

Social Media Sentiment Analysis of Covid-19: Reopening the Nations

Dr. Suboh Alkushyani

*Minnesota State University, Mankato,
Minnesota, 56001*

Dr. Du'a Alzaleq

*Minnesota State University, Mankato,
Minnesota, 56001*

Rishabh Yata

*Minnesota State University, Mankato,
Minnesota, 56001*

Shaharia Chowdhury

*Minnesota State University, Mankato,
Minnesota, 56001*

Abstract

The seriousness and severity of COVID-19 have prompted an increase in extraordinary sentiments and caused emotional as well as mental health challenges to people all around the world. Individuals have been utilizing Twitter to announce their sentiments and offer insights on a few issues identified with the COVID-19 during the lockdown and now after the gradual reopening of countries around the world. This paper examines the feelings of people specifically in most populated countries and countries with lower populations on the topic of reopening. Twitter data was used for investigation and studying the Tweets to find the nostalgic point of view, passionate viewpoint, and setting off words towards the reopening. This paper also intrigued to examine public sentiments after the countries slowly reopened their economies. The significant finding here is that when most countries turned to lockdown, individuals demonstrated a prevailing feeling of dread, however as reopening began individuals, even in worst-hit counties have expressed less dread. While this might be valid, because of this reopening stage and even with most caution and safety, every day positive cases are rising contrasted with the lockdown circumstance. Generally, individuals have a more positive assessment of the circumstance of resuming. Furthermore, we used four machine learning models and also one deep learning model to analyze their adequacy in classifying and predicting the number of infected and number of deaths since the reopening in specific nations like the United States, India, China, Puerto Rico, Denmark, and Finland. Strong classification accuracy is seen with Logistic Regression, Naïve Bayes, and decision tree classifiers. Additionally, it's been observed that Keras deep learning model classification method provides a weaker accuracy. This study gives bits of knowledge into Coronavirus dread assumption movement and provides related suggestions and limitations.

Keywords: Sentiment Analysis, Twitter, Covid-19, Pandemic, Corona Virus, Reopen, Textual Analytics, feelings. Twitter, Covid-19, Pandemic.

1. INTRODUCTION

Coronavirus, another strain of COVID-19, started in Wuhan, China in December 2019. On the eleventh of March 2020, the World Health Organization declared the COVID-19 episode as a pandemic. This exceptionally irresistible COVID-19 has crossed limits at a speed anyone might have envisioned flipping around the world. COVID19 has just guaranteed more than 0.3 million lives, with around 6 million individuals contaminated more than 200 nations or regions up to May 31, 2020. Ever since, It has shaken the worldwide financial and social structure so gravely, that continuing typical life is a thing to imagine. As a measure to control the circumstance, various nations have depended on complete lockdown. At the point when the entire world is COVID-19 stricken, the lockdown has been the main answer for getting a grip on the situation. From that point forward a large part of the discussion has been going on how the nations should facilitate the lockdown program. As lockdown is not a definitive arrangement, facilitating the limitations is the main approach to limit monetary misfortune and keep the public practical.

As we have established that feelings and estimations are the main impetus that individuals are partaking in online media. If we dissect the web-based media posts, at that point we can discover the experiences of those posts alongside feelings and conclusions. From this sort of examination, we may discover what individuals like, what they need, and what are the principal worries of them as explained in these studies [1, 2]. Twitter has an enormous number of every day dynamic clients all around the globe where individuals share their musings and data for any subject of late worry thereby. It is an accessible chronicle of human ideas. The wealth of freely accessible information shared through Twitter has urged analysts to decide the conclusions on nearly everything, including slants towards any item, administration. Many Analysts and experts are as of now giving valuable data on issues identified with this COVID-19 flare-up. We have mostly added to the issue of the returning stage giving some knowledge into individuals' responses. Coordination of reactions from on the web also,

from this present reality makes certain to uncover promising outcomes to address the current emergency. Our major finding is that during this return stage dread among individuals is less predominant contrasted with the past times when the states settled on lockdown choices. Another finding is that as the return stage begins, the quantity of new cases is likewise expanding. Individuals are imparting their musings and insights to battle this circumstance.

As the pandemic is still in full effect across the world and not showing any signs of slowing down, we have decided to explore the topic of sentiment analysis and discover where people stand based on the effects of reopening the countries with Covid-19 still in effect. Although there have been some similar studies done on this topic [1], we have tried to address this in a different view which is by comparing people's feelings on life after reopening in lower populated countries compared with the higher populated countries. We think researching this topic can explore today's challenges in governance, people's mental health, economy, financial status, and many other indirect effects of the pandemic. While it is true that reopening the countries have seen a spike in new cases, we plan to analyze how people's emotions differ among populated countries to formulate a solution based on the findings. proceedings are the records of the conference.

Contributions:

For this study, we used datasets from Kaggle and other sources like social media platforms to interpret people's sentiments about reopening their countries and how they feel and cope with the new normal. We believe this idea is new as there are very few research studies about sentiment analysis that studied the feelings of people about the reopening after a long lockdown period. We also want to go one step further and explore the sentiments of people today with the new public rules and restrictions in place. Therefore, the main questions we address in this study include people's feelings and reactions in populated countries compared to less populated on reopening, people's sentiments on covid-19 rules, and look at how these rules affect people's day to day life in general.

This paper is arranged and assembled as follows. Section 2 presents a writing audit or a literature review and a few contemplations of the past work. Section 3 contains all our methods for data seeking, data collection, cleaning of our dataset, sentiment analysis, emotional analysis, and machine learning regression and classification tasks. Section 4 explains the outcomes and discoveries of our study. Section 5 talks about the limitations and challenges of our study. Section 6 is the conclusion of this study. Lastly, Section 7 concludes the paper with future works.

2. RELATED WORK

Most of the related research studies and articles until now have done sentiment analysis on Twitter or related social media to find the true emotions of people during the quarantine and the lockdown phases. Some studies were limited by data from just one specific country, while some were only limited to 4 or 5 most effected and developed countries. We have also reviewed a study that is

closely related to our objective, but it had limitations of analyzing data from just one specific country and extremely limited data which the authors specified may have affected their research results and conclusion. Therefore, we built on the approach of those studies and increase the scale of the study significantly and implement a different method to get more exact results based on the Automatic sentiment analysis that will use Supervised machine learning classification algorithms. We also used unsupervised machine learning algorithms to explore data.

This research [2] distinguishes public supposition in midst of dangerous financial outcomes of the lockdown and investigates the following four potential public notion related situations. Since the pandemic has caused extreme feelings, and emotional and mental healthcare challenges, this research focused on elements like emotional consequences and find dominant public sentiment trends for effective decisions and policies. This study [2] investigated public notions utilizing Twitter Data, time-adjusted to the COVID-19 resuming banter, to recognize predominant feeling patterns related to the push to return to the economy. This research uses textual analytics methodologies to analyze public sentiment support for two potential divergent scenarios - an early opening and a delayed opening, and the consequences of each. Based on textual data analytics, including textual data visualization and statistical validation, the study concluded that tweet tweeted by American Twitter users shows more sure assumption uphold than negative, for reopening the US economy. This exploration [2] builds up a novel feeling extremity based public slant situations (PSS) system, which will stay valuable for future emergencies examination, well past COVID-19. With extra approval, this examination stream could introduce significant time-delicate open doors for state governments, the government, enterprises, and cultural pioneers to direct neighborhood and local networks, and the country into an effective new ordinary future.

This study [3] centers around examining the basic financial factor relationship for feelings about post-Covid-19 resuming. This examination endeavors to investigate the components related to positive and negative opinions of individuals about reopening the economy, in the United States. It considers the situational vulnerabilities like changes in work and travel design because of lockdown strategies, monetary decline, and related injury, and passionate factors, for example, misery. Twitter information was gathered, speaking to the 51 states including Washington DC of the US. The outcomes from the logit model exhibit that family families, individuals with low instruction levels, individuals in the workforce, low-pay individuals, and individuals with higher house lease are keener on resuming the economy. In contrast, families with a high number of individuals and high income are less intrigued to resume the economy. This investigation [3] gives a reasonable sign to the policymakers were to assign assets and what strategic alternatives they can attempt to improve the financial circumstances of individuals and moderate the effects of pandemics in the current circumstance and as well as in the future. This study's significant methods and approaches include gathering of state-wide financial qualities of individuals like schooling, pay, family size, and work status as well as built environment data like population density, and the quantity of COVID-19 related cases and coordinated with

Twitter information to play out the investigation. A binary logit model was utilized to distinguish the components that impact individuals toward a positive or negative estimation.

This pandemic has incredibly influenced general wellbeing and prosperity. Because of the pandemic danger of the COVID-19 pandemic, a few nations, including China, received lockdown and isolate arrangements, which may cause mental misery. This investigation [4] planned to investigate the mental effect of region-wide lockdown and individual isolate during the COVID-19 flare-up in China just as the relating hazard factors and defensive elements. The study analyzed the prompt (2-week) and postponed (2-month) effect of area-wide lockdown and individual isolation on mental pain in a public example of 1390 Chinese inhabitants. No quick effect of territory-wide lockdown on mental pain was noticed, though close to home isolate expanded people's nervousness, dread, and outrage. Notwithstanding the absence of beginning affiliation, mental pain expanded among those in region-wide lockdown. Self-shame and individual control both essentially directed the relationship between lockdown and mental pain, yet in various headings. Those with higher self-disgrace and lower individual control were more affected by the lockdown. The government uphold directed the effect of isolation on mental pain, yet not that of lockdown. The postponed impacts of lockdown and isolation on mental misery were noticed, and self-shame, social help, and saw control moderate the connections. This examination [4] is argued by the authors as the first to exhibit the mental expenses of region-wide lockdowns on people's emotional wellness, giving proof of the need for relief methodologies and ideal public psychological well-being readiness in nations with late flare-ups of COVID-19.

It is not just Covid-19 which is destroying lives. There is another emergency that has shown itself as mass dread and panic, energized by deficient and frequently inaccurate data as explained in this study [5]. There is subsequently a huge need to address and better comprehend COVID-19's enlightening emergency and measure public assessment so that fitting informing, and strategy choices can be executed. In this exploration article [5], work was done to distinguish public notions related to the pandemic utilizing Coronavirus explicit Tweets and R factual programming, alongside its notion examination bundles. The study shows experiences into the advancement of dread opinion after some time as COVID-19 moved toward top levels in the United States, utilizing unmistakable printed investigation upheld by fundamental literary information perceptions. Moreover, the study shows a methodological diagram of two fundamental AI/ML arrangement techniques, with regards to textual analysis, and analyze their adequacy in grouping Coronavirus Tweets of differing lengths. Results showed a solid grouping accuracy of 91% for short Tweets, with the Naïve Bayes classifier. And the logistic regression classifier gave an accuracy of 74% with shorter Tweets, and both demonstrated generally more fragile execution for longer Tweets.

There is another case study [6] that tried to characterize public emotions and sentiments in the country of India. As of May 3, 2020, the Ministry of Health and Family Welfare affirmed an aggregate of 39,980 positive COVID-19 cases and 1,301 deaths in India. To lower the rate, India has been securing its nation

from March 24 to May 17, 2020. This examination gathered "Coronavirus in India" related tweets adding up to 410,643 tweets in English from March 22 to April 21, 2020, to check the questions and settings related to public estimations during the lockdown. This work adds to the developing group of studies on COVID-19 online media mining by separating feelings and notions after some time, which might shed a light on the settings of articulations during the pandemic.

3. METHODOLOGY

There are two parts to this section. When it comes to social media analysis, we usually tend to do the Twitter analysis using the Twitter developer account. Sentimental analysis on Twitter was done using the word cloud which was only intended to focus on the word corona. The output came with corona word as the focus. After collecting the data, we used Jupiter notebook with python to do the pre-processing of data by eliminating the unwanted rows and columns and later ran the algorithms like Logistic regression, Naïve Bayes, Keras, and Decision tree.

3.1 Data collection

We have used multiple sources to collect our data. We have used a commonly utilized Tweets data collection, cleaning, and data preprocessing process. We downloaded Tweets using a Twitter API, the rTweet package in R, which was utilized to accumulate tweets from May to August of 2020, applying the watchword "Corona." The raw information was handled and arranged for tweet analysis using the R programming language and related R packages. The information was subset to target Tweets labeled by the nation as this study is specifically looking at countries like the United States, India, China, Puerto Rico, Denmark, and Finland. We could do this since Twitter tweet information contains two unique data types based on location for each tweet. One is an area for the tweet, demonstrating where the Tweet was posted from, and the other is the user's usual location, and user's location when the Twitter account was made. For the COVID-19 infection Tweets, we analyzed fear as well as negative feelings and discovered some irrational bits of knowledge, indicating moderately lower levels of fear in places which were essentially influenced by a high number of COVID-19 cases. Various R packages were utilized in the cleaning cycle, to make a clean dataset for additional investigation. Since the plan was to utilize the information for scholarly examination, we supplanted all recognizable offensive words with a novel alphanumeric label word, which contained the content "abuvsv", yet was blended in with numbers to try not to utilize a bunch of characters that might have preexisted in the Tweets. As we understand erasing offensive words would deny the information of potential examination openings, and thus an explicitly coded calculation was used to make a tweaked substitution. The dataset was additionally assessed to recognize the most helpful factors, clear and unessential qualities were erased to make a cleaned dataset. In addition to this, we have also collected a dataset

word 'Joy' showed that people show a positive attitude towards the situation. As we move on towards the beginning of May people still talked about the word 'Fear' through their posts, but this kept fading away over time and people had less fear as they started adopting how to deal with the situation.

3.6 Machine Learning Models

The focus of this analysis is on showing how generally utilized Machine learning techniques can be applied and used to contribute to classification and performing predictive analysis to find the rate of confirmed cases and death rates in the countries like United States, India, China, Puerto Rico, Denmark, and Finland since the reopening. This study does not develop any contributions to new ML theory or algorithms. In contrast to regression, which is utilized for assessing the likelihood of quantitative boundaries, classification can be successfully used for assessing the likelihood of subjective boundaries for twofold or multi-class factors which is when the prediction variable of interest is binary, absolute, or ordinal in nature. There are numerous classification techniques or classifiers for subjective information; among the most popular are Naïve Bayes, decision tree, random forest, logistic regression, and Keras deep learning model. The estimations are set from the training dataset. If either the indicators and additional reactions are not continuous quantitative factors, at that point the structure of this model is unfit and needs changes. The variables X and Y become intermediary factors and their significance relies upon the settings they are used; with regards to the current study, X means the features and Y is the class to be assessed, for which the model is being trained.

3.6.1 Application of the models: Process Workflow

The development and application of the models were done in the order shown in Figure 1.

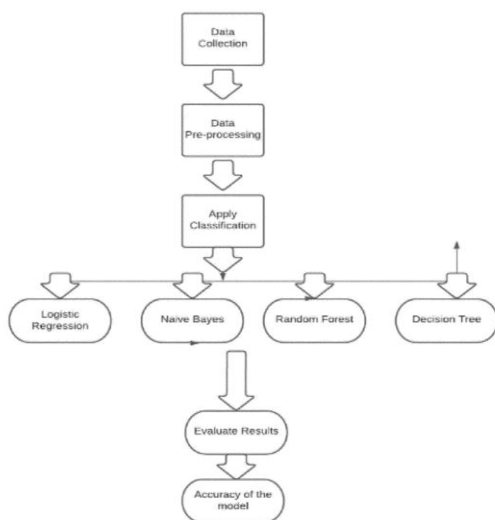


Figure 2. The process workflow here shows the total process from collecting data to the accuracy of the model to validate the proposed topic.

The code and source materials can be found on our GitHub link². We have also included a readme file for a walkthrough of our code and files.

3.6.2 Application of Naïve Bayes for COVID-19 Dataset

Naive Bayes is a classification method dependent on Bayes' Theorem with an assumption of independence among indicators. It is a set of supervised learning algorithms that assume conditional independence between every pair of features given the value of the class variable. In straightforward terms, a Naïve Bayes classifier expects that the presence of a specific element in a class is disconnected from the presence of some other element. The decoupling of the class conditional feature distribution means that each distribution can be independently estimated as one-dimensional distribution. As Naïve Bayes is known to be helpful for exceptionally large data collections, we have chosen it to predict total deaths and total new cases from the available training dataset. Results show that since the reopening we notice a reasonably strong classification accuracy of 76% with Naïve Bayes compared to others. The algorithm is not affected when a new data point is introduced on categorical and numerical features with missing values or not, it works well.

3.6.3 Application of Logistic Regression for COVID-19 Dataset

It is the appropriate regression analysis to conduct when the dependent variable is binary. It is mainly a predictive analysis that is used to describe data and to explain the relationship between the dependent variable and independent variable. It is also a probabilistic grouping technique that can be utilized for supervised machine learning. We utilize logistic regression and sigmoid capacity to build a binary classifier. We used logistic regression to again predict total deaths and total new cases from the available training dataset. Results show that since the reopening we notice a stronger accuracy of 80% with logistic regression compared to others.

3.6.4 Application of Decision Tree for COVID-19 Dataset

A decision tree uses a tree-like model which has branches, one represents the outcome of the test while the other leaf node represents a class label. The construction of a decision tree is about finding attributes that return the highest information gain. Each branch is representing a decision option that is showing all the possible alternatives, probabilities, and outcomes that might be benefited from taking this decision analysis. There were missing values in our dataset, but we have could have handled in two ways in decision tree one by either classifying missing values as a separate category that can analyze or building up the model that sets the variables with lots of missing values and later replace the missing one with predicted values.

² https://github.com/rishabhyata/CIS631_Project

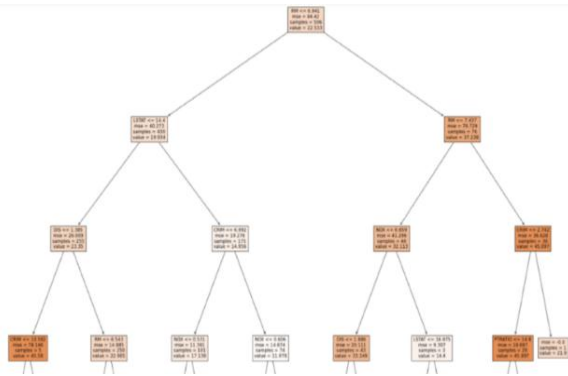


Figure 3. This is the decision tree chart that each node has a decision rule that mainly splits the data. Gini referred to as the Gini ratio, measures the impurity of the node. On the other hand, it can be said the node or leaf node is pure when all the records belong to the same class.

In Figure 3, we have a root node which is called the decision node, and more mutually subsets that are internal nodes. Then the variables and target variables were dividing the original input variable into significant subgroups. The goal of this decision tree is to identify the combinations of factors between the populated and lowest populated country on reopening the economy in the middle of an ongoing pandemic. A decision tree is a powerful tool that helps to predict, interpret the data that has several potential applications in Covid-19 research. Application of Random Forest for COVID-19 Dataset. Random forests help to leverage the ability of multiple varied analysis, organization strategies, and ensemble learning to supply correct models, preceptive variable importance ranking, record by record basis for deep data understanding. Its strengths are recognizing outliers and anomalies in knowledge data, predicting future outcomes, discovering data patterns, and exchange missing values with imputations. Random forest algorithms are normally not biased as multiple trees are mainly trained on a subset of data making the algorithm very stable. The overall algorithm is not affected when a new data point is introduced as it may only affect one tree but not all the trees and both on categorical and numerical features with missing values or not, it works well as Figure 3 shows.

3.6.5 Application of Random Forest for COVID-19 Dataset

Random forest is a supervised learning algorithm that helps to leverage the ability of multiple varied analysis, organization strategies, and ensemble learning to supply correct models, preceptive variable importance ranking, record by record basis for deep data understanding. Its strengths are recognizing outliers and anomalies in knowledge data, predicting future outcomes, discovering data patterns, and exchange missing values with imputations. These random forest algorithms are normally not biased as multiple trees are mainly trained on a subset of data making the algorithm very stable. The overall algorithm is not affected when a new data point is introduced as it may only affect one tree but not all the trees and both on categorical and numerical features with missing values or

not, it works well. One big advantage of doing this in our project as it can be used for both classification and regression problems.

3.6.6 Application of Keras for COVID-19 Dataset

Lastly, we have experimented with Keras which is the neural network's library. It is an incredible and simple to-utilize free open-source Python library for creating and assessing profound learning models. It wraps the productive mathematical calculation libraries Theano and TensorFlow and permits you to characterize and prepare neural organization models in only a couple of lines of code. We have used this on a binary classification problem where the target is either 1 or not 0. As most of the data in our dataset are numerical, it is easier to use directly with neural networks that anticipate numerical values both input as well as output. We used Keras to predict deaths and recoveries. The results from Keras showed an accuracy score of 44% which is understandable as our dataset was not large enough for the model. Keras and other neural networks generally require large data to process.

3.6.7 Plots and Visual representations

We have created several different visual representations to back up our claims as shown and explained in Figures 4, 5, 6.

Case Timeline: May 1st - December 7th

Countries: United States, India, China, Puerto Rico, Denmark, and Finland

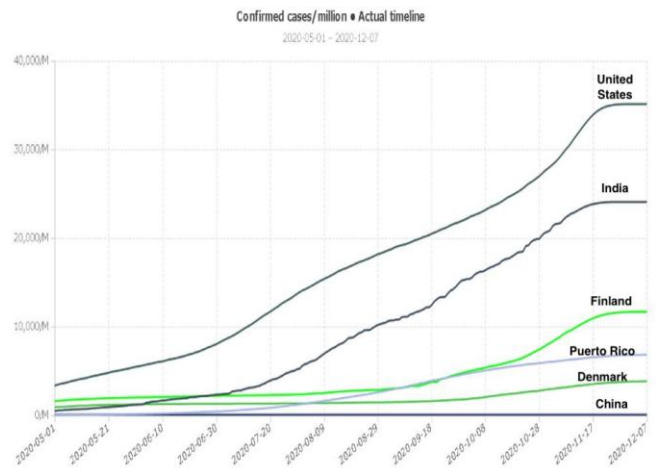


Figure 4. Confirmed cases since reopen and prediction for new cases until December 2020

Case Timeline: May 1st - December 7th

Countries: United States, India, China, Puerto Rico, Denmark, and Finland.

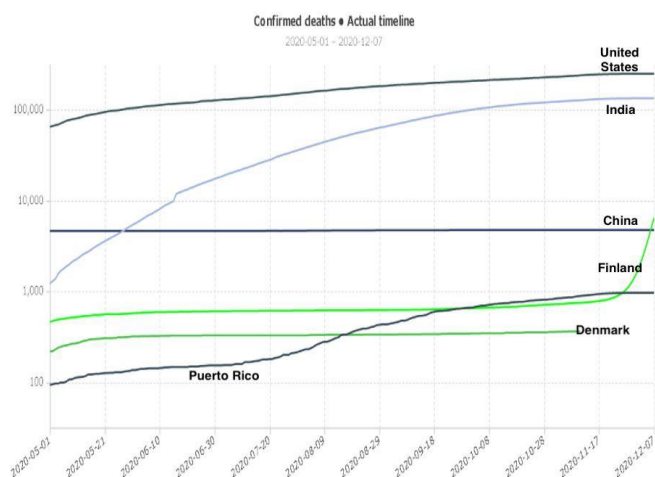


Figure 5. Confirmed deaths in the world since reopen and prediction for more deaths until December 2020

Case Timeline: May 1st - December 7th

Countries: United States, India, China, Puerto Rico, Denmark, and Finland.

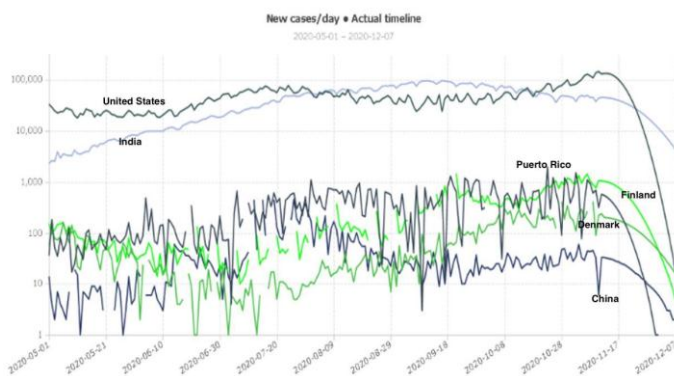


Figure 6: New cases/day in the world since reopen and prediction for new cases until December 2020

We discuss more these plots and other results in Section 4.

4. RESULTS AND DISCUSSION

The grouping results in this study are fascinating and shows a requirement for extra approval and observational model advancement with more Coronavirus information and extra strategies. Models created with extra information strategies while utilizing Naïve Bayes and Tweet Classification techniques. This would then be used as autonomous components for mechanized grouping of Coronavirus sentiment. The model and the discoveries can likewise be additionally reached out to compare worldwide pandemic bits of knowledge. Textual analytics has increased huge

consideration in recent years with the appearance of enormous information examination, unstructured information investigation, and expanded computational abilities at diminishing costs, which empowers the analysis of huge datasets. Our significant finding is that when all states turned to lockdown in March, individuals demonstrated a prevailing feeling of dread, however as reopening began individuals around the world, even in worst-hit counties have expressed less dread. While this might be valid, because of this reopening stage and even with most caution and safety, every day positive cases are rising contrasted with the lockdown circumstance. Generally, individuals have a more positive assessment of the circumstance of reopening. Furthermore, we have implemented four machine learning models (Naïve Bayes, Logistic Regression, Decision Tree, Random Forest) and one deep learning (Keras) classification model and analyze their adequacy in classifying and predicting the number of infected and number of deaths since the reopening in specific nations like United States, India, China, Puerto Rico, Denmark, and Finland. We noticed a strong classification accuracy with logistic regression, Naïve Bayes, and decision tree classifiers. Additionally, we also observe that Keras deep learning model classification method provides a weaker, but reasonable accuracy.

5. LIMITATIONS

One major limitation of this study was the primary datasets that we used was not rich enough to run on many algorithms and cannot do large scale analysis. In addition to that more measures and contexts are needed to be associated with the topic and these can be improved in our further studies. The current investigation has zeroed in on a textual collection comprising of Tweets filtered by "Corona" as the catchphrase. Accordingly, the analysis and the strategies are explicitly applied to information about a specific pandemic as an emergency circumstance, and subsequently, it very well may be contended that the structure sketched out in this paper is generalized. On the other hand, using Twitter data has some limitations too as it does not represent the complete section of the population and very few people use Twitter based on which part of the world they are from. The ever-changing data categories can impact human convictions and decision making, henceforth it is imperative to consider various web-based media stages with varying data arrangements to increase a comprehensive viewpoint. Also, as our paper topic is broad to compare the population in different countries and their thoughts for reopening, the data we collected from different sources is not enough to get the perfect accuracy sentiments of people around the world regarding the to reopen of the economy. Lastly, these restrictions are recognized from an overall viewpoint, they do not lessen the commitments made by this study, as the general shortcomings are not related to the essential purpose of this study.

6. CONCLUSION

People have spent most of their time on social media in this pandemic due to lockdown and normal social interactions went

online from meeting in person. This change in behavior of turning towards social media for advice and information is very helpful while studying the sentiments of the people towards the lockdown, corona virus and other factors too. This study explored the sentiments in people towards the re-opening and their opinion on social media platforms specially twitter which became a trending research in recent years. The dataset contains various words combinations taken from twitter. The study uses the best knowledge to use the datasets taken from authenticated source and applied algorithms to understand the sentiments and emotions towards the reopening. Naïve bayes, Random forest, Decision tree and keras deep learning models were used to find the accuracy rate of the models. The results and analysis taken from the twitter data showed that people in populated countries reacted differently than the lower populated countries. People in populated countries are taking it positive to reopen as the cases of Covid were still in rise despite the close of the economy, therefore they believe to have the reopening of the economy to have their financial state on the better side than worse. On the other hand, the least populated countries have more control on their supply of food, financial help from government and made them think re-opening would worsen the situation more. People's sentiments on Covid-19 rules were tough and effected their everyday life as almost most countries all around the world had to face shutdown of the economy and a complete lockdown for months which effected severely their mental state and financial state as they have high monthly housing rents to pay etc. Re-opening worldwide economy would cause a rise in the mortality, people will take extra caution and more people will follow government policies and rules.

7. FUTURE SCOPE

Further research can be conducted by pointing to specific reasons of why more populated countries tries to reopen the economy than less populated. Some of the factors comes in are due to lack of healthcare system, high rent, food crisis due to high population etc. And some of these populated countries are not using social media platforms to express their sentiments. They sometimes do not have any expression platforms thus this can make our result and analysis not most perfect. Some of the things that can be done to improve our work is to have more datasets from all social medias that each country uses and to extract the results of those emotions with our deep learning models. On the other hand, sentiments and human behavior can be affected by a wide range of factors like information propagation formats so in future more research could be done on the relevant time matching with the articles and responses to the tweets data for analyzing sentiments. This study opens a stream of research in identifying factors that contributes public sentiments towards the government decisions on re-open, public mental health and so on. Using the twitter data economy analysis can be analyzed as further research and the positive and negative sentiments insights on Covid 19 for effective communication and policy making for the nations. Also, we will use different sentiment analytics from twitter data, different classification algorithms to improve the accuracy

rate. In future we plan to continue this research and work to get richer dataset in future.

REFERENCES

- [1] Ahmed, R. (2020). COVID-19: Social Media Sentiment Analysis on Reopening. Ding, W., and Marchionini, G. 1997. *A Study on Video Browsing Strategies*. Technical Report. The University of Maryland at College Park.
- [2] J. Samuel *et al.*, "Feeling Positive About Reopening? New Normal Scenarios From COVID-19 US Reopen Sentiment Analytics," in *IEEE Access*, vol. 8, pp. 142173-142190, 2020, DOI: 10.1109/ACCESS.2020.3013933
- [3] Rahman, A. (2020). Twitter and Census Data Analytics to Explore Socioeconomic Factors for Post-COVID-19 Reopening Sentiment.
- [4] Gan, M. (2020). Immediate and delayed psychological effects of province-wide lockdown and personal quarantine during the COVID-19 outbreak in China. *Psychological Medicine*, 1–12. <https://doi.org/10.1017/S0033291720003116>
- [5] Samuel, J., Ali, G. G. M. N., Rahman, M. M., Esawi, E., & Samuel, Y. (2020). COVID-19 Public Sentiment Insights and Machine Learning for Tweets Classification.
- [6] Das, S., & Dutta, A. (2020). Characterizing public emotions and sentiments in COVID-19 environment: A case study of India.
- [7] Jimenez-Sotomayor, M. R., Gomez-Moreno, C., & Soto-Perez-de-Celis, E. (2020). Coronavirus, Ageism, and Twitter: An Evaluation of Tweets about Older Adults and COVID-19
- [8] "Evaluating the initial impact of COVID-19 containment measures on economic activity," OECD, Tech. Rep., 2020, accessed: 2020-05-15. [Online]. Available: <http://www.cica.net/oecd-evaluating-the-initial-impact-of-covid-19-containment-measures-on-economic-activity>
- [9] S. E. Saad and J. Yang, "Twitter sentiment analysis based on ordinal regression," *IEEE Access*, vol. 7, pp. 163 677–163685, 2019.
- [10] L. Li, Q. Zhang, X. Wang, J. Zhang, T. Wang, T. Gao, W. Duan, K. K. Tsoi, and F. Wang, "Characterizing the propagation of situational information in social media during COVID-19 epidemic: A case study on Weibo," *IEEE Transactions on Computational Social Systems*, vol. 7, no. 2, pp. 556–562, 2020
- [11] Alkhashayni, Suboh M. "mHealth technology: Towards a new persuasive mobile application for caregivers that addresses motivation and usability." (2016).
- [12] Alkhashayni, Suboh, and Susan McRoy. "mHealth technology: towards a new mobile application for caregivers of the elderly living with multiple chronic conditions (ELMCC)." *Proceedings of the 6th International Conference on Digital Health Conference*. 2016.

- [13] Kruse, Ryan, and Suboh Alkhushayni. "Identifying regional COVID-19 presence early with time series analysis." *IOP SciNotes* 1.2 (2020): 024003.
- [14] Alkhushayni, Suboh M., Daniel C. Zellmer, and Ryan J. DeBusk. "Text emotion mining on Twitter." *IOP SciNotes* 1.3 (2020): 035001.
- [15] Alkhushayni, Suboh M., M. Alzaleq Du'a, and Nadine L. Gadjou Kengne. "Blockchain Technology applied to Electronic Health Records." *Proceedings of 32nd International Conference on. Vol. 63. 2019.*
- [16] Al-zaleq, Du. "Optical Fiber Communication with Vortex Modes." (2019).
- [17] Kruse, Ryan, Tharindu Lokukatagoda, and Suboh Alkhushayni. "A framework for association rule learning with social media networks." *IOP SciNotes* (2021).
- [18] Alkhushayni, Suboh, Tejaswi Vuyyur, and Kayode O. Ayorind. "Analyzing the Trend and Forecasting of Covid-19 Outbreak Using Machine Learning Techniques." *International Journal of Applied Engineering Research* 16.3 (2021): 181-191.
- [19] Alkhushayni, Suboh Mohammad Shehadah. *Intrusion Detection by Port Scan Using Snort*. Diss. Universiti Utara Malaysia, 2008.
- [20] Alkhushayni, Suboh M., Taeyoung Choi, and M. Alzaleq Du'a, "Data analysis using representation theory and clustering algorithms" *International Journal of Engineering & Technology* 9(2227-524X):887-899. (2020)
- [21] Alkhushayni, Suboh M., M. Alzaleq Du'a, and Austin FitzGerald "Health tracker: data acquisition and analysis for monitoring health trends and assessing disease risk" *International Journal of Engineering & Technology*, 10 (1) (2021) 72-80