

Sentiment Analysis on the Basis of Tweeter Comments of Application of Drugs by Customary Language Toolkit and TextBlob Opinions of Distinct Countries

Dr. Rajesh Bose¹, Dr. P. S. Aithal², Dr. Sandip Roy³

¹Post Doctorate Research Fellow, Srinivas Institute of Management Studies, Srinivas University, Mangalore – 575 001, India, bose.raj00028@gmail.com

²Srinivas Institute of Management Studies, Srinivas University, Mangalore – 575 001, India
psaithal@srinivasgroup.com

³Brainware University, India, sandipro86@gmail.com

ABSTRACT

Twitter is now one of the greatest podiums all over the world through which anyone can present his opinion regarding a particular issue. It is a social networking site. Every day, it generates almost 500 million of tweets and the total volume of tweets contains 8TB of data. This data is very much significant if examined, because we are able to draw out salient facts through opinion mining. We can keep an eye at the augmentation of a product or any important affairs or events or a novel fashion in twitter data. The prime objective of this sentiment analysis or opinion mining is to explore emotion, opinion, subjectivity and perspective from a customary text on application of a medicine for the treatment of COVID-19(Corona Virus Disease 2019). We can classify the tweets into positive and negative sentiments in twitter sentiment analysis. The expression 'Cluster' refers to an accomplished method where homogeneous substances are kept in a distinct class and thus a pack of cluster is created. We went on with an analysis and concluded that the application of clustering can hurriedly and precisely differentiate tweets on account of their sentiment scores and thus it gets weekly and firmly positive or negative tweets, if they are clustered, with outcomes of distinct dictionaries. The objective of this paper is to scrutinize discrete viewpoints of clustering sentiment analysis and thus devises a method to make connection with the tweets of the statesmen of some first world, second world and third world countries. The frame of reference in this case is the polarity and subjectivity on the application of drugs for the treatment of COVID-19.

Key words: Cluster, Opinions, Sentiments, Twitter.

1. INTRODUCTION

After advent of internet, mankind has a great exposure to express themselves in various ways. Internet has taught human beings how to establish their opinions, views etc. before the world [9]. Twitter is such an exposure. It is a social networking site in which people can give their opinions regarding a particular issue. Now, to serve the purposes of various businesses too, twitter is frequently used and especially celebrities from different fields share their opinions, thoughts etc. regarding a particular event. The word limit of a tweet is not more than 140. Consequently, a user posts a twitter in brief [10]. Thus, the users of this podium are able to make common folks feel the sentiments of their tweets easily.

These sentiments that we watch in the tweets are defined in three ways: - positive, negative and neutral. The tweets that shower positive vibes for the concerned sentiments are called positive tweets. The negative sentiments are just opposite of them. Such sentiments are made in insulting or criticizing tone about the concerned issue, and neutral sentiments are those which are neither in appreciating tone nor in criticizing or insulting tone [9]. These sentiments as well as the tweets can be classified in some groups but the objective of this paper is somewhat different, and that is clustering with different sentiment analysis.

Clustering is a procedure that offers the allocation of particular classes to particular substances and in this procedure the substances in the same group are akin, whereas in the other divergent groups, they are not so [11].

In the sentiment analysis, distinct objects are taken as a whole in a group. They are outcomes of sentiments positivity, subjectivity, objectivity etc. Here clustering one another is made without any kind of supervision.

1.1 Sentiment Analysis

Sentiment analysis refers to determining the frame of mind of one's outlook written in normal language. After spotting

the frame of mind, the positive or negative polarity is allocated [12]. Several means are used to a normal write-up for analyzing the sentiments such as derivation of attributes, study of emojis, tokenization etc. When sentiment analysis is done, customarily, positive and negative words are derived from the write-up and designated an outcome from the dictionary of words. The outcomes of this analysis are considered with the reference frame of two dictionaries. These are used after the preprocessing and tokenization of the tweets. Such dictionaries are AFINN and TextBlob. Using these dictionaries how sentiment analysis is done, is revealed in this paper afterwards.

1.2 Clustering

Clustering is a method which is done without any kind of supervision. It is executed on the basis of identical data to create a model which is able to produce proper outcomes following the procedure of training data. In this method, the identical substances are allocated to a distinct class from them which are not akin to them [13].

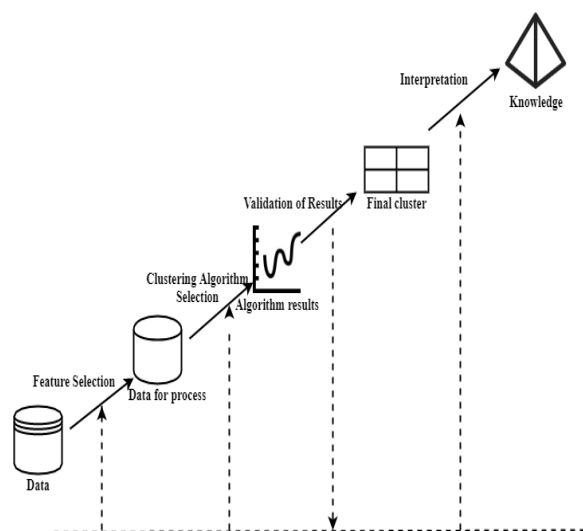


Figure 1: Stages of clustering

Outcomes of sentiments are computed in this study, for each tweet. Then we implement unsupervised means to them. At this, classes are created automatically on the basis of resemblance such as clustering. Survey on clustering is a necessary distinct evidence of classes with the identical substances or models [2]. We can present an instance as follows: - If we designate a data in a packet that is akin to other packet of groups we may be clearly inclined to that packet and create those packets of groups into a real cluster. Thus, the authors can increase the number of groups within identical and other intra groups. Consequently, they either obstruct or allocate the clusters which create a monitoring with precise volume devoid of any kind of excessive stress. Figure 1 represents of junctures of clustering.

2. RELATED WORK

With the help of the presentation of clustering on detectable guidelines like name, address and friend list based on twitter data, Bergsma *et al.*, we could predict a few attributes like identity and gender [1]. The necessity guidelines for any data analysis project were stated by Romero *et al* [2]. Using PCA, Vanessa Friedmann *et al.*, shortened the data and changed it to a lower dimensional attribute capacity. They are able to implement this attribute to alter to the K-means algorithm to split the specimens into clusters. In 2009, Go *et al.* [4] implemented the emoticons as a section of distant learning to achieve emotion or sentiment. The tweets that one finish with “☹” was called negative tweets. Unseen Bergsma *et al.* on the basis of twitter data were able to forecast some invisible features such as ethnicity and gender by performing clustering on visible parameters such as name, location and friend-list [1]. Romero, *et al.* [2] emphasizes the need to classify data and study of prediction parameters for any data analysis project. Vanessa Friedmann *et al.* [3] with the help of PCA, pruned the data and altered it into a lower dimensional feature space. This feature can be used to be passed into the k-means algorithm to segment the samples into clusters. Go *et al.* [4] in 2009 used the emoticons as part of distant learning to obtain emotion or sentiment. The tweets that people end with “☹” were called positive tweets and the tweets that people end with “☹” were called negative tweets. When we talk about other prominent endeavors, for the categorization of twitter data sentiments, we recall Barbosa *et al.* [5] suggested a polarity assumption that applied the data from three websites as obnoxious tags in an effort to inculcate a prototype and again apply almost thousand manually tagged tweets for verifying its accuracy. Applying C3ESL, LuizF.S Coleita *et al.* [6] contemplated amalgamation of classifiers and clusters in the categorization of tweet sentiments. They also contemplated that the additional data managed by data altercations created from “bag-of-words” with lexicon scores is able to modify the accuracy of categorization produced by classifiers. Thus, the clusters are able to create a pattern of “topic structure” which is used for encoding the data. It remains in the mode of meta-information. Phan *et al.* [7] showed how a structure is used for categorizing task with the participation of short and sporadic text chunks. The chief elements of their structure are the amalgamation of the relevant segment “universal dataset”. This segment is used to regulate the concealed subject-matters and to allocate categorizing task. A POS tagger was launched by Gipel *et al* [8] for twitter data. At first the modification of the attributes for twitter POS tagging was done by them, and then they went on with the acid tests to assess them. Finally, they came up with their elucidated complication and guided POS tagger to the research world.

3. CLASSIFICATION ALGORITHM

3.1 Naïve Bayes

Naïve Bayes classifiers are the teams of easy "probabilistic classifiers" that implements Bayes' theorem with solid (naïve) independence predictions between the attributes of machine learning. These are the easiest Bayesian network prototypes. [14] Naïve Bayes has been experimenting vigorously from 1960s. This was manifested anonymously into the retrieval group in the early 1960s, [15] and still it is a standard prototype for text classification, to one of the classes or the other like as spam or legitimate, sports or politics, etc. with global frequencies as a customary attributes. It is broadly used in automatic clinical detections. [16]

When the attribute has a specific tag, then Bayes equation is provided as –

$$P(l | f) = \frac{P(l) * P(f | l)}{P(f)} \quad (1)$$

$P(l)$ denotes the prior probability of a label or likelihood that a random feature set the label.

$P(f|l)$ denotes the prior probability of a given feature that is being classified as label.

$P(f)$ refers to prior probability that a given feature has occurred. In general, above equation could be rewritten as –

$$P(f | lb) = \frac{d(w, e(f | lb))}{sl(d(w, e(f | l))f)} \quad (2)$$

Kang and yoo [17] suggested an impromptu in Naïve bayes algorithm to augment the standard precision. They succeeded in reaching their goal to distinguish between the positive precision and the negative precision, introducing unigrams and bigrams as attribute due to augmentation of the standard precision.

For the categorization of positive and negative tweets Pak and Paroubek [18] suggested a prototype, implementing Naïve Bayes algorithm. Implementing twitter API for the amalgamation of tweets that have sentiments, twitter compilation was formed. POS-labels and N-gram attribute derivation methods were implemented. But the prototype gets less methodical, and it happens as training group regards the tweets taking in sentiments.

Polarity identification, which was created by Po-Wei liang and Bi-Ru Dai [19], blends machine learning methods and domain determined data of acquired tweets implementing Twitter API. In this event, there are applications of Unigram Naive Bayes algorithm and Mutual Information and Chi square attribute derivation methods. Even after that the suggested prototype did not produce reliable precision.

3.2 K-Mean

K-Mean is one of the easiest autonomous methods for clustering. Here the specimen information is categorized into K clusters. These remain far from one another. Next each data point is allocated to the cluster keeping the closest distance from it such as with the shortest Euclidean distance. Thus each data gets allocated to one of the opted clusters. [20]. Then the hub of the clusters are computed further and the homogeneous method is continued until the hubs cease to shift locations.

At first, MacQueen suggested that K-Means is an autonomous learning algorithm, in 1967 [21] It is a prominent technique of cluster analysis. This is a means of regulating the itemized substances into distinct groups named clusters. It is regulated on the basis of resemblance among substances which are based on distinct benchmarks. With the application of distinct characteristics and conduction of an insistent rotating apt technique, it elucidates the popular clustering problem. This algorithm separates a dataset X into k disjoint clusters such that each observation belongs to the cluster with the nearest

mean. Let \bar{x}_i be the centroid of cluster c_i and let $d(x_j, \bar{x}_i)$

be the dissimilarity between \bar{x}_i and object $x_j \in c_i$. Then the role minimized by the k-means is given by the following equation:

$$\min_{x_1, \dots, x_k} E = \sum_{i=1}^k \sum_{x_j \in c_i} d(x_j, \bar{x}_i) \quad (3)$$

We can use K-Means clustering in numerous applications that take in machine learning, identification of errors, understanding specimens, operating pictures, statistics and artificial intelligence [22,23,24] It is granted one of the quickest clustering algorithms with several alternatives. These alternatives are receptive to the preference of primary stages and are inclined to make solution of numerous matters of K-Means such as the assessment of the number of clusters [25], the procedure of start up of the axis of the clusters [26] and the tempo of the algorithm [27].

3.3 Fuzzy C-Mean

This algorithm allows the implementation of distance between the cluster hub and data point to choose membership task to each data point in connection with each cluster hub. Growth of the number of the membership tasks towards the cluster hub is directly proportional to the closeness of the data set to the hub of the cluster [28]. The total number of membership values should be integrity in each data point. After each renewal, the memberships and cluster hubs are modified.

The development of the number of the fundamental c-means impartial task or some improvement of the impartial task is responsible for most systematic fuzzy clustering algorithms. The development of the c-means operation produces a

nonlinear minimization dilemma which we can settle applying a range of techniques combining renewal minimization [29]. Implementation of the easy Picard renewal via the first order circumstances for fixed points, called FCM algorithm, is the most noteworthy approach. Bezdek [30] has shown the merging of FCM algorithm. A flawless C separation within class sum of squared mistake impartial task:

$$(4) \quad J = \sum_{i=1}^n \sum_{j=1}^c (u_{ij})^m d^2(y_i, c_j)$$

where $Y = [y_1, y_2, \dots, y_n]$ is the data set in a d-dimensional vector space, n is the number of data items, c is the number of clusters which is defined by the user where $2 \leq c \leq n, u_{ij}$ is the degree of membership of y_i in the jth cluster, m is a weighted exponent on each fuzzy membership, c_j is the center of cluster $j, d^2(y_i, c_j)$ is a square distance measure between object y_i and cluster c_j .

3.4 Decision Tree

It refers to a model that is like a flowchart, where each central node signifies an experiment on an attribute such as whether a coin flip appears with heads or tails. Here each node refers to a group which makes decision after calculating all attributes, and its branches exemplify unions of attributes which promote those group tags. The entire track from root to leaf exemplifies categorization guide lines.

Statistics, data mining and machine learning are some fields in which Decision Tree is applied as an anticipating designing means. Decision Trees are created as a means of an algorithmic model which determines tracks to break a dataset on the basis of divergent states. Decision Trees are used in a wide area for administered learning approach which is applied for both classification trees which are such prototypes where destination variable is able to take a distinct group of values. On the contrary, those decision trees are regression trees in which the destination variable is able to take uninterrupted values which are customarily real numbers. The generic name for it is Classification and Regression Tree (CART).

Decision Trees are used rapidly. C5.0 algorithm [31] is noteworthy to make Decision Tree. It has gained such worldwide fame that several industries use it for generating decision Trees because it can solve most of the problems precisely without any kind of intervention. In comparison with more modified and refined machine learning prototypes such as Neutral Networks and Support Vector machines, the decision trees under c5.0 algorithm customarily carry out in much simpler form to comprehend and utilize.

C5.0 is a Decision Tree algorithm that is applied to estimate the dilemma in the amalgamation of feature and

efficiency of a feature implementing break up and information achievements consecutively. The implementation of C5.0 on the dataset is classified into two equations:

1. Computing the breakup value of the data applying the equation as follows:

$$G(S, C) = E(S) - \sum_{\omega \in \text{value}(c)} \frac{S_w}{S} E(S_w) \quad (5)$$

Here $E(S)$ – breakup of an accumulation of datasets, c – stands for the number of groups in the device and p_i – stands for the number of example proportion that group I possesses.

2. Computing the information achievement for a feature C, in an accumulation S, in which $E(S)$ is the breakup of the entire accumulation and S is the series of examples which have value w for feature C.

$$E(S) = \sum_{i=1}^c - p_i \log_2 p_i \quad (6)$$

4. SUGGESTED PROTOTYPE

This accumulation illustrates each element of the suggested prototype. The prototype as follows is illustrated in a flowchart in Figure 1. The suggested algorithm in Algorithm 1 has also been illustrated below:

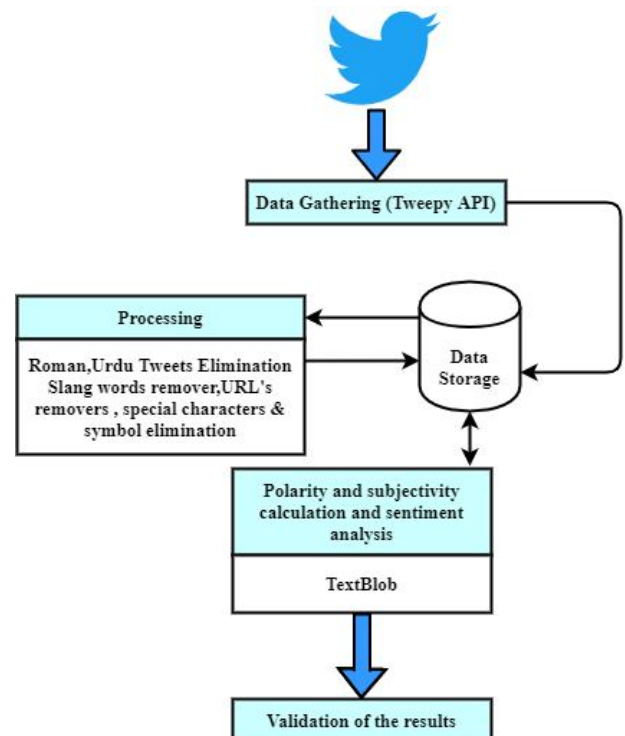


Figure 1: Suggested Prototype

ALGORITHM 1: Sentiment Classification of Tweeter Comments using Text Blob

Input: Text File (Tweeter comments which include Nouns, Adjectives, Adverbs)

Output: Values > 0 (Positive), Values < 0 (Negative), Values = 0 (Neutral)

Begin:

```

1. Sentiment Analysis () ← File
2. For each row in rows
3.   if Sentiment Polarity Score > 0 then
4.     Sentiment ← Positive
5.   else
6.     if Sentiment Polarity Score < 0 then
7.       Sentiment ← Negative
8.     else
9.       if Sentiment Polarity Score = 0 then
10.        Sentiment ← Neutral
11.      else
12.        end
13.    end
14.  end
15. end
    
```

Figure 2: Suggested Algorithm Flowchart

Table 1 is a depiction of a specimen of the junked data in a tabular form in which date and tweeted texts are mentioned for having an assessment of the outcomes of our research.

5. ILLUSTRATION OF DATA

Twitter is nothing but a social networking podium which provides its user with the opportunity of sending and reading micro-blogs whose character limit is not more than 280. It is called a ‘tweet’. Only the registered users are able to read and post their tweets via the web, short message service (SMS) and mobile apps.

Twitter being a cosmopolitan real-time communication podium, more than 400 million users visit monthly and 255 million active users of it are there all over the world. Distinguished personalities such as entire world’s statesmen, actors, sportspersons etc. are twitter’s active groups of registered users. Now, it offers more than 35 languages in it. Jack Dorsey, Evan Williams, Biz Stone, and Noah Glass augmented Twitter in 2006. Twitter’s headquarters are in USA and its offices are located in SanTwitter was launched in 2006 by Jack Dorsey, Evan Williams, Biz Stone, and Noah Glass. Twitter is headquartered in San Francisco, California, USA, with local offices in Atlanta, Austin, Boston, Boulder, Chicago, Detroit, Los Angeles, New York, Sunnynvale, and Washington.

Twitter's international offices are located in Amsterdam, Berlin, Dublin, London, Madrid, Paris, Rio de Janeiro, Sao Paulo, Singapore, Sydney, Seoul, Tokyo, Toronto, and Vancouver.

Table 1: Sample Tweet comments with post date and time

Date	Tweet Text
20-04-20 04:39	#COVID?19 #Kashmirtoday 15 Persons who were earlier COVID-19 Positive have been discharged from 03 Hospitals in Kashmir today. They have recovered and turned Negative. 07 from Baramulla Hospital, 06 from SKIMS Soura and 02 from CD Hospital Srinagar.
20-04-20 04:39	COVID-19 is a matter of grave concern now all over the world. It’s observed here the way by which the negotiations on corona virus have prevailed in social networking sites and how many brands take part in those negotiations.
20-04-20 04:39	In conclusion, Brexit, Britain and the European Union (EU) are expecting to formulate improvement towards a business deal, eleven weeks later. The span of the discussions is three weeks. The first one commences on Monday via video conferencing. #Britain #Coronavirus #EU

20-04-2 020 04:39	If you've a few mins and it applies to you, please take part in this research survey on the effects COVID-19 on gay/bi in the community who use dating apps. RT @TC_Witzel: ??????WE NEED YOU!????? We're recruiting UK folks from the LGBT community who have sex with men - esp gay & bi men (cis & trans), non-binary ppl & trans women - to a survey about hook-up apps, sexual behaviour, HIV & services during lockdown. RTs appreciated!
20-04-2 020 04:39	Due to the on-going uncertainty of the situation relating to COVID-19, @BBallBallarat has sadly had to cancel the 2020 Junior Tournament which was to be held on June 6-8, 2020. The association thanks everyone for their understanding during this difficu...
20-04-2 020 04:39	Health progress in Wajir the regional Covid-19 testing centre ready for testing. The County referral hospital of Wajir has also been adequately equipped to ensure that such issues and suspected cases are handled in a manner that doesn't endanger the lives of health workers.
20-04-2 020 04:39	A Number of Covid-19 Cases Confirmed
20-04-2 020 04:39	We're donating devices #Kindle #FireTablets #Echo #Alexa to [DE] @Caritas [FR] @hopitauxdeparis [IT] @savethechildren and others across #Europe providing extra resources during #COVID-19. Check full article:
20-04-2 020 04:39	.@NicolaSturgeon announces she will spend Barnet bung £ on handing out free food in the Yes favelas despite these people never working so #COVID19 makes fuck all difference to their daily routine #PresserCOVID19 #COVID19 #PressConference
20-04-2 020 04:39	#PatHeads. You might be a redneck if you have COVID-19.
20-04-2 020 04:39	Have you lost your job due to #COVID19? Follow the @EmmittSmith22 game plan to find your next opportunity at #SayYestoDallas #DRCWorks

20-04-2 020 04:39	A Number of Covid-19 Cases Confirmed
20-04-2 020 04:39	As part of the Global Health Crisis Reporting Forum, we're working with @ICFJ to bring you a series of webinars with global experts on COVID-19. Webinars are on the record, and experts can be quoted in stories. Check out this week's offerings:
20-04-2 020 04:39	Gold:"when the COVID-19 pandemic breached their borders +they asked for a loan to help battle the virus, the International Monetary Fund refused. Thankfully, the WHO (as well as China and Russia) have provided the embattled country with medical supplies and technical assistance"
20-04-2 020 04:39	A Phase 1 #clinical trial evaluating an investigational vaccine designed to protect against #coronavirus disease 2019 (COVID-19) has begun at Kaiser Permanente Washington Health #Research Institute (KPWHRI) in Seattle.
20-04-2 020 04:39	Fine tally of yesterday #COVID19 +ve case's in Srikalahasthi is 35 as of now In Last 24 hrs 25 new #COVID19 positive cases in Srikalahasthi
20-04-2 020 04:39	When quarantine is over???? #coronavirus #COVID #QuarantineDays #Covid_19 #CoronaLockdown
20-04-2 020 04:39	An interesting read on the positive impact #COVID19 will have on the digital side of the NHS long-term, and things they've managed to achieve during the pandemic so far.
20-04-2 020 04:39	NEXT: We've got another life hack for you this morning. This time from @travisdolynny. He'll teach us how to make life feel a little more normal by feigning a work 'commute.' We'll catch up with him at 7:45. #ldnont #COVID19

6. ILLUSTRATION OF THE ELEMENT OF SUGGESTED PROTOTYPE

This segment offers a concise illustration of each element of the suggested prototype:

6.1 Sentiment Analysis Element:

The following table 2 is a detailed illustration of the suggested prototype, based on steps:

Table 2: Stepwise Sentiment analysis Components

Steps	Illustration
Tokenization	It signifies the separation of texts in a series of tokens. They practically correlate with “words”. It is one of the fundamental jobs of NLP. For doing this, applying TextBlob, two steps are taken: i) Forming a TextBlob component and making proceeding a string with it. Ii) Designating tasks of TextBlob for performing a distinct job.
Derivation of Noun Phrase	As derivation of the words was made in the foregoing segments, here abstaining from that, we can derive noun phrase from the TextBlob. Derivation of noun phrase is specifically significant when one intends to examine the “who” in a sentence.
Adjoining Part of Speech	Adjoining part of speech or grammatical adjoining is a procedure to symbolize words which exist in a specific text concerning its clarity and subject-matter. If it is an easy word, it just says what part of speech it is. It is a comprehensive category of derivation of noun phrase in which we look for each of the parts of speech in a sentence.
Conjugation of Words	Conjugation is a method of creating a word, where letters are connected with the fundamental form of a word to give the idea of the grammatical meaning. It is a very easy method. The words that tokenized from a Text Blob can be simply converted into singular or plural form of the concerned one.
N-grams	It is an amalgamation of profuse words together. In relation with words, N grams (N>1) is better means in terms of information. It can be applied as attributes for

	designing of language. We are able to obtain it in a simple manner, in TextBlob applying the tasks of N-grams. It fetches an assembly of n sequential words.
Sentiment Analysis	Sentiment analysis is all about identifying the frame of mind or the emotion of the writer which we can categorise into positive, negative or neutral sentiments. Polarity and subjectivity are the two characteristics which are paid back by the sentiment function of TextBlob. Polarity in sentiment analysis refers to determining the sentiment orientation as positive, negative or neutral in written or spoken language. The different kinds of sentiment analysis comprise in compact sentiment analysis which offers more accuracy in the point of polarity such as very positive, positive, very negative, negative and neutral and emotion analysis that focuses on identifying emotions within the expressions as happiness, sadness, frustration, surprise etc. Polarity varies from the range -1 and 1, where 1 indicates a positive statement and -1 represents a negative statement. Language can consist of expressions that are objective or subjective. Objective expressions refer to pieces of evidence and subjective expressions signify conclusions which describe folks’ notions towards a certain subject or topic. It is subjectivity when the content is an illustrative article which must be examined in a context. Subjective sentences typically refer to personal opinion, emotion or judgement when evidently objective sentences refer to realistic information. Subjectivity also varies from the range of 0 and 1.

6.2 Programming Language and simulator

Python:

Python is an elucidated, high-level programming language for serving customary objectives. It's made by Guido van Rossum in 1991. This highlights code clarity with its prominent application of remarkable whitespace. This beneficial and sensual method focuses on and assists programmers to write lucid, consistent code for small and large-scale projects.[32]

Anaconda Navigator:

It's a liberal [33] sharing for Python and R programming languages. This is applied in scientific calculations such as data science, machine learning applications, large-scale data processing, predictive analytics, etc. The objective of it is to facilitate package management and exploitation. Package interpretations are handled in the package management systemconda.[34] This Anaconda sharing contains data-science packages that are fitting for Windows, Linux, and MacOS. Anaconda sharing offers more than 1,500 packages along with the conda package and virtual atmosphere manager. This comprises in too a GUI, Anaconda Navigator [35], for graphical substitute in command line interface (CLI). Hydroxychloroquine and COVID-19

Now, when statesmen of different countries wish to have hydroxychloroquine from India to cope up with COVID-19, WHO (World Health Organization) is recently deeming the application of hydroxychloroquine on COVID-19 patients within the agreement probation. The hydroxychloroquine wing has been ceased as a safeguard and the security data is being analyzed.

Now, the countries which have supervisory power, can provide guidance to their people about the application of any medicine. It's true that HQ and CQ are authorized products for the treatment of other diseases, but right now, these have not been considered fruitful for the treatment for the treatment of COVID-19. Evidently, in many countries, alarming information has been announced on the aftermath of the medicines and their application has been restricted in several countries to medical tests under vigorous surveillance in the arrangement of specialized health centers.

The WHO has suggested to doctors that they should not go on with erroneous treatments to the patients affected with COVID-19. The WHO has also warned people not to make experiments with themselves; in taking self-chosen drugs. The specialists all over the globe agree on this point that the possibility remains but more experiments are required at this time to come to this conclusion whether existing anti-viral medicines can be fruitful for the treatment of COVID-19 or

not. If they get prospective outcome then they will be able to mitigate the hazards of COVID-19.

COVID-19 has taken shape of a pandemic. Every day the graph shows how the number of patients attacked by COVID-19, increases day by day. Consequently, substantial death toll is being observed in a great many countries. It's needless to say that the whole world is in a turbulent situation now. So, kudos to those physicians and their support staff who have devoted themselves to cure the patients attacked by COVID-19. Among the treatments, one has drawn a great attention, and it is the factual application of CQ/HCQ. Although there is no evidence that can support that it's application is fruitful and based on medical science, many countries have already incorporated CQ/HCQ in COVID-19 treatment protocols [36], [37]. They are applying it for both the treatment of critically ill patients and setting of the benchmark of preventive measure.

In this paper, a huge dissimilarity in the results of the studies is manifested in the all-inclusive analysis of the antiviral of HCQ and CQ on SARS-COV-2 (Severe Acute Respiratory Syndrome Corona Virus 2) and a few other viruses. CQ and HCQ determined to generate artificial outcomes; still these kinds of data have not been converted into substantial in vivo researches. Only in some hospitals, some efforts have proposed a few fruitful results of CQ and HCQ in COVID-19 patients. Most of the cases are still elementary [38] [39] [40]. Moreover, not less than 7 of the current efforts were dismissed and it has not been clear till now whether it was for the probability of negative outcomes or other issues.

Numerous noxious elements are connected with such medicines [41] [42] [43]. Among them that which is having the priority, is the probability of QT supplement and the peril of Torsades de pointes, that is an implicitly life sinister arrhythmia [44] [45] [46]. However, when our analysis on literature displayed that it is almost unusual, it is not still clear if there would be some preservative or probable harmonious danger when such drugs are amalgamated with other drugs like AZT [46]. Evidently, it's risky to take a treatment choice if there is not an accomplished research work and an evident insight of medicine potency and security. In this state of ambiguity, we should keep in mind that it will be fatal in using such medicines in COVID-19 patients if comprehensive high standard incessant medical trials cannot be confirmed to explain their function for the treatment of COVID-19 [41].

7. SURVEYING OUTCOME

In this paper, all the outcomes are analyzed on the basis of the database of the information derived from twitter. Who doesn't know now that it is a social networking site in which one can give one's opinion or view on a novel, past or forthcoming event? Here sentiment outcomes are produced concerning the comments/reviews made by users on the concerned twitter

comments. The outcome is prospective or fruitless, that is considered on the basis of positive and negative reviews respectively. Summing up from the majority of the concerned comments, we consider it. In this research work, TextBlob algorithm is used on the twitter comments. This we can see in the model Table 1 and have the polarity or subjectivity score concerning the outcome in Table 3.

Table 3: Sample sentiment result of Tweeter comments

blob.sentiment/Polarity	blob.sentiment/Subjectivity
0.75	0.8
0	0
0.75	0.8
0.8	1
0	1
0.65	0.75
0.483333333	0.6
0.525	0.8
0.525	0.666666667
1	0.3
0.5	0.6
1	0.75
0	0
0	0.1
0.8	1
0.9	1
0.392857143	0.567857143
0.8	1
-0.15	0.05
1	1
0.289285714	0.417857143
0.8	1
0.1	0.125
0.8	1
0.25	1
0	0
0.25	0.625

The following graph displays the juxtaposition of the different statesmen of concerned countries with the outcomes. It expresses too the craze of HCQS (Hydroxy Chloro Quine Sulfate) implementation concerning the treatment of COVID-19. From this, we can sum up that the outcomes regarding polarity classification and subjectivity classification have a fruitful aspect for the implementation of HCQS for the treatment of COVID-19. We are hopeful to go on with more researches on the opinions or sentiments implementing distinct, upgraded methods.

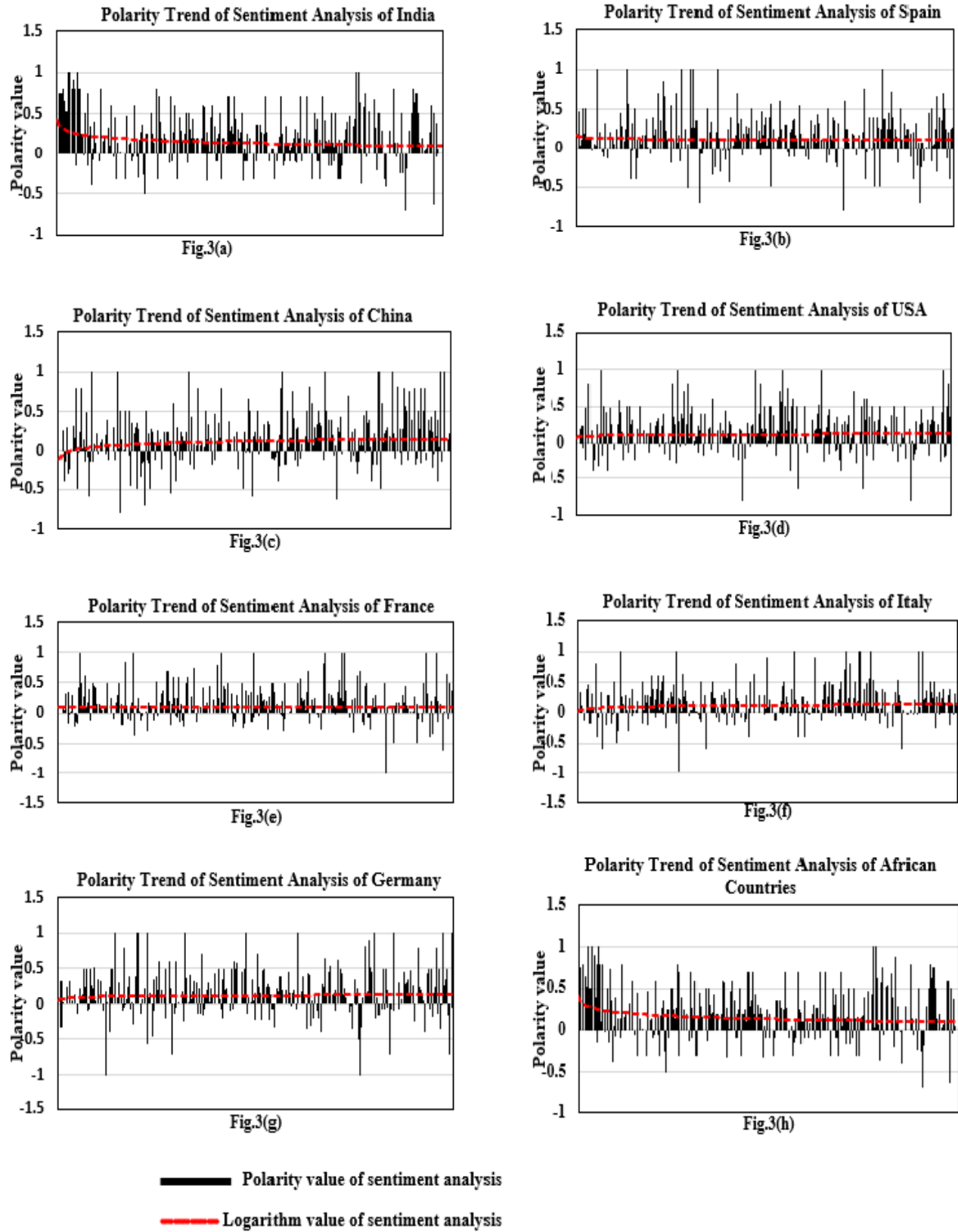


Figure 3: Polarity comparison of Tweeter sentiment analysis of political heads of different countries

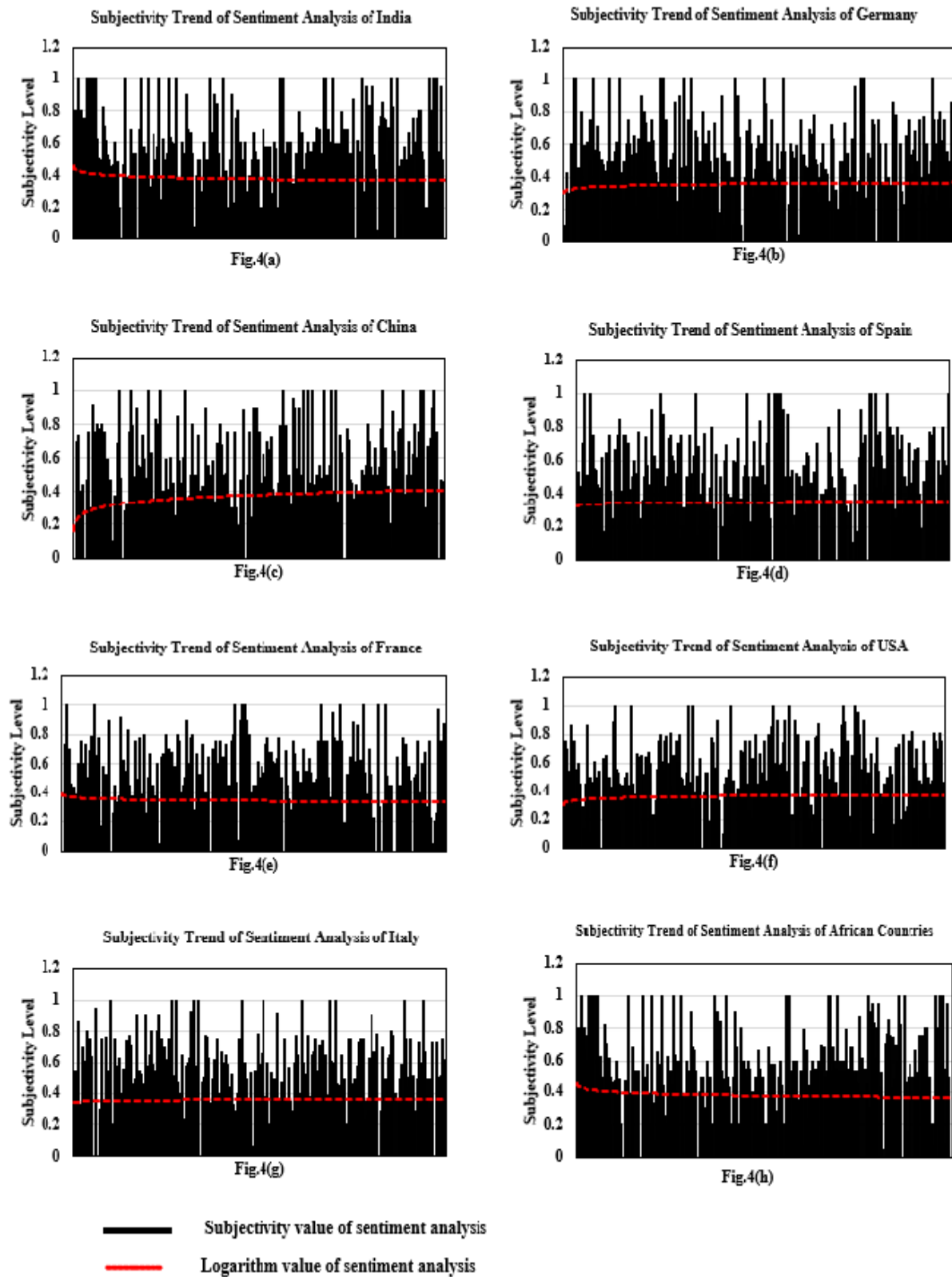


Figure 4: Subjectivity comparison of Tweeter sentiment analysis of political heads of different countries

8. CONCLUSION

This paper's objective is to correlate the distinct twitter comments of the statesmen of distinct countries for having a concrete knowledge on the application of drugs to the patients attacked by COVID-19. The prototype has dealt with the outlooks of the twitter's perspective of the specialists' interest to the businessmen inclined to invest in the drug market. For the evidence of the outcome, this correlative analysis is made. TextBlob approaches produce fruitful outcomes. The outcomes establish a solid relationship between twitter comments and at the apex of or downgrade sentiment polarity and opinion. Besides, analysis regarding this is executed to explore more modified outcomes in this sphere.

REFERENCES

1. S. Bergsma, M. Dredze, "Broadly improving user classification via communication-based name and location clustering on twitter," 2013.
2. C. Romero, S. Ventura, P. de Bra, and C. Castro, "Discovering prediction rules in aha! courses," Proceedings of the International Conference User Modelling, 25-34. https://doi.org/10.1007/3-540-44963-9_6
3. Vanessa Friedmann, "Clustering a customer base using twitter data," CS229, 2015.
4. Alec Go, Richa Bhayani, and Lei Huang, "Twitter sentiment classification using distant supervision," Technical report, Stanford, 2009.
5. Luciano Barbosa and Junlan Feng, "Robust sentiment detection on twitter from biased and noisy data," Proceedings of the 23rd International Conference on Computational Linguistics: Posters, pages 36-44, 2010.
6. Luiz F. S. Coletta, N'adia F. F. da Silva, Eduardo R. Hruschka, and Estevam Rand Hruschka Jr, "Combining classification and clustering for tweet sentiment analysis," Brazilian Conference on Intelligent Systems. 2014. <https://doi.org/10.1109/BRACIS.2014.46>
7. X. Phan, L. Nguyen, and S. Horiguchi, "Learning to classify short and sparse text & web with hidden topics from largescale data collections," In Proceeding of the 17th International conference on World Wide Web (WWW '08, ACM, New York, NY, USA, 91-100, 2008.
8. Kevin Gimpel, Nathan Schneider, Brendan O'Connor, Dipanjan Das, Daniel Mills, Jacob Eisenstein, Michael Heilman, Dani Yogatama, Jeffrey Flanigan, and Noah A. Smith, "Part-of-Speech tagging for twitter: annotation, features and experiments," CMU, Pittsburg, USA.
9. Manisha Rani and Jyoti Arora, "A review of data analysis of twitter," International Journal of Advanced Research in Computer Science and Software Engineering, Volume 6, Issue 5, May 2016 ISSN: 2277 128X.
10. Alexander Pak, Patrick Paroubek, "Twitter as a corpus for sentiment analysis and opinion mining".
11. P.-N.Tan, "Introduction to data mining," Pearson, 2006.
12. Svetlana Kiritchenko, Xiaodan Zhu and Saif M. Mohammad "Sentiment analysis of short informal texts," Journal of Artificial Intelligence Research 50 (2014) 723-762.
13. J. C. Dunn, "A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters," Journal of Cybernetics 1973, 3: 32-57.
14. Rish, Irina (2001). "An empirical study of the naive Bayes classifier". IJCAI Workshop on Empirical Methods in AI.
15. McCallum, Andrew. "Graphical Models, Lecture2: Bayesian Network Representation" (PDF). Retrieved 22 October 2019.
16. Maron, M. E. (1961). "Automatic Indexing: An Experimental Inquiry"(PDF). Journal of the ACM. 8 (3): 404-417. doi:10.1145/321075.321084.
17. Kang H., Yoo S.J. "Senti-Lexicon and improved Naïve Bayes algorithms for sentiment analysis of restaurant reviews".2011.
18. Pak, A., & Paroubek, P. 2010. "Twitter as a corpus for sentiment analysis and opinion mining". In LREc (Vol. 10, No. 2010).
19. Liang, P. W., & Dai, B. R. 2013. "Opinion mining on social media data. In Mobile Data Management (MDM)", 2013 IEEE 14th International Conference on (Vol. 2, pp. 91-96). IEEE.
20. J. B. MacQueen, "Some methods for classification and analysis of multivariate Observations," Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, University of California Press, 1:281297.
21. S. N. Sulaiman and N. A. M. Isa, "Adaptive fuzzy-K-means clustering algorithm for image segmentation," IEEE Transactions on Consumer Electronics, vol. 56, no. 4, pp. 2661-2668, 2010. <https://doi.org/10.1109/TCE.2010.5681154>
22. J. B. MacQueen, "Some methods for classification and analysis of multivariate observations," in Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability, pp. 281- 297, University of California Press, 1967.
23. R. Ebrahimipour, R. Rasoolinezhad, Z. Hajiabolhasani, and M. Ebrahimi, "Vanishing point detection in corridors: using Hough transform and K-means clustering," IET Computer Vision, vol. 6, no. 1, pp. 40-51, 2012.
24. P. S. Bishnu and V. Bhattacharjee, "Software fault prediction using quad tree-based K-means clustering algorithm," IEEE Transactions on Knowledge and Data Engineering, vol. 24, no. 6, pp. 1146-1150, 2012.
25. T. Elomaa and H. Koivisto, "On autonomous K-means clustering," in Proceedings of the International Symposium on Methodologies for Intelligent Systems, pp. 228-236, May 2005. https://doi.org/10.1007/11425274_24
26. J. M. Pena, J. A. Lozano, and P. Larrañaga, "An empirical comparison of four initialization methods for the

K-Means algorithm,” *Pattern Recognition Letters*, vol. 20, no. 10, pp. 1027– 1040, 1999.

26. T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu, “An efficient k-means clustering algorithms: analysis and implementation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 881– 892, 2002.

27. J. C. Dunn, “A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters,” *Journal of Cybernetics* 1973, 3: 32-57.

<https://doi.org/10.1080/01969727308546046>

28. Babuska R (2014) Fuzzy clustering lecture. <http://homes.di.unimi.it/valenti/SlideCorsi/Bioinformatica05/Fuzzy-Clusteringlecture-Babuska>. Accessed Oct 2014

29. Bezdek JC (1981) “Pattern recognition with fuzzy objective function algorithms”. Kluwer Academic Publishers, Norwell

30. I. S. Amiri, O. A. Akanbi, and E. Fazeldelhordi, “A Machine-learning Approach to Phishing Detection and Defense”. Syngress, 2014.

31. Kuhlman, Dave. "A Python Book: Beginning Python, Advanced Python, and Python Exercises". Section 1.1. Archived from the original (PDF) on 23 June 2012

32. "Open Source Community". Anaconda. Retrieved 2019-02-27.

33. "Conda – Conda documentation". Retrieved February 25, 2016.

34. "Anaconda Navigator". docs.anaconda.com.

35. Lenzer J. “Covid-19: US gives emergency approval to hydroxychloroquine de-spite lack of evidence”. *BMJ* 2020;369:m1335. <https://doi.org/10.1136/bmj.m1335>.

36. Indian Council for Medical Research. Recommendation for empiric use of hydroxychloroquine for prophylaxis of SARS-CoV-2 infection. <https://www.icmr.gov.in/index.html> Accessed 9 April 2020.

37. Gao J, Tian Z, Yang X. “Breakthrough: chloroquine phosphate has shown apparent efficacy in treatment of COVID-19 associated pneumonia in clinical studies”. *Biosci Trends* 2020; 14:72–3.

<https://doi.org/10.5582/st.2020.01047>.

38. Gautret P, Lagier JC, Parola P, Hoang VT, Meddeb L, Mailhe M, et al. “Hydroxychloroquine and azithromycin as a treatment of COVID-19: results of an open-label non-randomized clinical trial”. *Int J Antimicrob Agents* 2020;105949.

<https://doi.org/10.1016/j.ijantimicag.2020.105949>.

39. Chen Z, Hu J, Zhang Z, Jiang S, Han S, Yan D, et al. “Efficacy of hydroxychloroquine in patients with COVID-19: results of a randomized clinical trial”. *medRxiv* 2020:20040758.

<https://doi.org/10.1101/2020.03.22.20040758>.

40. Munster T, Gibbs JP, Shen D, Baethge BA, Botstein GR, Caldwell J, et al. “Hydroxychloroquine concentration–response relationships in patients with rheumatoid arthritis”. *Arthritis Rheum* 2002; 46:1460–9.

<https://doi.org/10.1002/art.10307>.

41. Makin AJ, Wendon J, Fitt S, Portmann BC, Williams R. “Fulminant hepatic failure secondary to hydroxychloroquine”. *Gut* 1994; 35:569–70.

<https://doi.org/10.1136/gut.35.4.569>.

42. Murphy M, Carmichael AJ. “Fatal toxic epidermal necrolysis associated with hydroxychloroquine”. *Clin Exp Dermatol* 2001; 26:457–8. <https://doi.org/10.1046/j.1365-2230.2001.00857-3.x>.

43. Chen CY, Wang FL, Lin CC. “Chronic hydroxychloroquine use associated with QT prolongation and refractory ventricular arrhythmia”. *Clin Toxicol (Phila)* 2006; 44:173–5.

<https://doi.org/10.1080/15563650500514558>.

44. Stas P, Faes D, Noyens P. “Conduction disorder and QT prolongation secondary to long-term treatment with chloroquine”. *Int J Cardiol* 2008;127: e80–2. <https://doi.org/10.1016/j.ijcard.2007.04.055>.

45. Chorin E, Dai M, Shulman E, Wadhwani L, Bar Cohen R, Barbhuiya C, et al. “The QT interval in patients with SARS-CoV-2 infection treated with hydroxychloroquine/azithromycin”. *medRxiv* 2020.

<https://doi.org/10.1101/2020.04.02.20047050>.