Sentiment analysis on Twitter data: a detailed comparison of TextBlob and VADER

Asderis Georgios-Alexandros

SID: 3308200002

SCHOOL OF SCIENCE & TECHNOLOGY

A thesis submitted for the degree of

Master of Science (MSc) in Data Science

JANUARY 2022

THESSALONIKI – GREECE

Sentiment analysis on Twitter data: a detailed comparison of TextBlob and VADER

Asderis Georgios-Alexandros

SID: 3308200002

Supervisor:	Assoc. Prof. C. Tjortjis
Supervising Committee Members:	Dr. K. Tzafilkou
	Dr. D. Karapiperis

SCHOOL OF SCIENCE & TECHNOLOGY

A thesis submitted for the degree of

Master of Science (MSc) in Data Science

JANUARY 2022

THESSALONIKI – GREECE

Abstract

Studying social media can help understand public perception on various subjects, such as healthcare, and provide valuable information for the time of the respective study. Nowadays, the global pandemic of COVID-19 has resurfaced the subject of vaccination and the contradiction of people against it. Insights from the analysis of social media networks can help researchers understand the extend of vaccination awareness among the public. A way to do so, is by using sentiment analysis. This process can be implemented by using TextBlob and VADER, which are two libraries in Python that can evaluate a given text and return its sentiment score for each tweet. This score ranges from -1,0 to +1,0, with -1,0 being extremely negative, +1,0 being extremely positive and tweets with a sentiment score ranging from -0,1 to +0,1 being neutral. This study uses 994.716 Twitter posts from eight hashtags about the top four COVID-19 vaccine production companies and two hashtags about the antivaccination movement. The time frame of the collection of tweets is from the 15th of July 2021 to the 7th of November 2021 (a total of 116 days). The aim is to analyze and compare the results of the two libraries, when applied on the same datasets. Results indicate that there were 92.064 special cases where the label was positive on one lexicon and negative on the other. Then, by arbitrarily selecting ten tweets (one for each hashtag) from these special cases, custom scores are suggested with the aim of understanding the function of these tools, and based on that, propose some potential improvements.

Keywords

COVID-19, TextBlob, VADER, Twitter, Sentiment Analysis, Data Mining on Twitter, COVID-19 Vaccines, Anti-Vaccination Movement

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my supervisor, prof. Christos Tjortjis, for his consistent support and guidance throughout this project. Furthermore, I would like to thank Dr. Paraskevas Koukaras for his useful advice and help in finalizing this project. At last, I would like to extend my special thanks to my family for their unconditional support throughout the process of completing my MSc studies during the past year.

Contents

List of Tables

List of Figures

1 Introduction

During the last century, the field of medical sciences has made significant steps towards the development and improvement of vaccines. This type of treatment appears to be useful and has helped combat, or even eliminate, life-threatening diseases throughout the course of the years. Hence, since there is the ongoing pandemic of COVID-19 (or SARS-CoV-2), vaccination is reasonably considered as a possible solution to limit the spread and eventually weaken the influence of the virus on humans. Since nowadays, people to tend to use social media on a daily basis, they can be considered as a reliable mean of extracting information about various subjects. These may include politics, entertainment, healthcare, etc. and provide valuable information for the time of the respective study. Nowadays, the global pandemic of COVID-19 has resurfaced the subject of vaccination and the reaction of people against it. People use social media to express their thoughts on vaccination campaigns, share information about COVID-19, criticize their government's handling of the pandemic and express worries about the vaccines.

The internet has facilitated communication on a global level and has become a mean of daily communication. It also constitutes the main source of information for billions of people. Therefore, studying public perception on current issues can provide valuable information to the researcher. Nowadays, due to the ever-growing popularity of social media, the existence and creation of datasets containing posts, reviews, blogs, etc. is continuous. For example, the utilization of this data can help a researcher, or an organization, gain knowledge on their subject of study or some of their products respectively. The data used for this study derive from Twitter which is the most popular microblogging platform worldwide and concern hashtags about the subject of COVID-19 vaccination. It should be noted that, this study's aim is not to be liable to ethical scrutiny since the data that are being used come from a social network platform. In addition, all the fetched tweets came from Twitter users that have consented for the privacy of their data to be public. Additionally, all of the processes used in this study did not use the username of the users, with the purpose of maintaining anonymity.

For the process of sentiment analysis, firstly an input text is read by the tool of each researcher's preference. Then, by using its own computation methods, the overall sentiment of the text can be evaluated and returned as the result. There are two main types of approaches for this task: lexical-based and supervised machine learning. This study was conducted using the first case. The second one needs training data, in order to fine-tune the model, which may be difficult to acquire. Furthermore, another important drawback of this approach is that the training process may take a considerable amount of time to be finished. On the other hand, lexical-based approaches do not need a training phase in their process. Also, since there are already existing libraries in many programming languages, researchers can rely on preexisting lexicons that have been manually constructed by developers. In this study, the tools used for this process are two Python libraries: TextBlob and VADER. By using these libraries, one can conduct sentiment analysis on a given text.

For the implementation of this study, firstly, the fetched tweets were preprocessed, and the text of the tweets was cleaned and prepared for the analysis. Then, by using two different lexicons, sentiment analysis was conducted. Both total and daily results of each lexicon individually will be presented in the following chapters. After that, the study focuses on cases that have been labeled diametrically opposite. This means that one of the two lexicons labeled a tweet's text as positive and the other lexicon as negative, or the opposite. Then, by arbitrarily selecting one tweet of the aforementioned category for each hashtag, a table was constructed containing ten tweets. There, we provide our view on the evaluation of these tweets by TextBlob and VADER, and suggest our own scores based on subjective rating. The purpose of this thesis is to suggest improvement to these tools by having an in depth look on how their results are produced.

1.1 Problem Definition

Analyzing social media content about opinions can be useful for a handful of cases, but sometimes the volume of data can make this process impossible for a human reader. Hence the process of sentiment analysis has to be automated. Handling natural language is a demanding task, and data originating from social media do not constitute an exception. The already existing sentiment analysis tools work in different ways, hence their results are usually not the same. In addition, they have been created for general purpose use and are not specialized in specific subjects, for example –in this instance- healthcare. This usually leads to inaccurate results when conducting the process of sentiment analysis. So, it would be reasonable to consider solving such issues though the analysis of experimental results, to create more efficient tools and processes for similar studies in the future. Additionally, the creators of these tools update the functions of the lexicons in their attempt to improve their tools every few years. Eventually, the following thesis can be proven to be useful for developing these tools further in the future.

1.2 Aims and Research Questions

This study aims to analyze the results of sentiment analysis from two natural language processing tools, TextBlob and VADER. The dataset that these tools will be used on is the same for both and contains user's tweets from ten different hashtags, about the topic of COVID-19, from the 15th of July 2021 to the 7th of November 2021. Then, after presenting and comparing their outcomes, personal -and subjective- sentiment scores are proposed, for a random set of tweets from the acquired data, with the aim of identifying malfunctions and providing potential improvements to these tools.

1.3 Thesis Outline

In this study, Chapter 1 contains an introduction to the main subject along with the problem's definition. Some fundamental concepts and tools, needed for the comprehension of the rest of the dissertation, are explained in Chapter 2. Chapter 3 contains related published work of other scientists about each of the subjects concerning the dissertation. Some details of the data, as well as the preprocessing part of the study are discussed in Chapter 4. The implementation of the sentiment analysis for both lexicons, their results and our proposal on the subject are in Chapter 5. Finally, Chapter 6 contains conclusions, challenges, and future work.

2 Background

This chapter contains fundamental concepts and tools for the process of sentiment analysis on tweets, that will be useful for understanding the rest of this study.

2.1 The COVID-19 Pandemic

In December 2019 a case of a new type of virus was detected in Wuhan, China, known as COVID-19 (or SARS-CoV-2) [1]. Due to its high transmissibility, it quickly raised concerns of the World Health Organization (W.H.O.) and three months later, in March 2020, it was declared as a global pandemic. This virus is mainly transmitted through air and its symptoms may vary, with the most common being fever, dry cough, and tiredness [2]. It is also believed that at least a third of the total COVID-19 infected individuals do not develop noticeable symptoms [3]. According to W.H.O, as of today (the end of December 2021), there have been more than 290 million confirmed cases of the virus worldwide with 5,4 million deaths attributed to COVID-19 [4]. A case is considered confirmed with a positive COVID-19 detection test.

2.2 COVID-19 Vaccines

Currently, scientists all over the world are developing medicines to inhibit the virus, with vaccines being the primary form of treatment so far. Several vaccines worldwide have already received approvement and have been distributed to various countries, with approximately 9 billion vaccine doses administered so far (end of December 2021) globally [4]. The aim of the vaccines is to provide immunity against the virus, reduce its spread and prevent severe illness and death. Several vaccines managed to reach an efficacy of 95% against symptomatic infection and so far, 22 have been authorized by national regulatory authorities for distribution and public usage, while 330 are still in various stages of testing and development [5]. Most of the countries' vaccination strategy is to give priority to high-risk groups (e.g., people with respiratory diseases and the elderly) and to people that work at healthcare or nursing homes.

2.3 Twitter & Twitter API

Twitter is a social media platform where users can post and interact with other individuals via a small text called "tweet". Twitter users can post, comment, like or retweet (meaning repost) another tweet through Twitter's website or its mobile app. On this platform, people openly express their opinions through tweets that relate to a topic with the use of hashtags (#). For example, if a user writes "#vaccine" anywhere in his tweet's text, the tweet will be grouped together in a subcommunity with other tweets that used the same hashtag. In this way, they can engage in conversation or post something new about a specific subject of their choice. A tweet's text is restricted to 240 characters, which obliges a user to be synoptic and precise about the point he is trying to get across. This makes Twitter suitable for natural language analysis and knowledge extraction about public perception on important issues. Also, it is notable that its users tend to tweet within a short period of time (or even simultaneously) about a certain event that takes place and generates some type of emotion. So, given its popularity, Twitter can be proven to be a valuable and reliable platform for data extraction about public perceptions. Furthermore, it can also be used by developers via one of its APIs (Application Programming Interface). The use of a Twitter API gives its users access to more advanced content and features compared to the regular platform. For example, users can analyze business data, use Twitter data for research purposes, learn new skills and teach other users, etc. Twitter API allowed the access and the extraction of data for this study.

2.4 Natural Language Processing (NLP)

Natural language processing refers to the "ability of computers to understand and process human data on text or speech format" [6]. Its functions involve computer science, the use of algorithms and artificial intelligence among others. Some examples of NLP are natural language generation, speech recognition and text summary generation. The purpose of NLP is not just to read or translate a script, but to be capable of extracting knowledge from it by understanding the tone, the sentiment, and the relationships between words or phrases. One of the features of NLP that is usually applied on social media data, is sentiment analysis.

2.5 Sentiment Analysis

Sentiment analysis is a combination of text analysis, natural language processing and computational linguistics with the purpose of extracting information, and identifying, studying, and quantifying the overall emotion of a given text. It is considered to be a computational technique for studying many and various opinions at the same time, while extracting their overall sentiment. Its applications may vary from social media, marketing, evaluating customer reviews or feedback and even healthcare purposes. Nowadays, companies can use Twitter and other social media platforms as a tool for analyzing customer satisfaction to improve their products or their marketing campaigns. After all, the ease of use and the immediate nature of these apps, make it an ideal tool for knowledge extraction. With the incorporation of social media in our everyday life and their ever-increasing use, apps like Facebook, Twitter, YouTube, Instagram, etc. can be used for making assumptions and evaluating products or services impact on consumers. The technique of sentiment analysis can also be used for predictive tasks.

2.6 phpMyAdmin, MySQL & Python

phpMyAdmin is a software build for administration of MySQL over the web [7]. MySQL (My Structured Query Language) is an open-source database management system that specializes in relational databases [8]. With phpMyAdmin, the user can modify databases by changing tables, columns, rows, create new databases to store data and export them in various formats (like CSV, XML, etc.). Python is a high-level open-source programming language that focuses on code readability with the use of simpler syntax [9]. It offers a variety of free scientific libraries, visualization tools, plotting, etc. With the use of tweepy, which is a library in Python, a user can access Twitter's API and work with Twitter data. Furthermore, with the use of mysql.connector, which is a self-contained driver in Python, a user can develop MySQL database applications through Python. For this study, a combination of the aforementioned tools produced the datasets with users' posts about the chosen hashtags.

2.7 KNIME Analytics Platform

KNIME is an open-source data analytics platform. It contains various functions of machine learning, data mining, reporting and visualizing data. With the use of nodes and edges, the user can choose how to handle multiple datasets at once, blend various external tools (like Python, R, Apache Spark, etc.), form his data, etc. while using little or no programming code at all. For this study, KNIME was used in the preprocessing and filtering of the datasets, which will be explained later, in Chapter 4.

2.8 Textblob

Textblob is a Python library used for processing natural language data and carries out its tasks through NLTK (Natural Language ToolKit) [10]. NLTK is Python library for various lexical tasks such as translation, tagging the part of speech and sentiment analysis among others. TextBlob calculates the polarity of a text given as input and returns a score ranging from -1.0 to +1.0 as an output (with -1.0 being very negative, +1.0 being very positive and 0.0 being neutral). Furthermore, it can calculate the subjectivity of a sentence, by using the intensity of words to decide if the text is personal opinion and emotions, or substantiated information. Subjectivity ranges between 0 and 1, where 0 is a personal opinion and 1 is substantiated information. TextBlob is frequently used for simpler tasks of NLP.

2.9 VADER

VADER stands for Valence Aware Dictionary and sEntiment Reasoner. It is an open-source lexicon sentiment analyzer that uses a list of lexical features to calculate the sentiment of a given text [11]. It is considered to be an ideal solution when used for data from social media. Similarly to TextBlob, it also assigns a score to the given text ranging from -1.0 to +1.0, with 0.0 again being neutral. But, VADER firstly assigns a valence score to each word of the text and then computes the compound score, ranging from -1.0 to +1.0, by summing up and normalizing the valence scores that have been assigned. Furthermore, VADER uses other heuristics for sentiment computation like degree modifiers, text in capitals and change of polarity in two clauses of the same sentence. More details about TextBlob's and VADER's method of calculating the polarity of a text will be given in Chapter 5.1.

3 Literature Review

This chapter contains an extensive study of publications concerning relevant subjects to those of the dissertation. These are: sentiment analysis, the antivaccination movement, Twitter bots and the two NLP tools used for this study. In the end there are some conclusions and challenges that emerged from studying the existing literature.

3.1 Introduction of the Literature Review

The following review summarizes published scientific research papers on matters relevant to the sentiment analysis on data from Twitter. It aims to present the work and the outcomes of several surveys about the public opinion on Twitter, concerning the subject of vaccination. In the studied papers, scientists used various text mining, natural language processing and machine learning techniques to process the fetched data and draw useful conclusions about people's stance towards vaccination worldwide.

Nowadays, people tend to use multiple types of social media on a daily basis for information, entertainment and expression of their opinions and feelings. This literature review will examine the latter case, since scientists have developed proper tools that can analyze the text (or the natural language) of a social media post. With the use of data mining and text mining researchers were able to access public tweets containing a hashtag of their choice. Then, with the appropriate processing of the retrieved data, they used proper natural language processing libraries in order to get the text's sentiment score of each tweet. This sentiment is produced by the process of sentiment analysis, and it is usually labeled as positive, neutral, or negative based on a sentiment score assigned by the lexicon or library of the researcher's choice.

However, the use of social media for information can also be misleading, since there is not actual control of what is being posted. This facilitates the spread of misinformation, disbelief, and fake news and can eventually lead to the creation of communities that are based on such information. The most typical example is the antivaccination movement. Its supporters seem to continuously increase during the last decade and this phenomenon has grown scientist's concern for public health [12]. This movement's impact is of great interest, especially now that there is an ongoing pandemic, and its actions are believed to be the main reason for lower-than-expected COVID-19 vaccination rates worldwide. In fact, the W.H.O. has declared vaccine hesitancy as one the top ten threats for humanity's health [13]. This subject has become of great interest for scientists and in this case, the study of antivaccination communities on Twitter and their behavior. For example, what they post, how they interact with each other and other communities, like people who are in favor of vaccines.

The results of the majority of the studies indicated that the sentiment is mostly -but not overwhelmingly- positive since there is a significant portion of tweets that are labeled as neutral and a smaller portion that are negative. Negative tweets derive mostly from antivaccination groups, where the main concerns that are discussed are safety issues, vaccine side effects and culture related issues. As indicated by the researchers, the activity and misinformation of anti-vaccine groups online, cannot yet be proven to be harmful for the general population, since they constitute the minority in online platforms. Concerning bot accounts it was concluded that their existence makes the combat of online misinformation even more difficult that it currently is.

The sources used in the composition of this study came from the Google Scholar website, which is an index online engine about academic literature. The papers that were chosen as references were found with the use of relevant search queries and were then filtered based on how informative and relevant they were, to the subject of this thesis. The review will consist of chapters that cover a wide range of topics that are related to the main subject of the dissertation, with the purpose of getting a more in depth look and understand the relevant work that has already been done. More specifically, the chapters concern sentiment analysis, the anti-vaccination movement on Twitter, bot accounts and NLP (natural language processing) tools. Furthermore, conclusions and future challenges that emerged from the study of existing literature will be presented.

3.2 Sentiment Analysis & Sentiment Analysis on Vaccination

3.2.1 Sentiment Analysis

Studying published work about sentiment analysis shows that its applications may vary from marketing, evaluating customer reviews, sociopolitical analysis and even healthcare purposes [14]. Companies worldwide tend to use Twitter for analyzing customer satisfaction or the success of marketing campaigns. The ease of use and the immediate nature of social media, makes them ideal for knowledge extraction on public opinion. With the incorporation of social media in our everyday life and their ever-increasing use, apps like Facebook, Twitter, YouTube, Instagram, etc. can be used for making assumptions and evaluating a product's or a service's impact on consumers [15]. The technique of sentiment analysis has also been used for predictive tasks. For example, in the past, it has been used for natural disaster prediction [16], severe epidemic outburst prediction [17] and the prediction of national election results [18].

In the previous decades, feedback data could only be found in physical form, which meant that only few people had access to those. Nowadays, with social media, people are free to post their thoughts and feelings and those who are interested can analyze posts without having geographic limitations [19]. However, even though data from social media are useful for sociopolitical event analysis, the nature of the data may cause some implications. Natural language in general is challenging to handle and data that consist of tweets are no different. Their text is usually unstructured, and may contain sarcasm, irony or emojis which are challenges, that the developers of these tools have not overcome so far. So, it is crucial for researchers to apply the best possible text classification techniques to deal with such obstacles and produce the best possible accuracy on the polarity of the text.

There are multiple types of sentiment analysis that come from either machine Learning or lexicon-based approaches. The most common ones include fine-grained (used for polarity detection), emotion based (detects emotions) and aspect based (uses NLP to recognize which aspects of a product attract a certain type of sentiment) [19]. In general, sentiment analysis models are built on sentiment dictionaries, meaning lexicons where each word is assigned a score that indicates if it is positive or negative based on the words subjective meaning. Some examples of such lexicons are SentiWordNet, ANEW(Affective Norms for English Words), LIWC(Linguistic Inquiry and Word Count), VADER(Valence Aware Dictionary for Sentiment Reasonable) and TextBlob. These tools can be applied to different sizes of text (on the whole document, on each sentence of the document, or on each phrase of the sentences) in order to have the best possible granularity that the researcher aims to achieve for his study.

Via Twitter's API, a developer can also obtain data that are not necessarily in text format. For example, each tweet has the time of posting, the location, the account's number of followers, the number of retweets, the number of likes, etc. This allows the researcher to conduct his analysis and answer various types of questions and is not restricting to only text analysis. Furthermore, via Twitter's API, one can access only 1% of the published tweets from the platform for personal processing purposes. That sample is collected randomly, and this is mainly done for personal data protection purposes. Nonetheless, it is estimated that the deviation of sentiment analysis is small (less than 1.8%), compared to the initial -complete- dataset that contains 100% of the tweets for a given hashtag [20].

3.2.2 Sentiment Analysis on vaccination

Using information deriving from social media can be proven to be useful for researchers and scientists from various fields of study, especially now that there is an ongoing pandemic. For example, by using this tool, a pharmaceutical company that produces these products can understand tendences of reluctancy and see any common side effects that people are talking about online. Also, it can help understand the stance of people towards the product and evaluate the effectiveness of their vaccination campaigns. This way, health experts can use such a tool for answering common questions, provide specialized guidelines and help people understand the importance of prevention against the COVID-19 virus. Sentiment analysis can also be combined with forecasting methods to help analyze healthcare related issues, like vaccination [21]. For example, a company, or a country's government, can estimate vaccination rates and act accordingly to the percentages. This way, they can plan their actions beforehand to prevent low vaccination percentages before they happen.

In the papers that were studied for this literature review, researchers gathered data from Twitter and applied various preprocessing and text mining techniques to get both the individual sentiment of each tweet and the overall sentiment of their collected tweets. The tools they used generated two types of results for each tweet. A word indicating its overall sentiment (positive, neutral or negative) and a decimal number (called compound) that indicated the score of the tweet in the sentiment range (from -1,0 to +1,0). Scores bellow 0,0 meant the sentiment was negative, scores larger than 0,0 meant it was positive and if the score was equal to 0,0, the tweet is considered neutral. Researchers studied questions that concerned the vaccination debate on Twitter like "What people think of vaccination", "Which company's vaccine is the most/least popular" and "What are the main concerns of people about the vaccine". Due to their nature, some may say that the data may not be equally distributed, but since they are sampled at random from Twitter, their partial distribution is assumed to be equal to the actual distribution.

In the studied papers, researchers chose different ways to work with the fetched tweets in order to deal with the large number of retweets Twitter's APIs include in the process of collection [21]. For example, in some cases, researchers chose to work only with unique tweets. This can be achieved by using the "tweet_id" column to group together tweets and only keep one copy of each tweet. This is possible due to the fact that retweets retain the same tweet_id number as the initial tweet. Other researchers chose to leave their sample of data intact and work with the ensemble of the acquired data. The third method that researchers chose, was to set wights on the retweets and contain them to a smaller number (e.g., keep only 1 or 2 of the retweets on each tweet that had been retweeted) to avoid getting invalid results due to bot account's behavior. This way, they tried to keep a balance between neglecting bots and letting spam behavior affect the integrity of their outcomes.

Results also indicate that influential accounts (like politicians and public figures) can affect public opinion on vaccination [22]. That is the reason, many countries promoted vaccination through campaigns with famous people. Furthermore, the results show that tweets with a negative stance towards major vaccine brands seem to constitute the minority, compared to the positive and neutral ones. Also, an analysis on thoughts of people around maintaining safety measures even after their vaccination, seems to also be positive for the most part [21]. Nevertheless, percentages of unvaccinated people remain at high levels in several countries worldwide. For example, a study found out that U.S. and U.K. had a vaccination coverage of only 36,31% and 43,99% on their population respectively, while at the same time the negative tweets about vaccines reached up to 40% for both countries [22].

People seem to be willing to express their thoughts and worries about vaccination through online platforms. The main concerns of people on this subject include side effects, the safety of the vaccines and some death cases that were reported right after vaccination of individuals. In a research that was conducted to identify the main words used using the LDA method (Latent Dirichlet Allocation), it was found that people tend to tweet about their vaccination status (using words like “people”,“first”,”dose”) [22]. Also, “distribution” was another word used by people, which can be interpreted as a worry of people about vaccine stock and distribution. Furthermore, people also expressed their concern about local governments, by using the names of their high-ranking politicians in their tweets and criticizing them about their handling of the pandemic.

3.3 The Antivaccination Movement on Twitter

Without a doubt, vaccines have been one of the most significant medical inventions of humanity, but nowadays they may take lot of years to be clinically approved. A vaccine can be considered safe after succeeding in various testing stages. These stages can take up to a couple of decades to be completed, in order to make a vaccine safe and available for the public. Nowadays, due to the ongoing pandemic of COVID-19, scientists that work worldwide on the development of vaccines for the virus’s combat, had to dramatically speed up their pace of work. This also led to the loosening of testing criteria and the acceleration of clinical trials.

Even though the vaccines that are being currently distributed worldwide have undergone extensive tests and inspections, a significant part of people remain hesitant about the safety of the vaccines. This movement is also known as the “Antivaccination Movement” and its members are called “Antivaxxers” [23]. Nowadays, people use social media on a daily basis as a source of information and many times they do not check the source, or the validity of the facts presented in a shared news article, or a user’s post. Social media spreads information more directly and fast than websites, while also encouraging the interaction between users and the forming of like-minded groups. But this also leads to the spread of misinformation, that affects public opinion on several matters, and the subject of vaccination does not constitute an exception. Due to the nature of a post on Twitter (being precise and having character limitations), it becomes an ideal platform for people that want to present false facts or data, and don’t want to explain in depth the presented information.

The existence of misinformation in health-related topics can have negative effects both on individual and public health, and creates disbelief, fear, and uncertainty to those that believe groundless facts. The methods of persuasion differ from source to source. The most common is the use of emotional words in invented and fake stories with the purpose of provoking emotions to the reader [24]. In another research it was observed that tweets with negative compound scores, contained more sentimental words than the positive ones and that negative tweets increased dramatically when news related to a pandemic’s outbreak were reported [25]. Furthermore, the use of conspiracy theories also seems to be a frequent mean of persuasion against the vaccines. Through it, various Twitter accounts, try to cause uncertainty and disbelief around scientifically proven facts, new discoveries, statements of doctors, etc. Also, studies indicate that anti-vaccine supporters tend to produce and post less original content than pro-vaccine supporters, by reposting or retweeting more often content that they found in online groups of like-minded people. Studies also indicated that, posts of these groups blame the vaccine for the creation of new strains of the virus [26].

Concerning COVID-19 vaccination, a common argument of these online groups is the profit of big pharmaceuticals. These groups support that the vaccination strategy is a fraud and it is something that has been invented by the governments with the sole purpose of economic gains by the pharmaceutical industry. Antivaxxers also tend to criticize political figures -and especially those who are against their personal beliefs- and blame them for mishandling the pandemic [27]. They also tend to exaggerate about the side effects of the vaccines and accuse them of being responsible about severe health issues like brain damage, autism, etc. or even present vaccine as a mean of control by the government that can change the human genetic code. Furthermore, it was also observed that these groups of people usually support alternative therapies like homeopathy and an organic way of living [26]. Another common behavior of antivaxxers is the refusal on vaccinating their children, even in their youngest ages. Furthermore, posts about the components of the vaccine's formula are quite popular in these groups. On these posts they either present that the vaccines contain substances that they actually do not, or analyze some of its actual components with false or exaggerated effects compared to their actual ones. They also present the disadvantages of vaccination to be riskier than the advantages, for their personal health.

The main reason behind the popularity of such tweets is the emotional language they use and the spread of rumors that remain active over time, causing the spread of fake information among users or the encouragement for people to ignore advice coming from health organizations. This leads to the deterioration of credibility at scientists and scientific institutions that promote innovations. Another survey conducted on tweets in Spanish, showed that the most influential type of account among antivaxxers are the common people that offer "some type of truth that nobody knew so far" [26]. This study also credited the health anxiety disorder that derives from the vast amount of information that is available online nowadays, without its facts necessarily being checked or being trustworthy. This disorder can cause unpredictable behaviors that derive from the lack of trust to healthcare authorities and can eventually influence public health. It can also lead to defensive behaviors against people from different social, political, or even ethnicity groups and may lead to the birth of new conspiracy theories that target groups that differ in their way of thinking.

An analysis on the external links (URLs) of the posts of these groups, showed that the content shared included videos of health professionals talking against vaccination. But, after more detailed research, it was found that most of them were not actually health professionals, and they were linked to anti-vaccine movement groups. Other links contained information that promoted the rights of unvaccinated citizens, advice on how to avoid mandatory vaccination and supposed new alternative therapies against COVID-19. When it comes to vaccine efficacy, around 54% of users believe that the vaccine is ineffective against the virus, approximately 37% that the vaccines did not work as promised by their manufacturers and 9% insisted on the existence of other, less harmful types of treatments [26].

In overall, antivaccination users involve a plethora of topics that blend safety concerns, unsubstantiated allegations, and conspiracy theories with the aim of protection of people against the vaccine of COVID-19 [28]. They may also propose alternative ways of treatment, to people that belong to likeminded groups, that suggest more effective immunity against the virus. After studying their online behavior, scientists were able to infer that the spread of information in these groups is not centralized, meaning that this need of people to oppose against vaccination is a global phenomenon and did not derive from a single action.

3.4 Twitter Bots

A Twitter bot is a software that controls a Twitter account and it is programmed to mimic human behavior on the platform [29]. A bot can autonomously tweet, re-tweet, like, follow, unfollow, and send a direct message to other Twitter users. Its behavior is controlled by a user, who can set them to act in a specific manner. Its use can either be labeled as proper, or improper. The first one suggests that a bot is sharing trustworthy and helpful information, gives immediate responses to users, or creates informative content on its own, while the second one is usually associated with sharing fake news, malware and spamming with large amounts of tweets. In 2017, a study estimated that approximately 15% of Twitter accounts globally were bots [30]. In matters of controversy, such as the subject of COVID-19 vaccination, the most common type of bots that were identified are the improper ones. Their online behavior can affect the real life of citizens since, for example, they can be the reason for a slightly vaccine-hesitant individual to turn into a vaccination denier.

In 2019, a published scientific paper conducted research with the aim of analyzing bot behavior on Twitter and their interaction with real user accounts, giving an emphasis on retweets [31]. After studying the sub-communities that were formed, results showed that both pro- and anti- vaccine users mostly tweeted within “echo chambers” (meaning groups of people with similar opinions to them). Furthermore, retweets from bots, were found to be at a much higher frequency compared to retweets from real human accounts, while always keeping the content of the retweet relevant to the sub-community they posted to (either pro- or anti- vaccination). It was estimated that around 1.5% of the accounts of their dataset were bots and that they produced 4.6% of the dataset's tweets. Even though 77% of the examined tweets were labeled as positive, the percentages of retweets in pro- and anti-vaccine groups were similarly low with 1.51% and 1.16% respectively. The conclusion of this research was that bot accounts do not have a significant influence on people since they constitute a minority on Twitter platform. Nevertheless, their existence makes the combat of online misinformation even more difficult than it is.

Another published research focused on examining the prevalence of bots and the effectiveness of moderating misinformation in online platforms, during the first days of the arrival of COVID-19 vaccine batches in the United States [32]. Results showed that the percentage of bot accounts was less than 1.5% of the total user accounts, even in hashtags created by antivaccination supporters. It was also calculated that, these accounts produced less than 5.5% of the fetched tweets. Their results also showed that, a lot of accounts that contained anti-vaccination content where deactivated and labeled as unavailable by Twitter. This indicates that Twitter took action against potentially harmful content, at least in their studied hashtags. Finally, it was stated that understanding the mechanisms behind the promotion of anti-vaccination material can help in the combat of misinformation and soothe the malign effects of this movement on public health.

3.5 NLP Tools (TextBlob vs. VADER)

The computation of sentiment of a text can be manual or automated. The first might be better quality-wise, but it can be time consuming, or even impossible to do in large sets of text. So, the development of automated ways of computing sentiment can be proven to be quite useful. In addition, since it is a process done by a machine that is set to value all the text input with a pre-set ensemble of rules, it is more biased. There are plenty of automated sentiment recognition tools. The most commonly used are TextBlob and VADER. As it has been already mentioned in Chapter 2.8 and Chapter 2.9, they calculate polarity using different methods. Some of their similarities include not taking into consideration stop words (like “a”, ”the”, ”is”) because they do not have semantic value and providing the calculated sentiment score in the range of -1,0 to +1,0. Furthermore, both lexicons can also perform other tasks except from sentiment analysis, like sentence tokenization, part of speech tagging and N-grams among others. Looking into the existing bibliography, these libraries and their features have been used for various tasks.

For instance, these sentiment analysis tools have been used for predicting the movement of the stock market [33], [34], [35]. With their work, researchers tried to predict the optimal time to buy or sell stocks by combining and correlating financial data with social media data. This was achieved by analyzing the sentiment of their natural language data and its fluctuation through time. Then, they tried to locate similar and dissimilar behaviors of the stock market movement and check which of those comply with the overall sentiment of people using social media.

Another study used VADER to evaluate the impact of social media sentiment analysis in the financial performance of an airline company [36]. The study used tweets from the company's hashtag to conduct the sentiment analysis and then used regression techniques to check for possible relationships between the two. Results showed a high correlation between sentiment analysis results and the company's passenger growth rate. In September 2021, one research was conducted to decide whether the adoption of e-learning was done effectively during the waves of the COVID-19 pandemic [37]. Once again, after acquiring their data from Twitter, researchers used TextBlob and VADER along with topic modeling techniques to conduct their experiment. Results showed that people were hesitant about certain matters like the opening day of campuses and the children's difficulty to understand the concept of online education.

In overall, these lexicons are useful because they do not limit their action in specific subjects and constitute an objective way of interpreting a text. So, anyone who wants to study online behaviors, thoughts, reactions, etc. can use these tools for the process of sentiment analysis.

3.6 Literature Review Conclusions

To summarize, even-though -currently- vaccines are the best way to achieve the so called "herd immunity", the spread of misinformation online can be proven to be dangerous not only on an individual, but also at a collective level. In overall, it is believed that the more evidence there is on the restriction of the virus' transmission through vaccination, the more it will help enhance the public’s trust on this type of preventive treatment. This study of literature showed that people with a negative stance against vaccination constitute the minority, without having a significant influence on people that are in favor of vaccination. The main concerns of antivaxxers are safety issues like luck of trust in medical experts, side-effects, and the deaths of recently vaccinated people. It is also notable that, in most cases, pro-vaccine users’ posts contained reliable scientific sources (like the W.H.O., scientific publications, etc.), while posts from anti-vaccine users contained unreliable sources and had lots of bot account activity within them. In general, social media can be used as a mean of promotion for healthcare marketing campaigns and can play an important role in times of doubt and fear, by restraining misinformation. Especially nowadays, that generations are being raised while using the internet and can easily be influenced when forming opinions about such issues. So, governments worldwide should take action and formulate sound policies to tackle the issue of propaganda and educate their citizens, with the aim of completing mass vaccination at least on a national level.

3.7 Literature Review Challenges

The study of scientific papers about sentiment analysis on Twitter, also revealed a handful of challenges that researchers faced and left them as open challenges for the researchers of the future. Most of them appeared to be similar among the various studies, even though studies took place in different countries, continents, and timeframes. For example, a system that can automatically detect and correctly classify the use of sarcasm or irony in a post’s text, has not yet been invented. This fact can also be enhanced by the lack of most systems to recognize and correctly classify the emotion of emojis, that constitute a common way of a user expressing a feeling (either positive or negative). The aforementioned issues need to be taken seriously, since the misclassification of sentiment in a group of tweets can be proven to be a vulnerability to the study and lead to imprecise results. Due to the nature of this subject, most studies were conducted using Twitter to fetch users’ posts data. So, it would be reasonable to suggest that the use of other social media can reinforce the draw of conclusions. Some studies even suggested cross-platform data collection and multi-modal sentiment analysis tools to analyze different types of content simultaneously (like pictures, text, and audio).

Another issue that was raised by researchers was the relativity of their results according to each of the studied periods since discussion on social media is considered to be time sensitive. This concept is called “concept drift” and it represents the change in the distribution of the data that can cause inaccurate results and degraded performance by the sentiment analysis tools. For example, examining a community’s interactions in a time of high controversy around a subject can be completely different compared to the study of the same subject in a period that it was not as trending among users. Furthermore, people in real life tend to change their opinion around a subject while receiving various stimuli. Unfortunately, this is something that -at least so far- cannot be a dynamic parameter in the study of social media. Additionally, another -more specialized to the subject- challenge that came up, was the lack of knowledge of medical terms by the sentiment analysis tools. For example, tweets that did not contain words directly expressing sentiment, but contained medical terms around vaccines, were found to be misclassified by the sentiment analysis tools. This could potentially be solved with the development of more specialized tools that will have a more advanced knowledge of medical terms that can conduct a more precise sentiment analysis.

In addition, since the pandemic of COVID-19 is still ongoing, a change in the trending topics that are discussed online should not be neglected. For example, on Twitter, hashtags that contained the word “vaccine” were not as popular in the beginning of the pandemic -in December 2019-, and they may not be as popular when scientists achieve the development and release of other types of medicines to cure COVID-19, like pills. Furthermore, by gathering data using a specific keyword or hashtag, might lead to missing other users’ content, that had posts or discussions relevant to that specific issue, but did not use this particular keyword. Even though this pandemic is a worldwide concern, the epidemic situation differs from country to country due to the measures taken be each government. Also, each country has different financial abilities, technological advancement, and international political status, which can lead to different vaccination opportunities for their citizens. So, studies should consider limiting their data sample to a specific language, region, etc. to get more accurate results. In other words, the promotion, the effectiveness, the distribution, and the safety of a vaccine, is linked directly to its acceptance by the people and this may vary in different parts of the world.

Summing up, studies that are currently being conducted, can become the foundation for building more efficient models that can be utilized to face similar threats in the future, by forming strategies and policies based on the analysis of data. So, it would be reasonable to continue conducting studies until the end of this pandemic and understand the overall sentiment of people and the actual effect of vaccination campaigns on people worldwide. Another important aspect that should not be neglected is the existence and the activity of bots on social media. With the right processing of the data, this will not be a problem, but it is an aspect of social media that should always be handled in relation to the respective objective. This leads to the need of developing public surveillance programs that will automatically detect and act against online misinformation. The longer it takes to respond to misinformation, the more difficult it becomes to deal with it later on, due to the fast spread of information online. So, it would be interesting to see if the development of such tools will take place in the future, since it concerns both governments and pharmaceutical companies, and can be proven to be the main way of dealing with misinformation.

4 Data Collection & Preprocessing

This chapter contains information about the data. This includes the timeframe of the study, the selection of hashtags and the way the data were retrieved. Furthermore, it explains the way fetched tweets were distributed in datasets, how the text of the tweets was preprocessed and the filtering process of tweets that ended up in the final datasets of the experiment.

4.1 Data Description

The collected data are from Twitter posts between the 15^{th} of July 2021 and the 7^{th} of November 2021 (116 days in total). The tweets were collected based on the hashtags (#) they contained. Specifically, these hashtags are: #antivax, #antivaxxers, #astrazeneca, #astrazenecavaccine, #johnsonandjohnson, #johnsonandjohnsonvaccine, #moderna, #modernavaccine, #pfizer and #pfizervaccine. The choice of these specific hashtags from COVID-19 vaccine production companies, was based mostly on the numbers of vaccine doses production worldwide. Another significant reason for this choice was the fact this study concerns tweets that were written only in English, so other vaccine companies that had high numbers of vaccines produced, like Sputnik V for example, could not provide a large amount of useful data, since they were mostly distributed to non-English speaking countries. The two hashtags about antivaccination were chosen due to the rise of the antivaccination movement's online presence in the past 2 years. For the sake of the study, hashtags were grouped in the following five pairs:

- #antivax & #antivaxxers
- #astrazeneca & #astrazenecavaccine
- #johnsonandjohnson & #johnsonandjohnsonvaccine
- #moderna & #modernavaccine
- #pfizer & #pfizervaccine

For each category the respective hashtags were more than just two, but the choice was made based on the hashtag's popularity on Twitter.

4.2 Data Collection

Data were collected from Twitter platform, by using Twitter's API (Application Programming Interface) based on the date and time they were posted on the platform. The implementation of fetching the data was done through Python and the storage of the retrieved data was done at the phpMyAdmin platform. In the aforementioned, with the use of MySQL, empty data tables were created and contained 47 columns with information about each tweet like the id number of a tweet, the number it has been retweeted, the username of its creator, the user's follower count, etc. Then, for every one of the chosen hashtags, these tables' rows were filled with tweets. After the datasets were completed, they were downloaded in CSV (Comma Separated Values) files and were ready to be processed.

4.3 Dataset Preprocessing

The existence of bots on online platforms makes the processing of data more challenging and may lead to imprecise results, if not considered. This leads to the necessity of filtering the retrieved data. This can be done by using their id number. Each tweet has a unique id number when it is created for the first time and keeps the same id number every time it is retweeted (meaning reposted). Since retweeting is a common behavior of bot accounts and in such controversial subjects as vaccination the presence of bots is high, the number of retweets had to be cut down. But, choosing to only keep one tweet for every tweet id number, means that all the retweeted content is not taken into consideration. In order not to completely neglect retweets, only one copy of each retweeted tweet was kept. Also, columns that were unnecessary for the analysis were removed from the dataset. Then, a new CSV file was exported and was ready for further preprocessing. Those steps were achieved with the use of KNIME Analytics Platform.

The table below (Table 1) contains the number of tweets for each hashtag. The low number of tweets that ended up in the final datasets, is due to the high number of retweets in the chosen hashtags. This can be explained by the fact that the subject of vaccination is trending during the time of this study and provokes the spamming behavior of bot accounts with high numbers of retweets.

Table 1: The number of total tweets, unique tweets and those that ended up being used after filtering the data

	Total Fetched Tweets	**Unique Tweets**	**Filtered Tweets (only 1 copy per RT)**	**Percentage Of Tweets Used In The Analysis**
#antivax	4.342.584	85.335	169.846	3,91%
#antivaxxers	7.089.197	139.515	277.635	3,92%
#astrazeneca	3.076.851	61.660	122.685	3,99%
#astrazenecavaccine	52.229	578	1.154	2,21%
#johnsonandjohnson	996.274	9.606	19.170	1,92%
#johnsonandjohnsonvaccine	75.880	1.234	2.459	3,24%
#moderna	4.143.463	82.200	163.594	3,95%
#modernavaccine	266.117	3.404	6.795	2,55%
#pfizer	5.030.119	100.596	200.178	3,98%
#pfizervaccine	1.521.606	15.641	31.200	2,05%
Total	26.594.320	499.769	994.716	3,74%

Furthermore, for the analysis to be meaningful, the text of the tweets had to be cleaned. Tweets usually contain URL links, mentions, hashtags, and other elements that are not useful in the study of natural language. Firstly, each tweet of the dataset that came from a retweet, contained a “RT” in the start of the text, which was removed. Then, all the hashtags were removed since a tweet may not contain only one. Also, punctuation and special characters (like @, !, &, $, etc.) were removed from the text of the tweets. Finally, all external links (in the form of “http\”) were removed.

After these steps, the text of each tweet only contained natural language, and was ready for the application of TextBlob and VADER to extract the sentiment of the text. The sentiment score for both ranges between –1,0 and +1,0. The three possible values that can be assigned to a tweet for its content characterization are: positive, neutral, and negative. Positive ranges from +0,1 to +1,0, negative from -1.0 to -0.1 and neutral is smaller but not equal to +0,1 and bigger but not equal to -0,1. The allocation of sentiment of the tweets for both lexicons can be seen in the following chapter (Chapter 5).

5 Experimentation

This chapter contains the main part of this study. Firstly, the is an explanation of how the two tools calculate polarity. Following, the results of sentiment analysis for both lexicons are presented with visualizations on an overall and a daily basis. Then, the study focuses on the cases where the results were opposite between the two lexicons.

5.1 How TextBlob and VADER Calculate Polarity

TextBlob and VADER present differences in the way they calculate the polarity scores of a text, which causes them to have different results. TextBlob's analysis is rule-based and uses a pre-defined dictionary that defines positive and negative words. In addition, word frequency, intensity and semantic relations between words are also taken into consideration. For calculations, TextBlob treats the text as a bag of words (BoW). This is a natural language processing technique that transforms and represents a text as a set of its words and does not take grammar and word order into consideration. So, after TextBlob assigns scores to each word, it calculates the overall sentiment by taking the average of these scores. Additionally, words that show negation reverse the polarity of a sentence.

Similarly to TextBlob, VADER is sensitive to polarity and intensity when calculating the sentiment score of a text and it is built using a dictionary with pre-defined emotion intensity score of words. When it is applied, it assigns emotion intensity to the elements of the text and then sums up the scores of the text's parts. It does not only return a decimal number, but a dictionary that contains three individual scores (positive, neutral, negative) and the overall compound score of the input. The compound score is calculated by normalizing the three individual scores. Also, VADER is able to reverse the meaning of words when they are accompanied by negation and can comprehend punctuation and capital letters.

5.2 TextBlob Sentiment Analysis Results

This part presents the outcomes of the analysis of TextBlob's results. First, the results are grouped based on sentiment type and then, the daily trajectory of the mean polarity is shown for each one of the five hashtag pairs.

5.2.1 TextBlob Number of Tweets by Sentiment Type

After merging the datasets in pairs, based on hashtags (as explained in Chapter 4.1), TextBlob was used to get the sentiment analysis results. The bar chart below (Figure 1) shows the number of tweets in each sentiment category per hashtag pair. The percentages shown above each bar, have been rounded.

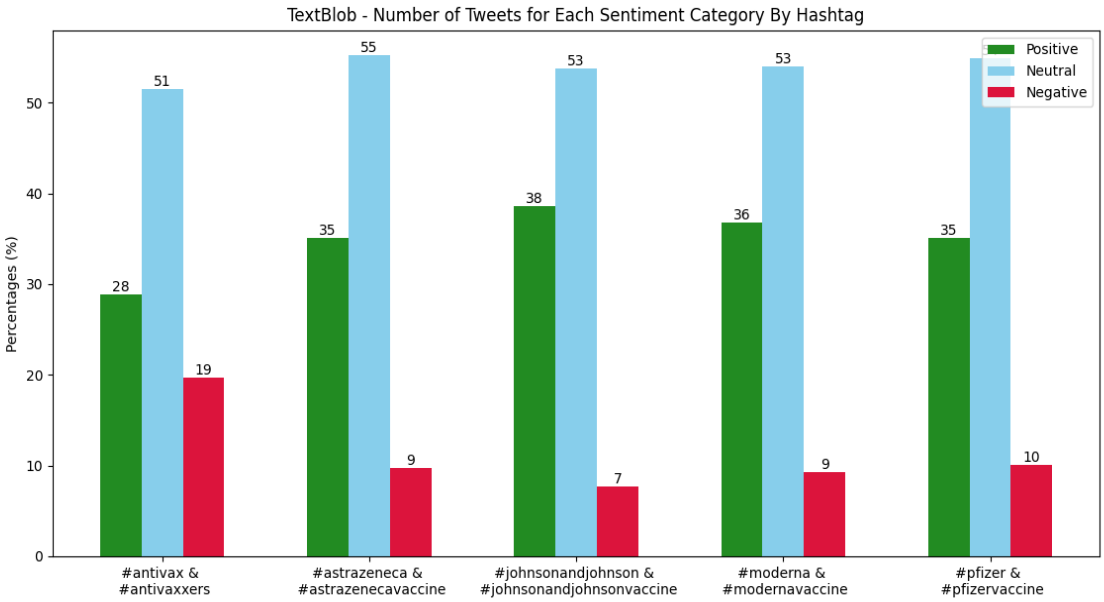

Figure 1: TextBlob's number of tweets for each sentiment by hashtag

As it can be seen from Figure 1, TextBlob returns neutral as the dominant sentiment in all five pairs. Furthermore, there is a similar distribution of sentiment in all five pairs concerning positive and negative tweets. The highest percentage of positive tweets is observed in hashtags about Johnson and Johnson, with 38,59%, followed by those of Moderna, with 36,83%. Tweets with antivaccination hashtags had the lowest percentage of positive tweets, with 28,80%. On the other hand, tweets containing hashtags about antivaccination, had the highest number of tweets categorized as negative, with 19,69%. All four hashtags about vaccine production companies had equal to or less than 10%.

Finally, the percentage of neutral tweets in all five pairs fluctuated between 51,51% and 55,21%.

5.2.2 TextBlob – Daily Mean Polarity

In order to study TextBlob's results throughout the studied time period, tweets were grouped together based on the date they were posted. Then, for each day, the average score of the tweets' polarity score was calculated. The following line graph (Figure 2) presents the daily trajectory of mean polarity for each pair of hashtags during the time period of study.

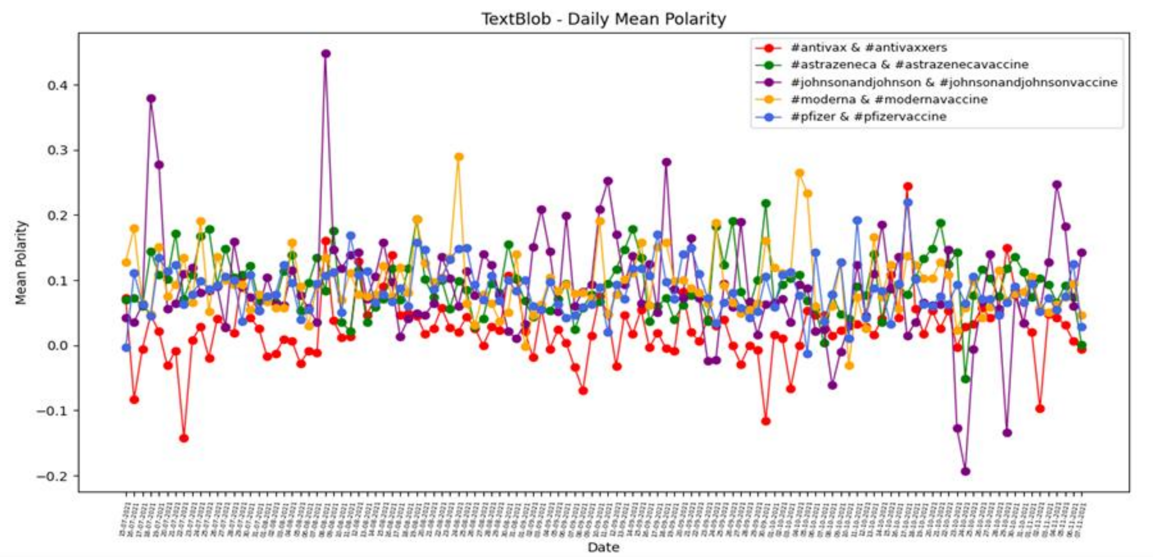

Figure 2: TextBlob's daily mean polarity trajectory

The main conclusion from Figure 2 is that sentiment polarity is mostly labeled as neutral since the mean polarity of all hashtags seems to range mainly from -0,1 to +0,2. This comes to correspondence with the results of Chapter 5.2.1. Even though hashtags of antivaccination get the lowest values daily, their polarity scores are not too negative, having their lowest observation at -0,15 on the 22nd of July. Pfizer and Astra Zeneca hashtags seem to simultaneously have the most stable and positive daily mean polarity scores, by only falling below 0,0 in two instances. Hashtags about Johnson and Johnson show the largest fluctuations in the graph. In fact, they seem to have both the highest and the lowest scores among all observations with +0,45 the 8th of August and -0,2 the 24th of October respectively.

5.3 VADER Sentiment Analysis Results

This part presents the results of the analysis of VADER. Similarly to Chapter 5.2, at first the results are grouped based on sentiment type and then, the daily trajectory of the mean polarity is shown for each one of the five hashtag pairs.

5.3.1 VADER Number of Tweets by Sentiment Type

Following the same procedure as Chapter 5.2.1, VADER was used to get the sentiment analysis results. The bar chart below (Figure 3) shows the number of tweets in each sentiment category, per hashtag pair. The percentages shown above each bar, have been rounded.

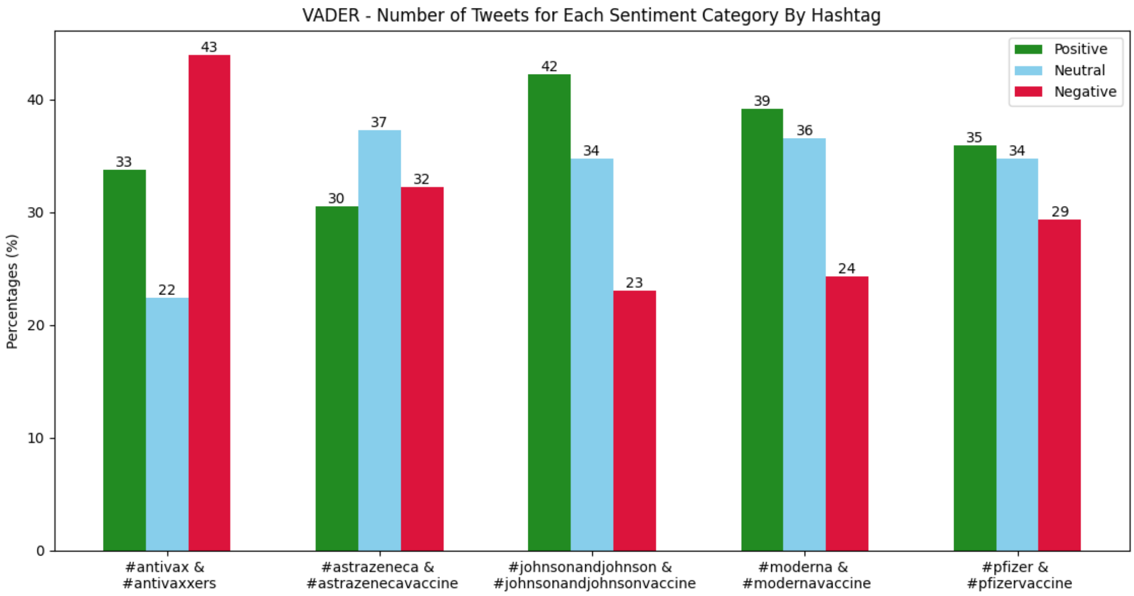

Figure 3: VADER's number of tweets for each sentiment by hashtag

As inferred from Figure 3, VADER yields different results compared to those from TextBlob. The neutral sentiment has the highest percentage in only one of the five hashtag pairs, that of Astra Zeneca, with 37,28%. For the remaining three companies, the allocation seems to be similar, with positive having the highest percentages, followed by neutral and then negative. Out of those three, the hashtags about Pfizer appear to be the most balanced, with the percentages of the three sentiments being relatively close, having a difference of only 6 units between them. Nevertheless, this is not the case in hashtags about antivaccination, where the negative sentiment seems to be the dominant one, with 43,90% followed by positive with 33,73% and then neutral with 22,37%. In overall, negative tweets seem to be almost threefold compared to those from TextBlob, while neutral are approximately 20% lower in all five cases.

5.3.2 VADER – Daily Mean Polarity

Again, similarly to Chapter 5.2.2, tweets were grouped together based on the date they were posted, with the aim of studying VADER's results throughout the studied period. Then, for each day, the average score of the tweets' polarity score was calculated. The following line graph (Figure 4) presents the daily trajectory of mean polarity for each pair of hashtags during the period of study.

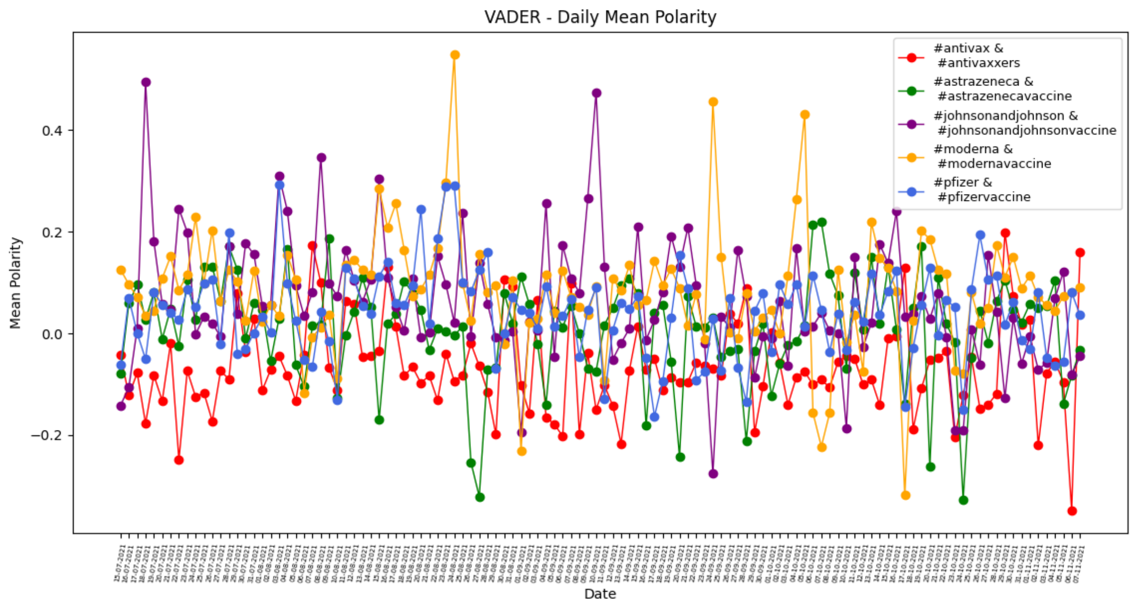

Figure 4: VADER's daily mean polarity trajectory

The results of daily mean polarity using VADER in Figure 4, present differences compared to those of TextBlob. First of all, results seem to have greater fluctuations in general and most of the values in this case range between +0,3 and -0,2. Once again, hashtags about Pfizer are the most stable ones, followed by those from Astra Zeneca which, however, present some highly negative values, with -0,27 on the 26th and -0,34 on the 27th of August and -0,35 the 24th of October. Moderna's hashtags had mostly neutral daily mean polarities, but on the 25th of August the highest positive mean was observed, with +0,57. Their lowest point was a -0,32, the 17th of October.

Johnson and Johnson's hashtags present continuously changing scores from day to day, like those from TextBlob. Their most notable values are in 18th of July with +0,50, the 10th of September with +0,48 while their lowest value was the 24th of September with -0,29. Similarly, to TextBlob's results, hashtags of antivaccination mostly occupy negative mean polarity scores and they even have the lowest mean polarity, with -0,39 the 6th of November. It also notable that antivaccination and Astra Zeneca hashtags where the only ones that did not surpass +0,2 on any of the 116 days of the study.

5.4 TextBlob vs. VADER: Comparison of Results

In order to be able to compare the results of the two lexicons with each other, first they have to be combined. This way, we can have an overview and identify the deviation of results. Then, to make the analysis more thorough, the allocation of each possible combination of results in pairs is presented.

5.4.1 Overview

The table below (Table 2) contains a breakdown of the allocation of tweets based on the sentiment characterization that both lexicons provided for each tweet. The first column contains the hashtags, and the second column contains the total number of tweets for the respective hashtag pair. The third column shows the number of tweets that have been identified with the same sentiment by both lexicons (either positive, neutral, or negative). The rest of the columns refer to the cases where the type of sentiment was identified as not the same. In these columns the sentiments that are mentioned in the column's title can be reciprocal. For example, in the fourth column "Positive and Neutral" there can either be instances that TextBlob recognized as positive and VADER as neutral at the same time, or the opposite, were TextBlob recognized the tweet's text as neutral and VADER as positive. The bottom two rows of the table contain a sum of each column and the percentage that each column occupies corresponding to the total amount of tweets.

Table 2: Allocation of tweets based on the sentiment characterization of TextBlob and VADER

	Total Number of Tweets	**Same Sentiment**	**Positive and Neutral**	**Negative and Neu-tral**	**Positive and Neg-ative**
#antivax #antivaxxers	447.481	211.624	83.715	105.267	46.875
#astrazeneca #astrazenecavaccine	123.839	60.613	26.884	23.743	12.599
#johnsonandjohnson #johnsonandjohnsonvaccine	21.629	13.086	4.040	2.955	1.548
#moderna #modernavaccine	170.389	88.032	43.921	26.192	12.244
#pfizer #pfizervaccine	231.378	115.023	55.293	42.264	18.798
Total	994.716	488.378	213.853	200.421	92.064
Percentage (%)	-	49.10%	21.50%	20.15%	9.25%

As it can be seen on Table 2, the two lexicons had the same outcome in less than half of the dataset's tweets (49.10%). The instances where the sentiment was either positive and neutral or negative and neutral had similar percentages, with 21,50% and 20,15% respectively. Finally, the rarest case was one being positive and the other being negative, with 9,25%. This percentage may not seem particularly high when compared to the whole but taking into consideration that both lexicons were given the exact same data, and that this was the most extreme case possible, it should be taken into consideration.

5.4.2 Comparison of Combined Results

For further analyzing the comparison of their results, the next graph (Figure 5) is a pie chart depicting the percentages of tweets in every possible combination of results. In the legend of the graph, TB stands for TextBlob, and VA stands for VADER. The percentages shown on side of each piece of the graph have been rounded.

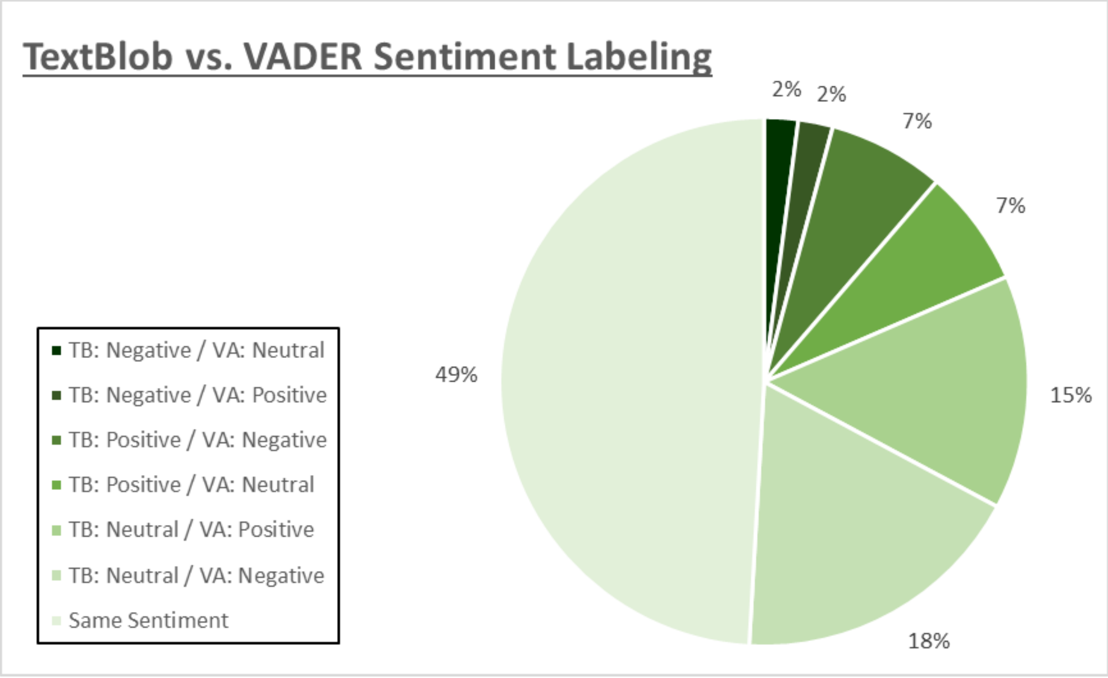

Figure 5: Allocation of TextBlob and VADER sentiment labels on tweets

From Figure 5, it can be deduced that same sentiment occupies the largest portion of tweets, when separated in those seven categories. Then, there is TextBlob neutral and VADER negative with 18% and TextBlob neutral and VADER positive with 15%, which also shows the tendency of TextBlob to categorize tweets as neutral that has already been discussed in Chapter 5.2.1. Following there is TextBlob positive and VADER both neutral and negative with 7% each. At last, there is TextBlob negative and VADER both positive and neutral with 2%. It is notable that even though the case of TextBlob being negative while VADER is neutral is not considered one of the extreme cases, it appears in a significantly low percentage. Also, out of the two extreme cases, where one is positive and the other one is negative, the case where TextBlob is positive and VADER is negative, is considerably more frequent than the opposite case.

5.5 TextBlob vs. VADER: Opposite Result Analysis

The next part of the study focuses on examples from the data, where the results of sentiment recognition were opposite for the same text. After filtering the data and only keeping those who had one lexicon being positive and the other being negative, a table was formed (Table 3) with the method of arbitrary selection, with one random tweet for every hashtag.

Eventually, the table below contains ten tweets, one for each of the hashtags of this study. The first column contains the respective hashtag, and the second column contains the tweet's text, cleaned as described in Chapter 4.3. The third column contains the sentiment score assigned by TextBlob and the fourth column the one assigned by VADER. The aforementioned column also contains the analytical scores (in braces "{ }") for each type of sentiment that VADER assigns to each text that is given as input. The fifth column shows the deviation between the score of TextBlob and the respective score of VADER for each case. For instance, in the first example, the deviation is 0,3 because TextBlob's score was -0,3 (meaning it is negative) and VADER's metric for the respective negative score is 'neg'=0,0. So, the difference between those scores is 0,3. The last column shows the suggested sentiment score and sentiment type, by using a proposed subjective evaluation method.

Regarding the last column of the table above, emotional evaluation of a text by humans is considered subjective, since each individual tends to weight the importance of parts of speech differently. This does not only apply in parts of speech, but also in the tone of a text. For example, irony, negation, punctuation, text in capital letters, etc. in a text can be received differently by multiple readers. For the assessment of each text of Table 2, some types of emotions should be defined, that comply with the given sentiment scores.

The existing literature from the sectors of linguistics and psychology has not yet defined the types of human sentiment in a universally accepted manner. Some scientists suggest that the main types of human emotions are four: Happiness, Sadness, Fear/Surprise, and Anger/Disgust [38], [39]. Others suggest that sentiment types come in pairs of opposites, like: Joy/Sorrow, Anger/Fear, Acceptance/Disgust, Surprise/Expectancy [40]. Others suggest that there are at least 27 types of emotions [41]. At last, a portion of scientists suggest that the main types of emotions are six: Happiness, Sadness, Fear, Surprise, Anger, Disgust [42], [43]. This thesis adopts, these six sentiment types which were assigned to each tweet on the "Suggested Score & Sentiment Type" column of the table above. Some instances contained one type of emotion, while others contained two types.

Table 3: Ten arbitrarily selected tweets with their TextBlob and VADER scores along with the suggested score and sentiment type

	Hashtag (#)	Tweet's Text	TextBlob Score	VADER Score	Difference	Suggested Score & Sentiment Type
1	#antivax	I m not gonna go off on antivax mentality for like at least 30 minutes Promise	-0.3	+0.586 {'neg': 0.0, 'neu':0.714, 'pos':0.286, 'compound': 0.5859}	0,3	-0,2 Anger
2	#antivaxxers	Raising doubts about mass vac-cination Think most of us knew anyway	+0.5	-0.296 {'neg': 0.18, 'neu': 0.82, 'pos': 0.0, 'compound': -0.296}	0,5	-0,3 Fear & Anger
3	#astrazene-ca	I m sooo happy I ve got my first doze of Astra-zeneca but my head hurts	+0.525	-0.421 {'neg': .237, 'neu': .629, 'pos': .134, 'compound': -0.4215}	0,391	-0,2 Happiness & Fear
4	#astrazene-cavaccine	AZ Vaccine Creator received standing ovation from the audi-ence at Wimble-don game recent-ly Thank you	-0.2	+0.361 {'neg': 0.0, 'neu':0.848, 'pos':0.152, 'compound': 0.3612}	0,2	+0,8 Happiness

5	**#john-sonandjohn-son**	Woman s Death Caused by Rare Side Effect	+0.3	-0.599 {'neg':0.328,'neu':0.672, 'pos': 0.0, 'compound': -0.5994}	0,3	-0,6 Fear & Sadness
6	**#john-sonandjohn-sonvaccine**	Any sane person knows better than to take a vaccine from the Nazis criminals	+0.5	-0.202 {'neg': 0.21, 'neu':0.625, 'pos':0.165, 'compound': -0.2023}	0,335	-1.0 Anger & Disgust
7	**#moderna**	Moderna could be implicated in experiments that created corona-virus	-0.4	+0.25 {'neg': 0.0, 'neu':0.818, 'pos':0.182, 'compound': 0.25}	0,4	-0,2 Anger & Fear
8	**#moderna-vaccine**	I got my covid 19 vaccine I signed up for the no waste list at my local phar-macy	+0.3	-0.612 {'neg':0.238,'neu':0.762, 'pos': 0.0, 'compound': -0.6124}	0,3	+0,2 Happiness
9	**#pfizer**	Pfizer is ap-proved Get vaxxed No more excuses	-0.25	+0.153 {'neg':0.183,'neu':0.583,'pos':0.233,'compound': 0.1531}	0,067	0,0 Happiness & Anger

10	**#pfizervac-cine**	Vaccination date is coming closer and closer and I am getting more and more unsure	+0.5	-0.373 {'neg':0.164,'neu':0.836, 'pos': 0.0, 'compound': -0.3729}	0,5	-0,3 Fear

The first tweet of Table 3, that of #antivax, contains irony, anger and maybe a bit of aggressiveness. This makes the positive score of VADER seem irrelevant to the actual text of the tweet. The suggested score is not highly negative since the text could also be interpreted as a compromise. The tweet from #antivaxxers has the highest value in the difference column (0,5), along with #pfizervaccine that has the same difference. That difference is created due to the relatively high score of TextBlob, but after reading its text, we can understand that it contains anger, doubt, and fear. So, even though the only word that could be considered negative is “doubts”, TextBlob’s score is not suitable to the text. On the other hand, the score given by VADER is significantly close to the suggested score. The third tweet, from #astrazeneca, contains a contrast of emotions. Its first part is completely positive, while its second part is negative. In this case, both scores of the lexicons should be closer to 0,0 since the conflict of emotions should balance the resulting score. Here, VADER is considered to have a better result because it labeled the tweet as negative, but the tweet in overall is not as negative as VADER shows (-0,421).

The #astrazeneca vaccine tweet is a congratulatory message to the creator of the vaccine. Eventually, the negative score that TextBlob provided as a result, is not suitable for this text. VADER gives a positive score, but it is relatively low, taking into consideration that the one and only sentiment of the text is happiness. The fifth tweet, from #johnsonandjohnson, is a statement of news. The fact that it concerns the loss of a person shows sadness, while the cause of the fact being a rare side effect shows fear. Here, the word “rare” could be interpreted as a way of convincing people that the cause of death was not one of the usual side effects, but most likely the text aimed to spread doubt and maybe express anger. Keeping that in mind, the suggested score is negative, but not as high as VADER suggested. The next tweet is from #johnsonandjohnsonvaccine and contains aggressive speech. In fact, it is probably the most aggressive tweet out of the random sample of the table above. This contradicts with TextBlob's score, which is positive, but judging by the content of the tweet, VADER's score should be much higher.

The text of #moderna tweet has a tone of disbelief, fear, and a bit of anger. The result shouldn't be positive as VADER suggests, but also not as negative as TextBlob suggests since, the word "could", reveals the uncertainty of its author. The eighth tweet, from #modernavaccine, is informative and has a positive tone in overall. Here, VADER not only does label the tweet as negative but gives it a rather high score (-0,621), which is the highest negative score out of all ten of the tweets. So, TextBlob's result is better and in fact it is closer to the suggested score. The next tweet, from #pfizer, also contains contradicting emotions, like the one from #astrazeneca. Furthermore, it has the lowest difference value out of all the sample's tweets, meaning that in this case, both lexicons did better at recognizing the sentiment of the text. The suggested score of this tweet is 0,0 because the text could either be interpreted as anger towards people that make excuses against vaccination, or as a prompt for everyone to be vaccinated. Moreover, it does not contain many words that reveal either positivity of negation. The last tweet of the table, from #pfizervaccine, is clearly a text revealing doubt and stress. So, TextBlob's score which is relatively high (+0,5), is not suitable for this text. On the other hand, once again, VADER assigned a score that is close to the suggested score.

In overall, VADER's scores seem to have a better match with those of our proposal. Given that the tweet from #pfizer was labeled as neutral (0,0) by us, in six out of the nine remaining cases VADER provided the same type of score as we did (either both positive, or both negative), while TextBlob only did three times. Nonetheless, this cannot constitute a safe conclusion on which of the two lexicons is better, since the sample of the Table 3 contained only 10 tweets out of 994.716 in total, and those 10 were selected at random.

5.6 A Suggestion for Improvement

As it can be seen on the websites of TextBlob and VADER, their creators continue to update these tools with new features. After the analysis conducted in Chapter 5.5, there are some observations, that could potentially lead to suggestions on the improvement of these tools. For instance, TextBlob's practice of reversing the polarity score when having a negation word in the middle of the sentence does not seem to always work. As it can be seen in Table 3, "Any sane person knows better than to take a vaccine from the Nazis criminals" was labeled as positive (+0,5) from TextBlob. Here, the word "Nazis" and "criminals" were probably perceived as negative, but the word "than" in the middle of the sentence made the polarity score change from negative to positive. Also, an update on the pre-defined dictionaries of both should be considered by their creators. This is due to the constant emergence of new scientific terminology, or even everyday language and expressions that will not be taken into consideration by the tools when calculating polarity if they are not properly updated.

Furthermore, as observed from the results of the random sample of Table 3, the score of sentences that do not contain many words expressing intense emotions should be influenced more by those few words. For example, in #pfizervaccine the text says: "Vaccination date is coming closer and closer and I am getting more and more unsure". Here TextBlob labeled the tweet as positive (with +0,5) probably by weighting the expression "more and more", but here the emphasis should be in the word "unsure". So, the creators of TextBlob, could potentially set a bar where if words that express sentiment are below a set percentage of the total number of words of the sentence (for example 10%), the score should lean towards their sentiment. This also could improve the results of #johnsonandjohnson tweets that says: "Woman s Death Caused by Rare Side Effect". Here, the only and most sentimental word is "death", but the tweet reveals negativity and fear. So, TextBlob's positive score (+0,3) is not fitting for this sentence and the aforementioned proposal could potentially improve the result.

6 Conclusions

This chapter contains the conclusions of this study along with challenges and limitations that came up during its implementation. Finally, future work is proposed.

6.1 Conclusion

In a broader context, as of today, vaccines are scientifically proven to be the best way of limiting transmission and severe illness from COVID-19. So, social media can be a mean of promotion for healthcare practices and play an important role in such times of doubt and uncertainty. However, the spread of information online is happening at such a fast pace, that the filtering of information that is being shared is almost impossible. This can be proven to be dangerous both for an individual and at a collective level. In order to be able to understand the extend of such phenomena, researchers will have to use all possible means in their battle against this historical ordeal, for the sake of humanity. The integration of social media in our everyday lives can make them a useful tool for extracting information on various matters of current affairs.

Data used in this study consist of Twitter posts, written in English, from ten different hashtags about COVID-19 vaccines. The choice of these specific hashtags from vaccine production companies, was based mostly on the numbers of vaccine doses production worldwide, meaning those were the four companies with the highest production rates. The two hashtags about antivaccination were chosen due to the rise of the antivaccination movement's online presence since the beginning of the pandemic. As already mentioned, the processing of natural language can be a challenging task since the available tools cannot always produce accurate results yet. The aim of this study was to analyze and compare the results of TextBlob and VADER for the same dataset. This was done with the purpose of understanding the way they work and proposing suggestions for their improvement.

Firstly, the fetched data were preprocessed and filtered. Then, both lexicons were used for the process of sentiment analysis and their results were presented both in total and on a daily basis for the studied period, from the 15th of July 2021 to the 7th of November 2021. Results showed that there were cases where the two lexicons provided diametrically opposite results for the same tweet’s text. This means that one of the two lexicons labeled a tweet’s text as positive and the other lexicon as negative, or the opposite. So, the study then focuses on these special cases. The last part of this thesis contains a table of ten arbitrarily selected tweets, one for each hashtag, along with our subjective rating for each tweet. This was done including a subjective polarity score that ranges from -1.0 to +1.0, and a characterization of the text’s sentiment based on six main emotions (Happiness, Sadness, Fear, Surprise, Anger, Disgust).

Through this process, one can understand why the results were different between the two, and which one provided better results, at least for the random sample of ten tweets. Based on this analysis, potential improvements of these tools are proposed. These include an update of their pre-defined dictionaries with expressions and scientific terminology. Furthermore, TextBlob’s practice of reversing the overall polarity of a sentence when there is a negation word in the middle, does not seem to always produce accurate results. Finally, as proposed, in sentences that contain few words that express emotion, the overall sentiment should lean more towards the sentiment type of these words (either positive or negative).

6.2 Challenges and Limitations

This study also revealed a handful of challenges and limitations that have to do with the implementation of sentiment analysis. For example, the tweets that were collected were limited to be in English. This leaves out a significant part of the world where English is not the primary language but, since COVID-19 is a global issue, people will still post online in other languages. Furthermore, the pandemic situation is still ongoing, which means that topics of discussion constantly change both online and offline. For instance, online conversation about the vaccines was much different during the first wave of the pandemic in 2020 when there were no vaccines available, and completely different now that there is vaccine availability. This dynamic concept cannot be dealt with currently, with the existing tools. Finally, another problem that is yet to be solved, is the correct understanding of sarcasm and irony by both TextBlob and VADER. This way of expression is closely related to anger and aggression which should classify the text automatically as negative, but most of the times, the two lexicons identified the text as positive.

6.3 Future Work

Concluding this research there are aspects that could potentially be modified with the purpose of improving the results. Some future work, that relates to the subject of this dissertation, would be combining multiple social media platforms to get data and then conduct a cross-platform analysis. This way, the conclusions can be more complete since content about a specific subject is rarely shared in only one social media platform. Furthermore, in the future, a more specialized tool can be developed that will conduct the process of analyzing sentiment analysis results. For instance, by taking into consideration other parameters like the number of retweets, account followers, geographic location, etc. At last, it would be beneficial for understanding this pandemic, to continue conducting similar studies until its end. Through this process we will be able to understand the actual footprint of vaccination in people's opinion and how this connects to the number of vaccinations worldwide. After all, studies conducted today can become the foundation of building a better future, and this might lead to humanity being better prepared for similar trials in the years to come.

References

[1] “In Hunt for Covid-19 Origin, Patient Zero Points to Second Wuhan Market @ www.wsj.com.” [Online]. Available: https://www.wsj.com/articles/in-hunt-for-covid-19-origin-patient-zero-points-to-second-wuhan-market-11614335404.

[2] CDC, “Symptoms @ Www.Cdc.Gov.” 2020, [Online]. Available: http://www.cdc.gov/norovirus/about/symptoms.html.

[3] D. P. Oran and E. J. Topol, “The Proportion of SARS-CoV-2 Infections That Are Asymptomatic : A Systematic Review,” *Ann. Intern. Med.*, vol. 174, no. 5, pp. 655–662, 2021, doi: 10.7326/M20-6976.

[4] WHO, “Index @ Covid19.Who.Int,” *Https://Covid19.Who.Int/*. p. 1, 2020, [Online]. Available: https://covid19.who.int/%0Ahttps://www.who.int/health-topics/coronavirus#tab=tab_3.

[5] “covid19.trackvaccines.org.” [Online]. Available: https://covid19.trackvaccines.org/vaccines/approved/.

[6] “natural-language-processing-NLP @ www.techtarget.com.” [Online]. Available: https://www.techtarget.com/searchenterpriseai/definition/natural-language-processing-NLP.

[7] “Index @ Www.Phpmyadmin.Net.” [Online]. Available: https://www.phpmyadmin.net/.

[8] MySQL, “What-Is-Mysql @ Dev.Mysql.Com.” 2020, [Online]. Available: https://dev.mysql.com/doc/refman/5.7/en/what-is-mysql.html.

[9] D. Kuhlman, “A Python Book,” *A Python B.*, pp. 1–227, 2013.

[10] “textblob.readthedocs.io.” [Online]. Available: https://textblob.readthedocs.io/en/dev/.

[11] E. Hutto, C.J. and Gilbert, “VADER: A Parsimonious Rule-based Model for,” *Eighth Int. AAAI Conf. Weblogs Soc. Media*, p. 18, 2014, [Online]. Available: https://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/viewPaper/8109.

[12] S. Tafuri, M. S. Gallone, M. G. Cappelli, D. Martinelli, R. Prato, and C. Germinario, “Addressing the anti-vaccination movement and the role of HCWs,” *Vaccine*, vol. 32, no. 38, pp. 4860–4865, 2014, doi:

10.1016/j.vaccine.2013.11.006.

[13] “ten-threats-to-global-health-in-2019 @ web.archive.org.” [Online]. Available: https://web.archive.org/web/20190627025209/http://www.who.int/emergencies/ten-threats-to-global-health-in-2019.

[14] D. Rousidis, P. Koukaras, and C. Tjortjis, *Social media prediction: a literature review*, vol. 79, no. 9–10. Multimedia Tools and Applications, 2020.

[15] P. Koukaras, C. Tjortjis, and D. Rousidis, *Social Media Types: introducing a data driven taxonomy*, vol. 102, no. 1. Springer Vienna, 2020.

[16] T. Sakaki, M. Okazaki, and Y. Matsuo, “Tweet analysis for real-time event detection and earthquake reporting system development,” *IEEE Trans. Knowl. Data Eng.*, vol. 25, no. 4, pp. 919–931, 2013, doi: 10.1109/TKDE.2012.29.

[17] L. Samaras, E. García-Barriocanal, and M. A. Sicilia, “Comparing Social media and Google to detect and predict severe epidemics,” *Sci. Rep.*, vol. 10, no. 1, pp. 1–11, 2020, doi: 10.1038/s41598-020-61686-9.

[18] O. Lazaros and C. Tjortjis, “Presidential Elections using Data extracted from,” *2018 South-Eastern Eur. Des. Autom. Comput. Eng. Comput. Networks Soc. Media Conf.*, pp. 1–8.

[19] A. Mishra, M. S. Wajid, and U. Dugal, “A Comprehensive Analysis of Approaches for Sentiment Analysis Using Twitter Data on COVID-19 Vaccines,” vol. 02, no. 009, pp. 1–10, 2021.

[20] Y. Wang, J. Callan, and B. Zheng, “Should we use the sample? Analyzing datasets sampled from Twitter’s stream API,” *ACM Trans. Web*, vol. 9, no. 3, 2015, doi: 10.1145/2746366.

[21] N. S. Sattar and S. Arifuzzaman, “Covid-19 vaccination awareness and aftermath: Public sentiment analysis on twitter data and vaccinated population prediction in the usa,” *Appl. Sci.*, vol. 11, no. 13, 2021, doi: 10.3390/app11136128.

[22] T. Na, W. Cheng, D. Li, W. Lu, and H. Li, “Insight from NLP Analysis: COVID-19 Vaccines Sentiments on Social Media,” pp. 1–10, 2021, [Online]. Available: http://arxiv.org/abs/2106.04081.

[23] “It’s the ‘vaccine hesitant’, not anti-vaxxers, who are troubling public health experts by Gaby Hinsliff - www.theguardian.com.” [Online]. Available:

https://www.theguardian.com/commentisfree/2020/nov/16/vaccine-hesitant-anti-vaxxers-public-health-experts-covid.

[24] S. L. Wilson and C. Wiysonge, "Social media and vaccine hesitancy," *BMJ Glob. Heal.*, vol. 5, no. 10, pp. 1–7, 2020, doi: 10.1136/bmjgh-2020-004206.

[25] Z. Xu and H. Guo, "Using Text Mining to Compare Online Pro- and Anti-Vaccine Headlines: Word Usage, Sentiments, and Online Popularity," *Commun. Stud.*, vol. 69, no. 1, pp. 103–122, 2018, doi: 10.1080/10510974.2017.1414068.

[26] I. Herrera-Peco *et al.*, "Antivaccine movement and covid-19 negationism: A content analysis of spanish-written messages on twitter," *Vaccines*, vol. 9, no. 6, pp. 1–14, 2021, doi: 10.3390/vaccines9060656.

[27] G. J. Rubin and S. Wessely, "The psychological effects of quarantining a city," *BMJ*, vol. 368, no. January, pp. 1–2, 2020, doi: 10.1136/bmj.m313.

[28] N. F. Johnson *et al.*, "The online competition between pro- and anti-vaccination views," *Nature*, vol. 582, no. 7811, pp. 230–233, 2020, doi: 10.1038/s41586-020-2281-1.

[29] Z. Chu, S. Gianvecchio, H. Wang, and S. Jajodia, "Detecting automation of Twitter accounts: Are you a human, bot, or cyborg?," *IEEE Trans. Dependable Secur. Comput.*, vol. 9, no. 6, pp. 811–824, 2012, doi: 10.1109/TDSC.2012.75.

[30] O. Varol, E. Ferrara, C. A. Davis, F. Menczer, and A. Flammini, "Online Human Bot Interaction," *Proc. Elev. Int. AAAI Conf. Web Soc. Media (ICWSM 2017)*, no. Icwsm, pp. 280–289, 2017, [Online]. Available: www.aaai.org.

[31] X. Yuan, R. J. Schuchard, and A. T. Crooks, "Examining emergent communities and social bots within the polarized online vaccination debate in Twitter," *Soc. Media+ Soc.*, vol. 5, no. 3, p. 2056305119865465, 2019.

[32] A. Egli, "Bad Robot : A Preliminary Exploration of the Prevalence of Automated Software Programmes and Social Bots in the COVID-19 # antivaxx Discourse on Twitter," no. July, 2021.

[33] A. Goel and A. Mittal, "Stock prediction using twitter sentiment analysis. Standford University, CS229," *Cs229.Stanford.Edu*, no. December, pp. 1–5, 2012, [Online]. Available: http://cs229.stanford.edu/proj2011/GoelMittal-StockMarketPredictionUsingTwitterSentimentAnalysis.pdf.

[34] T. H. Nguyen, K. Shirai, and J. Velcin, "Sentiment analysis on social media for

stock movement prediction," *Expert Syst. Appl.*, vol. 42, no. 24, pp. 9603–9611, 2015, doi: 10.1016/j.eswa.2015.07.052.

[35] S. Sohangir, N. Petty, and Di. Wang, "Financial Sentiment Lexicon Analysis," *Proc. - 12th IEEE Int. Conf. Semant. Comput. ICSC 2018*, vol. 2018-Janua, pp. 286–289, 2018, doi: 10.1109/ICSC.2018.00052.

[36] I. Nasiara, "" The Impact of Twitter Sentiment on Ryanair ' s Business Performance "," no. December, 2019.

[37] M. Mujahid *et al.*, "Sentiment analysis and topic modeling on tweets about online education during covid-19," *Appl. Sci.*, vol. 11, no. 18, 2021, doi: 10.3390/app11188438.

[38] S. Gu, F. Wang, N. P. Patel, J. A. Bourgeois, and J. H. Huang, "A model for basic emotions using observations of behavior in Drosophila," *Front. Psychol.*, vol. 10, no. APR, pp. 1–13, 2019, doi: 10.3389/fpsyg.2019.00781.

[39] R. E. Jack, W. Sun, I. Delis, O. G. B. Garrod, and P. G. Schyns, "Four not six: Revealing culturally common facial expressions of emotion.," *J. Exp. Psychol. Gen.*, vol. 145, no. 6, pp. 708–730, Jun. 2016, doi: 10.1037/xge0000162.

[40] A. Ben-Zeev, "The nature of emotions," *Philos. Stud.*, vol. 52, no. 3, pp. 393–409, 1987, doi: 10.1007/BF00354055.

[41] A. S. Cowen and D. Keltner, "Self-report captures 27 distinct categories of emotion bridged by continuous gradients," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 114, no. 38, pp. E7900–E7909, 2017, doi: 10.1073/pnas.1702247114.

[42] S. An, L. J. Ji, M. Marks, and Z. Zhang, "Two sides of emotion: Exploring positivity and negativity in six basic emotions across cultures," *Front. Psychol.*, vol. 8, no. APR, pp. 1–14, 2017, doi: 10.3389/fpsyg.2017.00610.

[43] D. Beleveslis, C. Tjortjis, D. Psaradelis, and D. Nikoglou, "A hybrid method for sentiment analysis of election related tweets," *2019 4th South-East Eur. Des. Autom. Comput. Eng. Comput. Networks Soc. Media Conf. SEEDA-CECNSM 2019*, pp. 1–6, 2019, doi: 10.1109/SEEDA-CECNSM.2019.8908289.