and M. N. Homsi, "Sentiment analysis in Twitter based on knowledge graph and deep learning classification," *Electronics*, vol. 10, no. 22, p. 2739, Nov. 2021.

[227] H. Yu, G. Lu, Q. Cai, and Y. Xue, "A KGE based knowledge enhancing method for aspect-level sentiment classification," *Mathematics*, vol. 10, no. 20, p. 3908, Oct. 2022.

[228] A. Naresh and P. V. Krishna, "An efficient approach for sentiment analysis using machine learning algorithm," *Evol. Intell.*, vol. 14, pp. 725–731, Jun. 2021.

[229] W. Muhammad, M. Mushtaq, K. N. Junejo, and M. Y. Khan, "Sentiment analysis of product reviews in the absence of labelled data using supervised learning approaches," *Malaysian J. Comput. Sci.*, vol. 33, no. 2, pp. 118–132, Apr. 2020.

[230] Z. Wang and Z. Lin, "Optimal feature selection for learning-based algorithms for sentiment classification," *Cogn. Comput.*, vol. 12, pp. 238–248, Jan. 2020.

[231] I. Safder, Z. Mahmood, R. Sarwar, S.-U. Hassan, F. Zaman, R. M. A. Nawab, F. Bukhari, R. A. Abbasi, S. Alelyani, N. R. Aljohani, and R. Nawaz, "Sentiment analysis for Urdu online reviews using deep learning models," *Expert Syst.*, vol. 38, no. 8, 2021, Art. no. e12751.

[232] O. Appel, F. Chiclana, J. Carter, and H. Fujita, "Cross-ratio uninorms as an effective aggregation mechanism in sentiment analysis," *Knowl.-Based Syst.*, vol. 124, pp. 16–22, May 2017.

[233] L. Chiruzzo, M. Etcheverry, and A. Rosá, "Sentiment analysis in Spanish tweets: Some experiments with focus on neutral tweets," *Procesamiento del Lenguaje Natural*, vol. 64, pp. 109–116, Jan. 2020.

[234] R. Dhanalakshmi and T. Sri Devi, "Adaptive cognitive intelligence in analyzing employee feedback using LSTM," *J. Intell. Fuzzy Syst.*, vol. 39, no. 6, pp. 8069–8078, Dec. 2020.

[235] Q. Cheng, Y. Ke, and A. Abdelmouty, "Negative emotion diffusion and intervention countermeasures of social networks based on deep learning," *J. Intell. Fuzzy Syst.*, vol. 39, no. 4, pp. 4935–4945, Oct. 2020.

[236] S. Rani, A. K. Bashir, A. Alhudhaif, D. Koundal, and E. S. Gunduz, "An efficient CNN-LSTM model for sentiment detection in #BlackLivesMatter," *Expert Syst. Appl.*, vol. 193, May 2022, Art. no. 116256.

[237] K. Pasupa and T. S. Na Ayutthaya, "Hybrid deep learning models for Thai sentiment analysis," *Cogn. Comput.*, vol. 14, pp. 167–193, Jan. 2022.

[238] Y. Long, R. Xiang, Q. Lu, C.-R. Huang, and M. Li, ''Improving attention model based on cognition grounded data for sentiment analysis,'' *IEEE Trans. Affect. Comput.*, vol. 12, no. 4, pp. 900–912, Oct. 2021.

[239] N. Liu and B. Shen, ''Aspect-based sentiment analysis with gated alternate neural network,'' *Knowl.-Based Syst.*, vol. 188, Jan. 2020, Art. no. 105010.

YONATAN MAMANI-COAQUIRA received the bachelor's degree in computer and systems engineering from the National University Micaela Bastidas of Apurimac, Peru, in 2013, and the master's degree in computer science from the National University of San Antonio Abad of Cusco, Peru, in 2021. Currently, he is pursuing the Ph.D. degree in engineering, with a specialization in computer science with the Pontifical Catholic University of Peru (PUCP).

He is a Conducting Researcher with the Artificial Intelligence Group, PUCP. His research interests include machine learning, deep learning, data analytics, computer vision, and natural language processing. He previously had an internship and collaborated as a Researcher with the Data Analytics Science and Engineering Laboratory, Pablo de Olavide University, Seville, Spain.

EDWIN VILLANUEVA received the degree in electrical engineering from the National University of San Agustin of Arequipa, Peru, in 2000, the master's degree from the University of São Paulo, Brazil, the master's degree in electrical engineering, in 2007, and the Ph.D. degree in science, in 2012

He completed postdoctoral stays with the Institute of Mathematical and Computer Sciences, University of São Paulo, and the Institute of Functional Genomics, Universität Regensburg, Germany. He was an Engineer in several public and private institutions in Peru, for more than six years, accumulating experience in the telecommunications and software development sectors. He participated as a Postdoctoral Researcher with the Institute of Mathematical and Computer Sciences, University of São Paulo, and Universität Regensburg. Since 2017, he has been a Senior Researcher with the Pontifical Catholic University of Peru. His research has mainly focused on areas of computer science, specifically artificial intelligence, machine learning, probabilistic graphical models, signal processing, and bioinformatics. He had been leading research and technological development projects in multidisciplinary fields, such as modeling and prediction of atmospheric phenomena, air quality monitoring and modeling, and modeling and prediction of financial markets.