TAF
Journal of Advances in Technology and Engineering Research
JATER
2023, **9**(2): 1-16

Check for updates

PRIMARY RESEARCH

Exploring the trends beyond sentiment analysis: Challenges and modern approaches in text mining

Aafia Gul Siddiqui[1], Sumbul Ghulamani[2*], Asadullah Shah[3], Kamran Khowaja[4]

[1,2,4] Department of Computer Science, SZABIST University, Hyderabad, Pakistan

[3] Kulliyah of Information and Communication Technology, International Islamic University Malaysia (IIUM), Kuala Lumpur, Malaysia

Keywords
Deep learning
Emotion detection
Machine learning
Text analytics
Unimodal sentiment analysis

Received: 8 February 2023
Accepted: 7 April 2023
Published: 27 June 2023

Abstract
In the ever-evolving landscape of information and communication, text mining plays a pivotal role in extracting valuable insights from vast textual data. While sentiment analysis has garnered substantial attention, this research delves into the broader spectrum of text mining, aiming to uncover emerging trends, challenges, and contemporary approaches that extend beyond traditional sentiment analysis. The study begins by scrutinizing the background of sentiment analysis in capturing the nuanced landscape of language, prompting an exploration into the types and classifications and compiling the available work of sentiment analysis based on text from the lexical approach to the deep learning approach. Researchers and practitioners grapple with multifaceted challenges that involve navigating the complexities of context, sarcasm, and ambiguity. This underscores the necessity for more advanced methodologies to effectively address the evolving intricacies of language. By synthesizing insights from the analysis of current trends and challenges, this research contributes to the ongoing dialogue in text mining, offering a comprehensive perspective beyond sentiment analysis. The findings of this study are anticipated to inform researchers, practitioners, and industry professionals in navigating the intricate landscape of text mining, fostering innovation and responsible deployment in an increasingly data-driven society.

I. INTRODUCTION

Sentiment analysis, a popular topic among social media is widely expanding as the big data is growing. Sentiment analysis, also referred to as opinion mining, is a Natural Language Processing (NLP) technique focused on discerning the emotional tone or sentiment conveyed within a given text. Sentiment analysis categorizes text into binary classes such as positive or negative, however, categorizing itself into multi-classes is more challenging as multiclass refers to classifications of emotions which can vary a lot [1]. Sentiment analysis has made significant progress since its inception as a natural language processing task nearly two decades ago [2]. The application of sentiment analysis comes to understanding public sentiment and opinion as it is increasingly vital for decision-makers. Public and customer sentiments serve as expressions of their opinions, encompassing both positive and negative emotions. Positive feelings like happiness, surprise, and love contrast with negative emotions such as sadness, anger, and disgust. Categorizing text emotions enhances sentiment analysis accuracy, contributing to a more refined opinion summary [1]. Sentiment analysis term was first found in [3] and opinion mining term was first found in [4].

In the current age of advanced technology, communication extends beyond verbal interactions. The contemporary generation leverages technology and social media to express feelings and ideas. Platforms like social media and the internet serve as outlets for diverse opinions and emo-

*Corresponding author: Sumbul Ghulamani
†email: Sumbul.ghulamani@hyd.szabist.edu.pk

tions, ranging from individual perspectives on specific topics to daily reflections. Given the global connectivity facilitated by the internet, the plethora of emotional expressions present online can be utilized for psychological and behavioral analysis, as well as for assessing public sentiment towards a product or topic [5]. Text-based emotion detection holds promise for various domains, including marketing, social behavioral analysis, public sentiment analysis, advertising, and human-computer interaction [6].

Text mining or text analytics is the process of extracting valuable insights, patterns, and information from large volumes of textual data. With the evolution in trends of text sentiment analysis, the challenges in data and its feature extraction are also increasing in complexity in the domain of NLP and deep neural architecture [1]. [7] describes sentiment analysis as "mini-NLP" because it was found discussing most of the contents of NLP. The new task now is to identify emotions from text which can be in different languages, grammar structures, and classifications [8, 9]. Beyond merely discerning the polarity of opinions, emotions can contribute to the overall weight of sentiment polarity, working in tandem to unveil the genuine interest of an individual or group [2]. However, sentiment analysis has now shifted from unimodality to multimodality with new data streams available such as images, audio, and video streams [2]. However, this research aims to provide a comprehensive review of compiling the trends available for text-based sentiment analysis.

II. BACKGROUND

Sentiment analysis is a process of using technology to determine the emotional tone or attitude expressed in a piece of text. It involves analyzing the words and context to understand whether the overall sentiment is positive, negative, or neutral [10]. Achieving this is frequently facilitated by employing algorithms and natural language processing techniques to extract meaning from written language, encompassing sources like social media posts, reviews, and other textual content. Sentiment analysis finds widespread application in areas such as social media monitoring, analysis of customer feedback, and market research.

Approaches of Sentiment Analysis

Unimodal sentiment analysis and multimodal sentiment analysis refer to different approaches in sentiment analysis based on the type of data or modalities used for analysis.

1) Unimodal Sentiment Analysis : Unimodal sentiment analysis involves analyzing sentiments using data from a single modality or source. This could be text-only data, such as reviews, comments, or social media posts [2].

2) Example : Analyzing the sentiment of customer reviews for a product based solely on the text content of those reviews.

Multimodal Sentiment Analysis : Multimodal sentiment analysis, on the other hand, involves analyzing sentiments using data from multiple modalities or sources. This can include a combination of text, images, audio, and video data [2, 11].

3) Example : Analyzing the sentiment of a video review that includes both spoken words and facial expressions. The analysis may consider the spoken words, facial expressions, and possibly the background music to determine the overall sentiment.

A. Data Types in Sentiment Analysis

Sentiment analysis can use various data types to analyze and determine the sentiment expressed in different forms of content. The most common data types include:

1) Text Data : Sentiment analysis is frequently applied to text data, such as reviews, social media posts, comments, and textual feedback [10].

2) Example : Analyzing product reviews on an e-commerce website to determine if customers express positive, negative, or neutral sentiments.

3) Image Data : Computer vision techniques can be employed to analyze sentiments in images, including facial expressions and visual cues.

4) Example : Assessing the emotional expressions of individuals in photos to understand the sentiment conveyed.

5) Audio Data : Sentiment analysis can be applied to analyze sentiments expressed through spoken words, voice tone, and other audio characteristics.

6) Example : Analyzing customer service calls to determine the sentiment of the callers based on their voice tone and language.

7) Video Data : Combining both visual and audio elements, sentiment analysis on video data can provide a more comprehensive understanding of sentiments.

8) Example : Analyzing sentiment in video reviews, considering both facial expressions and spoken words.

9) Social Media Data : Sentiment analysis is widely used on social media platforms to understand public opinion expressed through text, images, and videos.

10) Example : Analyzing Twitter posts to gauge the sentiment of users regarding a specific topic or event. In addition to it, another famous dataset of IMDb movie reviews is also widely used and analyzed in the recent years, as shown in figure 1.

ISSN: 2414-4592
DOI: 10.20474/jater-9.2.1

11) Survey and Feedback Forms : Sentiment analysis can be applied to responses in surveys, feedback forms, and questionnaires to understand the sentiments of participants.
12) Example : Analyzing open-ended survey responses to determine the overall sentiment of participants toward a product or service.
13) Sensor Data : In certain applications, sentiment analysis can be applied to data collected from sensors, such as sentiment in environmental monitoring or user sentiment in smart devices.
14) Example : Analyzing user sentiments expressed through interactions with a smart home system.
15) Chat and Conversational Data : Sentiment analysis can be used to understand the sentiment expressed in chat messages, emails, or other forms of digital conversations.
16) Example : Analyzing the sentiment of customer support chat interactions to assess customer satisfaction.
By leveraging these diverse data types, sentiment analysis can provide insights into the sentiments expressed across a wide range of contexts and modalities. The choice of data type depends on the specific application and the nature of the content being analyzed [12].

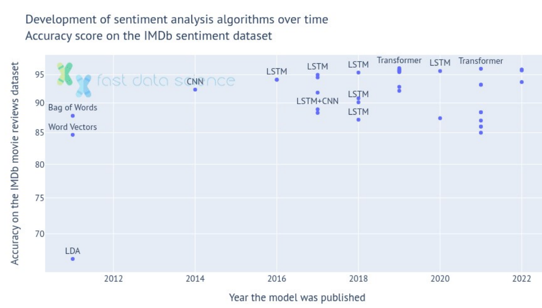

Fig. 1. IMDb movie reviews sentiment analysis [13]

B. Types of Sentiment Analysis

Sentiment analysis can be categorized into several types based on the scope and nature of the analysis. Here are some common types [10].
1) Document-Level Sentiment Analysis : Analyzing the sentiment of an entire document, such as an article, review, or blog post.
2) Use Case : Understanding the overall sentiment expressed in a piece of text.
3) Sentence-level sentiment analysis : Analyzing the sentiment of individual sentences within a document or text.
4) Use Case : Identifying the sentiment variations within a document.
5) Aspect-based sentiment analysis : Analyzing the sentiment of specific aspects or components within a piece of text.
6) Use Case : Understanding sentiments about different features or aspects mentioned in a product review.
7) Entity-level sentiment analysis : Focusing on the sentiment expressed towards specific entities, such as people, products, or companies.
8) Use case : Determining public sentiment about a particular brand or individual.

C. Multilingual sentiment analysis

Analyzing sentiment in text written in multiple languages.
1) Use case : Handling data from diverse linguistic sources.
2) Real-time sentiment analysis : Analyzing sentiments in real-time as data is generated.
3) Use case : Monitoring social media or news feeds for immediate reactions.
4) Fine-grained sentiment analysis : Providing more nuanced sentiment labels beyond just positive, negative, or neutral. For example, distinguishing between different levels of positivity or negativity.
5) Use case : Offering a more detailed understanding of sentiments expressed.
6) Emotion detection : Identifying specific emotions expressed in text, such as joy, anger, sadness, etc. In recent years, a trending interest has shown the work of sentiment analysis with emotion detection, as shown in Figure 1 from the website lens.org. this reveals that there is an considerable interest in the application of sentiment analysis in combination with emotion detection.
7) Use case : Understanding the emotional tone behind user comments or reviews.
8) Cross-domain sentiment analysis : Applying sentiment analysis models trained in one domain to analyze sentiment

ISSN: 2414-4592
DOI: 10.20474/jater-9.2.1

in a different domain. Use Case: Using sentiment models trained on product reviews to analyze sentiments in movie reviews. These types of sentiment analysis cater to various needs and scenarios, allowing businesses and researchers to gain insights into public opinion and attitudes in different contexts.

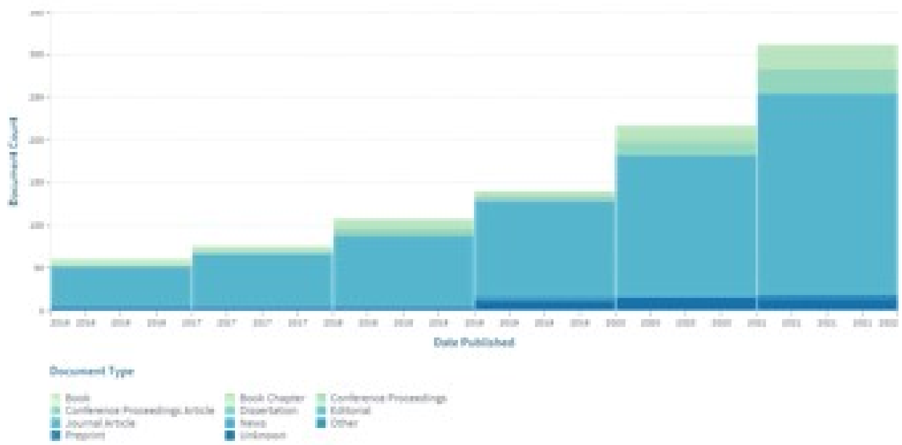

Fig. 2. Number of scholarly works by their publication in Sentiment Analysis and Emotion Detection [14]

D. *Text Sentiment Analysis*

Text Sentiment Analysis, a Natural Language Processing (NLP) technique, revolves around the identification and extraction of the sentiment or emotional tone embedded in a given text. Its primary objective is to ascertain whether the text conveys a sentiment that is positive, negative, or neutral. This analytical approach is widely employed in various applications to assess public opinion, analyze customer feedback, or discern the sentiment conveyed in textual content [13]. Here's a breakdown of key points related to Text Sentiment Analysis:

1) Sentiment categories : Positive Sentiment: Expresses favorable or positive opinions, emotions, or attitudes.

2) Negative sentiment : Conveys unfavorable or negative opinions, emotions, or attitudes.

3) Neutral sentiment : Indicates a lack of strong positive or negative emotions.

E. *Methods and Techniques*

1) Natural Language Processing (NLP) : Entails the utilization of algorithms and linguistic models to scrutinize and comprehend human language [14].

Machine Learning: Text sentiment analysis often employs machine learning models trained on labeled datasets to predict sentiment in unseen text.

F. *Common Applications*

1) Product reviews : Analyzing sentiments expressed in reviews to understand customer satisfaction.

2) Social media monitoring : Gauging public opinion on platforms like Twitter, Facebook, or Instagram.

3) Customer feedback analysis : Assessing sentiments in surveys, comments, and feedback forms.

4) Brand Monitoring : Evaluating sentiments associated with a brand in online discussions.

G. *Challenges*

1) Ambiguity : Some expressions may be context-dependent and challenging to interpret accurately.

2) Sarcasm and Irony : Textual nuances like sarcasm or irony can be difficult for models to identify.

3) Context understanding : Models may struggle with understanding the overall context in which certain sentiments are expressed.

H. *Techniques for Text Sentiment Analysis*

1) Rule-based approaches : Applying pre-established rules and patterns to detect sentiment.

2) Machine learning models : Training models on labeled datasets to acquire patterns and make predictions.

3) Deep learning models : Using deep neural networks for more complex sentiment analysis tasks.

A more detailed understanding of techniques of text senti-

ISSN: 2414-4592
DOI: 10.20474/jater-9.2.1

ment analysis is given in the figure 3.

I. Output

1) Sentiment scores : Quantitative measures indicating the intensity of positive or negative sentiment.
2) Sentiment labels : Categorizing text as positive, negative, or neutral [13].

Text Sentiment Analysis plays a pivotal role in assisting businesses in comprehending customer sentiment, overseeing brand reputation, and making informed, data-driven decisions grounded in public opinion. This tool is highly valuable in the realm of natural language processing and finds application across diverse industries.

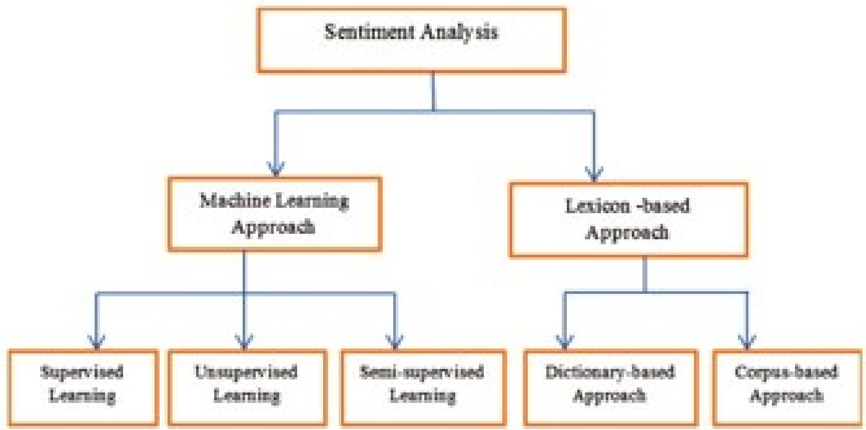

Fig. 2. Detailed techniques for sentiment analysis [15]

J. Classification of Text Sentiment Analysis

Text Sentiment analysis text is subdivided into three levels: document-level sentiment analysis, sentence-level sentiment analysis, and aspect-level sentiment classification.

K. Document-Level Sentiment Analysis

Document-level sentiment analysis involves treating the entire document as the focus of information, which focuses on a specific theme or topic and categorizes it into either positive or negative polarity. Pioneered by [15], on the dataset of IMDB movie reviews, it applied maximum entropy classifier, SVM, and Naive Bayes on the approaches of machine learning. Their experiments on numerous feature engineering revealed that the SVM classifier having the features of unigram achieved 82.9% of the highest accuracy. They suggested that accuracy could be further improved by incorporating coherence resolution, discourse analysis, and focus detection.
In a similar [2] conducted a classification on opinion polarity on a news dataset of Bengal with the application of SVM classifier, employed on the subjective portion of sentences. Extending beyond the classifiers of machine learning, lexicon entities and linguistic syntactic features were integrated to create a hybrid model. This system recognized the sentiment polarity (positive or negative) of opinionated phrases, achieving a precision of 70.04% and a recall of 63.02%.
[16] explored a framework for Manipuri language on document-level sentiment analysis, leveraging data extracted from letters to the editor published in local newspapers. The pre-processing of text involved part-of-speech tagging (POS), and verb identification was performed using a conditional random field (CRF) and modified verb lexicon. The modified verb lexicon was successful in predicting the sentiments in positive, negative, or neutral polarity, yielding a precision of 78.1%, a recall of 72.1%, and an F-measure of 75%.
Meanwhile, [17] introduced a sentiment analysis framework for Assamese news documents employing lexical features and machine learning algorithms. News data were manually collected from the local newspapers of Assamese. Random forest classifier, combined with verb lexical features, adverb, and adjective, achieved a highest accuracy of 67% among other classifiers. However, limitations of their model were identified, including an insufficient set of feature set and functional words ambiguity.

L. Multilingual Sentiment Analysis

Numerous approaches have been devised for sentiment analysis across diverse languages. A primary challenge in this field is the significant shortage of resources [18]. To tackle this issue, researchers frequently transfer knowledge

ISSN: 2414-4592
DOI: 10.20474/jater-9.2.1
TAF Publishing

from languages with ample resources to those with limited resources. Another approach for translating texts into English uses the machine translation systems [19]. However, there are other issues with translation systems, such as noise and sparseness in the data, and may not always capture essential parts of a text, leading to potential issues.

[20] investigated a lexicon and corpus-based approach for identifying subjectivity in multiple languages. They utilized [21] for English term lemmatization, translating it into Romanian terms. Applying a Naive Bayes classifier to the Romanian training dataset resulted in an F-score of 67.85. [22] introduced a framework for predicting a text polarity in a multilingual context, utilizing the [23] lexical resource. They utilized standard translation software to translate documents into various languages and then compared a statistical polarity classifier with an n-gram-based method for multilingual polarity detection. This evaluation was conducted on German movie reviews sourced from Amazon.

The multilingual nature of social media data implies that relying on a single official language for analysis may not adequately capture the overall sentiment expressed in online content [24]. To address this, the concept of multilingual sentiment analysis becomes crucial. Typical techniques involve automatically translating the target language into English and then applying methods and resources available in English. Alternatively, the use of parallel corpora is another approach. [25] developed a sentiment analysis system for tweets in English, which they then translated into four other languages using a machine translation system. Combining training datasets from languages with similar structures was observed to improve results compared to using individual languages in isolation. [26] concentrated on multilingual sentiment analysis without resorting to machine translation, instead, they analyzed emotion tokens or leveraged SentiLexicon.

[27] proposed a generative cross-lingual mixture model (CLMM) to leverage unlabeled bilingual parallel data. Experimental results indicated that multilingual sentiment analysis using a parallel corpus instead of machine translation can improve classification accuracy. [28] introduced an instance-level transfer learning scheme for cross-lingual sentiment analysis, translating important markup languages into the target language for additional training data.

[29] conducted a comprehensive study on multilingual sentiment analysis, reviewing various approaches and tools, identifying challenges, and offering recommendations, particularly for dealing with resource-poor languages. [24] discussed existing approaches for multilingual sentiment analysis, comparing them on the same datasets. [30] evaluated sentence-level sentiment polarity classification methods proposed for English and other specific languages. [31] reported a survey on multilingual sentiment analysis, encompassing different languages, resources, lexicons, corpora ontologies, and datasets in the Portuguese context.

M. Sentence-Level Sentiment Analysis

This confines the analysis of individual sentences. In their work, [32] introduced a lexical approach for mining and summarizing product reviews. They focused on creating feature-based summaries by first extracting the reviewed product features. Subsequently, they identified positive and negative opinion sentences and summarized the outcomes. Despite achieving promising results, the model exhibited limitations in handling pronoun resolution.

[33] opted for a machine learning approach employing unigram features for binary sentiment analysis on Twitter data. They calculated polarity and subjectivity using SentiWordNet, W-WSD, and TextBlob libraries. The Naive Bayes classifier, coupled with unigram features, attained highest accuracy of 79% in binary classification. Additionally, they also had Urdu Tweets into English translation for classification, with the help of SVM classifier finding N-gram features and was proved to be the most helpful model.

[34] proposed a Weakly-Supervised Deep Embedding (WDE) sentiment analysis framework on a review rating dataset. They employed a convolutional neural network to develop WDE-CNN and LSTM to create WDE-LSTM, extracting feature vectors from sentences in rating reviews. Assessing the system on an Amazon dataset covering three domains, namely, cell phones, digital cameras, and laptops —they obtained an accuracy of 87.9% for WDE-LSTM and 87.7% for WDE-CNN. The results of experiments revealed the models of deep learning, particularly when dealing with information-rich inputs, outperformed baseline models in achieving high accuracy.

1) Aspect-level sentiment analysis : Aspect-level sentiment analysis, also referred to as feature-based or entity-based sentiment analysis, involves the identification of aspects or features in a sentence and their categorization as negative or positive. Aspects can be either stated directly (explicit) or implied within the text (implicit), with direct aspects are easily found in the sentence, while indirect aspects are deduced from expressions of sentiment. The process of aspect-based sentiment analysis encompasses the identification, extraction, and grouping of aspects, aspects summarization, and sentiment classification [31].

[35] proposed a framework for implicit aspect extraction and polarity detection on a dataset of Chinese reviews with the help of explicit topic model. They extended the widely used Latent Dirichlet Allocation (LDA) topic modeling method to develop an explicit topic model. Prior to employing the topic modeling, the topics were assigned with the explicit aspects to various topics using LDA. Extraction was carried out for the explicit features through POS tagging, word segmentation, and feature clustering. Feature grouping based on the same domain features was carried out on phrases and words, and the chosen word clusters served as the attributes of training for classifiers, employing a support vector machine (SVM) classifier. Combination of SVM classifier and topic model algorithm achieved an F-measure of 77.78% for explicit and implicit classification features.

[36]loped a framework of an aspect-level sentiment detection called Sent_Comp, utilizing the techniques for sentence compression. The syntactic and extractive compression methods were applied to remove unnecessary sentiment information from the sentences. For the purpose of compressing automatic sentence information, they employed a model for discriminative conditional random field. While Sent_Comp being domain-independent, loss of information through the compression technique remains a significant challenge. The research identified a robust correlation between polarity and aspect words in aspect sentiment detection, ultimately influencing overall efficiency.

[37] introduced a framework for multitask learning for aspect term classification and identification within a unified model. A Bi-LSTM network was utilized, supported by a self-attention method, for aspect sentiment classification. The method involved identifying aspect terms using BiLSTM and self-attention, followed by sentiment prediction using a CNN framework. A frequent limitation noted in this research was the failure to prune aspects, leading to inaccuracies in aspect identification. [38] employed attention with LSTM mechanisms for aspect sentiment analysis, focusing on various aspects of the sentence. The attention mechanism focused on different segments of a sentence when various aspects were taken into account as input. Additionally, the model was trained in an end-to-end manner using backpropagation, incorporating the cross-entropy loss function.

N. Text Mining

Text mining, also recognized as text analytics, is a technique for extracting meaningful information and insights from large volumes of unstructured text data. Unorganized or unstructured text data refers to information that is not well defined and structured in a predefined manner, such as articles, emails, social media posts, reviews, and other forms of written content. The objective of text mining is to convert unstructured data into a structured format, facilitating analysis to reveal patterns, trends, and valuable knowledge. Here are key aspects of text mining [39]:

1) *Text data processing* : Text Extraction: Collecting and gathering text data from various sources.

2) *Tokenization* : Segmenting the text into individual units, such as words or phrases, commonly referred to as tokens.

3) *Stemming and lemmatizations* : Reducing words to their base or root form to simplify analysis.

4) *Information Retrieval* : Document Retrieval: Recognizing pertinent documents based on specific criteria.

5) *Keyword extraction* : Identifying and extracting important keywords or phrases from the text.

O. Text Analysis Techniques

1) *Text classifications* Categorizing documents or texts into predefined categories or topics.

2) *Named Entity Recognition (NER)* : Recognizing and categorizing entities, including names of individuals, organizations, locations, and more.

3) *Sentiment analysis* : Determining the emotional tone or sentiment expressed in the text.

4) *Topic modeling* : Discovering topics or themes present in a collection of documents.

5) *Clustering* : Grouping similar documents or texts based on their content.

Applications of Text Mining

6) *Customer feedback analysis* : Analyzing customer reviews, feedback, and comments to understand sentiment and preferences.

7) *Market research* : Extracting insights from textual data to identify market trends and consumer behavior.

8) *Healthcare and biomedical research* : Analyzing scientific literature for research trends and insights.

9) *Social media monitoring* : Tracking and analyzing conversations on social media platforms.

10) *Legal and compliance* : Analyzing legal documents for information extraction and compliance purposes.

P. Tools and technologies

1) *Natural Language Processing (NLP)* : Utilizing NLP techniques to understand and process human language.

2) *Machine learning* : Training models to recognize patterns and make predictions based on textual data.

ISSN: 2414-4592
DOI: 10.20474/jater-9.2.1

3) Text mining software : Using specialized software and tools designed for text mining tasks.

Text mining is an interdisciplinary field that integrates techniques from linguistics, computer science, and statistics to convert unstructured text data into actionable insights. It is widely used in various industries to unlock the value hidden in large volumes of textual information.

Q. How is Text Mining Related to Sentiment Analysis?

Text mining and sentiment analysis are closely related, with sentiment analysis being a specific application or task within the broader field of text mining [40]. Here's how they are connected:

1) Text Mining as a broad field : Text mining encompasses the entire process of extracting valuable information, patterns, and knowledge from unstructured text data.

2) Tasks : Text mining encompasses a range of tasks, including document retrieval, keyword extraction, text classification, named entity recognition, and sentiment analysis.

R. Sentiment Analysis as a Text Mining Task

Sentiment analysis is a distinct task within text mining that primarily focuses on determining the sentiment or emotional tone present in a text.

1) Objective : The principal objective of sentiment analysis is to categorize sentiment as positive, negative, or neutral, thereby offering insights into opinions and attitudes.

S. Role of Sentiment Analysis as a Text Mining Task

1) Insight extraction : Sentiment analysis is employed to extract insights related to the emotional tone or opinions expressed in textual data.

2) Decision support : Businesses use sentiment analysis results to make informed decisions, such as improving products based on customer feedback or managing brand reputation.

T. Integration of Sentiment Analysis Techniques

Text Classification: Sentiment analysis frequently entails the classification of text into sentiment categories (positive, negative, neutral), constituting a text classification task within the broader domain of text mining.

1) Feature extraction : Sentiment analysis utilizes features like words or phrases indicative of sentiment, a technique commonly used in various text mining tasks.

U. Applications of Sentiment Analysis in Text Mining

Customer Feedback Analysis: Understanding customer sentiments from reviews, comments, and feedback forms using sentiment analysis.

1) Social media monitoring : Analyzing sentiments expressed on social media platforms to gauge public opinion.

2) Market research : Extracting sentiment-related insights to understand consumer preferences and market trends.

In summary, sentiment analysis is a specific application within the broader context of text mining. While text mining involves a variety of tasks aimed at extracting valuable information from unstructured text data, sentiment analysis specifically focuses on uncovering the sentiments or opinions expressed in the text [41]. Sentiment analysis is a powerful tool within the realm of text mining, providing businesses and researchers with valuable insights into public sentiment and attitudes.

III. LITERATURE REVIEW

[1] has reviewed a detailed analysis of the recent shift from sentiment analysis using text to emotion detection, outlining the challenges in these works. They summarized key works from the last five years, examining the methods employed and exploring prevalent models of emotion classes. The evolution of text-based emotion detection has transitioned from initial keyword-based approaches to the application of sophisticated machine learning and deep learning algorithms. This transition has not only enhanced performance but also introduced greater task flexibility.

[2] has provided a comprehensive sentiment analysis from unimodality to multimodality, presenting new opportunities for enhanced sentiment detection beyond text-based analysis. Multimodal sentiment analysis, incorporating audio and video channels, broadens the scope and improves accuracy. Researchers are exploring diverse approaches, including the use of complex deep neural architectures, to enhance sentiment analysis system performance. Notably, transformer-based models have recently achieved significant success. The paper has conducted detailed research covering various sentiment analysis approaches, applications, and challenges ultimately emphasizing its tremendous potential. The survey's primary motivation is to underscore the shifting trends from unimodality to multimodality in addressing the tasks of sentiment analysis.

A Multi-Tier Social Media Sentiment Analysis was conducted by [42] in which he discussed a finer-grained classification of sentiment analysis which offers more nuanced insights into sentiments, though it poses greater challenges compared to binary classification. Conversely, multi-class classification leads to a notable decline in performance. The study investigates pre-processing methods and machine

ISSN: 2414-4592
DOI: 10.20474/jater-9.2.1

learning models employed for achieving multi-class sentiment classification, introducing a multi-layer classification model to enhance overall performance. Using a movie reviews dataset for implementation, which shares similarities with social media text, they employed supervised machine learning models such as SVM, Decision Tree, and Naive Bayes for sentiment classification. A comparison between single-layer and multi-tier models revealed slight improvements in the performance of the latter. Additionally, these multi-tier models exhibited improved recall, enabling the introduced model to grasp more contextual information.

Another study in which the dataset was taken from social media memes. Researched by [11] meme serves as a visual representation conveying a concept, gaining prominence in today's era of rapidly expanding social media platforms. For the areas of emotion and meme analysis, the detection of offensive content is a crucial yet challenging task due to the multimodal nature of meme content. The paper introduced a descriptive and balanced dataset to aid in offensive meme detection. Combining two models of deep semantic, hateXplain-BERT and baseline BERT, with numerous other deep Resnet architectures, gauging the intensity of offensive memes using the Meme-Merge collection. The outcomes demonstrated the effectiveness of model in categorizing offensive memes, attaining F1 scores of 0.7315 and 0.7140 for both Meme-Merge and baseline datasets.

[43] Discussed how mental health is impacted by social media through studying emotional sentiment analysis. The dataset was obtained from user posts through the Twitter API, and the Natural Language Understanding API tool was applied to extract and categorize emotions into five fundamental emotional categories: Sadness, Joy, Anger, Disgust, and Fear. Consequently, a method was proposed to filter emotionally detrimental social media content for users. Subsequently, an ideal emotion array based on a comprehensive list of about 450 English language words was defined. The primary objective of this extensive research article was to scrutinize the proposed solution aimed at enhancing the emotional quality of content presented to users.

[8] proposed a study “Sentiment Analysis is a Big Suitcase”. Her study acts as a collective container for our assorted ideas on the way in which our minds express emotions and opinions using natural language. To address the multi-faceted nature of this problem, the study proposed a three-layer structure with the inspiration from the 'jumping NLP curves' paradigm. They assert that, at the very least, 15 NLP problems must be addressed to achieve sentiment analysis performance akin to human capabilities.

Based on the ever-increasing scale of data on social media applications such as Facebook and Twitter. [44] studied text classification and created a hybrid model of this data to analyze insights into students' educational experiences, including their concerns, opinions, emotions, and feelings for the process of learning. Through qualitative analysis of approximately 25,000 tweets, they identified common problems such as diversity issues, a heavy study load, lack of social engagement, sleep deprivation, and negative emotions. Building on these findings, they implemented a hybrid model combining a Support Vector Machine and Naive Bayes classifier, to classify tweets related to students' challenges. The results indicated a reduction in training time and an improvement in classification accuracy compared to using individual Naive Bayes classifier and Support Vector Machine models.

Another study used twitter posts to analyze emotions. [6] wrote an article “Emotion and Sentiment Analysis from Twitter Text”. Online social networks serve as platforms for individuals to share views, thoughts, and perspectives using various media. Despite the diverse communication options, text remains prevalent. This paper focuses on detecting sentiment and emotion in Twitter posts, aiming to generate recommendations. A dataset was created from tweets and replies on specific topics, encompassing text, user, emotion, and sentiment information. Sentiment and emotion were analyzed, and user influence scores were determined. Novel features include incorporating replies, introducing scores for agreement, sentiment, and emotion in influence calculations, and generating recommendations based on users with similar sentiments on specific topics.

Das R. and Singh T conducted a study titled "Advancing Sentiment Analysis of Assamese News Articles Using Lexical Features" [17]. Their research focuses on employing machine learning classifiers and lexical features, such as adjectives, for sentiment analysis in Assamese news. The suggested model demonstrates superior performance over the baseline in F1-score on a standard dataset, thereby improving sentiment classification in a low-resource language.

IV. RECENT WORKS

Table 1 highlights the text sentiment analysis has gone through the recent works keeping in view the features, methods, and results.

ISSN: 2414-4592
DOI: 10.20474/jater-9.2.1

TABLE 1
RECENT WORKS IN TEXT SENTIMENT ANALYSIS

Authors	Emotions/Sentiments	Features	Method	Results
[45]	Fear, sad, anger, joy	N-grams, smiles, exclamation mark, question mark, curse words, greeting words, sentiment polarity	Keyword/Lexicon with rule-based	Avg. recall: 0.6 -0.7 Avg. precision: 0.6–0.7 Avg. F1: 0.6–0.7
[46]	love, joy sadness, anger, fear, surprised	Words	Keyword based with ontology	Avg. accuracy: 0.8
[47]	sadness, happiness, anger, fear, disgust, surprise	Words	Unsupervised learning – normalized PMI	Avg. accuracy: 0.68 Avg. recall: 0.72 Avg. precision: 0.92
Perikos and [48]	sadness, happiness, anger, fear, disgust, surprise	Words	Supervised learning ensemble – Naive Bayes, maximum entropy	Avg. accuracy: 0.7–0.8 Avg. precision: 0.8 Avg. sensitivity: 0.7–0.8 Avg. specificity: 0.7–0.8
[49]	joy, fear, sadness, anger – intensities	word, character	Supervised learning – SVM	
[50]	joy, sadness, fear, anger	Words	Supervised learning	Avg. accuracy: 0.8–0.9 Avg. recall: 0.8–0.9 Avg. precision: 0.9 Avg. F1: 0.8–0.9
[51]	joy, sadness, anger, neutral	Word, character, NLP feature	Supervised learning – Logistic regression, Naive Bayes, CNN	Avg. accuracy: 0.8–0.9 Avg. recall: 0.8–0.9 Avg. precision: 0.8–0.9 Avg. F1: 0.7–0.9
[52]	happiness, sadness, anger, surprise, fear, disgust	Word, emoticons	Hybrid, supervised learning (SMO, J48), word scoring	Avg. accuracy: 0.8–0.9
[53]	anger, disgust, fear, joy, sadness, surprise	Word	Deep learning – ensemble of CNN	Avg. accuracy: 0.6 Avg. F1: 0.5
[54]	anger, fear, joy, sadness	word, word embeddings	Machine learning – random forest, support vector machine deep learning – deep neural network	Avg. F1: 0.3–0.6 Avg. sensitivity: 0.3–0.6 Avg. specificity: 0.7–0.9
[6]	anger, disgust, fear, joy, sadness, surprise	words (noun, adjective, verb, adverb)	Supervised learning – Naive Bayes	Avg. accuracy: 0.1–0.6

ISSN: 2414-4592
DOI: 10.20474/jater-9.2.1

CONT.....

Authors	Emotions/Sentiments	Features	Method	Results
[55]	anger, fear, joy, sadness	word, emotion lexicon score	Deep learning – CNN	Avg. recall: 0.08–0.7 Avg. precision: 0.6–0.8 Avg. F1: 0.1–0.7
[56]	anger, disgust, fear, joy, sadness, surprise	Word	Deep learning – CNN with bidirectional long-short term memory (biLSTM)	Best accuracies: 0.5–0.9
[57]	anger, fear, joy, sadness	word, character	Machine learning – Naive Bayes, support vector machines, multilayer perceptron deep learning – CNN	Avg. accuracy: 0.9 Avg. recall: 0.9 Avg. precision: 1.0 Avg. F1: 0.9
[58]	happy, sadness, love, fear, anger, surprise.	N-gram	Semi-supervised learning with SVM	Avg. accuracy: 0.8
[41]	anger, disgust, fear, guilt, joy, shame, sadness	Word	SVM	Avg. accuracy: 0.8–0.9
[59]	joy, anger, fear, sadness, or neutral	Word	Unsupervised labelling	Avg. accuracy: 0.5-0.8 Avg. F1: 0.5–0.8
[60]	anger, afraid, happy, excited, sadness, bored, relax	Word	CNN	Avg. accuracy: 0.9 Avg. recall: 0.9 Avg. precision: 0.9 Avg. F1: 0.9
[61]	anger, sadness, fear, joy	Word	Supervised machine learning	Avg. accuracy: 0.6–0.9 Avg. recall: 0.6–0.9 Avg. precision: 0.7–0.9
[12]	Text, emoticons	TF-IF, Bag of words, N-gram, emoticon laxicons	Machine Learning, Deep Learning	Deep Learning (LSTM 89% and CNN 81%) are better and more trending than machine learning (SVM 78%, Logistic Regression 78%, Random Forest 76%, Naive Bayes 52%).

ISSN: 2414-4592
DOI: 10.20474/jater-9.2.1

Fig. 1. Trends in sentiment analysis [13]

CHALLENGES

Text classification faces the challenge of identifying features for accurate class distinction. In sentiment analysis, it's positive or negative classes; in emotion detection, it's various emotions. The common approach involves treating each word in the text as a feature or individual data point. This scheme is known as Bag-of-Words model, which treats each word as an isolated element [62]. While this approach is straightforward, it has inherent flaws and limitations. A standalone word may not adequately capture the essence of an entire text, as its meaning can vary significantly in different contexts [63]. To enhance information richness and classification accuracy, it is advisable to incorporate additional text features that convey the context of the text [64].

Challenges arise in text-related classification, including issues like text context, ambiguity, sarcasm, and implied semantics. As mentioned earlier, context plays a crucial role in interpreting words within a given text. To capture the overall information, relying solely on single words is insufficient; markers or word pairings that provide context are necessary [65]. Word ambiguity, often tied to context, is another challenge. Certain words can have multiple meanings based on their pairing or usage context. For instance, the word "break" in phrases like "break a glass," "take a break," and "break a leg" conveys entirely different meanings and sentiments. Sarcasm poses a particular challenge in speech and text-based tasks, as it often involves words with meanings opposite to their literal sense. Detecting sarcasm relies on understanding speech tones and topic context, but in textual form, this requires reliance on other markers. Moreover, determining the genuine emotion and sentiment of an author becomes intricate when they are implied rather than explicitly expressed. This poses the challenge of interpreting implied meanings rather than explicit ones. People express opinions and emotions differently, some directly and others more subtly. Direct expressions are easier to detect using specific keywords and punctuation. Nonetheless, implicit sentiment or emotion detection poses greater challenges due to the absence of clear sentiment or emotion keywords. Interpretation necessitates taking into account context, metaphors, and even cultural and social backgrounds, as speech patterns and word usage can vary significantly across individuals.

The diverse and dynamic nature of language gives rise to various challenges. Due to the distinct speech patterns exhibited by different individuals, devising general rules for categorizing ideas expressed in text becomes a complex undertaking. Consequently, a combination of numerous text features is employed to offer a more comprehensive understanding of the author's ideas, meaning, emotions, and sentiments. While single words serve as the fundamental feature, additional supporting elements such as phrases, punctuation, and emoticons, commonly found in informal texts like those on social media, contribute significantly to conveying meaning.

Text classification faces a challenge related to language usage, specifically, the distinction between formal written text and social media text. Formal text adheres to established grammar rules, displaying a recognizable pattern often found in articles, newsletters, reports, and scientific materials. On the other hand, social media text does not necessarily conform to standard language patterns and is considered informal. Despite this informality, social media texts are valuable resources due to their abundance and the prevalence of public opinions and expressions. Platforms like Twitter, Facebook, Instagram, and Reddit provides individuals with a platform to openly express feelings and opinions, and harnessing this textual data can assist in tasks such as sentiment analysis, emotion detection, and identification of issues like depression and suicide.

It is noteworthy that numerous natural language processing tools, including lexicons and annotators, have been developed based on the rules and grammar patterns of formal language. Consequently, they may prove insufficient when handling tasks related to informal text. However, the prevalence of informal text, being more extensively analyzed due to its abundance, presents a challenge in its own right.

[65] identified distinctive characteristics of social media text that set it apart from formal text. Firstly, social media texts often incorporate non-standard word forms, such as abbreviations, acronyms, redundant letter repetition, and misspelled words. Examples include expressions like "LOL," "looooooooove," "ur," and "wht" [44]. This phenomenon can be attributed to the

ISSN: 2414-4592
DOI: 10.20474/jater-9.2.1

lack of a formal requirement for proofreading in social media posts. Since these posts reflect individual musings without strict adherence to formal language rules, unconventional word forms naturally emerge. Certain platforms even enforce word count restrictions, indirectly promoting the use of short forms for typing convenience. Additionally, the repetition of letters may function as a way to express exaggeration and intense emotion [66].
Secondly, social media text diverges from formal text in sentence structure. While formal text usually consists of complete sentences with proper beginnings and endings, social media text may include fragments, half-constructed sentences, or rhetorical questions.
Thirdly, social media text exhibits site-related markups, such as hashtags, labels, mentions, and URLs. Hashtags like "depressed" and "love" act as labels, offering insights into the content of the posted text. However, in practical classification tasks, these labels are frequently removed, or at the very least, the "" sign is eliminated as it provides limited information [47]. Mentions or tagging of other users using the "" symbol is another aspect of social media text that doesn't contribute to the task's knowledge and is usually excluded.

V. CONCLUSION

In the pursuit of understanding and extracting insights from the ever-expanding realm of textual data, this research has journeyed beyond the confines of traditional sentiment analysis, unraveling a landscape rich with challenges and innovative approaches in text mining. The exploration of trends beyond sentiment analysis has illuminated the multifaceted nature of language and the evolving demands placed upon text mining methodologies.
Through a critical examination of the challenges, it becomes evident that the intricacies of context, sarcasm, and ambiguity pose formidable obstacles to the accuracy and depth of sentiment analysis. The acknowledgment of these challenges serves as a catalyst for the development and adoption of modern approaches that transcend sentiment analysis, ensuring a more nuanced and comprehensive understanding of textual data.
The study's survey of contemporary methodologies, including topic modeling, named entity recognition, and deep learning techniques, underscores the dynamism within the field of text mining. These approaches demonstrate a capacity to unveil layers of meaning that extend beyond the polarities captured by sentiment analysis alone. Moreover, the integration of domain-specific knowledge and linguistic resources emerges as a pivotal strategy for enhancing the precision and relevance of text mining outcomes.
As we conclude this exploration into the trends beyond sentiment analysis in text mining, it is evident that the field is poised for continual evolution. The synthesis of insights from this research contributes to a holistic understanding, providing guidance for researchers, practitioners, and industry professionals navigating the intricate landscape of textual data. By fostering innovation, embracing modern methodologies, and adhering to ethical considerations, the future of text mining holds promise for uncovering deeper layers of meaning and knowledge within the vast tapestry of human language.

VI. ACKNOWLEDGMENT

In the journey of completing this research, the deepest gratitude goes to Almighty Allah (S.W.T) for granting the strength, guidance, and unwavering faith throughout this endeavor. The heartfelt appreciation goes to my parents, whose prayers have been a constant support and motivation. A special acknowledgment goes to my dedicated supervisor, Ma'am Sumbul Ghulamani, for her invaluable guidance, expertise, and encouragement. Your mentorship has shaped the trajectory of my research and enriched my academic journey. I would also like to thank the esteemed faculty members for their insightful feedback and scholarly contributions. I acknowledge the SZABIST University Hyderabad for providing the necessary resources. This work is a product of collective efforts, and I am grateful for the collaborative spirit that has enriched its outcomes.

REFERENCES

[1] L. P. Hung and S. Alias, ``Beyond sentiment analysis: A review of recent trends in text based sentiment analysis and emotion detection,'' *Journal of Advanced Computational Intelligence and Intelligent Informatics*, vol. 27, no. 1, pp. 84-95, 2023. doi: https://doi.org/10.20965/jaciii.2023.p0084

[2] R. Das and T. D. Singh, ``Multimodal sentiment analysis: A survey of methods, trends, and challenges,'' *ACM Computing Surveys*, vol. 55, no. 13s, pp. 1-38, 2023. doi: https://doi.org/10.1145/3586075

[3] T. Nasukawa and J. Yi, ``Sentiment analysis: Capturing favorability using natural language processing,'' in *Proceedings of the 2nd international conference on Knowledge capture*, 2003.

[4] K. Dave, S. Lawrence, and D. M. Pennock, ``Mining the peanut gallery: Opinion extraction and semantic classification of product reviews,'' in *Proceedings of the 12th international conference on World Wide Web*, 2003.

[5] S. Shayaa, N. I. Jaafar, S. Bahri, A. Sulaiman, P. S. Wai, Y. W. Chung, A. Z. Piprani, and M. A. Al-Garadi, ``Sentiment analysis of big data: Methods, applications, and open challenges,'' *Ieee Access*, vol. 6, pp. 37 807-37 827, 2018. doi: https://doi.org/10.1109/access.2018.2851311

[6] K. Sailunaz and R. Alhajj, ``Emotion and sentiment analysis from twitter text,'' *Journal of Computational Science*, vol. 36, p. 101003, 2019.

[7] Q. Liu, Z. Gao, B. Liu, and Y. Zhang, ``Automated rule selection for aspect extraction in opinion mining,'' in *Twenty-Fourth International Joint Conference on Artificial Intelligence*, 2015.

[8] E. Cambria, S. Poria, A. Gelbukh, and M. Thelwall, ``Sentiment analysis is a big suitcase,'' *IEEE Intelligent Systems*, vol. 32, no. 6, pp. 74-80, 2017. doi: https://doi.org/10.1109/mis.2017.4531228

[9] A. Ghafoor, A. S. Imran, S. M. Daudpota, Z. Kastrati, S. Shaikh, and R. Batra, ``SentiUrdu-1M: A large-scale tweet dataset for Urdu text sentiment analysis using weakly supervised learning,'' *Plos One*, vol. 18, no. 8, p. e0290779, 2023.

[10] E. Cambria, D. Das, S. Bandyopadhyay, and A. Feraco, *A practical guide to sentiment analysis*. Springer, 2017.

[11] A. Alzu'bi, L. Bani Younis, A. Abuarqoub, and M. Hammoudeh, ``Multimodal deep learning with discriminant descriptors for offensive memes detection,'' *ACM Journal of Data and Information Quality*, vol. 15, no. 3, pp. 1-16, 2023. doi: https://doi.org/10.1145/3597308

[12] M. A. Ullah, S. M. Marium, S. A. Begum, and N. S. Dipa, ``An algorithm and method for sentiment analysis using the text and emoticon,'' *ICT Express*, vol. 6, no. 4, pp. 357-360, 2020. doi: https://doi.org/10.1016/j.icte.2020.07.003

[13] B. Liu, *Sentiment analysis and opinion mining*. Springer Nature, 2022.

[14] R. Kibble. (2013) Introduction to natural language processing. [Online]. Available: https://shorturl.at/GztHgs

[15] B. Pang, L. Lee, and S. Vaithyanathan. (2002) Thumbs up? sentiment classification using machine learning techniques. [Online]. Available: https://shorturl.at/u2PX8

[16] K. Nongmeikapam, D. Khangembam, W. Hemkumar, S. Khuraijam, and S. Bandyopadhyay, ``Verb based manipuri sentiment analysis,'' *International Journal on Natural Language Computing*, vol. 3, no. 3, pp. 113-118, 2014. doi: https://doi.org/10.5121/ijnlc.2014.3311

[17] R. Das and T. D. Singh, ``A step towards sentiment analysis of assamese news articles using lexical features,'' in *Proceedings of the International Conference on Computing and Communication Systems: I3CS 2020, NEHU, Shillong, India*. Springer, 2021.

[18] A. Balahur and M. Turchi, ``Multilingual sentiment analysis using machine translation?'' in *Proceedings of the 3rd workshop in computational approaches to subjectivity and sentiment analysis*, 2012.

[19] K. Kanclerz, P. Miłkowski, and J. Kocoń, ``Cross-lingual deep neural transfer learning in sentiment analysis,'' *Procedia Computer Science*, vol. 176, pp. 128-137, 2020. doi: https://doi.org/10.1016/j.procs.2020.08.014

[20] R. Mihalcea, C. Banea, and J. Wiebe, ``Learning multilingual subjective language via cross-lingual projections,'' in *Proceedings of the 45th annual meeting of the association of computational linguistics*, 2007.

[21] T. Wilson, P. Hoffmann, S. Somasundaran, J. Kessler, J. Wiebe, Y. Choi, C. Cardie, E. Riloff, and S. Patwardhan, ``OpinionFinder: A system for subjectivity analysis,'' in *Proceedings of HLT/EMNLP 2005 Interactive Demonstrations*, 2005.

[22] K. Denecke, ``Using sentiwordnet for multilingual sentiment analysis,'' in *2008 IEEE 24th International Conference on Data Engineering Workshop*. IEEE, 2008.

[23] S. Baccianella, A. Esuli, F. Sebastiani *et al.*, ``Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining,'' in *Lrec*, 2010.

[24] K. Dashtipour, S. Poria, A. Hussain, E. Cambria, A. Y. Hawalah, A. Gelbukh, and Q. Zhou, ``Multilingual sentiment analysis: State of the art and independent comparison of techniques,'' *Cognitive Computation*, vol. 8, pp. 757-771, 2016.

[25] A. Balahur and M. Turchi, ``Improving sentiment analysis in twitter using multilingual machine translated data,'' in *Proceedings of the International Conference Recent Advances in Natural Language Processing RANLP 2013*, 2013.

[26] A. Cui, M. Zhang, Y. Liu, and S. Ma, ``Emotion tokens: Bridging the gap among multilingual twitter sentiment analysis,'' in *Asia information retrieval symposium*. Springer, 2011.

[27] X. Meng, F. Wei, X. Liu, M. Zhou, G. Xu, and H. Wang, ``Cross-lingual mixture model for sentiment classification,'' in *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 2012.

[28] R. Xu, J. Xu, and X. Wang, ``Instance level transfer learning for cross lingual opinion analysis,'' in *Proceedings of the 2nd Workshop on Computational Approaches to Subjectivity and Sentiment Analysis (WASSA 2.011)*, 2011.

[29] S. L. Lo, E. Cambria, R. Chiong, and D. Cornforth, ``Multilingual sentiment analysis: From formal to informal and scarce resource languages,'' *Artificial Intelligence Review*, vol. 48, pp. 499-527, 2017.

[30] M. Araújo, A. Pereira, and F. Benevenuto, ``A comparative study of machine translation for multilingual sentence-level sentiment analysis,'' *Information Sciences*, vol. 512, pp. 1078-1102, 2020.

[31] D. A. Pereira, ``A survey of sentiment analysis in the portuguese language,'' *Artificial Intelligence Review*, vol. 54, no. 2, pp. 1087-1115, 2021. doi: https://doi.org/10.1007/s10462-020-09870-1

[32] M. Hu and B. Liu, ``Mining and summarizing customer reviews,'' in *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, 2004.

[33] A. Hasan, S. Moin, A. Karim, and S. Shamshirband, ``Machine learning-based sentiment analysis for twitter accounts,'' *Mathematical and Computational Applications*, vol. 23, no. 1, p. 11, 2018.

[34] W. Zhao, Z. Guan, L. Chen, X. He, D. Cai, B. Wang, and Q. Wang, ``Weakly-supervised deep embedding for product review sentiment analysis,'' *IEEE Transactions on Knowledge and Data Engineering*, vol. 30, no. 1, pp. 185-197, 2017. doi: doi: 10.1109/TKDE.2017.2756658

[35] H. Xu, F. Zhang, and W. Wang, ``Implicit feature identification in Chinese reviews using explicit topic mining model,'' *Knowledge-Based Systems*, vol. 76, pp. 166-175, 2015. doi: doi: 10.1109/TASLP.2015.2443982

[36] W. Che, Y. Zhao, H. Guo, Z. Su, and T. Liu, ``Sentence compression for aspect-based sentiment analysis,'' *IEEE/ACM Transactions on audio, speech, and language processing*, vol. 23, no. 12, pp. 2111-2124, 2015. doi: doi: 10.1109/TASLP.2015.2443982

[37] M. S. Akhtar, T. Garg, and A. Ekbal, ``Multi-task learning for aspect term extraction and aspect sentiment classification,'' *Neurocomputing*, vol. 398, pp. 247-256, 2020. doi: doi: 10.1016/J.NEUCOM.2020.02.093

[38] J. Ma, X. Cai, D. Wei, H. Cao, J. Liu, and X. Zhuang, ``Aspect-based attention LSTM for aspect-level sentiment analysis,'' in *2021 3rd World Symposium on Artificial Intelligence (WSAI)*. IEEE, 2021.

[39] S. M. Weiss, N. Indurkhya, T. Zhang, and F. Damerau, *Text mining: Predictive methods for analyzing unstructured information*. Springer Science & Business Media, 2010.

[40] R. Kibble. (2013) Introduction to natural language processing. [Online]. Available: https://shorturl.at/B3gwh

[41] Z. Erenel, O. R. Adegboye, and H. Kusetogullari, ``A new feature selection scheme for emotion recognition from text,'' *Applied Sciences*, vol. 10, no. 15, p. 5351, 2020. doi: doi: 10.3390/APP10155351

[42] H. Rahman, J. Tariq, M. A. Masood, A. F. Subahi, O. I. Khalaf, and Y. Alotaibi, ``Multi-tier sentiment analysis of social media text using supervised machine learning,'' *Comput. Mater. Contin*, vol. 74, no. 3, pp. 5527-5543, 2023. doi: doi: 10.32604/cmc.2023.033190

[43] F. Benrouba and R. Boudour, ``Emotional sentiment analysis of social media content for mental health safety,'' *Social Network Analysis and Mining*, vol. 13, no. 1, p. 17, 2023. doi: doi: 10.1007/S13278-022-01000-9/METRICS.

[44] P. Ingole, S. Bhoir, and A. Vidhate, ``Hybrid model for text classification,'' in *2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA)*. IEEE, 2018.

[45] Y. Wang and A. Pal, ``Detecting emotions in social media: A constrained optimization approach,'' in *Twenty-fourth international joint conference on artificial intelligence*, 2015.

[46] S. N. Shivhare, S. Garg, and A. Mishra, ``EmotionFinder: Detecting emotion from blogs and textual documents,'' in *International Conference on Computing, Communication & Automation*. IEEE, 2015.

[47] M. Hajar *et al.*, ``Using youtube comments for text-based emotion recognition,'' *Procedia Computer Science*, vol. 83, pp. 292-299, 2016. doi: doi: 10.1016/J.PROCS.2016.04.128

[48] I. Perikos and I. Hatzilygeroudis, ``Recognizing emotions in text using ensemble of classifiers,'' *Engineering Applications of Artificial Intelligence*, vol. 51, pp. 191-201, 2016. doi: doi: 10.1016/J.ENGAPPAI.2016.01.012

[49] S. M. Mohammad and F. Bravo-Marquez, ``Emotion intensities in tweets,'' in *6th Joint Conference on Lexical and Computational Semantics, Proceedings*, 2017.

[50] S. X. Mashal and K. Asnani, ``Emotion intensity detection for social media data,'' in *2017 International Conference on computing methodologies and communication (ICCMC)*. IEEE, 2017.

[51] H. AlBalooshi, S. Rahmanian, and R. V. Kumar, ``EmotionX-SmartDubai_NLP: Detecting user emotions in social media

text,'' in *Proceedings of the Sixth International Workshop on Natural Language Processing for Social Media*, 2018.

[52] B. Gaind, V. Syal, and S. Padgalwar. (2019) Emotion detection and analysis on social media. [Online]. Available: https://shorturl.at/7D9xh

[53] W. Witon, P. Colombo, A. Modi, and M. Kapadia, ``Disney at iest 2018: Predicting emotions using an ensemble,'' in *Proceedings of the 9th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, 2018.

[54] B. Kratzwald, S. Ilić, M. Kraus, S. Feuerriegel, and H. Prendinger, ``Deep learning for affective computing: Text-based emotion recognition in decision support,'' *Decision Support Systems*, vol. 115, pp. 24-35, 2018. doi: doi: 10.1016/J.DSS.2018.09.002

[55] S. Ibrahiem, S. Ismail, K. Bahnasy, and M. Aref, ``Convolutional neural network multi-emotion classifiers,'' *Jordanian Journal of Computers and Information Technology*, vol. 5, no. 2, p. 50, 2019.

[56] E. Batbaatar, M. Li, and K. H. Ryu, ``Semantic-emotion neural network for emotion recognition from text,'' *IEEE Access*, vol. 7, pp. 111 866-111 878, 2019. doi: doi: 10.1109/ACCESS.2019.2934529

[57] M. Baali and N. Ghneim, ``Emotion analysis of arabic tweets using deep learning approach,'' *Journal of Big Data*, vol. 6, no. 1, p. 89, 2019. doi: doi: 10.1186/s40537-019-0252-x

[58] K. S. N. Tan, T. M. Lim, and Y. M. Lim, ``Emotion analysis using self-training on malaysian code-mixed twitter data,'' in *International Conferences ICT, Society, and Human Beings 2020; Connected Smart Cities 2020; and Web Based Communities and Social Media 2020*. IADIS Lisbon, Portugal, 2020.

[59] S. Zad and M. Finlayson, ``Systematic evaluation of a framework for unsupervised emotion recognition for narrative text,'' in *Proceedings of the First Joint Workshop on Narrative Understanding, Storylines, and Events*, 2020.

[60] C. S. A. Razak, S. H. Ab Hamid, H. Meon, N. B. Anuar *et al.*, ``Two-step model for emotion detection on twitter users: A Covid-19 case study in Malaysia,'' *Malaysian Journal of Computer Science*, vol. 34, no. 4, pp. 374-388, 2021.

[61] A. Chowanda, R. Sutoyo, S. Tanachutiwat *et al.*, ``Exploring text-based emotions recognition machine learning techniques on social media conversation,'' *Procedia Computer Science*, vol. 179, pp. 821-828, 2021. doi: 10.1016/J.PROCS.2021.01.099

[62] F. Song, S. Liu, and J. Yang, ``A comparative study on text representation schemes in text categorization,'' *Pattern Analysis and Applications*, vol. 8, pp. 199-209, 2005. doi: doi: 10.1007/S10044-005-0256-3/TABLES/14

[63] L. P. Hung, R. Alfred, M. H. Ahmad Hijazi, A. Ibrahim, and A. Asri, ``A review on the ensemble framework for sentiment analysis,'' *Advanced Science Letters*, vol. 21, no. 10, pp. 2957-2962, 2015. doi: doi: 10.1166/ASL.2015.6494

[64] O. Bruna, H. Avetisyan, and J. Holub, ``Emotion models for textual emotion classification,'' *Journal of Physics*, vol. 772, no. 1, p. 012063, 2016.

[65] K. O. ans Dacher Keltner and J. M. Jenkins. (2023) Understanding emotions, 2nd ed.''. [Online]. Available: https://shorturl.at/EhMUO

[66] E. Cambria, B. Schuller, Y. Xia, and C. Havasi, ``New avenues in opinion mining and sentiment analysis,'' *IEEE Intelligent Systems*, vol. 28, no. 2, pp. 15-21, 2013. doi: 10.1109/MIS.2013.30

ISSN: 2414-4592
DOI: 10.20474/jater-9.2.1