IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 4 June 2023, accepted 28 June 2023, date of publication 7 July 2023, date of current version 13 July 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3293041

SURVEY

Challenges and Issues in Sentiment Analysis: A Comprehensive Survey

NILAA RAGHUNATHAN AND SARAVANAKUMAR KANDASAMY
Department of Computer Science, Vellore Institute of Technology, Vellore 632014, India

Corresponding author: Saravanakumar Kandasamy (ksaravanakumar@vit.ac.in)

ABSTRACT Sentiment analysis, a specialization of natural language processing (NLP), has witnessed significant progress since its emergence in the late 1990s, owing to the swift advances in deep learning techniques and the abundance of vast digital datasets. Though sentiment analysis has reached a relatively advanced stage in the area of NLP, it is erroneously assumed that sentiment analysis has reached its pinnacle, leaving no room for further improvement. However, it is important to acknowledge that numerous challenges that require attention persist. This survey paper provides a comprehensive overview of sentiment analysis, including its applications, approaches to sentiment classification, and commonly used evaluation metrics. The survey primarily focuses on the challenges associated with different types of data for sentiment classification, namely cross-domain data, multimodal data, cross-lingual data, and small-scale data, and provides a review of the state-of-the-art in sentiment analysis to address these challenges. The paper also addresses the challenges faced during sentiment classification irrespective of the type of data available. It aims at a better understanding of sentiment analysis to enable practitioners and researchers select suitable methods for sentiment classification depending on the type of data being analyzed.

INDEX TERMS Machine learning, sentiment analysis, natural language processing, cross-domain data, multimodal data, cross-lingual data, small-scale data.

I. INTRODUCTION

Sentiment analysis, also known as emotion AI or opinion mining, is a topic of study that seeks to comprehend the sentiment behind unstructured information. For instance, in 'Liam likes apples', Liam is the opinion holder, and his emotional state or sentiment on the aspect - apples is positive. The emotional state of the opinion holder is extracted from the given text by using machine learning (ML), natural language processing (NLP), and lexicon-based techniques.

A. GROWTH OF SENTIMENT ANALYSIS

In the internet age, driven by the younger generation, the amount of data generated and shared is growing exponentially. The data generated from the year 2020 onwards is nineteen percent more than all the data generated before the year 2020 [1]. This rapid increase in data generation will most likely continue in the near future, owing to the rapid rise in the number of web applications, almost infinite connections, and an unquenchable thirst for knowledge extraction.

Another aspect is that the data generated and shared is diversifying over time. When the internet started in the 1970s, it was exclusively used for military purposes by governments, but as of 2023, the internet and social media have 5.16 and 4.62 billion users worldwide, respectively [2], [3]. It has caused a proliferation of opinion sharing through blogs and social media posts on virtually anything and everything. Hence, data generated on the internet are acknowledged as a valuable source of data with an additional benefit for a broad range of application disciplines.

Another key driver of the growth of sentiment analysis is the increasing use of chatbots and virtual assistants. These conversational agents rely on NLP and sentiment analysis to comprehend and respond to user queries and feedback. The role of sentiment analysis is to recognize the state of the mind of the customer and provide a befitting response.

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

Sentiment analysis has witnessed significant growth in terms of techniques used in the last decade, with traditional approaches such as lexicon-based and rule-based methods being replaced by more sophisticated machine learning and deep learning techniques. Among these techniques, supervised machine learning, deep learning, and transfer learning have been the most widely adopted.

The growth of deep learning, supervised machine learning, and transfer learning techniques in the last decade has enabled more accurate and nuanced sentiment analysis results, providing businesses and organizations with deeper insights into customer opinions and sentiment trends.

B. APPLICATIONS

Sentiment classification has numerous applications as it enables automated analysis of large volumes of textual data, providing valuable insights that can be used to inform decision-making processes.

1) PUBLIC OPINION VIA SOCIAL MEDIA

Before the arrival of social networks, stakeholders, like non-profit organizations and corporations had to rely solely on media outlets to know public opinion. This gave the power to media outlets to control the narrative and flow of public opinion and partially prevented all the other stakeholders from verifying the raw data themselves. With the rise of social media, all stakeholders not only have ready admittance to the generated data in social media through a huge range of databases but also to collect them using technologies like tweepy and textblob [4]. Content generated from social media has been one of the best compilations of material from which to extricate public sentiment on a particular topic [5], [6], [7], [8]. The availability of raw data has also enabled them to detect fake news generated [9], [10], which would be very hard in traditional methods such as newspapers and magazines.

2) PRODUCT ANALYSIS

The e-commerce market has witnessed a compounded annual growth rate (CAGR) of 14.70% since the year 2020 [11]. This has led to an increasing number of customers inclined to post their individual opinions about a product, and hence the commercial significance of reviews has also been growing. A contentious issue nowadays is how to recognize and gather customer thoughts and attitudes from voluminous casual remarks. This requirement is addressed by sentiment analysis. Several works have been proven to have successfully classified reviews on e-commerce sites [12], [13], [14], [15]. Such sentiment analysis on reviews has proven beneficial not only for e-commerce sites but also for other applications like movie reviews as well [16], [17], [18].

3) STOCK MARKET

Trends in the stock market tend to reflect the state of the economy and the capital movement. These tendencies typically change in response to real-life events like parliamentary elections, industrial breakthroughs, or some form of calamity, but the most impactful factor is the statistics of these events and the public reaction. Therefore, the stock market is extremely volatile and prediction of the direction of the market is complicated. However, past works have proven that the trend of the stock market can also be predicted successfully [19], [20], [21], [22], [23], [24].

4) ENTERPRISE MANAGEMENT

Due to the increased competition among enterprises, traditional enterprise management techniques no longer have the capability of meeting the demands of sustainable development of businesses or making use of social data effectively. That being the case, organizations have turned to sentiment analysis to comprehend the emotional state that business employees experience on a regular basis [25], [26].

C. OUTLINE OF SENTIMENT ANALYSIS

Opinion mining or emotion AI, more commonly referred to as sentiment analysis, is the study of the emotional state of subjective data through systematic identification, extraction, and quantification. The foundation of sentiment categorization is based on three levels:

i. Document Level: The assignment at this level is to ascertain if the overall emotion the paper exhibits is favorable, unfavorable, or neutral.
ii. Sentence Level: This work at this level focuses on the sentences and evaluates if each one represents a favorable, unfavorable, or neutral viewpoint.
iii. Feature Level: This level intends to recognize the sentiment about a particular feature or entity of a given text. For instance, in the expression, 'The film had a good story'- the film is the entity and story is the aspect of the entity.

Having established the level of sentiment analysis classification, it is necessary to explore the mechanism by which this classification is carried out. Classification of any text has two stages. The first stage comprises pre-processing the data and the second stage comprises the classification of data. Data pre-processing in turn consists of two major steps-tokenization and token normalization. Tokenization is the operation of dividing a chain of input characters into the so-called individual tokens. Token normalization is the operation of reducing the derivational and inflectional constructs of a word to a common base construct. Token normalization consists of stemming, lemmatization, and removal of stop words. Stemming is removing the prefix, suffix, circumfix, and infix of a word to get it to the root form. Lemmatization is the conversion of the several forms of a word into a single form based on a word lemma. Lastly, stop words are words like 'is', 'the', 'a', etc., that have no impact on how sentiment is classified, hence they are removed. The second stage of text classification involves applying a chosen algorithm to the pre-processed data to consign it to predetermined classes or categories based on the task requirements.

D. FOCUS OF THIS SURVEY

Sentiment Analysis is a well-established specialization of NLP, and several survey papers have been published on this topic [27], [28], [29]. However, most of these surveys focus on a specific technique or a specific type of sentiment classification. In this survey paper, we take a different approach that is based on the types of sentiment classification that are best suited for different types of data.

Approach 1 - Based on the technique: One common approach employed in surveys is to base the survey paper on a particular technique for sentiment classification. For example, [30] explores deep learning techniques for sentiment classification. However, if a researcher intends to tackle a real-world problem, the approach to the problem would depend on the available data. If a problem has voluminous data from a different domain, it may require the use of deep learning and transfer learning techniques. In such cases, the researcher may need to refer to two survey papers, where the first survey focuses on aspect-level sentiment classification, and the second survey deals with cross-domain sentiment classification. Our survey takes an approach that depends on the type and source of available data to solve these types of issues. We discuss the challenges of the given data along with the most effective techniques to solve the problem, regardless of whether it is a machine learning or deep learning algorithm that should be used.

Approach 2 - Based on the type of sentiment classification: Some surveys focus on a particular type of sentiment classification [31], [32], [33], [34], however in real-world scenarios, obtaining clean and comprehensive datasets for sentiment analysis can be challenging. Ideally, we would like to have a single dataset containing all the necessary reviews, but this is rarely the case. This can lead to challenges related to multi-source cross-domain, and aspect-level cross-lingual sentiment classification. While some works focus on aspect-level sentiment analysis, they may not address the challenges of cross-domain sentiment classification. To address these challenges, researchers may need to refer to multiple surveys, each addressing a specific type of sentiment classification. Our survey discusses the challenges of different types of sentiment classification, including those that are often overlooked in other surveys.

To summarize, our goal is a survey that covers both the 'Approach 1' and 'Approach 2' discussed above. In other words, we aim to cover both the techniques for and the types of sentiment classification, so that the challenges in the real-world problems are more efficiently addressed.

Fig. 1 shows the breakup of the papers surveyed. To ensure a fair representation, the span was chosen to be two years. To give more weightage to more recent work, 32.7% of the papers are those published in the last two years (2021-22), and 50% of the papers are those published within the last four years (2019-22). However, at 17.6%, significant contributions prior to the year 2015 have also been included in the paper.

E. STRUCTURE OF THIS SURVEY

This survey is structured to provide a comprehensive review of the field of sentiment classification. Beginning with an overview of the various sentiment classification approaches, the evaluation metrics used to assess the performance of these models are examined next. The challenges associated with different kinds of sentiment classification categorized on the type of data are then delved into, and the past methods used to address these challenges are highlighted. The paper culminates with a broad outlook on the field of sentiment classification.

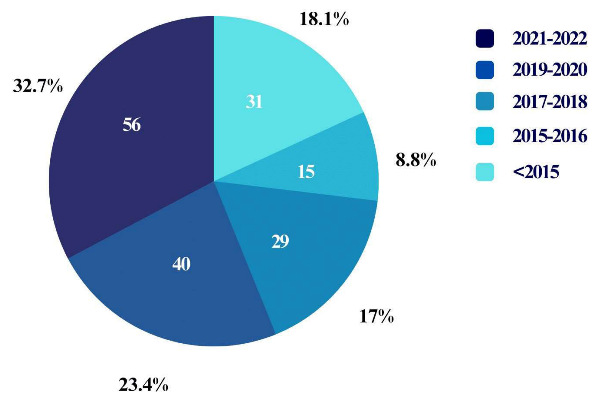

FIGURE 1. Timeline of papers surveyed.

II. APPROACHES TO SENTIMENT CLASSIFICATION

Two primary perspectives to sentiment analysis include machine learning and lexicon-based methods, which are explored in further detail below.

A. MACHINE LEARNING-BASED APPROACH

Machine learning (ML) allows a system to learn from its experiences and enhance its performance. This approach requires training a machine learning model on a labeled dataset, where each text sample is associated with a sentiment label. The model learns to recognize features and patterns in the text which are indicative of each sentiment class. Machine learning has proven its ability to handle the intricacies of human language [35], [36], [37], [38]. The ML-based approach can be further classified into four subgroups, explained next.

1) SUPERVISED LEARNING

Supervised learning is a specific category of machine learning techniques that involves instructing a computer program to recognize and categorize patterns in labeled datasets. These labeled datasets have pre-assigned labels that represent the correct output or desired response. The primary objective of supervised learning is to teach the model to associate input data with the corresponding labeled output data, which in turn enables the model to make accurate predictions on

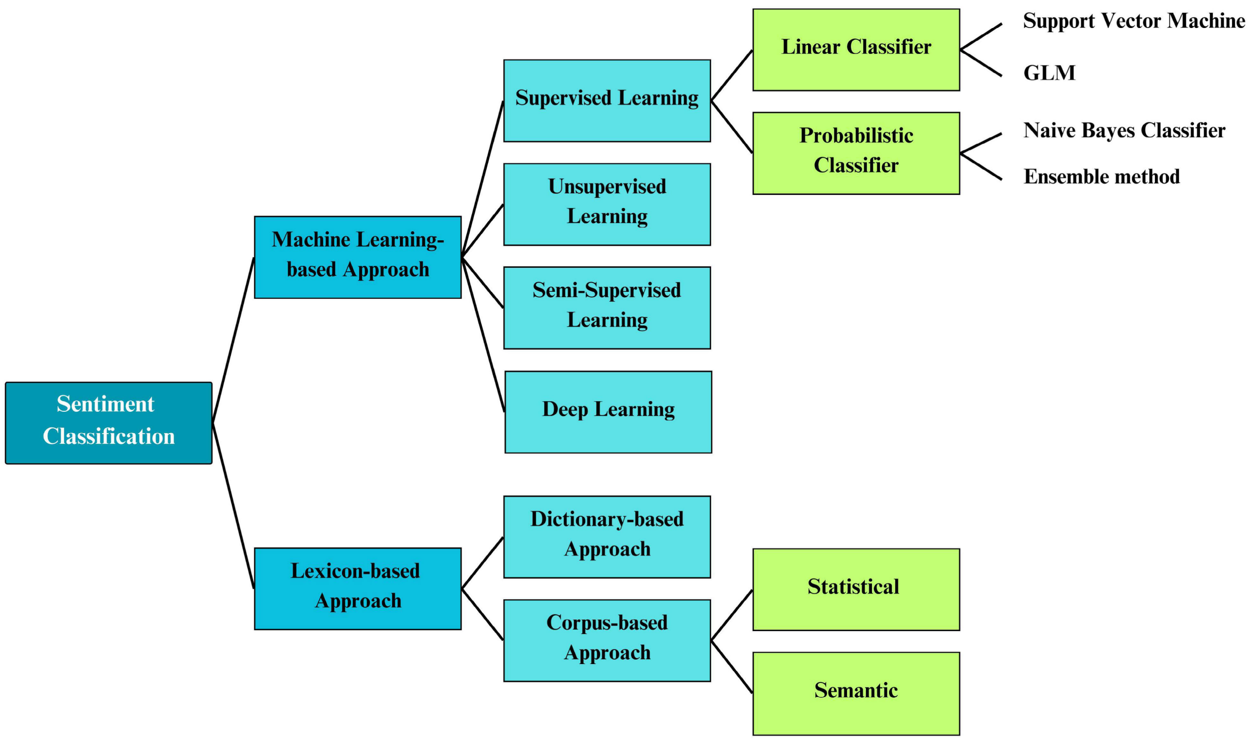

FIGURE 2. **Approaches to sentiment classification and their sub-classifications.**

new data that it has not seen before. Supervised machine learning algorithms are often preferred because they can leverage the labeled examples in the training dataset [39], especially when such data is available. These algorithms are very popular in sentiment classification and were among the earliest approaches to the task. Supervised learning can be further classified into the following two categories:

a: LINEAR CLASSIFIER

A linear classifier is a model that uses a linear combination of explanatory variables to classify data points into discrete categories. This approach divides the data into two spaces based on a linear boundary, with each space representing a different class label.

Two popular linear classifiers are explained next.

Support Vector Machine (SVM): SVM techniques attempt to create an ideal hyperplane that correctly divides the data points into the appropriate classes. To achieve the best generalization capability, SVM identifies the optimal trade-off between complexity and learning potential. For more than a decade, the SVM technique has proven to be successful on sentiment classification tasks [40] and is a very popular supervised machine learning classifier [41].

Generalized Linear Model (GLM): GLM is an augmentation of the traditional linear regression model that maximizes the log-likelihood to fit generalized linear models to the data [42], [43]. It is a versatile, reliable, and easy-to-understand model that offers a valuable method for modeling categorization [44].

b: PROBABILISTIC CLASSIFIER

A probabilistic generative model is a statistical method that leverages hidden random variables to model latent structure, providing a stochastic approach to data generation. These models use a predetermined dataset to perform statistical inference and calculate probability functions over the hidden variables, which are often approximated with higher-order uncertainty. Higher-order uncertainty refers to the additional uncertainty or risk that arises from considering higher-order moments beyond the mean and variance of a probability distribution.

The probabilistic-based approach involves converting text data into numerical vectors by mapping each word or document to a unique point in a high-dimensional vector space. This mapping is done in a way that captures the semantic and contextual meaning of the text. By using this vector representation, sentiment classification models can identify patterns and relationships in text data, allowing them to accurately categorize text into neutral, negative, and positive sentiment categories.

As with the linear classifier, two popular probabilistic classifiers are explained next.

Naïve Bayes Classifier (NB Classifier): Constructed on the Bayes Theorem and the simplistic independent hypotheses across features, the Naïve Bayes (NB) classifier is a straightforward stochastic-based supervised machine learning classifier. The NB classifier is predicated on the assumption of feature independence for a given class, which implies that it considers the features to be uncorrelated with one another. The NB classifier utilizes the independent feature assumption to calculate the posterior probability of a specific category based on the given text data in sentiment classification [45], [46], [47].

Ensemble Method: Ensemble learning is a technique that has greatly influenced the field of machine learning. To achieve higher levels of efficiency and generalization capacity than the individual basis learners, the ensemble learning approach trains many base learners and aggregates their predictions [48]. This technique combines the strengths of several models to enhance classification performance and is often used as the core of hybrid models [49], [50], [51]. Ensemble learning techniques can be widely categorized into three types - boosting, bagging, and stacking - each with its own unique approach to combining the base models.

2) UNSUPERVISED LEARNING

Unsupervised machine learning algorithms are used where there are no labeled data to train the classifier. These techniques rely on self-learning and have been demonstrated to be effective in the field of NLP, particularly in sentiment classification [52], [53]. The majority of the current unsupervised sentiment categorization approaches can be divided into two stages [54], [55]. In the first stage, the sentiment intensity of the text is calculated by estimating the sentiment strength of the terms and expressions used to express emotions. In the second stage, the sentiment categorization of data is achieved by referring to the sentiment strength of the data against the baseline value of '0'.

3) SEMI-SUPERVISED LEARNING

Supervised learning has shown a lot of success in many sentiment analysis tasks, but to boost the generalizability of the learning model, a substantial quantity of labeled data is required [56]. To leverage large amount of unlabeled data, unsupervised learning is presented as a feasible option. By blending unsupervised and supervised learning techniques, semi-supervised learning provides a suitable solution. Semi-supervised learning is accomplished by using the unlabeled data for unsupervised learning and the labeled data for supervised learning and then integrating both to enhance the learning model's performance [57], [58], [59].

4) DEEP LEARNING

Deep learning has ascended the emotion categorization ladder during the last decade. To overcome the challenge of handling many hidden layers in a neural network, deep learning adopts a multilayer approach. While conventional machine learning techniques rely on feature selection techniques or explicit feature extraction methods, deep learning models automatically learn and retrieve features, thereby enhancing accuracy and performance. Additionally, hyperparameters of classifier models are also automatically evaluated in most cases. Consequently, deep learning has gained widespread adoption in sentiment classification tasks [60], [61], [62], [63].

B. LEXICON-BASED APPROACH

A lexicon is an aggregation of words that are related to a specific emotional orientation. A sentiment lexicon is used in the lexicon-based approach to ascertain the polarity of a text document. In the pre-processed data, each word is identified and assigned a part of speech (POS), such as a verb, adverb, noun, pronoun, etc. This POS labeling is then utilized as a feature to extract the emotional content of a given text statement. Finally, the polarity is calculated by searching the tokenized words through any lexical resource to get their score to determine their polarity. The final score for the input data is obtained by totaling the individual scores of each word [64], [65], [66].

The lexicon-based approach can be further classified into two sub-groups explained below.

1) DICTIONARY-BASED APPROACH

The dictionary-based technique for sentiment analysis is an unsupervised method that involves utilizing a lexicon of terms to determine the sentiment of a given text [67]. This approach is rule-based and relies on a sentiment dictionary to statistically determine the sentiment weights of the words that express various emotions in the text [68]. However, the effectiveness of this approach relies on the accuracy and completeness of the pre-existing sentiment lexicons used.

2) CORPUS-BASED APPROACH

As opposed to the dictionary-based method, the corpus-based approach makes use of co-occurrence metrics or language structures in a corpus. This approach employs two different techniques:

a: STATISTICAL APPROACH

This corpus-based approach utilizes statistical methods to identify opinion seed words based on their frequency in writings with positive or negative tones. If their frequencies are equal, they are considered neutral. Contemporary techniques rely on the observation that words with similar sentiments often co-occur in corpora, and their polarity can be estimated by measuring their relative frequency of recurrence with other words in the same situation [69], [70], [71].

b: SEMANTIC APPROACH

This approach allows for semantically related phrases to be assigned similar emotional evaluations by utilizing a database of emotional words that can be recursively expanded with synonyms and antonyms. The sentiment polarity of a lexical item is ascertained by analyzing the proportional number of positive and negative counterparts for that term [72], [73], [74].

III. EVALUATION METRICS

For any growing field of research, it is necessary to establish a commonly accepted evaluation methodology that is widely used within the field. This holds true for sentiment classification as well. At present, the bulk of the studies surveyed adopt the following standard measures.

A. ACCURACY

Accuracy is a measure of the number of correct predictions, given the total number of predictions. It is calculated by taking the ratio of the number of true positives and true negatives to the total number of samples. Mathematically, the accuracy, ACC, is given by

$$\mathrm{ACC} = \frac{\mathrm{TN} + \mathrm{TP}}{\mathrm{TN} + \mathrm{TP} + \mathrm{FP} + \mathrm{FN}} \tag{1}$$

where TN is the tally of true negatives, TP is the tally of true positives, FP is the tally of false positives, and FN is the tally of false negatives.

A greater accuracy indicates that the model is better and will anticipate the text's emotion more exactly.

B. PRECISION

Precision is a measure of accuracy, indicating the extent to which the model's predictions are correct. It is the fraction of true positives to the total number of positives. Mathematically, the precision, PRE, is given by

$$\mathrm{PRE} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{2}$$

A lower value of precision indicates a higher value of false positives, and a higher value of precision indicates a lower number of false positives. Consequently, a greater precision score denotes better performance.

C. RECALL

Recall is a measure of the totality of the model's predictions, indicating the extent to which all relevant samples have been correctly classified. It is the ratio of true positive predictions to the sum of true positive and false negative predictions. Mathematically, the recall, REC, is given by

$$\mathrm{REC} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{3}$$

A higher recall means that more of the relevant samples have been correctly classified, while a lower recall indicates that there are many false negatives in the model's predictions.

D. F-MEASURE

F-measure is a metric that balances recall and precision in sentiment categorization. It is calculated as the harmonic mean of precision and recall, given by

$$\text{F1-score} = 2 \times \frac{\mathrm{PRE} \times \mathrm{REC}}{\mathrm{PRE} + \mathrm{REC}} \tag{4}$$

A bigger F1-score indicates a better performance of the model in capturing both the true positive cases and avoiding false positive cases. In other words, a big F1-score means that the model is well-balanced between correctly identifying the relevant results and having a low number of false positive results.

E. K-FOLD CROSS-VALIDATION

Machine learning models can be assessed by the resampling technique of cross-validation. In k-fold cross-validation, a limited sample of data is shuffled and split into k number of groups. Later the model is trained on a single group and tested on $k-1$ group. k-fold cross-validation is used to avoid underfitting and overfitting.

F. CONFUSION MATRIX

The performance of a model is visualized using a confusion matrix (Table 1). The various metrics that can help evaluate the model performance can be computed by analyzing the confusion matrix.

TABLE 1. Confusion matrix.

		Actual Values	
Predicted Values		Negative (0)	Positive (1)
	Negative (0)	TN	FN
	Positive (1)	FP	TP

IV. CHALLENGES IN SENTIMENT ANALYSIS

The sentiment classification approach in this paper is discussed in Sections IV-A to IV-D. These Sections cover different aspects, challenges, and their specific issues. Each Section concludes with two tables. The first table lists the challenges discussed, techniques used to overcome the challenges, and the most optimal technique. The second table lists the evaluation metrics of the most optimal techniques in the first table. Additionally, Section IV-E outlines the challenges that apply to sentiment classification regardless of the classification type.

A. CROSS-DOMAIN SENTIMENT CLASSIFICATION

While sentiment analysis techniques typically classify the sentiment of an entire text, such as a paragraph or a sentence, in practical applications, it is often required to analyze the sentiment towards multiple target entities within a piece of text. For example, the sentence "The film has great songs, but poor screenplay" is positive regarding the songs and negative regarding the screenplay. Cross-domain sentiment classification (CDSC) involves building a model that transfers knowledge from one domain to another. Without proper aspect information, it can be challenging to accurately determine the sentiment of a target phrase, which leads to a significant number of mistakes in sentiment classification [75].

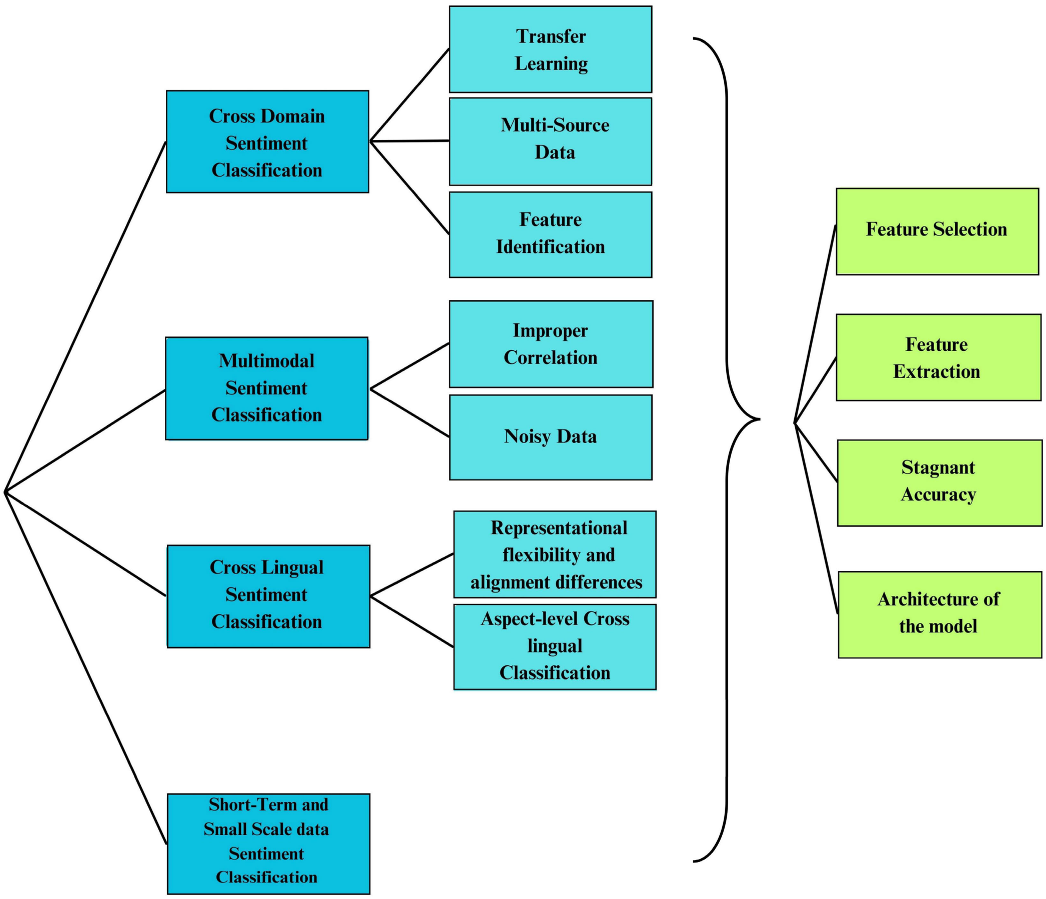

FIGURE 3. Obstacles specific to different types of sentiment classification and the shared challenges encountered by all of them.

In the past, target-dependent sentiment categorization has primarily focused on feature engineering to improve the performance of classifiers, such as SVM [75], [76], [77]. These techniques require complex feature engineering or/an substantial language riches, which are error-prone, time-consuming, and demand deep domain expertise. To address these challenges, several sentiment lexicons have been constructed [78], [79], [80], which require a significant amount of human effort. However, transferring these lexicons to another domain is a difficult task. Deep learning techniques have become widely used for these types of problems. One example is [81], which uses a neural attentive model for aspect-level sentiment classification across domains (NAACL). In NAACL, a weakly supervised latent dirichlet allocation (wsLDA) model is used to identify domain-specific aspects of a piece of text, using a multi-view attention mechanism and an aspect-level classifier. The following are the major difficulties in cross-domain sentiment classification.

1) TRANSFER OF SENTIMENT FEATURES

Cross-domain sentiment categorization requires an understanding of pivot elements and non-pivot elements.

Pivot elements are sentiment features that can be transferred across different domains, while non-pivot elements are domain-specific features that cannot be easily transferred. Pivot elements are words that have matching sentiment orientation irrespective of the domain they appear in, such as 'sad', 'happy', 'bad', and 'good'. Hence, pivot elements are used for comparing sentiments across domains. Non-pivot elements, on the other hand, are words that are domain-specific, like 'sweet', that can refer to taste, or quality of the human, or happiness. They are used for accurate sentiment classification in that particular domain and cannot be easily compared.

Therefore, it is important to understand the difference between pivot and non-pivot elements, as the use of pivot elements can increase the performance across different domains and non-pivot elements improve the accuracy in that specific domain. Fu and Liu [82] proposed a novel approach to improve cross-domain sentiment classification adopting a hierarchical attention network called KPE-net.

The KPE-net extracts the transferable pivot elements from different domains and constructs a joint attention learning network called non-KPE-net (NKPE-net) using pivot elements as a bridge. The NKPE-net captures the domain-specific non-pivot elements to enhance the sentiment classification performance. Finally, a Sentiment-Sensitive Network Model (SSNM) is built by combining both the KPE-net and NKPE-net to improve the sentiment classification accuracy across different domains. This approach addresses the challenges of transfer learning and domain adaptation in cross-domain sentiment classification, and the results from experiments establish its effectiveness when compared against other state-of-the-art methods (see Table 2).

2) MULTI-SOURCE

In many cases, data sources used for sentiment analysis are highly diverse, originating from different source domains. Traditional domain adaptation techniques are often used to bridge the gap between the target and source domains, but they are generally not effective in selecting critical domain sources and do little to mitigate the negative transfer that can result, ultimately leading to decreased model performance. To handle this concern, Fu and Liu [83] proposed a novel contrastive transformer-based domain adaptation method (CTDA). This method utilizes a mixed selection technique to choose the top-k sources and an adaptor to obtain domain-invariant information. Additionally, a discriminator is employed to extract domain-variant information. Overall, a classifier grading constituent is employed to determine the caliber of various source classifiers.

3) FEATURE IDENTIFICATION

In the sentences, 'The songs in the movie are very bad' and 'The songs are very pleasant to listen to', the word 'songs' is regarded as negative in the first expression and positive in the second expression. For such problems, deep learning models show poor performance. Tang et al. suggested using the graph domain adversarial transfer network (GDATN) in their study [84]. This approach predicts the sentiment of target domain data without labels by leveraging labeled data from a different domain. In the proposed method, features are obtained through the utilization of both a graph attention (GAT) network and bidirectional long short-term memory (BiLSTM) network. The gradient reversal layer (GRL) is then employed to identify and extract shared text features across domains. By using this method, the sentiment classification model can improve its performance in cross-domain sentiment analysis.

Table 2 shows the challenges encountered in cross-domain sentiment classification, the techniques used to address these challenges, and the most optimal of these techniques.

Table 3 presents an in-depth examination of the metrics employed to determine the most optimal technique for addressing each challenge in cross-domain sentiment classification mentioned in Table 2.

TABLE 2. Review of challenges to cross-domain sentiment classification.

Challenge	Techniques	Most Optimal Technique	Comment
Transfer of sentiment features	• DANN [85] • DAmSDA[85] • CNN-aux [85] • AMN [85] • HATN [85] • IATN [86] • CCHAN [87]	SSNM [82]	SSNM has better accuracy than the other existing methods
Multi-Source	• ASP[1], ASP[2] [88] • MAN [89] • Meta - LSTM [90] • WS-UDS, 2ST-UDA [91] • BERT [92] • Distil [93] • SDA, TOE [55]	CTDA [83]	CTDA has higher accuracy than existing methods.
Feature-Identification	• SCL [94] • DANN [95] • AMN [96] • HATN [85] • HAW, HAW+ [97]	GDATN [84]	GDATN has higher accuracy than the existing other methods

TABLE 3. Complementary analysis of Table 2.

Reference ID	Technique	Accuracy	Precision	Recall	F-measure
[82]	SSNM	✓	-	-	-
[83]	CTDA	✓	-	-	-
[84]	GDATN	✓	-	-	-

B. MULTIMODAL SENTIMENT CLASSIFICATION

Verbal communication is the primary mode of exchanging information in the real world, but people also express their feelings and opinions through non-verbal cues such as physical postures, facial gestures, and vocal intonation. The role of non-verbal communication in conveying perspectives is crucial and cannot be ignored. Therefore, comprehending the context of what someone is communicating through verbal communication alone is challenging. Likewise, extracting the sentiment of a user from texts alone is equally difficult, hence the use of multimodal sentiment classification (MSC) has widely increased. MSC makes use of multimodal information and extracts the textual, visual, and audio features to predict the sentiment.

In MSC, there are three fusion methods used: decision-level fusion, feature-level fusion, and hybrid fusion. Feature-level fusion involves obtaining characteristics from several media sources and merging them into one unified vector, which is used as the input for the classification process. Decision-level fusion, on the other hand, collects features from various media and feeds them into the classifier to generate a final vector. The hybrid fusion approach integrates decision-level and feature-level techniques, by first performing feature-level fusion between the two modalities, and then using decision-level fusion as a secondary step to integrate the results from the first stage of feature-level fusion with the remaining modality.

For visual feature extraction, convolutional neural networks (CNNs) or related models are commonly used. The output from the hidden layers of these networks is then utilized as image features for additional processing [98], [99]. Meanwhile, the widely used method for text feature extraction is bidirectional encoder representation (BERT), which utilizes a self-supervised learning approach on a large compilation to generate a distributed feature representation for words. Researchers have used different methods to integrate visual and textual features [92]. For instance, some studies use a visual feature to direct attention to an LSTM model for text feature extraction [100], while others use the cached LSTM model to extract comprehensive semantic data for an entire document [101]. Multilayer perceptron (MLP) is used in some studies to extract multimodal features [102], while an attention-based modal gated network (AMGN) is used to perform sentiment classification [103]. The AMGN uses a modal gate LSTM to integrate the text and image information, allowing for the selection of the strongest sentiment modal features.

The following are the major challenges with respect to multimodal sentiment classification.

1) IMPROPER CORRELATION

Improper correlation refers to the scenario in multimodal sentiment classification where the correlation between the different modalities used (such as textual, visual, and audio features) is not properly considered or accounted for. Concatenating modalities without proper correlation may not adequately represent properties within and across modalities. To address this issue, a cross-modal complementary network (CMCN) with hierarchical fusion has been proposed in [104] for multimodal sentiment classification. The CMCN utilizes three key modules for feature extraction in a hierarchical framework. The first module uses BERT to learn text features, while the second module uses the visual geometry group network (VGGNet) to learn image features. The third module is a cross-modal hierarchical fusion (CMHF) module, which fuses attributes from text and image modalities using joint optimization techniques to capture both the complementary features across different modalities and implicit characteristics within a single modality. The CMHF module allows for the efficient fusion of features and accounts for the proper correlation between modalities, leading to more accurate and robust sentiment classification.

2) NOISY DATA

Noisy data refers to data that contains errors or irrelevant information that can negatively affect the performance of the sentiment classification model. In the framework of multimodal sentiment classification, noisy data can arise from various sources, such as inaccuracies in the text, visual or audio data, or inconsistencies between different modalities. Convolutional neural networks (CNNs) are very popular to extricate features from images, which are then combined with text features using attention or concatenation mechanisms for multimodal sentiment classification. However, the existing methods do not fully exploit the complementary relationship between text and images or the exquisite details in the images. Moreover, these algorithms fail to consider the detrimental impact of noisy images on sentiment categorization. To combat these concerns, [105] has proposed a novel multimodal sentiment classification model that utilizes a gated attention mechanism. This gating mechanism empowers the model to selectively focus on the important image features while ignoring the noise introduced by the integration of text and image features.

Table 4 shows the challenges encountered in multi-modal sentiment classification, the techniques used to address these challenges, and the most effective/optimal of these techniques.

TABLE 4. Summary of challenges to multimodal sentiment classification.

Challenge	Techniques	Most Optimal Technique	Comment
Improper Correlation	• MFB [106] • GMU [107] • SentiBank + SentiStrength [108] • MVAN-M [109] • CFF-ATT [110] • MN-Hop2 + img2text [111] • CoMN-Hop4[111] • CoMN-Hop6[111] • HSAN [112] • DNN-LR [113] • CNN-Multi [114], and • CBOW + DA + LR [115]	CMCN [104]	CMCN has better accuracy and Weighted F1 and Macro F1 Score
Noisy Data	• TextCNN [116] • TextCNN_VGG16 [113] • BiGRU • BiGRU_VGG16 • VistaNet [114]	GAFN [105]	GAFN has better accuracy and F1 Score.

Table 5 presents an in-depth examination of the metrics employed to determine the most optimal technique for addressing each challenge in multimodal sentiment classification.

C. CROSS-LINGUAL SENTIMENT CLASSIFICATION

Sentiment analysis is challenged by the reality that data is kept in various languages due to the diverse linguistic backgrounds of people in various regions of the world. In a real-world scenario, if one were to analyze the sentiment of a specific topic based on tweets by users all over the globe, the tweets would be in several languages. Cross-lingual models are employed in these kinds of scenarios. Dealing with data

TABLE 5. Complementary analysis of Table 4.

Reference ID	Technique	Accuracy	Precision	Recall	F-measure
[104]	CMCN	✓	-	-	✓*
[105]	GAFN	✓	-	-	✓

✓* - Weighted F1 and Macro F1 Score

that spans many languages presents significant difficulties since each language has its own word resources and unique grammatical framework.

The following are the major challenges with respect to cross-lingual sentiment classification.

1) ASPECT-LEVEL CROSS-LINGUAL CLASSIFICATION

Aspect-level sentiment classification is a challenging exercise that requires establishing the sentiment inclination of each aspect in a multi-aspect phrase. This task also involves analyzing the relationships between important details and phrases, which is a critical part of aspect-based sentiment categorization. For instance, "The theatre had a huge screen, but the sound effects were awful." This phrase contains two aspects that require polarity classification: "screen" and "sound." The opinion words "awful" and "huge" are associated with the polar words "negative" and "positive", respectively.

Academic attention has been focused on developing models for aspect-level sentiment classification. In the context of supervised learning, ML algorithms are commonly employed by researchers to build classifiers [117], [118], [119]. One of the orthodox approaches is the attribute-based support vector machine (SVM) [120], [121]. Another promising technique is back-to-back neural network-based word representation learning, that utilizes an attention mechanism [122]. This approach is preferred by many researchers as it avoids the need for time-consuming feature engineering and produces high-quality results. Among these methods, the LSTM fusion attention mechanism has shown promising results in previous studies [123], [124].

BERT provides multilingual word embeddings for more than a hundred languages and is more effective in sentiment prediction than other models, especially for languages with limited resources [125]. Attention mechanisms are used among various neural networks to extract specific contexts in some studies [124], [126], [127]. Despite several research studies on aspect-based sentiment analysis, nouns, adjectives, and verbs, which are significantly related to aspects and their sentiments, are often overlooked as POS in these studies. Additionally, aspect-based sentiment classification across languages is not well-focused on homographic terms because of the limited volume of the cross-lingual data vocabulary.

In sentiment analysis, there have been several attempts to develop models that can classify sentiments in multiple languages. One such approach, described in [128], utilizes a multi-layer CNN architecture that builds upon the mono-lingual techniques outlined in [129]. Additionally, [130] attempted to categorize sentiments in Spanish, German, and French, using three distinct machine translation techniques: Google, Bing, and Moses.

Cross-lingual word embedding has been explored through an adversarial learning process using a supervised technique [131]. This approach involves mapping embeddings of the target and source languages into a shared vector space. This multilingual word approach can support up to thirty languages. However, it has been observed that the accuracy of mapping target and source languages in a common vector space decreases significantly when transitioning from resource-affluent to resource-impoverished languages, mainly as a result of insufficient data for languages with inadequate resources.

A multilingual n-gram-based approach was proposed in [132] for aspect category detection of online reviews, using multilingual word embedding to handle multilingual data. This approach divides aspect category detection into three sub-activities: aspect category detection, attribute detection, and entity detection, and operates independently of translation techniques. While one of its key advantages is its ability to recognize pre-established aspect categories, the approach is less successful in dealing with languages that have a shortage of resources, such as limited digital language resources.

By using multi-lingual BERT and bilingual dictionaries, a deep learning strategy for cross-lingual aspect-based sentiment categorization is proposed, which extracts POS tagging data [133]. The reviews are pre-processed, turned into tokens, and mapped from one language to another using bilingual dictionaries. The multilingual BERT is then utilized to generate vectors, which are then fed into the deep-learning classifier for training. The effectiveness of the method put forward is assessed using a multilingual dataset for aspect-based sentiment classification.

2) REPRESENTATIONAL FLEXIBILITY AND ALIGNMENT DIFFERENCES

Feature representation learning-based methods aim to mitigate distributional discrepancies by inducing a suitable feature characterization between the target and source language domains. Several variations of subject models have been suggested to address cross-lingual categorization issues. To achieve perfect alignment between the required domains, it has been suggested that the domains should share the same underlying subjects [134], [135], [136], [137] and to comprehend a projection matrix among several domains, it is proposed to use common subjects [138].

The techniques mentioned above suffer from an inherent flaw, namely, the requirement of a perfect layout of themes over language domains. The optimal layout is attained through either one-to-one topic layout or matrix extrapolation. However, imposing strict layout constraints can limit the representation adjustability and result in decreased model performance [139], particularly when there are significant dissimilarities in the distributions between the target and source language domains. Put differently, the presumption of perfect alignment is often erroneous since language domains often have varying underlying distributions. Therefore, it is essential to relax the constraints on perfect alignment to address cross-lingual categorization challenges.

Ref. [140] proposed a coarse alignment technique that uses group-to-group topic alignment to improve the model's representation and then fine-tune it into a fine-grained model at aspect level - aspect, opinion, and sentiment unification model (AOS), an unsupervised model that unifies aspects, opinions, and sentiments of reviews from various domains, uses coarse alignment to capture more accurate latent feature representation. To enhance AOS further, a partially supervised AOS model (ps AOS), is employed, in which tagged source language data are used in conjunction with logistical regression to diminish the variation in feature depictions across two language realms. Finally, a framework for expectation-maximization with Gibbs sampling is suggested as a way to improve the model's performance.

Table 6 shows the challenges encountered in cross-lingual sentiment classification, the techniques used to address these challenges, and the most effective/optimal of these techniques.

TABLE 6. Summary of challenges to cross-lingual sentiment classification.

Challenge	Techniques	Most Optimal Technique	Comment
Aspect-level Cross lingual classification	• LCF-BERT [133] • AEN-BERT [133] • SPC-BERT [133] • CNN-BERT without Attention [133]	CNN-BERT with Attention [133]	CNN-BERT with Attention has a better precision, recall, and f1 score.
Representational Flexibility and Alignment Difference	• LR • SVM • TRiTL [141] • DTL [142] • CL-SCL [143] • Co-Training [144] • TSU [145] • TCA [138] • PSCCLDA [136] • AOS [140]	Ps-AOS [140]	ps-AOS has better accuracy than the other models

Table 7 presents an in-depth examination of the metrics employed to determine the most optimal technique for addressing each challenge in cross-lingual classification.

D. SHORT-TERM AND SMALL-SCALE SENTIMENT CLASSIFICATION

The information exchange facilitated by the widespread use of the Internet has led to unparalleled convenience, enabling hot topics to generate massive online debates.

TABLE 7. Complementary analysis of Table 6.

Reference ID	Technique	Accuracy	Precision	Recall	F-measure
[133]	CNN-BERT with Attention	-	✓	✓	✓
[140]	Ps-AOS	✓	-	-	-

For small-scale short-term sentiment classification tasks, classifiers with smaller number of layers such as TextCNN [146], FastText [147], TextRNN [148], and TextCRNN [149] are commonly employed. These classifiers have the advantage of not needing extensive training or massive amounts of data. However, they tend to perform poorly in classification due to their limited structure, notwithstanding efforts aimed at optimization and other modifications.

Various techniques have been employed to augment data information to improve sentiment categorization in such situations. Data size expansion is a simple solution, which was achieved by introducing a hybrid neural network model that uses data expansion methods [150]. This technique can enlarge the data size, enhance the model's generalization ability, and ultimately increase its accuracy. Although intensive training on large-scale data is still required, this approach artificially increases the data scale. Another popular and practical strategy is to use multi-modal information. A hybrid classifier trained on a combination of images and text from social media was proposed instead of relying on a single input [151]. To achieve high accuracy, [152] incorporated video information on top of this foundation and extracted features from data of different modalities.

Finally, user attributes convolutional and recurrent neural networks (UCRNN), a sentiment categorization method based on text data from multi-modal social media, uses parallel RNN and CNN to analyze text information and user attributes, respectively.

E. OTHER CHALLENGES

The earlier sections focused on challenges that are specific to certain forms of sentiment classification. In addition, there are other issues that are relevant to sentiment classification in general, and these are discussed in the following subsection.

1) FEATURE SELECTION

Feature selection is a key step in improving the accuracy and reducing the training time of models used for prediction tasks. Typically, data used for prediction has numerous features, some of which may be unrelated or even damaging to the model's performance. To address this issue, three strategies have been developed for feature selection: filter, wrapper, and hybrid. Recent studies have demonstrated the advantages of the hybrid strategy for sentiment classification tasks compared to the wrapper and filter strategies [153].

The evaluation of the relevance of words to a document is a crucial aspect of sentiment analysis. One commonly used technique for this exercise is term frequency-inverse document frequency (TF-IDF). However, this method can be improved by incorporating Next Word Negation (NVM) [154]. This modification addresses common word negations like "yes," "yep," "yeah," and "sure".

To further enhance the technique, [155] integrated the chi-squared statistic selection method and used SVM as the classifier. The chi-squared statistic calculates the correlation between a word feature and its associated class or category. Later, Mohd Nafis and Awang [156] proposed a model that combines TF-IDF with support vector machine-recursive feature elimination (SVM-RFE) as an improved hybrid feature selection method. TF-IDF selects the features, and SVM-RFE ranks them in the order of their importance. Khan et al. [157] used a different hybrid strategy by integrating the wrapper-based backward feature selection (BFS) method with the ensemble of multiple filters feature selection (EMFFS) method. Another approach for feature selection is information gain (IG), in which each feature is given a weight and used for selection [158], [159]. Sparsity-adjusted information gain (SAIG) is a modified information gain (IG) algorithm for feature selection that outperforms the traditional IG algorithm in terms of accuracy and efficiency, especially in datasets with high sparsity [160].

2) FEATURE EXTRACTION

Statistical machine learning algorithms cannot be extended to text categorization issues although they are effective for less sophisticated applications of sentiment classification [161], [162]. Deep learning algorithms, on the other hand, produce noteworthy outcomes in sentiment analysis [163]. Deep learning allows CNNs to grasp intricate and non-linear structures, which enables the CNN to learn high-dimensional complexities.

However, CNN is not capable of association, and the success of the CNN model largely depends on the appropriate window size [164]. RNN is effective in learning sequential models, however, it cannot extricate local features simultaneously. Consequently, RNN can be used in conjunction with CNN.

The LSTM is an RNN-based model that aims to address the issue of the inability to achieve local feature extraction and sequential learning concurrently. RNN's structure is altered by LSTM - it transforms the RNN layer into a structure with a memory cell and a gate. The LSTM's goal is to preserve the data in the memory cell for future use and updation. The gradient exploding and disappearing difficulties in RNN [164] are resolved by LSTM with the help of this new structure. Additionally, since LSTM versions may capture extended short-term dependencies, applying them to tackle sentiment analysis issues is more promising.

In text sentiment categorization, the text is often defined as vectors in high-dimensional space. Bi-LSTM cannot emphasize crucial information while extracting context from features [165]. To overcome these limitations, a new deep-learning text classification model merging the CNN and Bi-LSTM structures has been proposed, which addresses the shortcomings of Bi-LSTM [166]. By including a convolutional layer into a CNN model, the new ConvBiLSTM structure seeks to address the restriction of Bi-LSTM. The one-dimensional convolutional layer shrinks the size of the input texts by extricating n-gram features at various sentence locations. These features are used to feed the Bi-LSTM, which extracts contextual data to categorize sentiment findings.

3) STAGNANT ACCURACY

The development of pre-trained language models (PLMs), such as ALBERT [167], BERT [92], and RoBERTa [168], has led to significant advances in various NLP applications, inclusive of document-level sentiment categorization. However, recent research has focused on improving text modeling further by incorporating user information into these models. While some studies have added user identity to traditional models, little investigation has been done on the combination of PLMs and user identity for enhanced attainment.

To deal with this problem, Cao et al. [169] proposed a new approach called user-enhanced pre-trained language models, which combines user identification with PLMs (U-PLMs). Two strategies, attention-based personalization, and embedding-based personalization were used to personalize the PLMs by incorporating user identification into different parts of the model. By injecting user identification into various aspects of the PLMs, the U-PLMs are enabled to achieve personalized text modeling from multiple perspectives, leading to improved performance.

4) ARCHITECTURE OF THE MODEL

The advancement of technology has led to the extensive use of deep learning methods for sentiment categorization. However, the increasing complexity of deep learning models has led to longer training times, which is a drawback for real-time applications that require computational time to be minimum. To address this issue, [170] put forward an elementary deep-learning architecture that uses a single-layered Bi-LSTM. Khasanah [171] also proposed two sentiment categorization models with a straightforward design, including a one-layer CNN model with fastText embedding and a BiGRU model. The study demonstrates that the CNN model can deliver comparable outcomes when used instead of the BiLSTM and BiGRU models, and also shows that the single-layer Bi-LSTM model can perform better when incorporated with fastText embedding.

V. CONCLUSION

In conclusion, sentiment analysis has grown into a crucial domain of research in natural language processing owing to its vast areas of application. Unlike many other survey papers in the field, this study adopted a unique approach of not limiting itself to one particular kind of sentiment classification or technique but instead focused on the kinds of sentiment classification that are best suited for different types of data. This approach allowed us to present a more comprehensive and nuanced understanding of sentiment analysis. The paper discussed two major approaches to sentiment analysis - machine learning-based and lexicon-based approaches. Standardized evaluation metrics were also presented to ensure consistency and comparability across studies. The central focus of our survey was to address the challenges that arise during the sentiment analysis of various kinds of data and to provide the latest overview of the ongoing research in sentiment analysis aimed at overcoming these challenges. The survey discusses challenges in four different types of sentiment classification - cross-domain sentiment classification, multimodal sentiment classification, cross-lingual sentiment classification, and short-term small-scale sentiment classification.

The main challenge of cross-domain sentiment classification (CDSC) is developing a model that would classify the sentiment across different domains. The major sub-challenges in CDSC discussed in this study are the classification and transfer of sentiment features, multisource data, and sentiment prediction of the target domain. The principal difficulty of multimodal sentiment classification (MSC) is developing a model that uses multimodal data to categorize the sentiment. The major sub-challenges in MSC discussed in this study are improper correlation and handling noisy data.

The main difficulty of cross-lingual sentiment classification (CLSC) is developing a model that uses data from different languages. The major sub-challenges of CLSC discussed in this study are aspect-level classification of cross-lingual data, representational flexibility, and alignment difference among different languages. The main challenge of short-term and small-scale sentiment classification is developing models that can classify sentiment in real-time with smaller scale data in a shorter time.

Additionally, the paper addressed the challenges associated with most of the sentiment classification techniques regardless of the type of data available. These challenges included feature selection, feature extraction, stagnant accuracy, and complex model architecture. Overall, this survey paper provided an in-depth analysis of sentiment analysis, its challenges, and the state-of-the-art in the field. The findings of this survey paper can be used as a reference for researchers and practitioners in sentiment analysis to overcome the challenges and improve the accuracy of their models. Further investigations in sentiment analysis is needed to address the challenges and make sentiment analysis more reliable and effective in different applications.

REFERENCES

[1] Statista. (2020). *Total Data Volume Worldwide 2010–2025*. [Online]. Available: https://www.statista.com/statistics/871513/worldwide-data-created/

[2] DataReportal—Global Digital Insights. (2018). *Digital Around the World*. [Online]. Available: https://datareportal.com/global-digital-overview

[3] D. Chaffey. (Jun. 2022). *Global Social Media Statistics Research Summary 2022 [June 2022]*. Smart Insights. [Online]. Available: https://www.smartinsights.com/social-media-marketing/social-media-strategy/new-global-social-media-research/

[4] O. Octaria, D. Manongga, A. Iriani, H. D. Purnomo, and I. Setyawan, ''Mining opinion based on tweets about student exchange with tweepy and TextBlob,'' in *Proc. 9th Int. Conf. Inf. Technol., Comput., Electr. Eng. (ICITACEE)*, Semarang, Indonesia, Aug. 2022, pp. 102–106, doi: 10.1109/ICITACEE55701.2022.9924013.

[5] V. Jagadishwari and N. Shobha, ''Sentiment analysis of COVID 19 vaccines using Twitter data,'' in *Proc. 2nd Int. Conf. Artif. Intell. Smart Energy (ICAIS)*, Feb. 2022, pp. 1121–1125, doi: 10.1109/ICAIS53314.2022.9742995.

[6] P. Vyas, M. Reisslein, B. P. Rimal, G. Vyas, G. P. Basyal, and P. Muzumdar, ''Automated classification of societal sentiments on Twitter with machine learning,'' *IEEE Trans. Technol. Soc.*, vol. 3, no. 2, pp. 100–110, Jun. 2022, doi: 10.1109/TTS.2021.3108963.

[7] F. Arias, M. Z. Nuñez, A. Guerra-Adames, N. Tejedor-Flores, and M. Vargas-Lombardo, ''Sentiment analysis of public social media as a tool for health-related topics,'' *IEEE Access*, vol. 10, pp. 74850–74872, 2022, doi: 10.1109/ACCESS.2022.3187406.

[8] L. Liu, Z. Cao, P. Zhao, P. J. Hu, D. D. Zeng, and Y. Luo, ''A deep learning approach for semantic analysis of COVID-19-related stigma on social media,'' *IEEE Trans. Computat. Social Syst.*, vol. 10, no. 1, pp. 246–254, Feb. 2023, doi: 10.1109/TCSS.2022.3145404.

[9] N. Seddari, A. Derhab, M. Belaoued, W. Halboob, J. Al-Muhtadi, and A. Bouras, ''A hybrid linguistic and knowledge-based analysis approach for fake news detection on social media,'' *IEEE Access*, vol. 10, pp. 62097–62109, 2022, doi: 10.1109/ACCESS.2022.3181184.

[10] S. Ni, J. Li, and H. Kao, ''MVAN: Multi-view attention networks for fake news detection on social media,'' *IEEE Access*, vol. 9, pp. 106907–106917, 2021, doi: 10.1109/ACCESS.2021.3100245.

[11] Grand View Research. (2020). *E-Commerce Market Share, Growth & Trends Report 2020–2027*. [Online]. Available: https://www.grandviewresearch.com/industry-analysis/e-commerce-market

[12] S. Maurya and V. Pratap, ''Sentiment analysis on Amazon product reviews,'' in *Proc. Int. Conf. Mach. Learn., Big Data, Cloud Parallel Comput. (COM-IT-CON)*, vol. 1. Faridabad, India, May 2022, pp. 236–240, doi: 10.1109/COM-IT-CON54601.2022.9850758.

[13] H. He, G. Zhou, and S. Zhao, ''Exploring e-commerce product experience based on fusion sentiment analysis method,'' *IEEE Access*, vol. 10, pp. 110248–110260, 2022, doi: 10.1109/ACCESS.2022.3214752.

[14] K. Chitra, A. Tamilarasi, S. G. Dharani, P. Keerthana, and T. Madhumitha, ''Opinion mining and sentiment analysis on product reviews,'' in *Proc. Int. Conf. Comput. Commun. Informat. (ICCCI)*, Coimbatore, India, Jan. 2022, pp. 1–7, doi: 10.1109/ICCCI54379.2022.9740777.

[15] C. Sindhu, S. Thejaswin, S. Harikrishnaa, and C. Kavitha, ''Mapping distinct source and target domains on Amazon product customer critiques with cross domain sentiment analysis,'' in *Proc. 2nd Int. Conf. Artif. Intell. Smart Energy (ICAIS)*, Coimbatore, India, Feb. 2022, pp. 782–786, doi: 10.1109/ICAIS53314.2022.9742732.

[16] S. Rathor and Y. Prakash, ''Application of machine learning for sentiment analysis of movies using IMDB rating,'' in *Proc. IEEE 11th Int. Conf. Commun. Syst. Netw. Technol. (CSNT)*, Indore, India, Apr. 2022, pp. 196–199, doi: 10.1109/CSNT54456.2022.9787663.

[17] B. Manikandan and S. Chakaravarthi, ''An approach of sentiment analysis for movie reviews,'' in *Proc. Int. Conf. Commun., Comput. Internet Things (IC3IoT)*, Chennai, India, Mar. 2022, pp. 1–4, doi: 10.1109/IC3IOT53935.2022.9767915.

[18] K. Soni, P. Yadav, and M. Rahul, ''Comparative analysis of rotten tomatoes movie reviews using sentiment analysis,'' in *Proc. 6th Int. Conf. Intell. Comput. Control Syst. (ICICCS)*, Madurai, India, May 2022, pp. 1494–1500, doi: 10.1109/ICICCS53718.2022.9788287.

[19] A. Gehlor and R. Singh, ''Estimation of sentiment analysis base stock market crisis,'' in *Proc. IEEE 3rd Global Conf. Advancement Technol. (GCAT)*, Bangalore, India, Oct. 2022, pp. 1–5, doi: 10.1109/GCAT55367.2022.9972059.

[20] F. Alzazah, X. Cheng, and X. Gao, ''Predict market movements based on the sentiment of financial video news sites,'' in *Proc. IEEE 16th Int. Conf. Semantic Comput. (ICSC)*, Laguna Hills, CA, USA, Jan. 2022, pp. 103–110, doi: 10.1109/ICSC52841.2022.00022.

[21] A. Nabil and N. Magdi, "A new model for stock market predication using a three-layer long short-term memory," in *Proc. 2nd Int. Mobile, Intell., Ubiquitous Comput. Conf. (MIUCC)*, Cairo, Egypt, May 2022, pp. 421–424, doi: 10.1109/MIUCC55081.2022.9781741.

[22] S. Sridhar and S. Sanagavarapu, "Analysis of the effect of news sentiment on stock market prices through event embedding," in *Proc. 16th Conf. Comput. Sci. Intell. Syst. (FedCSIS)*, Sofia, Bulgaria, Sep. 2021, pp. 147–150, doi: 10.15439/2021F79.

[23] T. T. Dang, Y. Cheng, and K. Hawick, "Market-aware sentiment analysis for stock microblogs," in *Proc. 26th Int. Conf. Autom. Comput. (ICAC)*, Portsmouth, U.K., Sep. 2021, pp. 1–6, doi: 10.23919/ICAC50006.2021.9594230.

[24] Y. Mehta, A. Malhar, and R. Shankarmani, "Stock price prediction using machine learning and sentiment analysis," in *Proc. 2nd Int. Conf. Emerg. Technol. (INCET)*, Belagavi, India, May 2021, pp. 1–4, doi: 10.1109/INCET51464.2021.9456376.

[25] A. Costa and A. Veloso, "Employee analytics through sentiment analysis," in *Proc. Brazilian Symp. Databases*, Oct. 2015, pp. 1–15, doi: 10.13140/RG.2.1.1623.3688.

[26] T. D. Chungade and S. Kharat, "Employee performance assessment in virtual organization using domain-driven data mining and sentiment analysis," in *Proc. Int. Conf. Innov. Inf., Embedded Commun. Syst. (ICIIECS)*, Coimbatore, India, Mar. 2017, pp. 1–7, doi: 10.1109/ICIIECS.2017.8276093.

[27] R. Liu, Y. Shi, C. Ji, and M. Jia, "A survey of sentiment analysis based on transfer learning," *IEEE Access*, vol. 7, pp. 85401–85412, 2019, doi: 10.1109/ACCESS.2019.2925059.

[28] F. Alattar and K. Shaalan, "A survey on opinion reason mining and interpreting sentiment variations," *IEEE Access*, vol. 9, pp. 39636–39655, 2021, doi: 10.1109/ACCESS.2021.3063921.

[29] K. Chakraborty, S. Bhattacharyya, and R. Bag, "A survey of sentiment analysis from social media data," *IEEE Trans. Computat. Social Syst.*, vol. 7, no. 2, pp. 450–464, Apr. 2020, doi: 10.1109/TCSS.2019.2956957.

[30] J. Zhou, J. X. Huang, Q. Chen, Q. V. Hu, T. Wang, and L. He, "Deep learning for aspect-level sentiment classification: Survey, vision, and challenges," *IEEE Access*, vol. 7, pp. 78454–78483, 2019, doi: 10.1109/ACCESS.2019.2920075.

[31] K. Schouten and F. Frasincar, "Survey on aspect-level sentiment analysis," *IEEE Trans. Knowl. Data Eng.*, vol. 28, no. 3, pp. 813–830, Mar. 2016, doi: 10.1109/TKDE.2015.2485209.

[32] A. Nazir, Y. Rao, L. Wu, and L. Sun, "Issues and challenges of aspect-based sentiment analysis: A comprehensive survey," *IEEE Trans. Affect. Comput.*, vol. 13, no. 2, pp. 845–863, Apr. 2022, doi: 10.1109/TAFFC.2020.2970399.

[33] P. K. Soni and R. Rambola, "A survey on implicit aspect detection for sentiment analysis: Terminology, issues, and scope," *IEEE Access*, vol. 10, pp. 63932–63957, 2022, doi: 10.1109/ACCESS.2022.3183205.

[34] H. Liu, I. Chatterjee, M. Zhou, X. S. Lu, and A. Abusorrah, "Aspect-based sentiment analysis: A survey of deep learning methods," *IEEE Trans. Computat. Social Syst.*, vol. 7, no. 6, pp. 1358–1375, Dec. 2020, doi: 10.1109/TCSS.2020.3033302.

[35] K. Dhola and M. Saradva, "A comparative evaluation of traditional machine learning and deep learning classification techniques for sentiment analysis," in *Proc. 11th Int. Conf. Cloud Comput., Data Sci. Eng.*, Noida, India, Jan. 2021, pp. 932–936, doi: 10.1109/Confluence51648.2021.9377070.

[36] A. Rawat, H. Maheshwari, M. Khanduja, R. Kumar, M. Memoria, and S. Kumar, "Sentiment analysis of COVID19 vaccines tweets using NLP and machine learning classifiers," in *Proc. Int. Conf. Mach. Learn., Big Data, Cloud Parallel Comput.*, vol. 1. Faridabad, India, May 2022, pp. 225–230, doi: 10.1109/COM-IT-CON54601.2022.9850629.

[37] D. Guan, "Sentiment classification of financial online reviews based on machine learning algorithm," in *Proc. Int. Conf. Educ., Netw. Inf. Technol. (ICENIT)*, Liverpool, U.K., Sep. 2022, pp. 289–293, doi: 10.1109/ICENIT57306.2022.00070.

[38] A. A. Wadhe and S. S. Suratkar, "Tourist place reviews sentiment classification using machine learning techniques," in *Proc. Int. Conf. Ind. 4.0 Technol. (I4Tech)*, Pune, India, Feb. 2020, pp. 1–6, doi: 10.1109/I4Tech48345.2020.9102673.

[39] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 3rd ed. Amsterdam, The Netherlands: Elsevier, 2012.

[40] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, pp. 273–297, Apr. 1995.

[41] C. C. Aggarwal and C. Zhai, *Mining Text Data*. Cham, Switzerland: Springer, 2012.

[42] M. Goldburd, A. Khare, and C. D. Tevet, "Generalized linear models for insurance rating," in *Proc. Casualty Actuarial Soc.*, 2016, pp. 1–122.

[43] P. McCullagh, "Generalized linear models," *Eur. J. Oper. Res.*, vol. 16, no. 3, pp. 370–384, 1984.

[44] I. Trofimov and A. Genkin, "Distributed coordinate descent for generalized linear models with regularization," *Pattern Recognit. Image Anal.*, vol. 27, no. 2, pp. 349–364, Apr. 2017.

[45] M. Abbas, K. Ali, S. Memon, A. Jamali, S. Memonullah, and A. Ahmed, "Multinomial Naïve Bayes classification model for sentiment analysis," *Int. J. Comput. Sci. Netw. Secur.*, vol. 19, no. 3, p. 62, Mar. 2019, doi: 10.13140/RG.2.2.30021.40169.

[46] A. A. Farisi, Y. Sibaroni, and S. A. Faraby, "Sentiment analysis on hotel reviews using multinomial Naïve Bayes classifier," *J. Phys., Conf. Ser.*, vol. 1192, Mar. 2019, Art. no. 012024, doi: 10.1088/1742-6596/1192/1/012024.

[47] M.-R. Amini and N. Usunier, "Semi-supervised learning," in *Learning With Partially Labeled and Interdependent Data*. Cham, Switzerland: Springer, 2015, pp. 3–18, doi: 10.1007/978-3-319-15726-9_3.

[48] J. Kazmaier and J. H. van Vuuren, "The power of ensemble learning in sentiment analysis," *Expert Syst. Appl.*, vol. 187, Jan. 2022, Art. no. 115819, doi: 10.1016/j.eswa.2021.115819.

[49] 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, doi: 10.1109/ACCESS.2022.3210182.

[50] M. S. Akhtar, D. Ghosal, A. Ekbal, P. Bhattacharyya, and S. Kurohashi, "All-in-one: Emotion, sentiment and intensity prediction using a multi-task ensemble framework," *IEEE Trans. Affect. Comput.*, vol. 13, no. 1, pp. 285–297, Jan. 2022, doi: 10.1109/TAFFC.2019.2926724.

[51] N. Aslam, F. Rustam, E. Lee, P. B. Washington, and I. Ashraf, "Sentiment analysis and emotion detection on cryptocurrency related tweets using ensemble LSTM-GRU model," *IEEE Access*, vol. 10, pp. 39313–39324, 2022, doi: 10.1109/ACCESS.2022.3165621.

[52] S. M. Al-Ghuribi, S. A. Mohd Noah, and S. Tiun, "Unsupervised semantic approach of aspect-based sentiment analysis for large-scale user reviews," *IEEE Access*, vol. 8, pp. 218592–218613, 2020, doi: 10.1109/ACCESS.2020.3042312.

[53] A. Yadav, C. K. Jha, A. Sharan, and V. Vaish, "Sentiment analysis of financial news using unsupervised approach," *Proc. Comput. Sci.*, vol. 167, pp. 589–598, Jan. 2020, doi: 10.1016/j.procs.2020.03.325.

[54] J. Rothfels and J. Tibshirani, "Unsupervised sentiment classification of English movie reviews using automatic selection of positive and negative sentiment items," *CS224N-Final Project*, vol. 43, no. 2, pp. 52–56, 2010.

[55] Y. Dai, J. Liu, J. Zhang, H. Fu, and Z. Xu, "Unsupervised sentiment analysis by transferring multi-source knowledge," *Cognit. Comput.*, vol. 13, no. 5, pp. 1185–1197, Sep. 2021, doi: 10.1007/s12559-020-09792-8.

[56] A. García-Pablos, M. Cuadros, and G. Rigau, "W2VLDA: Almost unsupervised system for aspect based sentiment analysis," *Expert Syst. Appl.*, vol. 96, pp. 229–241, Jan. 2018, doi: 10.1016/j.eswa.2017.08.049.

[57] J. J. E. Macrohon, C. N. Villavicencio, X. A. Inbaraj, and J.-H. Jeng, "A semi-supervised approach to sentiment analysis of tweets during the 2022 Philippine presidential election," *IEEE Access*, vol. 10, pp. 52217–52228, 2022.

[58] R. W. A. Caicedo, J. M. G. Soriano, and H. A. M. Sasieta, "Bootstrapping semi-supervised annotation method for potential suicidal messages," *Internet Intervent.*, vol. 28, Apr. 2022, Art. no. 100519, doi: 10.1016/j.invent.2022.100519.

[59] N. H. Cahyana, S. Saifullah, Y. Fauziah, A. S. Aribowo, and R. Drezewski, "Semi-supervised text annotation for hate speech detection using K-nearest neighbors and term frequency-inverse document frequency," *Int. J. Adv. Comput. Sci. Appl.*, vol. 13, no. 10, pp. 147–151, 2022, doi: 10.14569/IJACSA.2022.0131020.

[60] V. Tyagi, A. Kumar, and S. Das, "Sentiment analysis on Twitter data using deep learning approach," in *Proc. 2nd Int. Conf. Adv. Comput., Commun. Control Netw. (ICACCCN)*, Dec. 2020, pp. 187–190, doi: 10.1109/ICACCCN51052.2020.9362853.

[61] B. Seetharamulu, B. N. K. Reddy, and K. B. Naidu, "Deep learning for sentiment analysis based on customer reviews," in *Proc. 11th Int. Conf. Comput., Commun. Netw. Technol. (ICCCNT)*, Kharagpur, India, Jul. 2020, pp. 1–5, doi: 10.1109/ICCCNT49239.2020.9225665.

[62] A. C. Mazari and A. Djeffal, "Deep learning-based sentiment analysis of Algerian dialect during Hirak 2019," in *Proc. 2nd Int. Workshop Hum.-Centric Smart Environ. Health Well-Being (IHSH)*, Boumerdes, Algeria, Feb. 2021, pp. 233–236, doi: 10.1109/IHSH51661.2021.9378753.

[63] X. Glorot, A. Bordes, and Y. Bengio, "Domain adaptation for large-scale sentiment classification: A deep learning approach," in *Proc. 28th Int. Conf. Mach. Learn.*, vol. 2011, pp. 513–520.

[64] R. D. Tan, K. Piad, A. Lagman, J. Victoriano, I. Tano, N. S. Gabriel, and J. Espino, "LMS content evaluation system with sentiment analysis using lexicon-based approach," in *Proc. 10th Int. Conf. Inf. Educ. Technol. (ICIET)*, Matsue, Japan, Apr. 2022, pp. 93–98, doi: 10.1109/ICIET55102.2022.9778976.

[65] M. Huang, H. Xie, Y. Rao, Y. Liu, L. K. M. Poon, and F. L. Wang, "Lexicon-based sentiment convolutional neural networks for online review analysis," *IEEE Trans. Affect. Comput.*, vol. 13, no. 3, pp. 1337–1348, Jul. 2022, doi: 10.1109/TAFFC.2020.2997769.

[66] W. Suwanpipob, N. Arch-int, and M. Wattana, "A sentiment classification from review corpus using linked open data and sentiment lexicon," in *Proc. 13th Int. Conf. Inf. Technol. Electr. Eng. (ICITEE)*, Oct. 2021, pp. 19–23, doi: 10.1109/ICITEE53064.2021.9611898.

[67] S. Lee, S. Ma, J. Meng, J. Zhuang, and T.-Q. Peng, "Detecting sentiment toward emerging infectious diseases on social media: A validity evaluation of dictionary-based sentiment analysis," *Int. J. Environ. Res. Public Health*, vol. 19, no. 11, p. 6759, Jun. 2022, doi: 10.3390/ijerph19116759.

[68] E. Okango and H. Mwambi, "Dictionary based global Twitter sentiment analysis of coronavirus (COVID-19) effects and response," *Ann. Data Sci.*, vol. 9, no. 1, pp. 175–186, Feb. 2022, doi: 10.1007/s40745-021-00358-5.

[69] A. Husna, H. Rahman, and E. K. Hashi, "Statistical approach for classifying sentiment reviews by reducing dimension using truncated singular value decomposition," in *Proc. 1st Int. Conf. Adv. Sci., Eng. Robot. Technol. (ICASERT)*, May 2019, pp. 1–5, doi: 10.1109/ICASERT.2019.8934507.

[70] A. P. Kumar and A. Nayak, "A statistical approach to evaluate the efficiency and effectiveness of the machine learning algorithms analyzing sentiments," in *Proc. IEEE Int. Conf. Distrib. Comput., VLSI, Electr. Circuits Robot. (DISCOVER)*, Manipal, India, Aug. 2019, pp. 1–6, doi: 10.1109/DISCOVER47552.2019.9008028.

[71] R. Gupta, J. Kumar, and H. Agrawal, "A statistical approach for sarcasm detection using Twitter data," in *Proc. 4th Int. Conf. Intell. Comput. Control Syst. (ICICCS)*, Madurai, India, May 2020, pp. 633–638, doi: 10.1109/ICICCS48265.2020.9120917.

[72] S. Choudhary and J. Godara, "Semantic analysis on social media," in *Proc. Int. Conf. Comput. Sci. (ICCS)*, Phagwara, India, Dec. 2021, pp. 239–243, doi: 10.1109/ICCS54944.2021.00054.

[73] Y. S. Mehanna and M. B. Mahmuddin, "A semantic conceptualization using tagged bag-of-concepts for sentiment analysis," *IEEE Access*, vol. 9, pp. 118736–118756, 2021, doi: 10.1109/ACCESS.2021.3107237.

[74] G. Gautam and D. Yadav, "Sentiment analysis of Twitter data using machine learning approaches and semantic analysis," in *Proc. 7th Int. Conf. Contemp. Comput. (IC3)*, Noida, India, Aug. 2014, pp. 437–442, doi: 10.1109/IC3.2014.6897213.

[75] L. Jiang, M. Yu, M. Zhou, X. Liu, and T. Zhao, "Target-dependent Twitter sentiment classification," in *Proc. 49th Annu. Meeting Assoc. Comput. Linguistics, Human Lang. Technol.*, vol. 1, 2011, pp. 151–160.

[76] S. Kiritchenko, X. Zhu, C. Cherry, and S. Mohammad, "NRC-Canada-2014: Detecting aspects and sentiment in customer reviews," in *Proc. 8th Int. Workshop Semantic Eval.*, 2014, pp. 437–442, doi: 10.3115/v1/S14-2076.

[77] J. Wagner, P. Arora, S. Cortes, U. Barman, D. Bogdanova, J. Foster, and L. Tounsi, "DCU: Aspect-based polarity classification for SemEval task 4," in *Proc. 8th Int. Workshop Semantic Eval.*, Dublin, Ireland, 2014, pp. 223–229, doi: 10.3115/v1/S14-2036.

[78] A. Neviarouskaya, H. Prendinger, and M. Ishizuka, "SentiFul: Generating a reliable lexicon for sentiment analysis," in *Proc. 3rd Int. Conf. Affect. Comput. Intell. Interact. Workshops*, Amsterdam, The Netherlands, Sep. 2009, pp. 1–6, doi: 10.1109/ACII.2009.5349575.

[79] G. Qiu, L. Bing, J. Bu, and C. Chen, "Expanding domain sentiment lexicon through double propagation," in *Proc. Int. Joint Conf. Artif. Intell.*, 2009, pp. 1199–1204.

[80] M. Taboada, J. Brooke, M. Tofiloski, K. Voll, and M. Stede, "Lexicon-based methods for sentiment analysis," *Comput. Linguistics*, vol. 37, no. 2, pp. 267–307, Jun. 2011, doi: 10.1162/COLI_a_00049.

[81] M. Yang, W. Yin, Q. Qu, W. Tu, Y. Shen, and X. Chen, "Neural attentive network for cross-domain aspect-level sentiment classification," *IEEE Trans. Affect. Comput.*, vol. 12, no. 3, pp. 761–775, Jul. 2021, doi: 10.1109/TAFFC.2019.2897093.

[82] Y. Fu and Y. Liu, "Cross-domain sentiment classification based on key pivot and non-pivot extraction," *Knowl.-Based Syst.*, vol. 228, Sep. 2021, Art. no. 107280, doi: 10.1016/j.knosys.2021.107280.

[83] Y. Fu and Y. Liu, "Contrastive transformer based domain adaptation for multi-source cross-domain sentiment classification," *Knowl.-Based Syst.*, vol. 245, Jun. 2022, Art. no. 108649, doi: 10.1016/j.knosys.2022.108649.

[84] H. Tang, Y. Mi, F. Xue, and Y. Cao, "Graph domain adversarial transfer network for cross-domain sentiment classification," *IEEE Access*, vol. 9, pp. 33051–33060, 2021, doi: 10.1109/ACCESS.2021.3061139.

[85] Z. Li, Y. Wei, Y. Zhang, and Q. Yang, "Hierarchical attention transfer network for cross-domain sentiment classification," in *Proc. 32nd AAAI Conf. Artif. Intelligence*, 2018, vol. 32, no. 1, pp. 1–14, doi: 10.1609/aaai.v32i1.12055.

[86] K. Zhang, H. Zhang, Q. Liu, H. Zhao, H. Zhu, and E. Chen, "Interactive attention transfer network for cross-domain sentiment classification," in *Proc. 33rd AAAI Conf. Artif. Intell. 31st Innov. Appl. Artif. Intell. Conf. 9th AAAI Symp. Educ. Adv. Artif. Intell.*, Jan. 2019, pp. 5773–5780, doi: 10.1609/aaai.v33i01.33015773.

[87] T. Manshu and Z. Xuemin, "CCHAN: An end to end model for cross domain sentiment classification," *IEEE Access*, vol. 7, pp. 50232–50239, 2019, doi: 10.1109/ACCESS.2019.2910300.

[88] P. Liu, X. Qiu, and X. Huang, "Adversarial multi-task learning for text classification," in *Proc. 55th Annu. Meeting Assoc. Comput. Linguistics*, 2017, pp. 1–12, doi: 10.18653/v1/p17-1001.

[89] X. Chen and C. Cardie, "Multinomial adversarial networks for multi-domain text classification," in *Proc. Conf. North Amer. Chapter Assoc. Comput. Linguistics, Hum. Lang. Technol.*, 2018, pp. 1226–1240, doi: 10.18653/v1/n18-1111.

[90] J. Chen, X. Qiu, and X. Huang, "Meta multi-task learning for sequence modeling," in *Proc. AAAI Conf. Artif. Intell.*, Feb. 2018, vol. 32, no. 1, pp. 12007–12014, doi: 10.1609/aaai.v32i1.12007.

[91] Y. Dai, J. Liu, X. Ren, and Z. Xu, "Adversarial training based multi-source unsupervised domain adaptation for sentiment analysis," in *Proc. AAAI Conf. Artif. Intell.*, 2020, vol. 34, no. 5, pp. 7618–7625.

[92] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," 2018, *arXiv:1810.04805*.

[93] V. Sanh, L. Debut, J. Chaumond, and T. Wolf, "DistilBERT, a distilled version of BERT: Smaller, faster, cheaper and lighter," 2019, *arXiv:1910.01108*.

[94] J. Blitzer, M. Dredze, and F. Pereira, "Biographies, bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification," in *Proc. 45th Annu. Meeting Assoc. Comput. Linguistics (ACL)*, vol. 1, 2007, pp. 440–447.

[95] Y. Ganin, E. Ustinova, H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, M. Marchand, and V. Lempitsky, "Domain-adversarial training of neural networks," in *Domain Adaptation in Computer Vision Applications*, G. Csurka, Ed. Cham, Switzerland: Springer, 2017, pp. 79–87, doi: 10.1007/978-3-319-58347-1_10.

[96] Z. Li, Y. Zhang, Y. Wei, Y. Wu, and Q. Yang, "End-to-end adversarial memory network for cross-domain sentiment classification," in *Proc. 26th Int. Joint Conf. Artif. Intell.*, Aug. 2017, pp. 3897–3903, doi: 10.24963/IJCAI.2017/311.

[97] Y. Du, M. He, L. Wang, and H. Zhang, "Wasserstein based transfer network for cross-domain sentiment classification," *Knowl.-Based Syst.*, vol. 204, pp. 1–11, Sep. 2020, doi: 10.1016/j.knosys.2020.106162.

[98] G. R. Wang, K. Z. Wang, and L. Lin, "Adaptively connected neural networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Long Beach, CA, USA, Jun. 2019, pp. 1781–1790.

[99] R. Cadne, C. Dancette, H. Ben-Younes, M. Cord, and D. Parikh, "RUBi: Reducing unimodal biases for visual question answering," in *Proc. Conf. Neural Inf. Process. Syst. (NeurIPS)*, Vancouver, BC, Canada, Dec. 2019, pp. 841–852.

[100] N. Xu and W. Mao, "MultiSentiNet: A deep semantic network for multimodal sentiment analysis," in *Proc. ACM Conf. Inf. Knowl. Manage.*, Singapore, Nov. 2017, pp. 2399–2402, doi: 10.1145/3132847.3133142.

[101] J. Xu, D. Chen, X. Qiu, and X. Huang, "Cached long short-term memory neural networks for document-level sentiment classification," in *Proc. Conf. Empirical Methods Natural Lang. Process.*, Austin, TX, USA, 2016, pp. 1660–1669.

[102] C. Yang, X. Wang, and B. Jiang, "Sentiment enhanced multi-modal hashtag recommendation for micro-videos," *IEEE Access*, vol. 8, pp. 78252–78264, 2020, doi: 10.1109/ACCESS.2020.2989473.

[103] F. R. Huang, K. M. Wei, J. Weng, and Z. J. Li, "Attention-based modality-gated networks for image-text sentiment analysis," *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 16, no. 3, p. 79, 2020, doi: 10.1145/3388861.

[104] C. Peng, C. Zhang, X. Xue, J. Gao, H. Liang, and Z. Niu, "Cross-modal complementary network with hierarchical fusion for multimodal sentiment classification," *Tsinghua Sci. Technol.*, vol. 27, no. 4, pp. 664–679, Aug. 2022, doi: 10.26599/TST.2021.9010055.

[105] Y. Du, Y. Liu, Z. Peng, and X. Jin, "Gated attention fusion network for multimodal sentiment classification," *Knowl.-Based Syst.*, vol. 240, Mar. 2022, Art. no. 108107, doi: 10.1016/j.knosys.2021.108107.

[106] Z. Yu, J. Yu, J. Fan, and D. Tao, "Multi-modal factorized bilinear pooling with co-attention learning for visual question answering," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Venice, Italy, Oct. 2017, pp. 1839–1848, doi: 10.1109/ICCV.2017.202.

[107] J. Arevalo, T. Solorio, M. Montes-y-Gómez, and F. A. González, "Gated multimodal units for information fusion," 2017, *arXiv:1702.01992*.

[108] D. Borth, R. Ji, T. Chen, T. Breuel, and S.-F. Chang, "Large-scale visual sentiment ontology and detectors using adjective noun pairs," in *Proc. 21st ACM Int. Conf. Multimedia*, Oct. 2013, pp. 223–232, doi: 10.1145/2502081.2502282.

[109] X. Yang, S. Feng, D. Wang, and Y. Zhang, "Image-text multimodal emotion classification via multi-view attentional network," *IEEE Trans. Multimedia*, vol. 23, pp. 4014–4026, 2021, doi: 10.1109/TMM.2020.3035277.

[110] K. Zhang, Y. Geng, J. Zhao, J. Liu, and W. Li, "Sentiment analysis of social media via multimodal feature fusion," *Symmetry*, vol. 12, no. 12, p. 2010, Dec. 2020, doi: 10.3390/sym12122010.

[111] N. Xu, W. Mao, and G. Chen, "A co-memory network for multimodal sentiment analysis," in *Proc. 41st Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, Jun. 2018, pp. 929–932, doi: 10.1145/3209978.3210093.

[112] N. Xu, "Analyzing multimodal public sentiment based on hierarchical semantic attentional network," in *Proc. IEEE Int. Conf. Intell. Secur. Informat. (ISI)*, Beijing, China, Jul. 2017, pp. 152–154, doi: 10.1109/ISI.2017.8004895.

[113] Y. Yu, H. Lin, J. Meng, and Z. Zhao, "Visual and textual sentiment analysis of a microblog using deep convolutional neural networks," *Algorithms*, vol. 9, no. 2, p. 41, Jun. 2016, doi: 10.3390/a9020041.

[114] G. Cai and B. Xia, "Convolutional neural networks for multimedia sentiment analysis," in *Natural Language Processing and Chinese Computing*, J. Li, H. Ji, D. Zhao, and Y. Feng, Eds. Cham, Switzerland: Springer, 2015, pp. 149–158, doi: 10.1007/978-3-319-25207-0_14.

[115] C. Baecchi, T. Uricchio, M. Bertini, and A. D. Bimbo, "A multimodal feature learning approach for sentiment analysis of social network multimedia," *Multimedia Tools Appl.*, vol. 75, no. 5, pp. 2507–2525, Mar. 2016, doi: 10.1007/s11042-015-2646-x.

[116] D. Hazarika, S. Poria, R. Mihalcea, E. Cambria, and R. Zimmermann, "ICON: Interactive conversational memory network for multimodal emotion detection," in *Proc. Conf. Empirical Methods Natural Lang. Process.*, 2018, pp. 2594–2604, doi: 10.18653/v1/D18-1280.

[117] A. Go, R. Bhayani, and L. Huang, "Twitter sentiment classification using distant supervision," *Stanford*, vol. 1, no. 12, p. 2009, 2009. [Online]. Available: https://cs.stanford.edu/people/alecmgo/papers/TwitterDistantSupervision09.pdf

[118] B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up? Sentiment classification using machine learning techniques," in *Proc. Conf. Empirical Methods Natural Lang. Process.*, 2002, pp. 79–86, doi: 10.3115/1118693.1118704.

[119] A. Kennedy and D. Inkpen, "Sentiment classification of movie reviews using contextual valence shifters," *Comput. Intell.*, vol. 22, no. 2, pp. 110–125, May 2006, doi: 10.1111/j.1467-8640.2006.00277.x.

[120] S. Mohammad, S. Kiritchenko, and X. Zhu, "NRC-Canada: Building the state-of-the-art in sentiment analysis of tweets," in *Proc. 2nd Joint Conf. Lexical Comput. Semantics, 7th Int. Workshop Semantic Eval.*, Atlanta, GA, USA, Jun. 2013, pp. 321–327.

[121] S. Liu, F. Li, F. Li, X. Cheng, and H. Shen, "Adaptive co-training SVM for sentiment classification on tweets," in *Proc. 22nd ACM Int. Conf. Conf. Inf. Knowl. Manage.*, 2013, pp. 2079–2088, doi: 10.1145/2505515.2505569.

[122] H. T. Nguyen and M. Le Nguyen, "Effective attention networks for aspect-level sentiment classification," in *Proc. 10th Int. Conf. Knowl. Syst. Eng. (KSE)*, Ho Chi Minh City, Vietnam, Nov. 2018, pp. 25–30, doi: 10.1109/KSE.2018.8573324.

[123] Y. Tay, L. A. Tuan, and S. C. Hui, "Learning to attend via word-aspect associative fusion for aspect-based sentiment analysis," in *Proc. 32nd AAAI Conf. Artif. Intell. 30th Innov. Appl. Artif. Intell. Conf. 8th AAAI Symp. Educ. Adv. Artif. Intell.*, Feb. 2018, pp. 5956–5963.

[124] D. Ma, S. Li, X. Zhang, and H. Wang, "Interactive attention networks for aspect-level sentiment classification," in *Proc. 26th Int. Joint Conf. Artif. Intell.*, Aug. 2017, pp. 4068–4074, doi: 10.24963/ijcai.2017/568.

[125] K. Karthikeyan, Z. Wang, S. Mayhew, and D. Roth, "Cross-lingual ability of multilingual BERT: An empirical study," in *Proc. Int. Conf. Learn. Represent.*, 2019, pp. 1–12.

[126] Y. Wang, M. Huang, X. Zhu, and L. Zhao, "Attention-based LSTM for aspect-level sentiment classification," in *Proc. Conf. Empirical Methods Natural Lang. Process.*, 2016, pp. 606–615, doi: 10.18653/v1/d16-1058.

[127] Y. Song, J. Wang, T. Jiang, Z. Liu, and Y. Rao, "Targeted sentiment classification with attentional encoder network," in *Proc. ICANN*, vol. 11730. Cham, Switzerland: Springer, Sep. 2019, pp. 82–91, doi: 10.1007/978-3-030-30490-4_9.

[128] J. Deriu, A. Lucchi, V. De Luca, A. Severyn, S. Müller, M. Cieliebak, T. Hofmann, and M. Jaggi, "Leveraging large amounts of weakly supervised data for multi-language sentiment classification," in *Proc. 26th Int. Conf. World Wide Web*, Apr. 2017, pp. 1045–1052, doi: 10.1145/3038912.3052611.

[129] K. Dashtipour, S. Poria, A. Hussain, E. Cambria, A. Y. A. Hawalah, A. Gelbukh, and Q. Zhou, "Multilingual sentiment analysis: State of the art and independent comparison of techniques," *Cognit. Comput.*, vol. 8, no. 4, pp. 757–771, Aug. 2016, doi: 10.1007/s12559-016-9415-7.

[130] A. Balahur and M. Turchi, "Multilingual sentiment analysis using machine translation," in *Proc. 3rd Workshop Comput. Approaches Subjectivity, Sentiment Social Media Anal. (WASSA)*, Stroudsburg, PA, USA: ACM, 2012, pp. 52–60.

[131] A. Conneau, G. Lample, M. Ranzato, L. Denoyer, and H. Jégou, "Word translation without parallel data," 2017, *arXiv:1710.04087*.

[132] E. Ghadery, S. Movahedi, H. Faili, and A. Shakery, "MNCN: A multilingual Ngram-based convolutional network for aspect category detection in online reviews," in *Proc. AAAI Conf. Artif. Intell.*, vol. 33, 2019, pp. 6441–6448, doi: 10.1609/aaai.v33i01.33016441.

[133] K. Sattar, Q. Umer, D. G. Vasbieva, S. Chung, Z. Latif, and C. Lee, "A multi-layer network for aspect-based cross-lingual sentiment classification," *IEEE Access*, vol. 9, pp. 133961–133973, 2021, doi: 10.1109/ACCESS.2021.3116053.

[134] Z. Lin, X. Jin, X. Xu, W. Wang, X. Cheng, and Y. Wang, "A cross-lingual joint aspect/sentiment model for sentiment analysis," in *Proc. 23rd ACM Int. Conf. Conf. Inf. Knowl. Manage.*, Nov. 2014, pp. 1089–1098, doi: 10.1145/2661829.2662019.

[135] M. Paul and R. Girju, "Cross-cultural analysis of blogs and forums with mixed-collection topic models," in *Proc. Conf. Empirical Methods Natural Lang. Process.*, 2009, pp. 1408–1417.

[136] Y. Bao, N. Collier, and A. Datta, "A partially supervised cross-collection topic model for cross-domain text classification," in *Proc. 22nd ACM Int. Conf. Conf. Inf. Knowl. Manage.*, 2013, pp. 239–248, doi: 10.1145/2505515.2505556.

[137] F. Zhuang, P. Luo, Z. Shen, Q. He, Y. Xiong, Z. Shi, and H. Xiong, "Collaborative dual-PLSA: Mining distinction and commonality across multiple domains for text classification," in *Proc. 19th ACM Int. Conf. Inf. Knowl. Manage.*, Oct. 2010, pp. 359–368, doi: 10.1145/1871437.1871486.

[138] L. Li, X. Jin, and M. Long, "Topic correlation analysis for cross-domain text classification," in *Proc. 26th AAAI Conf. Artif. Intell.*, Jul. 2012, pp. 998–1004.

[139] G. Zhou, T. He, J. Zhao, and W. Wu, "A subspace learning framework for cross-lingual sentiment classification with partial parallel data," in *Proc. IJCAI*, 2015, pp. 1426–1433.

[140] D. Wang, B. Jing, C. Lu, J. Wu, G. Liu, C. Du, and F. Zhuang, "Coarse alignment of topic and sentiment: A unified model for cross-lingual sentiment classification," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 2, pp. 736–747, Feb. 2021, doi: 10.1109/TNNLS.2020.2979225.

[141] F. Zhuang, P. Luo, C. Du, Q. He, Z. Shi, and H. Xiong, ''Triplex transfer learning: Exploiting both shared and distinct concepts for text classification,'' *IEEE Trans. Cybern.*, vol. 44, no. 7, pp. 1191–1203, Jul. 2014, doi: 10.1109/TCYB.2013.2281451.

[142] M. Long, J. Wang, G. Ding, W. Cheng, X. Zhang, and W. Wang, ''Dual transfer learning,'' in *Proc. SIAM Int. Conf. Data Mining*, Apr. 2012, pp. 47–58, doi: 10.1137/1.9781611972825.47.

[143] P. Prettenhofer and B. Stein, ''Cross-lingual adaptation using structural correspondence learning,'' *ACM Trans. Intell. Syst. Technol.*, vol. 3, no. 1, pp. 13:1–13:22, Oct. 2011, doi: 10.1145/2036264.2036277.

[144] X. Wan, ''Co-training for cross-lingual sentiment classification,'' in *Proc. Joint Conf. 47th Annu. Meeting ACL 4th Int. Joint Conf. Natural Lang. Process.*, 2009, pp. 235–243.

[145] C. Ma, M. Wang, and X. Chen, ''Topic and sentiment unification maximum entropy model for online review analysis,'' in *Proc. 24th Int. Conf. World Wide Web*, May 2015, pp. 649–654, doi: 10.1145/2740908.2741704.

[146] M. Xuanyuan, L. Xiao, and M. Duan, ''Sentiment classification algorithm based on multi-modal social media text information,'' *IEEE Access*, vol. 9, pp. 33410–33418, 2021, doi: 10.1109/ACCESS.2021.3061450.

[147] A. Joulin, E. Grave, P. Bojanowski, and T. Mikolov, ''Bag of tricks for efficient text classification,'' in *Proc. 15th Conf. Eur. Chapter Assoc. Comput. Linguistics*, Valencia, Spain, 2017, pp. 427–431.

[148] F. Li, M. Zhang, G. Fu, T. Qian, and D. Ji, ''A Bi-LSTM-RNN model for relation classification using low-cost sequence features,'' 2016, *arXiv:1608.07720*.

[149] R. Wang, Z. Li, J. Cao, T. Chen, and L. Wang, ''Convolutional recurrent neural networks for text classification,'' in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Budapest, Hungary, Jul. 2019, pp. 1–6, doi: 10.1109/IJCNN.2019.8852406.

[150] X. Sun and J. He, ''A novel approach to generate a large scale of supervised data for short text sentiment analysis,'' *Multimedia Tools Appl.*, vol. 79, nos. 9–10, pp. 5439–5459, Feb. 2018, doi: 10.1007/s11042-018-5748-4.

[151] A. Kumar, K. Srinivasan, W.-H. Cheng, and A. Y. Zomaya, ''Hybrid context enriched deep learning model for fine-grained sentiment analysis in textual and visual semiotic modality social data,'' *Inf. Process. Manage.*, vol. 57, no. 1, Jan. 2020, Art. no. 102141, doi: 10.1016/j.ipm.2019.102141.

[152] S. Bairavel and M. Krishnamurthy, ''Novel OGBEE-based feature selection and feature-level fusion with MLP neural network for social media multimodal sentiment analysis,'' *Soft Comput.*, vol. 24, no. 24, pp. 18431–18445, Dec. 2020, doi: 10.1007/s00500-020-05049-6.

[153] T. Sabbah, M. Ayyash, and M. Ashraf, ''Hybrid support vector machine-based feature selection method for text classification,'' *Int. Arab J. Inf. Technol.*, vol. 15, no. 3, pp. 599–609, 2018.

[154] L. Dey, S. Chakraborty, A. Biswas, B. Bose, and S. Tiwari, ''Sentiment analysis of review datasets using Naïve Bayes- and K-NN classifier,'' *Int. J. Inf. Eng. Electron. Bus.*, vol. 8, no. 4, pp. 54–62, Jul. 2016, doi: 10.5815/ijieeb.2016.04.07.

[155] U. I. Larasati, M. A. Muslim, R. Arifudin, and A. Alamsyah, ''Improve the accuracy of support vector machine using chi square statistic and term frequency inverse document frequency on movie review sentiment analysis,'' *Sci. J. Informat.*, vol. 6, no. 1, pp. 138–149, May 2019, doi: 10.15294/sji.v6i1.14244.

[156] N. S. Mohd Nafis and S. Awang, ''An enhanced hybrid feature selection technique using term frequency-inverse document frequency and support vector machine-recursive feature elimination for sentiment classification,'' *IEEE Access*, vol. 9, pp. 52177–52192, 2021, doi: 10.1109/ACCESS.2021.3069001.

[157] J. Khan, A. Alam, and Y. Lee, ''Intelligent hybrid feature selection for textual sentiment classification,'' *IEEE Access*, vol. 9, pp. 140590–140608, 2021, doi: 10.1109/ACCESS.2021.3118982.

[158] A. Pratiwi and I. Adiwijaya, ''On the feature selection and classification based on information gain for document sentiment analysis,'' *Appl. Comput. Intell. Soft Comput.*, vol. 2018, pp. 1–5, Jan. 2018, doi: 10.1155/2018/1407817.

[159] R. Maulana, P. A. Rahayuningsih, W. Irmayani, D. Saputra, and W. E. Jayanti, ''Improved accuracy of sentiment analysis movie review using support vector machine based information gain,'' *J. Phys., Conf. Ser.*, vol. 1641, no. 1, Nov. 2020, Art. no. 012060, doi: 10.1088/1742-6596/1641/1/012060.

[160] B. Y. Ong, S. W. Goh, and C. Xu, ''Sparsity adjusted information gain for feature selection in sentiment analysis,'' in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Santa Clara, CA, USA, Oct. 2015, pp. 2122–2128, doi: 10.1109/BIGDATA.2015.7363995.

[161] Q. Huang, R. Chen, X. Zheng, and Z. Dong, ''Deep sentiment representation based on CNN and LSTM,'' in *Proc. Int. Conf. Green Informat. (ICGI)*, Aug. 2017, pp. 30–33, doi: 10.1109/ICGI.2017.45.

[162] M. S. Neethu and R. Rajasree, ''Sentiment analysis in Twitter using machine learning techniques,'' in *Proc. 4th Int. Conf. Comput., Commun. Netw. Technol. (ICCCNT)*, Tiruchengode, India, Jul. 2013, pp. 1–5, doi: 10.1109/ICCCNT.2013.6726818.

[163] J. Schmidhuber, ''Deep learning in neural networks: An overview,'' *Neural Netw.*, vol. 61, pp. 85–117, Jan. 2015, doi: 10.1016/j.neunet.2014.09.003.

[164] G. Xu, Y. Meng, X. Qiu, Z. Yu, and X. Wu, ''Sentiment analysis of comment texts based on BiLSTM,'' *IEEE Access*, vol. 7, pp. 51522–51532, 2019, doi: 10.1109/ACCESS.2019.2909919.

[165] G. Liu and J. Guo, ''Bidirectional LSTM with attention mechanism and convolutional layer for text classification,'' *Neurocomputing*, vol. 337, pp. 325–338, Apr. 2019, doi: 10.1016/j.neucom.2019.01.078.

[166] S. Tam, R. B. Said, and Ö. Ö. Tanriöver, ''A ConvBiLSTM deep learning model-based approach for Twitter sentiment classification,'' *IEEE Access*, vol. 9, pp. 41283–41293, 2021, doi: 10.1109/ACCESS.2021.3064830.

[167] P. Chi, P. Chung, T. Wu, C. Hsieh, Y. Chen, S. Li, and H. Lee, ''Audio Albert: A lite BERT for self-supervised learning of audio representation,'' in *Proc. IEEE Spoken Lang. Technol. Workshop (SLT)*, Jan. 2021, pp. 344–350, doi: 10.1109/SLT48900.2021.9383575.

[168] Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, and V. Stoyanov, ''RoBERTa: A robustly optimized BERT pretraining approach,'' 2019, *arXiv:1907.11692*.

[169] X. Cao, J. Yu, and Y. Zhuang, ''Injecting user identity into pretrained language models for document-level sentiment classification,'' *IEEE Access*, vol. 10, pp. 30157–30167, 2022, doi: 10.1109/ACCESS.2022.3158975.

[170] Z. Hameed and B. Garcia-Zapirain, ''Sentiment classification using a single-layered BiLSTM model,'' *IEEE Access*, vol. 8, pp. 73992–74001, 2020, doi: 10.1109/ACCESS.2020.2988550.

[171] I. N. Khasanah, ''Sentiment classification using fasttext embedding and deep learning model,'' *Proc. Comput. Sci.*, vol. 181, pp. 682–688, May 2021, doi: 10.1016/j.procs.2021.05.103.

NILAA RAGHUNATHAN is currently pursuing the B.Tech. degree in computer science (data science) with the School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, India.

Her current research interests include data science, machine learning, and natural language processing.

SARAVANAKUMAR KANDASAMY received the M.Tech. degree in computer science engineering from the Indian Institute of Technology, Guwahati, India, and the Ph.D. degree in computer science and engineering from the Vellore Institute of Technology, Vellore, Tamil Nadu, India.

Since 2008, he has been with the Vellore Institute of Technology, where he is currently an Associate Professor with the Department of Software Systems. His current research interests include natural language processing, machine learning, and advanced database systems.