A Comprehensive Review of Approaches, Methods, and Challenges and Applications in Sentiment Analysis

BV Pranay Kumar[1], M Sadanandam[2]
[1]Research Scholar, Department of CSE, Kakatiya University, Warangal & Assistant Professor, CJITS Jangaon, JNTU Hyderabad.
[2]Associate Professor, Department of CSE, KU College of Engineering and Technology, Kakatiya University, Warangal.
[1]pranaybv4u@gmail.com,[2]sadanb4u@yahoo.co.in

Abstract The rapid growth of Internet-based applications like social media platforms and blogs has resulted in comments and reviews regarding everyday routines. The process of collecting and analyzing people's opinions, thoughts, and impressions concerning various topics, products, subjects, and services are known as sentiment analysis. People's opinions can help corporations, governments, and individuals obtain information and make decisions based on those opinions. However, the sentiment analysis and evaluation procedure are full of challenges. These hardships make it challenging to accurately interpret sentiments and figure out the appropriate sentiment polarity. Sentiment analysis employs natural language processing and text mining to recognize and extract subjective information from text. This article offers a thorough explanation of the method for completing this task as well as sentiment analysis applications. Then, to fully comprehend the benefits and drawbacks of each method, it assesses, contrasts, and investigates them. To define future directions, the difficulties of sentiment analysis are finally examined.
Keywords Sentiment analysis, Text analysis, Methods, Approaches, Word embedding, Machine learning, Social media.

1. Introduction

Sentiment analysis has gained widespread acceptance in recent years among researchers and businesses, governments, and organizations The growing popularity of the Internet has lifted the web to the rank of the principal source of universal information. Lots of users use various online resources to express their views and opinions. To constantly monitor public opinion and aid decision-making, we must employ user-generated data to analyze it automatically. As a result, sentiment analysis has increased its popularity across research communities in recent years. Sentiment analysis is also called Opinion analysis or Opinion mining. We have seen a recent growth in the sentiment analysis task. In (Piryani et al. 2017) discuss the study topic from 2000 to 2015 and provides a framework for computationally processing unstructured data with the primary goal of extracting views and identifying their moods. Additionally, In the work of Yue et al. (2019) and Liu et al. (2012) researched the effectiveness of internet reviews.

The growth of social network sites has generated a slew of fields devoted to analyzing these networks and their contents to extract necessary information. Sentiment analysis is concerned with deriving the sentiments communicated by a piece of text from its content. Sentiment analysis is a sub-field of NLP and, given long and illustrious public opinion for decision-making, there must be multiple early works addressing it. However, it still works going on sentiment analysis develop till the new millennium. Several real-world applications require sentiment analysis for detailed investigation. for example, product analysis, discovering which components or qualities of a product appeal to customers in terms of product quality.

The two strategies for constructing two dynamic lexicons to aid in the classification of sentiments depending on their aspects: are a strategy based on statistics and genetic algorithms. The business sector has always utilized sentiment analysis for its improvement. Sentiment analysis for various applications like reputation management, market research, competitor analysis, product analysis, customer voice, etc. Various issues are associated with sentiment analysis and natural language processing, such as individuals' informal writing style, sarcasm, irony, and language-specific challenges. There are many words in different languages whose meaning and orientation change depending on the context and domain in which they are employed. Therefore, there are not many tools and resources available for all languages. Sarcasm and irony are two of the most critical challenges that have recently attracted the attention of researchers. There has been much development in detecting sarcasm and irony in the text. There are many challenges in sentiment analysis. In this work, we will analyze the various challenges, methodologies, applications, and algorithms that are employed in sentiment analysis.

This study provides a comprehensive investigation of sentiment analysis by discussing this area from various perspectives since it encompasses numerous research components connected to sentiment analysis, such as problems, applications, tools, and approaches.

The literature survey paper is organized into Sect. 2, Level of Sentiment Analysis, Sect. 3, which contain the Data Collection, Feature Extraction, and Feature Selection Method, explaining all the steps from data extraction to various task of Sentiment Analysis, Sect. 4 contain General Methodology for Sentiment Analysis and its Summary, Sect. 5, discusses the performance evaluation parameters, Sect. 6, Contain the Sentiment Analysis Application in Various Domains, Sect. 7, Contain the Challenges in Sentiment analysis, In the final Sect. 8, We Conclude our research work.

2. Levels of Sentiment Analysis

Sentiment analysis has been investigated on several levels: Document Level, Sentence Level, Phrase Level, and Aspect Level. Sentiment analysis in each level such as document, sentence and phrase, and aspect level.

2.1. Document-level sentiment analysis

Document-level: Document-level sentiment analysis is performed on a whole document, and single polarity is given to the whole document. This type of sentiment analysis is not used a lot. It can be used to classify chapters or pages of a book as positive, negative, or neutral. At this level, both supervised and unsupervised learning approaches can be utilized to classify the document (Bhatia et al. 2015). Cross-domain and cross-language sentiment analysis are the two most significant issues in document-level sentiment analysis. Domain-specific sentiment analysis has been shown to achieve remarkable accuracy while staying highly domain-sensitive.

2.2. Sentence-Level Sentiment Analysis

Sentence level: In this level of analysis, each sentence is analyzed and found with a corresponding polarity. This is highly useful when a document has a wide range and mix of sentiments associated with it (Yang and Cardie 2014). This classification level is associated with subjective classification. Each sentence's polarity will be determined independently using the same methodologies as the document level but with greater training data and processing resources. The polarity of each sentence may be aggregated to find the sentiment of the document or used individually. Occasionally, document-level sentiment analysis is insufficient for specific uses (Behdenna et al. 2018).

2.3. Phrase-Level Sentiment Analysis

Phrase level: Sentiment analysis also be performed where opinion words are mined at the phrase level, and classification will be done. Each phrase may contain multiple aspects or single aspects. This may be useful for product reviews of multiple lines; here, it is observed that a single aspect is expressed in a phrase (Thet et al. 2010). It has been a hot topic for researchers in recent times. While document-level analysis concentrated on categorizing the entire document as subjective, either positively or negatively, sentence-level analysis is more beneficial as a document contains both positive and negative statements. Word is the most basic unit of language; its polarity is intimately related to the subjectivity of the sentence or document in which it appears.

2.4. Aspect Level Sentiment Analysis

Aspect level: sentiment analysis is performed at the aspect level. Each sentence may contain multiple aspects; therefore, Aspect level sentiment analysis. Primary attention to all the aspects used in the sentence and assigns polarity to all the aspects after which an aggregate sentiment has calculated for the whole sentence (Schouten and Frasincar 2015; Lu et al. 2011).

3. Data Collection and Feature Selection

3.1. Data Collection

Data can be collected from the internet via web scraping, social media, news channels, E-commerce websites, Forums, Weblog, and some other websites. Data Collection is an important stage in Sentiment Analysis. Depending on the task sentiment analysis of text data can be combined with other types of data like video, audio, location, etc.

A few essential sources of data collection are *Social media*: Social data refers to information gathered via social media networks. It demonstrates how consumers interact with the product by accessing, posting, and exchanging.

Forums: Users can use message boards to discuss various topics, exchange opinions and ideas, and solicit assistance via text messages. Forums are an intriguing source for sentiment analysis due to the dynamic nature of user-generated information.

Weblog: A short weblog consists of paragraphs conveying a viewpoint, facts, personal diary entries, or links.

Electronic Commerce website: Electronic Commerce websites where users can give evaluations and express their opinions about a particular business or organization.

3.2. Feature Selection

It is important to remember that developing a classification model requires first identifying relevant features in the dataset. Thus, a review can be decoded into words during model training and appended to the feature vector. For a single word is considered, the technique is called a “Uni-gram”; when two words are considered, the technique is called a “Bi-gram”; and if three words are considered, the technique is referred to as a “Tri-gram.” combination of unigram and bigram helpful for analysis (Razon and Barnden 2015); the context feature which helpful for getting results most accurate.

Pragmatic features are those that emphasize the application of words rather than a methodological foundation. Pragmatics is the study of how context relates to perception in linguistics and related sciences. Pragmatics is the study of phenomena such as implicature, speech acts, relevance, and conversations.

Emoji are facial expressions used in sentiment analysis to convey emotions. Various emoticons are used to depict a wide variety of human emotions (Tian et al. 2017). Emoticons aid in conveying a person’s tone when composing a sentence and so aid in sentiment analysis. Substitute their meaning for the emoticons: The review contains a range of emotions, including happiness, sadness, and rage. Emoticons are classified into two categories: positive and negative sentiment emotions. Positive emotions are formed of positive emotions such as love, happiness, and joy, while negative emoticons are composed of negative emotions such as sadness, depression, and wrath.

Punctuation marks, or exclamation marks, serve to highlight the force of a positive or negative remark. Similarly, the apostrophe and the question mark are other punctuation marks.

Words in slang, such as *lol* and *rofl*. These are frequently used to introduce a sense of humor into a remark. Given the nature of opinion tweets, it is plausible to assume that a slang expression in the text suggests sentiment analysis. Substitute their meaning for the slang term.

Punctuation marks, like exclamation marks, serve to highlight the force of a positive or negative remark. Similarly, the apostrophe and the question mark are other punctuation marks.

3.3. Feature Extraction

Feature extraction is a key task in sentiment classification as it involves the extraction of valuable information from the text data, and it will directly impact the performance of the model. The approach tries to extract valuable information that encapsulates the text's most essential features. In most cases, punctuations are removed from the text after lowering it in the pre-processing stage, but they are used to extract features and hashtags, and emoticons are commonly used techniques for feature extractions listed below.

Terms frequency It is one of the simplest ways to express features that are more frequently used in various NLP applications, including Sentiment Analysis, for information retrieval. It considers a single word, i.e., uni-gram or group of two-three words, which can be in bi-gram and tri-gram, with their terms count representing features (Sharma et al. 2013). The term's presence gives the word a value of either *0 or 1*. Term frequency is the integer value, which is its count in the given document. TF-IDF can be used as a weighted scheme for better results that will measure the importance of any token in the given document.

Parts of Speech tagging The process of tagging a word in a text (corpus) based on its definition and context is also known as grammatical tagging. Tokens are categorized as nouns, verbs, pronouns, adverbs, adjectives, and prepositions.

Negations These are the words that can change or reverse the polarity of the opinion and shift the meaning of a sentence. Commonly used negation words include not, cannot, neither, never, nowhere, none, etc. Every word appearing in the sentence will not reverse the polarity; therefore, removing all negation words from stop-words may increase the computational cost and decrease the model's accuracy. Negation words must be handled with at most care (George et al. 2013). Negation words such as not, neither, nor, and so on are critical for sentiment analysis since they can revert the polarity of a given phrase. However, reversing the polarity is not straightforward because negation words might occur in a sentence without affecting the text's emotion. In some cases, neutral sentiment is also included, and neutral evaluations are frequently ignored in many sentiments analysis tasks due to their vagueness and lack of information.

Bag of Words (BoW) BoW is one of the simplest approaches for extracting text features. BoW will describe the occurrence of words in a document. The bag represents the vocabulary of words using which a vector is formed for each sentence. The main problem with this model is that it does not consider the syntactic meaning of the text. For instance, consider two sentences s1= "the food was good", s2= "the service was bad". The vocabulary is created for two sentences where v= {'the', 'food', 'was', 'service', 'bad', 'good'} and the length of the vector is 6 and is represented as v1= [1 1 1 0 0 1] and v2= [1 0 1 1 1 0]. BoW approach performance was evaluated using (TF-IDF) which performs better in most cases.

3.3.1. Word Embedding

Word embeddings represent words in a vector space by clustering words with similar meanings together. Each word is assigned to a vector, which is then learned like neural networks. It learns and chooses a vector from a predetermined vocabulary. The dimension of the words may be chosen by passing it as a hyperparameter. SG model and the continuous CBOW model are two of the most well-known algorithms for word embeddings. Both of these are shallow window approaches methods in which a short window of some size, such as four or six, is specified, and the current word is anticipated using context words in CBOW, while context words are forecasted using the current word in the SG model. Word embeddings are concerned with learning about words in the context of their local usage, which is specified by a window of nearby terms.

Word2vec: word2vec is a 2-layer neural network that is used for vectorizing the tokens. It is one of the famous and widely used vectorizing techniques developed by Mikolov et al. (2013). Word2vec mainly has two models CBOW and SG. The CBOW model predicts the target word using context words, whereas the SG model predicts the target word using context words. With a larger dataset, the SG model performs better.

Global Vectors (GloVe) Global Vectors for word representation have been developed by an unsupervised learning approach to generate word embeddings from a corpus word-to-word co-occurrence matrix. The gloVe is a popularly used method as it is straightforward and quick to train the GloVe model because of its parallel implementation capacity.
Fast Text It is an open-source and free library developed by FAIR (Facebook AI Research) mainly used for word classifications, vectorization, and creation of word embeddings. It uses a linear classifier to train the model, which is very fast in training the model (Bojanowski et al. 2017). It supports a CBOW and SG model. Semantic similarities may be found using this model.

ELMo ELMo is a deep contextualized text representation. ELMo contributes to overcoming the limitations of conventional word embedding approaches such as LSA, TF-IDF, and n-grams models. ELMo generates embeddings to words based on the contexts in which they are used to record the word’s meaning and retrieve additional contextual information. Through pretraining, Elmo can more accurately represent polysemous words in a variety of contexts and is more informative about the text’s higher-level semantics.

3.4. Feature Selection Approach

Feature Selection Approach is evaluated to identify a data characteristic. A characteristic can be insignificant, significant, or redundant. Various feature selection approaches are used to eliminate irrelevant and superfluous characteristics. Feature Selection is a procedure that identifies and eliminates superfluous and irrelevant characteristics from the feature list and thus increases sentiment classification accuracy. The benefit of these approaches is their efficacy, as they carefully address aspects. Choosing handcrafted features is a lengthy and complicated process.

SentiWordNet is a sentiment lexicon built from the WordNet database, with each term accompanied by numerical values indicating positive and negative sentiment. A well-known example is a SentiWordNet lexicon. contrarily, statistical processes are entirely automated and are widely used for feature selection, although they typically fail to distinguish between sentimental and non-sentimental features (Poria et al. 2014; Varelas et al. 2005).

Statistical techniques for feature selection are typically categorized into four categories: filter, embedding, wrapper, and hybrid.

Filter approach This is the most often used technique of selecting features. It selects features without utilizing any machine learning technique based on the general properties of the training data. The feature is ranked using several statistical metrics, and then the features with the highest rankings are chosen (Adomavicius and Kwon 2011). They are computationally inexpensive and well-suited for datasets with a high number of attributes. The words “Information Gain”, “Chi-square”, “Document Frequency”, and “Mutual information” are all used to refer to fundamental filter algorithms.

Wrapper approach This approach is based on machine learning algorithms since it relies on the output of the machine learning algorithm. Approaches are often iterative and computationally demanding due to this dependency, but they can determine the optimal feature set for that particular modeling algorithm. Wrapper techniques include creating feature subsets (forward or backward selection) plus various learning algorithms(such as NB or SVM).

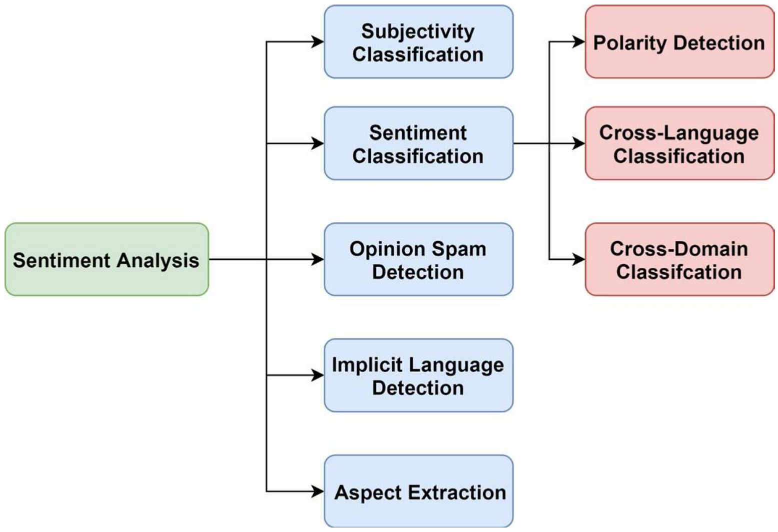

Fig. 1. Task of sentiment analysis

Embedded approach This method combines the feature selection procedure with the execution of the modeling algorithm. It employs classification methods that have a built-in feature selection capability. As a result, it is more computationally efficient than the wrapper approach.

Hybrid approach This strategy combines filter and wrapper approaches; hybrid methods generally utilize multiple approaches to produce the optimum feature subset. Hybrid techniques typically achieve excellent performance and accuracy through the use of many approaches. Numerous hybrid feature selection algorithms for sentiment analysis have been developed.

3.5. Task of Sentiment Analysis

Overview of the various task of sentiment analysis is shown in Fig.1 and explain as follows.

Subjectivity classification This is frequently assumed to be the first stage in sentiment analysis. Subjectivity classification recognizes subjective hints, emotional phrases, and subjective ideas. Subjectivity classification aims to keep undesirable objective data items out of subsequent processing.

Sentiment classification Sentiment categorization is a well-known researched task in sentiment analysis. Polarity determination is one of the subtasks of sentiment classification, and the term “Opinion analysis” is frequently used when referring to Sentiment Analysis. It is a little dusty and aimed at determining the sentiment of each piece of text. Polarity is traditionally either positive or negative (Wang et al. 2014). In the work of Xia et al. (2015), the opinion-level context is investigated, with intra-opinion and inter-opinion aspects being finely characterized. Neutral is also included in some cases. With a trained classifier, the cross-domain analysis predicts the sentiment of a target domain. Extracting the domain invariant features and where they are distributed is a commonly used approach. The cross-language analysis is done similarly by training the model on a dataset from a source language and then evaluating it on a dataset from a different language with limited data. The ambiguity of word polarity is one of the obstacles that sentiment analysis must overcome. Affective computing and sentiment analysis also have tremendous potential as subsystem technology for other systems (Cambria et al. 2017). They can augment the capabilities of customer relationship management and recommendation systems by enabling the discovery of which features customers particularly enjoy or the omission of items that have received highly unfavorable feedback from the suggestions.

Opinion Spam Detection Spam Detection has become a significant challenge in sentiment analysis because of the rising interest in e-commerce and review platforms. Opinion spams, often known as fraudulent or phone reviews, are well-written comments supporting or criticizing a product for their benefit. Opinion spam detection seeks to identify three distinct characteristics of a phone review: the review’s content, the review’s metadata, and real-world product expertise. Machine learning algorithms are frequently used to assess review material to detect dishonesty. The star or point ratings, IP address of the user, geolocation of the user, and other information are a few Metadata used in detecting spam opinions. In many circumstances, though, it is inaccessible for analysis, Real-world experience and knowledge are included thirdly. For example, if a product with bad ratings and reviews is being rated high for a period, that can be put under suspicion and analyzed for opinion spam detection.

Implicit Language Detection Sarcasm, irony, and humor are generally referred to as Implicit Language. These equivocal and ambiguous form is speech is an arduous task to detect, even by humans sometimes. However, this implicit language is an essential aspect of a sentence and can completely flip the meaning and polarity of the sentence. For instance, consider the phrase ’Brilliant, I am fired’. The word Brilliant is very positive, but it describes irony or sarcasm when combined with later parts, i.e., "I am fired" which makes the phrase "I am fired" more negative. Investigating signs such as emoticons, laughter emoticons, and extensive punctuation mark utilization are classic approaches for detecting implicit language.
Aspect Extraction: Aspect-level sentiment analysis is mainly composed of three steps aspect extraction, polarity classification, and aggregation.

The process of aspect-based sentiment analysis starts with the extraction of aspect, one of the key processes as this differentiates usual sentiment analysis. Aspects can be extracted using a predefined set of aspects which should be carefully predefined based on the domain on which it is used. Other approaches are more sophisticated approaches like Frequency-based methods, syn- tax-based methods supervised and unsupervised machine learning approaches. It has been seen that in reviews, few words are used more frequently than others and these most frequent terms are more likely to turn out as aspects; this straightforward method can turn out into quite a powerful approach fact that a significant number of approaches. This approach has a few shortcomings because all frequent nouns do not refer to aspects, terms like ’bucks,’ ’dollars,’ ’rupees,’ etc.

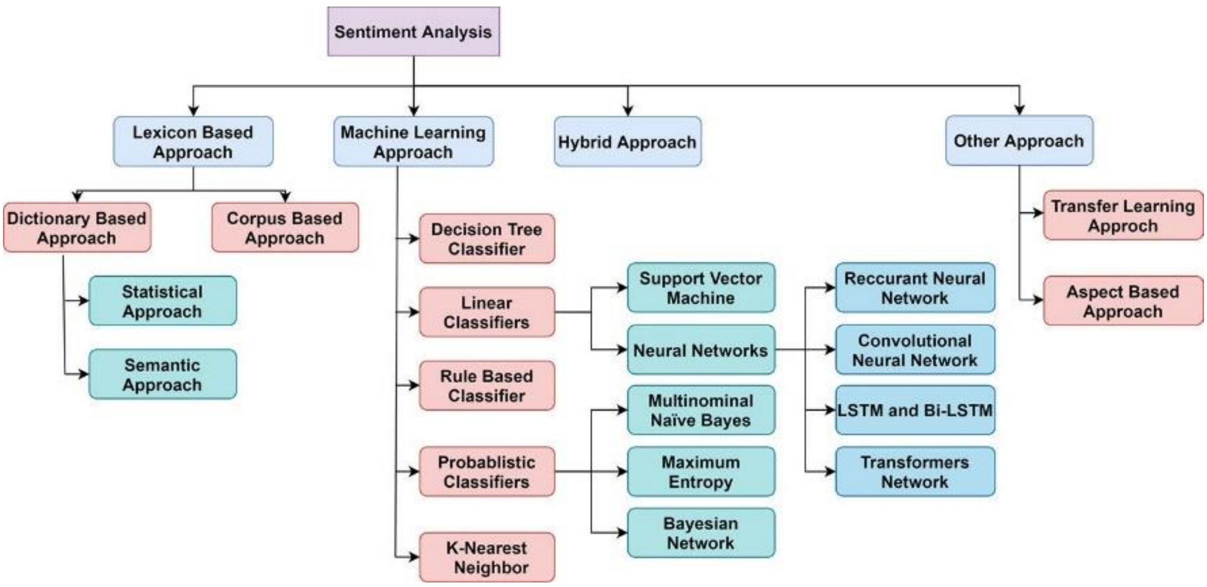

Fig.2. Approach of sentiment analysis

Also, aspects that are not mentioned frequently can be missed by this method. A set of rules can be supplemented with a frequency-based approach to overcome these problems, but these manually crafted rules tend to come from parameters that need to be tuned manually, which is a hectic and time-consuming task. Instead of focusing on the Frequency-based approach. The syntax-based approach can be used as this approach covers the flaws of the frequency-based approach of not detecting less frequent aspects (Bai et al. 2020). In this approach, For example, here, 'Awesome' refers to an adjective referring to the aspect "food" in 'Awesome food.' For this approach, many annotated data covering all syntactical relations should be collected for training the algorithm.

3.5.1. Need for Sentiment Analysis

Sentiment analysis is incredibly significant since it helps businesses understand consumer sentiment toward their brand. By automatically classifying the emotions behind social media interactions, reviews, and more, organizations can make informed decisions. Sentiment Analysis refers to the methods and strategies that enable firms to examine data about how their customer base feels about a given service or product.

Sentiment Analysis is a process that analyzes natural language utterances automatically, discovers essential claims or opinions, and classifies them according to their emotional attitude. In business needs sentiment analysis has increased consumer happiness through enhanced products, real-time problem detection, and market distinctiveness. Customer satisfaction analysis through sentiment analysis is enhanced. Identify and act on real-time problems.

4. Methodology

Three mainly used approaches for Sentiment Analysis include Lexicon Based Approach, Machine Learning Approach, and Hybrid Approach. In addition, researchers are continuously trying to figure out better ways to accomplish the task with better accuracy and lower computational cost.

4.1.1. Lexicon-Based Approach

Lexicons are the collection of tokens where each token is assigned a predefined score which indicates the neutral, positive, and negative nature of the text. A score is assigned to tokens based on polarity such as $+1, 0, -1$ for positive, neutral, and negative, or the score may be assigned based on the intensity of polarity and its values range from $[+1, -1]$ where $+1$ represents highly positive, and -1 represents highly negative. In Lexicon Based Approach, for a given review or text, the aggregation of scores of each token is performed, i.e., positive, negative, and neutral scores are summed separately. In the final stage, overall polarity is assigned to the text based on the highest value of individual scores. Thus, the document is first divided into tokens of single words, where-after the polarity of each token is calculated and aggregated in the end.

The lexicon-based technique is extremely feasible for sentiment analysis at the sentence and feature levels. Because no training data is required, it might be termed an unsupervised technique. On the other side, the primary disadvantage of this technique is domain dependence, as words can have several meanings and senses, and therefore a positive word in one domain may be negative in another.

The advantage of the lexicon-based approach is that not require any training data and is considered an unsupervised approach by some experts (Yan-Yan et al. 2010). The main disadvantage of the lexicon-based approach is that it is highly domain orientated and words about one domain cannot be used in another domain. For instance, consider the word huge it may be positive or negative based on the domain in which it is being used. In “the queue for the movie was huge” the word may be considered positive whereas, in “there was a huge lag in network” the word can be considered negative. Therefore, the polarity should be assigned to words carefully, considering the domain. There are mainly two approaches used in Lexicon-Based Approaches: Corpus-Based and Statistical Approaches.

4.1.2. Corpus-Based Based Approach

The approach employs semantic and syntactic patterns to ascertain the sentence’s emotion. This approach begins with a predefined set of sentiment terms and their orientation and then investigates syntactic or similar patterns to discover sentiment tokens and their orientation in a huge corpus. This is a situation-specific method that requires a significant amount of labeled data to train. However, it aids in resolving the issue of opinion words with context-dependent orientations.

Statistical Approach The seed opinion words or co-occurrence patterns can be found using a statistical approach. The rough idea behind this approach is that if it appears in positive texts more than negative texts, then it is more likely to be positive or vice versa. The key premise of this approach is that if comparable sentiment tokens are frequently observed in the same environment, they will likely have the same orientation. As a result, the orientation of the new token is determined by the frequency with which it appears alongside other tokens detected in a similar context. Turney and Littman’s approach for calculating mutual information can be used to calculate the frequency of co-occurrences of tokens.

A statistical approach is mostly used in several sentiment analysis applications. One such application is detecting manipulated reviews by running a statistical test of randomness popularly known as a training test. In work of Hu et al. (2012) expected that reviews written by customers would have random writing styles due to the random backgrounds of customers. They used a book review dataset from amazon.com to confirm their results but, it was found that close to 10.3 percent of products were subjected to online review manipulations.

LSA is another statistical technique for analyzing links between papers and tokens referenced in the documents to generate essential patterns connecting to the documents and phrases. The work of Cao et al. (2011) used LSA to fit semantic qualities from reviews to investigate the effect of various features. They engaged the program user feedback dataset from the CNETdownload.com website. Their main objective was to find out why few reviews received helpful votes while few reviews helpful votes. They determined various factors which may affect the helpful voting pattern for reviews.

Semantic Approach In this approach, the similarity score is calculated between tokens that are used for Sentiment Analysis. Wordnet is commonly used for this task. Antonyms and synonyms can be easily found using this approach as similar words have a positive score or higher value. Maks and Vossen proposed that the semantic approach can be used in various applications to build a lexicon model that can be used to describe adjectives, verbs, and nouns to use in Sentiment Analysis. They described, the in-depth description of subjectivity relations among the characters in a statement conveying distinct attitudes for each character. subjectivity tagged with the knowledge relating to both identity and orientation of attitude holder.

4.1.3. Dictionary-Based Based Method

Dictionary-based approach consists of a list of predefined set opinion words collected manually (Chetviorkin and Loukachevitch 2012; Kaity and Balakrishnan 2020). The primary assumption behind this approach is that synonyms have the same polarity as the base word, while antonyms have opposite polarity. Large corpora like thesaurus or wordnet are looked upon for antonyms and synonyms, after which it is appended to a group or seed list prepared earlier. In the first stage, an initial set of words is collected manually with their orientation. Later the list is expanded by looking at the antonyms and synonyms in the available lexical resources (Singh et al. 2017; Ho et al. 2014). Then the words are iteratively added to the list, and the list is expanded. Manual evaluation or correction may be done in the last stage to ensure its quality of it. Another disadvantage of all lexicon-based approaches (Hajek et al. 2020), including the dictionary-based approach, is finding opinion words specific for each domain as the polarity may vary.

Lexicon Method-Based Tools Summary Analysis of Lexicon Based method tools and available Dictionary as explained below:

Pre-define Dictionary Utilize a pre-defined list of positive and negative words to determine the polarity of texts based on the frequency with which each category is represented.

SentiWordNet assigns numerical sentiment scores to WordNet synsets that are either positive or negative.

Bing Liu's Sentiment Lexicon A dictionary has 4783 helpful positive and negative words.
SentiStrength Unless modified by any additional classification rules, texts are categorized according to the highest positive or negative score for any constituent word and highlight several characteristics of their subjectivity, such as the source (holder) of the subjectivity and terms included in phrases indicating positive or negative views.

National Taiwan University Sentiment Dictionary contains 2812 Positive and 8276 negative words.

WordNet-Affect is a WordNet Domains extension that includes a subset of synsets that are appropriate for representing affective notions associated with emotional words.

Affective Norms for English Words: Affective Norms for English Words are a collection of normative emotional ratings for a large number of English words. This collection of linguistic materials has been graded for enjoyment, arousal, and dominance to establish a baseline for future research on sentiment and attentiveness.

LingPipe can work on a wide range of activities, such as identifying topics, identifying named entities, parsing and indexing documents, database text mining, word segmentation, sentiment analysis, and language identification.

Apache OpenNLP provides support for parsing sentences, tokenization, part-of-speech tagging, segmentation, chunking, named entity extraction, language recognition, and coreference resolution.

Lexicon Sentiment Dictionary A language used in politics.

4.2. Machine Learning Approach

Machine Learning Algorithms can be used to categorize sentiments. Sentiment analysis is the process of identifying and quantifying the sentiment of text or audio using natural language processing, text analysis, computational linguistics, and other techniques. There are two primary in Machine Learning approaches to sentiment analysis:

- Supervised machine learning
- Lexicon-based unsupervised learning

This task can be accomplished using both supervised and unsupervised learning methodologies. Unsupervised strategies for sentiment analysis by utilizing knowledge bases, ontologies, databases, and lexicons that include detailed knowledge that has been selected and prepared specifically for sentiment analysis. Supervised learning methods are more commonly used due to their accurate results. These algorithms need to be trained on a training set before it is applied to the actual data. Features may be extracted from text data.
The machine learning technique utilizes syntactic and/or linguistic factors to address sentiment classification as a standard text classification issue utilizing syntactic and/or linguistic factors. The categorization model associates the underlying record's features with one of the class labels. The model is then used to predict a class label for a given instance of an unknown class. When an instance is assigned only one label, we have a difficult categorization challenge. When a probabilistic value of labels is assigned to an instance, this is referred to as the soft classification issue. Machine learning enables systems to acquire new abilities without being explicitly programmed to do so. Commonly used algorithms include:

4.2.1. Naive Bayes (NB)

NB technique is utilized for both categorization and training. NB is a Bayesian classification approach based on the theorem of Bayes. NB is a probabilistic classifier that uses the Bayes theorem to predict the probability of a given set of features as part of any particular label. The conditional probability that event A occurs given the individual probabilities of *A* and *B* and the conditional probability of occurrence of event *B*. Here it is assumed that features are not dependent. BoW model may be used for feature extraction. Generally, NB is applied when the training data size is small. NB was classified as positive 10% more accurately than negative classification. This led to a decrease in average accuracy when it was taken. In the work of Kang et al. (2012) solved this problem using an improved version of the NB classifier. They tested this model on to restaurant review dataset. In work of Tripathy et al. (2015) used machine learning for the classification of reviews. They proposed an NB model along with an SVM model (Hajek et al. 2020; Bordes et al. 2014). They used a movie review dataset for training and testing the models. Two thousand reviews were trained after pre-processing and vectorization of the training dataset. Count Vectorizer and *TF-IDF* were used before training the machine learning model. NB model proposed by Tripathy et al. (2015) gave an accuracy of 89.05 percent in a K-fold Cross-validation. The performance was better when compared to other models using the probabilistic NB algorithm (Calders and Verwer 2010).

4.2.2. Support Vector Machine (SVM)

The SVM approach, which uses hyper-planes, is used to analyze data and define decision limits in this technique. SVM is a type of non-probabilistic supervised learning technique that is frequently used for classification tasks. SVM's primary objective is to determine the hyperplane that best separates the data into distinct classes. As a result, SVM seeks out the hyperplane with the highest feasible margin. In the work of Li and Li (2013) used Support Vector Machines for sentiment polarity Classifier. Classifying reviews based on their quality is one of the many purposes for which SVM is utilized. Chen and Tseng (2011) used two multiple class SVM-based approaches. First is One-vs- all SVM and Multi-class SVM to classify reviews. Second, a method was proposed to evaluate the quality of the product review dataset quality by considering it a classification problem.

4.2.3. Logistic Regression (LR)

A machine learning technique known as logistic regression works by multiplying an input value by a weight value. It is a classifier that learns which input properties are most helpful in identifying positive and negative classes. Logistic regression is a probabilistic regression analysis used for classification tasks. For binary classification applications, logistic regression is commonly deployed. When there are multiple explanatory variables, logistic regression calculates the ratio of odds. Logistic regression uses Maximum-likelihood to calculate the best parameters. The independent variables may belong to any category i.e., Continuous, Discrete (ordinal and nominal). LR model (Hamdan et al. 2015) that the dependent variable is binary, and there is little or no multicollinearity between the predicting variables.

4.2.4. Decision Tree (DT)

DT Classifier is a supervised learning technique where a tree is built using the training example to classify the polarity of the text. DT uses a condition to divide data into parts recursively. RF is used more frequently than DT which combines multiple DTs to avoid over-fitting and improve accuracy. DT may be built using several algorithms like CART, ID3, C5.0, and C4.5 (Revathy and Lawrance 2017; Hssina et al. 2014; Singh and Gupta 2014; Patel and Prajapati 2018). These are used the identify the best-fitting attribute which needs to be placed in the root (Gower 1966; Revathy and Lawrance 2017; Patil et al. 2012). Yan- Yan et al. (2010)using a graph-based strategy, proposed a propagation strategy for integrating sentence-level and sentence-level features.

4.2.5. Maximum entropy (ME)

Conditional Exponential Classifiers: Conditional exponential classifiers encode labeled feature sets as vectors or arrays of integers. This vector is then used to compute feature weights, which can be used to select the most likely label for the feature set. Entropy is a measure of unpredictability. The Entropy is the maximum for uniformly distributed data. The input data consists of texts and ratings from *1-5* polarity assigned to it. The most popularly used algorithms include SVM, NB, and ME (Khairnar and Kinikar 2013; Kaufmann 2012) used ME Classifier to detect parallel sentences in any two-language pair, which have less training data. The other models used either required a massive amount of training data- set or used a language-specific technique (Bergsma et al. 2012), but their model showed improved results could be produced using any pair of languages. This will enable the establishment of parallel corpora for various languages.

4.2.6. K-Nearest Neighbors (KNN)

KNN algorithm is not extensively used in sentiment analysis but has been shown to produce good results when trained carefully. It operates on the fact that the classification of a test sample will be similar to nearby neighbors. The K value may be selected on any hyper-parameter tuning algorithms like Grid search or Randomized search cross-validation. The polarity may be hard-voted based on K nearest neighbors values, or soft addition may be done to find overall polarity.

4.2.7. Semi-Supervised Learning

In this case, where the training dataset contains both labeled and unlabeled data, semi-supervised learning appears to be a viable option (Zhu and Goldberg 2009). It is motivated that while gathering unlabeled data is relatively easy in many real-world applications, such as collecting articles from various blogs, labeling is expensive or labor-consuming because labeling the training dataset is typically performed by humans. Ortigosa-Hernán- dez et al. (2012) introduced the work of a real-world situation in which the user attitude is defined by three distinct (but related) target variables: subjectivity, sentiment polarity, and will to influence.

4.3. Hybrid Approach

The hybrid approach combines machine learning and lexicon-based approaches. Hybrid is a term that refers to the combination of machine learning and lexicon-based techniques for sentiment analysis. The hybrid technique combines the two and is extremely popular, with sentiment lexicons playing a significant role in the majority of systems. Sentiment analysis is a hybrid approach, including both statistical and knowledge-based methods for polarity recognition. In the work of Hassonah et al. (2020a) proposed a hybrid machine learning approach using SVM and two feature selection techniques using the multi-verse optimizer and Relief algorithms (Chang et al. 2020). The sentiment analysis task (Al Amrani et al. 2018) proposed using a machine learning-based hybrid approach including RF and SVM. They have shown that the individual models of SVM and RF had an accuracy of 81.01 and 82.03 percent, respectively, whereas the hybrid model combining both algorithms had an accuracy of close to *84%* in the product review dataset provided by *amazon.com.* Few researchers have proposed a hybrid architecture involving both lexicon-based and automated learning techniques to enhance the results. This is still a hot topic for researchers, and lots of research needs to be done.

4.4. Neural Network

Neural Network- In work of Van de Camp and Van den Bosch (2012) presented the use of Neural networks and SVM in supportive relationships. They used biographical texts to confirm their results. They were successful in marking relations between two individuals as neutral, positive, or negative. They concluded that an SVM and a single-layer Neural Network had shown improved results. In work of Moraes et al. (2013a) presented a comparative empirical analysis between SVM and ANN in document-level sentiment analysis. The motive of this comparison is that SVM was widely used as an algorithm for opinion mining as it had shown its capacity of getting accuracy. ANN, even though with good potential, did not have much attention. In Moraes et al. (2013b) they discussed all the aspects related to both ANN and SVM, including their requirements, their accuracies, and other contexts in which each model can perform the best. They have also implemented a consistent evaluation framework using well-known participants in supervised methods for selecting features and weights in orthodox BoW models.

The traditional RNN (Liu et al. 2016) were used for various NLP tasks as they used the previous time step information to predict the current time step, which ensures the usage of previous information and acts as memory as it remembers some information about a sequence. The most significant achievement or advantage of RNN was that it used previous information, thus remembering the previous information, which acted as memory. The main disadvantage of a traditional RNN is that it suffers from vanishing and exploding gradient descent, which means it cannot remember long-term relationships in the sequence.

In the case of Bi-LSTM (Plank et al. 2016) use the previous time step information along with the next time step information to predict the current time step, and pass the sequence in both ways forward as well as backward.

Deep learning has identified new avenues for emulating the peculiarly human potential, for example-based learning. While this method of bottom-up learning is successful for picture classification and object recognition, it is ineffective for NLP (Cambria et al. 2020). They blend top-down and bottom-up learning in their work using an array of symbolic and sub-symbolic AI tools and apply them to the intriguing challenge of text polarity detection.

Deep learning-based techniques are becoming highly popular due to their outstanding performance in recent times.

4.5. Other Approaches

4.5.1. Aspect-based based sentiment analysis (ABSA)

ASBA is a valuable and rapidly growing part of sentiment analysis that has gained prominence in recent years. Three critical phases compose aspect-level sentiment analysis: aspect detection, polarity or sentiment categorization, and aggregation. Aspect detection is a critical stage in Aspect-based Sentiment analysis, as it is followed by sentiment calculation. Aspects are mined either by using pre-defined implicit aspects or can be mined explicitly (Rana and Cheah 2016). Machine learning techniques, along with NLP techniques, are used to mine aspects out of a sentence.

Aspect-level sentiment analysis has many challenges as it to identify the individual aspect(implicit or explicit) and classifying as per sense is challenging to mine aspects (Tubishat et al. 2018), Therefore, complex algorithms like LSTM, Bi-LSTM or pre-trained models like BERT, GPT-2 may be used to accomplish the task. The researchers avoid vanilla RNN as it faces many problems like vanishing and exploding gradient descent. It is seen that recently attention-based models are being used in aspect detection. The next step after aspect detection is polarity assignment to those mined aspects.

4.5.2. Transfer Learning

Transfer learning is one of the advanced techniques in AI, where a pre-trained model can use its acquired knowledge to transfer to a new model. Transfer learning uses the similarity of data, distribution, and task. The new model directly uses the previously learned features without needing any explicit training data. Training data may be used to fine-tune the model for a new task. This technique can be used to transfer knowledge from one domain to another domain. This methodology has grown as a transfer learning technique because it can produce great accuracy and results while requiring significantly less training time than training a new model from scratch (Celik et al. 2020).

Transfer learning is frequently used in sentiment analysis to classify sentiments from one field to another field. In Meng et al. (2019) developed a multiple-layer CNN-based transfer learning approach. They used the weights and biases of a convolutional and pooling layer from a pre-trained model to the model. They used the features from a pre-trained model and fine-tuned weights of Fully connected layers. This approach can produce good results when large labeled data sets are absent and similarities in the tasks accomplished by the models.

In the work of Bartusiak et al. (2015), applied Transfer Learning to propose the sentiment analysis challenge. They used this technique to evaluate the sentiment at the document level in the Polish language. They used N-gram and Bi-gram to encode complex words and phrases. They used two different data- sets from two different domains to provide evidence that knowledge gained from the training model using the dataset of one domain can be used for a dataset of another domain.

4.5.3. Multimodal Sentiment Analysis (MSA)

MSA adds a new level to standard text-based sentiment analysis by incorporating additional modalities such as audio and visual data. Several studies have attempted to discern sentiment analysis in social multimedia using a variety of multimodal inputs, including visual, audio, and textual data (Soleymani et al. 2017). Social multimedia sites such as YouTube, video blogs (vlogs), or spoken evaluations contain expressions of sentiment, such as a video portraying a person discussing a product or a movie. Typically, spoken transcripts are examined separately from face and voice expressions, and the results of unimodal, text-based sentiment analysis are combined in post to create an "MSA" system. It may be bimodal, consisting of various combinations of two modalities, or trimodal, consisting of three modalities.

5. Performance Evaluation Parameter

The majority of state-of-the-art sentiment analysis makes use of accuracy, F1 score, and precision. Sentiment analysis using deep learning architectures: a review utilizes recall and accuracy as performance metrics. These metrics are as follows:

True Positive (TP): The number of positive reviews that have been correctly classified.
True Negative (TN): The number of negative reviews correctly classified as negative.
False Positive (FP): Number of incorrectly classified positive reviews.
False Negative (FN): Number of incorrectly classified negative reviews.

Precision is defined as the ratio of correctly classified positive samples to the total number of samples predicted as positive. This metric can be used to indicate the strength of the prediction. i.e., if a model has 100 percent precision, all the samples evaluated as positive are confidently positive.

The recall is also known as sensitivity. It is defined as the ratio of actual positive instances out of the total number of positive instances present in the classification. It measures the misclassifications done by the model. Precision and recall are inversely proportional to each other. Therefore, it is impossible to increase both Precision and Recall at the same time.
The F1 score is the harmonic mean of Recall and Precision. It is the most used metric after

Accuracy. It is used when we are unable to choose between Precision or Recall. F1 score manages the trade-off between recall and precision.

Accuracy This is the most commonly used metric in all the classification tasks. Accu- racy defines how accurate the model is. It is the ratio of correct classification to total predictions done by the model. Accuracy is a good metric to use for sentiment classification for a balanced dataset.

Specificity Is the opposite of sensitivity. It is not popularly used by researchers but is helpful in a few domains. It Is the ratio of the total number of correctly classified negative samples to negative classes present in the confusion matrix.

Confusion matrix A confusion matrix is a table that is frequently used to evaluate a categorization model's (or "classifier's") effectiveness on a set of training test data values are known. While the confusion matrix itself is rather straightforward to comprehend, the associated language might be perplexing.

6. Applications of Sentiment Analysis

Sentiment analysis has many applications, ranging from analyzing customer opinion, analyzing patient mental health status based on posts done on social media. Furthermore, technological advances such as Blockchain, IoT, Cloud Computing, and Big Data have broadened the range of applications for Sentiment Analysis, allowing it to be used in practically any discipline.

6.1. Business analysis

Sentiment analysis in the field of business intelligence offers several benefits. Additionally, firms can utilize sentiment analysis data to improve products, investigate client feedback and develop an innovative marketing strategy. The most typical use of sentiment analysis in the field of business intelligence is analyzing customers' impressions of services or products. These studies, however, are not limited to product producers; consumers may use them to review items and make more informed decisions. Sentiment analysis in business intelligence has various benefits.

6.1.1. Product Reviews

As the e-commerce business is burgeoning, so is the number of products sold and reviews given by the customers. Sentiment analysis will help customers choose better products (Paré 2003). Phrase level or aspect level (Schouten and Frasincar 2015) sentiment analysis performed on product reviews. Sentiment analysis can determine what the customer thinks about its latest product after launching or examining comments and reviews. Keywords for a specific product feature (food, service, cleanliness) can be chosen, and a sentiment analysis framework (Mackey et al. 2015) can be trained to identify and analyze only the necessary information.

6.1.2. Market research and competitor analysis

Market research is perhaps the most common sentiment analysis application, besides brand image monitoring and consumer opinion investigation. The purpose of sentiment analysis is to determine who is emerging among competitors and how marketing campaigns compare. It can be utilized to acquire a complete picture of brands and its competitor's consumer base from the ground up. Sentiment analysis may collect data from several platforms Twitter, Facebook, and blogs, deliver tangible results, and overcome difficulties in business intelligence.

6.2. Healthcare And Medical Domain

This is one of the industries where sentiment analysis is being utilized in recent times. Data can be collected from various sources like surveys, Twitter, blogs, news articles, reviews, etc. This data can then be analyzed for various use cases, one of them being an evaluation of standards and analysis of new updates in the medical field. In work of Clark et al. (2018) used Twitter tweets concerning patients' experiences as an add-on to analyze public health. Over a year, they generated roughly fi e million breast cancer-related tweets using Twitter's Streaming API. After pre-processing, the tweets were classified with a standard LR classifier and a CNN model. Positive treatment experiences, rallying support, and expanding public awareness were all linked. In conclusion, applying sentiment analysis to analyze patient-generated data on social media can help determine patients' needs and views.

6.2.1. Reputation Management

The application of sentiment analysis in diverse markets is brand monitoring and reputation management. Evaluating how customers view their brand, product, or service is a benefit to fashion companies, marketing agencies, IT companies, hotel chains, media channels, and other businesses. The sentiment analysis tool adds more variety and intelligence to the brand's and its product's portrayal. It enables businesses to track how their customers perceive their brands and highlight the precise data about their attitudes. Look for trends and changes, and pay attention to influence presentations. Altogether, sentiment analysis can be utilized in automating the media surveillance system as well as the alarm system that goes with it. Keep track of the brand's discussions and ratings on various social media platforms.

7. Challenges in Sentiment Analysis

Sentiment analysis comes with various challenges ranging from computational cost to informal writing and the presence of variations in languages. A few significant challenges faced in sentiment analysis are:

7.1. Structured Sentiments

Structured sentiments are found in formal sentiment reviews, they are more focused on formal problems such as books or research. Because the authors are professionals, they are capable of writing thoughts or observations concerning scientific or factual concerns.

7.2. Semi-Structured Sentiments

Semi-Structured Sentiments fall between structured and unstructured sentiments. These require an awareness of numerous review-related concerns.

7.3. Unstructured Sentiment

Unstructured Sentiment is an informal and free-flowing writing type in which the writer is not constrained by any rules (Mukherjee et al. 2013). The text may comprise multiple sentences, each of which could potentially include both pros and cons. For example, unstructured reviews off more opinion information than their formal counterparts (Levashina et al. 2014). A feature explicitly stated: If a feature occurs in a review sentence’s segment/chunk, the feature is referred to as an explicit feature of the product.

7.4. Methodological Challenges

The majority of sentiment analysis in the modern day is data-driven machine learning models adapting a sentiment analysis algorithm developed for product evaluations to evaluate microblog postings is an unanswered question. Additionally, how to deal with ambiguous situations and irony are key differences in sentiment analysis. For instance, a sarcastic remark about an object is intended to communicate a negative sentiment; yet, conventional sentiment analysis algorithms frequently miss this meaning. Numerous methods have been proposed (Castro et al. 2019; Medhat et al. 2014) for detecting sarcasm in language. However, the problem is far from resolved, as comedy is very culturally particular, and it is challenging for a machine to understand unique(and frequently fairly detailed) cultural allusions. In the work of Poria et al. (2018a) suggests incorporating vocal and facial expressions into multimodal sentiment analysis; Which can improve its success rate in identifying sarcastic comments. Furthermore, individuals express sentiment for social reasons unrelated to their fundamental dispositions.
Sentiment analysis applies to different types of data, each of which presents particular challenges. Sentiment analysis of human-to-machine and human-to-human interactions requires very similar datasets to those used for emotion recognition. As a result, it has the same limitations in terms of size and unreliable ground truth. Additionally, there is the issue of labeling confidential laboratory data, which prohibits those permitted to examine the data from performing the time-consuming operation of labeling. As a result, they are restricted in terms of the amount of data they can collect in the laboratory and our ability to label huge volumes of data. There are several methods for assessing feelings, but word embedding algorithms such as word2vec and GloVe turn words into meaningful vectors.

Multimedia information on websites is the second source of multi-modal sentiment data. Social media provides us with a wealth of data that helps us to scale. The issue is that the data acquired vary in terms of quality and context, and the data is limited to specific populations that are more prevalent on the internet. However, because the data is publicly available, crowdsourcing may be utilized to categorize it easily. According to the available data on MSA, people are more prone to communicate positive or negative ideas online, resulting in a scarcity of neutral opinions represented in all MSA studies evaluated.

Sarcasm People tend to use sarcasm when they do not meet their expectations. It is very tough for machines to pick up sarcasm as many factors affect sarcasm, such as tone, situation, background information, etc. Sarcasm is a satirical remark that may look like praising but in reality. Sarcasm is used by people to criticize. Sarcasm is a type of sentiment in which people express implicit information, usually the polar opposite of the message content, to emotionally hurt someone or mock something. Sarcasm detection in text mining is one of the most challenging tasks in NLP, but it has lately become an interesting research subject due to its usefulness in enhancing social media sentiment analysis (Eke et al. 2020). *The informal style of writing:* Informal style of writing is the biggest challenge to all NLP tasks, including sentiment analysis. People are very casual about writing reviews or texts; they tend to use acronyms, emojis, and shortcuts in their text which is very hard to pick up. Acronyms can be handled if they are universal. There are a lot of regional acronyms which change and grow day by day.

Grammatical errors: Grammatical errors are very common in informal texts and can be handled, but only to some extent; spelling errors can also be corrected limited. It is very difficult to burgeoning the spelling mistake of users uniquely every time. The accuracy of sentiment analysis and NLP tasks may be improved if these errors can be handled and corrected.

Computational cost To get better accuracy, we need to increase the training data size and complicate the model, which will exponentially increase the computational cost of the model for training; a high-end GPU may be required to train a model with a huge corpus. Models like SVM, and NB are not computationally costly, but neural networks and attention models have shown that they are computationally costly.

Availability of data As NLP and sentiment analysis is a recently boomed technology, the Availability of data may also be a challenge in some cases. Although data is available on Twitter for sentiment analysis, high-quality training data is challenging for supervised learning algorithms. Training data for ABSA is challenging to find online therefore needs to be prepared manually. The training data of one domain may not be applicable and valuable to other domains. For instance, a model trained on a hotel review dataset does not help predict the sentiments of a stock or mutual fund dataset and vice versa.

Adaptations of language Languages change as they move to different regions and places; although the base language remains the same, many factors influence language, such as language prominence, pronunciation, literacy rate, etc. Similarly, different spellings for the same word, such as *"color"* and *"color,"* mean the same but are spelled differently in different regions. This will create duplicates and may affect the accuracy and computational cost of the model. The language barrier is the hardest of the challenges for NLP. There are thousands of languages spoken worldwide, although NLP techniques are hardly available in *5-10* languages, and resources are widely available for English.

Phrases containing degree adverbs and intensifiers Adverbs such as slightly, barely, and moderately are used to quantify the sentiments. For instance, consider review r1= "The food is barely good" and r2= "the food is good". *r*1 is considered neutral or slightly positive, whereas *r*2 is considered to be highly positive. The adverbs *'barely'* and *'really'* decide the extent of positiveness and the word *'good'.* Similarly, intensifiers also quantify the sentiment of the sentences. Intensifiers like very, too are used to increase the positiveness or negative ness of the token. For instance, *"too good"* is considered to be more positive than *"good."* Intensifiers and degree adverbs impose a challenge on aggregating the sentiment values and comparing two sentences of the same sentiment rather than differentiating between two sentences of opposite polarity.

Mixed Code Data Code-mixing is the employment of vocabulary and grammar from different languages in the same sentence Code Mixing is a linguistic phenomenon that can occur in a multilingual situation where speakers speak multiple languages. This phenomenon is becoming increasingly common as communication between groups of people who speak different languages grows. Code-Mixing: A review of Facebook posts created by Hindi-English users revealed a high code-mixing level in the posts. The problems in the Hindi-English code-mixed text were reported using a PoS tag annotated corpus. It's frequent in multilingual societies and presents considerable difficulty to NLP tasks like sentiment analysis. The lack of formal grammar for code-mixed phrases makes it challenging to identify compositional semantics, which is critical for conducting sentiment analysis using rule-based and machine learning-based techniques. Furthermore, because mixing is up to the individual, there are no predetermined mixing guidelines, which is one of the significant drawbacks.

8. Conclusion

This article discussed sentiment analysis and associated techniques. The primary objective of this work is to investigate and complete classification methods with their advantage and disadvantages in sentiment analysis. To begin, several levels of sentiment analysis were discussed, followed by a quick overview of necessary procedures such as data collection and feature selection. Next, methods of sentiment categorization systems were classified and compared in terms of their advantages and disadvantages. Due to their simplicity and excellent accuracy, supervised machine learning methods are often the widely utilized technique in this discipline. Classification using NB and SVM algorithms are commonly used as benchmarks against which newly proposed approaches can be compared. Several of the most common application areas are discussed then the survey examines the significance and consequences of sentiment analysis challenges in sentiment evaluation. The comparison investigates the relationship between the structure of sentiment reviews and the difficulties associated with sentiment analysis. This comparison reveals domain dependence, which is essential for identifying sentiment issues. The future work will consist of continuously expanding the comparison area with additional findings. The subsequent challenges illustrate that sentiment analysis is still a relatively unexplored subject of study.

References

[1] Piryani R, Madhavi D, Singh VK (2017) Analytical mapping of opinion mining and sentiment analysis research during 2000–2015. Inf Process Manag 53(1):122–150

[2] Yue L, Chen W, Li X, Zuo W, Yin M (2019) A survey of sentiment analysis in social media. Knowl Inf Syst 60(2):617–663

[3] Hu N, Bose I, Koh NS, Liu L (2012) Manipulation of online reviews: an analysis of ratings, readability, and sentiments. Decis Support Syst 52(3):674–684

[4] Bhatia P, Ji Y, Eisenstein J (2015) Better document-level sentiment analysis from rst discourse parsing arXiv preprint arXiv: 15090 1599

[5] Han T, Liu C, Yang W, Jiang D (2019) A novel adversarial learning framework in deep convolutional neural network for intelligent diagnosis of mechanical faults. Knowl-Based Syst 165:474–487

[6] Behdenna S, Barigou F, Belalem G (2018) Document level sentiment analysis: a survey. EAI Endorsed Trans Context-aware Syst Appl 4(13):e2

[7] Thet TT, Na JC, Khoo CS (2010) Aspect-based sentiment analysis of movie reviews on discussion boards. J Inf Sci 36(6):823–848

[8] Schouten K, Frasincar F (2015) Survey on aspect-level sentiment analysis. IEEE Trans Knowl Data Eng 28(3):813–830

[9] Razon A, Barnden J (2015) A new approach to automated text readability classification based on concept indexing with integrated part-of-speech n-gram features. In: Proceedings of the international conference recent advances in natural language processing, pp 521–528

[10] Tian Y, Galery T, Dulcinati G, Molimpakis E, Sun C (2017) Facebook sentiment: reactions and emojis. In: Proceedings of the fifth international workshop on natural language processing for social media, pp 11–16

[11] Sharma A, Lyons J, Dehzangi A, Paliwal KK (2013) A feature extraction technique using bi-gram probabilities of position-specific scoring matrix for protein fold recognition. J Theor Biol 320:41–46

[12] George DR, Rovniak LS, Kraschnewski JL (2013) Dangers and opportunities for social media in medicine. Clin Obstet Gynecol 56(3)

[13] Mikolov T, Chen K, Corrado G, Dean J (2013) Efficient estimation of word representations in vector space. arXiv preprint arXiv: 13013 781

[14] Bojanowski P, Grave E, Joulin A, Mikolov T (2017) Enriching word vectors with subword information. Trans Assoc Comput Linguist 5:135–146

[15] Poria S, Cambria E, Winterstein G, Huang GB (2014) Sentic patterns: dependency-based rules for concept-level sentiment analysis. Knowl-Based Syst 69:45–63

[16] Adomavicius G, Kwon Y (2011) Improving aggregate recommendation diversity using ranking-based techniques. IEEE Trans Knowl Data Eng 24(5):896–911

[17] Wang G, Sun J, Ma J, Xu K, Gu J (2014) Sentiment classification: the contribution of ensemble learning. Decis Support Syst 57:77–93

[18] Xia Y, Cambria E, Hussain A, Zhao H (2015) Word polarity disambiguation using Bayesian model and

[19] opinion-level features. Cognit Comput 7(3):369–380

[20] Cambria E, Das D, Bandyopadhyay S, Feraco A (2017) Affective computing and sentiment analysis. In: A practical guide to sentiment analysis. Springer, pp 1–10

[21] Bai X, Liu P, Zhang Y (2020) Investigating typed syntactic dependencies for targeted sentiment classification using graph attention neural network. IEEE/ACM Trans Audio Speech Lang Process 29:503–514

[22] Yan-Yan Z, Bing Q, Ting L (2010) Integrating intra-and inter-document evidences for improving sentence sentiment classification. Acta Autom Sinica 36(10):1417–1425

[23] Hu N, Bose I, Koh NS, Liu L (2012) Manipulation of online reviews: an analysis of ratings, readability, and sentiments. Decis Support Syst 52(3):674–684

[24] Cao Q, Duan W, Gan Q (2011) Exploring determinants of voting for the “helpfulness’’ of online user reviews: a text mining approach. Decis Support Syst 50(2):511–521

[25] Cao Q, Duan W, Gan Q (2011) Exploring determinants of voting for the “helpfulness’’ of online user reviews: a text mining approach. Decis Support Syst 50(2):511–521

[26] Cao Y, Zhang P, Xiong A (2015) Sentiment analysis based on expanded aspect and polarity-ambiguous word lexicon. Int J Adv Comput Sci Appl 6(2):97–103

[27] Carvalho J, Plastino A (2021) On the evaluation and combination of state-of-the-art features in twitter sentiment analysis. Artif Intell Rev 54(3):1887–1936

[28] Castro S, Hazarika D, Pérez-Rosas V, Zimmermann R, Mihalcea R, Poria S (2019) Towards multimodal sarcasm detection (an _obviously_ perfect paper). arXiv preprint arXiv: 19060 1815

[29] Celik Y, Talo M, Yildirim O, Karabatak M, Acharya UR (2020) Automated invasive ductal carcinoma detection based using deep transfer learning with whole-slide images. Pattern Recogn Lett 133:232–239

[30] Chang JR, Liang HY, Chen LS, Chang CW (2020) Novel feature selection approaches for improving the performance of sentiment classification. J Ambient Intell Humaniz Comput pp 1–14

[31] Chatterjere A, Guptha V, Chopra P, Das A (2020) Minority positive sampling for switching points-an anecdote for the code-mixing language modeling. In: Proceedings of the 12th language resources and evaluation conference, pp 6228–6236

[32] Chen CC, Tseng YD (2011) Quality evaluation of product reviews using an information quality framework. Decis Support Syst 0(4):755–768

[33] Chen X, Wang Y, Liu Q (2017) Visual and textual sentiment analysis using deep fusion convolutional neural networks. In: 2017 IEEE international conference on image processing (ICIP). IEEE, pp 1557–1561

[34] Hershcovich D, Donatelli L (2021) It's the meaning that counts: the state of the art in NLP and semantics.KI-Künstliche Intelligenz pp 1–16

[35] Ho C, Murad MAA, Doraisamy S, Kadir RA (2014) Extracting lexical and phrasal paraphrases: a review of the literature. Artif Intell Rev 42(4):851–894

[36] Hssina B, Merbouha A, Ezzikouri H, Erritali M (2014) A comparative study of decision tree id3 and c4. 5. Int J Adv Comput Sci Appl 4(2):13–19

[37] Hu N, Bose I, Koh NS, Liu L (2012) Manipulation of online reviews: an analysis of ratings, readability, and sentiments. Decis Support Syst 52(3):674–684.

[38] Hu X, Tang J, Gao H, Liu H (2014) Social spammer detection with sentiment information. In: 2014 IEEE international conference on data mining. IEEE, pp 180–189.

[39] Jain PK, Yekun EA, Pamula R, Srivastava G (2021) Consumer recommendation prediction in online reviews using cuckoo optimized machine learning models. Comput Electr Eng 95:107397

[40] Janjua F, Masood A, Abbas H, Rashid I, Khan MMZM (2021) Textual analysis of traitor-based dataset through semi-supervised machine learning. Futur Gener Comput Syst 125:652–660

[41] Jiménez-Zafra SM, Martín-Valdivia MT, Molina-González MD, Urena-López LA (2019) How do we talk about doctors and drugs? sentiment analysis in forums expressing opinions for the medical domain. Artif Intell Med 93:50–57

[42] Juraska J, Walker M (2021) Attention is indeed all you need: semantically attention-guided decoding for data-to-text nlg. arXiv preprint arXiv: 21090 7043

[43] Kaity M, Balakrishnan V (2020) Sentiment lexicons and non-English languages: a survey. Knowl Inf Syst 1–36

[44] Kamal A (2013) Subjectivity classification using machine learning techniques for mining feature-opinion pairs from web opinion sources. arXiv preprint arXiv: 13126 962

[45] Kanapala A, Pal S, Pamula R (2019) Text summarization from legal documents: a survey. Artif Intell Rev 51(3):371–402

[46] Kang H, Yoo SJ, Han D (2012) Senti-lexicon and improved Naive Bayes algorithms for sentiment analysis of restaurant reviews. Expert Syst Appl 39(5):6000–6010.

[47] Mukherjee A, Venkataraman V, Liu B, Glance N (2013) What yelp fake review filter might be doing? In: Proceedings of the international AAAI conference on web and social media, vol 7

[48] Naseem U, Razzak I, Musial K, Imran M (2020) Transformer based deep intelligent contextual embedding for Twitter sentiment analysis. Futur Gener Comput Syst 113:58–69

[49] Ortigosa-Hernández J, Rodríguez JD, Alzate L, Lucania M, Inza I, Lozano JA (2012) Approaching sentiment analysis by using semi-supervised learning of multi-dimensional classifiers. Neurocomputing 92:98–115

[50] Oueslati O, Cambria E, HajHmida MB, Ounelli H (2020) A review of sentiment analysis research in Arabic language. Futur Gener Comput Syst 112:408–430

[51] Paré DJ (2003) Does this site deliver? B2B e-commerce services for developing countries. Inf Soc 19(2):123–134

[52] Park HW, Park S, Chong M (2020) Conversations and medical news frames on Twitter: infodemiological