Université Paris 1 Panthéon Sorbonne

Exploring Emotions through Sentiment Analysis on the Twitter Social Network

A Systematic Literature Review

Mémoire présenté par :
JOTI Anxhela

Pour l'obtention d'un master MIAGE parcours
Système d'Information et Innovation (S2I)

Tuteur enseignant : KORNYSHOVA Elena
Maître d'apprentissage : DESDOUITS Nathan
Membre du jury : DENECKÈRE Rébecca
Année : 2022/2023

Acknowledgments

As I go back in time, I'm reflecting on the people who supported me throughout this journey, realizing that they are indeed a considerable number. I extend my heartfelt gratitude to the entire academic staff of Panthéon-Sorbonne for their exceptional efforts and incredible dedication.

I express sincere appreciation to my thesis supervisor, Madame Elena Kornyshova, for her constant availability and devotion. I can proudly attribute the quality of my work to her precise guidance and the challenges she encouraged me to overcome.

A special thanks is owed to the Tinyclues and Splio teams, who provided an environment during my apprenticeship where I could showcase my core values and grow both professionally and personally. Nathan Desdouits, in particular, has not only been my manager but also a true mentor, guiding me through all professional challenges. Today, I express my gratitude for the trust he placed in me and for his significant role in my professional development over the past year.

I am grateful to have shared this experience with my fellow colleagues from the university, who made this journey all the more fascinating and adventurous.

The achievement of this milestone is dedicated to my cherished family, who have always been my cheerleaders, trusting and supporting me, never landing me on earth when I'm dreaming of the stars.

Thank you to everyone who stood by me! As I journey back in time, I feel compelled to express this louder and in my native language: *Ia dola!*

Table of contents

Abstract

This paper presents a comprehensive systematic review of 60 papers published between 2018 and 2022, investigating the applications and advancements in Sentiment Analysis on Twitter. The study aims to offer valuable insights to researchers by analyzing evolving trends, applications, and methodologies in Sentiment Analysis during this period. The analysis revealed a surge of interest in SA, peaking in 2021, particularly focused on health domains, notably during the pandemics, and mental health. Machine learning and deep learning methodologies gained preference over traditional lexicon-based approaches due to their ability to handle complexity and capture nuanced sentiments. The study highlights the need for further research to analyze the new methodologies proposed in the realm of SA aiming to enhance the SA performance and accuracy.

Glossary

SA	Sentiment Analysis
AT	Accuracy Test
QB	Questionnaire-based
ML	Machine Learning
DL	Deep Learning
NB	Naive Bayes
LR	Linear Regression
SVM	Support Vector Machine
RF	Random Forest
DT	Decision Tree
LSTM	Long Short-Term Memory
CNN	Convolutional Neural Network
GCN	Graph Convolutional Network
SVC	Support Vector Classifier
NN	Neural Network
KNN	K-Nearest Neighbor
DCNN	Deep Convolutional Neural Networks

1. Introduction

1.2. Context

In today's digitally interconnected world, social media platforms have become an integral part of people's daily lives. Among these platforms, Twitter has gained immense popularity as a microblogging platform where users share their thoughts, opinions, and experiences in real time. According to [1], Twitter reached 396.5 million users in April 2023, where 206 million among them are active.

The rise of Online Social Media (OSM) platforms has revolutionized the way people communicate and express themselves online. These platforms provide an unique opportunity to share one’s thoughts, feelings, and opinions on a wide range of topics that capture the public’s attention by enabling communication through text, image, audio, and video.

It has therefore been a key area of interest, analyzing human emotions through the data retrieved from different OSMs. Understanding and interpreting emotions from textual and unstructured data has preserved significant attention among scientists and linguistics, leading to a new field of Natural Language Processing (NLP), known as Sentiment Analysis.

This field has gained a lot of interest in the last 10 years due to its potential applications across various domains and found a valuable application in the realm of business and marketing [2,3], healthcare [4-8], politics [75, 76] , social issues [10, 11], traumatic events [9, 12], etc.

Given the exponential growth of data on Twitter and the increasing interest in understanding human emotions at scale, the exploration of emotions through sentiment analysis on Twitter seems significantly relevant. This doesn't just offer us insights into individuals’ feelings but also lets us learn about the overall feelings of the groups and what they might mean in different areas.

1.3. The research problem

The Sentiment Analysis on Twitter has seen notable growth from 2018 until now [fig.2]. There have been significant contributions in the literature for the following purposes:

1. Applying Sentiment Analysis in a specific domain as well as trying to get answers related to a real-life problem.
2. Finding the issues with the existing methodology and/or algorithms and proposing an enhanced method.
3. Comparing the existing methodology and/or algorithms to find the one with the highest performance.

The list of contribution types is surely not exhaustive, nor mutually exclusive. There exists a lot of papers in the literature that try to tackle more than one problem (more elaborated in 4. Related work). For example, they strive to propose a new methodology of Sentiment Analysis while analyzing a dataset relevant to an existing issue. Meanwhile, some other papers apply Sentiment Analysis in a specific domain while comparing the existing methodologies through several tests to find out the one with the higher performance.

What the literature lacks until now is a detailed synthesis of all the available methodologies and algorithms used for Sentiment Analysis on Twitter, in order to have a clear understanding of the contributions that have been made in the last few years. For that reason, this thesis tries to tackle the following research problem:

What approaches have been used for Sentiment Analysis on Twitter over the last years?

To tackle this problem, this thesis aims to develop a comprehensive framework that provides a structured approach for analyzing and evaluating articles and the methodologies used in the context of sentiment analysis on the Twitter social network. The framework aims to assess the methodology and contributions of articles in this domain by defining relevant criteria that can be applied to all the articles. By creating this framework, the thesis seeks to enable researchers and practitioners to gain a deeper understanding of the field and identify gaps for further research.

2. Theoretical background

2.1 Twitter

Twitter is a free social networking site where users broadcast short posts known as tweets. These tweets can contain text, videos, photos, or links [20]. Users can connect on their account whether by using the Twitter application or on Twitter's website www.twitter.com.
According to [1], Twitter reached 396.5 million users in April 2023, where 206 million among them are active.

Twitter, a widely recognized social media platform, is commonly referred to as a microblogging platform. The term "microblogging" emerged in the early 2000s and has since become a prevalent concept in research studies related to social media and communication technologies.
[21] tries to explore how Twitter is used by scholars in academic conferences and its potential benefits for enhancing collaboration and knowledge co-construction. In their study Ross et al. define microblogging as below:

> "Microblogging is a variant of blogging which allows users to quickly post short updates, providing an innovative communication method that can be seen as a hybrid of blogging, instant messaging, social networking, and status notifications."

Because of these features, Twitter is known as a platform where people may express themselves spontaneously. This is the main reason why Twitter has been an interesting medium among scientists, particularly in understanding mass opinion.
Twitter is also widely used by celebrities, including artists, athletes, politicians, and even religious figures.
The platform provides an API that allows third parties to retrieve information from Twitter as well as to take action in their account.
According to the official webpage of Twitter API, here are some of its capabilities [21]:

- Fetching & posting tweets: Developers can utilize the API to fetch tweets from Twitter. It is also possible to filter by posting time, account, and particular keywords or hashtags. They can post tweets as well.

- Fetching user information: Twitter API allows the developers to access user-related information, such as profiles, followers, friends, etc.

- Manage direct messages: Twitter API allows creating Direct Messages and listing Direct Message conversation events.

- Account management: It is also possible to manage Twitter accounts, such as following or unfollowing users, updating user profiles, and handling user settings.

- Accessing trends and metadata: The API provides access to trending topics, locations, and other metadata related to Twitter content.

In 2022 Twitter was acquired by Elon Musk, the CEO of Tesla Motors. In late January he started to invest in the company and by mid March became the biggest shareholder with 9.2% of the shares [112].

Figure 1. X (ex. Twitter) new logo

On 27th of October 2022 he succeeded to conclude the deal and acquire Twitter [112]. It has been reported that during the following months Twitter has been under a lot of structural changes and layoffs causing a lot of noise in the industry. The company went through a rebranding phase

leading to its name changing to 'X' and logo as presented to Fig. 1. [112]. Twitter is now known as X, and this is also reflected on the Twitter application.
On the 30th of March 2023, Twitter announced the new fees for the tiers using its API.
There are four subscription tiers available for Twitter API access, according to their official website https://developer.twitter.com/en [22]:

- Free: Provides write-only access, enabling users to post up to 1,500 tweets per month at no cost.

- Basic: A $100 per month subscription suitable for hobbyists or prototypes, offering the option to post 3,000 tweets per month at the user level or 50,000 tweets per month at the application level. The read limit is set at 10,000 tweets.

- Pro: A $5,000 per month subscription providing the start-ups the option to read up to 1,000,000 tweets and to post not more than 300,000 tweets at an application level.

- Enterprise: Offers commercial-level access tailored to specific needs, with managed services provided by a dedicated account team. While the exact price was not specified, it was reported that a "low-cost enterprise plan" could cost up to $42,000 per month.

This raised financial discomfort within the enterprises and the developers' community, which didn't hesitate to express their revolt [23], [24], [25]. Moreover, academics all over the world are invited by Oxford Internet Institute Researchers to sign the appeal, advocating for unrestricted access to Twitter data and facilitating future research [26].

2.2. Sentiment Analysis

Before having a clear definition of Sentiment Analysis, it is crucial to have a better comprehension of the concept of *Sentiment* and how it differs from its synonyms:

In [12] it is highlighted the definition of several synonyms of *sentiment*, provided by Pang and Lee [13] in order to clear any ambiguity:

- Opinion implies a conclusion thought out yet open to dispute (“each expert seemed to have a different opinion”).

- View suggests a subjective opinion (“very assertive in stating his views”).

- Belief often implies deliberate acceptance and intellectual assent (“a firm belief in her party's platform”).

- Conviction applies to a party’s firmly and seriously held belief (“the conviction that animal life is as sacred as human”).

- Persuasion suggests a belief grounded on assurance (as by evidence) of its truth (“was of the persuasion that everything changes”).

- Sentiment suggests a settled opinion reflective of one's feelings (“her feminist sentiments are well-known”).

[12] uses *sentiment* and *opinion* interchangeably. In the research conducted by [13], it is stated that:

> "The body of work we review is that which deals with the computational treatment of (in alphabetical order) *opinion*, *sentiment*, and *subjectivity* in text. Such work has come to be known as *opinion mining*, *sentiment analysis*, and/or *subjectivity analysis*."

That implies that opinion, sentiment, and subjectivity are used interchangeably.

Remarkably, the notion of *emotion* is not in the synonym list defined by Pang and Lee (2008), nor is it used to express any research question in [13].

The book published by B Liu (2012) [14] underlines a considered number of definitions between ambiguous or synonym words.

[14] states that *subjectivity* and *emotion* are two important concepts that are closely related to *sentiment* and *opinion*. It is relevant to say that he never uses emotion as a synonym of sentiment, but more as a related concept of it.

Meanwhile, the definition of emotion, according to [14], is as follows:

> *Emotions* are our subjective feelings and thoughts that are closely related to *sentiments*. The strength of *sentiment* or *opinion* is typically linked to the intensity of certain *emotions*, e.g., joy and anger. *Opinions* that we study in sentiment analysis are mostly evaluations (although not always). [14]

In the years when [13][14] were conducted (2008-2012), the Sentiment Analysis was more related to public opinion expressed through reviews or personal blogs. With the rise of Online Social Media (OSM), another purpose of Sentiment Analysis was to study public *emotions*. Even though there are papers that use Sentiment Analysis and Emotion Analysis interchangeably, [15] highlights a difference between them. According to [15] Sentiment Analysis is a means of assessing if data is *positive*, *negative*, or *neutral*. In contrast, Emotion detection is a means of identifying distinct human emotion types such as *furious*, *cheerful*, or *depressed*. "Emotion detection," "affective computing," "emotion analysis," and "emotion identification" are all phrases that are sometimes used interchangeably.
[17] applies the same concept when making the difference between the emotional lexicon (anger, fear, anticipation, trust, surprise, sadness, joy, and disgust) and the sentiment lexicon (positive, negative).
Meanwhile, some other papers do not distinguish these two concepts. Emotion Analysis, according to [16], is a field of study based on the evaluation of the *sentiments*, *opinions*, and *emotions* of people regarding certain entities, events, subjects, and institutions.

Sentiment Analysis has found a valuable application in the realm of business and marketing [2,3], by providing notable information on customer opinion and experience, brand perception, and/or product feedback. Moreover, a high interest was found in healthcare [4-8], especially during the pandemic where a lot of scientific articles strived to understand the public opinion regarding the situation and the mindset that they were on. In politics and public opinion research, sentiment analysis on Twitter offers a unique opportunity to capture and analyze public sentiment towards political candidates [76], policies, or social issues [9]. Additionally, sentiment analysis on Twitter has found applications in areas such as customer support [2], public health monitoring, and stock market prediction [35], among others.

3. Research methodology

3.1. An overview of Systematic Literature Review

A Systematic literature review (SLR) is a systematic way of collecting, critically evaluating, integrating, and presenting findings from across multiple research studies on a research question or topic of interest. It is “systematic” since it adopts a consistent, widely accepted, methodology [18]. As scientific inquiries, literature reviews should be valid, reliable, and repeatable [19].

In general, literature reviews can be categorized into two main types: (1) a review that provides the background for an empirical study, and (2) a standalone piece [18].

In this thesis, the process of conducting a Systematic Literature Review (SLR) is outlined, drawing upon the key steps highlighted in [18]. The left side of Table 1. below presents these essential steps, while the right side demonstrates their specific application within this study.

Key Steps [18]	The application in the thesis
Formulate a study question that is precise and free of ambiguity.	1.3. The research problem
Clearly define each term in your study question and the identified concepts of interest.	2. Theoretical background

Determine the concepts that need to be considered in your literature search.	The search query in 3.2. Research Questions
Choose databases that are relevant to your topic to conduct your search.	MIAGE Scholar, mentioned in 3.2. Research Questions
Establish inclusion and exclusion criteria for selecting articles.	Inclusion and Exclusion criteria in 3.4. Selection criteria
Review the abstracts of the identified articles and screen them based on the exclusion criteria.	
Conduct a critical review of the full articles that meet the inclusion criteria.	6. Results Appendix 1

Table 1. Key steps of developing an SLR, according to [18]

3.2. Research questions

This paper aims to develop a systematic framework for analyzing and structuring the existing papers in the field of emotion analysis through Sentiment Analysis on Twitter. This framework will consist of a set of well-defined criteria that can be used to assess the methodology, the algorithm, the evaluation method, and the contributions of the articles under review. By employing a rigorous and systematic approach, the methodology aims to provide a reliable and comprehensive evaluation of the literature, facilitating a deeper understanding of the field and identifying areas for further research.
To reach this objective our general research question would be:

How do we analyze emotions on Twitter?

In order to break down this question and to have a structured response, we have set some detailed research questions relevant to our study, which are presented as follows:

RQ 1 - What terminology is used for expressing emotion on Twitter?

RQ 2 - In which domains is Sentiment Analysis on Twitter applied?

RQ 3 - How is the Sentiment Analysis on Twitter applied?

Research sub-question 3.1 - What algorithms/lexicon are used for Sentiment Analysis?

Research sub-question 3.2 - What are the steps of the Sentiment Analysis process?

RQ 4 - How is the methodology for Sentiment Analysis on Twitter evaluated?

3.3. Research query

To gather relevant articles for this study, a literature search was conducted using the MIAGE Scholar website.

In order to have a relevant query we followed the steps below:

1. Definition of Key Keywords:
    - Sentiment Analysis: This term represents the primary focus of this study.
    - Emotion (with a focus on public emotions): The intention was to include papers that specifically addressed the study of *emotions* within the public domain.
    - Twitter: As the scope of the study is focused on sentiment analysis within the context of social media, only Twitter was taken under analysis.
2. Incorporation of Synonyms:
    - Opinion Mining: Recognizing that the terms *Sentiment Analysis* and *Opinion Mining* are often used interchangeably, this synonym was included to expand the search coverage.
3. Determination of Subject Area:
    - Computer Science: Considering that Sentiment Analysis is a multidisciplinary field encompassing linguistics, psychology, and computer science, the research scope was defined within the domain of Computer Science to provide a targeted search.
4. Specification of the Publication Year Range:

- 2018 to 2022: While sentiment analysis has been a subject of research for several years, the period between 2018 and 2022 witnessed a notable increase in studies focusing on sentiment analysis on Twitter. Consequently, to capture the latest advancements and trends, the search was limited to articles published within this specific timeframe.

Considering the criteria mentioned above, we constructed the query as follows:

(TITLE-ABS-KEY("opinion mining") OR TITLE-ABS-KEY("sentiment analysis")) AND TITLE-ABS-KEY(emotion) AND TITLE-ABS-KEY(Twitter) AND SUBJAREA("COMP") AND PUBYEAR > 2017 AND PUBYEAR < 2023

3.4. Selection criteria

By running the query, the result was 736 articles. The whole selection process was carried out in two big phases.

1- *The first one*, as suggested by [18], consists of reading the title and the abstract of the article only. During the initial phase, the primary goal was to collect all the **potential** pertinent papers by using essential criteria.
Below are presented the inclusion and exclusion criteria used in order to carry out a first selection of the bibliography:

Inclusion criteria:

- The paper is primary research.
- The paper studies emotional reactions or public opinions.
- The dataset used contains English tweets only.
- The paper studies text-only tweets. (Emojis, hashtags, images, and videos are out of the scope of this study.)

Exclusion criteria:

- The paper is secondary research.
- The paper studies tweets in another language than English.
- The paper considers in its study audiovisual data, and/or emojis and/or hashtags.

Among 736 papers, only **187** fulfilled the criteria mentioned below.
Among them 1 was duplicated, so we were left with 186 papers in total.

2- *In the second phase*, the selected papers underwent a more in-depth filtering process for further refinement. It consists of reading the paper more carefully, by paying attention to the objective of the paper, the methodology used, and the obtained result.
During this process, we were especially interested in papers that had a business or social objective and a well-defined methodology.
In order to have the most pertinent articles for our framework, we applied the following criteria.

Inclusion criteria:

- The paper explains in a well-defined and structured way the methodology followed.
- The paper studies text-only tweets.
- The paper has a social or business objective behind the Sentiment Analysis.

Exclusion criteria:

- The paper doesn't explain in a structured way the methodology followed.
- Sentiment Analysis is not the main focus of the paper.
- Twitter is not the main social media studied in the paper.
- The paper studies also unstructured data such as images, video, etc.
- The paper considers hashtags in its study.
- The paper considers emojis in its study.
- The paper studies tweets in languages other than English.

- The study aims to compare the performance of several algorithms without having a social or business goal.
- The study aims to propose a new method/algorithm without having a social or business goal.

After applying the exclusion criteria mentioned above, 60 articles were considered for the framework of this thesis.

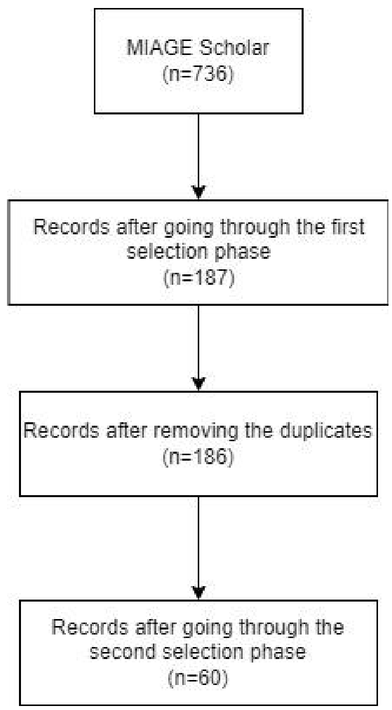

Figure 2. The process of selecting the papers

3.5. Data extraction

To organize our data, we developed a framework that contains the selected papers. The objective behind this is to create a system that makes it easy to comprehend and compare the methodologies that have been followed to apply Sentiment Analysis on Twitter during the last few years. It also facilitates the structure and visualization of the findings and the answers to our research questions.

This framework is presented in a table chart containing 60 scientific papers retrieved from Miage Scholar and it is found on Appendix 1. As mentioned in 3.4. Selection criteria section we have retrieved the papers published **from 2018 to 2022**, that **have a business or social objective** behind their Sentiment Analysis and study **text-only tweets**.

[30] is a secondary study that has analyzed papers published from 2004 to 2020 that study emotion detection. To present their results, the authors have developed a comparative table with several criteria organized by the domain in which the paper has contributed. The framework that they have developed consists of the following criteria: *emotion model*, *dataset*, *algorithm/technique/method*, *objective*, *advantages*, *disadvantages*, *evaluation measures and emotion detected*. The emotion model is out of scope for our study. Advantages and disadvantages are out of the scope as well. We are concentrated in the algorithm/technique/method field, but we want a more detailed flow of the process. [29] has gone through the same process by analyzing methodologies of sentiment analysis of 60 papers. Their focus is on soft computing techniques and in their final framework they included criteria such as the *domain* where the sentiment analysis was applied, the *techniques* used (the algorithms) and the *tools* (lexicons, softwares, python libraries etc.)
Considering the criteria from [29] and [30], the framework built for this study is developed through 14 fields, described as below:

- Paper title: The title of the paper
- Publication year: The year of publication
- Application domain**:** The field of study to which the paper contributes
- Objective: The business or social objective of the paper
- Data source: Where the data of the study is retrieved from
- Data collection: Detailed insights into the data collection methodology, including the filters used, or the database name in the case open data.
- Emotion terminology: The terms applied in literature to express emotion.

- Detected emotions: The emotions identified through the Sentiment Analysis.
- Preprocessing: The steps undertaken in data preprocessing.
- Processing: The methods used to prepare data for classification.
- Classification Approach: The approach used for classification, larger view.
- Algorithm/ML Model: The specific algorithm or ML model used for classification.
- Evaluation Method: The approach for evaluating the employed methodology.
- Contribution type: A categorization of whether the paper introduces an innovative methodology or relies on established traditional methods.

4. Related work

Numerous literature reviews have been conducted on the subject of Sentiment Analysis.

In [27], the authors analyzed 19 articles published between 2019 and 2020. They developed a framework that considers various criteria, including the dataset used, the technique employed, feature representation, general observations, and accuracy. While the paper notes that the majority of the data came from Twitter, it lacks specific clarity on the other social media platforms considered in the study.

[28] analyzed 26 articles to develop a comparative and statistical study of research on sentiment analysis in the last few years. The paper is highly centered in the preprocessing phase of Sentiment Analysis, by explaining in detail how this process is carried out. They also considered the translation of non-English papers into English as a step of preprocessing. The authors compared articles based on criteria such as the objective and the domain of the paper, the size and the source of the dataset, preprocessing steps, language before preprocessing and after preprocessing, approach, and accuracy. It is not specified the timeframe of the studied papers, nor the social media on which the papers were based.

[29] has conducted a thorough and complete analysis of the existent literature on Sentiment Analysis on Twitter. The authors studied 60 papers published between 2012-2017 and they built a comparative table with the following criteria: the techniques used, dataset, tools, domain, cross-validation, performance parameters, and accuracy. The paper is concentrated only on papers that have used supervised Machine Learning algorithms, therefore excluding all the Sentiment Analysis lexicon-based.

[30] is another completed study carried out to review the emotion modeling approaches for text-based sources. The results are presented through the qualitative analysis of the papers published during 2004-2020. The authors have developed a comparative table organized by the domain in which the paper has contributed. The paper does not show any particular interest in the social media domain.

[31] provides a comprehensive review of sentiment analysis in social media, covering methods, platforms used, and applications. The study analyzed articles published between 2014 and 2019, selecting 24 out of 77 based on their objectives. Most articles that were considered in this study utilized the opinion-lexicon method, focusing on Twitter data and applications in world events, healthcare, politics, and business.

This thesis aims to answer the research questions presented in 3.2. Research questions. 60 papers were reviewed and categorized in several criteria. The results of the paper are presented in the following section 6. Results, where the answers of the research questions are elaborated, starting from a description of the data and some general statistics, the domains where the Sentiment Analysis is applied, the application of SA and the model evaluation.
Finally, the conclusions of the paper are presented as well as the potential future research points in the field. An appendix is associated with this thesis, that represents the developed framework.

6. Results

6.1. Framework statistics

The chart in Fig. 3 illustrates the number of papers that we retrieved per year, highlighting an increase in publication over the years. Notably, the peak was reached in 2021 with 22 papers. This might be attributed to the significant interest of the researchers in the impact of Covid-19 on the population. A thorough analysis of the reasons why Sentiment Analysis has been applied in the last years and the specific field to which the papers contribute is found in section 6.3. The domains where Sentiment Analysis on Twitter is applied.

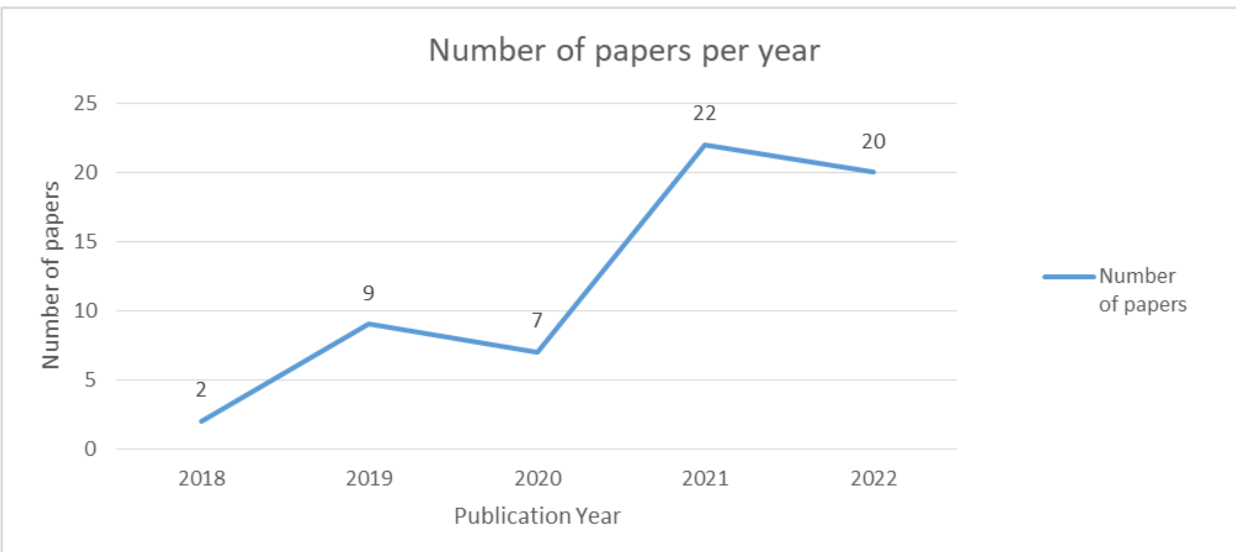

Figure 3. Number of papers per year

Most of the papers (32) used Twitter API, the official API provided by Twitter, to fetch their data. Kaggle, a data science competition platform, offers various public databases, including tweets on different topics. 12 of the papers' datasets were based on Kaggle and the rest are public datasets provided by other platforms, developers, or Python libraries like GetOldTweets3 or Snscrape. There are also 7 papers whose dataset is not explicitly mentioned, they are therefore marked as Unknown. A visual representation is found on Fig. 4.

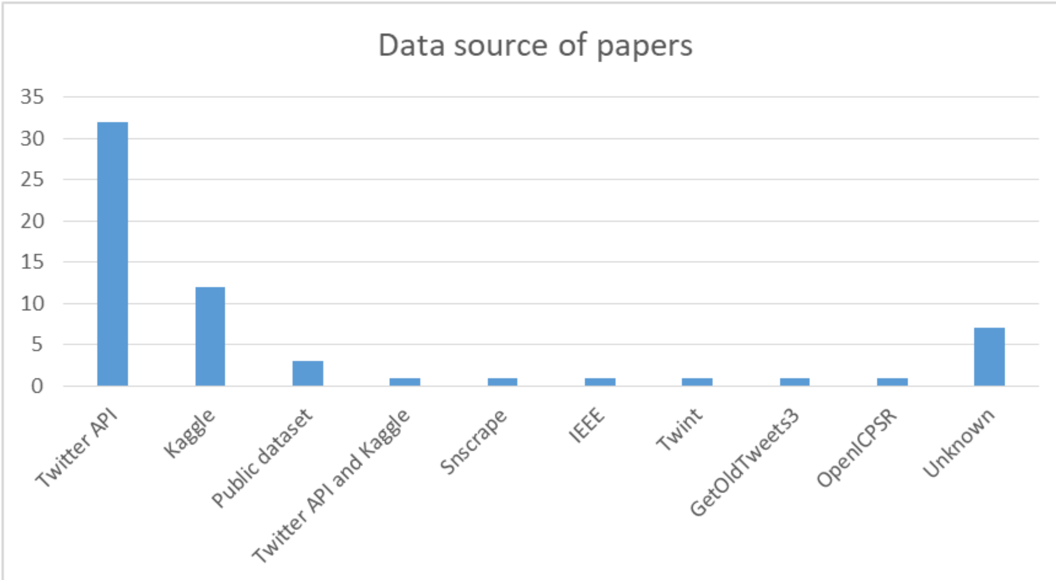

Figure 4. The data sources of the papers

6.2. The terms used for expressing emotions on Twitter

This part of the research responds thoroughly to the first research question, RQ1. As it was mentioned in 2.2. Sentiment Analysis section, different terms are used to express the outcome of the Sentiment Analysis process. These terms include *sentiment*, *emotion*, *feeling*, *opinion*, etc.

Sometimes, the words have been used interchangeably. To ensure clarity in our thesis and to have a systematic answer for our RQ1, in order to define the specific notion of expressing emotion in each paper, we have adopted the term that the authors themselves used to express their findings. For example, the authors in [32] use both *sentiment* and *emotion* in their study. They are mentioned in the following citations respectively:

"The tweets were assigned identifiers based on the probability scores utilized to assess and quantify the strength of the public ***sentiment***."

“The proposed work was further augmented with a lexicon-based emotion analyzer to categorize the ***emotions*** associated with the tweets into positive, negative, and neutral”

As we may notice in the second citation, the term *emotion* is used to describe the categorization of tweets into three emotional classes, which presents the outcome and the finding of [32]. For that reason, in our framework, we designated *emotion* as the relevant term of this paper. This approach was constantly applied to all the papers.

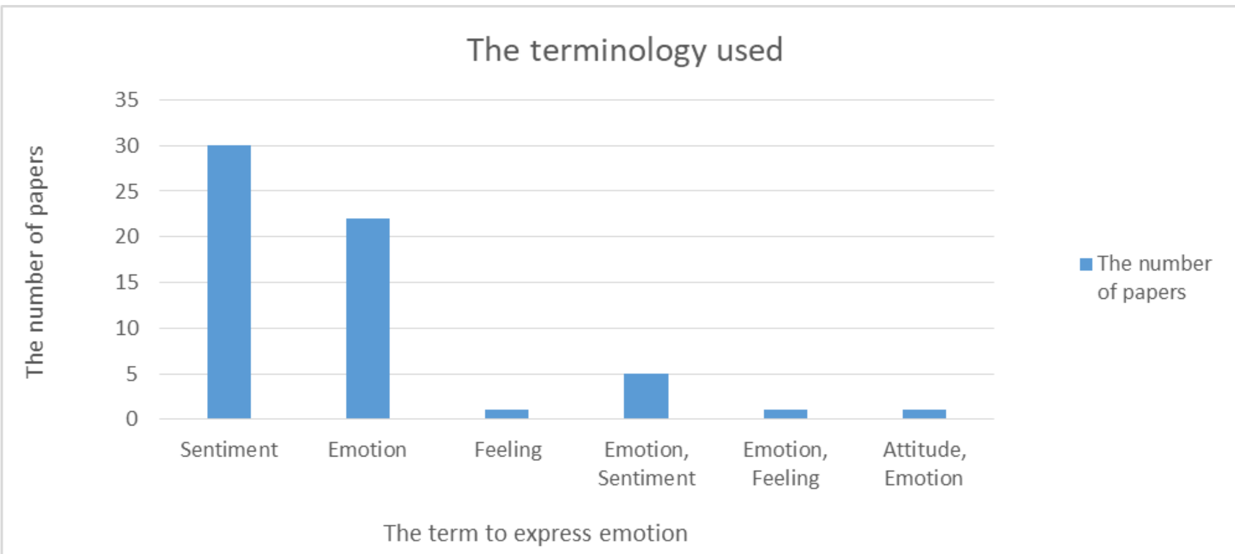

Figure 5. The terminology used to express emotions

As we may notice, from Fig. 5 the most used term in our dataset is *sentiment* with 30 occurrences followed by *emotion* with 22.

The questions regarding the relationship between *sentiment* and *emotion* are not recent. In 1954 the English philosopher Charlie Dunbar Broad published a paper [33] presenting the emotion classifications and the difference between *sentiment* and *emotion*.

Broad [33] highlights that emotions are immediate and direct responses to specific stimuli or situations. They are characterized by their emotional quality and intensity, often arising from attributes of an object. For example, fear or anger felt in response to a sudden loud noise or perceiving a threat. Whereas sentiment is a more complex and broad concept. [33] highlights that sentiments are formed through repeated interactions with an object. They involve multiple emotional tones that become associated with an idea of the object. For example, a sentiment of love for a person that is formed through repeated experiences and associated emotions.

This also applies in our study, where in the majority of the papers *sentiment* is a broad concept and it takes value like *positive* or *negative*. Optionally, they may or may not include *neutral*. Whereas emotions take several nuances within a sentiment, for example, *anger*, *anticipation*, *disgust*, *fear*, *joy*, *sadness*, *surprise*, and *trust*.

Papers like [11], [34], [35] highlight a strong dependent relationship between *polarity* and *sentiment*. Polarity is the technical result after processing data. It is presented by a score (for example [-1,1]) attributed to a word, sentence, or document that allows one to define the sentiment as positive, negative, or neutral.

Among our dataset, 5 of the papers distinguish sentiment from emotion, by using them both in their findings. [11], [38] and [39] classified the users' tweets into 8 emotions: *anger, anticipation, disgust, fear, joy, sadness, surprise, and trust*, and 2 sentiments: *positive, and negative.*
[37] using the Machine Learning approach classified tweets into 3 sentiments: *positive, negative*, and *neutral*, and 6 emotions: *anger, disgust, fear, joy, sadness, surprise,* and *neutral.*

It is also noticed a usage of the term *feeling* in the studied papers. Even though a term very closely related to emotion and sentiment [14], *feeling* is rarely used in Sentiment Analysis literature. The authors of [40] don't explain the difference between these concepts, either they hold a strong opinion on what term should be used. However, they use *feeling* as a relevant term for their study. [41] uses numerous synonyms to express emotions such as sentiment, feelings, and opinion, but they define their output after processing data in two categories; 2 feelings: *negative* and *positive,* and 8 emotions: *anger, anticipation, disgust, fear, joy, sadness, surprise, and trust.* It is the same system as [11], [38] and [39] but the term changes. They also used the same lexicon as [11] which is the NRC Word-Emotion Association Lexicon.

One paper [42] has also used the term *attitude* as a synonym for sentiment. Without giving any further information regarding the difference, the authors of [42] use these terms interchangeably. However, while explaining their output after data processing, they use the term *attitude* to refer to their finding. Same as [11], [38], [39], and [41] the authors classify their data into 2 attitudes: *negative* and *positive,* and 8 emotions: *anger, anticipation, disgust, fear, joy, sadness, surprise, and trust.*

6.3. The domains where Sentiment Analysis on Twitter is applied

The application domain is the field study to which the paper contributes and it is closely aligned with the study's objective. The studied papers were from a wide range of domains, from healthcare and marketing to social movements and natural disasters. This was useful to answer comprehensively the research question number two, RQ2.

All topics related to customers, such as customer behavior, customer experience, customer satisfaction, and customer reviews were considered to belong in "Marketing", as a higher-level domain. Brand management and product review were also classified within the marketing domain as well. On the other hand, the topics related directly to technology service, transport service, or any other consumer service were classified under the "Service" domain.

"Healthcare" encompasses subjects related to health, including mental well-being, COVID-19, and other medical conditions.

Furthermore, some papers study wars, mass shootings, or other urgent scenarios that were categorized under the "Traumatic event" domain.

Noticeably, the application domain with the highest number of papers is Healthcare totaling 25 papers. During 2021-2020 there was a significant interest in comprehending the emotional consequences of COVID-19 on individuals. This is evident from the 19 papers directly related to COVID-19, constituting 76% of the papers within the Healthcare category and 32% of the overall paper count [3][5][7][46-61]. [3] and [49] focus on vaccines only. Their objective is to analyze and evaluate the opinions expressed on Twitter regarding different vaccines for COVID-19 and explore the key topics and themes discussed about it.

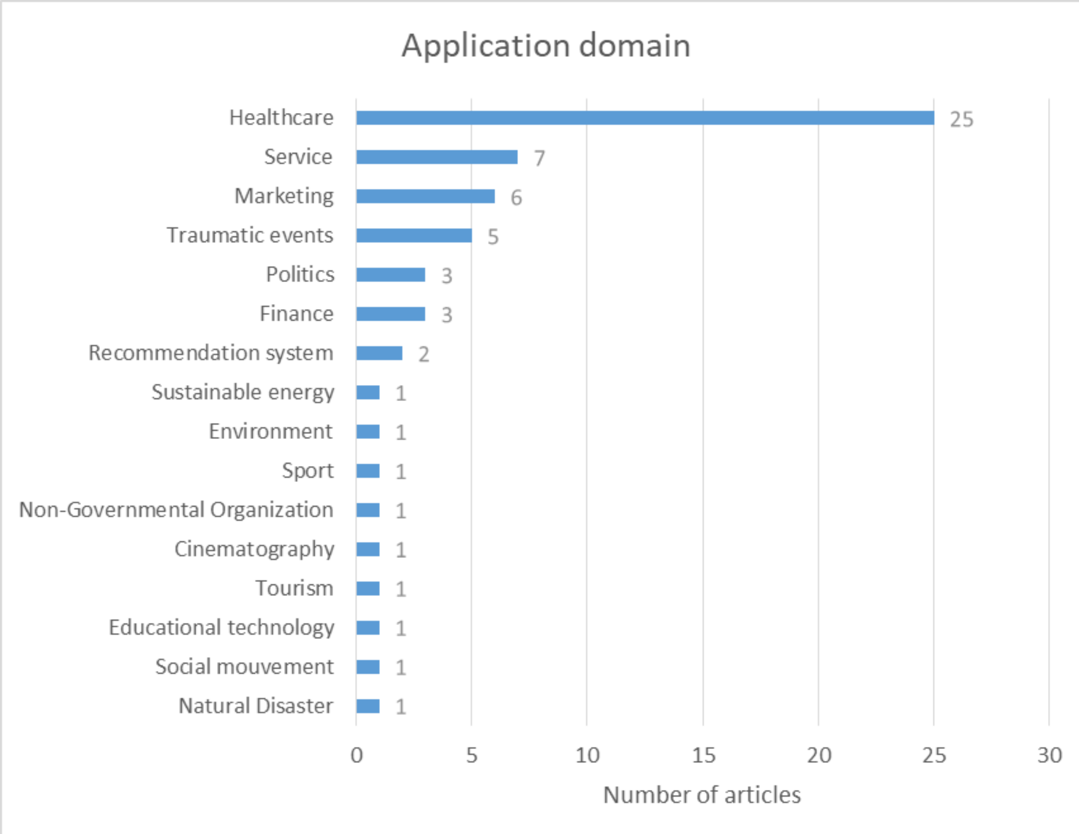

Figure 6. The application domain of the reviewed papers

[46] classifies worry levels in COVID-19-related discussions on Twitter, by presenting three levels as output: *no-worry, worry,* and *high-worry.* [54] and [58] have filtered tweets by geographic location, respectively New Zealand and India. The information and the output provided are relevant for these locations only. [53] tries to predict the popularity of tweets, measured by retweets, by analyzing public emotions of Covid-19. According to [53], tweets with higher emotional intensity are more popular than tweets presenting information on the COVID-19 pandemic.

Other papers within the Healthcare domain pay attention to mental health. [62] classifies suicide-related content on Twitter, while [6], [63], and [64] try to understand the depressed patient's social behavior in order to detect depression among the population.

Furthermore, it has been shown an interest in domains such as Service and Marketing. It is always a point of interest for enterprises to understand the customers of their brand [2] [45], the satisfaction of their customers [43], and analyze product reviews [44]. The same interest applies to services.

[2][43][65][66] use Sentiment Analysis for analyzing customer reviews. [43] proposed an enhanced method to conduct SA whereas [65] created a web application where users may get Twitter reviews for any skincare product.

Within the Service domain, [67] and [68] analyze public opinion regarding the 4G and 5G technology. [69] and [70] use Sentiment Analysis to classify public opinion towards transport services (Uber & Ola cab services) and the Airbnb business model. Meanwhile [84] and [71] analyze sentiment within airlines and flight services.

Traumatic events are another interesting domain where Sentiment Analysis takes hold. The Ukraine-Russia war has taken a high interest in researchers trying to understand the emotions of the population touched by this conflict [72][73]. Other papers, [8] and [74] comprehend emotions related to mass-shooting and terror-related events.

Sentiment Analysis takes hold in the political field where researchers explore the feelings of fear and anger during the 2020 [76] and 2016 US presidential elections and Brexit [75].

Moreover, research has been carried out to explore the influence of Twitter polarity on the stock market movements [77], as well as on the NFT market [38]. [35] tries to predict the stock volatility of giant companies like Apple, Microsoft, and Amazon, with an accuracy below average but another potential to be explored.

6.4. The application of Sentiment Analysis on Twitter

There are two approaches used for Sentiment Analysis: Lexicon-based approach and Machine Learning approach.

6.4.1 Lexicon-Based Approach

The lexicon-based approach uses the information that a sentiment lexicon holds [80] which is a predefined list of words where each word is associated with a specific sentiment [78]. A lexicon score is calculated based on the polarity of the words. The sentiment of a text is calculated from the polarity of the words or phrases in that text [79]. The sentiment score of the text can be computed as the average of the polarities conveyed by each of the words in the text [79]. There are several lexicons and tools developed by researchers and used in the Natural Language Processing field, especially in Sentiment Analysis.

The tools and lexicons used for Sentiment Analysis on Twitter:

a. LIWC - Linguistic Inquiry and Word Count

LIWC-22, the Linguistic Inquiry and Word Count 22, is a text analysis tool composed of software and a dictionary. [81] This dictionary links words in analyzed texts to psychological concepts. Words in analyzed texts are known as target words, while those in the LIWC-22 dictionary are called dictionary words. These words are categorized into groups representing different domains, termed "sub dictionaries" or "word categories." [81] For instance, the term "cried" is part of categories like affect, emotion, and communication. LIWC22 is the newest version of LIWC and it contains over 12,000 words, phrases, and stems, evaluating psychosocial constructs. This version introduces new, revised, and removed categories to suit improved analytical methods and language diversity [81].

b. Sentistrength

SentiStrength uses a collection of sentiment words, emoticons, and stems from its dictionary to assign either positive or negative values to words, by having as an output a sentiment score for the entire text [82]. Each text assessed by SentiStrength is assigned a positive sentiment score between 1 and 5, as well as a negative score ranging from -1 to -5 [82]. The overall sentiment score for a sentence is the highest positive and negative values assigned to its words. In multiple sentences, the sentiment score for the entire text is determined by the maximum scores of the individual sentences [82].

c. NRC - National Research Council Canada

The NRC Emotion Lexicon is a list of English words and their associations with eight basic emotions (anger, fear, anticipation, trust, surprise, sadness, joy, and disgust) and two sentiments (negative and positive) [83]. Also known as EmoLex (its original name), the lexicon contains 14,182 unigrams (words).

d. TextBlob

TexBlob is a Python library for text data processing. It provides an API that can be used for several natural processing research tasks [54]. TextBlob labels the tweets based on semantic orientation and intensity of every word, by using lexicon-based approaches and its output is two properties: polarity and subjectivity of a sentence [55]. The polarity score is given between -1 and 1, with -1 being the most negative and 1 being the most positive [50]. Each tweet is also assigned a subjectivity score between 0 and 1, with 0 being the least subjective and factual and 1 the most subjective and highly opinionated [50]. [100] used TextBlob and VADER to retrieve the polarity of tweets and they also used NRC to detect the emotions of the tweets.

e. VADER - Valence Aware Dictionary for Sentiment Reasoning

VADER, a widely employed lexicon and rule-based model, finds common application in sentiment analysis on social media platforms. Within VADER, words are categorized based on two key factors: their polarity, which can be either positive or negative, and the degree of polarity intensity [56]. This not only enables the determination of sentiment in tweets but also provides insights into the emotional depth of the expressed sentiment [56]. In a study referenced as [101], VADER was utilized to assess tweets, assigning ratings ranging from -4, denoting "Extremely Negative," to 4, signifying "Extremely Positive," with a score of 0 indicating "Neutral" emotions. One of the major advantages of employing VADER lies in its ability to incorporate additional words and features commonly found in social media, such as acronyms, emoticons, and slang [56].

There is a visual representation in Fig. 7 that displays the usage of every lexicon tool among the reviewed papers. There are papers that have used more than one lexicon, to cover a wide range of emotions [75]. It is noticed that the NRC lexicon is the most used lexicon followed by TextBlob and VADER.

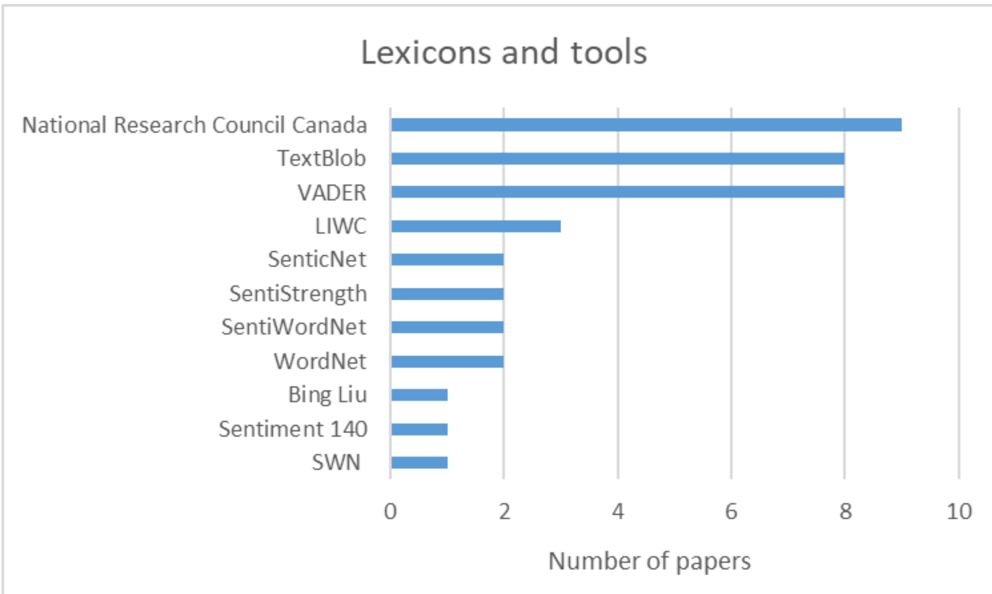

Figure 7. Lexicons and tools used

6.4.2. Machine Learning Approach

The Machine Learning approach uses Machine Learning algorithms to classify sentiment. The ability to learn from data sets empowers these algorithms to draw conclusions about the properties of data and predict future outcomes of new, unseen data [86]. Machine Learning encompasses the Classical Machine Learning approach and Deep Learning approach [30].
Machine learning algorithms categorize text into distinct emotion classes. Two classifications exist for these algorithms: supervised and unsupervised [30]. In the reviewed literature, supervised learning finds extensive usage.

Unsupervised algorithms include Hierarchical, Cmeans, Kmeans clustering, etc.
Supervised learning is made of a two-step process [30]:

- Learning (training): Learn a model using the training data.
- Testing: Test the model using unseen test data to assess the model's accuracy.

The algorithms used for Sentiment Analysis

Among the reviewed articles the most used supervised models are Naive Bayes classifier, Support Vector Machine, Logistic Regression, and The K-Nearest Neighbor.

- **Naive Bayes Classifier** is a probabilistic classifier that is based on the Bayes theorem with the assumption that features are independent. The Naive Bayes classifier defines the probability of the document belonging to a particular class [87]. It is easy to implement and it has been widely used by the reviewed articles, yet there are other algorithms that have higher performance.
  [47] the lowest accuracy compared to Support Vector Machine, Logistic Regression, Random Forest, Decision Tree, and XGBoost. [54] also demonstrated a deteriorated performance of NB compared to SVM. [35] used Naive Bayes to classify tweets before processing them into RNN (Recurrent Neural Network) to predict stock price movements.

- **Support Vector Machine** is a supervised machine learning algorithm that acts like a hyper-plane dividing various data types. In a two-dimensional space the hyper-plane is represented as a single line [89]. It takes as an input a vector space and its output is two values, whether positive or negative [88]. In general, SVM demonstrates high performance in terms of accuracy. In [7] SVM outperforms other algorithms K-Nearest Neighbor and Naïve Bayes Logistic Regression. [47] also compared the performance between SVM and Multinomial Naïve Bayes, Logistic Regression, Random Forest, Decision Tree, and XGBoost where it proved an accuracy, F1 score, and precision of higher than 90% over four different datasets. In the majority of reviewed articles SVM had quite a good performance [40] [46][65][92]

- **Logistic Regression** finds the output in scenarios when having one or multiple independent variables. The resulting value is commonly represented as either 0 or 1, following a binary format [90]. Mostly the sentiment output is in binary format, negative or positive. Its performance has been compared with other algorithms [46][47][72][55]. In general, it performs quite well, in [92] has the highest accuracy compared to other classifiers, while in [71] has a lower accuracy than Random Forest.
  [53] retrieved a higher performance when combining LR with TF-IDF vectorizer while [66] evaluated that the most accurate "model-vectorizer" combination was Logistic Regression with Count Vectorizer with an 87.2% model accuracy.

- **The K-Nearest Neighbor** classification method relies on evaluating the distance between the unknown data point and the training data point that share similarities. This distance is quantified using the Euclidean distance metric, where closer data points signify greater similarity between them [91]. The KNN algorithm has been compared to other classifiers but in none of the reviewed papers had the highest performance. [7], [54], [66], [92], [67]

Fig. 8 illustrates statistics on the classification approach used. Among the reviewed articles, the majority, around 32% used the machine learning approach to conduct Sentiment Analysis, followed by 28% using a deep learning approach, and finally, 30% used a lexicon-based approach. The final classification, deep learning, and machine learning is not a hybrid approach. These papers have used both methods separately in order to compare the performance of each algorithm. [6][2][40][35][51][67]

Fig. 9 illustrates the used machine learning algorithms in the reviewed literature. Some articles have used several of them to compare their performances, every applied algorithm is counted. The most used algorithms are Naive Bayes used by 25% of papers, Support Vector Machine by 22% and Logistic Regression by 18%.

Fig. 10 demonstrates the utilization of deep learning algorithms among the reviewed papers. Some articles have used several of them to compare their performances, every applied algorithm is counted. The most used algorithms are Convolutional Neural Networks used by 5 papers [8][54][69][68][59], followed by the BERT algorithm used by 4 papers [93][8][63][52].

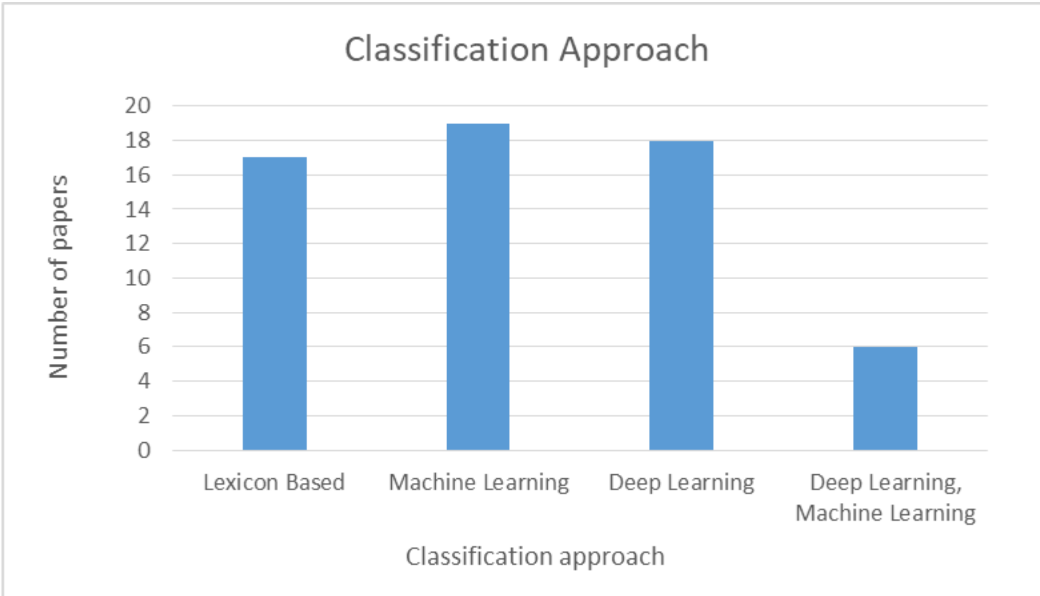

Figure 8. The classification approach of the reviewed papers

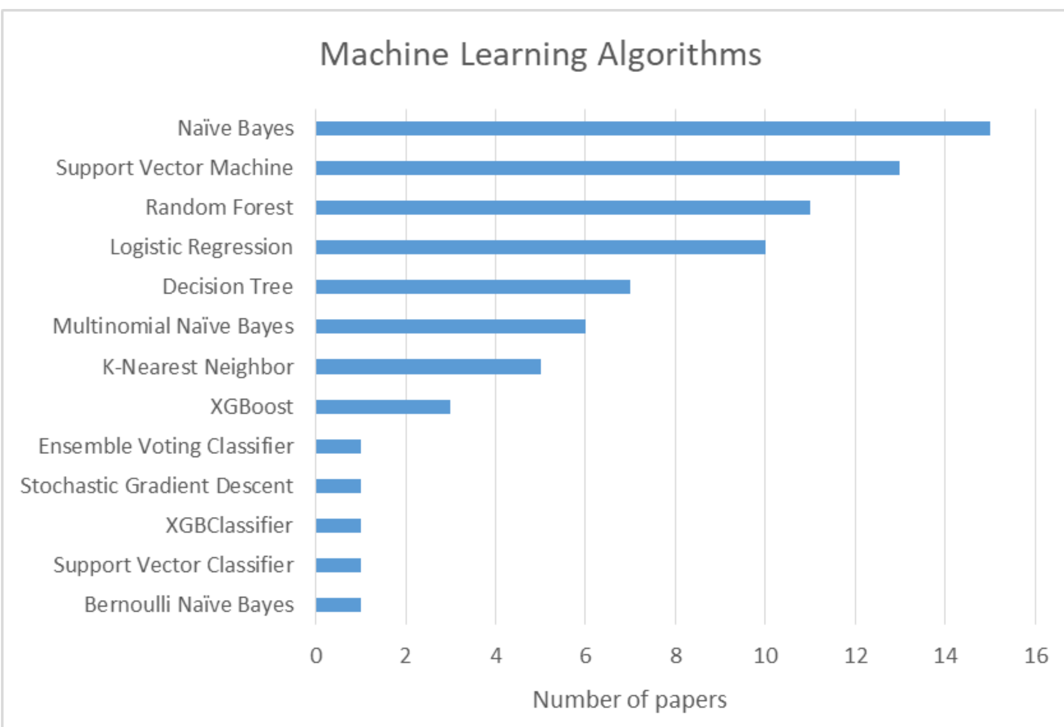

Figure 9. The machine learning algorithms used

In their study, [8] utilized three distinct algorithms: CNN, BERT, and Bi-LSTM. Their findings demonstrated that CNN achieved the highest accuracy among these algorithms. Conversely, a different study by [54] compared CNN with various Machine Learning algorithms and determined that SVM exhibited superior performance. According to the authors of the paper [69], CNN is known for image processing. Motivated by this, they sought to evaluate CNN in the domain of text analysis. However, their results indicated suboptimal performance when compared to both Deep and Non-Deep Neural Networks.

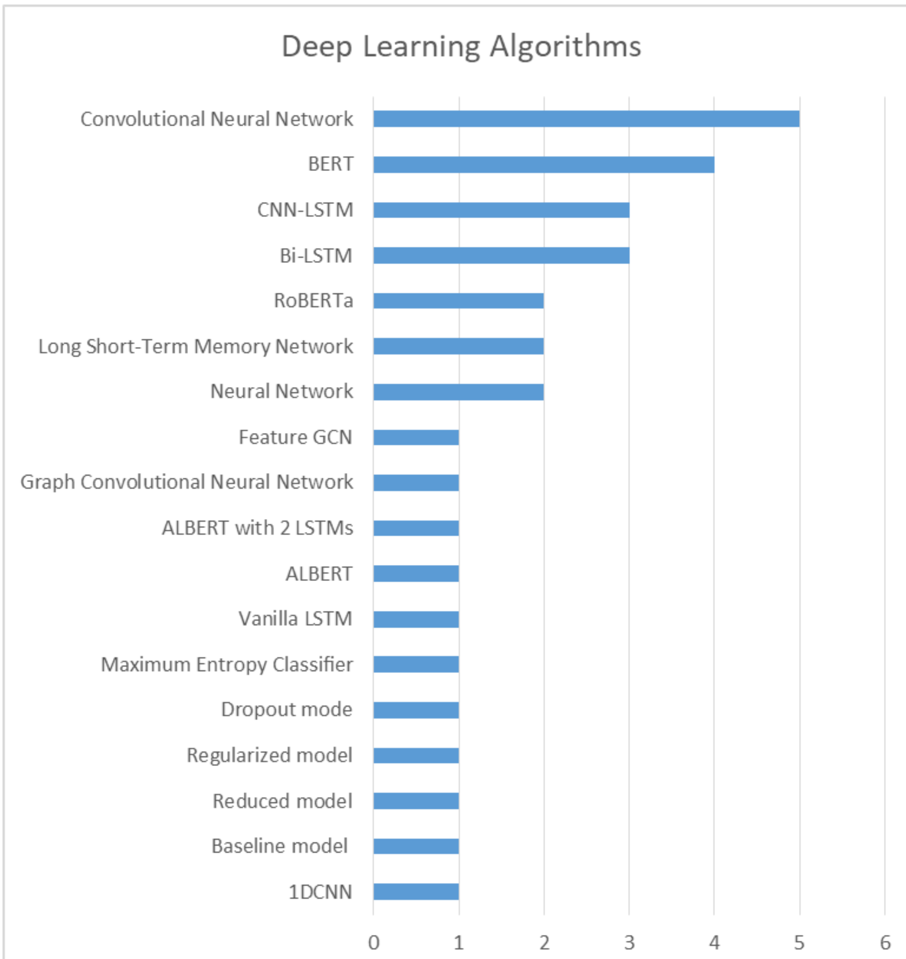

Figure 10. The deep learning algorithms used

BERT, a state-of-the-art language model presented by Google, is used to learn the context of a word in a given text segment [52]. [63] uses BERT to assign a tweet as depressive or not. Data is passed through several layers and the output layer gives an array of probabilities for each class of input text and the final class is chosen as the one with the highest probability. The authors of [63] tested other deep learning algorithms such as CNN-LSTM and LSTM, resulting in a higher accuracy of BERT. [48] compared specific models of BERT with one another. The authors used BERT models pre-trained on Covid tweets and RoBERTa, an improved BERT model that is capable of predicting empty words in a text.

Long short-term memory networks (LSTMs) are a unique type of computer program that excels at learning and retaining information over extended periods. LSTMs stand out because they possess a mechanism to retain crucial information while discarding unnecessary details. This ability allows them to effectively handle information that follows a recurring pattern over a long duration [67]. The authors of the paper [67] compared LSTM with traditional algorithms Naive Bayes, SVM, KNN and as a result LSTM had the highest accuracy.

6.4.3. The steps of Sentiment Analysis process on Twitter

Both the lexicon-based and machine learning methods follow similar initial steps to process their data. They both begin by collecting and preparing the data through a preprocessing phase.

In the lexicon-based approach, once the data is cleaned, it can be directly processed using pre-built lexicons. This allows each tweet to be assigned a sentiment.

In the machine learning approach, the data also undergo preprocessing. However, before classifying they must be formatted as suitable inputs for a machine learning algorithm. This involves a feature extraction step to convert the input into a vector format. Usually, the public databases retrieved from Kaggle contain the assigned sentiment for each tweet. This is useful to train the model. If the sentiment hasn't been previously assigned, a lexicon tool is used to find the sentiment, train the model, and apply the algorithm [55]. As a result, we retrieve the classified sentiments. Finally, the models are evaluated.

In summary, both methods start by preparing data, with the lexicon-based approach using built-in lexicons for processing, and the machine learning approach using feature extraction and potential lexicon application before classification.

Fig. 11 represents a BPMN diagram explaining the general flow of both approaches.

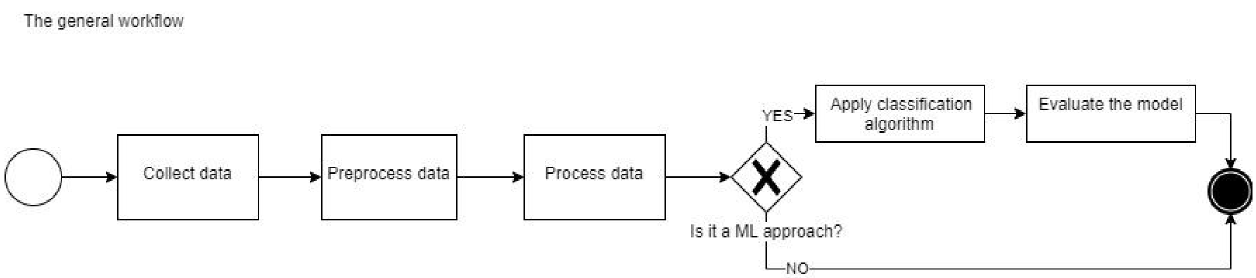

Figure 11. The general flow of both approaches

We will break down the detailed steps for both approaches, starting with the common initial steps of data collection and preprocessing. Afterward, we'll give an overview of the specific steps for each methodology.

Data collection

As illustrated in Fig.4 the most used data source is Twitter API followed by public databases provided by Kaggle.

The Twitter API allows filtering data by using keywords, publication date, and location before fetching it. Papers [2][3][49][41][94][53] have been filtered by keyword and date range. [37][50][54][55] have used hashtags and a date range to filter their tweets. For datasets retrieved from Kaggle or other open sources, usually, the dataset is already filtered and the authors select it by topic [84][40][96][32][57].

Data preprocessing

The data preprocessing phase encompasses several steps, represented as in Fig.12. The whole process is described below.

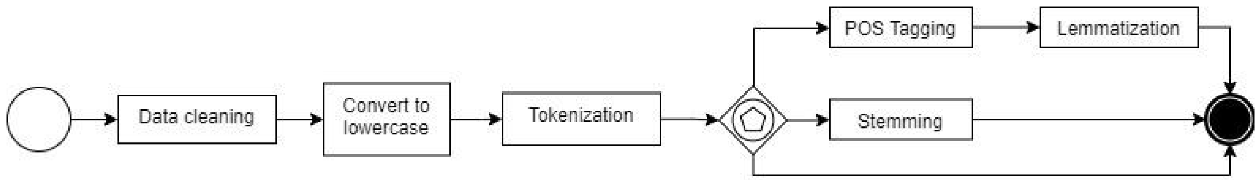

Figure 12. Data preprocessing phase

- **Data cleaning**

This is an important step before starting to process data. Raw data usually contains a lot of noise, whitespaces, stop words, slang words, URLs and hashtags. This makes the data inconsistent impacting the performance and the result of the model [55].

Data cleaning includes the removal of special symbols, hyperlinks, mentions, stop words, and punctuation marks [2][3][11][46][47][48][37]. [97] converted @username to username. [66] removed tweets having less than two words. [49] filtered non-English tweets.

- **Lowercase**

In order to avoid conflict with the case sensitivity, it is suggested to convert all letters in lowercase [51][53][97][50][47][46][11][97][51]. Some papers decide to follow the same logic by transforming all the letters in uppercase [8].

- **Tokenization**

Tokenization consists of splitting the phrase into smaller pieces called tokens. These tokens might be words, symbols, punctuation, etc. [56][98][57][58][92][75][72][64]

- **Part-Of-Speech - POS**

POS tags serve the purpose of distinguishing words and elucidating their grammatical functions [61]. This tagging process identifies the word class of each word in the sentences, such as whether it's a noun, adjective, verb, etc.[38]. The assigned POS tags are crucial for achieving accurate lemmatization results [61]. Among the reviewed papers, three of them have used POS tagging in their preprocessing phase [60][61][106] followed by lemmatization.

- **Stemming**

Stemming is a process of text normalization into standard language that removes the prefixes and suffixes of the word by keeping only its stem [70]. It is faster than lemmatization by not as advanced, as it only removes the suffixes without having a morphology context of the word and it is not smart when processing irregular verbs [110]. For example, the verb *leaving* it is transformed as *leav*, without *e* in the end. Moreover, the word *leaving* may be used as a noun, for example *your leaving*, or as an adjective *the leaving party*. In every case the stemming process will return *leav*, without considering the function of the word.

- **Lemmatization**

Lemmatization is the process of finding the normalized form of the word, by adding or removing its suffix. For example, the words *stops, stopped,* and *stopping* should be cleaned from their suffixes and tuned in their normalized form which is *stop* [99]. Lemmatization is a more advanced process since it considers the function of the word but also transforms the verbs in their infinitive form [110].

Among the reviewed papers, stemming and lemmatization are used interchangeably. The only exceptions are papers referred as [38] and [54] that have included both lemmatization and stemming in their preprocessing diagram without explaining in detail. According to [38], lemmatization is preferred over stemming because it produces better results by performing an analysis based on the word's part and producing true dictionary words.

Other steps that have been applied from a minority of the reviewed papers are: the correction of spelling mistakes [97], removal of very short tweets (less than two words) [66], removal of non-English tweets [57], contraction expansion [51], expansion of abbreviated words [100].

Once the preprocessing is finished, the data are cleaned and in the right format to continue in the processing phase. The latter is slightly more complicated for the machine learning approach. Fig. 13 explains visually through a BPMN diagram, the steps followed during Processing and Evaluation for machine learning and lexicon-based approach. This phase is elaborated in the following paragraphs for each approach.

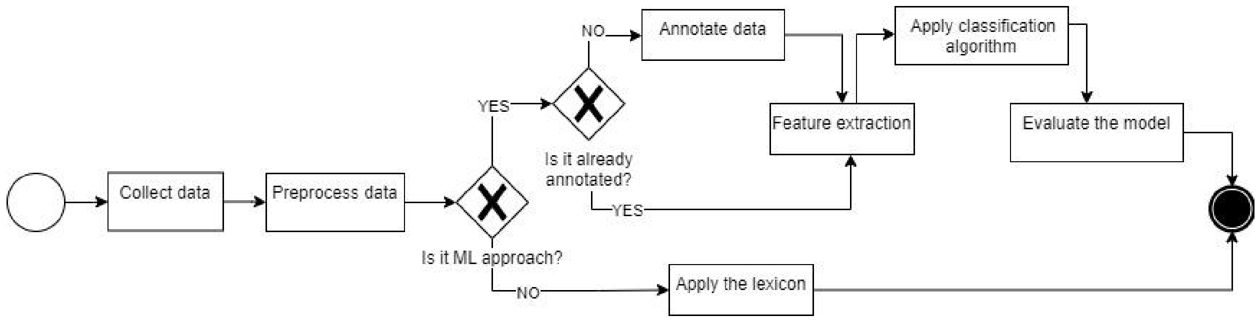

Figure 13. A detailed view of Process data phase

Lexicon-based approach

Processing

Following the preprocessing phase, the text goes under processing steps where each word is compared to the lexicon selected. The lexicon is initially provided with the polarity of each word individually and it contains words that are used to express an emotional nuance [30][79]. When the word from the text is found in the lexicon, then the indicated polarity is attributed to it. Otherwise, if it doesn't exist, then the word is automatically assigned as neutral. This process continues for every word of the sentence. In the end, the polarity of the sentence is determined by calculating the average of the polarities associated with its individual words [79].

This process is done automatically by using Python libraries. The lexicon is chosen according to two main criteria:

- **The spectrum of emotions**. For instance, when aiming to have a spectrum of 8 emotions, the NRC lexicon is employed [11][39][100][41][42][102]. Alternatively, when the goal is having two or three emotions detected (positive, negative, neutral), lexicons like TextBlob [3][50][58], VADER [3][49][95][56], SenticNet [94] come in handy.

- **The objective of Sentiment Analysis.** Some lexicons are created for a specific domain. For example, VADER was developed for social media-style text aiming at real-time streaming data and incorporating elements, such as emoticons [101]. SentiHealth is a health-related sentiment lexicon [103]. Lexicons such as NRC [83] or SentiWordNet are general sentiment lexicons [101].

This is the last step of lexicon-based Sentiment Analysis. In general, this approach is not evaluated at the end. The evaluations are carried out beforehand when creating the lexicon and they are manually checked by language experts and reviewers [83].

Machine Learning Approach

Processing

When using the Machine Learning approach, the processing phase implicates some extra steps than the lexicon-based, that are useful to prepare the data as the right input for the machine learning algorithm.

- **Data annotation**

Data annotation in sentiment analysis involves the process of labeling text data with sentiment labels or polarity values to create a labeled dataset for training and evaluating machine learning models. This annotated dataset serves as the foundation for teaching the machine learning model to understand and predict sentiment in text [111]. The public databases, for example the databases on Kaggle, are already based with a sentiment [66]. This facilitates the process of sentiment analysis while following the machine learning approach. For that reason, the majority of the reviewed papers haven't gone through this process. The paper referred as [7] retrieved the data from Kaggle already labeled, however they applied VADER and replaced the existing labels. Other papers, especially the ones that have used Twitter API have annotated their tweets before passing to other phases. [38] uses NRC lexicon to annotate the data before using GloVe for word vectorization. [6] used TF-IDF to get the most important words then used LIWC to get the sentiment and classified the words in 14 psychological attributes. [64] used Sentistrength to annotate their sentiment and train the NN model. [63] utilized VADER to label their tweets and assign scores within the range of -1 to 1. This particular study aimed to identify two emotions, namely depression and non-depression; thus, following the application of VADER, only negative tweets were retained.

Additionally, there exist some research papers that omit details regarding their annotation process. For instance, [37] states that all tweets and comments underwent annotation, categorizing them into positive, negative, or neutral sentiments, without specifying the annotator or the followed process.

- **Feature extraction**

In the field of natural language processing, it is essential to convert textual data into a format compatible with algorithms. Typically, this involves the transformation of textual data into a numerical representation, followed by vectorization, which can result in the creation of large and sparse datasets [104]. These datasets may expand in size, potentially reaching the dimensions of the entire vocabulary, thus incurring significant computational costs.

Feature extraction plays a crucial role in addressing this issue by reducing the dimensionality of the data. This process reshapes the input space into a lower-dimensional form while retaining the most relevant information [105]. The primary objective is to distill a smaller, more focused set of features, emphasizing sentiment-related data while eliminating non-relevant elements [105]. Hence, it is important to choose features carefully to improve classification because the quality of the feature selection may have a direct impact on the accuracy of prediction [72].

There exist several techniques to carry out feature extraction such as TF-IDF, Word2Vec, GloVe, and countvectorizer.

a. TF-IDF

TF-IDF, which stands for "term frequency-inverse document frequency" represents a widely employed methodology for understanding the significance of words within a document [105].

Term Frequency, referred to as TF, quantifies how often a specific term t appears in a given document d. This value increases when t occurs multiple times within the document. TF is computed by determining the ratio of the frequency of term t in document d to the total number of terms found in that specific document [105].

$$TF(t,\ d)\ =\ \frac{Number\ of\ times\ term\ t\ appears\ in\ a\ document\ d}{Total\ number\ terms\ in\ a\ document\ d}$$

While TF offers insights into the frequency of a term t, it may not necessarily reflect its importance. Some terms, such as stop words, may occur frequently but lack significance. Therefore, the Inverse Document Frequency (IDF) comes into play to assess the importance of a term. IDF assigns greater weight to terms that occur rarely in the document d.

$$IDF(t)\ =\ log_e \frac{Total\ number\ of\ documents}{Total\ number\ of\ documents\ with\ the\ term\ t\ in\ it}$$

The final weight TF-IDF for a term *t* in document *d* is calculated as

$$TF\ -\ IDF(t,d)\ =\ T\,F(t,d) \times IDF(t)$$

Among the reviewed papers, 13 of them used TF-IDF as a vectorizer. [53] studied the performance of a combination of five feature extraction and four ML algorithms where the authors concluded that adding TF-IDF weighting in topics modeling enhanced the accuracy and F1-score of all the classifier algorithms, particularly the accuracy of the Ensemble Voting Classifier, which was raised by 26.43% to the level of 0.9504, and its F1-score increased by 32% to a value of 0.95. [55] compared the classifiers' performance and the obtained result proved that the combination of TextBlob and CountVectorizer gives a higher accuracy than VADER and TF-IDF Vectorizer, especially when applying Logistic Regression. All of the papers using TF-IDF vectorizer applied a machine learning algorithm, and three of them [6][59][67] used some extra deep learning algorithms.

b. Word2Vec

Word2Vec has the capability to transform tokenized words into vectors that encapsulate the vocabulary found in textual documents [53]. It is a shallow neural network developed by Google whose training data is based on a dataset gathered from tweets and Facebook posts [106]. Word2Vec has the ability to group highly similar words together in a common vector space. This grouping is achieved by identifying word similarities by applying cosine similarities [40]. Each word within the vocabulary is encoded into a 200-dimensional vector. The sentences are then transformed from a sequence of words to a sequence of vectors, where every word is a 200-dimensional vector. These vectors subsequently serve as input for the neural network [106].

Of the reviewed articles, 7 of them have used Word2Vec in their study [5][11][40][106][69][68]. According to [11] the vectorization model accuracy may be improved by combining these word representations with other lexicons such as WordNet and EmoLex which contain more semantically supported information. The authors propose a method to improve word vector representations by using information from other lexicons (WordNet and EmoLex) to support words with similar meanings and representations. This improved model promotes a more nuanced understanding of emotions and therefore provides valuable feedback to emergency responders for improved crisis management and service enhancement.

c. GloVe

GloVe is a global log-bilinear regression model that provides a vector space representation of words [107]. GloVe was published as an improved method, compared to word2vec, that incorporates global statistics in order to understand not only the word, but the context it is in [107]. GloVe is used by five papers among 60 papers reviewed. All of the papers that use the GloVe model applied Deep Learning algorithms to classify the sentiment.

[8] and [38] have used GloVe as a first layer of CNN that corresponds to the non-static model of embeddings.

According to [5], using GloVe before a CNN algorithm is the most well-performing model with 85% accuracy, compared to other machine learning algorithms. Moreover [51] achieved high results in terms of accuracy while using different machine and deep learning algorithms such as Naïve Bayes, Support Vector Machine, Random Forest and Bi-LSTM.

d. CountVectorizer

The Count Vectorizer is another feature extraction model. It relies on the frequency of each word in the document to transform the provided tweets into vectors [55]. CountVectorizer is building a matrix with each unique word that is represented by a column and each text sample from the used document represented by a row [66].

Among the reviewed papers, 5 of them have used CountVectorizer. [66] evaluated three machine learning algorithms Support Vector Machine, K-Nearest Neighbor, Logistic Regression with the application of CountVectorizer and TF-IDF respectively. The authors of [66] concluded that overall the “model-vectorizer” combination with the highest accuracy was Logistic Regression with CountVectorizer with an 87.2% model accuracy.

[55] proves that Random Forest gives maximum accuracy of 92.91% using CountVectorizer and TextBlob for a bi-class classification.

Classification

After feature extraction, the data will be ready and compatible to be processed into a machine learning or deep learning model. 60% of the papers that have used a machine learning approach have utilized several algorithms in order to compare the performance and the result of each one.

7. The evaluation of Sentiment Analysis on Twitter

The final stage in the sentiment analysis process is evaluation.

Among reviewed papers none of the studies following the lexicon based approach conducted any evaluation test. The evaluations are carried out beforehand when creating the lexicon and they are manually checked by language experts and reviewers [83].

In a machine learning approach, the data is typically divided into two parts: one for training and the other for testing. While the specific ratio may vary between studies, it is common to allocate 70% to 80% of the data for training and the remaining 20% to 30% for testing.

Based on the training data, the testing data are used to check if the classification model is reliable and accurate. Among the reviewed papers, the majority of the machine learning approach retrieved information from the confusion matrix [4]. Confusion matrix is a performance measurement for classification algorithms, where the output may be two or more classes [4].

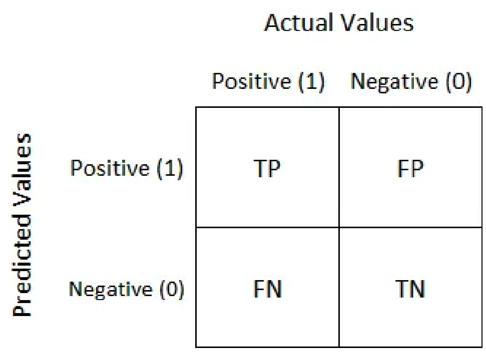

Figure 14. Confusion matrix

Fig. 14 is a presentation of a confusion matrix of a binary classification.

True Positive (TP) - A value predicted as positive and it's true

True Negative (TN) - A value predicted as negative and it's true

False Positive (FP) - A value predicted as positive and it's false

False Negative (FN) - A value predicted as negative and it's false

By using the confusion matrix we may extract four useful metrics that help to measure the performance of the classification.

Precision defines the ratio between correctly predicted positive observations and the total predicted positive observations [7].

$$Precision = \frac{TP}{TP + FP}$$

Recall is defined as the ratio of correctly predicted positive observations and all the observations in actual positive class [7].

$$Recall = \frac{TP}{TP + FN}$$

F1-score is the harmonic mean of precision and recall. It takes both false positives and false negatives into account [108].

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

Accuracy measures the ratio between all correctly predicted observations and the total observations [109].

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Among the reviewed papers that followed a machine learning or deep learning approach presented by 43 papers, 39 (90.7%) of them used accuracy testing based on the confusion matrix to evaluate their classification model.

[93] tries to create a recommendation system on books based on users' reviews on twitter. To evaluate the model, the authors launched a ResQue questionnaire, which is a widely used evaluation framework for recommender systems. The questionnaire was based on book recommendations by the model that they proposed and by Amazon. The aim was to compare these two recommendation systems by measuring criteria such as *accuracy, novelty, intention, content, reflection, sentiment, reflection, diversity, usefulness, satisfaction*.

Three papers out of 42 (7%) do not go through any evaluation process [2][35][73]. Fig. 15 dictates visually the distribution of the evaluation of the papers that have followed the ML/DL approach.

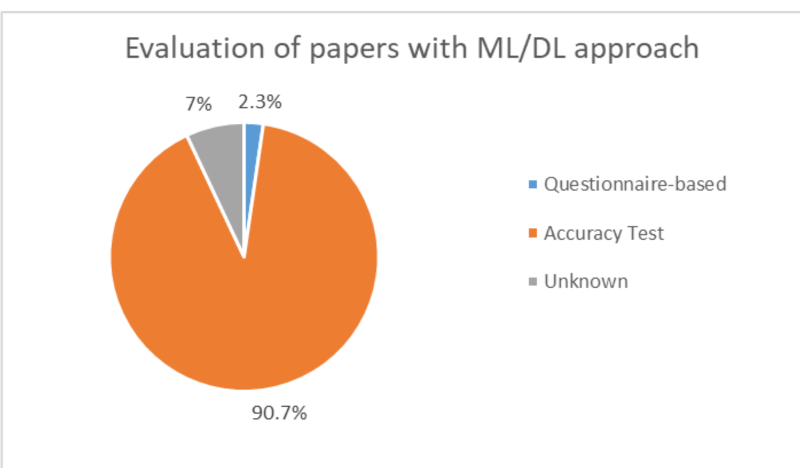

Figure 15. The evaluation of the papers that have used ML/DL approach

Conclusion

In this comprehensive systematic review, we analyzed 60 papers published between 2018 and 2022. The objective was to provide valuable insights for researchers regarding the evolving landscape of sentiment analysis utilization on Twitter during this period, shedding light on its applications and the procedural aspects involved. Notably, a remarkable surge of interest in sentiment analysis was observed, reaching its peak in 2021. The dominant focus of these papers was on employing sentiment analysis within the health domain, driven by the necessity to understand public emotions during the ongoing pandemics. Mental health emerged as another prominent domain of interest.

In terms of analytical methodologies, machine learning and deep learning techniques were favored over traditional lexicon-based approaches, owing to their ability to handle complexity and yield more nuanced results. Numerous studies were undertaken to enhance these models, aiming to capture nuanced sentiments such as 'good' or 'very good', necessitating the expansion of lexicons to encompass a broader spectrum of emotional states. Of special significance was the attention directed towards depression estimation and exploring the correlation between sentiment analysis outcomes and stock prices.

However, it is worth noting that this review lacked an analysis of the methodologies proposed to enhance the performance and accuracy of sentiment analysis. Future studies should pay attention to the ongoing advancements, particularly within the realm of Machine Learning, to assess the latest proposals and methodologies that drive the field forward.

Appendix 1

Paper, Year	Application domain	Objective	Data source	Data collection	Definition of emotion	Emotions detected	Preprocessing	Processing	Classification Approach	Algorithm/ ML Model	Evaluation Method	Contribution type
93, 2022	Recommendation system	Create a book recommendation system	Twitter API		Emotion, Sentiment	10 joy, anger, sadness, fear, shame,fondness, dislike,excitement, relief, and surprise.		ML-Ask	DL	BERT	QB	Improved Method
16, 2021	Natural Disaster	Analyze tweets related to Cyclone Fani	Unknown	Developing a Tweeter crawler	Emotion	3 positive, neutral, and negative	Data cleaning Lemmatization	Word2Vec	ML	NB	AT	Application
8, 2020	Traumatic events	Analyze reactions to mass shooting events and compare the performance of several DL algorithms	Twitter API	Developing a Tweeter crawler	Emotion	5 Anger, Fear, Sadness, Disgust, and Surprise	Uppercase conversion Data cleaning	GloVe	DL	CNN, Bi-LSTM, BERT,	AT	Improved Method
7, 2021	Healthcare	Analyze public sentiment and opinions regarding the COVID-19 pandemic and the vaccination process	Kaggle	Dataset 1: Corona Virus NLP Tweets – Text Classification Dataset 2: COVID-19 World Vaccine Adverse Reactions	Sentiment	3 Positive, negative, neutral	Data cleaning	VADER	ML	SVM,KNN, Multinomial NB, LR	AT	Application
6, 2020	Healthcare	Detect depression among tweets	Twitter API	From 2016 to 2019	Emotion	3 Low, Medium, High level of depression	Data cleaning Tokenization	TF-IDF LIWC	DL, ML	1DCNN, NN, SVM, RF	AT	Application

5, 2022	Healthcare	Analyzing the behavior of the people during the COVID-19 lockdown	Kaggle	Collecting tweets from hashtag's keywords	Emotion	3 Positive, Negative, Neutral	Tokenization	Word2vec	DL	CNN-LSTM	AT	Application
84, 2021	Service	Improve customer experience in the airline industry by analyzing sentiment in tweets, identifying areas for improvement, addressing concerns, and enhancing services.	Kaggle	US airlines tweets database	Sentiment	3 Positive, Negative, Neutral	Data cleaning Tokenization	One-hot-econding	DL	Baseline model, Reduced model, Regularized model, Dropout mode	AT	Application
9, 2022	Social movement	Analyze the sentiments and emotions expressed by people on Twitter regarding the #BlackLivesMatter movement.	Kaggle	Keywords: #BlackLivesMatter Location: Minnesota or Washington D.C	Sentiment	6 Trust, Anger, Anticipation, Fear, Disgust, and Surprise.	Data cleaning Tokenization		DL	CNN-LSTM	AT	Application
2, 2021	Marketing	Analyze user sentiment on Twitter towards prominent beverage brands and provide visualization and future prediction of customer behavior based on categorized tweets.	Twitter API	Keywords: Coca-Cola, Pepsi and Fanta Period: January 1, 2015, to October 1, 2020	Emotion	7 Positive, WeaklyPositive, StronglyPositive, Negative, WeaklyNegative, Strongly Negative, Neutral	Data cleaning		DL, ML	NB, DT, LSTM,		Application
3, 2021	Healthcare	Analyze and evaluate the opinions expressed on Twitter regarding the COVID-19 vaccine and explore the key topics and themes discussed in relation to it.	Twitter API	Keywords: AZD1222, BNT162, BNT162b2, coronavirus vaccine, Covaxin, CoronaVac, covid vaccine, covid19vaccine, covidshield, covid19 vaccine, covidvaccine, covishield, mRNA-1273	Sentiment	3 Positive, Negative, Neutral	Data cleaning Tokenization Lemmatization	VADER TextBlob	LB			Application

11, 2019	Traumatic events	Analyze the sentiments and emotions expressed on Twitter during emergency events and provide valuable feedback to emergency responders for improved crisis management and service enhancement.	Twitter API	Keywords #LasVegas, #LasVegas and #LasVegasShooting Period: October 1st, 2017 toOctober 6th, 2017	Emotion, Sentiment	8 emotions: anger, anticipation, disgust, fear, joy, sadness, surprise, trust 2 sentiments: positive, negative	Lowercase conversion Data cleaning	WordNet NRC Glove Word2vec	LB			Improved Method
100, 2022	Educational technology	Identify and comprehend public perspectives, sentiments, attitudes, and discourses regarding the adoption and integration of augmented reality and virtual reality in education.	Unknown	Keywords: “augmented reality” OR #AR OR augmentedreality; (“augmented reality” OR #AR OR augmentedreality) AND (learn OR teach OR train OR education OR university OR college OR school OR class OR student OR pupil); “virtual reality” OR #VR OR virtualreality; (“virtual reality” OR #VR OR virtualreality) AND (learn OR teach OR train OR education OR university OR college OR school OR class OR student OR pupil) Period: January 2010 to December 2020.	Emotion	The pairs of emotions are: (i) surprise and anticipation, (ii) joy and sadness, (iii) fear and anger as well as (iv) acceptance and disgust.	Data cleaning Lowercase conversion Expansion of abbreviated words	TextBlob VADER NRC	LB			Application
46, 2022	Healthcare	Analyze and classify worry levels in COVID-19-related discussions on Twitter	Twitter API	Keywords: coronavirus, covid-19, #coronavirus, and #covid-19. Period: from January 30 to February 28, 2020; from March 29 to April 29, 2020; from May 10 to June 30, 2020	Emotion	3 “no-worry”, “worry”, and “high-worry”	Data cleaning Lowercase conversion Lemmatization	TF-IDF	ML	Multinomial NB, LR, SVM, RF	AT	Application

47, 2021	Healthcare	Analyze the impact of the COVID-19 pandemics	Twitter API	Dataset 1 #covid19, #coronavisrus Dataset 2 #covid19australia, #covid19india, #covid19usa, #covid19uk, #covid19russia, #covid19china Dataset 3 #StayHome, #StayHomeStaySafe, #SocialDistancing, #QuarantineandChill Dataset 4 : Dataset 1 + Dataset 2 + Dataset 3	Sentiment	3 Positive, Negative, Neutral	Data cleaning Lowercase conversion Tokenization	TF-IDF	ML	Multinomial NB, LR, SVM, RF, DT, XGBoost	AT	Application
62, 2022	Healthcare	Distinguish between different types of suicide-related content on Twitter and developing the TWISCO dataset.	Twitter API	Period: July 2021 - September 2021	Emotion	7 Anger, Fear, Disgust, Surprise, Sadness and Joy, Neutral	Data cleaning	LIWC RFE Paladin VAD Annotation	DL	Max. Entropy Class., Vanilla LSTM, ALBERT, ALBERT 2 LSTMs, Graph CNN, Feature GCN	AT	Improved Method
48, 2021	Healthcare	Detect the informative tweets of COVID-19 and classify their sentiments	Public dataset	Dataset: COVID-19 pandemic (2020), WNUT-2020	Sentiment	3 Positive, Negative, Neutral	Data cleaning	TextBlob	DL	RoBERTa, CT-BERT, BERTweet, Majority Voting	AT	Application
49, 2021	Healthcare	Analyze and classify emotions regarding COVID-19 vaccine	Twitter API	Keywords: China's COVID-19, Chinese COVID-19, Sinopharm COVID-19 vaccine, Sinovac COVID-19 vaccine Sinopharm vaccine, Sinovac vaccine Period: December 15th, 2020 to February 10th, 2021 Language: English	Emotion	3 Positive, Negative, Neutral	Remove duplicated tweets Remove non-English symbols Data cleaning	VADER	LB			Application

63, 2022	Healthcare	Detect depression	Kaggle	Keywords: #depressed, #depression, #loneliness, #hopelessness, #antidepressants, #suicide	Emotion	3 Positive, Negative, Neutral	Remove duplicated tweets Data cleaning	Text augmentation VADER	DL	BERT	AT	Application
64, 2019	Healthcare	Detect depression	Twitter API	Keywords: Depression, Sadness, Tired, Guilt, Suicidal, Anxiety, Mental health, Random	Emotion	2 Depressed, Non depressed	Data cleaning Tokenization Lemmetization	SentiStrength	DL	Back, Propagation NN	AT	Application
77, 2022	Finance	Explore the influence of Twitter polarity on stock market movements and identifies correlations between influential Twitter accounts' sentiments and market behavior	Twitter API		Emotion	3 Positive, Negative, Neutral	Data cleaning Tokenization Stemming	Bing Liu Sentiment 140 NRC Affin SenticNet	LB			Application
37, 2019	Recommend ation system	Generate recommendations that align with users' shared sentiments and emotions on specific topics	Twitter API	Keywords: #Syria, #DonaldTrump, #SchoolShooting, #Christmas2017, #NewYear, #ValentinesDay2018, #Terrorism, #olympicgames2018, #WomensDay, #Oscars2018.	Emotion, Sentiment	3 Sentiments: Positive, Negative and Neutral 6 Emotions: Anger, Disgust, Fear, Joy, Sadness, Surprise and Neutral	Data Cleaning		ML	NB	AT	Application
72, 2022	Traumatic events	Identifying the emotions related to the Ukraine-Russia war topic	Twitter API and Kaggle		Emotion	6 anger, fear, surprise, joy, love, and sadness	Data cleaning Segmentation Tokenization Stemming	Word n-grams TF-IDF CountVectoriz er	ML	LR DT NB SVC	AT	Improved Method
50, 2022	Healthcare	Understand changing perceptions of mental health over 7 years, focusing on the COVID pandemic's impact	Snscrap e	Keywords #mentalhealth, #anxiety, #depression, #ptsd Period: January 2015 - December 2021. Geolocation: United States.	Sentiment	3 Positive, Negative, Neutral	Data cleaning Lowercase conversion Lemmatize	TextBlob	LB			Application

102, 2021	Tourism	Explore opinions on halal tourism	Twitter API	Period: October 2008 to October 2018	Emotion	8 fear, disgust, anger, sadness, trust, joy, surprise, and anticipation	Data cleaning Lowercase conversion Stemming	NRC	LB			Application
75, 2022	Politics	Explore the feelings of fear and anger during the 2016 US presidential election and Brexit	Unkno wn		Emotion	4 Fear, anxiety, anger and disgust	Tokenization Lemmetization	WordNet NRC	DL	NN	AT	Application
40, 2020	Cinematogra phy	Analyze user sentiments in IMDB movie reviews	Public dataset	Dataset: Large Movie Review dataset of Stanford University	Feeling	2 positive, negative	Data cleaning	Bag of Words Word2Vec	DL, ML	SVM, RF, Artificial NN	AT	Application
96, 2021	Non-Govern mental Organization	Provide assistance through NGOs based on the people's emotional conditions.	Kaggle	Dataset: Emotion in Text data set,' 'ISEAR,' 'Amazon Reviews,' 'US airlines data set,' 'EmmoBank,' 'TREC'	Emotion	3 Positive, Negative, Neutral	Data cleaning	TF-IDF	ML	NB, RF, DT, XGBClassif ier	AT	Application
43, 2020	Marketing	Improving sentiment analysis for customer reviews	Public dataset	Datasets: Drugs, Electronics, Mobile phones	Sentiment	7 extremely positive, very positive, positive, neutral, negative, very negative, and extremely negative	Data cleaning Tokenization Lemmatization	SWN Lexicon Linguistic Hedge	LB	Fuzzy Based Rules	AT	Improved Method
76, 2021	Politics	Analyze emotional aspects of the 2020 U.S. election	Unkno wn	Tweets published by: @realDonaldTrump and @JoeBiden Period: during 2021	Emotion	8 trust, anger, disgust, fear, anticipation, joy, sadness, and surprise.		SenticNet	LB			Application
97, 2021	Sport	Understand fans' sentiments during FIFA World Cup.	Twitter API	2018 FIFA world-cup	Sentiment	3 positive, negative, and neutral	Data cleaning Corrected spelling mistakes	GloVe	DL	CNN-LST M	AT	Application

32, 2021	Environment	Assess global public awareness of climate change by analyzing sentiment in tweets.	Kaggle	Tweets concerning climate change	Emotion	3 positive, negative, and neutral	Data cleaning Tokenization		DL	Bi-LSTM	AT	Application
94, 2018	Traumatic events	Understand social media behavior during terror-related events	Twitter API	Keywords: error, terrorism, terrorist, ISIS) from Period: 18 May 2017 to 5 June 2017 29 September 17 to 7 October 2017	Sentiment	3 positive, negative, and neutral	Data cleaning	SentiWordNet SentiStrength	LB			Application
65, 2019	Marketing	Create a web application for mining skincare opinions, benefiting both consumers and entrepreneurs	Twitter API	Keywords “#skincare”	Sentiment	5 very positive, positive, moderate, negative, and very negative	Data cleaning	SentiWordNet	ML	NB, SVM	AT	Application
41, 2022	Sustainable energy	Understand how sentiments have changed during the Ukrainian-Russo conflict and its impact on global energy policies	Twitter API	Keyword: “green energy” Period: 16 February 2022 to 3 March 2022	Emotion, Feeling	2 Feelings: negative and positive 8 Emotions: anger, anticipation, disgust, fear, joy, sadness, surprise, and trust	Data cleaning	NRC	LB			Application
106, 2018	Politics	Analyze public sentiments on political topics, aiding better decision-making in elections and campaign	Twitter API	Web Scrapping	Sentiment		Data cleaning Tokenization Part-of-Speech	Word2vec	DL	Artificial NN	AT	Application
35, 2020	Finance	Extract sentiment and technical features for predicting stock volatility of specific companies	Unknown	Keywords: $AAPL, $AMZN, $MSFT Period: 1 August 2009 and 1 August 2019	Emotion	3 positive, negative, and neutral	Data cleaning	CountVectorizer	DL, ML	NB, Recurrent RNN		Application

73, 2022	Traumatic events	Understand emotions related to the Russia-Ukraine conflict	Twitter API	Post February 24, 2022	Emotion	5 Joy, Surprise, Anger, Sadness, Fear, Disguise	Data cleaning		ML	NB		Application
95, 2021	Marketing	Assess consumer perceptions towards selected automotive brands on Twitter	Unkno wn	Keywords: #VW, #Peugeot, #Citroen, #Kia, #Skoda, and #Toyota Period: 23 October 2018 to 30 October 2018	Emotion	2 Positive, negative	Data cleaning	VADER	LB			Application
51, 2022	Healthcare	Analyze sentiments in COVID-19-related tweets	Kaggle		Sentiment	5 Extremely Negative, Negative, Neutral, Positive, and Extremely Positive	Lowercase conversion Expanded contractions Data cleaning	GloVe FastText	DL, ML	NB, SVM, RF, Bi-LSTM	AT	Improved Method
52, 2022	Healthcare	Analyze sentiments in COVID-19 vaccination tweets	Kaggle	Covid-19 Vaccine Tweets Sentiment	Sentiment	3 positive, negative, and neutral			DL	Majority Voting, Stacking, BERT, RoBERTa	AT	Improved Method
53, 2022	Healthcare	Predict tweet popularity based on emotional intensity during different phases of the Covid-19 pandemic	Twitter API	Keywords: Covid, corona, pandemic (and synonyms) Period: January 20, 2020, to May 29, 2021. Only tweets that had at least one retweet	Emotion	4 fear, anger, joy, sadness	Data cleaning Lowercase conversion Removed stop words Tokenization	- Topic Modeling LDA - Topics + TF-IDF - BOW by TF-IDF - Doc2vec - Doc2vector + TF-IDF	ML	RF, Stochastic Gradient Descent, LR, Ensemble Voting	AT	Improved Method

98, 2022	Healthcare	Analyze how ASD and COVID-19 are discussed on Twitter to understand people's attitudes	Twitter API	Keywords: # autism; #ASD; #ActuallyAutistic; #AutismSpeaks; #AutismParent; #Autchat; #AutismAwareness; #WorldAutismAwarenessDay; aspergers; autism; autistic; autism spectrum; Disorder; spectrumDisorder; autism disorder	Sentiment	3 positive, negative, and neutral	Tokenization Lemmatization Stemming	AFiNN	LB			Application
54, 2021	Healthcare	Analyze sentiments in COVID-19-related tweets	Twitter API	Keywords: #COVID 19 and #coronavirus Period: March 21 2020, to April 23 2020	Sentiment	2 Positive, Negative		TextBlob	ML	NB, KNN, CNN, SVM	AT	Application
55, 2021	Healthcare	Analyze sentiments in COVID-19-related tweets	IEEE	Keywords: #2019-ncov, #ncov, #covid_19, #sarscov2, #covid19, #ncov2019, #corona, #covid, #lockdown Period: 1 January 2021	Sentiment	2 Positive, Negative 3 positive, negative, and neutral		TF-IDF CountVectoriz er TextBlob VADER	ML	XGBoost, SVM, RF, DT, Multinomial NB, LR	AT	Application
66, 2021	Marketing	Analyze sentiment of product reviews onTwitter	Twitter API		Sentiment	2 Positive, Negative	Data cleaning Remove tweets < 02 words Lowercase conversion	TF-IDF CountVectoriz er	ML	SVM, KNN, LR	AT	Application
56, 2021	Healthcare	Analyze sentiments in COVID-19-related tweets	Twitter API		Sentiment	3 positive, negative, and neutral	Lowercase conversion Tokenization Data cleaning	VADER	LB			Application
57, 2021	Healthcare	Analyze sentiments in COVID-19-related tweets	Kaggle	COVID-19 dataset	Sentiment	3 positive, negative, and neutral	Data cleaning Tokenization	Bag of Words TF-IDF	ML	XGBoost, SVM, RF, DT, Multinomial NB, LR	AT	Application
58, 2021	Healthcare	Analyze sentiments in COVID-19-related tweets	Twitter API		Sentiment	3 positive, negative, and neutral	Data cleaning Tokenization	TextBlob	LB			Application

92, 2021	Marketing	Evaluate websites based on customer reviews	Kaggle		Sentiment	3 positive, negative, and neutral	Data cleaning Tokenization Stemming	TF-IDF	ML	LR, NB, SVM, DT, RF, KNN	AT	Application
69, 2020	Service	Analyze sentiments in tweets about Uber and Ola cab services to better understand user needs	Twitter API		Sentiment	2 Positive, Negative	Data cleaning	Word2Vec	DL	Non-Deep Feed Forward NN, Deep Feed Forward NN, CNN	AT	Application
67, 2020	Service	Analyze public sentiment towards 5G technology	Twint	Keyword: #5G Period: 01 January 2019 to 30 May 2020	Sentiment	2 Positive, Negative	Data cleaning	TextBlob TF-IDF	DL, ML	NB, SVM, KNN, RNN, LSTM	AT	Application
70, 2019	Service	Understand public opinions on the sharing economy business model of Airbnb in the hospitality sector	Twitter API	Keyword: Airbnb Period: January 7th to March 28th, 2017	Sentiment	3 positive, negative, and neutral	Data cleaning Tokenization Stemming		ML	NB	AT	Application
68, 2019	Service	Analyze public sentiment towards 4G Smartfren technology	GetOld Tweets 3		Sentiment	2 Positive, Negative	Data cleaning Tokenization Stemming	Word2Vec	DL	CNN	AT	Application
71, 2019	Service	Understand passenger sentiments about airplane services	Unkno wn		Sentiment	3 positive, negative, and neutral	Data cleaning Lowercase	CountVectoriz er	ML	RF, LR	AT	Application
42, 2019	Service	Identify sentiments that consumers have about health insurance	Twitter API	Keywords: “health insurance,” “health plan,” “health provider” or “doctor” Period: 1 November 2016 until 31 January 2017	Attitude, Emotion	2 attitudes: positive, negative 8 Emotion: anger,anticipatio n, disgust, fear, joy, sadness, surprise,and trust	Data cleaning Lemmatization	Bag of Words NRC	LB			Application

59, 2021	Healthcare	Predict the symptoms effect based on Sentiment Analysis of tweets related to COVID-19	OpenIC PSR		Sentiment	2 fear, sadness	Data cleaning Correct spelling	TF-IDF	DL	CNN	AT	Improved Method
60, 2022	Healthcare	Analyze public sentiment regarding the new normal during the COVID-19 pandemic	Twitter API	Keywords: “face mask”, “hand hygiene”, “Sejahtera scan”, “new normal COVID-19”, “stay at home”, “work from home” and “social distancing” Period: March to June 2022	Sentiment	2 Positive, Negative	Data cleaning Tokenization Part-Of-Speech Lemmatization	Bag of Words	ML	NB	AT	Application
61, 2022	Healthcare	Analyze sentiments in COVID-19-related tweets	Kaggle		Sentiment	4 joy, fear, anger, and sadness	Data cleaning Part-of-Speech Lemmatization	TF-IDF	ML	NB, Multinomial NB,Bernoulli NB	AT	Application
38, 2022	Finance	Analyze public sentiments and emotions on Twitter related to non-fungible tokens (NFTs) to understand their impact on the NFT market.	Twitter API	Period: 15th October 2021 to 20th June 2022.	Emotion, Sentiment	8 emotions: anger, fear, anticipation, trust, surprise, sadness, joy, and disgust 2 sentiments: positive, negative	Data cleaning Stemming Lemmatization	GloVe	DL	Trac (Google Lib)	AT	Improved Method
39, 2019	Healthcare	Develop an effective method for assessing depression levels in Twitter	Twitter API	Keywords: #abuse, #anxiety, #addict, #addiction, and #bullying	Emotion, Sentiment	8 emotions: anger, anticipation, disgust, fear, joy, sadness, surprise, and trust 2 sentiments: positive, negative	Data Cleaning	AFINN BING NRC	LB			Improved Method

References

1. https://backlinko.com/twitter-users

2. Ali, Tahir Mohammad, Ali Nawaz, and Attique Ur Rehman. "An Integrated Framework for Sentiment Analysis of Giant Cold Drinks Companies." In *2021 International Conference on Innovative Computing (ICIC)*, pp. 1-5. IEEE, 2021.

3. Rahul, Kumar, Bhanu Raj Jindal, Kulvinder Singh, and Priyanka Meel. "Analysing public sentiments regarding COVID-19 vaccine on twitter." In *2021 7th international conference on advanced computing and communication systems (ICACCS)*, vol. 1, pp. 488-493. IEEE, 2021.

4. https://towardsdatascience.com/understanding-confusion-matrix-a9ad42dcfd62

5. Dumka, Ankur, Parag Verma, Rajesh Singh, Anil Kumar Bisht, Divya Anand, Hani Moaiteq Aljahdali, Irene Delgado Noya, and Silvia Aparicio Obregón. "A novel deep learning based healthcare model for COVID-19 pandemic stress analysis." *Computers, Materials & Continua* 72, no. 3 (2022): 6029-6044.

6. Mustafa, Raza Ul, Noman Ashraf, Fahad Shabbir Ahmed, Javed Ferzund, Basit Shahzad, and Alexander Gelbukh. "A multiclass depression detection in social media based on sentiment analysis." In *17th International Conference on Information Technology–New Generations (ITNG 2020)*, pp. 659-662. Springer International Publishing, 2020.

7. Rahman, Tasnia, and Sakifa Aktar. "A Machine Learning Approach to Track COVID-19 Pandemic using Sentiment Analysis." In *2021 3rd International Conference on Electrical & Electronic Engineering (ICEEE)*, pp. 145-148. IEEE, 2021.

8. Harb, Jonathas GD, Régis Ebeling, and Karin Becker. "A framework to analyze the emotional reactions to mass violent events on Twitter and influential factors." *Information Processing & Management* 57.6 (2020): 102372.

9. Rani, Shalli, Ali Kashif Bashir, Adi Alhudhaif, Deepika Koundal, and Emine Selda Gunduz. "An efficient CNN-LSTM model for sentiment detection in# BlackLivesMatter." *Expert Systems with Applications* 193 (2022): 116256.

10. Hasan, Asif, Tripti Sharma, Azizuddin Khan, and Mohammed Hasan Ali Al-Abyadh. "Analysing hate speech against migrants and women through tweets using ensembled deep learning model." *Computational Intelligence and Neuroscience* 2022 (2022).

11. Singh, Neha, Nirmalya Roy, and Aryya Gangopadhyay. "Analyzing the emotions of crowd for improving the emergency response services." *Pervasive and mobile computing* 58 (2019): 101018.

12. Mejova, Yelena. "Sentiment analysis: An overview." *University of Iowa, Computer Science Department* (2009).

13. Pang, Bo, and Lillian Lee. "Opinion mining and sentiment analysis." *Foundations and Trends® in information retrieval* 2.1–2 (2008): 1-135.

14. Liu, B. (2012). Sentiment analysis and opinion mining. Synthesis lectures on human language technologies, 5(1), 1–167.

15. Nandwani, Pansy, and Rupali Verma. "A review on sentiment analysis and emotion detection from text." *Social Network Analysis and Mining* 11, no. 1 (2021): 81.

16. Ansar, Wazib, Saptarsi Goswami, and Amit Kumar Das. "A data science approach to analysis of tweets based on cyclone Fani." In *Data Management, Analytics and Innovation: Proceedings of ICDMAI 2020, Volume 1*, pp. 243-261. Springer Singapore, 2021.

17. Singh, Neha, Nirmalya Roy, and Aryya Gangopadhyay. "Analyzing the emotions of crowd for improving the emergency response services." *Pervasive and mobile computing* 58 (2019): 101018.

18. Pati, Debajyoti, and Lesa N. Lorusso. "How to write a systematic review of the literature." *HERD: Health Environments Research & Design Journal* 11, no. 1 (2018): 15-30.

19. Xiao, Yu, and Maria Watson. "Guidance on conducting a systematic literature review." *Journal of planning education and research* 39, no. 1 (2019): 93-112.

20. https://www.techtarget.com/whatis/definition/Twitter

21. https://developer.twitter.com/en/docs/twitter-api

22. https://developer.twitter.com/en

23. https://publer.io/blog/twitter-enterprise-api-price-may-kill-smaller-apps

24. https://www.theverge.com/2023/3/30/23662832/twitter-api-tiers-free-bot-novelty-accounts-basic-enterprice-monthly-price

25. https://mashable.com/article/twitter-new-api-pricing-kills-indie-developer-apps

26. https://www.oii.ox.ac.uk/news-events/news/appeal-to-twitter-make-the-data-accessible/

27. Garg, Samarth, Divyansh Singh Panwar, Aakansha Gupta, and Rahul Katarya. "A literature review on sentiment analysis techniques involving social media platforms." In *2020 Sixth International Conference on Parallel, Distributed and Grid Computing (PDGC)*, pp. 254-259. IEEE, 2020.

28. Nassr, Zineb, Nawal Sael, and Faouzia Benabbou. "A comparative study of sentiment analysis approaches." *Proceedings of the 4th International Conference on Smart City Applications*. 2019.

29. Kumar, Akshi, and Arunima Jaiswal. "Systematic literature review of sentiment analysis on Twitter using soft computing techniques." *Concurrency and Computation: Practice and Experience* 32.1 (2020): e5107.

30. Kusal, Sheetal, et al. "AI based emotion detection for textual big data: techniques and contribution." *Big Data and Cognitive Computing* 5.3 (2021): 43.

31. Drus, Zulfadzli, and Haliyana Khalid. "Sentiment analysis in social media and its application: Systematic literature review." *Procedia Computer Science* 161 (2019): 707-714.

32. THUKRAL, Tishya, VARSHNEY, Ashwani, et GAUR, Vibha. Intensity quantification of public opinion and emotion analysis on climate change. *International Journal of Advanced Technology and Engineering Exploration*, 2021, vol. 8, no 83, p. 1351.

33. BROAD, Charlie Dunbar. Emotion and sentiment. *The Journal of Aesthetics and Art Criticism*, 1954, vol. 13, no 2, p. 203-214.

34. Lakshmi Praveena, T., and N. V. Muthu Lakshmi. "Predicting Sentiment and Emotion on ASD Using Machine Learning and Deep Neural Networks Based on Twitter Data." *Intelligent Sustainable Systems: Selected Papers of WorldS4 2021, Volume 1*. Springer Singapore, 2022.

35. Archary, Diren, and Marijke Coetzee. "Predicting Stock Price Movement with social media and Deep Learning." *2020 International Conference on Artificial Intelligence, Big Data, Computing and Data Communication Systems (icABCD)*. IEEE, 2020.

36. Reagan, Kenna, Aparna S. Varde, and Lei Xie. "Evolving Perceptions of Mental Health on Social Media and their Medical Impacts." *2022 IEEE International Conference on Big Data (Big Data)*. IEEE, 2022.

37. Sailunaz, Kashfia, and Reda Alhajj. "Emotion and sentiment analysis from Twitter text." *Journal of Computational Science* 36 (2019): 101003.

38. Qian, Cheng, et al. "Understanding public opinions on social media for financial sentiment analysis using AI-based techniques." *Information Processing & Management* 59.6 (2022): 103098.

39. Stephen, Jini Jojo, and P. Prabu. "Detecting the magnitude of depression in Twitter users using sentiment analysis." *International Journal of Electrical and Computer Engineering* 9.4 (2019): 3247.

40. Rachiraju, Sai Chandra, and Madamala Revanth. "Feature extraction and classification of movie reviews using advanced machine learning models." *2020 4th international conference on intelligent computing and control systems (ICICCS)*. IEEE, 2020.

41. Ibar-Alonso, Raquel, Raquel Quiroga-García, and Mar Arenas-Parra. "Opinion mining of green energy sentiment: A russia-ukraine conflict analysis." *Mathematics* 10.14 (2022): 2532.

42. van den Broek-Altenburg, Eline M., and Adam J. Atherly. "Using social media to identify consumers' sentiments towards attributes of health insurance during enrollment season." *Applied Sciences* 9.10 (2019): 2035.

43. Khattak, Asad, et al. "Fine-grained sentiment analysis for measuring customer satisfaction using an extended set of fuzzy linguistic hedges." (2020).

44. Jayakody, J. P. U. S. D., and B. T. G. S. Kumara. "Sentiment analysis on product reviews on twitter using Machine Learning Approaches." *2021 International Conference on Decision Aid Sciences and Application (DASA)*. IEEE, 2021.

45. Babčanová, Dagmar, et al. "Qualitative and quantitative analysis of social network data intended for brand management." *Wireless Networks* 27.3 (2021): 1693-1700.

46. Alharbi, Tahani Soud, and Fethi Fkih. "Building and Testing Fine-Grained Dataset of COVID-19 Tweets for Worry Prediction." *International Journal of Advanced Computer Science and Applications* 13.8 (2022).

47. Banik, Sagar, et al. "Classification of COVID19 Tweets based on sentimental analysis." *2021 International Conference on Computer Communication and Informatics (ICCCI)*. IEEE, 2021.

48. Malla, SreeJagadeesh, and P. J. A. Alphonse. "COVID-19 outbreak: An ensemble pre-trained deep learning model for detecting informative tweets." *Applied Soft Computing* 107 (2021): 107495.

49. Xu, Han, et al. "COVID-19 vaccine sensing: Sentiment analysis from Twitter data." *2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2021.

50. Reagan, Kenna, Aparna S. Varde, and Lei Xie. "Evolving Perceptions of Mental Health on Social Media and their Medical Impacts." *2022 IEEE International Conference on Big Data (Big Data)*. IEEE, 2022.

51. Bhandari, Aakash, et al. "Sentiment analysis of COVID-19 tweets: Leveraging stacked word embedding representation for identifying distinct classes within a sentiment." *International Conference on Artificial Intelligence and Big Data in Digital Era*. Cham: Springer International Publishing, 2021.

52. Ismail, Qusai, et al. "Sentiment analysis of covid-19 vaccination responses from twitter using ensemble learning." *2022 13th International Conference on Information and Communication Systems (ICICS)*. IEEE, 2022.

53. Mahdikhani, Maryam. "Predicting the popularity of tweets by analyzing public opinion and emotions in different stages of Covid-19 pandemic." *International Journal of Information Management Data Insights* 2.1 (2022): 100053.

54. Baker, Oras, et al. "Twitter sentiment analysis using machine learning algorithms for COVID-19 outbreak in New Zealand." *2021 IEEE 11th International Conference on System Engineering and Technology (ICSET)*. IEEE, 2021.

55. Shanthi, S., et al. "Machine learning based twitter sentiment analysis on COVID-19." *2021 5th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*. IEEE, 2021.

56. Dumre, Rutvik, Shubham Mishra, and Aashka Dave. "Twitter Sentiment Analysis on COVID-19." *2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT)*. IEEE, 2021.

57. Mahbub, Nosin Ibna, et al. "Sentiment analysis of microblogging dataset on coronavirus pandemic." *2021 5th International Conference on Electrical Information and Communication Technology (EICT)*. IEEE, 2021.

58. Afroz, Nida, et al. "Sentiment analysis of COVID-19 nationwide lockdown effect in India." *2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS)*. IEEE, 2021.

59. Phan, Huyen Trang, et al. "Tweet sentiment analysis for predicting the symptoms effect level regarding COVID-19." *2021 IEEE International conference on fuzzy systems (FUZZ-IEEE)*. IEEE, 2021.

60. Samsudin, Siti Hajar Aishah, et al. "Sentiment Analysis on Acceptance of New Normal in COVID-19 Pandemic using Naïve Bayes Algorithm." *International Journal of Advanced Computer Science and Applications* 13.9 (2022).

61. Ressan, Murtadha B., and Rehab F. Hassan. "Naive-Bayes family for sentiment analysis during COVID-19 pandemic and classification tweets." *Indonesian Journal of Electrical Engineering and Computer Science* 28.1 (2022): 375.

62. Schoene, Annika M., et al. "Classifying suicide-related content and emotions on Twitter using Graph Convolutional Neural Networks." *IEEE Transactions on Affective Computing* (2022).

63. Hasan, Mushrifah, and Sunil Ghane. "Data-driven Depression Detection System for Textual Data on Twitter using Deep Learning." *2022 2nd Asian Conference on Innovation in Technology (ASIANCON)*. IEEE, 2022.

64. Biradar, Abhilash, and Shashikumar G. Totad. "Detecting depression in social media posts using machine learning." *International conference on recent trends in image processing and pattern recognition*. Singapore: Springer Singapore, 2018.

65. Pugsee, Pakawan, Vasinee Nussiri, and Wansiri Kittirungruang. "Opinion mining for skin care products on twitter." *Soft Computing in Data Science: 4th International Conference, SCDS 2018, Bangkok, Thailand, August 15-16, 2018, Proceedings 4*. Springer Singapore, 2019.

66. Jayakody, J. P. U. S. D., and B. T. G. S. Kumara. "Sentiment analysis on product reviews on twitter using Machine Learning Approaches." *2021 International Conference on Decision Aid Sciences and Application (DASA)*. IEEE, 2021.

67. Seçkin, Tuğkan, and Zeynep Hilal Kilimci. "The evaluation of 5G technology from sentiment analysis perspective in Twitter." *2020 Innovations in Intelligent Systems and Applications Conference (ASYU)*. IEEE, 2020.

68. Aldiansyah, Muhammad Radifan, and Priyo Sidik Sasongko. "Twitter Sentiment Analysis About Public Opinion on 4G Smartfren Network Services Using Convolutional Neural Network." *2019 3rd International Conference on Informatics and Computational Sciences (ICICoS)*. IEEE, 2019.

69. Indulkar, Yash, and Abhijit Patil. "Sentiment Analysis of Uber & Ola using Deep Learning." *2020 International Conference on Smart Electronics and Communication (ICOSEC)*. IEEE, 2020.

70. Alamsyah, Andry, Wachda Yuniar Rochmah, and Ditya Dwi Adhi Nugroho. "Understanding public opinion towards new sharing economy business model using content analysis." *2019 International Conference on Information Management and Technology (ICIMTech)*. Vol. 1. IEEE, 2019.

71. Pandey, Hari Mohan, et al. "Twitter-based opinion mining for flight service utilizing machine learning." *Informatica (Slovenia)* 43.3 (2019): 381-386.

72. Al Maruf, Abdullah, et al. "Emotion Detection from Text and Sentiment Analysis of Ukraine Russia War using Machine Learning Technique." *International Journal of Advanced Computer Science and Applications* 13.12 (2022).

73. Garcia, Manuel B., and Armi Cunanan-Yabut. "Public sentiment and emotion analyses of Twitter data on the 2022 Russian invasion of Ukraine." *2022 9th international conference on information technology, computer, and electrical engineering (ICITACEE)*. IEEE, 2022.

74. Kostakos, Panos, et al. "Meta-terrorism: identifying linguistic patterns in public discourse after an attack." *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*. IEEE, 2018.

75. Friedrichs, Jörg, Niklas Stoehr, and Giuliano Formisano. "Fear-anger contests: Governmental and populist politics of emotion." *Online Social Networks and Media* 32 (2022): 100240.

76. Valle-Cruz, David, Asdrubal Lopez-Chau, and Rodrigo Sandoval-Almazan. "How much do Twitter posts affect voters? Analysis of the multi-emotional charge with affective computing in political campaigns." *DG. O2021: The 22nd Annual International Conference on Digital Government Research*. 2021.

77. Valle-Cruz, David, et al. "Does twitter affect stock market decisions? financial sentiment analysis during pandemics: A comparative study of the h1n1 and the covid-19 periods." *Cognitive computation* (2022): 1-16.

78. Sarlan, Aliza, Chayanit Nadam, and Shuib Basri. "Twitter sentiment analysis." *Proceedings of the 6th International conference on Information Technology and Multimedia*. IEEE, 2014.

79. Kolchyna, Olga, et al. "Twitter sentiment analysis: Lexicon method, machine learning method and their combination." *arXiv preprint arXiv:1507.00955* (2015).

80. Bonta, Venkateswarlu, Nandhini Kumaresh, and N. Janardhan. "A comprehensive study on lexicon based approaches for sentiment analysis." *Asian Journal of Computer Science and Technology* 8.S2 (2019): 1-6.

81. Boyd, Ryan L., et al. "The development and psychometric properties of LIWC-22." *Austin, TX: University of Texas at Austin* (2022): 1-47.

82. Thelwall, Mike. "The Heart and soul of the web? Sentiment strength detection in the social web with SentiStrength." *Cyberemotions: Collective emotions in cyberspace* (2017): 119-134.

83. https://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm

84. Gupta, Meenu, Rakesh Kumar, Harshit Walia, and Gagandeep Kaur. "Airlines based twitter sentiment analysis using deep learning." In *2021 5th International Conference on Information Systems and Computer Networks (ISCON)*, pp. 1-6. IEEE, 2021.

85. Chen, Yixing, and JingTao Yao. "Sentiment analysis using part-of-speech-based feature extraction and game-theoretic rough sets." *2021 International Conference on Data Mining Workshops (ICDMW)*. IEEE, 2021.

86. Hamad, Rebeen Ali, Saeed M. Alqahtani, and Mercedes Torres Torres. "Emotion and polarity prediction from Twitter." *2017 Computing Conference*. IEEE, 2017.

87. Parveen, Huma, and Shikha Pandey. "Sentiment analysis on Twitter Data-set using Naive Bayes algorithm." *2016 2nd international conference on applied and theoretical computing and communication technology (iCATccT)*. IEEE, 2016.

88. Rahat, Abdul Mohaimin, Abdul Kahir, and Abu Kaisar Mohammad Masum. "Comparison of Naive Bayes and SVM Algorithm based on sentiment analysis using review dataset." *2019 8th International Conference System Modeling and Advancement in Research Trends (SMART)*. IEEE, 2019.

89. M. Ahmad, S. Aftab, and I. Ali, "Sentiment Analysis of Tweets using SVM," Int J Comput Appl, vol. 177, no. 5, pp. 25–29, Nov. 2017, doi: 10.5120/IJCA2017915758

90. Tyagi, Abhilasha, and Naresh Sharma. "Sentiment analysis using logistic regression and effective word score heuristic." *International Journal of Engineering and Technology (UAE)* 7.2 (2018): 20-23.

91. Hota, Soudamini, and Sudhir Pathak. "KNN classifier based approach for multi-class sentiment analysis of twitter data." *Int. J. Eng. Technol* 7.3 (2018): 1372-1375.

92. Ganesh, R., et al. "Website Evaluation Using Opinion Mining." *2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA)*. IEEE, 2021.

93. Fujimoto, Takumi, and Harumi Murakami. "A Book Recommendation System Considering Contents and Emotions of User Interests." *2022 12th International Congress on Advanced Applied Informatics (IIAI-AAI)*. IEEE, 2022.

94. Kostakos, Panos, et al. "Meta-terrorism: identifying linguistic patterns in public discourse after an attack." *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*. IEEE, 2018.

95. Babčanová, Dagmar, et al. "Qualitative and quantitative analysis of social network data intended for brand management." *Wireless Networks* 27.3 (2021): 1693-1700.

96. Hemanth, D. J. "Finding State of Mind Through Emotion and Sentiment Analysis of the Twitter Text." *Advances in Parallel Computing Technologies and Applications* 40 (2021): 244.

97. Hegde, Siddhanth U., A. S. Zaiba, and Y. Nagaraju. "Hybrid cnn-lstm model with glove word vector for sentiment analysis on football specific tweets." *2021 international conference on advances in electrical, computing, communication and sustainable technologies (ICAECT)*. IEEE, 2021.

98. Corti, Luca, et al. "Social media analysis of Twitter tweets related to ASD in 2019–2020, with particular attention to COVID-19: topic modelling and sentiment analysis." *Journal of big Data* 9.1 (2022): 113.

99. Plisson, Joël, Nada Lavrac, and Dunja Mladenic. "A rule based approach to word lemmatization." *Proceedings of IS*. Vol. 3. 2004.

100. Lampropoulos, Georgios, et al. "Augmented reality and virtual reality in education: Public perspectives, sentiments, attitudes, and discourses." *Education Sciences* 12.11 (2022): 798.

101. Min, Wan Nur Syahirah Wan, and Nur Zareen Zulkarnain. "Comparative evaluation of lexicons in performing sentiment analysis." *Journal of Advanced Computing Technology and Application (JACTA)* 2.1 (2020): 14-20.

102. Feizollah, A., et al. "Exploring halal tourism tweets on social media. J Big Data. 2021."

103. Asghar, Muhammad Zubair, et al. "SentiHealth: creating health-related sentiment lexicon using hybrid approach." *SpringerPlus* 5 (2016): 1-23.

104. Chen, Yixing, and JingTao Yao. "Sentiment analysis using part-of-speech-based feature extraction and game-theoretic rough sets." *2021 International Conference on Data Mining Workshops (ICDMW)*. IEEE, 2021.

105. Avinash, M., and E. Sivasankar. "A study of feature extraction techniques for sentiment analysis." *Emerging Technologies in Data Mining and Information Security: Proceedings of IEMIS 2018, Volume 3*. Springer Singapore, 2019.

106. Dorle, Saurabh D., and Nitin N. Pise. "Political Sentiment Assessment through Social Media." *2018 IEEE Punecon*. IEEE, 2018.

107. Pennington, Jeffrey, Richard Socher, and Christopher D. Manning. "Glove: Global vectors for word representation." *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*. 2014.

108. https://towardsdatascience.com/performance-metrics-confusion-matrix-precision-recall-and-f1-score-a8fe076a2262

109. https://towardsdatascience.com/taking-the-confusion-out-of-confusion-matrices-c1ce054b3d3e

110. https://www.baeldung.com/cs/stemming-vs-lemmatization

111. Kenyon-Dean, Kian, et al. "Sentiment analysis: It's complicated!." *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*. 2018.

112. https://abcnews.go.com/Business/timeline-elon-musks-tumultuous-twitter-acquisition-attempt/story?id=86611191

113. https://www.nytimes.com/2023/08/03/technology/twitter-x-tweets-elon-musk.html