Sentiment Analysis of Covid-19 Tweets using Logistic Regression Count Vectorizer with GSO

Anjali Yadav[1], Prof. Chetan Agrawal[2], Prof. Pawan Meena[3]

[1]Research Scholar, [2,3]Assistant Professor

Department of CSE, RITS, Bhopal, India

Abstract: Sentiment analysis (SA) is the process of carefully going over content that has been stored online to identify and categorize the viewpoints expressed in a text. This process is to assess the author's point of view about a certain object, movie, topic, etc. There might be positive, negative, or neutral results. Examples of several SA techniques for identifying and evaluating emotions connected to the topic's polarity—positive, negative, or neutral—were given in this research. social media platforms SA might be a useful source of knowledge and data. Social media allows for fast electronic material exchange since it is built on the Internet. people may openly express their thoughts on social media sites like Twitter, and those opinions can then be shared with other people. During the most recent COVID-19 pandemic, public opinion analytics proved useful in determining the best course of action for public health. The propagation of misleading information, aided by social media and other digital platforms, has been shown by the COVID-19 pandemic to be a greater danger to global public health than the virus itself. By examining well-written tweets, one may determine the general public's thoughts on social distance. One method for automatically locating and classifying subjective data in text data is sentiment analysis. Regression using Logistic Regression The tweets are classified using Coder Vectorization with Grid Search Optimization (LRCV-GSO) and pseudo labelling. The keywords are correctly extracted from the tweets by it. The findings demonstrate that when combined with COVID tweets, the LRCV-GSO classifier outperformed other models by 79%. In light of their respective viewpoints on the pandemic, the proposed technique may assist medical experts and decision-makers in selecting the best course of action for each location. The constructed models can assist healthcare professionals and governments in making the best judgements in the event of future pandemic outbreaks.

Key Terms: Covid-19, Tweets, Accuracy, LRCV-GSO, Sentiment Analysis.

IJSMRT-25100102

I. INTRODUCTION

With two thirds of internet users accessing social networks, the COVID-19 epidemic has increased social media use. The virus-caused reduction in newspaper supply has resulted in a notable change in social media use. With users from a wide range of socioeconomic backgrounds and interests, Twitter is an online microblogging platform that has grown to be a major source of opinions. Researchers interested in studying Twitter data during the outbreak have been gaining attention.

The subject of sentiment analysis, often known as opinion mining (SAOM), automates the process of obtaining public views from written content on blogs, social media, and other platforms. It examines attitudes towards things like goods, people, or services as well as individual experiences, feelings, and opinions. At several levels, SAOM makes use of data mining, NLP, and text mining.

An emotional spectrum among Italians is shown by sentiment analysis of COVID-19 tweets [1][2]. The research analyses 179,108 tweets, displaying a range of emotions like fear, happiness, sorrow, rage, and delight, using machine learning techniques such

Random Forest, Decision Tree, and Support Vector Machine [3]. This research demonstrates how attitudes towards the epidemic have changed over time, with the majority of people feeling negatively about it but also responding well to the vaccination campaign [4]. Sentiment analysis reveals a lack of favourable thoughts towards the coronavirus despite the broad negativity, mirroring the public's views and actions regarding the current pandemic [5]. Sentiment analysis plays a critical role in decision-making, social consensus-building, and analysing public views and feelings during epidemic times.

The research investigates whether adding these elements enhances sentiment analysis models and how linguistic and cultural heterogeneity affects sentiment analysis of COVID-19 tweets. It focuses on language structure, grammar, and semantics as well as cultural norms, values, and expressions. In order to increase sentiment analysis accuracy and dependability, the research examines language elements and cultural quirks in COVID-19 tweets, concentrating on individuality vs. collectivism, power distance, and uncertainty avoidance. By tackling the difficulties caused by language and cultural variety in the context of COVID-19 tweets, this study seeks to develop sentiment analysis algorithms. Our knowledge of public sentiment during global health emergencies may be improved, and more effective public health communication methods can be made possible by using the results to influence the creation of more robust and culturally appropriate sentiment analysis models.

The next sections follow the introduction; part II describes the background aspects of sentiment analysis. Section III provides research question in sentiment analysis using Twitter data. Section IV discusses the research objectives. Section V explain the methodology with appropriate parameters. A results and analysis mentioned in Section VI. Section VII provides a brief explanation of potential conclusions.

II. RELATED WORK

There is a publishing delay in detecting psychological hotspot problems when bibliometric data is the exclusive emphasis. In order to address this problem, we propose to find participants in psychological research by examining the present academic debate on Twitter. We gathered all 69,963 tweets from 139 accounts of German-speaking psychology professors, departments, and research institutions, as well as divisions of the German Psychological Society (DGPs), that were sent between August 2007 and July 2020 [1].

All around the globe, the COVID-19 pandemic has had a significant impact on people's everyday lives. Governments use a variety of tactics and plans to fight the disease during the epidemic, including social distance and lockdowns. Together with the virus itself, these practices and policies may cause people to experience a range of mental health problems, including depression, anxiety, melancholy, and so forth. In this study, we leverage enormous text data provided by Twitter users to investigate the sentiment dynamics of residents of the Australian state of New South Wales (NSW) throughout the epidemic [2].

A kind of RNA disease known as corona viruses affects animals with warm blood and feathers. These diseases may induce respiratory plot pollution in people that ranges from mild to fatal. There are some milder cases of the common cold (which may also be caused by other illnesses, such rhinoviruses), but more serious cases of COVID-19, MERS, and SARS might occur. The adverse effects vary by species: they induce loose stools in pigs and dairy animals, but upper respiratory tract illnesses in hens. Antibodies or antiviral drugs are not yet available to treat or prevent human coronavirus infections [3].

A novel viral ailment that surfaced in 2019 is called coronavirus disease (COVID 19). Since the virus has now spread over the whole planet, almost every country is fighting it and making every effort to stop it. As they search for a treatment, the World Health Organization has designated it a pandemic (World Health Organization, 2020) and is taking all necessary steps to control it [4].

The current coronavirus illness (COVID-19) epidemic is having a negative impact on human welfare, the global health system, and the social, economic, and psychological aspects of society. Social media is being used by people, groups, and governments to discuss the COVID-19 pandemic. The issues surrounding COVID-19 that are being addressed on social media sites are not well-documented. Policymakers and healthcare institutions may find it easier to assess stakeholder demands and respond to them with the help of such data analysis. This research

aims to determine which COVID-19 pandemic issues are most commonly discussed on Twitter [5].

A growing number of data owners are outsourcing their data to cloud servers for easier data management and mining because cloud computing offers people and businesses massive processing power and scalable storage capacities to support a wide range of big data applications in fields like health care and scientific research [6].

As more individuals get online, social media is becoming into a platform for opinion sharing. On September 18, 2016, a group of four terrorists attacked security personnel. In this paper, we use Twitter extracts to study the attitudes and survival of tweets after a terrorist occurrence. The information flow of data published on Twitter is examined using variables like the amount of likes, retweets, and the most recent retweet. The reach increases with the quantity of retweets. Social media users were quite outraged by the occurrence. We analyse both the survival of tweets and our sentiment statistics [7].

Social media is becoming into a valuable tool for supporting business decisions. This paper presents a novel approach to social ETL design by combining sentiment analysis techniques with ETL design to include social opinion data into multidimensional designs. The definition of a lexical opinion analysis method that extracts the sentiment polarity of informal language offered in the Twitter social network is the main contribution of this research [7].

One of the primary factors influencing the sequence in which AI-based opinion inquiry frameworks are executed is the extremity shift problem. In this study, we solve the extremity shift problem in report level sensation grouping by proposing a three-stage overflow model. First, we break up each record into many subsentences, and then we build a hybrid model that combines factual techniques and rules to discern between express and certain extreme changes independently. Additionally, in order to remove extremity shift in refutations, we provide a disposal strategy. Finally, we train base classifiers on prepared subsets separated by different types of extremities motions, and use a weighted combination of the component classifiers for sentiment classification. The results of a number of studies show that our approach for extremity shift identification and

disposal substantially outperforms a few other alternatives [8].

Emotional analysis, which addresses the computational handling of judgement, opinion, and subjectivity in text, has received a great deal of attention lately. We are interested in the fine-grained perspective-based feeling inquiry that implies to discern perspectives that customers comment on and these angles' polarities, as opposed to the traditional coarse-grained feeling examination tasks, such as report level opinion order. Angle assembled opinion research is highly dependent on syntactic components. All the same, this project focuses on regular and unconstrained surveys, which in turn serve as a test for syntactic parsers. To tackle this problem, we provide a method in this study that adds a feeling sentence pressure (Sent_Comp) phase before doing the viewpoint-based opinion analysis. Sent_Comp, which is different from the previous sentence pressure model for regular news sentences, attempts to remove the opinion meaningless data for feeling analysis. This simplifies and reduces the complexity of a complex feeling sentence into one that is easier to understand. In order to subsequently pack emotional phrases, we use a discriminative contingent irregular field model with particular distinctive features. Sent_Comp basically focuses on the display of the angle-based sensation analysis by using the Chinese corpus of four item spaces. Sentence pressure may be felt with the aid of the Sent_Comp components, especially the predicted semantic highlights [9].

Web 2.0 developments and applications, together with the emergence of the so-called Social Web, have altered the nature of the Internet in the last few years. with the traditional Web, customers were just purchasers of data; but, with the Social Web, they now play an even more active role as information providers. Due to the widespread use of the most popular method of creating Web material, many public and private organizations are focusing on analyzing this information to ascertain the opinions of the general public about different issues [10].

In order to generate workable and flexible solutions, computerized approaches are essential given the ongoing size and development rate of the Web. The purpose of assessment mining, a very active research area, is to remove different kinds of instructional and added-value components from the viewpoints of consumers by combining natural language processing,

computational semantics, and text analysis approaches. However, a number of limitations limit flow assessment mining techniques, such as the absence of novel numerical strategies in sensation examination processes or the lack of semantic linkages between ideas in highlight search processes. In this study, we propose a unique evaluation mining approach that leverages new Semantic Web-directed responses to improve on the outcomes of conventional regular language handling techniques and opinion analysis protocols. The two main objectives of the proposed philosophy are to: (1) improve highlight-based evaluation mining by using ontologies during the element selection phase; and (2) provide an additional vector investigation-based method for opinion analysis. In a real-world scenario relating to a movie survey, the theory has been implemented and extensively tested, yielding extremely positive outcomes in comparison to other traditional methods [11].

During the COVID pandemic, many people experienced a range of mental health issues, which led to emotional swings. People expressed their emotions on social media. Social media platforms therefore provide an enormous amount of data for comprehending people's attitudes and reactions to the situations they encountered throughout the pandemic. The COVID-19 study's data sources are shown in Table 1. It proves that during COVID, Twitter served as the main source of data [12].

Thirty-one major studies used Twitter as a research source; twenty-four of them. However, a WeChat account, Yelp, Reddit, and other media forums constituted the remaining data sources. Twitter: others use Twitter to express their thoughts and ideas with others all around the world. It is the most downloaded app, with 81.47 million users [9]. "Tweets" are a way that people express themselves. About 200 billion tweets are published annually, according to [10].

In order to stop the COVID virus from spreading, social distance was used, which reduced human-to-human interaction. Numerous countries-imposed lockdowns and limited access to buildings, including schools. Students in particular were anxious and tense since they were compelled to stay away from their families, stay in their dormitories, and stop their academic endeavors as a result of the lockdown. Through social media, students express their

emotions, and researchers have tried to learn more about these emotions and students [13].

The coronavirus has affected billions of people's lives either directly or indirectly. Reopening has become challenging because of the global economic crisis it has caused. Any nation's ability to survive is threatened by a protracted economic shutdown. These circumstances are forcing people to restart their businesses and go back to their usual lives. Consequently, the researchers concentrated on finding out what people felt of reopening after COVID-19 [14].

In today's digital environment, consumers may provide their thoughts and feedback about the calibre of goods or services they use from different businesses. When other customers decide which product or service to use, these reviews help them make an informed decision. The star ratings seen in online reviews have an impact on the restaurant's revenue. People were quite concerned about the COVID-spread, and eateries were provided special SOPs during the pandemic. Consequently, several restaurants got negative reviews for their cold outside spaces and inadequate service. Researchers looked at what customers thought of restaurants, which helped managers maintain high-caliber food and atmosphere [15].

The development of a COVID vaccine could help stop the virus's spread. As a consequence, several businesses are working hard to develop different kinds of vaccines [16]. But vaccination uptake and reception are crucial for using immunizations to reduce COVID. The ability of people to help themselves will be compromised if COVID refuses to assist them. Researchers studied into public perceptions about immunizations in [17]. Additionally, COVID raised cross-border perceptions of prejudice, which in turn raised racist behavior [18].

III. PROBLEM IDENTIFICATION

The purpose of this study is to look at how linguistic and cultural diversity affects sentiment analysis of tweets about COVID-19. The project specifically aims to investigate the following research questions:

- What effect does linguistic variability—that is, differences in syntax, semantics, and language structure—have on the sentiment analysis accuracy of COVID-19 tweets in various languages?

- How much do cultural variations in expressions, norms, and values influence how emotion in COVID-19 tweets is interpreted?
- Can sentiment analysis models for COVID-19 tweets across linguistic and cultural backgrounds perform better when taking linguistic and cultural context into account?

IV. RESEARCH OBJECTIVES

- Examine linguistic patterns and traits in COVID-19 tweets in a variety of languages to see how language-specific traits affect sentiment analysis results.
- Look at cultural subtleties and variations in how emotion is expressed in tweets connected to COVID-19 by analysing culture aspects including power distance, uncertainty avoidance, and individuality vs. collectivism.
- To improve the precision and dependability of sentiment analysis for COVID-19 tweets across various linguistic and cultural demographics, create or modify sentiment analysis models that take linguistic and cultural context clues into account.

V. METHODOLOGY

A statistical method used in sentiment analysis and text classification in natural language processing (NLP) is called count-based logistic regression. In order to predict binary outcomes based on predictor variables, like positive or negative emotion, logistic regression is used. Count Another use of this method is Vectorizer, which converts text input into numerical properties for machine learning algorithms. It counts the occurrences of every word in the vocabulary using a vector representation of each page. In doing so, a vocabulary of words from the whole corpus of text data is produced, and each document is then represented by this vocabulary. Grid Search Optimisation, a pre-training procedure for determining the model's hyperparameters, is often used to train logistic regression models for sentiment analysis or text classification applications. The Logistic Regression Count Vectorizer with Grid Search Optimisation (LRCV-GSO) approach for sentiment analysis of tweets operates as follows, in a condensed, step-by-step overview:

Input: Covid-19 Tweets Dataset

Output: Precision, Recall, F1-Score, Accuracy

Step 1- Data Preparation: Compile a dataset of tweets that have been classified as neutral, negative, or positive.

Step 2- Preprocessing: Make sure the text data is clean by eliminating extraneous characters, URLs, punctuation, and changing the content's case to lowercase.

Step 3-Feature Extraction: To create numerical feature vectors from the pre-processed text data, use Count Vectorizer. Every tweet is visualised as a vector, with each element denoting the number of that term in the tweet.

Step 4- Model Training: Utilising the feature vectors acquired by Count Vectorizer, train a logistic regression model. You train a model on the training set and use an evaluation measure (accuracy, F1-score, etc.) to assess the model's performance on the validation set for each combination of hyperparameters in the grid.

Step 5- Choose the greatest Hyperparameters: Based on your selected evaluation measure, you analyse all possible combinations of hyperparameters and choose the one that yields the greatest performance.

Step 6- Model Testing: In order to get an objective assessment of the model's performance, you lastly assess the model using the optimal hyperparameters on a different test set.

Step 7- Model Evaluation: To determine the efficacy of the trained logistic regression model in sentiment analysis, analyse its performance on a different test set using measures like accuracy, precision, recall, and F1-score.

VI. RESULTS AND DISCUSSION

The following observations are performed on anaconda navigator with python 3.11.1 with jupyter lab toolbox. The proposed procedure LRCV-GSO perform on (Covid 19 Dataset) Coronavirus Tweets.csv and calculate precision, recall, F1-Score and accuracy parameters are calculated as follows:

Figure 1: Load dataset and display their details

Figure 5: Calculation of confusion matrix, precision, recall, F1-Score and accuracy among different models and LRCV-GSO (Proposed Model)

Table 1: Estimation of Sentiment Parameters using LRCV-GSO (Proposed Model)

Sentiment Parameter	Count
Positive	11422
Negative	9917
Neutral	7713
Extremely Positive	6624
Extremely Negative	5481

Figure 2: Text Preprocessing of covid tweets for LRCV-GSO (Proposed Model)

Figure 6: Graphical Analysis of Sentiment Parameters using LRCV-GSO (Proposed Model)

Figure 3: Evaluation of removing stop words and stemming in text preprocessing for LRCV-GSO (Proposed Model)

Figure 4: Evaluation of sentiment analysis of LRCV-GSO (Proposed Model)

Figure 7: Spherical Analysis of Sentiment Proportions LRCV-GSO (Proposed Model)

Table 2: Estimation of Precision, Recall and F1-Score among different models and LRCV-GSO (Proposed Model) for Negative Tweets

Models	Precision	Recall	F1-Score
KNN [4]	0.55	0.40	0.46
Decision Tree [6]	0.69	0.68	0.68
SVM [7]	0.77	0.75	0.76
XG Boost [11]	0.7	0.6	0.79
LRCV-GSO (Proposed Model)	0.79	0.79	0.79

Figure 8: Graphical Analysis of Precision among different models and LRCV-GSO (Proposed Model) for Negative Tweets

The above graph show that the proposed model gives better precision for negative tweets as compare than other models. The precision of LRCV-GSO (Proposed Model) is improve by 0.02 as compare than SVM prediction model.
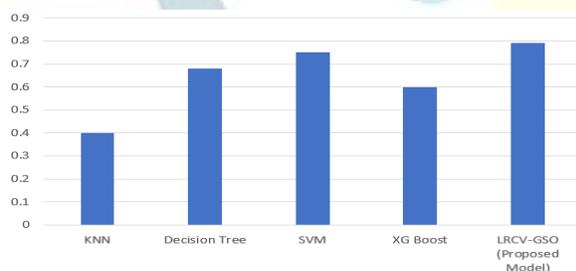

Figure 9: Graphical Analysis of Recall among different models and LRCV-GSO (Proposed Model) for Negative Tweets

The above graph show that the proposed model gives better recall for negative tweets as compare than other models. The recall of LRCV-GSO (Proposed Model) is improve by 0.04 as compare than SVM prediction model.

Figure 10: Graphical Analysis of F1-Score among different models and LRCV-GSO (Proposed Model) for Negative Tweets

The above graph show that the proposed model gives better F1-Score for negative tweets as compare than other models. The F1-Score of LRCV-GSO (Proposed Model) is similar to XG Boost Classifier model.

Table 3: Estimation of Precision, Recall and F1-Score among different models and LRCV-GSO (Proposed Model) for Neutral Tweets

Models	Precision	Recall	F1-Score
KNN [4]	0.25	0.72	0.37
Decision Tree [6]	0.62	0.67	0.64
SVM [7]	0.65	0.66	0.66
XG Boost [11]	0.75	0.65	0.68
LRCV-GSO (Proposed Model)	0.68	0.68	0.68

Figure 11: Graphical Analysis of Precision among different models and LRCV-GSO (Proposed Model) for Neutral Tweets

The above graph show that the XG-Boost model gives better precision for neutral tweets as compare than other models.
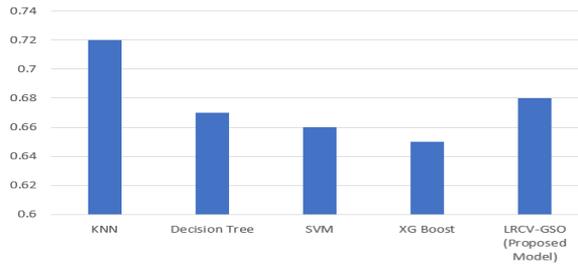
Figure 12: Graphical Analysis of Recall among different models and LRCV-GSO (Proposed Model) for Neutral Tweets

The above graph show that the KNN model gives better recall for neutral tweets as compare than other models.
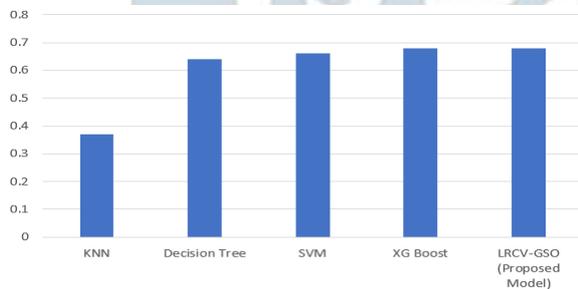

Figure 13: Graphical Analysis of F1-Score among different models and LRCV-GSO (Proposed Model) for Neutral Tweets

The above graph show that the proposed model gives better F1-Score for neutral tweets as compare than other models. The F1-Score of LRCV-GSO (Proposed Model) is similar to XG Boost Classifier model.

Table 3: Estimation of Precision, Recall and F1-Score among different models and LRCV-GSO (Proposed Model) for Positive Tweets

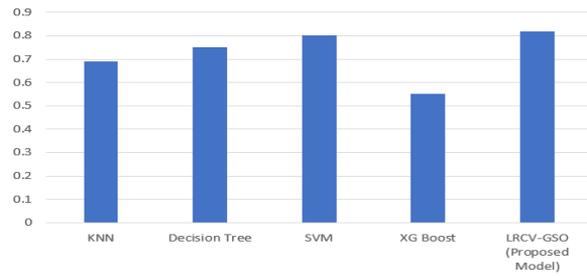| Models	Precision	Recall	F1-Score
KNN [4]	0.69	0.29	0.41
Decision Tree [6]	0.75	0.74	0.75
SVM [7]	0.8	0.81	0.8
XG Boost [11]	0.55	0.86	0.67
LRCV-GSO (Proposed Model)	0.82	0.83	0.82

Figure 14: Graphical Analysis of Precision among different models and LRCV-GSO (Proposed Model) for Positive Tweets

The above graph show that the proposed model gives better precision for positive tweets as compare than other models. The precision of LRCV-GSO (Proposed Model) is better than SVM Classifier model.

Figure 15: Graphical Analysis of Recall among different models and LRCV-GSO (Proposed Model) for Positive Tweets

The above graph show that the XG-Boost model gives better recall for positive tweets as compare than other models.

Figure 16: Graphical Analysis of F1-Score among different models and LRCV-GSO (Proposed Model) for Positive Tweets

The above graph show that the proposed model gives better F1-Score for positive tweets as compare than other models. The F1-Score of LRCV-GSO (Proposed Model) is better than SVM model.

Table 4: Estimation of Accuracy among different models and LRCV-GSO (Proposed Model) for Positive Tweets

Models	Accuracy
KNN [4]	0.41
Decision Tree [6]	0.7
SVM [7]	0.76
XG Boost [11]	0.6
LRCV-GSO (Proposed Model)	0.78

Figure 17: Graphical Analysis of Accuracy among different models and LRCV-GSO (Proposed Model) for Covid-19 Tweets

The above graph show that the proposed model gives better accuracy for covid 19 tweets as compare than other models. The accuracy of LRCV-GSO (Proposed Model) is better than SVM Classifier model.

VII. CONCLUSION

Opinion identification has to be improved by library professionals due to the exponential expansion of electronic text data. Increasing recognisable evidence competence across various assessment boundaries is the goal of a suggested model. By using word references, the model pre-handles the data and extracts designs and catchphrases. When less information is needed, such in geo-roused calculations, the accuracy of the model rises. The suggested method identifies text information in organisations, accepts human addition for recognisable evidence, employs stemming and lemmatizing procedures for great opinion content arrangement, and improves the learning model by using the Grid Search Optimisation (GSO) technology. The model may be modified for usage in the future, however further testing with other datasets and AI calculations will be necessary.

REFERENCES

[1] Bittermann A, Batzdorfer V, Müller SM, Steinmetz H. Mining Twitter to detect hotspots in psychology. Zeitschrift für Psychologie. 2021;.

[2] https://en.wikipedia.org/wiki/Fuzzy_logic, Anonymous, retrieved 8th August, 2020.

[3] Abd-Alrazaq A, Alhuwail D, Househ M, Hamdi M, Shah Z Top Concerns of Tweeters During the COVID-19 Pandemic: Infoveillance Study J Med Internet Res 2020;22(4):e19016DOI: 10.2196/19016 PMID: 32287039.

[4] Li, S., Wang, Y., Xue, J., Zhao, N., & Zhu, T. (2020). The impact of COVID-19 epidemic declaration on psychological consequences: a study on active Weibo users. International journal of environmental research and public health, 17(6), 2020.

[5] Koyel Chakraborty, Surbhi Bhatia, Siddhartha Bhattacharyya, Jan Platos, Rajib Bag and Aboul Ella Hassanien, "Sentiment Analysis of COVID-19 tweets by Deep Learning Classifiers – a study to show how popularity is affecting accuracy in social media", Elsevier Journal of Applied Soft Computing, Volume 97, Part A, December 2020.

[6] https://en.wikipedia.org/wiki/Naive_Bayes_classifier, Anonymous, retrieved 12th September, 2020

[7] Abd-Alrazaq A, Alhuwail D, Househ M, Hamdi M, Shah Z. Top concerns of tweeters during the COVID-19 pandemic: infoveillance study. Journal of medical Internet research. 2020.

[8] Chakraborty K, Bhatia S, Bhattacharyya S, Platos J, Bag R, Hassanien AE. Sentiment Analysis of COVID-19 tweets by Deep Learning Classifiers—A study to show how popularity is affecting accuracy in social media. Applied Soft Computing. 2020; 97: 106754.

[9] Barkur G, Vibha GBK. Sentiment analysis of nationwide lockdown due to COVID 19 outbreak: Evidence from India. Asian journal of psychiatry. 2020; 51:102089.

[10] Zhou J, Yang S, Xiao C, Chen F. Examination of community sentiment dynamics due to COVID-19 pandemic: A case study from Australia. arXiv preprint arXiv:200612185. 2020.

[11] Organization WH, et al. Coronavirus disease 2019 (COVID-19): situation report, 72; 2020. Available from: https://apps.who.int/iris/handle/10665/331685.

[12] GuixianXu ,Ziheng Yu , Haishen Yao, Fan Li, Yueting Meng, And Xu Wu. "Chinese Text Sentiment Analysis Based On Extended Sentiment Dictionary". Ieee Access April 13, 2019. Digital Object Identifier 10.1109/Access.2019.2907772.

[13] Farkhund Iqbal, Jahanzeb Maqbool Hashmi, Benjamin C. M. Fung, Rabia Batool, Asad Masood Khattak, Saiqa Aleem, And Patrick C. K. Hung. "A Hybrid Framework for Sentiment Analysis Using Genetic Algorithm Based Feature Reduction". IEEE Access volume 7 February 8, 2019.

[14] Trușcă, M. M. (2019, October). Efficiency of SVM classifier with Word2Vec and Doc2Vec models. In Proceedings of the International Conference on Applied Statistics (Vol. 1, No. 1, pp. 496-503). Sciendo.

[15] Sharma, R., Mondal, D., and Bhattacharyya, P. (2018). A comparison among significance tests and other feature building methods for sentiment analysis: A first study. In Gelbukh, A., editor, Computational Linguistics and Intelligent Text Processing, pages 3–19, Cham. Springer International Publishing.

[16] Tyagi, Abhilasha & Sharma, Naresh. (2018). Sentiment Analysis using Logistic Regression and Effective Word Score Heuristic. International Journal of Engineering and Technology (UAE). 7. 20-23. 10.14419/ijet.v7i2.24.11991.

[17] Tyagi, Abhilasha & Sharma, Naresh. (2018). Sentiment Analysis using Logistic Regression and Effective Word Score Heuristic. International Journal of Engineering and Technology (UAE). 7. 20-23. 10.14419/ijet.v7i2.24.11991.

[18] Ying Fang, Hai Tan, And Jun Zhang. "Multi-Strategy Sentiment Analysis of Consumer Reviews Based On Semantic Fuzziness". Ieee Access May 2, 2018.Digital Object Identifier 10.1109/Access.2018.2820025.

[19] Lei Wang, JianweiNiu, And Shui Yu. "Sentidiff: Combining Textual Information and Sentiment Diffusion Patterns For Twitter Sentiment Analysis".

Journal Of Latex Class Files, Vol. 14, No. 8, August 2018.

[20] Ali Sadollah (October 31st 2018). Introductory Chapter: Which Membership Function is Appropriate in Fuzzy System?, Fuzzy Logic Based in Optimization Methods and Control Systems and Its Applications, Ali Sadollah, IntechOpen, DOI: 10.5772/intechopen.79552.

[21] A. M. Shah, X. Yan, A. Qayyum, R. A. Naqvi, S. J. Shah, Mining topic and sentiment dynamics in physician rating websites during the early wave of the COVID-19 pandemic: Machine learning approach, Int. J. Med. Inform. 149 (2021). doi:10.1016/j.ijmedinf. 2021.104434.