REVIEW

A review of sentiment analysis: tasks, applications, and deep learning techniques

Neeraj Anand Sharma[1] · A. B. M. Shawkat Ali[1] · Muhammad Ashad Kabir[2]

Abstract

Sentiment analysis, a transformative force in natural language processing, revolutionizes diverse fields such as business, social media, healthcare, and disaster response. This review delves into the intricate landscape of sentiment analysis, exploring its significance, challenges, and evolving methodologies. We examine crucial aspects like dataset selection, algorithm choice, language considerations, and emerging sentiment tasks. The suitability of established datasets (e.g., IMDB Movie Reviews, Twitter Sentiment Dataset) and deep learning techniques (e.g., BERT) for sentiment analysis is explored. While sentiment analysis has made significant strides, it faces challenges such as deciphering sarcasm and irony, ensuring ethical use, and adapting to new domains. We emphasize the dynamic nature of sentiment analysis, encouraging further research to unlock the nuances of human sentiment expression and promote responsible and impactful applications across industries and languages.

Keywords Analysis · Deep learning · Machine learning · Mining · Natural language processing · Opinion · Sentiment

1 Introduction

In recent years, sentiment analysis has emerged as a pivotal field within natural language processing (NLP), driven by the exponential growth of digital data and the increasing need to extract insights from textual information. With the prevalence of social media platforms, online reviews, and customer feedback, understanding the sentiments expressed in a text has become essential for various applications, ranging from business intelligence to social sciences and beyond.

Sentiment analysis, also known as opinion mining, refers to the computational process of identifying, extracting, and analyzing subjective information from a text to determine the overall sentiment or attitude expressed toward a particular entity, product, topic, or event [1–3]. This process involves classifying text into pre-defined categories such as positive, negative, or neutral sentiments, or more nuanced emotions and opinions [1–3].

Traditionally, sentiment analysis was performed using machine learning algorithms such as support vector machine (SVM), Naive Bayes, Logistic Regression, and Random Forest. However, as the scope of sentiment analysis has widened and new gaps have been identified within sentiment analysis classifications, there has been a shift toward utilizing more complex algorithms to address emerging challenges [4]. Deep learning algorithms, which are part of the machine learning architecture, offer promising solutions to these challenges, with architectures such as convolutional neural networks (CNNs), long short-term memory networks (LSTMs), and recurrent neural networks (RNNs) demonstrating efficacy in handling complex sentiment analysis tasks [4].

Furthermore, recent advancements in deep learning have introduced a new algorithm known as bidirectional encoder representations from transformers (BERT) [5]. Developed by Google, BERT has garnered attention for its ability to capture contextual relationships within text and has shown superior performance compared to other deep learning techniques in sentiment analysis tasks.

The proliferation of social media platforms has significantly contributed to the availability of vast amounts of

✉ Neeraj Anand Sharma
  neerajs@unifiji.ac.fj

  A. B. M. Shawkat Ali
  shawkata@unifiji.ac.fj

  Muhammad Ashad Kabir
  akabir@csu.edu.au

[1] Department of Computer Science and Mathematics, School of Science and Technology, The University of Fiji, Lautoka, Fiji

[2] School of Computing, Mathematics, and Engineering, Charles Sturt University, Bathurst, NSW 2795, Australia

textual data for sentiment analysis. For instance, data from Facebook and Twitter indicate a substantial increase in active users over the past few years, with billions of users collectively expressing their opinions through comments, reviews, posts, and statuses on various topics of interest [6, 7]. This abundance of data presents unprecedented opportunities for research and analysis, underscoring the importance of effective sentiment analysis methodologies [8].

This paper aims to provide a comprehensive review of sentiment analysis tasks, applications, and the utilization of deep learning techniques in this domain. We begin by clarifying the fundamental concepts of sentiment analysis, including various tasks such as document-level, sentence-level, aspect-based, and emotion detection, among others. Subsequently, we delve into the diverse applications of sentiment analysis across domains such as business, social media, finance, politics, and disaster management, highlighting its significance in informing decision-making processes and driving actionable insights.

Furthermore, we explore the landscape of deep learning techniques employed in sentiment analysis, discussing their advantages over traditional approaches and presenting a detailed overview of architectures such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), long short-term memory networks (LSTMs), gated recurrent units (GRUs), bidirectional encoder representations from transformers (BERT), large language models (LLMs), and graph neural networks (GNNs). By synthesizing existing literature and empirical studies, we aim to elucidate the strengths, limitations, and performance characteristics of these deep learning models in sentiment analysis tasks.

Through an extensive methodology encompassing literature search, data extraction, and analysis, we examine the current state-of-the-art in sentiment analysis, identify common challenges, and assess the performance of deep learning techniques across different domains and languages. Additionally, we discuss implications for future research directions and potential avenues for overcoming existing limitations in sentiment analysis methodologies.

In essence, this paper serves as a comprehensive guide for researchers, practitioners, and enthusiasts alike, seeking to deepen their understanding of sentiment analysis, explore its diverse applications, and harness the power of deep learning techniques to extract valuable insights from textual data.

1.1 Related works

To gain a deeper understanding of existing research and identify potential gaps in knowledge, this section reviews relevant scholarly literature on sentiment analysis tasks, applications, and deep learning methods. We were able to identify three survey papers that try to focus on the same issues as ours.

Firstly, the paper [9] explores various deep learning architectures for sentiment analysis. It highlights the effectiveness of deep learning techniques in this domain and predicts future advancements in deep learning for sentiment analysis. Deep learning architectures explored in the paper were CNN, RNN, LSTM, attention mechanism RNN, memory network (MemNN), and recursive neural network (RecNN). Sentiment analysis tasks highlighted in the survey paper were document-level sentiment classification, sentence-level sentiment classification, aspect-level sentiment classification, aspect extraction and categorization, opinion expression extraction, sentiment Composition, opinion holder extraction, temporal opinion mining, sentiment analysis with word embedding, sarcasm analysis, emotion analysis, and multimodal data for sentiment analysis. Although the paper mentions deep learning, no experimental analysis proves which deep learning architecture is superior to the other, and the paper fails to mention relevant field application examples. Following this, the paper [10] presents a comprehensive, state-of-the-art review of research work completed on various aspects of sentiment analysis from 2002 to 2014. The paper analyzes research work across six key areas: subjectivity classification, sentiment classification, review usefulness measurement, lexicon creation, opinion word and product aspect extraction, and various applications of opinion mining. While the paper primarily focuses on sentiment analysis tasks, it refers to some field application examples but fails to acknowledge deep learning architectures. Finally, paper [11] is similar to the above two papers where the primary goal of this paper was to disseminate knowledge on the different types of tasks involved in sentiment analysis and opinion mining. Although the majority of the tasks were looked into, research still needs to look into popular tasks and capture if any experiments have been done on these tasks or not.

Table 1 highlights the contributions and limitations of recent existing survey papers that are similar to this review paper.

Having reviewed the current landscape of sentiment analysis research, the following section details our contributions to this field.

1.2 Our contributions

This research aims to illuminate promising avenues for future advancements in sentiment analysis by exploring six key questions. By addressing the limitations identified in existing literature, this study seeks to bridge critical knowledge gaps and propel the field toward more robust and effective sentiment analysis techniques. The following sections delve into each research question, highlighting the limitations in current research and outlining how this study will contribute to a deeper understanding and practical application of sentiment analysis with deep learning techniques.

Table 1 Limitations of recent survey papers in the area of sentiment analysis tasks and deep learning

References	Year	Contribution	Limitation
[12]	2024	The paper argues that deep learning offers superior accuracy compared to other methods, particularly in handling large datasets, complex features, and contextual information. It highlights deep learning's ability to overcome limitations like negation and modifiers in sentences The authors conduct a critical analysis of existing deep learning sentiment analysis methods, including their tools, advantages, disadvantages, research gaps, and findings	The paper acknowledges the ongoing need for exploring new computational techniques to further improve the accuracy of sentiment analysis, particularly in social media and related domains The paper identifies the need for advancements in handling multimodal data (text, image, audio), multiple data sources, live data streams, and feedback data The study recognizes that sentiment analysis based solely on text can struggle with factors like coherence, negation, intensifiers, and semantic meaning
[13]	2024	The emergence of models like BERT and GPT is identified as a major breakthrough. These models leverage pre-training and fine-tuning, leading to significant improvements in NLP benchmarks compared to previous methods The paper emphasizes the ability of these models to tackle challenges like limited data, adapting to new domains, and handling multiple languages. This broadens the applicability of NLP techniques	The study identifies the need to integrate symbolic knowledge, potentially combining deep learning with rule-based approaches, for further advancements There's a call for developing more efficient deep learning models to improve scalability and potentially reduce the computational resources required The paper highlights the potential of exploring multimodal approaches, where NLP is combined with other data types like images or audio, for a richer understanding of language
[14]	2020	The paper addresses the lack of comparative analysis between common deep learning models (CNN, RNN, LSTM) for sentiment polarity analysis. The authors conduct experiments to compare the performance of these models on different datasets The paper highlights the benefit of combining deep learning models with word embedding techniques over TF-IDF for sentiment analysis tasks	The experiments focus on sentiment polarity analysis, which is a simplified form of sentiment analysis. Future work is needed to explore more nuanced aspects like aspect-based sentiment analysis The study uses datasets with specific characteristics. The authors acknowledge the need for testing deep learning methods on a wider variety of datasets to ensure the generalizability of findings The paper fails to highlight the different types of sentiment analysis tasks
[15]	2019	This paper contributes to the field of sentiment analysis by providing a comprehensive survey of deep learning architectures used for both sentence-level and aspect-level sentiment analysis. The authors delve into the advantages and drawbacks of various state-of-the-art methodologies However, the study reveals a clear trend in aspect-level sentiment analysis: a shift from traditional word-based approaches toward a semantic concept-centric approach. This paves the way for the development of next-generation algorithms that can reason with language and context at a more sophisticated level, combining deep learning with semantic concepts	The optimal deep learning architecture hinges on the specific sentiment analysis task, making it challenging to recommend a single best method for all scenarios The absence of standardized evaluation methods hinders a definitive comparison of different deep learning approaches across various domains
[16]	2019	This paper offers a comprehensive survey of sentiment analysis using deep learning architectures. It presents a well-structured classification of sentiment analysis approaches, including handcrafted and machine-learned features, alongside a detailed analysis of deep learning models like CNNs, RNNs, LSTMs, and more It identifies the potential of sentiment analysis in various domains like implicit sentiment detection, spam filtering, and medical applications. Additionally, it explores the possibilities of multimodal data analysis for sentiment understanding	The major challenge identified is the lack of sufficient training data for specific sentiment analysis tasks. Deep learning models often require large datasets for optimal performance The paper highlights the issue of limited datasets containing both sentiment and topic information for text retrieval tasks The paper identifies a strong trend towards transfer learning approaches, emphasizing the need for robust dataset creation The study acknowledges the need for extensive training data to achieve improved performance in fine-grained sentiment analysis

Table 1 (continued)

References	Year	Contribution	Limitation
[17]	2019	Deep learning models outperform shallow models (like SVMs and basic neural networks) due to their complex architecture with multiple hidden layers. This allows them to capture intricate patterns in data, leading to more accurate sentiment prediction Deep learning models can be trained in supervised or unsupervised ways, eliminating the need for manual feature engineering, which saves time and effort	Deep learning models require vast amounts of data for training. This can be a barrier for applications with limited datasets Training these complex models can be very expensive due to the need for powerful hardware like GPUs and the extended training time involved

First, by identifying well-established sentiment analysis tasks and those with limited research, We aim to illuminate unexplored territories. While several studies [9–11, 15, 16] provide a comprehensive overview of sentiment analysis tasks, limitations exist in identifying unexplored territories [14]. This research aims to address this gap by thoroughly examining sentiment analysis tasks and pinpointing areas with a limited research focus. This will guide future researchers toward unaddressed challenges and potentially lead to breakthroughs in sentiment analysis techniques.

Second, We will investigate the performance of various deep learning techniques in sentiment analysis tasks. This evaluation will provide valuable insights into which technique excels in specific scenarios. The effectiveness of deep learning for sentiment analysis is well-established [12, 17], but a clear understanding of which technique excels in specific scenarios is still needed. Studies like [9, 16] offer a broad survey but lack in-depth performance comparisons. This research aims to bridge this gap by conducting a rigorous evaluation of various deep learning techniques across different sentiment analysis tasks. By identifying the strengths and weaknesses of these techniques, the research can contribute to the development of more effective models, ultimately leading to more accurate and reliable sentiment analysis tools.

Third, the research explores how deep learning models can be effectively applied in real-world applications, such as customer feedback analysis. While the potential of sentiment analysis in various domains is acknowledged [10, 16], limited research explores its practical application. This research addresses this limitation [15] by investigating how deep learning models can be effectively applied in real-world scenarios, such as customer feedback analysis. Demonstrating the practical value of sentiment analysis can pave the way for its wider adoption across various industries.

Fourth, We will analyze how deep learning techniques perform across datasets from different domains and languages. This is crucial for their broader applicability. While the importance of generalizability across domains and languages is acknowledged [16], limited research explores how deep learning techniques perform in these diverse contexts.

This research will analyze how deep learning techniques perform on datasets from diverse domains and languages. By revealing their strengths and weaknesses in these contexts, the research can inform strategies for improving their adaptability and broadening their applicability.
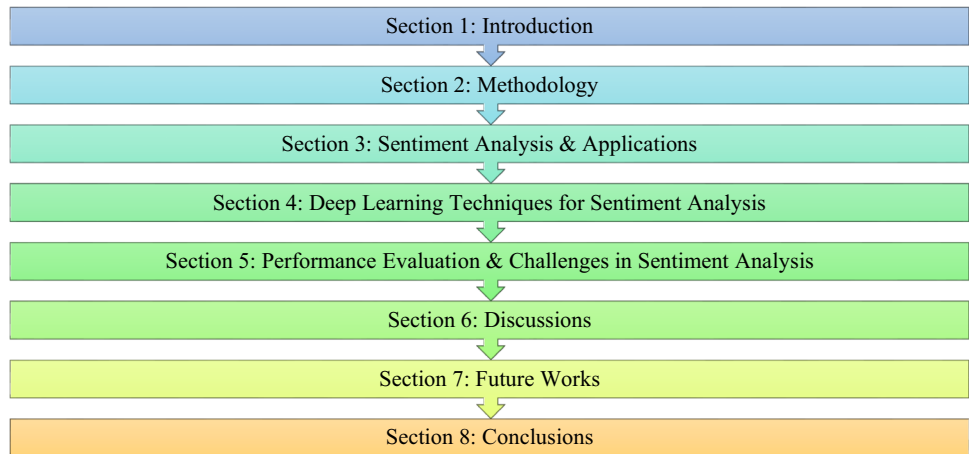
Fifth, by exploring commonly used datasets for sentiment analysis with deep learning, the research can contribute to the standardization of practices in this field. The lack of standardized practices regarding datasets used for sentiment analysis with deep learning is a recognized limitation [13]. This research aims to contribute to dataset standardization by exploring commonly used datasets and their characteristics. This analysis will facilitate data sharing, collaboration, and the development of more robust and generalizable deep learning models.

Finally, the research will identify and address common challenges in sentiment analysis, such as handling negation, sarcasm, and complex emotions. Shedding light on these persistent issues can pave the way for the development of more nuanced and accurate sentiment analysis techniques, significantly improving the field's overall effectiveness. Sentiment analysis faces challenges such as handling negation, sarcasm, and complex emotions [11, 12, 14]. This research will address these challenges by identifying effective techniques for sentiment analysis models to overcome them. By shedding light on these persistent issues, the research can pave the way for the development of more nuanced and accurate sentiment analysis techniques, significantly improving the field's overall effectiveness.

1.3 Paper structure

The paper is distributed into eight sections. Section 2 contains the methodology that has been applied to this review paper, Sect. 3 delves into the sentiment analysis tasks and applications, and Sect. 4 elaborates on the seven deep learning techniques for sentiment analysis with a look at their strengths and limitations. In Sect. 5, we perform evaluations on all seven deep learning techniques and identify common sentiment analysis challenges, Sect. 6 includes a discussion that looks into the main objectives of this study, and Sect. 7

Fig. 1 Paper layout

Section 1: Introduction
Section 2: Methodology
Section 3: Sentiment Analysis & Applications
Section 4: Deep Learning Techniques for Sentiment Analysis
Section 5: Performance Evaluation & Challenges in Sentiment Analysis
Section 6: Discussions
Section 7: Future Works
Section 8: Conclusions

highlights the future works and latest trends that need to be explored. Finally, Sect. 8 concludes the paper. Figure 1 highlights how the paper will be structured.

2 Methodology

2.1 Objectives

This paper aims to provide a comprehensive review of sentiment analysis tasks and applications, with a particular emphasis on the utilization of deep learning algorithms. Through a thorough literature review, we aim to achieve the following:

- Identify the various types of sentiment analysis tasks that have been addressed and those with potential for further exploration.
- Analyze deep learning algorithms used in sentiment analysis, exploring their strengths, weaknesses, and suitability for different tasks.
- Highlight key challenges associated with sentiment analysis in general.
- Identify real-world applications of sentiment analysis, particularly those that can significantly impact the customer service industry.

2.2 Research questions

Since my research focuses on sentiment analysis, the following six research questions will guide my investigation:

1. What are the main sentiment analysis tasks that have been addressed in existing literature, and which areas offer potential for further research?

2. What are the different deep learning algorithms used in sentiment analysis, and how do their performances compare across various tasks and datasets?
3. How can deep learning models be effectively utilized in real-world applications, such as customer feedback analysis, to extract valuable insights?
4. How does the performance of deep learning algorithms vary when applied to sentiment analysis tasks using datasets from different domains or languages?
5. What are the common datasets used to evaluate deep learning models for sentiment analysis tasks?
6. What are the major challenges still faced in sentiment analysis, and how can these challenges be addressed?

These questions will provide a focused framework for my literature review and analysis.

2.3 Literature search procedure and criteria

Since this research paper focuses on reviewing existing research on sentiment analysis using deep learning algorithms, a comprehensive literature search is crucial. This section outlines the search criteria and process followed to identify and review high-quality papers.

2.3.1 Constructing the search terms

Our research questions guided the development of specific keywords for our literature search. These keywords included terms like sentiment analysis, opinion analysis, sentiment mining, sentiment applications, sentiment tasks, and combinations with machine learning and deep learning algorithms (e.g., "sentiment analysis with deep learning"). Additionally, we included specific algorithm names such as CNN, RNN, LSTM, BERT, GRU, LLM, and GNN.

Table 2 Databases and results for literature search

No	Databases	Total search results	Primary selection	Final selection
1	IEEE Xplore digital library	13,890	210	51
2	Science direct	11,093	115	29
3	Springer	5129	68	26
4	Google scholar	17,672	112	45
5	Wiley online library	1562	27	6
6	ACM digital library	7341	59	24

The search terms were refined through two strategies. First, we identified synonyms for key terms. For example, "sentiment analysis" and "opinion mining" are often used interchangeably. Second, we combined terms to narrow down the search. For instance, the combination of "sentiment analysis using deep learning algorithms" or "sentiment analysis using deep learning techniques" ensured that retrieved papers contained both aspects relevant to our research. This approach helped us identify relevant studies while excluding irrelevant ones.

2.3.2 Search strategy

We searched a variety of academic databases, including IEEE Xplore, ScienceDirect, and Springer, as detailed in Table 2. The initial search yielded a large number of papers.

2.3.3 Publication selection

We employed a two-phase selection process to identify relevant and high-quality research papers. The first phase involved a primary selection based on titles, keywords, and abstracts. Papers that appeared relevant to our research questions were selected for further review.

The second phase involved a final selection after thoroughly reading the shortlisted papers. During this phase, we focused on the research objectives, experimental design, results, gaps, and limitations identified by the authors. This in-depth review allowed us to select papers that would significantly contribute to our own research. Mention about text data used and not image data.

2.3.4 Range of research papers

The literature review primarily encompasses research papers published between 2014 and 2024. In total, 178 papers were utilized in this review paper and Fig. 2 highlights those papers selected from a specific year. The focus of the experimental results analysis is on papers published between 2019 and 2024 which is also illustrated in Fig. 3. While the focus is on recent research, a few key papers from earlier years have
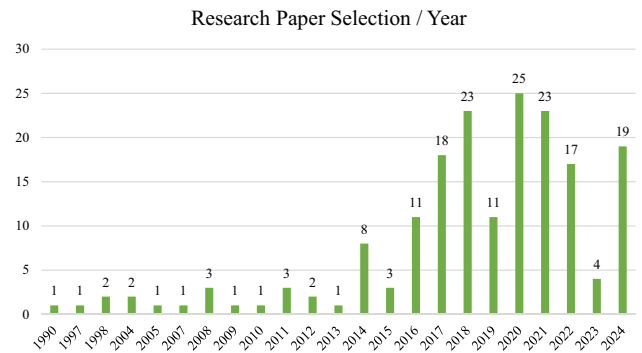
Fig. 2 Number of research paper selection per year

also been included, particularly those that discuss significant challenges in sentiment analysis.

By following this rigorous literature search and selection process, we were able to gather a comprehensive collection of high-quality research papers that will form the foundation for our review of sentiment analysis using deep learning algorithms.

2.4 Data extraction and analysis

A standardized approach was used to extract relevant data from the reviewed papers to answer our research questions. These data included details such as:

- Sentiment analysis tasks addressed
- Real-world application focus
- Deep learning algorithms or models used
- Gaps and challenges highlighted
- Datasets and data size
- Language used
- Performance measures (e.g., F1 score, accuracy)

The following sections summarize the key findings related to each research question:

Fig. 3 A simplified paper extraction process
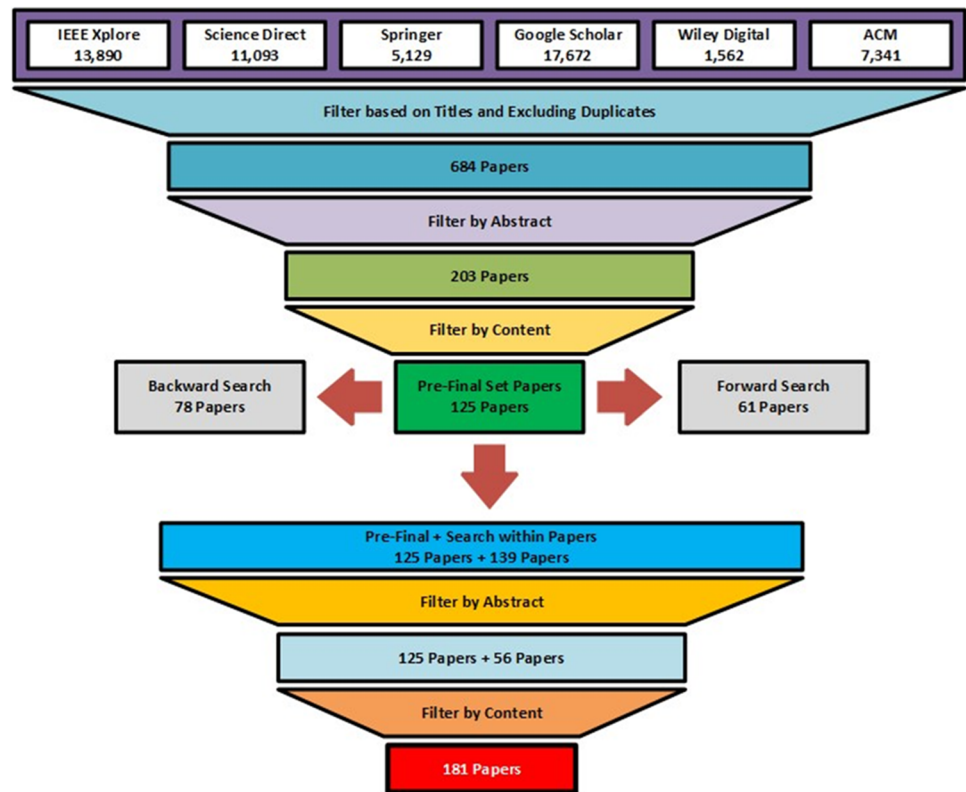
2.4.1 Main sentiment analysis tasks

We were able to identify several sentiment analysis tasks from the literature review. These tasks include document-level sentiment analysis, sentence-level sentiment analysis, aspect-based sentiment analysis, aspect extraction, emotion detection, multi-domain sentiment classification, multilingual sentiment analysis, multimodal sentiment analysis, opinion summarization, opinion spam detection, opinion holder extraction and classification, time extraction and standardization, visual sentiment analysis, and graded sentiment analysis. All these mentioned sentiment analysis tasks have appeared in several papers where the majority of the research focuses on general sentiment analysis and now research is being done extensively on the first six mentioned tasks [9–11, 15, 16]. From the literature review, we found out that more sentiment analysis tasks might be looked into in the near future and no research has been done on them as of now [9–11, 15, 16]. These are entity extraction and categorization, aspect extraction and categorization, sentiment or opinion classification, sentiment intersubjectivity, lexicon expansion, financial volatility prediction, opinion recommendation, stance detection, sarcasm analysis, subjectivity detection, opinion polarity classification, argument expression detection, emotion detection, emotion polarity classification, emotion classification, and emotion cause detection.

2.4.2 Deep learning techniques in sentiment analysis

From our literature review, we found out that there are seven common deep-learning algorithms used in sentiment analysis. These include convolutional neural networks (CNNs), recurrent neural networks (RNNs), long short-term memory (LSTM), gated recurrent unit (GRU), transformer-based model known as BERT, large language models (LLMs), and graph neural networks (GNNs). LSTM/BiLSTM and GRU/Bi-GRU are variants of RNN. To identify which algorithm performed the best, we looked into several experimental papers where a specific sentiment analysis task is mentioned and noted the F1 and accuracy scores together with the algorithm/model used and the respective datasets that were used in the experiments [12, 17].

2.4.3 Real-world applications of sentiment analysis

While going through several research articles, it was found that there are several use cases and application-based scenarios where sentiment analysis-based applications were used to tackle certain issues [15]. This spans several domains such as education, brand monitoring and business intelligence, social media, finance, and stock marketing, enhancing customer service and experience, market research and analysis, politics, crime prediction, disaster assessment, response, and management, and demonetization. These application

domains will enable us to identify some research gaps and scope.

2.4.4 Performance variations across different domains or languages

We looked into research articles that focused on different domains or languages. There was a total of ten different languages that were identified from the literature search. These languages were English, Hindi, Chinese, Turkish, Sinhala, Slovene, Arabic, Persian, Indonesian, and Vietnamese. Since English language datasets were mostly used by researchers, it becomes very difficult to compare and contrast which language domain would provide the best result. There is scope for research using other languages as well not only English [16]. Determining the best language would alternatively depend on the language section, the algorithm used, the dataset size, and the preprocessing tasks used.

2.4.5 Common datasets for sentiment analysis

We identified common datasets used for sentiment analysis tasks with deep learning algorithms. Examples include IMDB, SemEval 2014–2016, Twitter, Sentiment140, Chinese Text Corpora, Amazon Review, Twitter, and some online review datasets. The dataset size ranged from 1000 to anywhere more than 1,000,000+ . From the literature review, we saw that several experiments were done using custom online review datasets which is not common as datasets could be manipulated to get the required accuracy [13].

2.4.6 Challenges in sentiment analysis

We identified and extracted common challenges associated with sentiment analysis in general and the tasks as well. Some of the major challenges found were sarcasm and irony detection, algorithm optimization, language model issues, subjectivity, tone-setting polarities, and data inconsistency among others. The majority of the research emphasized the need for improved models that could classify sarcasm and irony [11, 12, 14].

By extracting and analyzing this data, we were able to gain valuable insights into the current state of sentiment analysis research using deep learning algorithms. This information will be crucial for developing a comprehensive review and identifying potential areas for further research.

2.5 Quality assessment

To ensure the quality and credibility of the selected research articles, we implemented a rigorous quality assessment process. Two independent reviewers evaluated the papers against the criteria established for our research questions. This evaluation included:

- Relevance to the research questions
- Appropriateness and soundness of the methodology
- Overall contribution to the field of sentiment analysis using deep learning
- Validity and significance of the experimental results

By employing this two-reviewer approach, we aimed to minimize bias and ensure the inclusion of high-quality research articles that would form the foundation for our review.

3 Sentiment analysis and applications

3.1 Definition and importance of sentiment analysis

Sentiment analysis, also known as opinion mining, is a rapidly growing field in natural language processing (NLP) that automatically identifies the sentiment (positive, negative, or neutral) expressed in textual data. This analysis is crucial for businesses to understand customer opinions, monitor brand perception, and gather valuable feedback on products and services. The origins of sentiment analysis can be traced back to ancient Greece [18], but its use has exploded in recent years, with applications spanning various domains like education, brand monitoring, social media analysis, finance, and market research [16, 19, 20].

Sentiment analysis can categorized into various tasks such as document-level sentiment analysis, sentence-level sentiment analysis, aspect-based sentiment analysis, aspect extraction, emotion detection, multilingual sentiment analysis, multimodal sentiment analysis, opinion summarization, opinion spam detection, opinion holder extraction and classification, time extraction and standardization, graded sentiment analysis, and visual sentiment analysis [1].

Sentiment analysis can be performed using various methods. Lexicon-based approaches rely on sentiment dictionaries that assign positive, negative, or neutral values to words [1]. Machine learning approaches train models on labeled data to automatically identify sentiment [1]. Deep Learning, a powerful sub-category of Machine Learning, has emerged as a leading approach due to its ability to handle complex data patterns [1].

3.2 Sentiment analysis tasks

3.2.1 Document-level sentiment analysis (DLSA)

Document-level sentiment analysis (DLSA) typically examines text within a document to determine its overall sentiment

(positive, negative, or neutral) [21]. This analysis is usually applied to text exceeding 40 characters. Currently, DLSA supports sentiment analysis in English, French, German, Spanish, and Italian [21].

Example: *"Tim buys a brand-new Samsung smartphone. He likes the phone, although he finds the screen size a bit large for his taste. The touchscreen feels responsive, and the voice quality is good. Overall, Tim enjoys his new smartphone."*

DLSA excels when a single author expresses their opinions or feelings about a single entity in a document. In this case, the entity is the Samsung smartphone [21].

3.2.2 Sentence-level sentiment analysis (SLSA)

Sentence-level sentiment analysis (SLSA) typically involves two key tasks: subjectivity classification and sentiment classification. Subjectivity classification: Sentences are categorized as either subjective (expressing opinions or feelings) or objective (presenting facts) [22]. Sentiment classification: Subjective sentences are further classified as positive or negative. Objective sentences provide factual information relevant to the topic, while subjective sentences focus on emotions, viewpoints, and personal feelings. In SLSA, algorithms like Naive Bayes can be used to identify and remove sentences that lack sentiment or relevant entities (targets of opinions) sentence-level sentiment analysis.

Example: *"Asco Motors are doing well in this bad economy"*.

While document-level and sentence-level sentiment analysis are valuable, they have limitations. They do not necessarily reveal what aspects people like or dislike, nor do they identify the specific targets of opinions.

3.2.3 Aspect-based sentiment analysis (ABSA)

Aspect-based sentiment analysis (ABSA) is a targeted approach that focuses on analyzing specific aspects within a sentence and determining the sentiment (positive, neutral, or negative) associated with those aspects. This analysis utilizes an aspect-based classifier to identify these aspects and their corresponding sentiment [16].

Example: *"The food at the restaurant was delicious, but the service was slow."*

ABSA would identify *"food"* and *"service"* as aspects. It would then classify the sentiment for *"food"* as positive and the sentiment for *"service"* as negative.

3.2.4 Aspect extraction (AE)

Aspect extraction (AE) focuses on identifying the specific elements (aspects) within a sentence that are being evaluated [23]. These aspects are often nouns or noun phrases that represent entities like products, services, or features.

Example: *"The food tasted great, but the service was really poor,"*

AE would identify *"food"* and *"service"* as the aspects being evaluated for sentiment. AE plays a crucial role in aspect-based sentiment analysis as it helps pinpoint the targets of opinions expressed in a sentence [23]. Generally, pre-trained models are used in AE to identify aspects accurately.

3.2.5 Emotion detection (ED)

Emotion detection is a more nuanced form of sentiment analysis compared to other approaches. It goes beyond basic sentiment (positive, negative, neutral) and aims to identify specific emotions such as happiness, sadness, anger, and frustration. Analyzing these emotions often requires advanced techniques like lexicons or complex machine learning algorithms. Lexicons are large databases of words with pre-assigned emotional values. However, they have limitations. Lexicons often struggle with nuanced language or sarcasm.

For example, a lexicon might categorize *"your product is bad"* and *"your customer support is killing me"* as negative, even though the latter expresses frustration through hyperbole.

Similarly, *"this is badass"* and *"you're killing it"* might be misconstrued as negative due to words like *"bad"* and *"killing"*. Machine learning algorithms offer a more sophisticated approach to emotion detection. These algorithms can learn from large datasets of text labeled with specific emotions, allowing them to handle complex language and identify emotions more accurately. The field of emotion detection emerged around 2005 [24].

3.2.6 Multi-domain sentiment classification

Multi-domain sentiment classification tackles the challenge of analyzing sentiment across different domains (areas of focus) in text data. Imagine a document containing topics relevant to both finance and sports. Multi-domain sentiment classification aims to understand the sentiment within each domain. This process typically involves transfer learning, where a model trained on one domain (source domain) is adapted to analyze sentiment in another domain (target domain). The knowledge gained from training on the source domain helps the model perform better in the target

domain. Research by [25] utilized a deep learning model with word embeddings to create NeuroSent, a tool for multi-domain sentiment analysis. Another study [26] proposed the domain attention model (DAM), which specifically focuses on modeling features relevant to sentiment analysis in different domains.

3.2.7 Multilingual sentiment analysis

Multilingual sentiment analysis presents the most significant challenge among the discussed analysis options. This is because it requires extensive preprocessing to handle the complexities of different languages. Additionally, it often relies on external resources such as sentiment lexicons (word lists with emotional values) specifically designed for each language, translated corpora (large collections of text), or noise detection algorithms to remove irrelevant data. Strong coding skills and familiarity with natural language processing techniques are beneficial for tackling this challenge. An alternative approach involves using a language detector. The field of multilingual sentiment analysis emerged around 2009, with pioneering work by Boiy and Moens [27].

3.2.8 Multimodal sentiment analysis (MMSA)

Multimodal sentiment analysis (MMSA) tackles the most complex form of sentiment analysis by considering information from various communication channels: spoken language, written text, and visuals. This approach goes beyond analyzing just words to understand the sentiment conveyed through different communication mediums. Soleymani et al. [28] identified three main categories of MMSA: analyzing spoken reviews and vlogs, analyzing images and social media tags, and analyzing human–machine interactions. While the field gained traction with the work of Morency et al. in 2011 [29], MMSA continues to be an evolving area of research.

3.2.9 Opinion summarization

Opinion summarization is a technique used in sentiment analysis to generate concise summaries of shared opinions or viewpoints on a specific topic. This process typically involves several steps, including data preprocessing, where the text is cleaned and formatted, tokenization (breaking the text into individual words or phrases), and stemming (reducing words to their root form) [30]. Opinion summarization is particularly valuable in analyzing social media data. For example, imagine a social media post about a new political candidate. Hu and Liu are credited with pioneering the field of opinion summarization in their work published in 2004 [31].

3.2.10 Opinion spam detection (OSD)

Opinion spam detection (OSD) tackles the issue of deceptive online activity that aims to mislead users and sentiment analysis tools. This includes tactics like creating fake reviews, blogs, and social media posts to promote a particular entity (product, service, person) or damage its reputation. Essentially, OSD focuses on identifying and filtering out bogus reviews, fake data, and misleading information circulating online. For example, imagine a news article claiming a *"20% Pay Rise to All Civil Servants in Fiji"*. OSD algorithms could help identify this as potential fake news. Jindal and Liu's 2008 scientific paper is credited as one of the first to explore Opinion Spam Detection [32].

3.2.11 Opinion holder extraction and classification

Opinion holder extraction and classification is a technique used in sentiment analysis to identify and categorize the entities expressing opinions within a sentence. These entities can be nouns, verbs, or adjectives.

Example: *"Tim had intentions to go out tonight"*.

Here, *"intentions"* is the opinion holder, suggesting Tim has an opinion about going out.

Another example: *"Jane was skeptical about the weather"*.

In this case, *"skeptical"* is the opinion holder, indicating Jane's doubt regarding the weather.

Similarly, *"Jane criticized Sarah"*.

Here, "criticized" is the verb acting as the opinion holder, implying Jane has a negative opinion of Sarah. Pioneering work by [33] laid the groundwork for models that can handle both opinion holder extraction and classification.

3.2.12 Time extraction and standardization

Time extraction and standardization play a vital role in sentiment analysis. By identifying and standardizing temporal references (mentions of time) within text data, analysts can track shifts in sentiment over time [33]. Extracting time references can be achieved through various methods: rule-based methods, machine learning algorithms, and natural language processing (NLP) techniques. Once extracted, time references need standardization to facilitate analysis and comparison across diverse texts and data sources. Standardization can be achieved either manually or using automated techniques. Manually: This involves human reviewers manually converting time references to a consistent format (e.g., all dates converted to YYYY-MM-DD format). Automated techniques: Machine learning or NLP techniques can automate the standardization process for improved efficiency.

The task of time extraction and standardization was first addressed in the work of [33].

3.2.13 Visual sentiment analysis

Visual sentiment analysis is a growing field within sentiment analysis that tackles the challenge of understanding emotions conveyed through visual content. This increasing interest stems from the rise of visual communication, with people expressing themselves more and more through photographs, emojis, and other visual elements [34]. This technology has valuable applications in various fields such as Social media marketing, Customer input analysis, and Advertising and market research [34]. This approach is particularly useful for analyzing content uploaded on social media platforms like Facebook, Twitter, and Instagram, where images play a prominent role in brand promotion and user feedback.

3.2.14 Graded sentiment analysis (GSA)

Graded sentiment analysis, also known as fine-grained sentiment analysis, provides businesses with a more nuanced understanding of customer sentiment. This technique goes beyond a simple positive, negative, or neutral classification. Instead, it typically uses a five-point scale: very positive, positive, neutral, negative, and very negative [35]. This approach is particularly useful for analyzing product ratings, often expressed as star ratings (1–5). A 1-star rating would be categorized as *"very negative,"* indicating the customer is highly dissatisfied. Conversely, a 5-star rating would be considered *"very positive,"* suggesting a high level of customer satisfaction.

3.3 Applications of sentiment analysis

The explosion of Web 2.0 ushered in a new era of online interaction and data generation. This era saw the rise of e-commerce platforms, social media, personalized company websites, cloud-based systems, and a significant increase in online traffic [36]. As a result, conducting business online has become the norm for many consumers, especially evident during the COVID-19 pandemic when businesses pivoted online for survival. This digital landscape has created a space for users to express their opinions and sentiments regarding products and businesses. Online reviews and ratings are now crucial factors for businesses, as today's consumers rely heavily on them when making purchasing decisions or choosing service providers. It is a common practice for people to conduct "background checks" online by reading reviews before buying a product or working with a company [36]. Positive experiences often lead to positive online reviews, while negative experiences can result in critical feedback. The applications of sentiment analysis extend beyond just products and e-commerce. Here, we'll explore how sentiment analysis is utilized in various industries and business areas.

3.3.1 Education

Sentiment analysis holds immense potential to improve educational quality at all levels. As highlighted by [19], it can be effectively utilized in several key areas within the education sector such as instructional evaluations, institutional decision-making and policy-making, intelligent information and learning systems enhancement, and assignment evaluation and feedback improvement.

Instructional Evaluation: Sentiment analysis can analyze student surveys, comments, and reviews to gauge their perception of teachers and courses [37–39]. This real-time feedback allows for prompt intervention and adjustments if necessary. Studies by [40, 41] demonstrate the effectiveness of using sentiment analysis on student feedback provided after lectures on platforms like Moodle.

Institutional Decision-Making and Policy-Making: Social media analysis can be used to understand student and parent sentiment regarding university operations and rank [40, 41]. Similarly, sentiment analysis of parent feedback during meetings can inform improvements to university programs and services [42].

Intelligent Information and Learning Systems Enhancement: Sentiment analysis can be integrated into learning platforms to identify student needs and provide targeted assistance. Research by Scaffidi [43] proposes a model that analyzes forum messages for sentiment and suggests solutions based on the identified issues. Similarly, research by [44] explores using sentiment analysis to provide emotional support and guidance within adaptive e-learning systems.

Assignment Evaluation and Feedback Improvement: Sentiment analysis models can be used to evaluate student work that relies heavily on opinion and sentiment, such as essays [45]. This approach can help identify areas for improvement and provide more targeted feedback. Cummins et al. [44] argue that sentiment analysis can also reveal student attitudes that might contribute to poor performance.

The analytical power of sentiment analysis opens doors for significant research opportunities in education. By mining educational data effectively, sentiment analysis can help address issues like early dropout rates, the impact of social media on student performance, and the correlation between student sentiment and academic achievement [46, 47].

3.3.2 Brand monitoring and business intelligence

In today's tech-driven world, artificial intelligence plays a central role in brand monitoring. Businesses use data analysis to gauge the performance of their products and services.

Sentiment analysis takes this a step further by analyzing online reviews, comments, and ratings to understand public perception. Positive reviews are a valuable indicator that a business is on the right track. They highlight customer satisfaction and areas where the business excels. Negative reviews, while sometimes disheartening, offer valuable insights for improvement. By analyzing these reviews, businesses can identify shortcomings and take corrective actions before problems escalate.

Research by Chaturvedi et al. [48] explored various approaches to sentiment analysis for product evaluation. They identified a crucial challenge: the lack of a universal solution for all businesses. This underscores the potential for developing new, more generalized approaches. Benedetto and Tedeschi presented a case study using an intelligent sentiment analysis approach [49]. They targeted prominent electronic products on Twitter and devised a system to score both the polarity (positive/negative) and popularity of the sentiment expressed.

3.3.3 Social media

Social media platforms such as Twitter and Facebook have become a powerful force for public opinion. Users can express their views on anything and everything, from product reviews to social commentary. The sheer volume of data generated on these platforms is constantly growing, as evidenced by statistics from the past three years [6, 7]. This data offers a rich resource for researchers like Cheng and Tsai [50]. They observed the emergence of new communication styles on social media, including shortened text, slang, emojis, and acronyms. Recognizing the potential for future research, they created a unique dataset capturing these elements alongside sentiment analysis. This dataset empowers other researchers to develop models and applications that can understand this evolving online language.

The COVID-19 pandemic significantly impacted the world, and social media became a platform for expressing a wide range of emotions. Research by Nemes and Kiss [51] employed sentiment analysis using a deep learning algorithm to analyze comments, hashtags, tweets, and posts related to COVID-19. Their findings revealed a mix of positive and negative emotions among the public, reflecting the complex nature of the pandemic. Similarly, Bhat et al. [52] conducted research in the medical domain, highlighting the importance of identifying public sentiment regarding the cost of living, lockdowns, hospital capacity, healthcare worker deaths, and other pandemic-related issues.

Social media plays a crucial role in e-commerce, with companies leveraging these platforms to advertise and reach a wider audience. Research in sentiment analysis and social media has opened doors for the development of recommender systems [53, 54]. By analyzing social media sentiment, these systems can identify well-received products and those that are not performing as well. Armed with this information, sellers can gain valuable insights into customer sentiment and make informed decisions to improve their products or marketing strategies.

3.3.4 Finance and stock monitoring

The financial world, including stock markets and digital currencies, can be significantly swayed by social media activity. A prime example unfolded in 2021, 2022, and 2023 with the surge of digital currencies. Tweets from influential figures like Elon Musk, even those using emojis instead of text, could dramatically alter market sentiment for these currencies. Digital currencies like Shiba Inu, Dogecoin, and Bitcoin experienced overnight price fluctuations, sometimes creating instant millionaires based on a single social media post. Critics argue that some influencers might manipulate the market by buying shares at low prices and then using their social media reach to encourage followers to invest, driving up the price before cashing out for a hefty profit.

Researchers are exploring the potential of sentiment analysis to predict stock market movements. Sousa et al. proposed a model using a BERT deep learning algorithm to analyze sentiment in news articles that could potentially impact the stock market [55]. Their model achieved an accuracy of 82.5% in predicting stock price movements, demonstrating its effectiveness. However, a limitation identified was the use of a relatively small dataset.

Similar research by Ren, Wu, and Liu developed a model that combines sentiment analysis with a support vector machine to forecast stock price changes [56]. Their model achieved an impressive accuracy of 89.93%, potentially aiding investors in making informed decisions. Pagolu et al. proposed a model that analyzes public sentiment on social media (using Twitter data) to understand how a company's performance is perceived after stock price fluctuations [57]. Their model, which employed Random Forest with Word2Vec and N-gram techniques, achieved an accuracy of 70.49%. Khedr, Salama, and Yaseen presented a model that predicts stock market behavior by considering both historical price data and sentiment analysis of financial news [58]. Their model boasts a prediction accuracy of up to 89.80%, suggesting its potential effectiveness compared to other existing models.

3.3.5 Enhancing the customer experience and service

For any business to thrive, understanding customer sentiment is crucial. Happy customers translate to a loyal customer base and business success. The key to achieving this lies in actively listening to customer feedback and opinions. The digital age

has ushered in a wealth of customer feedback readily available online. Platforms like product review sites, business rating services, and social media provide a treasure trove of customer opinions and sentiment. Tools like "web scrapers" can be used to efficiently extract this valuable data from various online sources. In-depth analysis of this data is key to gaining valuable insights into customer sentiment and identifying areas for improvement. The rise of e-commerce and the COVID-19 pandemic have further amplified the importance of online customer reviews. Businesses can leverage these reviews, often posted on their websites or social media pages, to gain valuable feedback on product quality, customer experience, and marketing effectiveness [59]. Recognizing the importance of online reviews, many businesses, both online and offline, actively monitor their ratings and reviews.

Researchers like Jain and Dandannavar proposed a sentiment analysis framework that addresses the need for a scalable, flexible, and fast approach to analyzing customer sentiment on platforms like Twitter [60]. This framework utilizes machine learning methods and Apache Spark to extract valuable insights from large datasets of customer reviews. By harnessing the power of sentiment analysis, businesses can transform customer feedback into actionable insights, ultimately leading to improved products, enhanced customer experiences, and effective marketing strategies.

3.3.6 Market research and analysis

Sentiment analysis enables businesses to conduct market research and identify competitive landscapes in different areas and regions. With this information, businesses can pinpoint locations that are likely to be successful and avoid those with intense competition. Additionally, sentiment analysis can be used to explore market opportunities, identify a competitive edge, and forecast future trends [20, 61]. This area of application is under-researched, with limited publications showcasing its use in real-world scenarios [62]. However, research by Rambocas and Pacheco highlights the challenges associated with this approach [62]. These challenges include technical limitations such as accuracy, reliability, and validity of sentiment analysis models, practical limitations such as cost concerns, miscoding sentiment due to sarcasm or slang, cross-cultural variations in expression, and the prevalence of deceptive reviews, and ethical considerations such as the right to user privacy and the potential for exploitation of user data.

3.3.7 Politics

The realm of politics is highly sensitive, demanding careful handling of data and analysis. During the 2016 US presidential election, traditional news media like the Washington Post and ABC News projected polls heavily favoring Clinton [1].

However, Trump emerged victorious, highlighting the significant influence of social media and public sentiment on traditional polling methods [1]. This unexpected outcome spurred a surge in research on sentiment analysis in politics. One such study by Haselmayer and Jenny [63] explored the creation of a dictionary-based technique for detecting negative sentiment. Their experiment utilized Australian political data from 2013. Human coders were employed to identify negative sentiment, and these findings were used to build a German political language dictionary specifically designed to analyze political party statements and media reportage. Their results revealed that media outlets primarily focused on larger parties, neglecting smaller ones. Additionally, the study found that negative statements often targeted the opposition, and attempts to manipulate language using certain English phrases were identifiable through the German dictionary [63].

Kušen and Strembeck [64] investigated the 2016 Austrian presidential election. They analyzed 343,645 Twitter messages, combining sentiment analysis techniques with network science to uncover key insights. Their research yielded five key findings: Neutral Tweets from the Winner, Dominance of Negative Sentiment, Follower Polarization, Engagement Differences, and Misinformation by Supporters. Further research by Kuamari and Babu in India [65] focused on identifying emotions related to the two major political parties, the Bharatiya Janata Party (BJP) and the Indian National Congress (INC), extracted from social media data. By understanding these classified emotions, political parties can identify areas for improvement and focus on policies that better address the needs of the people [65].

3.3.8 Crime prediction

Security intelligence in today's world is crucial for every nation. The rise of Web 2.0 has blurred the lines between traditional and online crimes. To effectively address both types of crime, artificial intelligence (AI) needs to be integrated into security systems. Research suggests that AI can be a valuable tool in this fight. For instance, Boukabous and Azizi proposed a hybrid solution that utilized the BERT deep learning algorithm and a lexicon-based dataset containing crime-related data [66]. Their real-time experiment using Twitter data yielded promising results, achieving an accuracy of 94.91% and an F1 score of 94.92% [66]. Similarly, Hannach and Benkhalifa proposed a hybrid model that leveraged WordNet semantic relations and a Term-Weighting scheme to identify specific aspects of crime data [67]. Their model outperformed Naive Bayes, SVM, and Random Forest classifiers when tested against two proposed approaches: implicit aspect sentence detection (IASD) and implicit aspect identification (ISI). These findings highlight the potential of WordNet for improved classification in crime-related text analysis [67].

Azeez and Aravindhar (2015) proposed two practical approaches that could benefit local police departments [68]. The first involves developing a visual analytics system to aid in decision-making, optimizing resource allocation, and officer deployment. The second approach focuses on utilizing real-time data from social media platforms like Twitter and Facebook. The authors emphasize the extensive research already conducted on the importance of data in crime analysis [68]. Gerber's research proposes a new model that utilizes tagged tweets for crime prediction [69]. This model incorporates Twitter-based linguistic analysis and statistical topic modeling to identify crime-related topics across different US cities. Experimental results demonstrated the model's effectiveness in reducing response times and improving crime identification. When compared to traditional methods, the proposed model achieved better results in 19 out of 25 crime prediction cases by leveraging Twitter data [69].

3.3.9 Disaster assessment, response, and management

Natural disasters are inherently unpredictable, causing immense distress due to their uncertain severity. While current technologies can predict disasters beforehand, pinpointing their exact intensity remains a challenge. Recent years have seen a rise in earthquakes, bushfires, cyclones, floods, and heat waves across the globe. Social media during these events is often filled with negative sentiments expressed by users. However, along with these emotions, users also share crucial information for authorities, such as the location of the disaster, its severity, people in distress, and the time of the incident. Research by Sufi and Khalil proposes a fully automated algorithm that leverages AI and NLP to extract location-based sentiment data related to global disasters [70]. This system can handle 110 languages and utilizes Twitter data for its analysis. The experiment yielded promising results, achieving an accuracy of 97% and an F1 score of 90%. The authors claim their study is the first of its kind to report on location-based disaster intelligence derived from sentiment analysis [70].

Similar research by Mendon et al. proposes a new hybrid framework that utilizes K-means clustering and TF-IDF to classify disaster-related sentiment [71]. Their experiment analyzed 243,746 tweets related to the 2018 Kerala floods in India. The findings shed light on the varying public emotions during the disaster, and how government and other stakeholders could leverage sentiment analysis to understand public sentiment based on location and time, ultimately leading to more effective disaster management [71]. However, the study does not include a user-friendly graphical interface (GUI) for government agencies to utilize in future disaster response efforts. Maharani investigated emergency response, situational awareness, and management during the Jakarta floods using sentiment analysis [72]. The research employed the BERT algorithm and a data crawler to collect tweets related to the flooding. The experiment focused on whether appropriate actions were taken based on public sentiment and how awareness was instigated. While achieving an accuracy of 90%, the author identified a significant limitation: the data used was not adequately cleaned [72].

Finally, another study introduced a big data approach for disaster response using sentiment analysis [73]. This model collects real-time data from Twitter and categorizes it based on the level of assistance required by people affected by the disaster. The experiment found that the lexicon-based approach for sentiment analysis outperformed all other methods, leading the authors to recommend it for future applications [73].

3.3.10 Demonetization

Demonetization refers to the process of withdrawing a particular currency denomination from circulation, rendering it invalid legal tender. This policy can have significant economic and social impacts, as seen in India's case when the government abruptly demonetized 500 and 1000 rupee notes in 2016 [74]. The stated objectives of demonetization were to curb terrorism financing, corruption, and inflation. However, the public perception was that the implementation was poorly planned and executed.

Research by Singh, Sawhney, and Khalon employed sentiment analysis to investigate the impact of the Indian government's demonetization policy [74]. While initial public opinion leaned heavily toward negativity, sentiment turned positive once new notes were released. The researchers analyzed Twitter data from all 30 Indian states. Nine states initially expressed strong opposition to demonetization, with residents highlighting challenges such as difficulty obtaining new notes due to rural location or agricultural background, lack of convenient banking facilities, and a perceived lack of warning, particularly in rural areas [74]. Similarly, Roy et al. conducted sentiment analysis on Twitter data to examine the public's perception of demonetization and identify areas for improvement in future policy implementation [75]. Their findings indicated that 45% of the public expressed positive sentiment, 22% were neutral, and 33% expressed negative sentiment [75].

4 Deep learning techniques for sentiment analysis

4.1 Deep learning versus other approaches

Deep learning has revolutionized sentiment analysis by offering several advantages over traditional methods. Unlike

approaches that rely on manually crafted features, deep learning models can automatically discover and learn relevant features directly from the data itself [76]. This eliminates the need for time-consuming and potentially error-prone feature engineering by human experts. Furthermore, deep learning models can capture complex relationships within the data that might be missed by handcrafted features.

This automatic feature learning capability is particularly beneficial when dealing with large datasets, which are commonplace in sentiment analysis tasks involving social media posts or customer reviews [77]. Deep learning algorithms excel at processing vast amounts of data due to their parallel processing capabilities and ability to learn from intricate patterns within the data. This makes them well-suited for handling the large datasets encountered in sentiment analysis.

Studies have shown that deep learning models can achieve superior accuracy compared to traditional methods [78]. Their ability to learn intricate relationships and adapt to new data leads to more robust and generalizable models that perform well on unseen data. Additionally, deep learning models offer versatility across various sentiment analysis tasks. They can be effectively applied for tasks beyond basic sentiment polarity classification, such as aspect-based sentiment analysis or even for identifying emotions within text data.

In contrast, deep learning offers a powerful and versatile approach to sentiment analysis. Its ability to automatically learn features, handle large datasets, achieve high accuracy, and adapt to various tasks makes it a valuable tool for researchers and practitioners working in this field.

4.2 Deep learning techniques

Extensive research has been conducted on sentiment analysis and traditional machine learning algorithms like random forest, Naive Bayes, and support vector machines (SVMs) are increasingly being outperformed by newer methods. This shift is driven by challenges related to data complexity, algorithm performance, and processing limitations [79].

According to [79], support vector machines and conventional neural networks struggle to compete with deep learning networks for sentiment analysis tasks. Deep learning networks excel at handling massive amounts of data, allowing them to outperform traditional algorithms. However, the optimal algorithm choice within deep learning depends on the specific application and dataset involved. Recent research suggests that recurrent neural networks (RNNs) are surpassing the previously preferred convolutional neural networks (CNNs) in sentiment analysis [79]. For tasks involving long-term dependencies within the data, two variations of RNNs are particularly well-suited: gated recurrent units (GRUs) and long short-term memory (LSTM) networks [79].

4.2.1 Convolutional neural network (CNN)

Introduced in 1989, convolutional neural networks (CNNs) are a type of feed-forward neural network [80]. Initially, CNNs were applied in the field of artificial intelligence (AI) to identify the visual cortex in animals. In this context, each neuron within the CNN analyzes a small receptive field. These receptive fields overlap, allowing the network to progressively build a representation of the entire object. These overlapping regions are also referred to as filters within CNNs [81]. A CNN architecture typically consists of three main layers. Input Layer: This layer receives raw input data and transforms it into a suitable format for further processing. Feature Extraction Layer: This layer is comprised of two sub-layers: convolutional layers and pooling layers. Convolutional layers identify and learn new features within the data, while pooling layers downsample the data, reducing its dimensionality. Classification Layer: This layer uses fully connected networks with a classifier to categorize the extracted features and make predictions.

Collobert and Weston (2008) were the first to successfully apply CNNs to natural language processing (NLP) tasks [82]. Since then, CNN models have achieved impressive results in various NLP domains, including sentiment analysis [83, 84]. Research by [85, 86] explored approaches and models utilizing CNNs for sentiment classification. One of the advantages of CNNs is their efficiency. Compared to traditional neural networks with fully connected layers, CNNs have fewer parameters, leading to faster training times. Additionally, CNNs excel at learning contextual features within text data through the use of filters. However, CNNs also have limitations in sentiment analysis tasks. When dealing with long-term dependencies within text data, CNNs require a deep network architecture, which can be computationally expensive [87, 88]. This limitation paved the way for the introduction of recurrent neural networks (RNNs).

4.2.2 Recurrent neural network (RNN)

Recurrent neural networks (RNNs) are a type of neural network that departs from the traditional feed-forward structure. In contrast to feed-forward networks, RNNs incorporate a memory element, allowing them to retain information from previous computations and use it to inform subsequent processing. This capability is rooted in Elman's principle (1980), which emphasizes the importance of memory in sequential information processing [89]. RNNs offer several advantages for tasks involving sequential data, such as text analysis. First, they are adept at modeling sequences of varying lengths, enabling them to handle long-range dependencies within the data [90]. This means that RNNs can effectively capture the context of a sequence, even if relevant information appears earlier in the sequence.

However, RNNs also have limitations. One challenge is the issue of vanishing or exploding gradients, which can occur when processing long sequences [91]. These gradient problems can hinder the network's ability to learn effectively from long-range dependencies. Additionally, RNNs struggle with sequential data that doesn't follow a linear structure, such as tree-like data. To address these limitations, several variations of RNNs have been developed, including long short-term memory (LSTM), bidirectional LSTM (BiLSTM), gated recurrent unit (GRU), and bidirectional GRU (Bi-GRU). These variants aim to overcome the shortcomings of the original RNN architecture [91]. Despite these limitations, RNNs have demonstrated exceptional performance in sentiment analysis tasks. Experimental results continue to support the effectiveness of RNNs in this domain [90].

4.2.3 Long short-term memory (LSTM)

In 1997, Hochreiter and Schmidhuber proposed the long short-term memory (LSTM) network to address the vanishing gradient problem that plagued traditional RNNs [92]. LSTMs achieve this by incorporating a gating mechanism into the RNN architecture. This gating mechanism allows memory cells to store information for longer durations and retrieve past computational results as needed. While LSTMs offer a powerful solution, their architecture can be complex. To address this complexity, the gated recurrent unit (GRU) was introduced as a simpler alternative [92].

LSTMs are widely used in sentiment analysis tasks due to their effectiveness. One advantage of LSTMs is their ability to process information in both directions, unlike traditional unidirectional RNNs. This bidirectional capability (known as BiLSTM) is particularly valuable for sentiment analysis, as it allows the network to identify aspects and entities that contribute to the overall sentiment within a text sequence [92]. Internally, LSTMs utilize three gating mechanisms, often referred to as gates. Input Gate: This gate determines what information from the current input and the previous cell state will be stored in the cell state. Forget Gate: This gate decides what information to forget from the cell state. Output Gate: This gate controls the information that gets passed on to the next step in the network.

4.2.4 Gated recurrent unit (GRU)

The gated recurrent unit (GRU) is a variant of the recurrent neural network (RNN) architecture, similar to the long short-term memory (LSTM) network [93]. However, unlike LSTMs with their three gates, GRUs utilize a simpler architecture with only two gates. These two gates are responsible for controlling the flow of information within the network. Reset Gate: This gate combines new input data with information from the previous cell state, effectively determining

what information needs to be updated. Update Gate: This gate controls which information from the previous cell state is retained and which information is discarded.

By streamlining the gating mechanism, GRUs offer a more computationally efficient alternative to LSTMs while maintaining a similar level of effectiveness [93]. This efficiency makes GRUs a valuable option for tasks where computational resources are limited. The viability of GRUs in sentiment analysis tasks has been explored through research experiments [93]. Based on the results of these studies, GRUs can be further modified and adapted for various research applications. Similar to LSTMs, GRUs also have a bidirectional variant known as Bi-GRU.

4.2.5 BERT

BERT (Bidirectional Encoder Representations from Transformers) is a groundbreaking natural language processing model developed by Google in 2018 [5]. Since its introduction, it has revolutionized our ability to understand language. BERT's secret weapon is its bidirectional architecture, allowing it to consider both the preceding and following words in a sentence, resulting in a deeper contextual understanding [5].

This powerful model leverages a unique transformer architecture that employs a self-attention mechanism. This mechanism allows BERT to capture and identify relationships between words within a sentence. BERT utilizes a two-step approach: pre-training and fine-tuning. Through pre-training on massive amounts of text data using tasks like masked language modeling and next-sentence prediction, BERT learns rich word representations. It can also tokenize text into subword units, enabling it to handle various linguistic complexities. BERT's applications extend to text classification, translation, sentiment analysis, and more, consistently setting new benchmarks across various NLP domains [5]. It has inspired subsequent models and remains a cornerstone in modern NLP, propelling advancements in understanding and generating human-like text.

4.2.6 Large language models (LLMs)

Large language models (LLMs) have emerged as a powerful force in the field of natural language processing (NLP) [94]. These are complex AI models trained on massive datasets of text and code, enabling them to understand and generate human-like language. They possess a remarkable ability to grasp intricate relationships between words, analyze context, and perform a variety of NLP tasks, including sentiment analysis. LLMs are trained on vast amounts of text data, encompassing diverse sources like books, articles, social media posts, and code repositories. This exposure allows them to learn complex language patterns and nuances that

are crucial for accurate sentiment analysis [95]. Unlike simpler models, LLMs excel at understanding the context within a sentence. They can analyze the surrounding words, sentence structure, and even cultural references to determine the intended sentiment, even when the explicit words might be ambiguous. This is particularly valuable for tasks like sarcasm detection, where the literal meaning of the text might differ from the underlying sentiment [95].

LLMs can be fine-tuned for specific sentiment analysis tasks by leveraging their pre-trained knowledge. This significantly reduces the time and resources required to develop task-specific sentiment analysis models, making them readily applicable to various domains [95]. The impact of LLMs on sentiment analysis is undeniable. They offer greater accuracy, adaptability, and efficiency compared to traditional methods. As LLM technology continues to evolve, we can expect even more sophisticated sentiment analysis capabilities, unlocking valuable insights from textual data across diverse fields like customer reviews, social media monitoring, and market research [95].

4.2.7 Graph neural networks (GNNs)

Sentiment analysis has traditionally relied on methods that analyze text as a sequence of words. However, a recent advancement in the field involves the use of graph neural networks (GNNs) for sentiment analysis tasks. GNNs are a type of deep learning architecture specifically designed to work with data represented as graphs, where nodes represent entities (like words in a sentence) and edges represent relationships between them [96]. GNNs can effectively capture the complex interdependencies between words within a sentence. By analyzing the relationships between words, GNNs can understand how the sentiment of one word can be influenced by others. This is particularly valuable for tasks like aspect-based sentiment analysis, where the sentiment toward different aspects of a product or service needs to be identified [96]. GNNs can seamlessly integrate external knowledge bases into the sentiment analysis process. These knowledge bases can include information about word ontologies, synonyms, and even sentiment lexicons. By incorporating this additional context, GNNs can achieve a more nuanced understanding and improve sentiment classification accuracy [97].

GNNs are well-suited for analyzing complex text structures that go beyond simple linear sequences. For example, they can effectively handle nested sentences, sarcasm detection (where sentiment can be conveyed through the structure of the sentence rather than individual words), and sentiment analysis in dialog settings (where the sentiment of one speaker can influence the sentiment of the next) [96]. While GNNs offer promising potential for sentiment analysis, there are also challenges such as scalability and interpretability that need to be addressed [97]. Despite these challenges, the

potential benefits of GNNs in sentiment analysis are significant. As research progresses and these challenges are addressed, GNNs are poised to become a powerful tool for extracting deeper insights from textual data and understanding the nuances of human sentiment.

4.3 Strengths and limitations of deep learning techniques

Deep learning algorithms offer a powerful approach to sentiment analysis, but it is crucial to understand their inherent strengths and limitations. Table 3 highlights the strengths and limitations of the different deep learning algorithms that were identified within this literature.

In the upcoming section, we will look into how the above-mentioned deep learning architectures perform against several attributes in an experimental environment.

5 Performance evaluation and challenges in sentiment analysis

Despite significant advancements, sentiment analysis still faces challenges that can hinder its accuracy. This section explores both performance evaluation methods and the key challenges encountered in sentiment analysis.

5.1 Sentiment analysis performance

Evaluating the effectiveness of a sentiment analysis model is crucial for ensuring its reliability and usefulness in real-world applications, here we explore some of the recent experiments conducted using the seven deep learning algorithms against sentiment analysis tasks, dataset and size, and language.

Sentiment analysis has been experimented with using classical machine learning algorithms, deep learning algorithms, and hybrid approaches. In this section, we will look into the five deep learning algorithms and tabulate them individually to better understand the algorithm performance based on a specific dataset. The table was designed to collect and compare publication year specifically, sentiment analysis tasks, language, algorithm or model, the dataset used, and the scores. All these are important in understanding how the experiment was conducted and how the accuracy and F1 scores were attained.

We prioritized research articles published between 2019 and 2024 to ensure our analysis reflects the latest advancements. The following tables provide an overview of the number of experimental papers published within the past year, categorized by specific deep learning algorithms. Table 4 showcases recent experimental papers that utilize convolutional neural networks (CNNs) for sentiment analysis. Our search identified only seven prominent papers in this

Table 3 Strengths and limitations of deep learning algorithms

Deep learning techniques	Strengths	Limitations
CNN [16, 98–100]	Fewer parameters need to be trained Computationally faster and less expensive compared to LSTM, GRU, and RNN Ability to extract features from sentences LSTM combined with CNN gets better performance Ability to identify patterns CNN performs best with emotion-based sentences and long sentiment Is non-linear	Has reservation issues with preserving long-term dependency Huge demand for data Ignores dependency features that help understand semantic and syntactic sentences
RNN [16, 98, 101]	Performs better than CNN by focusing on the word order (sequential data) in the data Ability to save past computations Maintains fewer parameters through weight sharing Has the ability to capture long-distance dependence traits that reflect syntactic and semantic information that CNN ignores Assists in maintaining the word order in an input sequence The ability to create a fixed-size vector that summarizes the sequence while taking into consideration the weighted combination of all words thanks to a distributed hidden state that can store previous calculations No need for a huge dataset	Inability to process extremely lengthy sequences It can be disadvantageous for NLP applications like sentiment analysis due to the vanishing gradient and expanding gradient problem Requires fewer parameters The final concealed state is chosen to represent the sentence, which could result in an inaccurate prediction Neglecting to identify long-term dependencies
LSTM [16, 102–104]	Can selectively forget and remember things, which is significantly better for categorizing feelings Capable of keeping track of inputs from earlier steps to prevent the vanishing and growing gradient problem in an RNN Outperforms CNNs by a wide margin thanks to its capacity to account for the text's sequence when predicting mood Text can assist in representing and encoding a sentence's semantics Can be utilized for aspect-based sentiment analysis to extract sequential information in the aspect target sequence	Applying back propagation after calculating the output each time is computationally intensive Finding the best approach could be more difficult because each weight needs to be trained separately. Additionally, this substantially slows down this network To obtain the label for the complete sentence, the output produced at each time step (for each input word) must be reconciled to the full phrase
GRU [16, 105, 106]	It has fewer gates, making it faster and more affordable computationally Addresses the issue of vanishing and exploding gradients that RNN has The structure is less complex than LSTM Can be helpful in document-level sentiment classification for capturing interdependencies across sentences of review documents Able to detect emotion in a clamorous speech	Lacks a memory unit, exposing all concealed information without control Due to their greater capacity for memory retention, LSTMs outperform GRUs on larger texts
BERT [5, 107]	BERT is effective for task-specific models Suitable for smaller, specific tasks due to massive training corpus Metrics are adjustable and immediately applicable Constant model updates for exceptional accuracy Achieved through effective fine-tuning Pre-trained BERT versions for 100 + languages are available Beneficial for non-English initiatives	The model is large due to the training framework and corpus Training time is lengthy due to size and weight updates Costly due to increased computing needs Needs adjustment for downstream tasks, designed for integration rather than standalone use

Table 3 (continued)

Deep learning techniques	Strengths	Limitations
LLMs [108, 109]	LLMs can learn some aspects of language structure and meaning without explicit symbolic rules Transformers have achieved excellent results on a broad range of NLP tasks, including sentiment analysis, natural language inference, and question answering While some argue LLMs struggle with generalization, the paper acknowledges they can perform well on new data sets in certain cases Even if LLM learning differs from humans, their success suggests alternative ways to acquire language knowledge from data Research is ongoing to improve LLM performance in areas like abstract reasoning and real-world knowledge tasks Transformers can be trained on various data types (text, images, etc.) potentially leading to more comprehensive understanding	LLMs are argued to mainly combine existing information from their training data, lacking true innovation LLMs lack semantic grounding, meaning they cannot connect words to real-world objects or intentions LLMs do not represent symbolic systems like grammar, which may be crucial for language understanding While some propose combining LLMs with symbolic systems, current efforts haven't shown significant improvement It's debated whether LLMs learn language similarly to humans. Even if different, it doesn't necessarily make them ineffective LLMs are evaluated based on task performance, not necessarily how they achieve those results. This can be seen as a weakness by those who prioritize understanding human cognition
GNNs [110, 111]	Well-suited for representing and reasoning about graphs (networks) Can capture complex relationships between nodes and edges Can learn generalizable patterns from graph data and be applied to various graph-structured tasks They can learn node and edge representations directly from graph data, with no need for pre-defined features Can model complex graph structures with various message-passing mechanisms Can be parallelized for efficient training and inference on large graphs (depending on the specific GNN architecture) Can be combined with other deep learning models (e.g., CNNs, RNNs) for tasks involving both graph and non-graph data	Difficulty in handling large and dynamic graphs May struggle with generalizing to unseen graph structures Learned features might be difficult to interpret May struggle with very deep or long-range dependencies in graphs Training complex GNNs can still be computationally expensive Requires careful design to ensure compatibility and efficient information flow

Table 4 CNN for sentiment analysis

References	Year	Sentiment analysis tasks	Language	Deep learning model	Dataset	Dataset size	Accuracy or F1 score
[112]	2021	Consumer sentiment analysis	English	CNN-LSTM	Twitter Airline Data	14,640	Acc = 91.3% & F1 = 87.5%
[113]	2021	Aspect-based sentiment classification	English	CNN	SemEval 2014	5000	Acc = 88.5% & F1 = 88.4%
[114]	2022	Document-based sentiment analysis	English	CNN	IMDB	50,000	Acc = 92.0%
[115]	2019	Sentiment analysis	Hindi	CNN	Hindi Movie Reviews	7354	Acc = 95.4%
[116]	2020	Sentiment analysis	English	CNN-LSTM	Twitter Dataset	1,600,000	Acc = 81.2%
[117]	2020	Sentiment analysis based on sentiment lexicon	Chinese	SLCABG	Chinese e-Commerce Site Review	100,000	Acc = 93.5% & F1 = 93.3%
[4]	2020	Sentiment classification	Turkish	CNN-LSTM	Twitter Dataset	17,289	Acc = 82.1%

category. Notably, some of these papers explore hybrid approaches, combining CNNs with other algorithms to improve model performance.

Table 5 highlights the use of recurrent neural networks (RNNs) in sentiment analysis. Based on the performance results, RNNs often rely on hybrid approaches to achieve improvements compared to other algorithms. According to the accuracy and F1 scores, RNN should not be considered one of the major deep learning techniques to apply.

Table 6 highlights the use of long short-term memory (LSTM), a variant of the recurrent neural network (RNN) architecture, in sentiment analysis. LSTMs are widely used for sentiment analysis tasks and have demonstrated competitive accuracy. However, the findings suggest that LSTMs often benefit from being combined with other algorithms to achieve optimal performance. While the introduction of BiLSTM (bidirectional LSTM) did not yield significant improvements in this specific case, its effectiveness may vary depending on the sentiment task being analyzed.

Table 7 explores the use of the gated recurrent unit (GRU) algorithm, another variant of the RNN architecture similar to LSTM. The findings suggest that GRUs are not currently the preferred choice for sentiment analysis experiments. This aligns with the observed trend of hybrid approaches, where GRU performance also appears to benefit from being combined with other algorithms.

Table 8 highlights the use of BERT, a currently trending algorithm in the NLP domain. This table showcases a significant number of published papers (20) that explore sentiment analysis using BERT. Notably, BERT demonstrates strong performance not only with English datasets but also exhibits comparable effectiveness with other languages. Furthermore, BERT achieves competitive results even without fine-tuning, surpassing other deep learning algorithms on its own. When combined with other algorithms in a hybrid approach, BERT's performance remains unmatched, setting a benchmark that other hybrid models struggle to reach.

Table 9 highlights the use of large language models (LLMs) and in this case, we can see that LLMs work well as a hybrid model. However, comparing the accuracy and F1 with the likes of BERT and GNN shows that this approach is not a suitable option.

Table 10 highlights the use of graph neural networks (GNNs). Based on the accuracy and F1 scores, GNN models show promising results and in some instances, results may be better than other deep learning techniques that offer similar experimental approaches.

Based on the findings presented in Tables 4–10, Section VI will discuss the performance variations of deep learning algorithms for sentiment analysis.

5.2 Challenges in sentiment analysis

In the past few years, we have seen an increase in the number of publications published in sentiment analysis using deep learning algorithms or publications trying to focus on a specific sentiment task. Within these publications, the experiments carried out have some limitations and these can be classified as challenges because these are some common challenges that have been noticed in other publications. Although publications and research are being carried out to solve the limitations and gaps identified, many feel that this will take an extended time. Table 11 highlights some of the recent challenges observed in various publications ranging from 2019 to 2024.

6 Discussions

In the below section, we look into the various attributes of sentiment analysis after analyzing them and pay attention to the six research questions.

6.1 Sentiment analysis tasks

Building on the findings from the literature review, this section delves into sentiment analysis tasks. The literature review highlighted various challenges in sentiment analysis [31]. These challenges have now been categorized into specific tasks. Addressing these tasks can significantly advance research outcomes.

Figure 4 illustrates the current research focus on sentiment analysis tasks. Notably, aspect-based sentiment analysis and general sentiment analysis remain key areas of investigation as researchers strive to improve model accuracy. Beyond these core tasks, researchers are exploring new and emerging sentiment analysis tasks that present exciting opportunities for further exploration.

While general sentiment analysis and aspect-based sentiment analysis remain dominant research areas, a significant opportunity exists to explore other sentiment analysis tasks. Among the reviewed studies on general sentiment analysis, research using CNNs with Hindi movie review data achieved an accuracy of 95.4%. This represents the highest reported accuracy within this category. However, it is important to consider that the chosen dataset (Hindi movie reviews) might not be generalizable to other domains like social media (Twitter) or movie reviews in a different language (IMDB). Therefore, the accuracy might vary depending on the dataset used. Similar considerations apply to the accuracy reported for aspect-based sentiment analysis.

Table 5 RNN for sentiment analysis

References	Year	Sentiment analysis tasks	Language	Deep learning model	Dataset	Dataset size	Accuracy or F1 score
[113]	2021	Aspect-based sentiment classification	English	RNN	SemEval2014	5000	Acc = 54.2% & F1 = 64.2%
[59]	2020	Sentiment analysis	English	RNN	Sentiment140	1,600,000	Acc = 90.5% & F1 = 94.1%
[118]	2021	Aspect-based sentiment analysis	Arabic	RNN-SVM	Arabic Online Review Dataset	1513	Acc = 95.4% & F1 = 93.4%
[119]	2020	Aspect-based sentiment analysis	English	RNN-RecNN	SemEval2014	5000	Acc = 81.4%
[120]	2024	Sentiment analysis	English	RNN + LSTM	Twitter	100,000	Acc = 91.0%
[121]	2024	Sentiment analysis	English	SA-RNN-BERT	University Online Comments	3820	Acc = 84.0% & F1 = 83.0%
[122]	2023	Deep sentiment analysis	English	D-RNN (DSA)	Twitter	10,000	Acc = 86.0% & F1 = 89.8%

Table 6 LSTM for sentiment analysis

References	Year	Sentiment analysis tasks	Language	Deep learning model	Dataset	dataset size	Accuracy or F1 score
[123]	2021	Sentiment analysis using sentence-state	Sinhala	S-LSTM	Sinhala Online Newspaper Reviews	4,500,000	Acc = 87.9% & F1 = 87.9%
[124]	2021	Transparent aspect-level sentiment analysis	English	BiLSTM-DGCN	Twitter Dataset	6051	Acc = 75.4% & F1 = 73.1%
[113]	2021	Sentiment analysis	English	LSTM-LDA	Online Phone Reviews of OPPO, Huawei, and Xiaomi	31,825	Acc = 89.5%
[125]	2020	Sentiment analysis	English	LSTM	Amazon Review Dataset	120,000	Acc = 70.0% & F1 = 70.3%
[8]	2022	Sentiment analysis	Chinese	LSTM	Chinese Text Corpora	3483	Acc = 83.2%
[126]	2020	Aspect-based sentiment analysis	English	BiLSTM + KGCapsAN	Twitter Dataset	6248	Acc = 88.5% & F1 = 70.7%
[127]	2024	Document-level model	English	LSTM + CNN	Maritime Law Legislation Data	98,000	Acc = 98.05%

The sentiment analysis tasks described above have been the subject of extensive research and experimentation. However, a significant portion of this research has focused on general sentiment analysis, which doesn't delve into the specific aspects being evaluated. Future research directions aim to move beyond basic sentiment classification. The goal is to develop techniques that can:

- Differentiate between finer-grained sentiment categories: This includes tasks like aspect extraction and categorization, where the system identifies not just the overall sentiment but also the specific aspects of an entity being evaluated (e.g., positive sentiment toward a phone's camera but negative sentiment toward its battery life).

Table 7 GRU for sentiment analysis

References	Year	Sentiment analysis tasks	Language	Deep learning model	Dataset	Dataset size	Accuracy or F1 score
[128]	2022	Aspect-level sentiment analysis	English	CPA-SA	SemEval2014	400,000	Acc = 89.0% & F1 = 72.4%
[129]	2021	Intent-based sentiment analysis	English	GRU	IMDB	50,000	Acc = 89.0%
[125]	2020	Sentiment analysis	English	GRU	Amazon Review Dataset	120,000	Acc = 71.0% & F1 = 71.3%
[130]	2019	Aspect-based sentiment analysis	English	B-RDAM	SemEval2015	2000	F1 = 87.2%
[131]	2021	Aspect-based sentiment analysis	Chinese	GRU-CNN	Chinese Online Review	120,000	Acc = 89.6% & F1 = 89.6%
[132]	2024	Sentiment analysis	English	MPNet-GRUs	IMDB	50,000	Acc = 94.71%
[133]	2023	Sentiment analysis	Chinese	T-E-GRU	Chinese Comments Dataset	700,000	Acc = 90.09% & F1 = 90.07%

- Analyze sentiment intersubjectivity: This involves understanding whether an opinion expressed is personal or reflects a broader sentiment.
- Predict financial volatility: Sentiment analysis can be used to analyze financial news and social media data to potentially predict fluctuations in the market.
- Extract and categorize entities: This involves identifying and classifying the entities (people, organizations, products) being discussed in text data.
- Expand sentiment lexicons: This refers to developing more comprehensive dictionaries of words with emotional associations.
- Perform advanced analysis tasks: This includes tasks like sarcasm analysis, opinion polarity classification (strength of positive or negative sentiment), emotion cause detection and classification, and even recommending opinions based on user preferences.

The research article by [31] identified these and other potential areas for future exploration in sentiment analysis.

6.2 Deep learning algorithms

The literature review indicates that traditional approaches, such as lexicon-based and rule-based methods, as well as machine learning algorithms like Naive Bayes, Support Vector Machines, and decision trees, are no longer considered the most suitable for sentiment analysis [1]. The review highlights that hybrid models and, particularly, deep learning algorithms are now the preferred approaches for sentiment analysis [1]. Figure 5 shows the most commonly used deep learning algorithms. BERT by far is the most preferred option due to it having the following attributes that include contextual understanding, the ability to train large corpora, transfer learning, word and sentence embeddings, handling out-of-vocabulary, and performing fine-grained sentiment analysis.

However, even BERT's performance is highly dependent on the chosen dataset. As mentioned earlier, datasets can vary in size, quality, and suitability for training a specific model. Figure 5 highlights that SemEval and Online Reviews datasets are generally considered as decent choices. While the Twitter dataset has not yet been explored with this model, Figs. 6 and 7 illustrate the impact of dataset selection on accuracy. The online reviews dataset yields BERT accuracy scores ranging from 84.0 to 97.7%, while the SemEval dataset achieves accuracy between 87.0 and 95.2%. These accuracy ratings are questionable because only three datasets (IMDB, Catering Industry Dataset, and SST) had data sizes that exceeded 50,000 sentences. IMDB had 50,000 with an accuracy of 95.2%, Catering Dataset had 105,000 with an accuracy of 88.6%, and SST had 400,000 with an accuracy of 88.6%. Other datasets had data anywhere from 2961 to 16,000 sentences while experimenting with BERT.

Similarly, when we compare CNN, RNN, LSTM, GRU, LLM, and GNN with the online reviews dataset we can expect to see similar results where CNN and RNN both achieve an accuracy score of 95.4%. LSTM achieves an accuracy of 89.5% while GRU yields 89.6%. BERT outperforms all these four models having achieved 97.7% accuracy. To get the best results, we need to experiment using the same dataset and the same preprocessing tasks to accurately compare the models and determine which is best for future use cases.

Table 8 BERT for sentiment analysis

References	Year	Sentiment analysis tasks	Language	Deep learning model	Dataset	Dataset size	Accuracy or F1 score
[134]	2022	Document-level sentiment analysis	English	BERT	Sentihood	3845	Acc = 93.8%
[135]	2022	Sentiment analysis	English	BERT	n/a	n/a	Acc = 94.0% % F1 = 83.0%
[136]	2022	Target-level sentiment analysis	Slovene	BERT	SentiCoref 1.0	10,427	Acc = 77.0 & F1 = 73.0%
[137]	2022	Fine-grained sentiment analysis	English	CRF-BERT	Online Reviews Crawled from Taobao, Tmall, & Jingdong	108,276	F1 = 94.1%
[129]	2021	Intent-based sentiment analysis	English	BERT	IMDB	50,000	Acc = 95.2%
[138]	2022	Sentiment analysis	English	SiBERT	SST-2 Dataset	400,000	Acc = 88.6%
[139]	2022	Sentiment orientation prediction	Chinese	BERT-BiLSTM	Chinese Hotel Reviews	16,000	Acc = 96.5% & F1 = 96.4%
[140]	2019	Aspect-based sentiment analysis	English	BERT	SemEval 2015 & 2016	6328	Acc = 89.8% & F1 = 85.0%
[141]	2020	Sentiment analysis	English	BERT4SentiSE	Stack OverFlow Posts	5500	F1 = 87.0%
[142]	2021	Sentiment analysis	Arabic	BERT	ASTD	10,000	Acc = 91%
[143]	2020	Sentiment analysis	Turkish	BERT	IMDB	50,000	Acc = 93.3%
[144]	2021	Aspect-based sentiment analysis	English	BERT-SAN	SemEval 2014	n/a	Acc = 87.6%
[145]	2021	Aspect-based sentiment analysis	Persian	Pars-BERT	Pars–ABSA Dataset	5602	Acc = 91.0% & F1 = 90.0%
[146]	2021	Text aspect-level sentiment analysis	English	BERT-CBiGRU + ATR	Catering Industry Dataset	105,000	Acc = 88.6% & F1 = 77.1%
[147]	2020	Aspect-based sentiment analysis	Vietnamese	m-BERT + vi-BERT	SemEval 2016	2961	F1 = 84.2%
[148]	2020	Aspect-based sentiment analysis	English	BERT-IAN	SemEval 2014	2328	Acc = 80.4%
[149]	2021	Implicit aspect extraction	English	BERT	SemEval 2014 — 2016	8290	Acc = 78.1% & F1 = 73.2%
[150]	2020	Aspect-based sentiment analysis	Indonesian	m-BERT	Airyrooms Dataset	9448	F1 = 97.7%
[151]	2021	Multi-domain aspect extraction	English	MDAE-BERT	SemEval 2014	3644	Acc = 90.0% & F1 = 65.9%
[152]	2024	Text sentiment analysis	Chinese	BERT-CNN-BiLSTM-Att	Chinese Online Movie Reviews	400,000	Acc = 95.7% & F1 = 94.6%

Table 9 LLMs for sentiment analysis

References	Year	Sentiment analysis tasks	Language	Deep learning model	Dataset	dataset size	Accuracy or F1 score
[153]	2024	Sentiment analysis	Arabic	LLM + Ensemble	SemEval2017	1823	Acc = 86.71% & F1 = 81.06%
[154]	2024	Financial sentiment analysis (domain specific)	English	BLOOMZ (HAD)	CMC	13,545	Acc = 87.67% & F1 = 92.95%
[94]	2023	Sentiment analysis	English	GPT3.5 + GPT4 + InstructGPT3.5 (LLM Negotiations)	Yelp	598,000	Acc = 96.3%
[155]	2023	Financial sentiment analysis (domain specific)	English	LLM + RAG	Twitter	2388	Acc = 88.1% & F1 = 84.2%
[156]	2023	Market sentiment analysis (domain specific)	English	PaLM-540B	Reddit Posts	20,000	Acc = 72.0%
[157]	2024	Sentiment classification	English	Llama 2	Amazon	N/A	Acc = 93.5%
[158]	2024	Emotion analysis	English	Distilled LLM	Financial Dataset	1,000,000	F1 = 84.0%

Table 10 GNNs for sentiment analysis

Reference	Year	Sentiment analysis tasks	Language	Deep learning model	Dataset	Dataset size	Accuracy or F1 score
[96]	2024	Sentiment analysis	Persian	RGCN	Digikala	100,000	Acc = 91.17% & F1 = 74.15%
[159]	2022	Sentiment analysis	Chinese	GNN-LSTM	Weibo	120,000	Acc = 95.25% & F1 = 95.22%
[160]	2021	Sentiment analysis	Chinese	GACNN	SMP-ECISA2019	19,917	F1 = 88.16%
[161]	2023	Aspect-based Sentiment analysis	English	RDGCN + BERT	SemEval2014	3044	Acc = 87.49% & F1 = 81.16%
[162]	2024	Sentiment analysis	English	BUGE	Movie Review	10,662	Acc = 80.44% & F1 = 79.73%
[163]	2024	Aspect-based Sentiment analysis	English	TextGT + BERT	SemEval2014	3044	Acc = 92.21% & F1 = 81.48%
[164]	2021	Sentiment analysis	English	TGNN	SST	9613	Acc = 83.1%

When it comes to selecting an online reviews dataset, we should be careful as these generally could be manipulated and yield a positive or higher accuracy. It is best to carry out experiments with different datasets to see if you are getting similar results. Not all deep learning architecture may be suitable for a specific dataset which is why people opt for hybrid solutions.

6.3 Applications

Sentiment analysis plays a pivotal role in various application domains, including education, brand monitoring and business intelligence, social media, finance, and stock market monitoring, customer experience and service enhancement, market research and analysis, politics, crime prediction, disaster management, and demonetization. It creates value in healthcare by analyzing patient feedback and mental health. Product reviews and brand impressions help e-commerce, while consumer sentiment insights help travel and hospitality improve services. The entertainment industry benefits

Table 11 Challenges identified in sentiment analysis tasks

| References	Sentiment analysis tasks	Challenges
[165]	Sentiment analysis	The author focuses on Sentiment Analysis of Code-Mixed Data, Anaphora, and Coreference Resolution, Spam Detection, Negation Handling, Word Sense Disambiguation, Low-resource Languages, and Sarcasm Detection
[166]	Sentiment analysis	Challenges include Fake Reviews, Accuracy improvement of subjectivity Detection, Subjectivity Detection, Context dependency, and Computational cost
[20]	Sentiment analysis	Subjectivity and Tone, Setting and Polarity
[112]	Sentiment analysis	Various optimization methodologies, as well as machine learning models for sentiment analysis, can be used to speed up the process. These techniques include feature selection, dimensionality reduction, model pruning, and the use of hardware accelerators. The goal is to reduce processing time while preserving or improving model performance, resulting in faster results and better decision-making
[167]	Sentiment analysis	In sentiment analysis, dealing with cynicism, mockery, and consequences is a significant problem. When addressing the opinions, subject or phrase variances must also be considered. If we assume that there is some relationship between words with equal polarity and reviews, then a list of keywords may help determine polarity. Datasets for sentiment analysis are extremely domain-specific, yet at the same time, assembling or creating them is difficult
[168]	Sentiment analysis	Heterogeneous characteristics of big data, analyzing sparse, uncertain, and incomplete data, semantic relations in multi-source data fusion Does sentiment analysis help in designing enterprise strategies? The influence of the post and the impact of social bots need to be looked into
[169]	Sentiment analysis	The time-consuming nature of CNN and RNN algorithms in temporal processing is a significant disadvantage, as character-level analysis takes longer than word- and sentence-level structures
[98]	Aspect-based sentiment analysis	The author highlights Domain Adaptation, Technical Requirements, Multilingual Application, and Linguistic Complications
[170]	Aspect-based sentiment analysis	Researchers need to focus on domain-dependent study, data preprocessing, multilingual sentiment analysis, classifying multiple targets, and multimodal Analysis
[171]	Aspect Category sentiment analysis	ACSA frequently lacks direct sentence presence, making it difficult for models to reliably detect their context and placement
[172]	Multi-classification sentiment analysis	There has been little research on identifying hidden emotions in brief text
[151]	Multi-domain aspect extraction	Inconsistency of aspects from target and source domains and context-based semantic distance between ambiguous aspects
[28]	Multimodal sentiment analysis	A problem that still has to be solved is domain adaptation, such as applying a model developed for sentiment analysis of product reviews to the analysis of microblog postings. How to handle ambiguous circumstances and irony is another key difficulty for sentiment analysis. As an illustration, a sarcastic compliment is intended to send a negative message
[173]	Multilingual sentiment analysis	Unexplored issues in MSA include less-studied levels, developing setups, different representations, sentiment-specific techniques, and low-resource languages and dialects
[30]	Opinion summarization	It is difficult to navigate through all the daily tweets to analyze important topics. Blogs and social media posts lack any established guidelines, are utterly chaotic, and noisy, and have a casual dialect. They frequently include emoticons, sarcasm, and terms that aren't found in dictionaries. Additionally, they contain many spelling, grammatical, punctuation, and capitalization errors as well as constantly changing rules, and are written by non-specialists. This suggests that we cannot use a dictionary or knowledge base

Table 11 (continued)

References	Sentiment analysis tasks	Challenges
[174]	Opinion mining	How can one tell which user reviews compare two applications? How can we interpret user preferences and opinions based on comparison reviews? Comparisons are made indirectly regarding the precision of comparative sentence mining, new applications, generic-sounding apps, improvement in precision, analysis of the market at large, topic distribution, user studies, and tool support
[124]	Transparent aspect-level sentiment analysis	First, a few models consider the texts' natural language semantic properties. Second, many models take into account the target words' locational features while ignoring the connections between individual target words and between entire sentences. Thirdly, many models in sentiment analysis lack transparency in data gathering, processing, and result generation. Through dependency syntax analysis, the proposed BiLSTM-DGCN model links the target words with the related emotive terms

Fig. 4 Sentiment Analysis Tasks
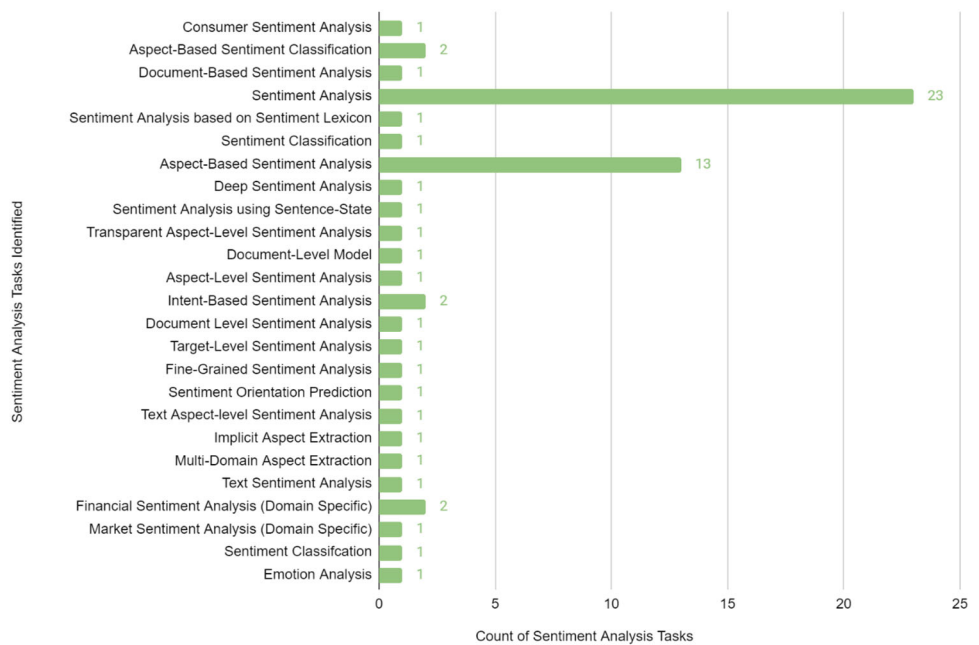
Fig. 5 Deep learning algorithms used in SA
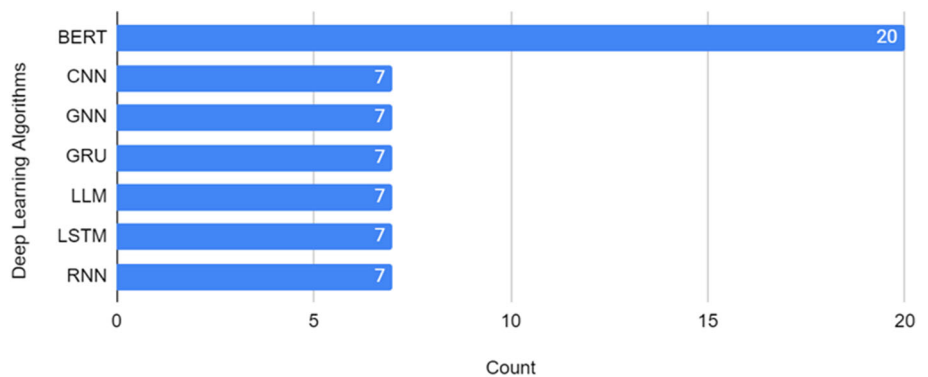
Fig. 6 Datasets used with BERT
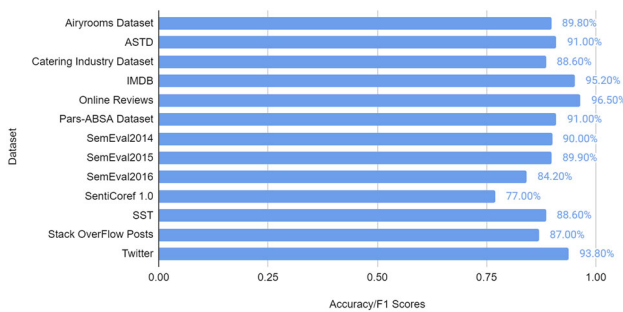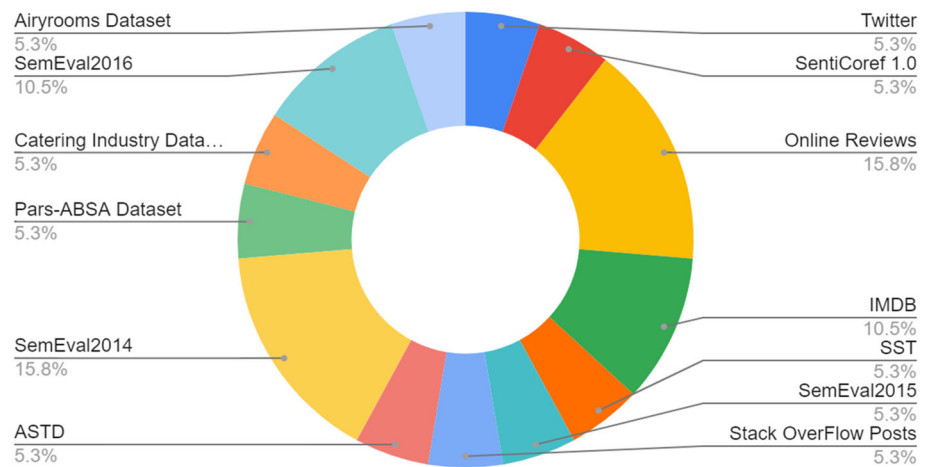
Fig. 7 Accuracy/F1 scores for BERT vs. datasets

by analyzing audience reactions to media material, while sentiment research helps the real estate and car industries understand consumer mood. Sentiment analysis impacts decisions across many domains, from environmental concerns and fashion trends to technology adoption and parenting guidance. However, rigorous domain-specific adaptation and ethical data handling are crucial for its responsible usage.

The applications of sentiment analysis reach far and wide, influencing the world in diverse ways. Some applications have the potential to make a positive impact, such as those focused on crime prevention or disaster response. Others serve crucial functions within specific industries, like credit risk assessment in banking or demand forecasting in manufacturing. Beyond these established applications, numerous unexplored areas are ripe for sentiment analysis exploration. These include:

- Finance and Risk Management: Optimizing markdown strategies in retail, identifying fraudulent claims and tax activities, and detecting credit card fraud and money laundering in the banking sector.
- Predictive Maintenance: Analyzing data to anticipate maintenance needs in the aerospace industry.

- Public Sector: Analyzing electronic health records to understand disease patterns.
- Video Surveillance: Extracting insights from video footage for security purposes.
- Vehicle-Based Services: Personalizing services based on customer sentiment.
- Travel and Tourism: Tailoring travel recommendations based on customer feedback.

Research in these areas has the potential to yield novel solutions and significant advancements [175, 176].

Sentiment analysis can be a powerful tool within the customer-business domain. Businesses can gain valuable insights into product performance and customer satisfaction by analyzing customer sentiment. Research has explored this concept, focusing on understanding customer sentiment and its impact on product success. However, a gap exists in the development of user-friendly applications. Ideally, graphical user interfaces (GUIs) could empower customers to make informed choices based on sentiment analysis data. Similarly, businesses lack real-time sentiment analysis tools to track the performance of popular products. By bridging these gaps, sentiment analysis can be leveraged to benefit both businesses and their customers.

6.4 Language scope

English is the most commonly used language in sentiment analysis research due to its relative ease of processing. However, even English, despite its perceived simplicity for humans, presents challenges for sentiment analysis models. Jargon, sarcasm, and idioms can be particularly difficult to interpret and categorize as positive, negative, or neutral. Figure 8 identifies and shows the popular languages used in the latest experiments that are also highlighted above as
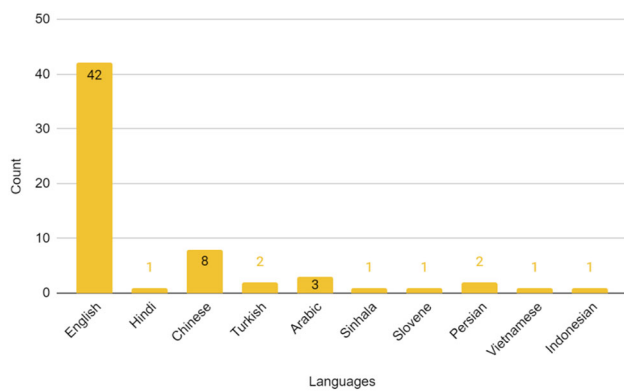
Fig. 8 Languages used in SA experiments

well. This dominance of English in research creates significant opportunities for exploring sentiment analysis in other languages.

6.5 Datasets

Datasets play a crucial role in training sentiment analysis models. The more data a system is trained on, the better it may understand and process future data. Figure 9 illustrates that "Online Reviews", "SemEval", "Twitter", and "IMDB" datasets are popular choices. Online review datasets can range from 1000 to over 1.6 m rows of data, while SemEval datasets typically contain 3000 to 15,000 + rows. It is important to note that all types of datasets, which are usually stored in editable CSV files, can be manipulated.

6.6 Challenges

Our research identified several persistent challenges in sentiment analysis, particularly regarding sarcasm and irony detection [62]. Some researchers propose that hybrid models incorporating the BERT transformer architecture may offer solutions to these limitations [5]. This area presents a promising avenue for further investigation. Other identified challenges include dataset inconsistency, emotion discovery, and negation handling.

From the above challenges, we can categorize key takeaways into four different challenge types: data and annotation challenges, modeling and computational challenges, social media and user-generated content challenges, and transparency and explainability challenges.

6.6.1 Data and annotation challenges

- Code-mixed data: Analyzing text sentiment that mixes multiple languages.
- Anaphora and Coreference Resolution: Understanding references back to previously mentioned entities.

- Spam Detection: Identifying and filtering out irrelevant or misleading content.
- Negation Handling: Accurately interpreting the sentiment when negation words are present.
- Word Sense Disambiguation: Determining the intended meaning of a word with multiple possible senses.
- Low-resource Languages: Limited training data for sentiment analysis in less common languages.
- Sarcasm Detection: Identifying text that conveys the opposite meaning through irony.
- Fake Reviews: Detecting fraudulent or misleading reviews that distort sentiment analysis results.
- Data Bias: Datasets can be skewed toward certain viewpoints or sentiment polarities.
- Domain Specificity: Models trained on one domain might not generalize well to others.
- Data Scarcity: Difficulty in acquiring and creating large, high-quality datasets for training models.
- Heterogeneous Big Data: Dealing with the complex and diverse nature of big data for sentiment analysis.
- Sparse, Uncertain, and Incomplete Data: Handling missing or unreliable information within data.

6.6.2 Modeling and computational challenges

- Accuracy Improvement: Continuously improving the accuracy of sentiment analysis, particularly for subjectivity detection.
- Context Dependency: Taking into account the surrounding context to understand sentiment accurately.
- Computational Cost: Balancing processing speed with model performance, especially for complex algorithms like CNNs and RNNs.
- Domain Adaptation: Adapting models trained on one domain to work effectively in another.
- Multilingual Applications: Developing sentiment analysis models that can handle multiple languages effectively.
- Linguistic Complications: Dealing with the complexities of human language, such as ambiguity and irony.
- Aspect-Based Sentiment Analysis (ABSA): Identifying sentiment toward specific aspects of a product or service within the text.
- Hidden Emotion Detection: Recognizing subtle emotions in short text snippets.
- Unexplored Issues in Multimodal Sentiment Analysis (MSA): Exploring sentiment analysis in non-text formats like audio or video.

Fig. 9 Popular datasets used in experiments
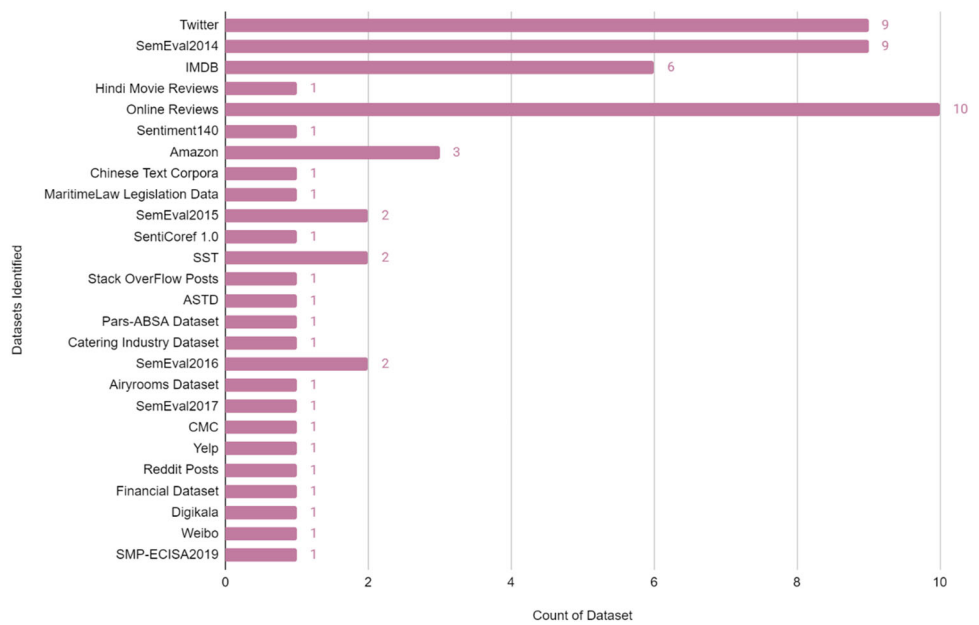
6.6.3 Social media and user-generated content challenges

- Information Overload: Filtering and analyzing vast amounts of social media data efficiently.
- Informal Language: Dealing with informal writing styles, slang, emoticons, and grammatical errors common in social media posts.
- Evolving Language: Keeping pace with the constantly changing nature of online language.
- Comparison Review Analysis: Understanding how users compare different products or services in their reviews.

6.6.4 Transparency and explainability challenges

- Model Transparency: Understanding how sentiment analysis models reach their conclusions.
- Data Bias: Mitigating biases present in the data used to train sentiment analysis models.

Furthermore, the above discussion has assisted me with answering my research questions:

RQ1: While general sentiment analysis and aspect-based sentiment analysis remain well-researched areas, several promising avenues exist for further exploration. Fine-grained sentiment analysis, which delves into more nuanced classifications, and irony/sarcasm detection are gaining traction. Additionally, aspect extraction, emotion analysis, multilingual sentiment comprehension, temporal sentiment tracking, and user-level sentiment analysis are emerging areas of interest. As technology evolves, exploring these nuanced tasks can significantly enhance our understanding of user opinions in diverse contexts.

RQ2: Various deep learning methods are employed in sentiment analysis, which focuses on identifying emotions from text. RNNs excel at processing sequences by remembering past information, with LSTM and GRU variants adept at capturing long-range dependencies. CNNs, originally designed for image processing, can also be applied to capture localized text patterns. Transformer models, which leverage attention mechanisms to understand global word relationships, have been a major breakthrough, offering impressive performance. While transformer-based models, particularly BERT variants, often shine due to their contextual awareness, the ultimate "best" algorithm depends on the specific dataset, available training data, and the unique problem being addressed. Experimentation is crucial for achieving optimal performance.

RQ3: Deep learning models offer significant value in real-world applications like customer feedback analysis. To leverage them effectively, here's a recommended approach:

- Data Preparation: Assemble a diverse dataset and preprocess the text for optimal model training.
- Model Selection: Choose a suitable deep learning architecture, potentially considering pre-trained models for their contextual understanding.
- Fine-tuning: Fine-tune the chosen model on your specific customer feedback data to tailor it to your domain.
- Aspect-Based Analysis: Consider using aspect-based sentiment analysis to dissect feedback sentiment toward specific aspects of your product or service.

- Real-Time Processing: Establish a real-time processing pipeline to continuously analyze new feedback data.
- Model Evaluation: Regularly evaluate the model's performance using appropriate metrics to identify areas for improvement.
- Challenge Mitigation: Address challenges like imbalanced data and customize the model for industry-specific language.
- User Interface: Develop a user-friendly interface for easy input and result visualization.
- CRM Integration: Integrate sentiment analysis with your CRM system for practical use of insights.
- Feedback Loop: Establish a feedback loop to continuously improve the model's performance.
- Ethical Considerations: Adhere to ethical considerations and data privacy regulations throughout the process.

By following these steps, deep learning models can be harnessed to power accurate sentiment analysis, empowering businesses with actionable insights from customer feedback.

RQ4: The performance of deep learning models in sentiment analysis can vary significantly when applied to different domain or language datasets. Language structure, cultural expressions, and contextual nuances all contribute to this disparity. Models trained on data from one domain may struggle when confronted with data from another domain due to variations in attitudes and language patterns. Similarly, model performance varies across languages due to distinct syntactical elements, idiomatic expressions, and cultural references. The availability and quality of training data also change across domains and languages, further impacting model efficacy. While pre-trained models offer a valuable starting point, their effectiveness can be influenced by domain and language shifts, often requiring further fine-tuning. Additionally, issues like code-switching, multilingualism, and resource constraints need to be addressed. Successful solutions, such as cross-lingual transfer learning and domain adaptation techniques, can alleviate these difficulties, resulting in improved sentiment analysis results across the board.

RQ5: Several established datasets serve as essential benchmarks in the field of sentiment analysis using deep learning techniques:

- IMDb Movie Reviews: This dataset offers a large collection of movie reviews labeled with positive or negative sentiment, making it a valuable resource for general sentiment analysis tasks.
- Amazon Product Reviews: This dataset contains reviews related to various products, facilitating evaluations within the product context.
- Twitter Sentiment Analysis Dataset: This dataset is specifically designed for social media sentiment analysis, encompassing brief and informal tweet expressions.

- SemEval Sentiment Analysis Datasets: These datasets encompass multilingual and multi-domain problems, reflecting the complexities of real-world scenarios.
- Online Reviews Sentiment Datasets: These datasets focus on sentiments regarding specific products and services.

Collectively, these datasets enable the training, assessment, and comparison of deep learning models in sentiment analysis. However, choosing datasets that are specifically tailored to the task at hand remains crucial for optimal performance.

RQ6: Sentiment analysis faces several challenges due to the inherent complexities of human language. Sarcasm and irony pose significant difficulties, as models often struggle to recognize these linguistic strategies where the intended meaning differs from the literal wording. Capturing the nuances of these expressions requires a deeper understanding of context, cultural references, and tone. Additionally, mitigating the influence of data imbalances, domain shifts, and multilingual complexities is crucial for improving sentiment analysis. Developing ethical and unbiased models, along with increasing contextual memory and addressing privacy concerns, are all critical areas that require further attention. By tackling these challenges, future research can enable models to effectively capture the vast spectrum of sentiments expressed in human language, leading to significant advancements in sentiment analysis capabilities.

7 Future works

This research has identified several promising avenues for further exploration in sentiment analysis. These future works can be broadly categorized into challenges related to tasks, data, and models.

7.1 Challenges related to tasks

The field of sentiment analysis can benefit from exploring new and emerging sentiment tasks that go beyond basic sentiment classification. Fine-grained sentiment analysis offers a more nuanced understanding by delving into detailed classifications beyond positive, negative, or neutral. Similarly, irony and sarcasm detection can help models capture the intended meaning behind words, even when they contradict the literal sentiment. Extracting aspects and understanding emotions from text can provide deeper insights into user opinions. Expanding sentiment analysis to handle multiple languages and track sentiment changes over time broadens its applicability in a globalized world.

Develop and evaluate hybrid deep learning models that integrate different architectures (e.g., transformers, recurrent neural networks) to tackle intricate sentiment analysis

tasks like sarcasm and irony detection. These models could leverage the strengths of each approach to achieve superior performance and capture the nuances of human expression. Finally, comprehending sentiment at the user level can be crucial for understanding individual perspectives within a larger dataset.

7.2 Challenges related to data

Data quality and availability present several challenges in sentiment analysis. Imbalanced datasets, where one sentiment might be heavily overrepresented, can skew model performance. Models trained on data from one domain may not perform well when applied to a different domain due to domain shift issues. Additionally, the increasing prevalence of code-switching and multilingual data necessitates solutions for handling these complexities. Integrate sentiment analysis with XAI techniques to gain insights into how models arrive at their sentiment classifications. This transparency can be crucial for building trust and ensuring the responsible use of sentiment analysis in various domains.

7.3 Challenges related to models

Developing more robust deep learning architectures specifically designed for sentiment analysis tasks can improve overall accuracy and effectiveness. Furthermore, increasing the contextual memory of models is crucial for capturing the intricate relationships within the text and accurately interpreting sentiment. However, advancements in model development must be accompanied by a focus on ethical considerations and mitigating bias to ensure fair and responsible sentiment analysis. Finally, addressing privacy and security concerns associated with user data is paramount for building trust and ensuring the responsible use of sentiment analysis in real-world applications.

Investigate and develop cutting-edge deep learning architectures specifically tailored for sentiment analysis tasks. This could involve advancements in areas like transformers and recurrent neural networks to achieve higher accuracy and efficiency.

7.4 Generative models for sentiment analysis

This research primarily focused on deep learning architectures for sentiment analysis. However, a recent and exciting development in the field involves the exploration of generative models like ChatGPT for sentiment analysis tasks. Generative pre-trained transformers (GPTs) like ChatGPT are capable of generating realistic and coherent text. Their potential application in sentiment analysis lies in two key areas:

Data Augmentation: One of the major challenges in sentiment analysis is the availability of large, high-quality labeled datasets. Generative models can create synthetic data with specific sentiment labels [177]. This artificially generated data can then augment existing datasets, improving the training process and generalizability of sentiment analysis models.

Sentiment Simulation: Generative models can simulate human sentiment expression in various contexts. This can be valuable for tasks like sarcasm detection, where models need to understand the subtle nuances of language that convey a sentiment different from the literal meaning [178]. By simulating different types of sarcasm and irony, researchers can develop models that are better equipped to identify these complex linguistic features.

While generative models offer promising potential for sentiment analysis, there are also challenges to consider:

Quality Control of Generated Data: The quality of synthetic data generated by GPTs is crucial for its effectiveness in sentiment analysis. Techniques to ensure the accuracy and consistency of generated sentiment labels are necessary [179].

Bias and Fairness: Generative models themselves can inherit biases present in the data they are trained on. It is essential to develop methods for mitigating bias in generated data to ensure fair and responsible sentiment analysis applications [180].

By tackling these future research directions, we can significantly enhance the capabilities of sentiment analysis and unlock its full potential for various real-world applications.

8 Conclusions

In conclusion, sentiment analysis has emerged as a transformative force with its applications revolutionizing diverse fields like education, brand management, finance, and disaster response. This analytical tool empowers businesses and organizations to shape strategies, improve customer relations, and make data-driven decisions with greater confidence.

Sentiment analysis has transcended basic sentiment classification, delving deeper into the intricacies of human expression. Tasks like fine-grained sentiment analysis and aspect-based sentiment assessment unveil nuanced insights, enabling a more comprehensive understanding of opinions and emotions.

Successfully navigating the landscape of sentiment analysis requires a strategic interplay between datasets, techniques, and language considerations. While established datasets like IMDb Movie Reviews and Twitter Sentiment Dataset provide valuable resources, the choice of data should be meticulously aligned with the specific analysis goals.

Deep learning techniques, particularly the versatile BERT model, offer a powerful combination of contextual understanding and transfer learning capabilities. However, the optimal algorithm for a given task depends on factors such as computational resources, domain relevance, and the intricacies of the target language. BERT on its own is a powerful algorithm but paired with the likes of RNN, LSTM, and GNN in hybrid ensures better accuracy for sentiment analysis.

Despite its advancements, sentiment analysis still grapples with challenges. The complex puzzle of sarcasm and irony detection necessitates the development of solutions that can decipher the subtle nuances of human communication. Additionally, ethical considerations surrounding privacy and bias remain paramount, especially when dealing with sensitive data in healthcare or criminal prediction. Furthermore, the exploration of novel sentiment tasks beyond conventional classifications presents exciting opportunities for researchers to unearth new insights and methodologies.

As sentiment analysis continues to evolve, its journey reflects a relentless pursuit of accuracy, relevance, and ethical integrity. This dynamic field indicates researchers and practitioners to delve deeper, harnessing the power of cutting-edge algorithms, expansive datasets, and cross-disciplinary collaboration. In essence, sentiment analysis is more than just an analytical tool; it is a transformative force driving innovation, fostering deeper understanding, and empowering informed decision-making across industries and languages.

Data availability Data sharing does not apply to this article as no datasets were generated or analyzed during the current study.

Declarations

Conflict of interest The authors declare that they have no competing interests.

Research involving human and animal participants This research did not involve any human or animal-based data.

References

1. Zhang, W., Xu, M., Jiang, Q.: Opinion mining and sentiment analysis in social media: Challenges and applications. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 536–548. Springer, Berlin (2018)

2. Pang, B., Lee, L.: Opinion mining and sentiment analysis, Foundations and Trends® in Information Retrieval: 2(1–2), 1-135. https://doi.org/10.1561/1500000011

3. Taboada, M., Brooke, J., Tofiloski, M., Voll, K., Stede, M.: Lexicon-Based Methods for Sentiment Analysis (2011). http://direct.mit.edu/coli/article-pdf/37/2/267/1798865/coli_a_00049.pdf

4. Salur, M.U., Aydin, I.: A novel hybrid deep learning model for sentiment classification. IEEE Access **8**, 58080–58093 (2020). https://doi.org/10.1109/ACCESS.2020.2982538

5. Devlin, J., Chang, M.W., Lee, K., Toutanova, K.: BERT: Pre-training of deep bidirectional transformers for language understanding (2018). https://doi.org/10.48550/arXiv.1810.04805

6. Daniel Ruby.: Twitter Statistics: New Trends, Figures & Data. Demand Sage (2022).

7. Simon Kemp.: Facebook Statistics and Trends, Datareportal

8. Tul, Q., et al.: Sentiment analysis using deep learning techniques: a review. Int. J. Adv. Comput. Sci. Appl.Comput. Sci. Appl. (2017). https://doi.org/10.14569/ijacsa.2017.080657

9. Zhang, L., Wang, S., Liu, B.: Deep learning for sentiment analysis: a survey. WIREs Data Min. Knowl. Discov. (2018). https://doi.org/10.1002/widm.1253

10. Ravi, K., Ravi, V.: A survey on opinion mining and sentiment analysis: Tasks, approaches and applications. Knowl. Based Syst. **89**, 14–46 (2015). https://doi.org/10.1016/j.knosys.2015.06.015

11. Liu, B., Zhang, L.: A survey of opinion mining and sentiment analysis. In: Mining Text Data. Springer US, Boston, pp. 415–463 (2012). https://doi.org/10.1007/978-1-4614-3223-4_13

12. Islam, Md.S., et al.: Challenges and future in deep learning for sentiment analysis: a comprehensive review and a proposed novel hybrid approach. Artif. Intell. Rev. Intell. Rev. **57**(3), 62 (2024). https://doi.org/10.1007/s10462-023-10651-9

13. Elsa, J., Koraye, J.: Deep learning techniques for natural language processing: recent developments (2024). https://easychair.org/publications/preprint_download/FPbH

14. Dang, N.C., Moreno-García, M.N., De la Prieta, F.: Sentiment analysis based on deep learning: a comparative study. Electronics (Basel) **9**(3), 483 (2020). https://doi.org/10.3390/electronics9030483

15. Prabha, M.I., Umarani Srikanth, G.: Survey of sentiment analysis using deep learning techniques. In: 2019 1st International Conference on Innovations in Information and Communication Technology (ICIICT), pp. 1–9 (2019). https://doi.org/10.1109/ICIICT1.2019.8741438

16. Yadav, A., Vishwakarma, D.K.: Sentiment analysis using deep learning architectures: a review. Artif. Intell. Rev.. Intell. Rev. **53**(6), 4335–4385 (2020). https://doi.org/10.1007/s10462-019-09794-5

17. Ain, Q.T., et al.: Sentiment analysis using deep learning techniques: a review. Int. J. Adv. Comput. Sci. Appl.Comput. Sci. Appl. (2017). https://doi.org/10.14569/IJACSA.2017.080657

18. Mäntylä, M.V., Graziotin, D., Kuutila, M.: The evolution of sentiment analysis—a review of research topics venues and top cited papers. Comput. Sci. Rev. **27**, 16–32 (2018). https://doi.org/10.1016/j.cosrev.2017.10.002

19. Dolianiti, F.S., Iakovakis, D., Dias, S.B., Hadjileontiadou, S., Diniz, J.A., Hadjileontiadis, L.: Sentiment analysis techniques and applications in education: a survey. In: Communications in Computer and Information Science, pp. 412–427. Springer, Berlin (2019). https://doi.org/10.1007/978-3-030-20954-4_31

20. Shah, P.V., Swaminarayan, P.R.: Sentiment analysis—an evaluation of the sentiment of the people: a survey. In: Kotecha, K., Piuri, V., Shah, H.N., Patel, R. (eds.) Data Science and Intelligent Applications, pp. 53–61. Springer, Singapore (2021)

21. Behdenna, S., Barigou, F., Belalem, G.: Document level sentiment analysis: a survey. EAI Endors. Trans. Context-Aware Syst. Appl. **4**(13), e2–e2 (2018)

22. Meena, A., Prabhakar, T.V.: Sentence level sentiment analysis in the presence of conjuncts using linguistic analysis. In: Advances in Information Retrieval: 29th European Conference on IR Research, ECIR 2007, Rome, Italy, April 2–5, 2007. Proceedings 29, pp. 573–580 (2007)

23. Mukherjee, A., Liu, B.: Aspect extraction through semi-supervised modeling. In: Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics, Long Papers, vol. 1, pp. 339–348 (2012)

24. Alm, C.O., Roth, D., Sproat, R.: Emotions from text: machine learning for text-based emotion prediction. In: HLT (2005)

25. Dragoni, M., Petrucci, G.: A neural word embeddings approach for multi-domain sentiment analysis. IEEE Trans. Affect. Comput.Comput. **8**(4), 457–470 (2017). https://doi.org/10.1109/TAFFC.2017.2717879

26. Yuan, Z., Wu, S., Wu, F., Liu, J., Huang, Y.: Domain attention model for multi-domain sentiment classification. Knowl. Based Syst. **155**, 1–10 (2018). https://doi.org/10.1016/j.knosys.2018.05.004

27. Boiy, E., Moens, M.-F.: A machine learning approach to sentiment analysis in multilingual Web texts. Inf. Retr. Boston **12**(5), 526–558 (2009). https://doi.org/10.1007/s10791-008-9070-z

28. Soleymani, M., Garcia, D., Jou, B., Schuller, B., Chang, S.F., Pantic, M.: A survey of multimodal sentiment analysis. Image Vis. Comput.Comput. **65**, 3–14 (2017). https://doi.org/10.1016/j.imavis.2017.08.003

29. Morency, L.-P., Mihalcea, R., Doshi, P.: Towards multimodal sentiment analysis: harvesting opinions from the web. In: Proceedings of the 13th International Conference on Multimodal Interfaces, in ICMI '11. New York, NY, USA: Association for Computing Machinery, pp. 169–176 (2011). https://doi.org/10.1145/2070481.2070509

30. Moussa, M.E., Mohamed, E.H., Haggag, M.H.: A survey on opinion summarization techniques for social media. Future Comput. Inform. J. **3**(1), 82–109 (2018). https://doi.org/10.1016/j.fcij.2017.12.002

31. Hu, M., Liu, B.: Mining and summarizing customer reviews. In: Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, in KDD' 04. Association for Computing Machinery, New York, NY, USA, pp. 168–177 (2004). https://doi.org/10.1145/1014052.1014073

32. Jindal, N., Liu, B.: Opinion spam and analysis. In: Proceedings of the 2008 International Conference on Web Search and Data Mining, in WSDM '08. Association for Computing Machinery, New York, NY, USA, pp. 219–230 (2008). https://doi.org/10.1145/1341531.1341560

33. Kim, S.-M., Hovy, E.: Determining the sentiment of opinions. In: Proceedings of the 20th International Conference on Computational Linguistics, in COLING '04. Association for Computational Linguistics, USA, p. 1367 (2004). https://doi.org/10.3115/1220355.1220555

34. Ortis, A., Farinella, G.M., Battiato, S.: Survey on visual sentiment analysis. IET Image Process. **14**(8), 1440–1456 (2020)

35. Habimana, O., Li, Y., Li, R., Gu, X., Yu, G.: Sentiment analysis using deep learning approaches: an overview. Sci. China Inf. Sci. (2020). https://doi.org/10.1007/s11432-018-9941-6

36. Jain, S., Gupta, V.: Sentiment analysis: a recent survey with applications and a proposed ensemble algorithm. In: Smart Innovation, Systems and Technologies, Springer Science and Business Media Deutschland GmbH, pp. 13–25 (2022). https://doi.org/10.1007/978-981-16-9447-9_2

37. Gottipati, S., Shankararaman, V., Gan, S.: A conceptual framework for analyzing students' feedback. In: 2017 IEEE Frontiers in Education Conference (FIE), pp. 1–8 (2017). https://doi.org/10.1109/FIE.2017.8190703

38. Dhanalakshmi, V., Bino, D., Saravanan, A.M.: Opinion mining from student feedback data using supervised learning algorithms. In: 2016 3rd MEC International Conference on Big Data and Smart City (ICBDSC), pp. 1–5 (2016). https://doi.org/10.1109/ICBDSC.2016.7460390

39. Koufakou, A., Gosselin, J., Guo, D.: Using data mining to extract knowledge from student evaluation comments in undergraduate courses. In: 2016 International Joint Conference on Neural Networks (IJCNN), pp. 3138–3142 (2016). https://doi.org/10.1109/IJCNN.2016.7727599

40. Altrabsheh, N., Gaber, M.M., Cocea, M.: SA-E: sentiment analysis for education. Frontiers in Artificial Intelligence and Applications, 255, pp. 353–362 (2013). https://doi.org/10.3233/978-1-61499-264-6-353

41. Colace, F., de Santo, M., Greco, L.: SAFE: a sentiment analysis framework for E-learning. Int. J. Emerg. Technol. Learn. (iJET) **9**(6), 37 (2014). https://doi.org/10.3991/ijet.v9i6.4110

42. Rani, S., Kumar, P.: A sentiment analysis system to improve teaching and learning. Computer (Long Beach Calif) **50**(5), 36–43 (2017). https://doi.org/10.1109/MC.2017.133

43. Scaffidi, C.: Mining online forums for valuable contributions. In: 2016 11th Iberian Conference on Information Systems and Technologies (CISTI), pp. 1–6 (2016). https://doi.org/10.1109/CISTI.2016.7521559

44. Cummins, S., Burd, L., Hatch, A.: Using feedback tags and sentiment analysis to generate sharable learning resources investigating automated sentiment analysis of feedback tags in a programming course. In: 2010 10th IEEE International Conference on Advanced Learning Technologies, pp. 653–657 (2010) https://doi.org/10.1109/ICALT.2010.186

45. Wen, M., Yang, D., Rosé, C.P.: Sentiment analysis in MOOC discussion forums: what does it tell us? In: EDM (2014)

46. Tucker, C., Pursel, B., Divinsky, A.: Mining student-generated textual data in MOOCS and quantifying their effects on student performance and learning outcomes. In: 2014 ASEE Annual Conference & Exposition Proceedings, ASEE Conferences, pp. 24.907.1–24.907.14. https://doi.org/10.18260/1-2--22840

47. Oliveiar, L., Figueira, A.: Visualization of sentiment spread on social networked content: Learning analytics for integrated learning environments in. IEEE Glob. Eng. Educ. Conf. (EDUCON) **2017**, 1290–1298 (2017). https://doi.org/10.1109/EDUCON.2017.7943014

48. Chaturvedi, S., Mishra, V., Mishra, N.: Sentiment analysis using machine learning for business intelligence. In: 2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI), pp. 2162–2166 (2017). https://doi.org/10.1109/ICPCSI.2017.8392100

49. Benedetto, F., Tedeschi, A.: Big data sentiment analysis for brand monitoring in social media streams by cloud computing. In: Pedrycz, W., Chen, S.-M. (eds.) Sentiment Analysis and Ontology Engineering: An Environment of Computational Intelligence, pp. 341–377. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-30319-2_14

50. Cheng, L.-C., Tsai, S.-L.: Deep learning for automated sentiment analysis of social media. In: Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, in ASONAM '19. Association for Computing Machinery, New York, NY, USA, pp. 1001–1004 (2019). https://doi.org/10.1145/3341161.3344821

51. Nemes, L., Kiss, A.: Social media sentiment analysis based on COVID-19. J. Inf. Telecommun. **5**(1), 1–15 (2021). https://doi.org/10.1080/24751839.2020.1790793

52. Bhat, M., Qadri, M., Beg, N.-A., Kundroo, M., Ahanger, N., Agarwal, B.: Sentiment analysis of social media response on the Covid19 outbreak. Brain Behav. Immun.Behav. Immun. **87**, 136–137 (2020). https://doi.org/10.1016/j.bbi.2020.05.006

53. Ashok, M., Rajanna, S., Joshi, P.V., Kamath, S.: A personalized recommender system using machine learning based sentiment analysis over social data. In: 2016 IEEE Students' Conference on Electrical, Electronics and Computer Science (SCEECS), pp. 1–6 (2016). https://doi.org/10.1109/SCEECS.2016.7509354

54. Sun, J., Wang, G., Cheng, X., Fu, Y.: Mining affective text to improve social media item recommendation. Inf. Process. Manag.Manag. **51**(4), 444–457 (2015). https://doi.org/10.1016/j.ipm.2014.09.002

55. Sousa, M.G., Sakiyama, K., de Souza Rodrigues, L., Moraes, P.H., Fernandes, E.R., Matsubara, E.T.: BERT for stock market sentiment analysis. In: 2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI), pp. 1597–1601 (2019). https://doi.org/10.1109/ICTAI.2019.00231

56. Ren, R., Wu, D.D., Liu, T.: Forecasting stock market movement direction using sentiment analysis and support vector machine. IEEE Syst. J. **13**(1), 760–770 (2019). https://doi.org/10.1109/JSYST.2018.2794462

57. Pagolu, V.S., Reddy, K.N., Panda, G., Majhi, B.: Sentiment analysis of Twitter data for predicting stock market movements. In: 2016 International Conference on Signal Processing, Communication, Power and Embedded System (SCOPES), pp. 1345–1350 (2016). https://doi.org/10.1109/SCOPES.2016.7955659

58. Khedr, A.E., Salama, S.E., Yaseen, N.: Predicting stock market behavior using data mining technique and news sentiment analysis. Int. J. Intell. Syst. Appl. **9**(7), 22–30 (2017). https://doi.org/10.5815/ijisa.2017.07.03

59. Dang, N.C., Moreno-García, M.N., de la Prieta, F.: Sentiment analysis based on deep learning: a comparative study. Electronics (Switzerland) **9**(3), 483 (2020). https://doi.org/10.3390/electronics9030483

60. Jain, A.P., Dandannavar, P.: Application of machine learning techniques to sentiment analysis. In: 2016 2nd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT), pp. 628–632 (2016). https://doi.org/10.1109/ICATCCT.2016.7912076

61. Chaturvedi, I., Cambria, E., Welsch, R.E., Herrera, F.: Distinguishing between facts and opinions for sentiment analysis: survey and challenges. Inf. Fus. **44**, 65–77 (2018). https://doi.org/10.1016/j.inffus.2017.12.006

62. Rambocas, M., Pacheco, B.G.: Online sentiment analysis in marketing research: a review. J. Res. Interact. Mark. **12**(2), 146–163 (2018). https://doi.org/10.1108/JRIM-05-2017-0030

63. Haselmayer, M., Jenny, M.: Sentiment analysis of political communication: combining a dictionary approach with crowdcoding. Qual. Quant. **51**(6), 2623–2646 (2017). https://doi.org/10.1007/s11135-016-0412-4

64. Kušen, E., Strembeck, M.: Politics, sentiments, and misinformation: an analysis of the Twitter discussion on the 2016 Austrian Presidential Elections. Online Soc. Netw. Media **5**, 37–50 (2018). https://doi.org/10.1016/j.osnem.2017.12.002

65. Kuamri, S., Babu, C.N.: Real time analysis of social media data to understand people emotions towards national parties. In: 2017 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT), pp. 1–6 (2017). https://doi.org/10.1109/ICCCNT.2017.8204059

66. Boukabous, M., Azizi, M.: Crime prediction using a hybrid sentiment analysis approach based on the bidirectional encoder representations from transformers. Indones. J. Electr. Eng. Comput. Sci. **25**(2), 1131 (2022). https://doi.org/10.11591/ijeecs.v25.i2.pp1131-1139

67. el Hannach, H., Benkhalifa, M.: WordNet based implicit aspect sentiment analysis for crime identification from Twitter. Int. J. Adv. Comput. Sci. Appl.Comput. Sci. Appl. (2018). https://doi.org/10.14569/IJACSA.2018.091222

68. Azeez, J., Aravindhar, D.J.: Hybrid approach to crime prediction using deep learning. In: 2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI), pp. 1701–1710. IEEE (2015). https://doi.org/10.1109/ICACCI.2015.7275858

69. Gerber, M.S.: Predicting crime using Twitter and kernel density estimation. Decis. Support. Syst. Support. Syst. **61**, 115–125 (2014). https://doi.org/10.1016/j.dss.2014.02.003

70. Sufi, F.K., Khalil, I.: Automated disaster monitoring from social media posts using AI-based location intelligence and sentiment analysis. IEEE Trans. Comput. Soc. Syst. (2022). https://doi.org/10.1109/TCSS.2022.3157142

71. Mendon, S., Dutta, P., Behl, A., Lessmann, S.: A hybrid approach of machine learning and lexicons to sentiment analysis: enhanced insights from Twitter data of natural disasters. Inf. Syst. Front. **23**(5), 1145–1168 (2021). https://doi.org/10.1007/s10796-021-10107-x

72. Maharani, W.: Sentiment analysis during jakarta flood for emergency responses and situational awareness in disaster management using BERT. In: 2020 8th International Conference on Information and Communication Technology (ICoICT), pp. 1–5 (2020). https://doi.org/10.1109/ICoICT49345.2020.9166407

73. Ragini, J.R., Anand, P.M.R., Bhaskar, V.: Big data analytics for disaster response and recovery through sentiment analysis. Int. J. Inf. Manag. **42**, 13–24 (2018). https://doi.org/10.1016/j.ijinfomgt.2018.05.004

74. Singh, P., Sawhney, R.S., Kahlon, K.S.: Sentiment analysis of demonetization of 500 & 1000 rupee banknotes by Indian government. ICT Express **4**(3), 124–129 (2018). https://doi.org/10.1016/j.icte.2017.03.001

75. Roy, K., Kohli, D., Kumar, R.K.S., Sahgal, R., Wen-Bin, Yu.: Sentiment analysis of Twitter data for demonetization in India ? A text mining approach. Issues Inf. Syst. **18**(4), 9–15 (2017). https://doi.org/10.48009/4_iis_2017_9-15

76. Zhang, L., Wang, S., Liu, B.: Deep learning for sentiment analysis: a survey. Wiley Interdiscip. Rev. Data Min. Knowl. Discov. (2018). https://doi.org/10.1002/widm.1253

77. Mumuni, A., Mumuni, F.: Automated data processing and feature engineering for deep learning and big data applications: a survey. J. Inf. Intell. (2024). https://doi.org/10.1016/j.jiixd.2024.01.002

78. O'Mahony, N., et al.: Deep learning vs. traditional computer vision. In: Advances in Computer Vision: Proceedings of the 2019 Computer Vision Conference (CVC), vol. 1, pp. 128–144 (2020)

79. Joseph, J., Vineetha, S., Sobhana, N.V.: A survey on deep learning based sentiment analysis. Mater. Today Proc. **58**, 456–460 (2022). https://doi.org/10.1016/j.matpr.2022.02.483

80. Lecun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. Proc. IEEE **86**(11), 2278–2324 (1998). https://doi.org/10.1109/5.726791

81. Patterson, J., Gibson, A.: Deep Learning: A Practitioner's Approach, 1st edn. O'Reilly Media Inc, New York (2017)

82. Collobert, R., Weston, J.: A unified architecture for natural language processing: deep neural networks with multitask learning. In: Proceedings of the 25th International Conference on Machine Learning, in ICML '08. New York, NY, USA: Association for Computing Machinery, pp. 160–167 (2008). https://doi.org/10.1145/1390156.1390177

83. Wang, J., Sun, J., Lin, H., Dong, H., Zhang, S.: Convolutional neural networks for expert recommendation in community question

answering. Sci. China Inf. Sci. **60**(11), 110102 (2017). https://doi.org/10.1007/s11432-016-9197-0

84. Collobert, R., Weston, J., Bottou, L., Karlen, M., Kavukcuoglu, K., Kuksa, P.: Natural language processing (Almost) from scratch. J. Mach. Learn. Res. **12**, 2493–2537 (2011)

85. Kalchbrenner, N., Grefenstette, E., Blunsom, P.: A convolutional neural network for modelling sentences (2014). arXiv preprint arXiv:1404.2188. https://doi.org/10.48550/arXiv.1404.2188

86. Kim, Y.: Convolutional neural networks for sentence classification. In: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), pp. 1746–1751. Association for Computational Linguistics, Stroudsburg, PA, USA (2014). https://doi.org/10.3115/v1/D14-1181

87. Johnson, R., Zhang, T.: Deep pyramid convolutional neural networks for text categorization. In: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Long Papers), vol. 1, pp. 562–570. Association for Computational Linguistics, Stroudsburg, PA, USA (2017). https://doi.org/10.18653/v1/P17-1052

88. Conneau, A., Schwenk, H., Barrault, L., Lecun, Y.: Very deep convolutional networks for text classification (2016). arXiv preprint arXiv:1606.01781. https://doi.org/10.48550/arXiv.1606.01781

89. Elman, J.L.: Finding structure in time. Cogn. Sci.. Sci. **14**(2), 179–211 (1990). https://doi.org/10.1016/0364-0213(90)90002-E

90. Ko, W.-J., Tseng, B.-H., Lee, H.-Y.: Recurrent Neural Network based language modeling with controllable external memory. In: 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 5705–5709 (2017). https://doi.org/10.1109/ICASSP.2017.7953249

91. Hochreiter, S.: The vanishing gradient problem during learning recurrent neural nets and problem solutions. Int. J. Uncertain. Fuzziness Knowl. Based Syst. **6**(2), 107–116 (1998). https://doi.org/10.1142/S0218488598000094

92. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Comput.Comput. **9**(8), 1735–1780 (1997). https://doi.org/10.1162/neco.1997.9.8.1735

93. Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., Bengio, Y. Learning phrase representations using RNN encoder-decoder for statistical machine translation (2014) arXiv preprint arXiv:1406.1078. https://doi.org/10.48550/arXiv.1406.1078

94. Sun, X., Li, X., Zhang, S., Wang, S., Wu, F., Li, J., Zhang, T. and Wang, G. Sentiment analysis through llm negotiations (2023) arXiv preprint arXiv:2311.01876. https://doi.org/10.48550/arXiv.2311.01876

95. Howard, J., Ruder, S.: Universal language model fine-tuning for text classification (2018). arXiv preprint arXiv:1801.06146. https://doi.org/10.48550/arXiv.1801.06146

96. Khosravi, A., Rahmati, Z., Vefghi, A.: Relational graph convolutional networks for sentiment analysis (2024). arXiv preprint arXiv:2404.13079. https://doi.org/10.48550/arXiv.2404.13079

97. Marra, G., Diligenti, M., Giannini, F.: Relational reasoning networks (2021). arXiv preprint arXiv:2106.00393. https://doi.org/10.48550/arXiv.2106.00393

98. Do, H.H., Prasad, P.W.C., Maag, A., Alsadoon, A.: Deep learning for aspect-based sentiment analysis: a comparative review. Expert Syst. Appl. **118**, 272–299 (2019). https://doi.org/10.1016/j.eswa.2018.10.003

99. Lee, G., Jeong, J., Seo, S., Kim, C., Kang, P.: Sentiment classification with word localization based on weakly supervised learning with a convolutional neural network. Knowl. Based Syst. **152**, 70–82 (2018). https://doi.org/10.1016/j.knosys.2018.04.006

100. Dos Santos, C., Gatti, M.: Deep convolutional neural networks for sentiment analysis of short texts. In: COLING (2014)

101. Hassan, A., Mahmood, A.: Convolutional recurrent deep learning model for sentence classification. IEEE Access **6**, 13949–13957 (2018). https://doi.org/10.1109/ACCESS.2018.2814818

102. Singhal, P. and Bhattacharyya, P., 2016. Sentiment analysis and deep learning: a survey. Center for Indian Language Technology, Indian Institute of Technology, Bombay. https://www.cfilt.iitb.ac.in/~cfiltnew/resources/surveys/sentiment-deeplearning-2016-prerna.pdf

103. Peng, H., Ma, Y., Li, Y., Cambria, E.: Learning multi-grained aspect target sequence for Chinese sentiment analysis. Knowl. Based Syst. **148**, 167–176 (2018). https://doi.org/10.1016/j.knosys.2018.02.034

104. Rao, G., Huang, W., Feng, Z., Cong, Q.: LSTM with sentence representations for document-level sentiment classification. Neurocomputing **308**, 49–57 (2018). https://doi.org/10.1016/j.neucom.2018.04.045

105. Rana, R.: Gated Recurrent Unit (GRU) for emotion classification from noisy speech (2016). arXiv preprint arXiv:1612.07778. https://doi.org/10.48550/arXiv.1612.07778

106. Verma, S., Saini, M., Sharan, A.: Deep sequential model for review rating prediction. In: 2017 Tenth International Conference on Contemporary Computing (IC3), pp. 1–6 (2017). https://doi.org/10.1109/IC3.2017.8284318

107. Acheampong, F.A., Nunoo-Mensah, H., Chen, W.: Transformer models for text-based emotion detection: a review of BERT-based approaches. Artif. Intell. Rev. Intell. Rev. **54**(8), 5789–5829 (2021). https://doi.org/10.1007/s10462-021-09958-2

108. Shojaee-Mend, H., Mohebbati, R., Amiri, M., Atarodi, A.: Evaluating the strengths and weaknesses of large language models in answering neurophysiology questions. Sci. Rep. **14**(1), 1–10 (2024)

109. Lappin, S.: Assessing the strengths and weaknesses of large language models. J. Logic Lang. Inf. **33**(1), 9–20 (2024). https://doi.org/10.1007/s10849-023-09409-x

110. Chen, T., Qiu, D., Wu, Y., Khan, A., Ke, X., Gao, Y.: View-based explanations for graph neural networks. Proc. ACM Manag. Data (2024). https://doi.org/10.1145/3639295

111. Liu, X., Zhang, L., Guan, H.: Uplifting message passing neural network with graph original information (2022). arXiv preprint arXiv:2210.05382. https://doi.org/10.48550/arXiv.2210.05382

112. Jain, P.K., Saravanan, V., Pamula, R.: A hybrid CNN-LSTM: a deep learning approach for consumer sentiment analysis using qualitative user-generated contents. ACM Trans. Asian and Low-Resource Lang. Inf. Process. **20**(5), 1 (2021). https://doi.org/10.1145/3457206

113. Brauwers, G., Frasincar, F.: A survey on aspect-based sentiment classification. ACM Comput. Surv.Comput. Surv. (2021). https://doi.org/10.1145/3503044

114. Alshuwaier, F., Areshey, A., Poon, J.: Applications and enhancement of document-based sentiment analysis in deep learning methods: systematic literature review. Intell. Syst. Appl. (2022). https://doi.org/10.1016/j.iswa.2022.200090

115. Rani, S., Kumar, P.: Deep learning based sentiment analysis using convolution neural network. Arab. J. Sci. Eng. **44**(4), 3305–3314 (2019). https://doi.org/10.1007/s13369-018-3500-z

116. Tyagi, V., Kumar, A., Das, S.: Sentiment analysis on twitter data using deep learning approach. In: Proceedings—IEEE 2020 2nd International Conference on Advances in Computing, Communication Control and Networking, ICACCCN 2020, Institute of Electrical and Electronics Engineers Inc., pp. 187–190 (2020). https://doi.org/10.1109/ICACCCN51052.2020.9362853

117. Yang, L., Li, Y., Wang, J., Sherratt, R.S.: Sentiment analysis for E-commerce product reviews in Chinese based on sentiment lexicon and deep learning. IEEE Access **8**, 23522–23530 (2020). https://doi.org/10.1109/ACCESS.2020.2969854

118. Obiedat, R., Al-Darras, D., Alzaghoul, E., Harfoushi, O.: Arabic aspect-based sentiment analysis: a systematic literature review. In: IEEE Access, Institute of Electrical and Electronics Engineers Inc., vol. 9, pp. 152628–152645 (2021). https://doi.org/10.1109/ACCESS.2021.3127140

119. Aydln, C.R., Gungor, T.: Combination of recursive and recurrent neural networks for aspect-based sentiment analysis using inter-aspect relations. IEEE Access **8**, 77820–77832 (2020). https://doi.org/10.1109/ACCESS.2020.2990306

120. Gothane, S., et al.: Sentiment analysis in social media using deep learning techniques. IJISAE (2024)

121. Zyout, I., Zyout, M.: Sentiment analysis of student feedback using attention-based RNN and transformer embedding. Int. J. Artif. Intell. **13**(2), 2173–2184 (2024)

122. Durga, P., Godavarthi, D.: Deep-sentiment: an effective deep sentiment analysis using a decision-based recurrent neural network (D-RNN). IEEE Access **11**, 108433–108447 (2023). https://doi.org/10.1109/ACCESS.2023.3320738

123. Demotte, P., Senevirathne, L., Karunanayake, B., Munasinghe, U., Ranathunga, S.: Sentiment analysis of sinhala news comments using sentence-state LSTM networks. In: 2020 Moratuwa Engineering Research Conference (MERCon), pp. 283–288 (2020). https://doi.org/10.1109/MERCon50084.2020.9185327

124. Wang, B., Guo, P., Wang, X., He, Y., Wang, W.: Transparent aspect-level sentiment analysis based on dependency syntax analysis and its application on COVID-19. J. Data Inf. Qual. **14**(2), 1 (2022). https://doi.org/10.1145/3460002

125. Sachin, S., Tripathi, A., Mahajan, N., Aggarwal, S., Nagrath, P.: Sentiment analysis using gated recurrent neural networks. SN Comput. Sci. (2020). https://doi.org/10.1007/s42979-020-0076-y

126. Zhang, B., Li, X., Xu, X., Leung, K.C., Chen, Z., Ye, Y.: Knowledge guided Capsule attention network for aspect-based sentiment analysis. IEEE/ACM Trans. Audio Speech Lang. Process. (2020). https://doi.org/10.1109/TASLP.2020.3017093

127. Abimbola, B., De La Cal Marin, E., Tan, Q.: Enhancing legal sentiment analysis: a convolutional neural network-long short-term memory document-level model. Mach. Learn Knowl. Extr. **6**(2), 877–897 (2024). https://doi.org/10.3390/make6020041

128. Huang, B., et al.: Aspect-level sentiment analysis with aspect-specific context position information. Knowl. Based Syst. (2022). https://doi.org/10.1016/j.knosys.2022.108473

129. Sudhir, P., Suresh, V.D.: Comparative study of various approaches, applications and classifiers for sentiment analysis. Glob. Transit. Proc. **2**(2), 205–211 (2021). https://doi.org/10.1016/j.gltp.2021.08.004

130. Rida-E-Fatima, S., et al.: A multi-layer dual attention deep learning model with refined word embeddings for aspect-based sentiment analysis. IEEE Access **7**, 114795–114807 (2019). https://doi.org/10.1109/ACCESS.2019.2927281

131. Zhao, N., Gao, H., Wen, X., Li, H.: Combination of convolutional neural network and gated recurrent unit for aspect-based sentiment analysis. IEEE Access **9**, 15561–15569 (2021). https://doi.org/10.1109/ACCESS.2021.3052937

132. Loh, N.K.N., Lee, C.P., Ong, T.S., Lim, K.M.: MPNet-GRUs: sentiment analysis with masked and permuted pre-training for language understanding and gated recurrent units. IEEE Access (2024). https://doi.org/10.1109/ACCESS.2024.3394930

133. Zhang, B., Zhou, W.: Transformer-encoder-GRU (T-E-GRU) for Chinese sentiment analysis on Chinese comment text. Neural. Process. Lett. **55**, 1847 (2021)

134. Abdullah, T., Ahmet, A.: Deep learning in sentiment analysis: a survey of recent architectures. ACM Comput. Surv. Comput. Surv. (2022). https://doi.org/10.1145/3548772

135. Obaidi, M., Nagel, L., Specht, A., Klünder, J.: Sentiment analysis tools in software engineering: a systematic mapping study. Inf. Softw. Technol.Softw. Technol. **151**, 107018 (2022). https://doi.org/10.1016/j.infsof.2022.107018

136. Žitnik, S., Blagus, N., Bajec, M.: Target-level sentiment analysis for news articles. Knowl. Based Syst. **249**, 108939 (2022). https://doi.org/10.1016/j.knosys.2022.108939

137. Xiao, Y., Li, C., Thürer, M., Liu, Y., Qu, T.: User preference mining based on fine-grained sentiment analysis. J. Retail. Consum. Serv.Consum. Serv. **68**, 103013 (2022). https://doi.org/10.1016/j.jretconser.2022.103013

138. Hartmann, J., Heitmann, M., Siebert, C., Schamp, C.: More than a feeling: accuracy and application of sentiment analysis. Int. J. Res. Mark. (2022). https://doi.org/10.1016/j.ijresmar.2022.05.005

139. Du, Y., Wang, Y., Wei, K., Jia, J.: The sentiment analysis and sentiment orientation prediction for hotel based on BERT-BiLSTM model. In: Lecture Notes in Electrical Engineering, pp. 498–505. Springer Science and Business Media Deutschland GmbH, Singapore (2022). https://doi.org/10.1007/978-981-16-9423-3_62

140. Hoang, M., Alija Bihorac, O., Rouces, J.: Aspect-based sentiment analysis using BERT. In Proceedings of the 22nd Nordic conference on computational linguistics (pp. 187-196). https://aclanthology.org/W19-6120

141. Biswas, E., Karabulut, M.E., Pollock, L., Vijay-Shanker, K.: Achieving reliable sentiment analysis in the software engineering domain using BERT. In: Proceedings—2020 IEEE International Conference on Software Maintenance and Evolution, ICSME 2020, Institute of Electrical and Electronics Engineers Inc., pp. 162–173 (2020). https://doi.org/10.1109/ICSME46990.2020.00025

142. Chouikhi, H., Chniter, H., Jarray, F.: Arabic sentiment analysis using BERT model. In: Communications in Computer and Information Science, pp. 621–632. Springer Science and Business Media Deutschland GmbH, Singapore (2021). https://doi.org/10.1007/978-3-030-88113-9_50

143. Acikalin, U.U., Bardak, B., Kutlu, M.: Turkish sentiment analysis using BERT. In: 2020 28th Signal Processing and Communications Applications Conference, SIU 2020—Proceedings, Institute of Electrical and Electronics Engineers Inc. (2020) https://doi.org/10.1109/SIU49456.2020.9302492

144. Wang, Y., Chen, Q., Wang, W.: Multi-task BERT for aspect-based sentiment analysis. In: Proceedings - 2021 IEEE International Conference on Smart Computing, SMARTCOMP 2021, Institute of Electrical and Electronics Engineers Inc., pp. 383–385 (2021). https://doi.org/10.1109/SMARTCOMP52413.2021.00077

145. Jafarian, H., Taghavi, A., Javaheri, A., Rawassizadeh, R.: Exploiting Bert To Improve Aspect-Based Sentiment Analysis Performance on Persian Language. https://github.com/hooshvare/parsbert

146. Xie, X., Qin, B., Wan, Z., Nie, W.: Text aspect-level sentiment analysis based on multi- task joint learning. In: Proceedings—2021 2nd International Symposium on Computer Engineering and Intelligent Communications, ISCEIC 2021, Institute of Electrical and Electronics Engineers Inc., pp. 127–131 (2021). https://doi.org/10.1109/ISCEIC53685.2021.00033

147. Tran, O.T., Bui, V.T.: A BERT-based hierarchical model for vietnamese aspect based sentiment analysis. In: Proceedings—2020 12th International Conference on Knowledge and Systems Engineering, KSE 2020, Institute of Electrical and Electronics Engineers Inc., pp. 269–274 (2020). https://doi.org/10.1109/KSE50997.2020.9287650

148. Zhang, H., Pan, F., Dong, J., Zhou, Y.: BERT-IAN Model for aspect-based sentiment analysis. In: Proceedings - 2020 International Conference on Communications, Information System and Computer Engineering, CISCE 2020, Institute of Electrical and Electronics Engineers Inc., pp. 250–254 (2020). https://doi.org/10.1109/CISCE50729.2020.00056

149. Wang, L., Yao, C., Li, X., Yu, X.: BERT-based implicit aspect extraction. In: Proceedings of 2021 IEEE 3rd International Conference on Civil Aviation Safety and Information Technology, ICCASIT 2021, Institute of Electrical and Electronics Engineers Inc., pp. 758–761 (2021). https://doi.org/10.1109/ICCASIT53235.2021.9633578

150. Azhar, A.N., Khodra, M.L.: Fine-tuning pretrained multilingual BERT model for indonesian aspect-based sentiment analysis. In: 2020 7th International Conference on Advance Informatics: Concepts, Theory and Applications (ICAICTA), pp. 1–6. IEEE (2020). https://doi.org/10.1109/ICAICTA49861.2020.9428882

151. dos Santos, B.N., Marcacini, R.M., Rezende, S.O.: Multi-domain aspect extraction using bidirectional encoder representations from transformers. IEEE Access 9, 91604–91613 (2021). https://doi.org/10.1109/ACCESS.2021.3089099

152. He, A., Abisado, M.: Text sentiment analysis of douban film short comments based on BERT-CNN-BiLSTM-Att model. IEEE Access 12, 45229–45237 (2024). https://doi.org/10.1109/ACCESS.2024.3381515

153. Miah, M.S.U., Kabir, M.M., Bin Sarwar, T., Safran, M., Alfarhood, S., Mridha, M.F.: A multimodal approach to cross-lingual sentiment analysis with ensemble of transformer and LLM. Sci. Rep. 14(1), 9603 (2024). https://doi.org/10.1038/s41598-024-60210-7

154. Xing, F.: Designing heterogeneous LLM agents for financial sentiment analysis (2024). arXiv preprint arXiv:2401.05799. https://doi.org/10.48550/arXiv.2401.05799

155. Zhang, B., Yang, H., Zhou, T., Babar, A., Liu, X.-Y.: Enhancing financial sentiment analysis via retrieval augmented large language models. In Proceedings of the Fourth ACM International Conference on AI in Finance (ICAIF '23). Association for Computing Machinery, New York, NY, USA, 349–356 (2023). https://doi.org/10.1145/3604237.3626866

156. Deng, X., Bashlovkina, V., Han, F., Baumgartner, S., Bendersky, M.: LLMs to the moon? Reddit market sentiment analysis with large language models. In: Companion Proceedings of the ACM Web Conference 2023, pp. 1014–1019. ACM, New York, NY, USA (2023). https://doi.org/10.1145/3543873.3587605

157. Krugmann, J.O., Hartmann, J.: Sentiment analysis in the age of generative AI. Cust. Needs Solut. Needs Solut. 11(1), 3 (2024). https://doi.org/10.1007/s40547-024-00143-4

158. Bhat, R.H.: Stock price trend prediction using emotion analysis of financial headlines with distilled LLM model (2024). https://mavmatrix.uta.edu/cse_theses/4/

159. Li, Y., Li, N.: Sentiment analysis of weibo comments based on graph neural network. IEEE Access 10, 23497–23510 (2022). https://doi.org/10.1109/ACCESS.2022.3154107

160. Yang, S., Xing, L., Li, Y., Chang, Z.: Implicit sentiment analysis based on graph attention neural network. Eng. Rep. (2022). https://doi.org/10.1002/eng2.12452

161. Zhao, X., et al.: RDGCN: reinforced dependency graph convolutional network for aspect-based sentiment analysis. In: Proceedings of the 17th ACM International Conference on Web Search and Data Mining, pp. 976–984. ACM, New York, NY, USA (2024). https://doi.org/10.1145/3616855.3635775

162. Jin, Y., Zhao, A.: Bert-based graph unlinked embedding for sentiment analysis. Complex Intell. Syst. 10(2), 2627–2638 (2024). https://doi.org/10.1007/s40747-023-01289-9

163. Yin, S., Zhong, G.: TextGT: a double-view graph transformer on text for aspect-based sentiment analysis. In: Proceedings of the AAAI Conference on Artificial Intelligence, pp. 19404–19412 (2024)

164. Chen, W., Zheng, X., Zhou, H., Li, Z.: Evaluation of logistics service quality: sentiment analysis of comment text based on multi-level graph neural network. Traitement du Signal 38(6), 1853–1860 (2021). https://doi.org/10.18280/ts.380630

165. Birjali, M., Kasri, M., Beni-Hssane, A.: A comprehensive survey on sentiment analysis: approaches, challenges and trends. Knowl. Based Syst. (2021). https://doi.org/10.1016/j.knosys.2021.107134

166. Jain, P.K., Pamula, R., Srivastava, G.: A systematic literature review on machine learning applications for consumer sentiment analysis using online reviews. Comput. Sci. Rev. 41, 100413 (2021). https://doi.org/10.1016/j.cosrev.2021.100413

167. Chakraborty, K., Bhattacharyya, S., Bag, R., Hassanien, A.A.: Sentiment analysis on a set of movie reviews using deep learning techniques. Soc. Netw. Anal. (2019). https://doi.org/10.1016/b978-0-12-815458-8.00007-4

168. Shayaa, S., et al.: Sentiment analysis of big data: methods, applications, and open challenges. IEEE Access 6, 37807–37827 (2018). https://doi.org/10.1109/ACCESS.2018.2851311

169. Beseiso, M., Elmousalami, H.: Subword attentive model for arabic sentiment analysis: a deep learning approach. ACM Trans. Asian Low Resource Lang. Inf. Process. 19(2), 1–7 (2020). https://doi.org/10.1145/3360016

170. Liu, H., Chatterjee, I., Zhou, M., Lu, X.S., Abusorrah, A.: Aspect-based sentiment analysis: a survey of deep learning methods. IEEE Trans. Comput. Soc. Syst. 7(6), 1358–1375 (2020). https://doi.org/10.1109/TCSS.2020.3033302

171. Zhu, P., Chen, Z., Zheng, H., Qian, T.: Aspect aware learning for aspect category sentiment analysis. ACM Trans. Knowl. Discov. DataKnowl. Discov. Data (2019). https://doi.org/10.1145/3350487

172. Liu, L., Chen, H., Sun, Y.: A multi-classification sentiment analysis model of chinese short text based on gated linear units and attention mechanism. ACM Trans. Asian Low-Resource Lang. Inf. Process. (2021). https://doi.org/10.1145/3464425

173. Agüero-Torales, M.M., Abreu Salas, J.I., López-Herrera, A.G.: Deep learning and multilingual sentiment analysis on social media data: an overview. Appl. Soft Comput.Comput. (2021). https://doi.org/10.1016/j.asoc.2021.107373

174. Li, Y., Jia, B., Guo, Y., Chen, X.: Mining user reviews for mobile app comparisons. Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. 1(3), 1–15 (2017). https://doi.org/10.1145/3130935

175. Sehgal, D., Agarwal, A.K.: Real-time sentiment analysis of big data applications using twitter data with Hadoop framework. In: Advances in Intelligent Systems and Computing, pp. 765–772. Springer Verlag, Singapore (2018). https://doi.org/10.1007/978-981-10-5699-4_72

176. Nasreen Taj, M.B., Girisha, G.S.: Insights of strength and weakness of evolving methodologies of sentiment analysis. Glob. Transit. Proc. 2(2), 157–162 (2021). https://doi.org/10.1016/j.gltp.2021.08.059

177. Karampatsis, R.-M., Babii, H., Robbes, R., Sutton, C., Janes, A.: Big code != big vocabulary: open-vocabulary models for source code (2020). https://doi.org/10.1145/3377811.3380342

178. She, J., Hu, Y., Shi, H., Wang, J., Shen, Q., Mei, T.: Dive into Ambiguity: latent distribution mining and pairwise uncertainty estimation for facial expression recognition (2021). https://doi.org/10.48550/arXiv.2104.00232

179. Li, X., et al.: OSLNet: deep small-sample classification with an orthogonal softmax layer. EEE Trans. Image Process. (2020). https://doi.org/10.1109/TIP.2020.2990277

180. Peterson, V., Rufiner, H.L., Spies, R.D.: Kullback-leibler penalized sparse discriminant analysis for event-related potential classification (2016). arXiv preprint arXiv:1608.06863. https://doi.org/10.48550/arXiv.1608.06863