

# TWITTER SENTIMENT ANALYSIS DURING COVID-19 IN FLORIDA

by

JINGYI LI

(Under the Direction of Tianming Liu)

## ABSTRACT

Sentiment analysis is a popular approach to figure out people's thoughts by digging into human-generated text content from online users. This study focuses on tweets from the Twitter platform to investigate people's sentiments and emotions related to COVID-19 pandemic. Tweets from Florida users have been collected and preprocessed on three periods of time, April 22-28, July 15-21, and October 14-20, to study the changes in sentiments, subjectivities, and emotions. Two sentiment analysis approaches, TextBlob and NRCLex, have been used and compared in this study: TextBlob has strength in sentiment classifications and defining subjectivities, while the NRC Lexicon method provides more detailed emotional states analysis. The results from both methods show that neutral sentiments are the majority and positive sentiments outweigh negative sentiments. However, TextBlob indicates a more significant difference between positive and negative than NRC. We also observe that the positive attitudes decrease during the week in July and increase back in October.

INDEX WORDS: Sentiment analysis, Coronavirus, COVID-19, Twitter, TextBlob, NRCLex, Tweepy

TWITTER SENTIMENT ANALYSIS DURING COVID-19 IN FLORIDA

by

JINGYI LI

BS, University of Washington, 2017

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial

Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2020

© 2020

Jingyi Li

All Rights Reserved

TWITTER SENTIMENT ANALYSIS DURING COVID-19 IN FLORIDA

by

JINGYI LI

Major Professor: Tianming Liu  
Committee: Adam Goodie  
Frederick Maier

Electronic Version Approved:

Ron Walcott  
Vice Provost for Graduate Education and Dean of the Graduate School  
The University of Georgia  
December 2020

## ACKNOWLEDGEMENTS

The completion of this undertaking could not have been possible without so many people's participation and assistance.

First, I want to thank my major professor, Dr. Tianming Liu. He always has the brightest sight to find interesting topics and the right directions for me. Even though at the beginning, I had a time of doubting my strength and reasons for doing this research, but it turned out to be meaningful for me and would benefit me so much in my career. I also want to appreciate my committee member, Dr. Adam Goodie, and Dr. Fred Maier. I learned a lot of professional skills and new knowledge in their classes. They have witnessed my growth and have always been there whenever I need them.

It is not an easy journey for me to pursue an MS degree in AI, especially without any previous computer science knowledge or coding experience. I want to give a ton of thanks to Zichu Zheng, Gengyu Zhang, Heng-Shiou Sheu, and Tsung-Wei Hu. Without their help, it's impossible for me to complete my degree.

I also wish to express my love to my fiancé Bomeng Li and MerMer, the cat. They are always warm and supportive for me to live away from home and accompany me through hard times. They are the most significant ones I met in my life journey, and I would wholeheartedly spend the rest of my life with them.

## TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS .....	iv
LIST OF TABLES .....	vii
LIST OF FIGURES .....	viii
CHAPTER	
1 INTRODUCTION .....	1
2 BACKGROUND AND RELATED WORK.....	5
Why we study Covid-19? .....	5
Why we focus on Florida, US?.....	6
Why we choose Twitter to carry out sentiment analyses?.....	12
3 DATA AND METHODOLOGY .....	16
Methods Review .....	16
Dataset .....	18
Implementation and Procedure .....	21
Understand the Algorithm Behind.....	24
4 RESULTS AND DISCUSSION.....	26
TextBlob - Polarity Analysis .....	27
TextBlob - Subjectivity Analysis .....	30
TextBlob - Emotion Analysis .....	31
NRC - Emotion Affects .....	35

NRC - Sentiment Classification .....	40
Discussion of Sentiment Shifts and Method Comparison .....	42
5 CONCLUSION .....	46
6 LIMITATION AND FUTURE WORK .....	48
REFERENCES .....	51

## LIST OF TABLES

	Page
Table 2.1: Coronavirus Cases and Per 100,000 Cases Data in Four States by October 23rd, 2020 .....	10
Table 3.1: Number of Tweets Collected on Each Day Related to COVID-19.....	20
Table 4.1: Most Frequently Appeared Words in the Cleansed Tweets from Different Periods .....	27
Table 4.2: Emotional Guidance Scale .....	32
Table 4.3: NRC Emotion Affects Percentage for the “Beginning Phase”.....	36
Table 4.4: NRC Emotion Affects Percentage for the “Burst Phase”.....	36
Table 4.5: NRC Emotion Affects Percentage for the “Recovery Phase” .....	37

## LIST OF FIGURES

	Page
Figure 2.1: Cumulative Curve by Date for the Ten Countries with the Highest Absolute Daily Number of Deaths.....	7
Figure 2.2: Distribution Map of Cumulative Confirmed Cases and Cumulative Death Cases in the US by October 23, 2020.....	9
Figure 2.3: Daily New Cases of COVID-19 Reported in Florida from March to October 2020 with 7-day Average Calculated .....	11
Figure 3.1: Illustration of Geographical Range of the Search.....	20
Figure 3.2: Flowchart for sentiment analysis procedures.....	23
Figure 4.1: Sentiment classification on tweets for the “Beginning Phase” by the TextBlob method .....	28
Figure 4.2: Sentiment classification on tweets for the “Burst Phase” by the TextBlob method .....	29
Figure 4.3: Sentiment classification on tweets for the “Recovery Phase” by the TextBlob method .....	29
Figure 4.4: Average subjectivity results for tweets from three weeks showing the composition of objective, subjective, and neutral in percentage.....	31
Figure 4.5: Distribution of emotions by matching the TextBlob polarities to the Emotional Guidance Scale for the “Beginning Phase” .....	33

Figure 4.6: Distribution of emotions by matching the TextBlob polarities to the Emotional Guidance Scale for the “Burst Phase” .....	33
Figure 4.7: Distribution of emotions by matching the TextBlob polarities to the Emotional Guidance Scale for the “Recovery Phase” .....	34
Figure 4.8: Rankings of average emotional affects for the “Beginning Phase” with the NRC method .....	38
Figure 4.9: Rankings of average emotional affects for the “Burst Phase” with the NRC method .....	39
Figure 4.10: Rankings of average emotional affects for the “Recovery Phase” with the NRC method .....	39
Figure 4.11: Sentiment classification on tweets for the “Beginning Phase” by the NRC method .....	41
Figure 4.12: Sentiment classification on tweets for the “Burst Phase” by the NRC method .....	41
Figure 4.13: Sentiment classification on tweets for the “Recovery Phase” by the NRC method .....	42
Figure 4.14: Sentiment classification of positive, negative, and neutral comparison over time with the TextBlob method based on polarity scores .....	43
Figure 4.15: Sentiment classification of positive, negative, and neutral comparison over time with the NRC method based on the count of positive and negative labels ...	43

## CHAPTER 1

### INTRODUCTION

A novel coronavirus swipes the world by the end of December 2019. Around 180 countries have been involved, with overall, more than 70 million people have been infected by the time of writing this document (John Hopkins University Coronavirus Research Center, 2020). There have been around 16 M confirmed cases and 300 K fatal cases among all infected people in the USA, and these numbers keep rising every day. The first case of Covid-19 was reported in Wuhan province in China; however, the virus was out of control and developed explosively worldwide in merely several months.

Many countries have taken regulations and policies to guide people to fight against this global pandemic. In the US, regulations such as social distancing and shelter-in-place have been most widely applied. This prevention guidance has impacted human life in every way because they forced people to stay away from each other and at the same time blocked the social connections in business, education, public events, and many other daily activities. Many businesses were affected by these policies and had to shut down, and people were unemployed. As a result, there were growing stresses and concerns among people and in the communities.

In recent years, online sentiment analysis has been considered one of the most popular methods to dig into people's thoughts. It relies on people's habits of expressing feelings, sharing ideas, and holding discussions through social media. The sentiment

analysis applications are widely covered by many areas, such as opinion mining, recommender systems, and event detection (Sailunaz & Alhadj, 2019).

As the coronavirus problem becomes severe, people worldwide care about it and send their thoughts through many forms like bloggers, comments, micro-bloggers, text messages, and others. Studying these posts' contents can help governments and organizations understand people's needs, concerns, satisfying requirements, and things needing improvement. Sentiment analysis is a method of reaching these goals and a way of studying online content. Tracking and analyzing tweets could be a reasonable method to dig into people's thoughts, behavior, and reactions regarding the Covid-19 problem because people's thoughts change every moment in different situations. Moreover, tweets reflect real-time responses from the masses to understand the sentiment trends during this pandemic better.

To reveal the viewpoints from people in Florida on how they think about the coronavirus problem, we have conducted sentiment analysis based on tweets related to COVID-19 and collected them in the range of Florida state in this study. We have used Twitter API to collect and process tweets related to coronavirus topics with hashtags #coronavirus, #covid-19, and #covid, from April 22nd to 28th, July 15th to 21st, and October 14th to 20th. The geographical location of the data source has been set to cover the range of the Florida state. The raw tweets have been sent to a data cleansing process and have been stored in CVS files. The files contain two columns of useful information, including timestamps of each tweet and cleansed tweet texts.

The datasets have been sent to do the sentiment analysis through two different methods. One of them uses the TextBlob with a pre-trained machine learning model to

predict each tweet's sentiments by returning its sentiment polarity and subjectivity. The other method uses the NRC Lexicon method with a pre-implemented Emotion Lexicon package to predict the emotional affects in each tweet by calculating the annotated emotion scores pre-defined for each word occurring in that tweet. Both methods process and classify all tweets into three classes: positive, negative, or neutral. The emotions associated with each tweet have also been conducted and discussed by matching the polarity scores to an emotional guidance chart in the TextBlob method. For the NRC Lexicon method, the emotions have been directly calculated with the semantics of the words in that tweet using the built-in emotional scale features. The emotions include anticipation, joy, surprise, trust, anger, sadness, disgust, and fear, and the results are returned in terms of counts of emotional states.

In this thesis, we have primarily,

- 1) Developed sentiment analysis on Florida tweets to study the percentages of positive, negative, and neutral comments from people by computing the polarity scores and conducting the emotions associated with the tweets during unpredictable times.
- 2) Observed and analyzed the sentiment changes from three different time slots.
- 3) Compared the differences in results from both experiment methods and discussed the advantages and disadvantages of each approach.

To achieve these goals, we expanded analyses in detail and provided the over-time discussions. The results concluded that the majority of tweets from people in Florida showed a neutral attitude. Despite the neutral tolls, positive sentiments outweighed the negative sentiments for all days in this study. However, we could observe that the positive tolls had a noticeable drop during the week in July and following with a rise back

in October. Although the sentiment was only conducted into three classes, the emotions behind the masses were more complicated. People expressed fear, anger, and sadness in their comments. However, many Florida citizens were still holding a positive and hopeful approach and showing trust in the government and themselves. The results from both approaches are generally consistent with each other but with minor differences, and both methods have been found useful in doing the sentiment analysis on tweets. Each of them has strengths with particular focuses: TextBlob has been found easier on classifying comments into positive, negative, and neutral classes, and it offers the subjectivity scores. We recommend using the NRC Lexicon method if the study focuses on more than sentiment classes, but with the needs of more complicated emotional affects behind the text.

The remainder of the paper is organized in the following chapters. In Chapter 2, we will go through the problem backgrounds and illustrate this research's reasons and meanings. Chapter 3 will go over methods of sentiment analysis and introduce our approach in detail. In Chapter 4, the results and analysis will be provided along with the visualized charts discussing the analysis results based on two different approaches and talking of sentiment changes among people during this weird time. Chapter 5 and 6 will provide conclusions and go over the limitations of this research paper and mention some possible future work improvements.

## CHAPTER 2

### BACKGROUND and RELATED WORK

Coronavirus became a popular topic in researches because it widely influenced people's lives in many ways. This section provides an overview of our problem selection criteria by answering the following questions:

- why it is essential to study COVID-19
- why we choose to focus our study on Florida state
- why we choose Twitter to carry out sentiment analyses

We will offer detailed discussions clarifying the problem and walking through related literature works to support our study.

#### **Why we study COVID-19?**

On the last day in December 2019, a cluster of extreme respiratory contaminations was reported in Wuhan, Hubei Province, China. The cases appeared to be related to Huanan Seafood Market because most of the patients had experience visiting or working in the market (Organization, 2020d). The cases were confirmed to be associated with a novel coronavirus, named 2019-nCoV or COVID-19, by health authorities in China on January 7th, 2020 (Organization, 2020c). In the beginning, experts believed that the diseases were generated from the seafood market and could not transmit among people; however, the later epidemiologic data proved it wrong and indicated that person-to-person transmission of the virus is more robust than expected (Chan et al., 2020; Phan et al., 2020).

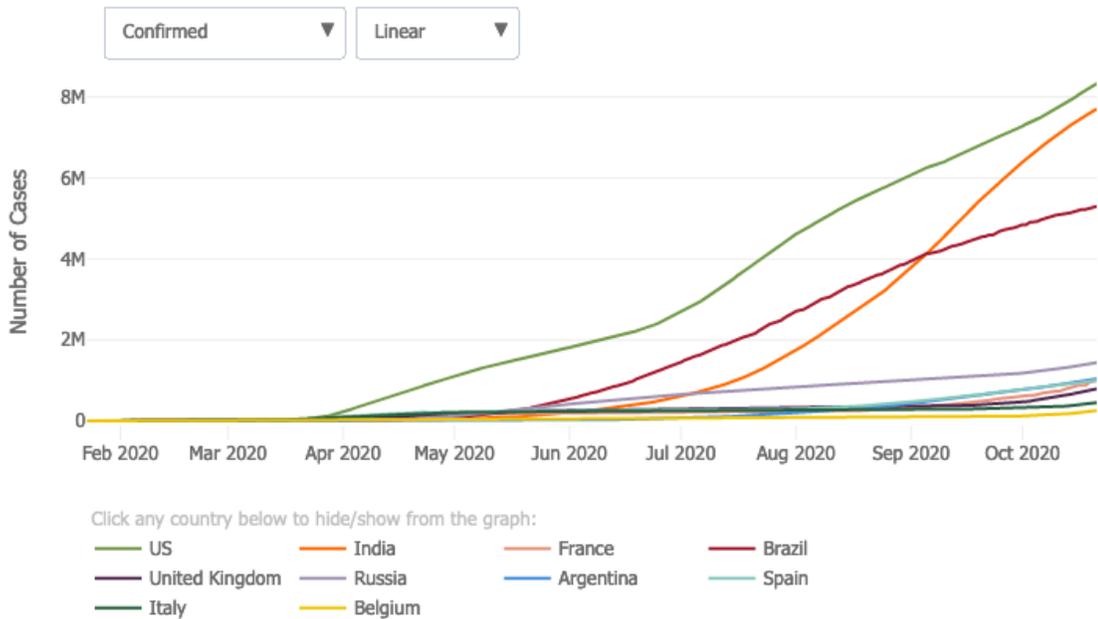
People were not ready for this pandemic, and the spread of the disease was out of control. Just three months from the very first case, there were 823,626 confirmed cases and 40,598 death globally as of April 1, 2020, and the number almost doubled to 1,610,909 confirmed cases and 99,690 death by ten days after that as of April 11 (Organization, 2020a, 2020b). According to John Hopkins University Coronavirus Resource Center, 189 countries and regions had been involved by October 22, 2020.

Before the coronavirus broke out worldwide, people lack awareness about this novel virus and gave it time to spread. Looking back at the epidemiologic history, the COVID-19 is not the first outbreak of flu, but the speed of its development was scary. The virus's diagnosis seemed similar to seasonal flu initially, but the seriousness level of symptoms went up as more patients involved. There is no approved medication for the treatment of COVID-19 even after millions of patients died (Jahangir, Muheem, & Rizvi, 2020). People started to panic, especially when more and more cases were reported with no travel history or contact with known patients (Bajema et al., 2020). The shortage of medical equipment also raised the panic level because masks were in short and very difficult to purchase. Masses became too worried; consequently, the screening kits were not enough to fill the demands. Many people tried to get testing equipment even with light symptoms similar to the coronavirus (Beetz et al., 2020).

### **Why we focus on Florida, US?**

The situation in the US is typically severe. The US's first infection case was confirmed on January 20, 2020, but the circumstances grew seriously from then (Holshue et al., 2020). Figure 2.1 shows the pandemic's cumulative trends over time in countries with the highest absolute daily number of deaths. The green curve showed how seriously

the US cumulative cases develop from late January 2020 to October 2020. We could tell from the plot that the confirmed cases rapidly grew up and infected more than 4 million people in less than half a year from the beginning. Although India's spread is growing fast recently, the US is still the top infected country in the world (John Hopkins University Coronavirus Research Center, 2020).



*Figure 2.1.* Cumulative Curve by Date for the Ten Countries with the Highest Absolute Daily Number of Deaths

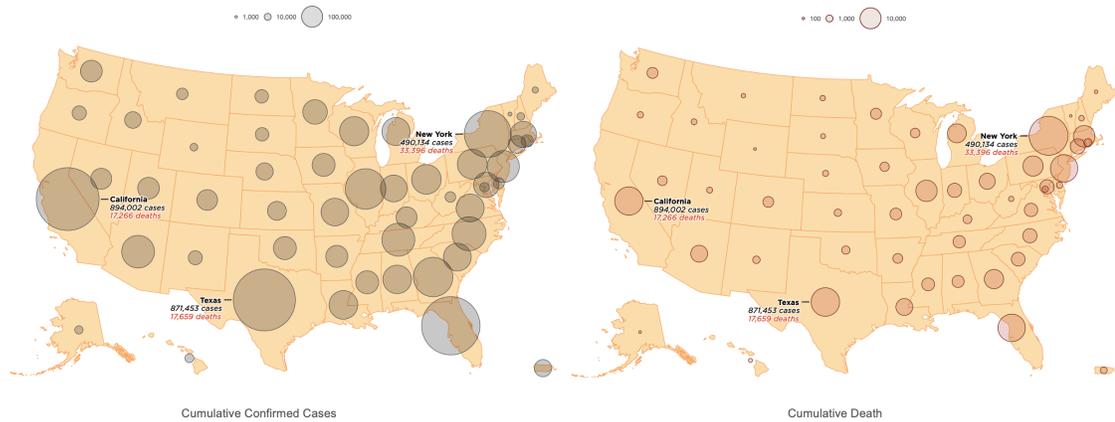
Current research carried out people's sentiment analysis from different places globally, analyzing people's thoughts toward pandemic. Most of them analyzed at a bigger scale, such as European countries, the Americas, or even several continents. Dubey conducted a country-wise sentiment analysis for twelve different countries and

inspect how different citizens from different parts of the world reacted to the virus in a specific period (Dubey, 2020). Some others were investigating a specific country. For example, Ogbuju et al. researched Nigeria by analyzing Nigerians' emotions within the lockdown exercise period to understand the policy's effectiveness and its impact on individuals (Ogbuju et al., 2020). Pran et al. studied the country of Bangladeshi, and the research focused on people's comments from several Facebook news in the Bangla language (Pran, Bhuiyan, Hossain, & Abujar, 2020). Pokharel's research targeted people sharing their locations as "Nepal" and collecting and mining social media posts to understand their attitude (Pokharel, 2020). Since sentiments are from people, choosing a good group of people to investigate becomes very important to make sure the research meaningful.

Although there have been tons of literature on this topic, the study of a more specific geographic area is still missing. By the end of August, the Americas countries account for about 13% of the world's population but produce more than 50% of officially reported global cases and deaths (World Health Organization, 2020). The United States are deeply involved in this pandemic with the highest number of confirmed cases and deaths and could be the right target country to study. However, individuals' thoughts vary a lot from each other depending on their backgrounds, experiences, religions, or occupations, even for people from the same country (Zarghami, 2020). There might be a problem when people from some parts of the country feel optimistic about the pandemic; however, people from the rest parts are taking a negative attitude and resulting in a neutral conclusion seeing the US as a whole. The pandemic situation in each state is very different from each other, and the policies and regulations also vary a lot. Therefore,

studying the entire US country without understanding the reason would not be helpful in the future. In this research, to eliminate the bias generated by taking a large corpus of people and produce more valuable thoughts for the future, we focus on a more specific region, Florida State.

The decision of choosing Florida as our research target state depends on information from several aspects, including population, cumulative confirmed cases, confirmed case ratio, and death cases. A distribution map of cumulative confirmed cases and cumulative death cases are shown in Figure 2.2 based on the data reported at the JHU COVID-19 research center, the data source is from WHO (World Health Organization, 2020), and the size of the circles on the map reflects the data's size.



*Figure 2.2.* Distribution Map of Cumulative Confirmed Cases and Cumulative Death Cases in the US by October 23, 2020

There are four regions with large circles in both maps worth a discussion: southwestern (CA), southern (Texas), northeastern (NY), and southeastern (FL). We have

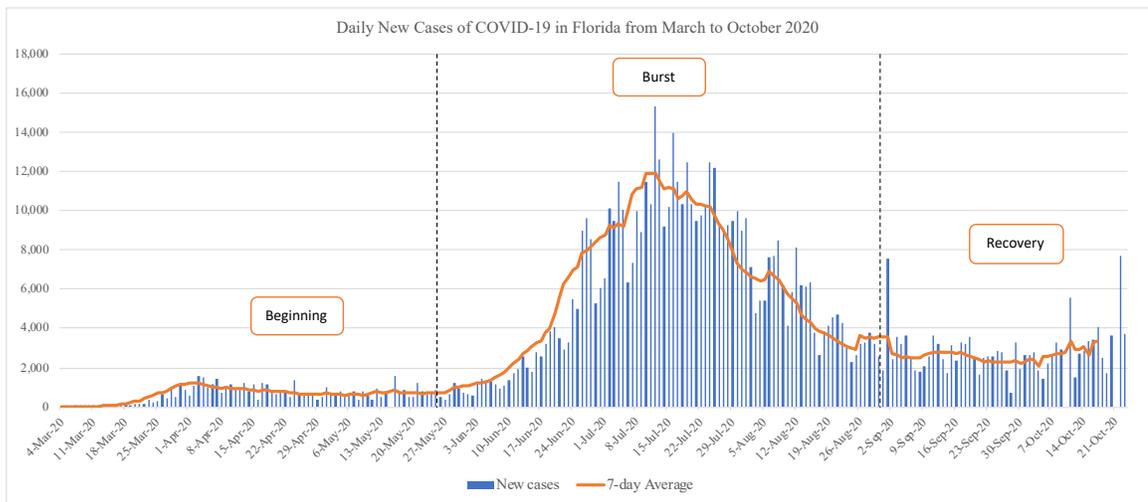
conducted a table with the confirmed cases, deaths, cases per 100,000 population, and deaths per 100,000 population in the four different states in Table 2.1 with the data aggregated from multiple official data sources in the COVID-19 Data Repository (Gardner & Dong, 2020) provided by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University by October 23, 2020. The seriousness of COVID-19 in Florida is not hard to see, making it a typical region to study. Considering the population and confirmed cases information, Florida has the highest confirmed cases per 100,000 of 3,576.2 among the four regions, meaning its infection chance is the highest. Although the death number of FL is the lowest among the four states, its death rate is much worse than the other two regions except NY.

Table 2.1  
*Coronavirus Cases and Per 100,000 Cases Data in Four States by October 23, 2020*

		Total	Per 100,000
<b>Florida</b>	Confirmed Cases	768,091	3,576.2
	Deaths	16,267	75.7
<b>New York</b>	Confirmed Cases	490,134	2,519.5
	Deaths	33,396	171.7
<b>California</b>	Confirmed Cases	894,002	2,262.6
	Deaths	17,266	43.7
<b>Texas</b>	Confirmed Cases	871,453	3,005.4
	Deaths	17,659	60.9

Extracting data from CDC (Centers for Disease Control and Prevention, 2020), a plot of confirmed daily new cases along with the calculated 7-day average curve in Florida is displayed from March to October shown in Figure 2.3. Different from the NY state, the pandemic in Florida did not happen as a burst in a short time; it experienced

three stages: the beginning, the burst, and the recovery period. The changes in the stages and the longer time intervals would give people more time to think and respond so that the opinions and sentiments expressed by their people could be more prosperous. Mining people’s sentiments in different phases could allow us to view the problem on a larger scale.



*Figure 2.3. Daily New Cases of COVID-19 Reported in Florida from March to October 2020 with 7-day Average Calculated*

Another factor that makes the study focusing on Florida more unbiased might be the political background, especially when 2020 is the election year. Johnson et al., in their paper, mentioned the political effects of COVID-19 and possible swing states of 2020 (Johnson, Pollock, & Rauhaus, 2020). So far, there is insufficient research on how different people support the two parties in the US regarding coronavirus problems. However, choosing Florida that has been a swing state for the recent 20+ years, would be

less biased than California (more Democratic) or Texas (Republican dominant) on the target problem (Savidge, 2017).

### **Why we choose Twitter to carry out sentiment analyses?**

Sentiment analysis (SA) or opinion mining is a subfield of Natural Language Processing in Artificial Intelligence using NLP and machine learning techniques to explicate and classify thoughts and emotions from subjective data (Hasan, Moin, Karim, & Shamshirband, 2018). The use of sentiment analysis is widely applied in various disciplines such as political and social news, science and technology, public health, history, arts, and others (Alessia, Ferri, Grifoni, & Guzzo, 2015). For example, in business fields, companies utilize SA to perceive sentiment in social data, gauge brand reputation, and understand their customers (Alessia et al., 2015). Product managers could detect product reviews' feedback to study the quality of the products or services to improve user experiences and satisfaction scores (Fang & Zhan, 2015).

Traditional sentiment analysis is commonly done by carrying out opinion polls or questionnaires to examine the client produced substance and text-based overview results with numbers, scores, and appraisals (Bishop, 2004). Although these analyses provide guidance and help manage efforts towards the most noteworthy changes to make, however, manual investigation of text is very tedious and almost difficult to do reliably. Without a proficient cycle set up, associations will remain to a great extent in obscurity about how to make their improvements. With the advancements in communication technology and better access to social networks, people tend to express their viewpoints and feelings on social media platforms, such as Twitter, Facebook, and YouTube, with friends either from reality or virtually connected with their accounts (Pokharel, 2020). In

the research of Oliveira et al. in 2017, they studied political preferences among people based on tweets from Twitter related to the Brazilian presidential candidates during the second round of elections in 2014. They investigate the performance of sentiment analysis of data extracted from social media compared to traditional public opinion surveys regarding the accuracy of the results. Their conclusion supported that there was no apparent disparity in results between the two data sources. The sentiment analysis based on social media data could reveal citizens' political preferences at high accuracy and be recommended as a good source (Oliveira, Bermejo, & dos Santos, 2017).

Consistently, many websites and online web-based media produce many data, and analyzing and classifying those data could bring more knowledge and information to us (Iglesias, Tiemblo, Ledezma, & Sanchis, 2016). For example, in web news mining, the use of "Big Data" and news mining-based automatic system could help track, assess, and categorize daily news and help organizers properly manage news articles (Guellil & Boukhalifa, 2015).

Social networks provide convenient and easy access to data. However, not all of the information is valuable, so that choosing appropriate data sources and domains from the vast data pool would become a fundamental problem in any research area. Ravindran and Garg, in their book *Mastering social media mining with R*, mentioned six types of Internet-based platform that facilitate people's emotions (Ravindran & Garg, 2015):

1. Networking platforms: services that allow members to interact with other people of similar interests and background, such as Facebook and LinkedIn.
2. Micro-blogging platforms: services that focus on short updates that are pushed out to followers and subscribers, such as Twitter and Tumblr.

3. Photo sharing platforms: services that allow users to share their self-generated photos, such as Instagram and Flickr.
4. Video sharing platforms: services that allow users to share their self-generated videos, such as YouTube and Vimeo.
5. Stack exchanging platforms: online forums that allow users to exchange ideas and hold conversations around the posts, such as GitHub.
6. Instant message platforms: services that allow users to communicate directly with their friends, such as WhatsApp and WeChat.

Micro-blogging, with its easy access, frequent posts, real-time responses, and less content-generating time, provide people a platform to share their opinions anytime from virtually any corner of the world. The Twitter social network is the largest and the most popular micro-blogging platform considered a liable and useful data source by researchers (Kouloumpis, Wilson, & Moore, 2011). According to Twitter officials' usage statistics data, there are more than 330 million monthly active users and 145 million daily users producing 500 million tweets per day by Q1 2019. Among the daily active users, 30 million are from the United States (21%), so the studies of tweets geographically based on the US would have an extra advantage with its solid user base.

Many recent works used Twitter as the source to carry out sentiment analysis in a variety forms of applications, such as to predict political preferences, to investigate the effectiveness of a service or policy, and to monitor the infectious disease and public health crisis (Russell, 2013; Tang, Chang, & Liu, 2014). For example, H. Wang et al. built a real-time SA system for the 2012 US presidential election cycle. They provided accurate online political landscapes based on public sentiments expressed on Twitter

(Wang, Can, Kazemzadeh, Bar, & Narayanan, 2012). Chew, and Eysenbach conducted an “infodemiology” study for public health with a content analysis of tweets in 2010 when H1N1 flu burst worldwide. They tracked the use of the terms “H1N1” and “swine flu” overtime to validate real-time public attention trends (Chew & Eysenbach, 2010). When the Ebola virus blew up, Fung et al. systematically reviewed twelve existing research pieces related to the Ebola virus and social media in a cross-sectional manner and proved the utility of using online-based SA in public health discipline (Fung et al., 2016). Of the 12 articles, 7 of them were from Twitter studies, and this also reflected the trends of sentiment analysis and emphasized Twitter’s preference over other social media among researchers on SA problems.

## CHAPTER 3

### DATA and METHODOLOGY

This section will walk through some of the most commonly used Twitter sentiment analysis methods employed in recent literature related to COVID-19 disease and generally discuss some limitations based on their research. Then we will describe our research process with improvements and explain our data analyzing approach in detail.

#### **Methods Review**

There are two major approaches for sentiment analysis: lexicon-based methods and machine learning-based methods. A dictionary of words with pre-defined meanings is annotated with associate semantic polarity and sentiment strength in lexical methods. There is no training process in this approach, and the polarity score in each sentence is solely based on the semantic of each word with calculations (Taboada, Brooke, Tofiloski, Voll, & Stede, 2011). In machine learning-based methods, it requires a dataset training step before applying for the sentiment classification. A model needs to be trained with pre-labeled examples to extract features to predict on test data and classify paragraphs with positive, neutral, and negative polarity in a contextual manner (Hasan et al., 2018).

Dubey's recent research was an excellent example of the lexicon-based method application. In his study, tweets were collected globally across 12 countries from March 11th to March 31st, and sentiments were analyzed and discussed in different countries. After the pre-processing step, an NRC Emotion Lexicon package was implemented in his study to conduct the text mining and sentiment analysis (Dubey, 2020). Their results

showed that each emotion's percentage took place among all tweets collected in that country, and the composition was compared across different countries. Word clouds were also provided to visualize the most commented words in each country to understand diverse topics in every country studied.

Many other researchers preferred to follow machine learning-based approaches because the development of TextBlob and NLTK gave this approach a big convenience. TextBlob is an open-source Python library built upon Natural Language Toolkit (NLTK) that provides an easy-to-use interface to the NLTK library. It provides useful tools for dealing with everyday tasks in Natural Language Processing problems such as sentiment analysis, classification, language detection and translation, and others (Loria, 2018). Many researchers had used TextBlob, and its sentiment analysis performance was particularly warranted with many practices. For example, Kaur and Sharma analyzed people's sentiment from different countries upon the coronavirus using the Twitter API and NLTK library to collect and process data and use TextBlob for further analyses and classifications (Kaur & Sharma, 2020). The tweets were collected at two particular times, although they did not mention their data source's exact dates. At the end of their paper, they raised a problem indicating the sentiment results (positive, negative, and neutral) were not showing the trends because the results were composed of too many neutral sentiments (43.9%). Manguri et al. did similar research, but their tweets were selected from a seven-day interval between April 9th to April 15th (the most spread week of the pandemic by the time of their research) with keywords “#COVID-19” and “#coronavirus” (Manguri, Ramadhan, & Amin, 2020). Another research from Pokharel followed this method but focused on the target Twitter users, specifically from Nepal

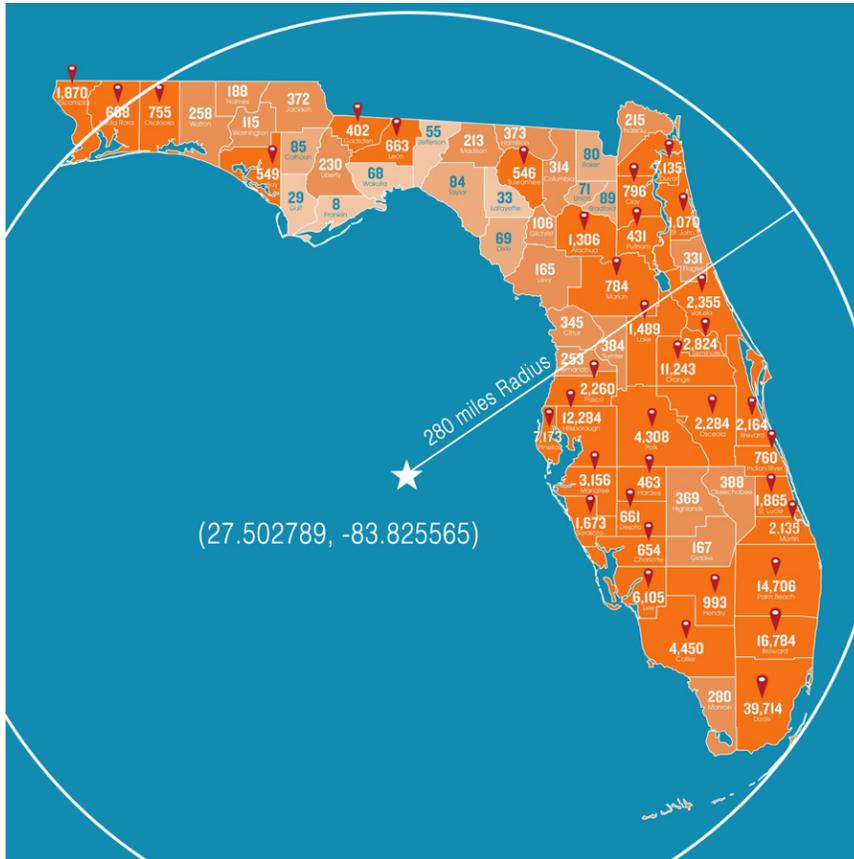
(Pokharel, 2020). His data were gathered for 11 consecutive days from May 21st to May 31st filtered by “Nepal” and with the same two hashtags used in Manguri’s research and followed with sentiment polarity and subjectivity analysis as well.

Of many kinds of research, the studies focused on only one domain of doing sentiment analysis but not providing a comparison using both learning-based and lexicon-based methods. In our research, we combine the two approaches and carry out our study using the most popular current method that relies on TextBlob and NLTK to perform learning-based analyses and use the NRC lexicon-based technique to do the same thing for comparison. Other than the methods improvements, we also realize some limitations of current studies of SA on coronavirus problems and make some enhancements on analyzing method designs. Most present research papers target a single day or several consecutive days collecting tweets, but it is hard to understand the extent to which the emotions are being positive or negative without comparing sentiment changes over a more extended period. To stand at a broader view of the problem and observe possible sentiment shifts, we extend our research to study three different time intervals and provide SA results over time.

### **Dataset**

With the Tweepy library’s help, researchers can easily extract data from the Twitter database using its API. Tweepy library offers many customizable filters for developers to sort out useful information for their researches (Roesslein, 2020). Functions can help sort keywords, languages, geographics, hashtags, date intervals, and preferences of tweets selection based on the most recent order or in a manner of popularity. The method *geocode* in Tweepy returns tweets from users located at a given latitude and

longitude within a preset radius. Our research targets Florida state in the US, so we set the *geocode* with the latitude and longitude values of (27.502789, -83.825565) at sea to roughly cover the entire state at a radius of 280 miles. Figure 3.1 below visualizes the geographical range we search with associated coordinates and radius. The keywords we have selected for filtering the search are “coronavirus,” “covid-19,” and “covid” because they are the top mentioned terms toward the discussions of COVID-19 problems based on a real-time GitHub dataset (Banda et al., 2020). Tweets have been gathered and stored in CSV format from three selected periods in 2020, where period one is April 22nd to April 28th, period two is July 15th to July 21st, and period three is October 14th to October 20th. As discussed earlier, three periods have been associated with the beginning, the burst, and the recovery periods of COVID-19 developments in Florida, respectively. The collected tweets are passed in for pre-processing to remove distracting contents, and Table 3.1 below shows the number of tweets retrieved from Twitter API with all filters passed on after the data cleansing.



	October 14	1015
	October 15	914
	October 16	803
<b>Recovery Period</b>	October 17	646
	October 18	394
	October 19	1155
	October 20	897

### **Implementation and Procedure**

This section will provide a further explanation of how we collect data and conduct sentiment analysis. A flow chart of complete procedures is shown in Figure 3.2 after we expand each stage in detail.

1. **Preparation:** Twitter developer accounts are free to apply and with which researchers can access the Twitter APIs to complete a lot of different tasks. With a developer account, we have generated API keys and tokens for the authentication process. A Tweepy library package has been implemented with the command “pip install tweepy” at the terminal to use Python to fetch tweets.
2. **Data Collecting:** to collect tweets, we have set search keywords as #covid-19, #covid, #coronavirus with geoID at (27.502789, -83.825565) with 280 mi radius, and time intervals as Apr 22-28, Jul 15-21, and Oct 14-20 respectively. The streams of tweets have been collected and stored in separate CSV files for all 21 days.
3. **Data Cleaning:** all the original tweets have been sent to a cleansing process, where the white spaces, HTML and links, punctuations, and junk terms have been removed, and the content has been converted into lowercase for more straightforward analyses. Word lists containing the top frequently appeared words have been conducted for

manually examining the possible problems. The words “positive” and “negative” have been removed for reducing the ambiguity.

4. **Sentiment Analysis:** TextBlob and NRC methods are applied to the same dataset separately. Both of them can be implemented by pip with the command “pip install textblob” and “pip install NRCLex.” TextBlob depends on NLTK 3, and it has been automatically installed as TextBlob has been installed. The dictionary Corpora also needs to be downloaded for analyzing tweets, and the command is “python -m textblob.download\_corpora” at the terminal. The previously cleaned tweets have been sent to two methods, and TextBlob automatically processes the sentiment polarity and subjectivity scores for all tweets and provides classifications of positive, negative, or neutral. On the other hand, NRC measures eight emotional affects associated with each word and counts each emotional occurrence on each day’s tweets. The classification of the sentiment of each tweet is also calculated in this method.
5. **Result Analysis:** further discussions and analyses have been performed on both methods. For TextBlob, the results are shown in terms of polarity outcomes, subjectivity scores, and perspective emotions; and for NRC, the percentage of default emotional affects and the classifications of sentiments are shown for each day to produce a clearer perception of people’s sentiment. The overtime analysis has also been conducted to observe possible sentiment and emotional changes. Then the comparison and discussion of the results from both approaches are provided.

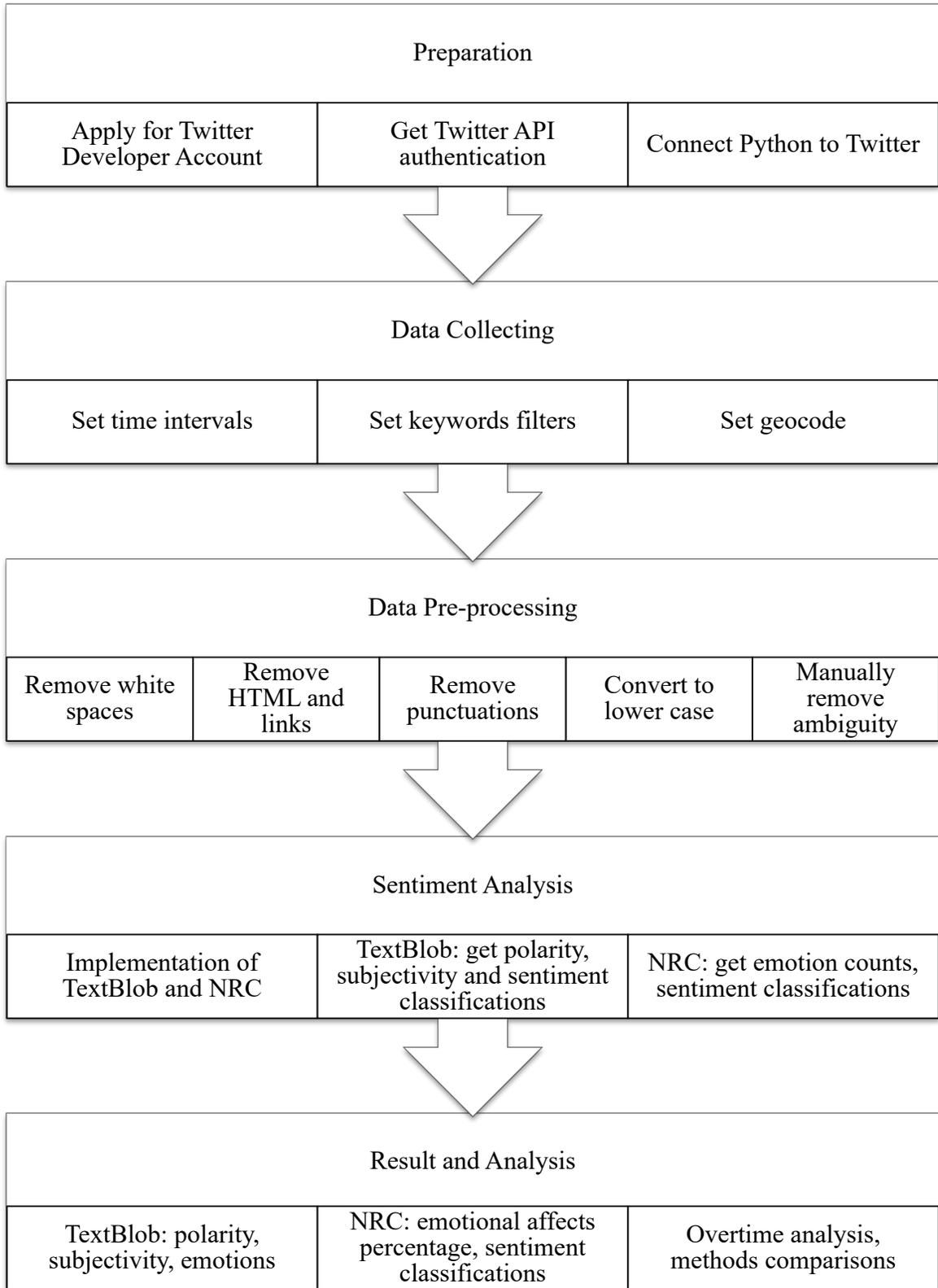


Figure 3.2. Flowchart for sentiment analysis procedures

## Understand the Algorithm Behind

TextBlob becomes a popular tool in sentiment analysis problems because of its simple access and good classification performance. The algorithm behind TextBlob's classification function is a Naïve Bayes (NB) model (Loria, 2020).

Naïve Bayes is a probabilistic algorithm that relies on the Bayes theorem and can be used to predict the sentiment classification of unknown datasets (Parveen & Pandey, 2016). The NB equation provides a method of computing posterior probability from priors using Bayesian probability terminology as shown in the equation:

$$P(c|x) = \frac{P(c)P(x|c)}{P(x)}$$

Where  $P(c)$  is the prior probability of an event,  $P(x|c)$  is the likelihood of event  $x$  given  $c$  happens, and  $P(x)$  is another prior probability showing the standalone chance that  $x$  happens.

Applying this theorem back on the sentiment classification problem, it is also not hard to understand. The Naïve Bayes classifier is created by training on a large data pool where tweets are pre-labeled with positive, negative, and neutral classes, and then is used to predict a new tweet regarding its features. Let us say tweet  $T$  has a set of *Features*  $F$ , where  $F = (f1, f2, f3, \dots, fn)$ , then the probability of the tweet  $T$  to be a *positive* tweet can be calculated as:

$$P(\text{positive}|Features) = \frac{P(\text{positive}) P(Features|\text{positive})}{P(Features)}$$

The example above calculates the probability of a new fed-in tweet to be positive with pre-estimated probabilities, and the equation returns a value ranges from 0 to 1, which is also known as a probability score representing the chance that the tweet is positive. It

works in the same way as calculating probabilities of a tweet being *negative* or *neutral* and gives the associated polarity to each tweet in the dataset.

The lexicon-based method NRC Word-Emotion Association Lexicon is easier to understand. NRC stands for National Research Council Canada, where Dr. Saif M. Mohammad works with Turney and creates this system for analyzing emotional affects in text with annotations manually done by crowdsourcing (Mohammad & Turney, 2013). The NRCLex package that Mark M. Bailey develops bases on the work done by Dr. Mohammad for easier access to sentiment analysis using Python (Bailey, 2019). The package expands the NRC lexicon to approximately 27,000 words based on the NLTK library's WordNet synonyms. It predicts sentiment and emotions of given texts by categorizing words into eight primary emotions, where anticipation, joy, surprise, and trust are four positive emotions, and anger, sadness, disgust, and fear are the other four negative emotions. Two sentiments, positive and negative, are also associated with each query, where the sentiment classification follows this rule: if the text contains more positive than negative words, it returns positive; otherwise, it is negative; however, if the text contains the same positive and negative words it returns both positive and negative, and in such cases, we associate them as neutral sentiments in this paper.

## CHAPTER 4

### RESULTS and DISCUSSION

In this section, we will show the results of sentiment analysis for both research methods. In the TextBlob method, the results of polarity, subjectivity, and emotions will be shown and discussed, respectively. For the NRC Lexicon method, we will provide emotional affects composition for each day in a similar way and follow with sentiment classification compositions. The results from both methods will be compared and contrasted to understand each method's possible advantages and weaknesses of this targeted problem. A further discussion of sentiment changes over time in different phases will also be provided by computing and comparing average results from each period.

Before we start analyzing tweets, we have added a step looking into the word lists that contain the most frequently appeared words from each week's dataset. This step is for manually examining if the words that we will study have any apparent concerns, and if there are any possible words, they should be better removed to achieve a better result. Table 4.1 below shows the ranks in word frequency, listing the top 10 words with the highest occurrence from each period.

Table 4.1  
*Most Frequently Appeared Words in the Cleansed Tweets from Different Periods*

	week 1	week 2	week 3
1	covid19	deaths	covid19
2	positive	cases	vaccine
3	trump	covid19	trump
4	new	more	positive
5	cases	positive	cases
6	more	trump	people
7	deaths	cdc	more
8	pandemic	home	biden
9	people	lockdown	restore
10	health	worse	new

We have observed a potential problem: the word “positive” frequently appears in tweets from both weeks 1 and 3. However, we believe the meaning of this word in most of the situations in our study is not associated with its common semantic that is being positive in attitude, but on the contrary, it refers to the coronavirus test as being positive. Since the appearance of the word “positive” is high, it is concerned it can potentially mislead the sentiment results because of an incorrectly assigned score. To avoid the possible influences of the word like this, we have decided to manually throw out the words “positive” and “negative” in our dataset before running the sentiment analysis.

### **TextBlob - Polarity Analysis**

Three different figures below demonstrate the sentiment polarity of tweets from three phases in April (Figure 4.1), July (Figure 4.2), and October (Figure 4.3), respectively. The tweets with keywords “coronavirus,” “covid-19”, and “covid” have been combined in one file for analysis. The polarities are shown in terms of positive, negative, and neutral and are measured in percentage.

For tweets processed on a specific date, if a tweet’s polarity is greater than 0, it would be labeled with positive sentiment. Every tweet needs to go through this process and receive classification as positive, negative, or neutral. The count of tweets in each class has been recorded. The percentages are then calculated by dividing the counts over the total number of tweets from that particular day. An average percentage score has also been given in each time slot by calculating each week’s data’s mean percentages.

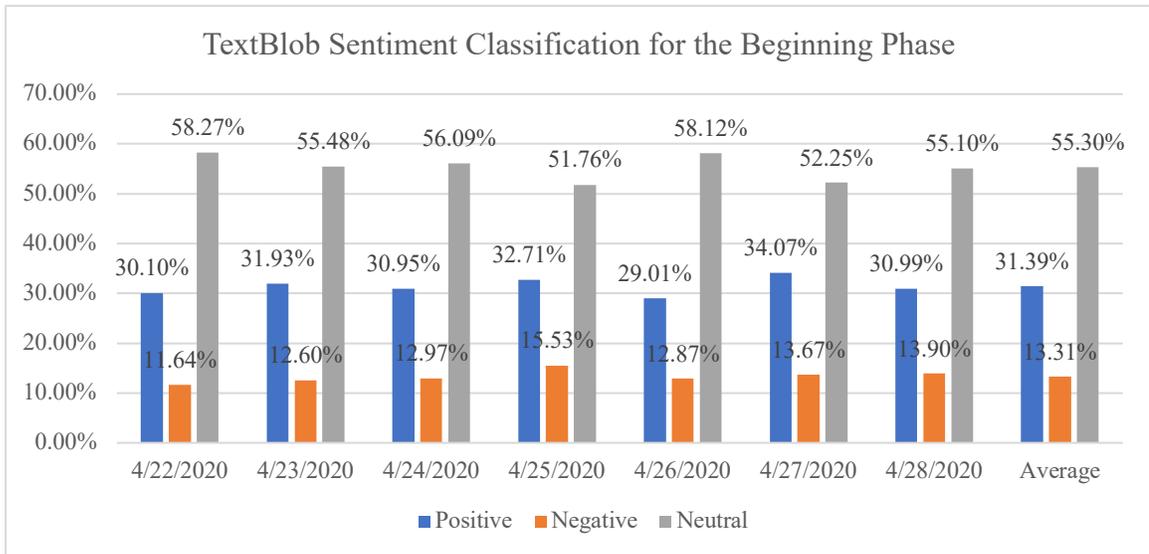


Figure 4.1. Sentiment classification on tweets for the “Beginning Phase” by the TextBlob method

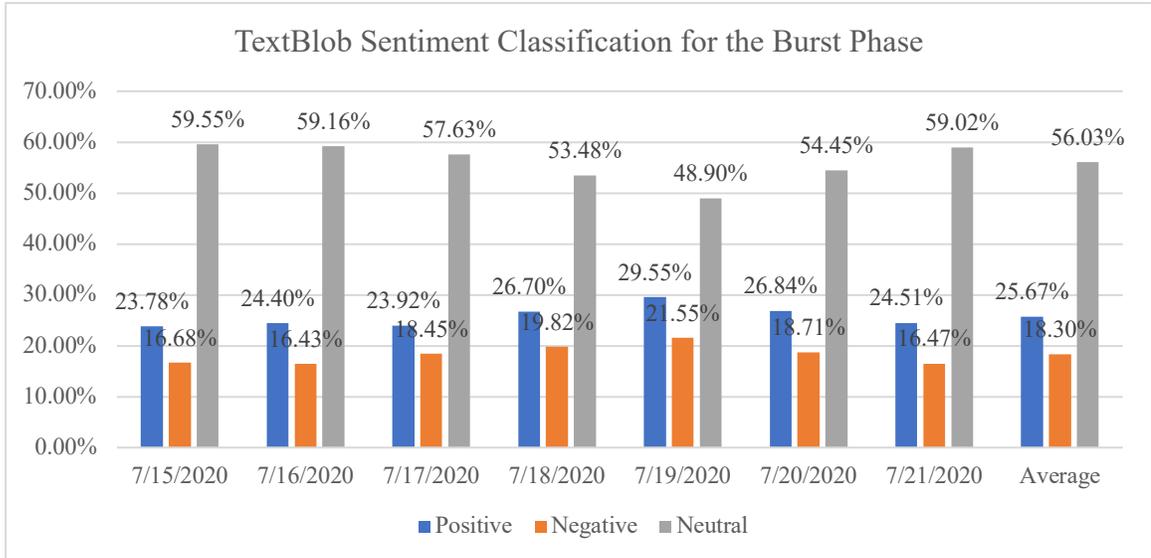


Figure 4.2. Sentiment classification on tweets for the “Burst Phase” by the TextBlob method

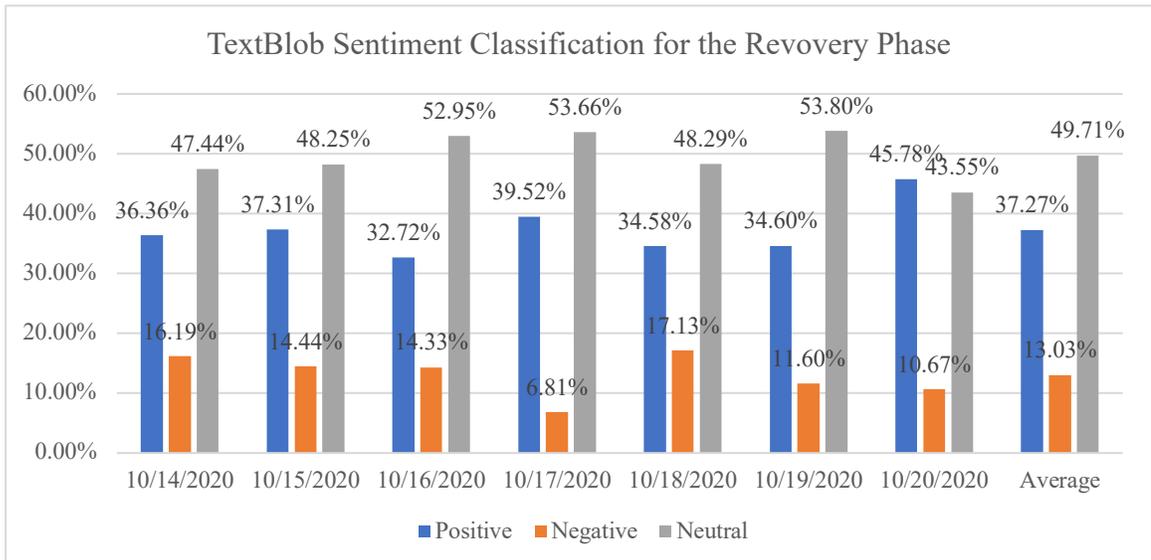


Figure 4.3. Sentiment classification on tweets for the “Revoverly Phase” by the TextBlob method

During the week from April 22nd to April 28th, around 31% of people published optimistic views, while 13% of tweets were negative, and the neutral toll was

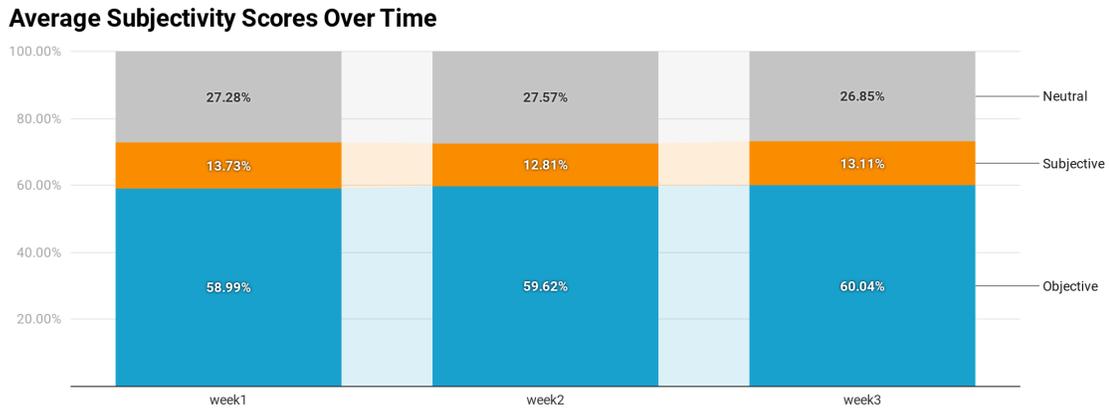
significantly high across all days at about 55%. For the second week, the average positive posts dropped to less than 26%, the negative sentiments rose by 5 percent to 18%, and the neutral comments during this week remained at the same rate, which was around 56%. For the recovery week, the tweets with negative sentiment fell back to the same level as in the first week, which was around 13%, but there was a significant rise in positive comments (more than 37%) and a drop in neutral tolls (50%) comparing to the earlier weeks.

### **TextBlob - Subjectivity Analysis**

With the subjectivity scores returned by TextBlob, the diagram of tweets subjectivity is provided as shown in Figure 4.4, and it reflects the level of the subjectivity of contents in the tweets produced by the users for the whole 21 days.

For tweets collected from a particular day, subjectivity scores have been conducted for every individual tweet and stored as a list of float numbers ranging from 0 to 1. If the number is smaller than or equal to 0.33, it would be accounted as an objective tweet; if the value is greater than or equal to 0.67, it will be subjective; otherwise, the subjectivity would be neutral (when it lies between 0.33 and 0.67).

The numbers of objective, subjective, and neutral counts have been converted into percentages, and then the average subjectivity scores are calculated and presented for each 7-day interval, as shown in Figure 4.4 below.



*Figure 4.4.* Average subjectivity results for tweets from three weeks showing the composition of objective, subjective, and neutral in percentage

Overall, the subjectivity scores for all 21 days were not showing any significant differences from each other. Calculating the average percentages on a full-time scale, many of the records were objective, and it took about 59.55%. Meanwhile, opinions that were subjectively expressed only took up 13.22%. The rest of the 27.23% tweets were not associated with precise characteristics, so they were classified as neutral in subjectivity.

Although the polarity changes from time, the subjectivity does not involve many fluctuations except for some subtle variances. The relatively stable subjectivity scores somehow reflect the consistency of our data source. The tweets collected from different dates are from different users with different contents, but our dataset’s contribution composition stays at the same level throughout the entire experiment.

### **TextBlob - Emotion Analysis**

For perceiving a clearer understanding of how people feel about this topic more than merely from the positive, negative, and neutral classifications, there is a method of connecting the polarity scores to emotional states with the use of Emotional Guidance

Scales. Manguri et al. created a chart of emotional scale regarding the sentiment analysis related to coronavirus disease (Manguri et al., 2020). The chart contains 11 emotions people commonly feel about this disaster, and the scale from -1 to 1 with 0.2 increments are assigned to each of the emotion, where -1 denotes the most depressed and fear feeling and +1 denotes the emotion of being happy and joyful. The emotions shift gradually from positive to negative, with the highest scale value to the lowest, shown in Table 4.2.

Table 4.2  
*Emotional Guidance Scale*

<b>Scale</b>	<b>Emotion</b>
1	Happy and joy
0.8	Confident
0.6	Optimistic
0.4	Hopeful
0.2	Calm and content
0	Neutral and relaxed
-0.2	Relieved
-0.4	Pessimistic and impatient
-0.6	Worry and boredom
-0.8	Discouraged and difficulty
-1	Depressed and fear

Our dataset’s polarity scores are rounded to the scale scores with one decimal place to assign the emotions to all tweets by matching two scores. For example, a tweet with polarity equals 0.436111 will be count into the emotional state as “Hopeful” because its polarity rounds to 0.4. Tweets from each 7-day interval have been combined, and the polarities are translated in terms of occurrences for each emotion. A distribution of emotions is interpreted in each stage as in Figures 4.5, 4.6, and 4.7.

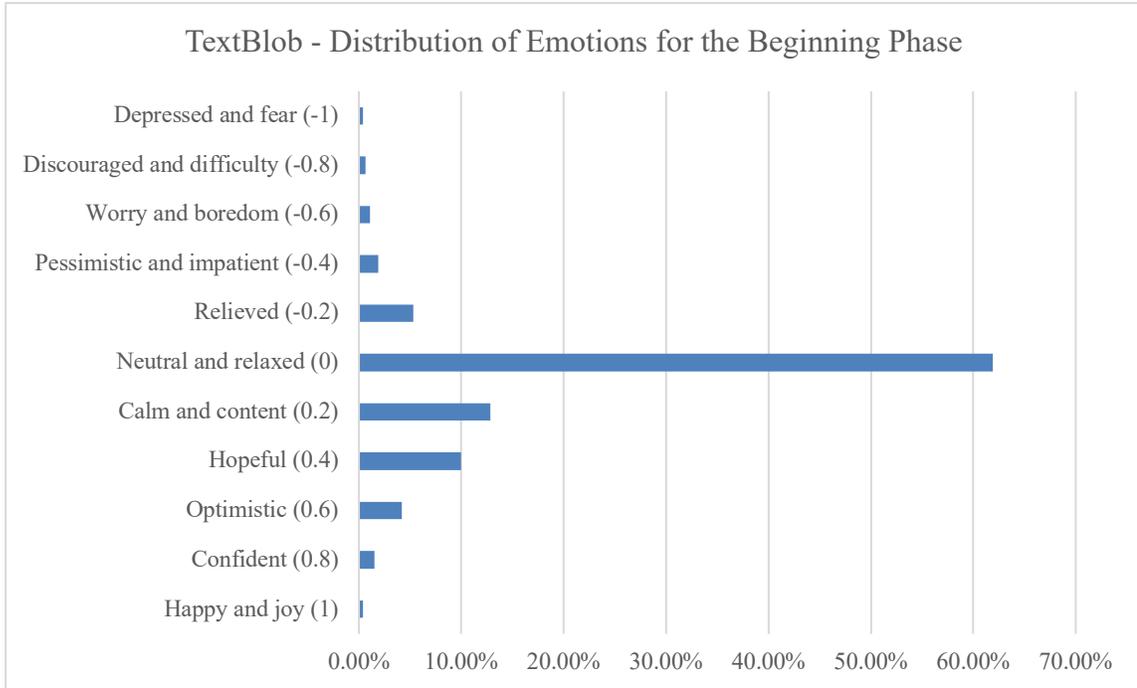


Figure 4.5. Distribution of emotions by matching the TextBlob polarities to the Emotional Guidance Scale for the “Beginning Phase”

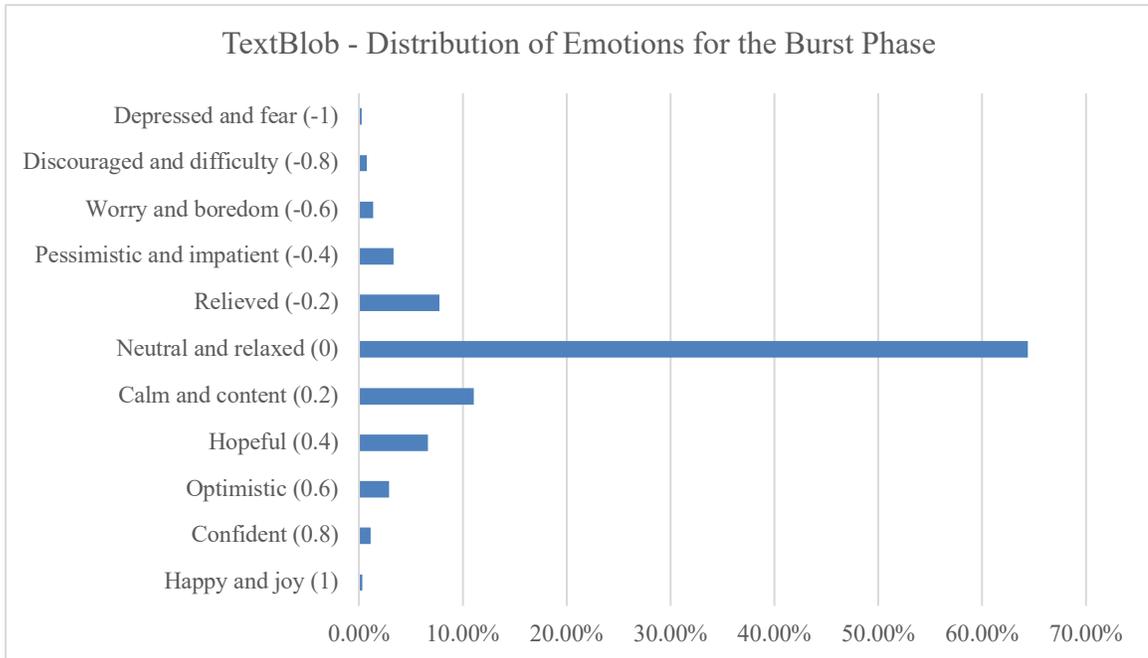


Figure 4.6. Distribution of emotions by matching the TextBlob polarities to the Emotional Guidance Scale for the “Burst Phase”

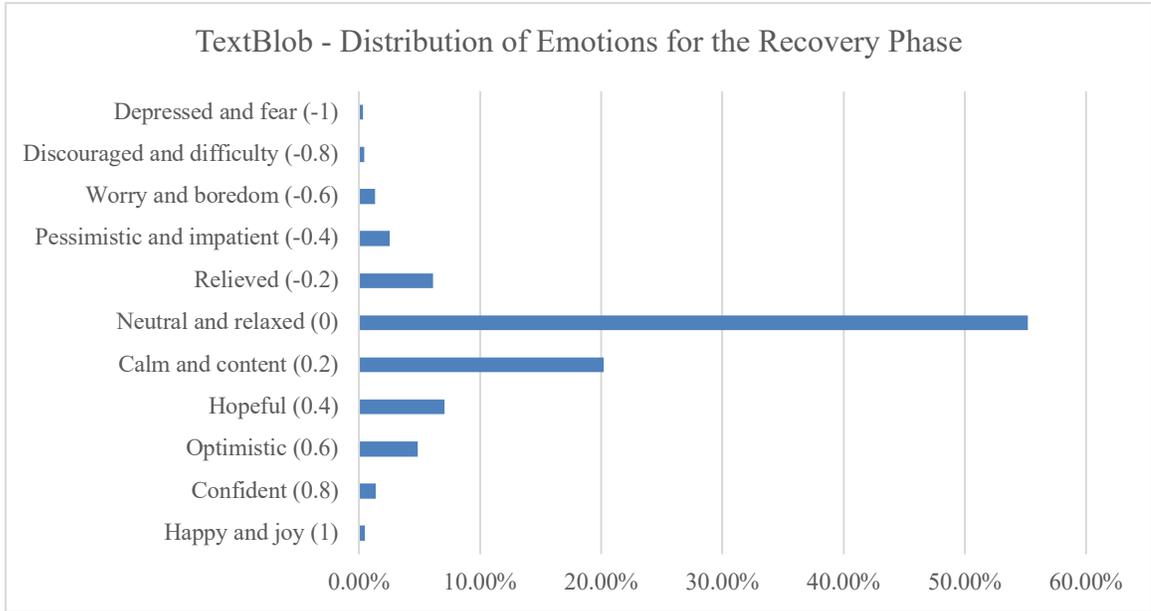


Figure 4.7. Distribution of emotions by matching the TextBlob polarities to the Emotional Guidance Scale for the “Recovery Phase”

From the results, the top counted emotional state among all weeks was “Neutral and Relaxed,” where polarity equaled to 0. Over 60% of this emotion was shown in the first two weeks, and more than 50% in the third week. The significant-high occurrence of this emotional class was consistent with a high level of neutral polarity compositions. As we can see, there was a drop in neutral tolls during the third week, and at the same time, the neutral and relaxed emotion state for that week decreased as well. Other than the neutral emotions, the second most feeling was “Calm and Content.” Around 10% of people expressed this emotion during weeks 1 and 2, and for week 3, this percentage doubled to 20%. The histogram graphs for all weeks were close to normal distribution shapes, where more emotions were with medium emotional states. There were general shifts of the distributions, where overall positive emotions were more than negatives.

## NRC - Emotional Affects

Using the same datasets, we deploy the NRC emotional lexicon method to observe the emotional affects related to this pandemic at different times.

For each tweet processed with the NRCLex package, each word in that tweet will be checked for possible emotional state. If the word does not associate with emotions in the pre-annotated library, it will be skipped. For example, a cleansed tweet in plain text as “one doctor said while precautions still need to be taken this is the first time that it seems the pandemic may come to an end,” the raw emotion scores we can get from running the NRCLex command would be like: {‘positive’: 1, ‘trust’: 1, ‘anticipation’: 1, ‘fear’: 1, ‘negative’: 1, ‘sadness’: 1}. It means considering all the words in this individual tweet: four emotions are being counted for one time, which are “Trust,” “Anticipation,” “Fear,” and “Sadness.” As there is the same number of positive and negative counts in this tweet, where trust and anticipation are associated with positive meanings and fear and sadness are for the negative meanings, the example of this tweet will be accounted for both “positive” and “negative” tags.

For a set of tweets stored in the same CSV file, we need to apply this analysis to all rows in the document. Then the numbers of occurrence of eight emotions are recorded along with the number of positive and negative counts. For example, for dataset on July 15th, we have acquired the raw emotion scores as {‘negative’: 828, ‘sadness’: 339, ‘anger’: 263, ‘anticipation’: 428, ‘joy’: 242, ‘positive’: 881, ‘surprise’: 216, ‘trust’: 512, ‘fear’: 434, ‘disgust’: 162}. By calculating the sum of all emotions from this day, we have computed each emotion’s occurrence in the form of percentages.

The preceding tables (Table 4.3-4.5) display each emotion’s percentage counts from the records on all 21 days we study. The mean percentages are computed at the bottom of each column in all three tables reflecting the average composition of each of the eight emotional affects for three weeks.

Table 4.3  
*NRC Emotion Affects Percentage for the “Beginning Phase”*

	Anticipation	Joy	Surprise	Trust	Anger	Sadness	Disgust	Fear
4/22	18.44	10.24	8.55	19.42	11.31	12.28	6.45	13.31
4/23	10.58	15.21	9.19	20.04	15.78	11.39	5.89	11.92
4/24	17.33	7.28	11.02	18.58	9.30	12.12	5.28	19.09
4/25	15.92	8.84	8.90	21.13	8.29	14.59	5.08	17.25
4/26	22.40	6.77	10.42	18.75	8.33	12.50	6.77	14.06
4/27	19.51	10.18	9.45	16.55	13.62	12.23	4.93	13.53
4/28	13.65	12.61	7.41	21.05	14.88	10.35	5.15	14.90
Average	16.83	10.16	9.28	19.36	11.65	12.21	5.65	14.87

Table 4.4  
*NRC Emotion Affects Percentage for the “Burst Phase”*

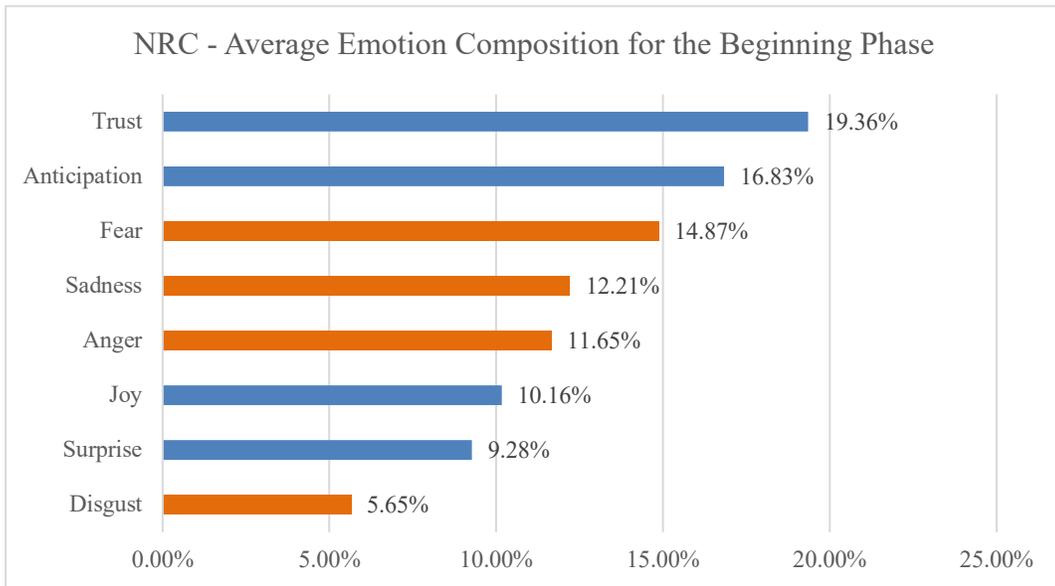
	Anticipation	Joy	Surprise	Trust	Anger	Sadness	Disgust	Fear
7/15	16.49	9.32	8.32	19.72	10.13	13.06	6.24	16.72
7/16	16.15	8.41	11.75	18.77	9.23	15.42	5.31	14.96
7/17	13.58	6.60	13.77	17.45	9.16	14.39	7.91	17.13
7/18	18.42	7.09	9.81	20.11	8.28	17.05	5.26	13.98
7/19	14.05	6.62	12.74	17.43	9.31	14.43	7.81	17.61
7/20	11.70	11.07	8.18	19.50	10.94	15.09	7.67	15.85
7/21	12.95	8.44	11.78	18.46	10.53	13.87	6.43	17.54
Average	14.76	8.22	10.91	18.78	9.65	14.76	6.66	16.26

Table 4.5  
*NRC Emotion Affects Percentage for the “Recovery Phase”*

	Anticipation	Joy	Surprise	Trust	Anger	Sadness	Disgust	Fear
10/14	12.88	19.82	5.56	26.64	7.83	9.47	4.92	12.88
10/15	11.61	17.74	6.13	28.97	5.41	8.52	5.66	15.96
10/16	12.31	15.63	6.72	22.18	9.69	13.71	4.63	15.13
10/17	18.42	11.29	10.81	21.42	8.12	6.29	7.14	16.51
10/18	12.71	20.06	5.46	26.42	7.94	9.53	5.36	12.51
10/19	13.52	18.15	5.34	26.69	6.41	10.32	7.47	12.10
10/20	15.49	14.33	7.62	29.03	5.97	8.63	5.46	13.47
Average	13.85	16.72	6.81	25.91	7.34	9.50	5.81	14.08

To see the ranks of emotions in each frame, we sort the percentage values from the largest to the smallest and represent them in Figures 4.8-4.10. Over the three weeks, “Trust” had always been the top counted emotion. It has a percentage of 19.36% for the beginning, 18.78% for the burst, and as high as 25.91% for the recovery. This result reflected that although the disease was widely spread, people still had faith in their governments and themselves in overcoming this global crisis. The rest of the emotion counts for each week differed from each other. In the beginning week, the seriousness of the disease was not yet seen and talked about by people, and that was probably the reason users showed more anticipation than fear toward this topic at that time. During the burst week, the situation changed, and the infected cases and deaths reached the maximum. Data from this time showed that more people expressed fear, sadness, surprise, and disgust emotions, and the comments showing “trust” was also the lowest among the three. When the time came to the third week when the situation got better and entered into a “recovery phase,” people’s comments also corresponded with this change. It was the first time the emotion “Joy” climbed up in the rankings that 16.72% of words were

associated with this feeling. The blue color in the figures corresponds to the positive emotions, and the orange color corresponds to the negative emotions. Different colors provide a more straightforward illustration of how an emotion climbs up or falls in the rank of percentage composition when observing the results from three weeks together.



*Figure 4.8.* Rankings of average emotional affects for the “Beginning Phase” with the NRC method

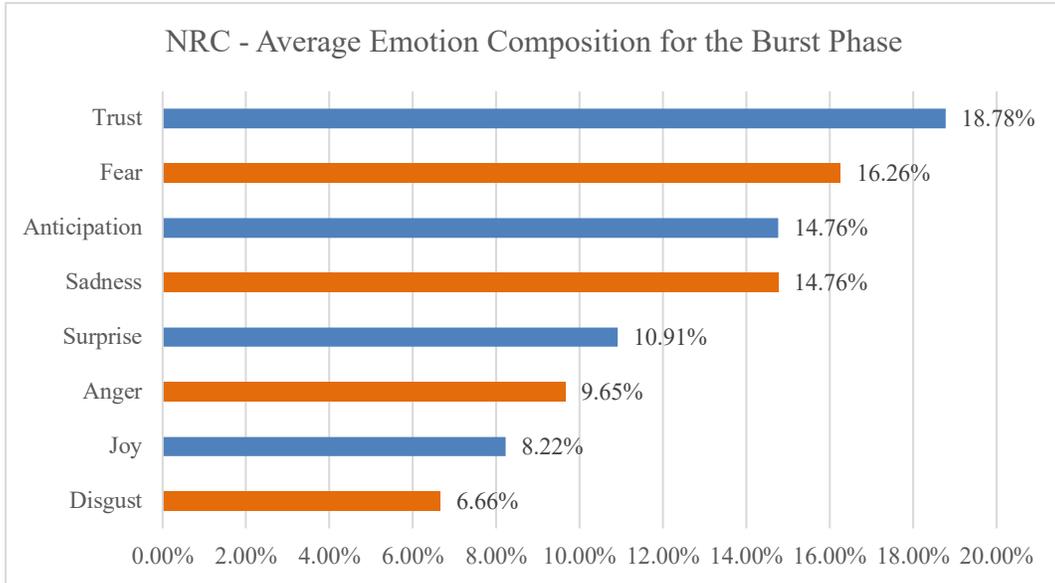


Figure 4.9. Rankings of average emotional affects for the “Burst Phase” with the NRC method

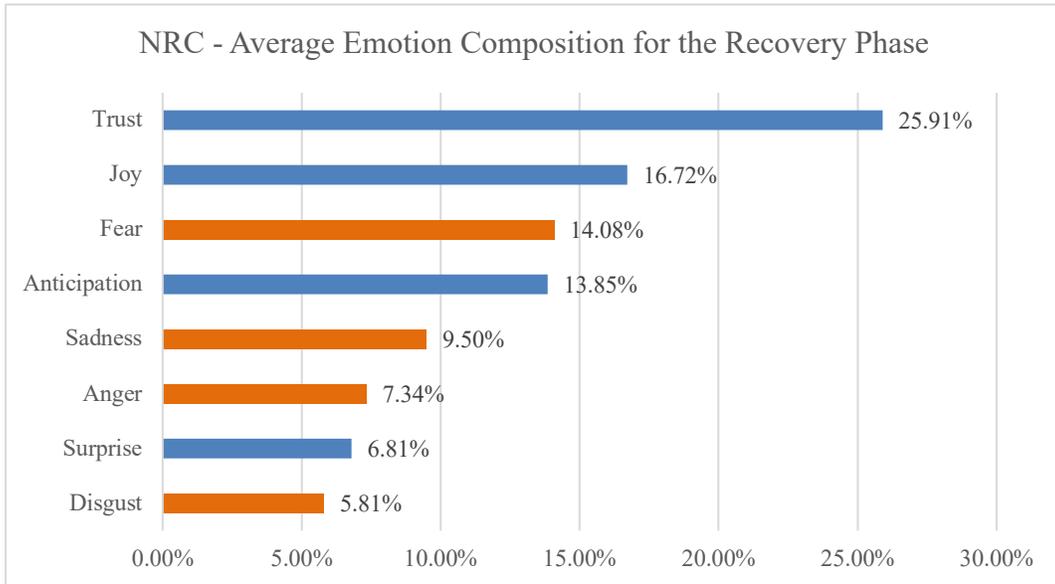


Figure 4.10. Rankings of average emotional affects for the “Recovery Phase” with the NRC method

## NRC - Sentiment Classification

Sentiments regarding positive, negative, and neutral classifications are also provided with the NRC Lexicon method.

Taking the same example of the dataset from July 15th as previous mentioned, we have obtained the raw emotion scores as {'negative': 828, 'sadness': 339, 'anger': 263, 'anticipation': 428, 'joy': 242, 'positive': 881, 'surprise': 216, 'trust': 512, 'fear': 434, 'disgust': 162}. We can observe that the positive tag is 881, and for the negative, it is 828. However, there have been many tweets being associated with more than one sentiment tag because there is the same number of positive and negative emotions associated with that specific tweet, like the example given previously, and they have been counted as neutral sentiment in this study. Calculating the number of neutral tweets for this date is the same problem as computing the union of the positive and negative sets, and it has been done by summing up the positive and negative counts then minus the data size (the total number of tweets collected on July 15th). We can then calculate how much each sentiment class takes place in terms of percentage on this particular day.

The Figures below show the percentage of word counts in each slot in a similar way as the polarity classification analysis done by TextBlob, and the average percentages are calculated and shown at the end of each figure. The sentiment composition for the beginning week was shown in Figure 4.11, overall, 32.72% of tweets were positive, and 24.00% were negative. During the second week (Figure 4.12), positive comments decreased to 27.45%, and negative increased a little to 24.93%. Coming to the last week (Figure 4.13), the positive sentiment rose back to 33.70%, and the negative tweets notably dropped to 17.93%. The neutral tweets across weeks were at a stable level of

around 46.43%, which was also about the same amount as the neutral tolls in polarity analysis by TextBlob.

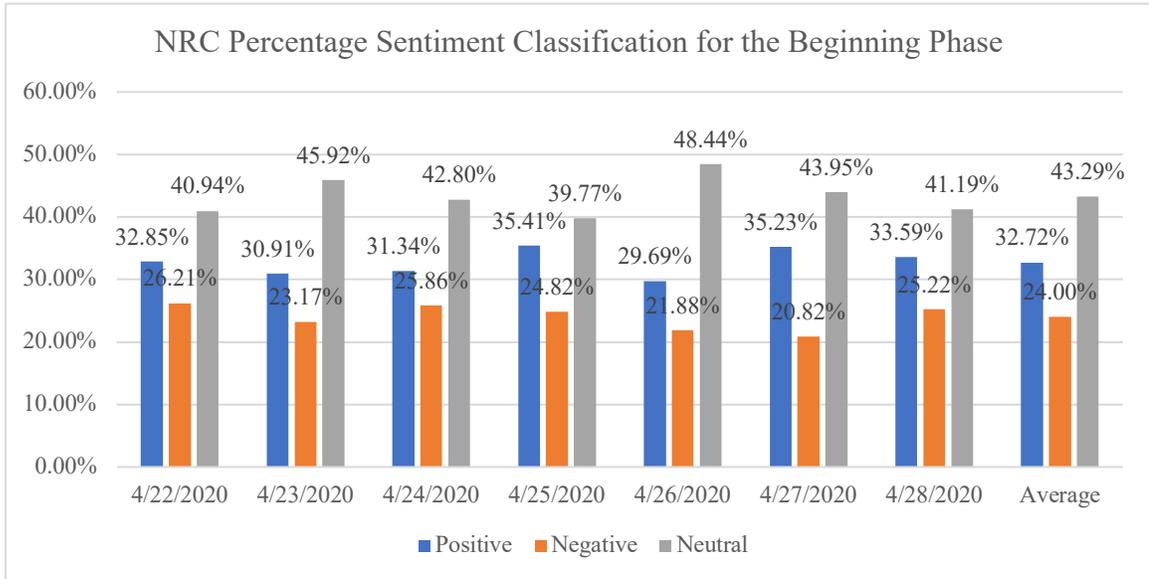


Figure 4.11. Sentiment classification on tweets for the “Beginning Phase” by the NRC method

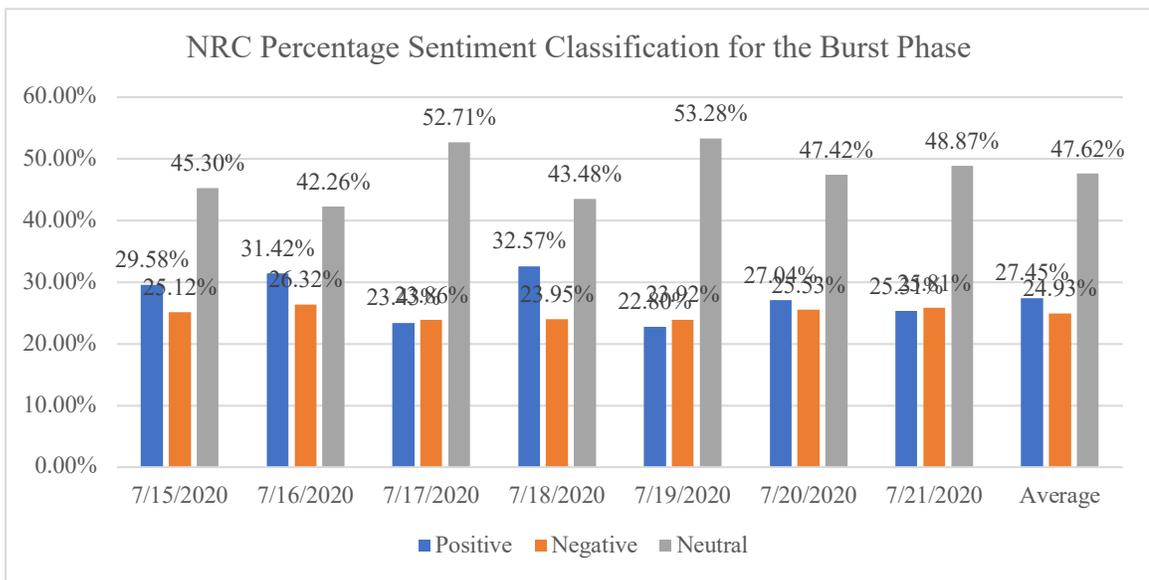


Figure 4.12. Sentiment classification on tweets for the “Burst Phase” by the NRC method

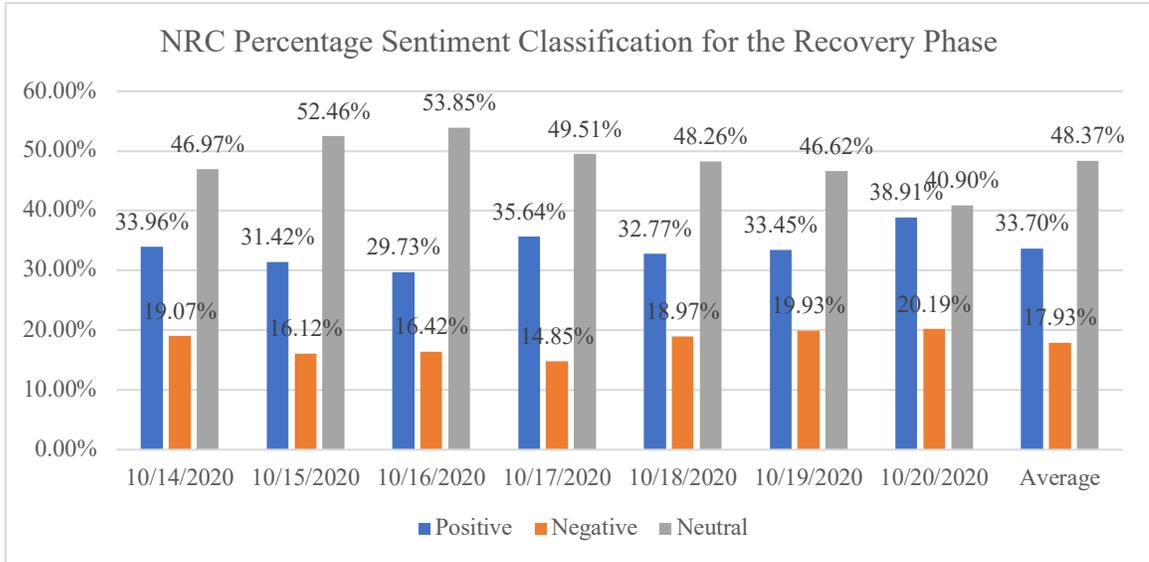


Figure 4.13. Sentiment classification on tweets for the “Recovery Phase” by the NRC method

### Discussion of Sentiment Shifts and Method Comparison

To investigate the change in emotions and sentiment of citizens from Florida state and whether there is a significant sentiment change over the long run during the coronavirus pandemic, we have conducted analyses comparing the results in three different periods.

To see the change in polarities by TextBlob, we have combined data from each 7-day interval and have produced a percent stacked bar chart comparing the amount of positive, negative, and neutral comments horizontally, as shown in the preceding Figure 4.14. Similarly, the NRC method’s figure of sentiment classification over time has also been conducted and shown in Figure 4.15.

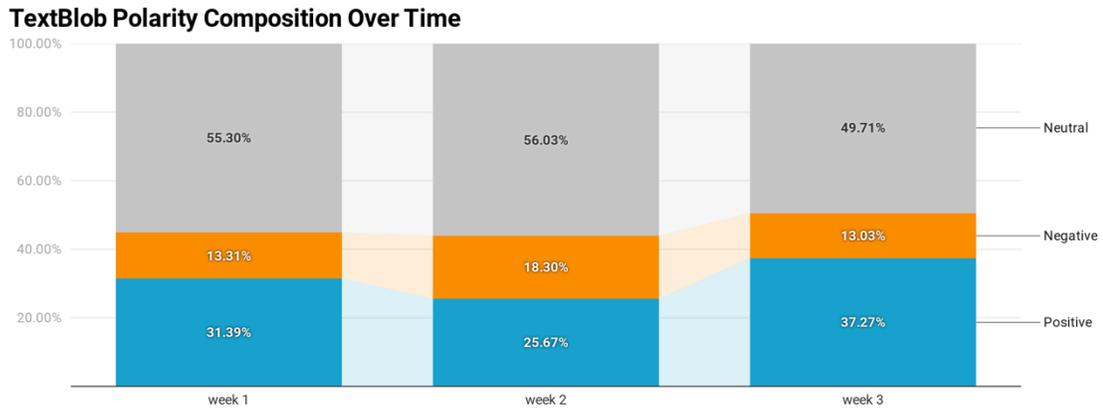


Figure 4.14. Sentiment classification of positive, negative, and neutral comparison over time with the TextBlob method based on polarity scores

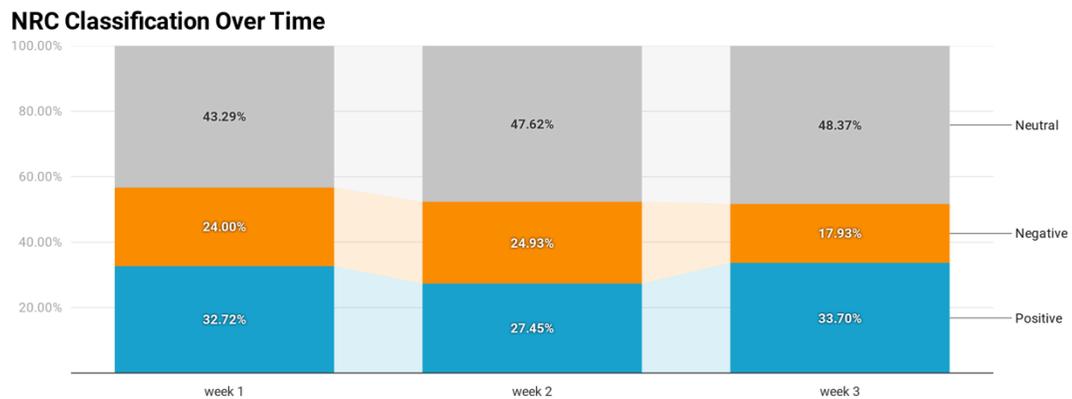


Figure 4.15. Sentiment classification of positive, negative, and neutral comparison over time with the NRC method based on the count of positive and negative labels

Both methods show similarities that the highest part of sentiment composition is neutral, and positive comes next then negative. There are slight differences in each week's configuration, but the tendencies are close. We can observe a noticeable drop of positive comments in the second week of our study and follow with a significant rise in the recovery week. The drop in positive comments during the burst week conformed with the most challenging situation during that time and the negative articles and news

published with terrifying deaths and cases during this period pushed people's panic and fear to the peak. In October, when the daily new cases and deaths reported gradually got controlled, people might feel relief from the past few months and tended to express a joyful feeling, so that the highest part of positive emotion was formed during this time.

Looking into each chart, we can also tell minor differences between the two methods. Although the neutral ratio calculated by both algorithms stated similar results, the difference between the negative and positive tweets counted with the TextBlob method appeared to be relatively more remarkable than this difference in the NRC method. It meant that regarding this specific problem domain, TextBlob classified more negative sentiments, while the NRC method tended to classify data in a more average way. A possible reason that causes this difference might be the difference in techniques behind the two approaches. With the TextBlob method, words are considered related, so the classification is based on combining all words appearing in that sentence. However, the tweets have been classified independently based on the words' definition with the NRC method. The weights of each sentence's emotion count have not been considered in the classification process with NRC, so this method's results will become more averagely distributed if the sample size increases. This observation is similar to the situation in the coin-flipping simulation: as the trial number grows, the probability of getting either side of the coin should tend to be closer to each other.

Taking the specific cleansed tweet that we discussed as an example for another time:

“one doctor said while precautions still need to be taken this is the first time that it seems the pandemic may come to an end,” we expect different sentiment results from these two methods. The results given by NRCLex for this tweet is neutral because the words

appearing in this tweet are associated with the same number of positive and negative emotions; however, with the TextBlob, the sentiment result is positive because it is returned with the polarity = 0.25 and subjectivity = 0.33.

## CHAPTER 5

### CONCLUSION

This research is about sentiment analysis on people's reaction to coronavirus outbreaks. The study is divided into three stages and based on Tweets gathered from the beginning week in April, the burst week in July, and the recovery week in October. Three keywords “#covid-19”, “#coronavirus,” and “#covid” are passed in searching with Twitter API in three stages, and the geographical region has been set to the area of Florida state in the US. The collected tweets with different keywords are combined for each day based on their timestamps and are preprocessed to remove useless contents. The lists of top-appeared words have been investigated manually for ambiguity checking, and the terms “positive” and “negative” have been removed from the dataset. The cleansed data has been sent to the analysis tool, and sentiment analyses of a total of 26541 tweets have been conducted by applying the TextBlob library and the NRC Lexicon techniques.

Although there are sentiment changes across time, the results have shown that the neutral sentiments are the majority. There are about 53.68% on average with the TextBlob and 46.43% with the NRC regarding the coronavirus problem in all weeks. Furthermore, a significantly large portion of comments is objective, with an average percentage of 59.33%. The portion of positive sentiments are larger than the negative sentiments in all weeks, but it is obvious to see a drop in positive sentiment during the Burst week in July when the disease in Florida was the most serious. From this study, we understand that people's feelings vary a lot from time to time, and the sentiment of posts

on social media reflect their emotions to some extent. This research uses two sentiment analysis methods regarding the outbreak of COVID-19 and gives us pictures of sentiments from citizens, governors, and organizations in Florida, telling how they broadcast the situations. From the results, we have observed TextBlob indicated a more significant difference between positive, negative, and neutral tolls than does the NRC Lexicon method. We cannot conclude either of the approaches produces better or more accurate results over the other because sentiment analysis of the unknown texts can hardly be associated with “correct values.” However, by practice, we have observed the differences in strengths in each approach. We would recommend TextBlob in doing the sentiment classification tasks because it has in-built functions that display the sentiment polarities. It also returns subjectivity scores that reflect the level of subjectivity in the specific text that the NRC method cannot do. However, the TextBlob does not have a better way to associate more complex emotions to each sentence than directly matching the polarity scores to some emotional guidance chart. When needing more complicated emotions, we recommend using the NCR Lexicon method; however, the NRC method recognizes the words equally, so there must be uncertainties when the sentences have multiple meanings and when the sentence is narrated to express opposite meanings of the annotated words.

## CHAPTER 6

### LIMITATION and FUTURE WORK

Although we have applied some improvements in our research approach comparing to other current works, there are still limitations in many aspects left for future studies.

- Due to time and computing capacity limitation, our study only focuses on Florida state as our target range, but there are more states and regions in the US and worldwide worth an investigation regarding this worldwide pandemic.
- There are more than internet users worldwide, but our sentiment analysis is entirely based on online comments, which would probably produce bias and uncertainty. It might be a good idea to carry out sentiment analysis in a hybrid way, taking part in traditional tolls and online-based text mining.
- Another problem with the online study is the diversity of the sentiments. Since not every user on Twitter wants to post an idea regarding a specific topic, approximately 80% of the contents are produced by only 20% of the most active users. Although we filtered out retweets, the dataset's unique users might still not enough to support the result solidly representing sentiment from Florida people. To reduce the uncertainty and enhance the diversity, researchers could focus on improving the data cleaning process and reduce comments from the same users to improve the contributions from unique users. Another way to this goal might be doing the sentiment analysis differently. Instead of illustrating a big picture of people's sentiment, research can be

based only on specific top active users by tracking their sentiment changes and understanding possible trends upon a time.

- We have discussed Microblogging's advantages and taking Twitter as our data source, but there are other reliable sources of information in different forms from other platforms. It is quite essential to study more data sources to get a more accurate sentiment illustration.
- We have selected English as our searching language during the data collection step, but users use more languages to express their feelings, especially in a multi-cultural environment like the US. To study sentiments based on different languages may also help understand the target problem's race and demographic composition.
- We have noticed a large portion of neutral tweets in the collected tweets with both methods, and it might be caused by junk and robot spam. The data preprocessing steps help filter out some meaningless contents but cannot figure out Ads, junks, and repeated spams that do not usually involve any ideas or thoughts. Some literature focuses on detecting and removing Twitter Bots, and adding an extra step to reduce the bots could reduce the portion of neutrals caused by this problem in the future.
- In the study of emotions with the TextBlob method, it is shown that the top emotion is “Neutral and Relaxed” in all-time intervals because it is associated with polarities from -0.1 to 0.1. However, the translation of the emotion scale might not be accurate in every way because 0 in polarity might only mean the sentiment in a specific comment is neither positive nor negative but does not mean the one who posts it feels calm and relax. It could mislead our understanding of people’s emotions, and it is better to develop another emotion scheme in the future.

- For the NRC study, the analysis method involves some uncertainties as well.

Although we manually remove the words “positive” and “negative” considering their opposite meanings in our study, there must be more situations that the words used in tweets produce ambiguity that we cannot detect. It is stated that irony and sarcasm are two complex emotional states that are difficult to define in formal terms, and they potentially become a challenge in analyzing the semantics. Although we include eight emotional states, it is still not enough to classify those emotions precisely.

- Our method is designed to study different phases during this pandemic to understand the possible changes of emotions over time regarding the coronavirus. However, coronavirus is not a single event, people’s feelings about it are not based merely on the development of the disease itself. In further studies, it will be more exciting and challenging if we correlate the changes in sentiment among people and the changes in public policies or events over time.

## REFERENCES

- Alessia, D., Ferri, F., Grifoni, P., & Guzzo, T. (2015). Approaches, tools and applications for sentiment analysis implementation. *International Journal of Computer Applications*, 125(3).
- Bailey, M. M. (2019). NRCLex. Retrieved from <https://pypi.org/project/NRCLex/#description>
- Bajema, K. L., Oster, A. M., McGovern, O. L., Lindstrom, S., Stenger, M. R., Anderson, T. C., . . . Chu, V. T. (2020). Persons evaluated for 2019 novel coronavirus—United States, January 2020. *Morbidity and Mortality Weekly Report*, 69(6), 166.
- Banda, J., Tekumalla, R., Wang, G., Yu, J., Liu, T., Ding, Y., & Chowell, G. (2020). *A Twitter Dataset of 100+ million tweets related to COVID-19*.
- Beetz, C., Skrahina, V., Förster, T. M., Gaber, H., Paul, J. J., Curado, F., . . . Weckesser, V. (2020). Rapid Large-Scale COVID-19 Testing during Shortages. *Diagnostics*, 10(7), 464.
- Bishop, G. F. (2004). *The illusion of public opinion: Fact and artifact in American public opinion polls*: Rowman & Littlefield Publishers.
- Centers for Disease Control and Prevention. (2020). *COVID-19 Response. COVID-19 Case Surveillance Public Data Access*.
- Chan, J. F.-W., Yuan, S., Kok, K.-H., To, K. K.-W., Chu, H., Yang, J., . . . Poon, R. W.-S. (2020). A familial cluster of pneumonia associated with the 2019 novel

- coronavirus indicating person-to-person transmission: a study of a family cluster. *The Lancet*, 395(10223), 514-523.
- Chew, C., & Eysenbach, G. (2010). Pandemics in the age of Twitter: content analysis of Tweets during the 2009 H1N1 outbreak. *PloS one*, 5(11), e14118.
- Dubey, A. D. (2020). Twitter Sentiment Analysis during COVID19 Outbreak. *Available at SSRN 3572023*.
- Fang, X., & Zhan, J. (2015). Sentiment analysis using product review data. *Journal of Big Data*, 2(1), 5.
- Fung, I. C.-H., Duke, C. H., Finch, K. C., Snook, K. R., Tseng, P.-L., Hernandez, A. C., . . . Tse, Z. T. H. (2016). Ebola virus disease and social media: a systematic review. *American journal of infection control*, 44(12), 1660-1671.
- Gardner, L., & Dong, E. (2020). COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University. Retrieved from <https://github.com/CSSEGISandData/COVID-19>
- Guellil, I., & Boukhalfa, K. (2015). *Social big data mining: A survey focused on opinion mining and sentiments analysis*. Paper presented at the 2015 12th international symposium on programming and systems (ISPS).
- Hasan, A., Moin, S., Karim, A., & Shamshirband, S. (2018). Machine learning-based sentiment analysis for twitter accounts. *Mathematical and Computational Applications*, 23(1), 11.
- Holshue, M. L., DeBolt, C., Lindquist, S., Lofy, K. H., Wiesman, J., Bruce, H., . . . Tural, A. (2020). First case of 2019 novel coronavirus in the United States. *New England Journal of Medicine*.

- Iglesias, J. A., Tiemblo, A., Ledezma, A., & Sanchis, A. (2016). Web news mining in an evolving framework. *Information Fusion*, 28, 90-98.
- Jahangir, M. A., Muheem, A., & Rizvi, M. F. (2020). Coronavirus (COVID-19): history, current knowledge and pipeline medications. *International Journal of Pharmaceutics & Pharmacology*, 4(1).
- John Hopkins University Coronavirus Research Center. (2020). Cumulative Cases. Retrieved from <https://coronavirus.jhu.edu/data/cumulative-cases>
- Johnson, A. F., Pollock, W., & Rauhaus, B. (2020). Mass casualty event scenarios and political shifts: 2020 election outcomes and the U.S. COVID-19 pandemic. *Administrative Theory & Praxis*, 42(2), 249-264.  
doi:10.1080/10841806.2020.1752978
- Kaur, C., & Sharma, A. (2020). *Twitter Sentiment Analysis on Coronavirus using Textblob* (2516-2314). Retrieved from
- Kouloumpis, E., Wilson, T., & Moore, J. (2011). *Twitter sentiment analysis: The good the bad and the omg!* Paper presented at the Fifth International AAAI conference on weblogs and social media.
- Loria, S. (2018). textblob Documentation. *Release 0.15*, 2.
- Loria, S. (2020). TextBlob: Simplified Text Processing. Retrieved from <https://textblob.readthedocs.io/en/dev/>
- Manguri, K. H., Ramadhan, R. N., & Amin, P. R. M. (2020). Twitter Sentiment Analysis on Worldwide COVID-19 Outbreaks. *Kurdistan Journal of Applied Research*, 54-65.

- Mohammad, S. M., & Turney, P. D. (2013). Crowdsourcing a word–emotion association lexicon. *Computational Intelligence*, 29(3), 436-465.
- Ogbuju, E., Oladipo, F., Yemi-Petters, V., Abdumalik, R., Olowolafe, T., & Aliyu, A. (2020). Sentiment Analysis of the Nigerian Nationwide Lockdown Due to COVID19 Outbreak. *Available at SSRN 3665975*.
- Oliveira, D. J. S., Bermejo, P. H. d. S., & dos Santos, P. A. (2017). Can social media reveal the preferences of voters? A comparison between sentiment analysis and traditional opinion polls. *Journal of Information Technology & Politics*, 14(1), 34-45.
- Organization, W. H. (2020a). Coronavirus disease 2019 (COVID-19): situation report, 72.
- Organization, W. H. (2020b). Coronavirus disease 2019 (COVID-19): situation report, 82.
- Organization, W. H. (2020c). Novel coronavirus—china. 2020. URL: <https://www.who.int/csr/don/12-january-2020-novel-coronavirus-china/en>.
- Organization, W. H. (2020d). Pneumonia of unknown cause—China. 2020. *Available at: who.int/csr/don/05-january-2020-pneumonia-of-unkown-cause-china/en/*. Accessed April, 1.
- Parveen, H., & Pandey, S. (2016). *Sentiment analysis on Twitter Data-set using Naive Bayes algorithm*. Paper presented at the 2016 2nd international conference on applied and theoretical computing and communication technology (ICATCCT).

- Phan, L. T., Nguyen, T. V., Luong, Q. C., Nguyen, T. V., Nguyen, H. T., Le, H. Q., . . .  
Pham, Q. D. (2020). Importation and human-to-human transmission of a novel coronavirus in Vietnam. *New England Journal of Medicine*, 382(9), 872-874.
- Pokharel, B. P. (2020). Twitter Sentiment Analysis During Covid-19 Outbreak in Nepal. *Available at SSRN 3624719*.
- Pran, M. S. A., Bhuiyan, M. M. R., Hossain, S. A., & Abujar, S. (2020). Analysis Of Bangladeshi People's Emotion During Covid-19 In Social Media Using Deep Learning. *2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, 1-6.
- Ravindran, S. K., & Garg, V. (2015). *Mastering social media mining with R*: Packt Publishing Ltd.
- Roesslein, J. (2020). Tweepy Documentation. Retrieved from <http://docs.tweepy.org/en/latest/index.html>
- Russell, M. A. (2013). *Mining the social web: data mining Facebook, Twitter, LinkedIn, Google+, GitHub, and more*: "O'Reilly Media, Inc."
- Sailunaz, K., & Alhaji, R. (2019). Emotion and sentiment analysis from Twitter text. *Journal of Computational Science*, 36, 101003.
- Savidge, M. (2017). Florida: The Swingiest Swing-State. In: CNN.
- Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-based methods for sentiment analysis. *Computational linguistics*, 37(2), 267-307.
- Tang, J., Chang, Y., & Liu, H. (2014). Mining social media with social theories: a survey. *ACM Sigkdd Explorations Newsletter*, 15(2), 20-29.

Wang, H., Can, D., Kazemzadeh, A., Bar, F., & Narayanan, S. (2012). *A system for real-time twitter sentiment analysis of 2012 us presidential election cycle*. Paper presented at the Proceedings of the ACL 2012 system demonstrations.

World Health Organization. (2020). *Coronavirus disease (COVID-19) Weekly Epidemiological Update*. Retrieved from [https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200831-weekly-epi-update-3.pdf?sfvrsn=d7032a2a\\_4](https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200831-weekly-epi-update-3.pdf?sfvrsn=d7032a2a_4)

Zarghami, M. (2020). Psychiatric aspects of Coronavirus (2019-nCoV) infection. *Iran J Psychiatry Behav Sci*, 14(1), e102957.