**Quantifying COVID-19 Content in the Online**

**Health Opinion War Using Machine Learning**

### A mini project report submitted in partial fulfilment of the requirements for the award of degree of

**BACHELOR OF TECHNOLOGY IN**

**COMPUTER SCIENCE AND ENGINEERING**

**Submitted by**

**Mr. K. HANUMANTH RAO**

### Asst. Professor



**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING GITAM**

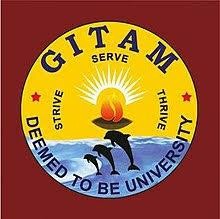
**(Deemed to be University) HYDERABAD CAMPUS APRIL - 2020**

## DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING GITAM SCHOOL OF TECHNOLOGY

**GITAM**

**(Deemed to be University)**

### HYDERABAD CAMPUS



**DECLARATION**

We, hereby declare that the major project report entitled **“SERENDIPITY – A MACHINE LEARNING APPLICATION FOR MINING SERENDIPITOUS DRUG USAGE**

**FROM SOCIAL MEDIA”** is an original work done in the Department of Computer Science and Engineering, GITAM School of Technology, GITAM (Deemed to be University) submitted in partial fulfillment of the requirements for the award of the degree of “Bachelor of Technology” in Computer Science and Engineering. The work had not submitted to any other college or University for the award of any degree or diploma.

Date:

### Registration No(s) Name(s) Signature(s) 2210316359 V. SRI KRISHNA CHAITANYA

**2210316338 G. NITESH KUMAR**

**2210316303 A. SANJAY**

**2210316366 M. NIKHIL NETHAJI REDDY**

**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING GITAM SCHOOL OF TECHNOLOGY**

**GITAM**

**(Deemed to be University) HYDERABAD CAMPUS**



**CERTIFICATE**

This is to certify that the project report entitled **“SERENDIPITY – A MACHINE LEARNING APPLICATION FOR MINING SERENDIPITOUS DRUG USAGE**

**FROM SOCIAL MEDIA”** is a bonafide record of work carried out by **V SRI KRISHNA CHAITANYA (2210316359), GADDALAY NITESH KUMAR (2210316338)**, **ANUMANDLA SANJAY (2210316303), NIKHIL NETHAJI REDDY MALLEPALLY**

**(2210316366)** submitted in partial fulfilment of the requirements for the award of degree of Bachelor of Technology in Computer Science and Engineering.

### Project Guide Head of the Department

**Mr. K. Hanumanth Rao Dr. S. Phani Kumar**

**Assistant Professor Professor**

**ACKNOWLEDGMENT**

Our project would not have been successful without the help of several people. We would like to thank the personalities who were part of our project in numerous ways, those who gave us outstanding support from the birth of the project.

We are extremely thankful to our honorable Pro-Vice Chancellor, Prof. **N. Siva Prasad** for providing necessary infrastructure and resources for the accomplishment of our project.

We are highly indebted to **Prof. N. Seetharamaiah**, Principal, School of Technology, for his support during the tenure of the project.

We are very much obliged to our beloved **Prof. S. Phani Kumar**, Head of the Department of Computer Science & Engineering for providing the opportunity to undertake this project and encouragement in completion of this project.

We hereby wish to express our deep sense of gratitude to **Mr. K. Hanumanth Rao, Asst. Professor**, Department of Computer Science and Engineering, School of Technology for the esteemed guidance, moral support and invaluable advice provided by him for the success of the project.

We are also thankful to all the staff members of Computer Science and Engineering department who have co-operated in making our project a success. We would like to thank all our parents and friends who extended their help, encouragement and moral support either directly or indirectly in our project work.

Sincerely,

# ABSTRACT

A huge amount of potentially dangerous COVID-19 misinformation is appearing online. Here we use machine learning to quantify COVID-19 content among online opponents of establishment health guidance, in particular vaccinations (``anti-vax''). We \_nd that the anti-vax community is developing a less focused debate around COVID-19 than its counterpart, the pro-vaccination (``pro-vax'') community. However, the anti-vax community exhibits a broader range of ``\_avors'' of COVID-19 topics, and hence can appeal to a broader cross-section of individuals seeking COVID-19 guidance online, e.g. individuals wary of a mandatory fast-tracked COVID-19 vaccine or those seeking alternative remedies. Hence the anti-vax community looks better positioned to attract fresh support going forward than the pro-vax community. This is concerning since a widespread lack of adoption of a COVID-19 vaccine will mean the world falls short of providing herd immunity, leaving countries open to future COVID-19 resurgences.We provide a mechanistic model that interprets these results and could help in assessing the likely ef\_cacy of intervention strategies. Our approach is scalable and hence tackles the urgent problem facing social media platforms of having to analyze huge volumes of online health misinformation and disinformation.

**TABLE OF CONTENTS**

1.[INTRODUCTION 1](file:///D:\College%20Wise%20Data\2020\Manoj%20Client%20Hyd\GITAM%202020%20MINI\MINI%20FINAL%20DOC'S%20MODIFIED\A%20Driving%20Decision%20Strategy(DDS)%20Based%20on%20Machine%20learning%20for%20an%20autonomous%20vehicle.docx#_TOC_250018)

* 1. [Motivation 1](file:///D:\College%20Wise%20Data\2020\Manoj%20Client%20Hyd\GITAM%202020%20MINI\MINI%20FINAL%20DOC'S%20MODIFIED\A%20Driving%20Decision%20Strategy(DDS)%20Based%20on%20Machine%20learning%20for%20an%20autonomous%20vehicle.docx#_TOC_250017)
  2. Existing System 2
  3. Objective 2
  4. Outcomes 2
  5. Applications 3
  6. [Structure of Project (System Analysis)](file:///D:\College%20Wise%20Data\2020\Manoj%20Client%20Hyd\GITAM%202020%20MINI\MINI%20FINAL%20DOC'S%20MODIFIED\A%20Driving%20Decision%20Strategy(DDS)%20Based%20on%20Machine%20learning%20for%20an%20autonomous%20vehicle.docx#_TOC_250016) 3
     1. [Requisites Accumulating and Analysis](file:///D:\College%20Wise%20Data\2020\Manoj%20Client%20Hyd\GITAM%202020%20MINI\MINI%20FINAL%20DOC'S%20MODIFIED\A%20Driving%20Decision%20Strategy(DDS)%20Based%20on%20Machine%20learning%20for%20an%20autonomous%20vehicle.docx#_TOC_250015) 4
     2. [System Design](file:///D:\College%20Wise%20Data\2020\Manoj%20Client%20Hyd\GITAM%202020%20MINI\MINI%20FINAL%20DOC'S%20MODIFIED\A%20Driving%20Decision%20Strategy(DDS)%20Based%20on%20Machine%20learning%20for%20an%20autonomous%20vehicle.docx#_TOC_250014) 4
     3. [Implementation](file:///D:\College%20Wise%20Data\2020\Manoj%20Client%20Hyd\GITAM%202020%20MINI\MINI%20FINAL%20DOC'S%20MODIFIED\A%20Driving%20Decision%20Strategy(DDS)%20Based%20on%20Machine%20learning%20for%20an%20autonomous%20vehicle.docx#_TOC_250013) 4
     4. Testing 4
     5. [Deployment of System and Maintenance 4](file:///D:\College%20Wise%20Data\2020\Manoj%20Client%20Hyd\GITAM%202020%20MINI\MINI%20FINAL%20DOC'S%20MODIFIED\A%20Driving%20Decision%20Strategy(DDS)%20Based%20on%20Machine%20learning%20for%20an%20autonomous%20vehicle.docx#_TOC_250012)
  7. [Functional Requirements 4](file:///D:\College%20Wise%20Data\2020\Manoj%20Client%20Hyd\GITAM%202020%20MINI\MINI%20FINAL%20DOC'S%20MODIFIED\A%20Driving%20Decision%20Strategy(DDS)%20Based%20on%20Machine%20learning%20for%20an%20autonomous%20vehicle.docx#_TOC_250011)
  8. [Non-Functional Requirements 5](file:///D:\College%20Wise%20Data\2020\Manoj%20Client%20Hyd\GITAM%202020%20MINI\MINI%20FINAL%20DOC'S%20MODIFIED\A%20Driving%20Decision%20Strategy(DDS)%20Based%20on%20Machine%20learning%20for%20an%20autonomous%20vehicle.docx#_TOC_250010)
     1. [Examples of Non-Functional Requirements](file:///D:\College%20Wise%20Data\2020\Manoj%20Client%20Hyd\GITAM%202020%20MINI\MINI%20FINAL%20DOC'S%20MODIFIED\A%20Driving%20Decision%20Strategy(DDS)%20Based%20on%20Machine%20learning%20for%20an%20autonomous%20vehicle.docx#_TOC_250009) 6
     2. Advantages of Non-Functional Requirements 6
     3. Disadvantages Non-Functional Requirements 6
     4. Key Learnings 7

2.[LITERATURE SURVEY 7](file:///D:\College%20Wise%20Data\2020\Manoj%20Client%20Hyd\GITAM%202020%20MINI\MINI%20FINAL%20DOC'S%20MODIFIED\A%20Driving%20Decision%20Strategy(DDS)%20Based%20on%20Machine%20learning%20for%20an%20autonomous%20vehicle.docx#_TOC_250008)

3.[PROBLEM IDENTIFICATION & OBJECTIVES 1](file:///D:\College%20Wise%20Data\2020\Manoj%20Client%20Hyd\GITAM%202020%20MINI\MINI%20FINAL%20DOC'S%20MODIFIED\A%20Driving%20Decision%20Strategy(DDS)%20Based%20on%20Machine%20learning%20for%20an%20autonomous%20vehicle.docx#_TOC_250007)8

* 1. [Existing Approach 1](file:///D:\College%20Wise%20Data\2020\Manoj%20Client%20Hyd\GITAM%202020%20MINI\MINI%20FINAL%20DOC'S%20MODIFIED\A%20Driving%20Decision%20Strategy(DDS)%20Based%20on%20Machine%20learning%20for%20an%20autonomous%20vehicle.docx#_TOC_250006)8

3.2 [Proposed System 1](file:///D:\College%20Wise%20Data\2020\Manoj%20Client%20Hyd\GITAM%202020%20MINI\MINI%20FINAL%20DOC'S%20MODIFIED\A%20Driving%20Decision%20Strategy(DDS)%20Based%20on%20Machine%20learning%20for%20an%20autonomous%20vehicle.docx#_TOC_250005)8

3.3 Modules 19

3.4 Algorithms 19

4.SYSTEM DESIGN 22

5.[IMPLEMENTATION 34](file:///D:\College%20Wise%20Data\2020\Manoj%20Client%20Hyd\GITAM%202020%20MINI\MINI%20FINAL%20DOC'S%20MODIFIED\A%20Driving%20Decision%20Strategy(DDS)%20Based%20on%20Machine%20learning%20for%20an%20autonomous%20vehicle.docx#_TOC_250002)

* 1. Flowchart 34
  2. Code 34

6.TESTING 45

7.[RESULTS AND DISCUSSIONS](file:///D:\College%20Wise%20Data\2020\Manoj%20Client%20Hyd\GITAM%202020%20MINI\MINI%20FINAL%20DOC'S%20MODIFIED\A%20Driving%20Decision%20Strategy(DDS)%20Based%20on%20Machine%20learning%20for%20an%20autonomous%20vehicle.docx#_TOC_250001) 53

8.CONCLUSION AND FUTURE SCOPE 59

9[REFERENCES](file:///D:\College%20Wise%20Data\2020\Manoj%20Client%20Hyd\GITAM%202020%20MINI\MINI%20FINAL%20DOC'S%20MODIFIED\A%20Driving%20Decision%20Strategy(DDS)%20Based%20on%20Machine%20learning%20for%20an%20autonomous%20vehicle.docx#_TOC_250000) 60

**LIST OF FIGURES**

|  |  |  |
| --- | --- | --- |
| S.NO | NAME | P.NO |
| 1 | Project SDLC | 3 |
| 2 | Use case diagram | 31 |
| 3 | Class diagram | 32 |
| 4 | Sequence diagram | 33 |

# 1.INTRODUCTION

Scientific experts agree that defeating COVID-19 will depend on developing a vaccine. However, this assumes that a sufficiently large proportion of people would receive a vaccine so that herd immunity is achieved. Because vaccines tend to be less effective in older people, this will require younger generations to have very high COVID-19 vaccination rates in order to guarantee herd immunity [1]. Yet there is already signi\_cant opposition to existing vaccinations, e.g. against measles, with some parents already refusing to vaccinate their children. Such vaccine opposition increased the number of cases in the 2019 measles outbreak in the U.S. and beyond [2]. Any future COVID-19 vaccine will likely face similar opposition [3], [4].

**1.1** **MOTIVATION**

Online social media platforms, and in particular the builtin communities that platforms like Facebook (FB) feature, have become popular fora for vaccine opponents (anti-vax) to congregate and share health (mis)information. Such misinformation can endanger public health and individual safety[1], [4]. Likewise, vaccine supporters (pro-vax) also congregate in such online communities to discuss and advocate for professional public health guidance.Well before COVID-19, there was already an intense online con\_ict featuring antivax communities and pro-vax communities. Within anti-vax communities, the narratives typically draw on and generate misinformation about establishment medical guidance and distrust of the government, pharmaceutical industry, and new technologies such as 5G communications [1], [4], [5]..

* 1. **Existing System**

The existing methods using Most models assume a standard SEIR structure.

Fraser and collegues to estimate size but make different changes on the nature of the different compartments and their respective residence times

**1.2.1 Limitations of existing system**

.

* 1. **Objectives**

The anti-vax community exhibits a broader range of “flavors” of COVID-19 topics, and hence can appeal to a broader cross-section of individuals seeking COVID-19 guidance online, e.g. individuals wary of a mandatory fast-tracked COVID-19 vaccine or those seeking alternative remedies. Hence the anti-vax community looks better positioned to attract fresh support going forward than the pro-vax community. This is concerning since a widespread lack of adoption of a COVID-19 vaccine will mean the world falls short of providing herd immunity, leaving countries open to future COVID-19 resurgences.

**1.4 Outcomes**

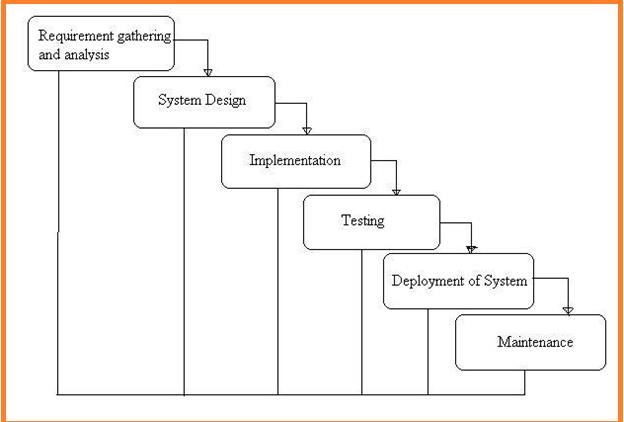
Overall, this approach shows that a machine-learning algorithm, the LDA algorithm, identi\_es plausible topics within collections of posts from online communities surrounding the vaccine and COVID-19 debate. In addition to being able to handle large quantities of data, its results emerge

quickly using statistical grouping techniques, instead of having to rely on potentially biased, slow and costly human labeling..

**1.5Applications**

This strategy used in

### STRUCTURE OF PROJECT (SYSTEM ANALYSIS)



**Fig: 1 Project SDLC**

* Project Requisites Accumulating and Analysis
* Application System Design
* Practical Implementation
* Manual Testing of My Application
* Application Deployment of System
* Maintenance of the Project

### REQUISITES ACCUMULATING AND ANALYSIS

It’s the first and foremost stage of the any project as our is a an academic leave for requisites amassing we followed of IEEE Journals and Amassed so many IEEE Relegated papers and final culled a Paper designated “Individual web revisitation by setting and substance importance input and for analysis stage we took referees from the paper and did literature survey of some papers and amassed all the Requisites of the project in this stage

### SYSTEM DESIGN

In System Design has divided into three types like GUI Designing, UML Designing with avails in development of project in facile way with different actor and its utilizer case by utilizer case diagram, flow of the project utilizing sequence, Class diagram gives information about different class in the project with methods that have to be utilized in the project if comes to our project our UML Will utilizable in this way The third and post import for the project in system design is Data base design where we endeavor to design data base predicated on the number of modules in our project

### IMPLEMENTATION

The Implementation is Phase where we endeavor to give the practical output of the work done in designing stage and most of Coding in Business logic lay coms into action in this stage its main and crucial part of the project

### 1.6.4TESTING UNIT TESTING

It is done by the developer itself in every stage of the project and fine-tuning the bug and module predicated additionally done by the developer only here we are going to solve all the runtime errors

### MANUAL TESTING

As our Project is academic Leave, we can do any automatic testing so we follow manual testing by endeavor and error methods

### DEPLOYMENT OF SYSTEM AND MAINTENANCE

Once the project is total yare, we will come to deployment of client system in genuinely world as its academic leave we did deployment i our college lab only with all need Software’s with having Windows OS .

The Maintenance of our Project is one-time process only

### FUNCTIONAL REQUIREMENTS

1.Data Collection

2.Data Preprocessing

3.Training And Testing

4.Modiling

5.Predicting

### NON FUNCTIONAL REQUIREMENTS

NON-FUNCTIONAL REQUIREMENT (NFR) specifies the quality attribute of a software system. They judge the software system based on Responsiveness, Usability, Security, Portability and other non-functional standards that are critical to the success of the software system. Example of nonfunctional requirement, *“how fast does the website load?”* Failing to meet non-functional requirements can result in systems that fail to satisfy user needs. Non- functional Requirements allows you to impose constraints or restrictions on the design of the system across the various agile backlogs. Example, the site should load in 3 seconds when the number of simultaneous users are > 10000. Description of non-functional requirements is just as critical as a functional requirement.

* Usability requirement
* Serviceability requirement
* Manageability requirement
* Recoverability requirement
* Security requirement
* Data Integrity requirement
* Capacity requirement
* Availability requirement
* Scalability requirement
* Interoperability requirement
* Reliability requirement
* Maintainability requirement
* Regulatory requirement
* Environmental requirement

### EXAMPLES OF NON-FUNCTIONAL REQUIREMENTS

Here, are some examples of non-functional requirement:

* + - 1. Users must upload dataset
      2. The software should be portable. So moving from one OS to other OS does not create any problem.
      3. Privacy of information, the export of restricted technologies, intellectual property rights, etc. should be audited.
    1. **ADVANTAGES OF NON-FUNCTIONAL REQUIREMENT**

Benefits/pros of Non-functional testing are:

* The nonfunctional requirements ensure the software system follow legal and compliance rules.
* They ensure the reliability, availability, and performance of the software system
* They ensure good user experience and ease of operating the software.
* They help in formulating security policy of the software system.
  + 1. **DISADVANTAGES OF NON-FUNCTIONAL REQUIREMENT**

Cons/drawbacks of Non-function requirement are:

* None functional requirement may affect the various high-level software subsystem
* They require special consideration during the software architecture/high-level design phase which increases costs.
* Their implementation does not usually map to the specific software sub-system,
* It is tough to modify non-functional once you pass the architecture phase.
  + 1. **KEY LEARNING**

# The character of the time period, the length of road, the weather, the bus speed and the rate of road usage are adopted as input vectors in Support Vector Machine

# 2.LITERATURE SURVEY

# 

**A. Kata, “A postmodern Pandora’s box: Antivaccination misinformation on the Internet,” Vaccine, vol. 28, no. 7, pp. 1709–1716, Feb. 2010, doi: 10.1016/j.vaccine.2009.12.022.**

The Internet plays a large role in disseminating anti-vaccination information. This paper builds upon previous research by analyzing the arguments proffered on anti-vaccination websites, determining the extent of misinformation present, and examining discourses used to support vaccine objections. Arguments around the themes of safety and effectiveness, alternative medicine, civil liberties, conspiracy theories, and morality were found on the majority of websites analyzed; misinformation was also prevalent. Themost commonly proposedmethod of combating thismisinformation is through better education, although this has proven ineffective. Education does not consider the discourses supporting vaccine rejection, such as those involving alternative explanatory models of health, interpretations of parental responsibility, and distrust of expertise. Anti-vaccination protestors make postmodern arguments that reject biomedical and scientific “facts” in favour of their own interpretations. Pro-vaccination advocates who focus on correcting misinformation reduce the controversy to merely an “educational” problem; rather, these postmodern discourses must be acknowledged in order to begin a dialogue. With morbidity and mortality from vaccine-preventable diseases [VPDs] having reached record lows [1], vaccines are one of the most successful tools for biomedical science and public health. Yet paradoxically, the effectiveness of vaccination has led to the reemergence of anti-vaccination sentiments. Vaccines may be seen as unnecessary or dangerous because incidence rates of VPDs in developed countries have plummeted. Vaccine “reactions” – negative health events following vaccination, attributed to the vaccine – then appear to be more common than the diseases themselves [2]. In this way, vaccines can be considered victims of their own success. The media plays a large role in disseminating and sensationalizing vaccine objections. Such objections are part of what has been called the “anti-vaccination movement”, which has had a demonstrable impact on vaccination policies, and individual and community health [3]. A common sequence to vaccination scares involves scientific debate about potential vaccine risks, which communication technology transmits via a rhetoric of doubt; parents incorporate this with personal experiences and spread their views to their social groups [4]. These social groups exert considerable pressure on vaccination decisions by creating a “local vaccination culture” [5]. With the prominence of the Internet in today’s world, ∗ Correspondence address: 110 Parkwood Cres., Hamilton, Ontario L8V 4Z7, Canada. Tel.: +1 905 387 3141. E-mail addresses: aniakata@gmail.com, kataa@mcmaster.ca. the attitudes, beliefs, and experiences of that local culture can quickly become global. Internet usage statistics show approximately 74% of Americans and 72% of Canadians are online [6]. An estimated 75–80% of users search for health information online [7]. Of these users, 70% say the information they encounter online influences their treatment decisions [8]. In 2006, 16% of users searched online for information on immunizations or vaccinations [9]. While online research is more convenient and accessible than reading medical literature or visiting health practitioners, too great a reliance on Internet-based information can be problematic. Over half (52%) of users believe “almost all” or “most” information on health websites is credible [8]; yet the availability of inaccurate and deceptive information online has labelled the Internet a “modern Pandora’s box” [10]. The nature of the Internet allows any and all opinions to spread widely and instantaneously. Individuals and groups gain exposure online without being filtered or reviewed – and anti-vaccination advocates have taken advantage of this fact. Anti-vaccination messages are more common on the Internet than in other forms of media, increasing the likelihood that vaccination decisions may be based on misleading information [11]. Indeed, parents who exempt children from vaccination are more likely to have obtained information from the Internet than parents who have their children vaccinated; they are also more likely to have used certain antivaccination websites [12]. This demonstrates the importance of understanding what messages are presented online and why they may be accepted. The body of research examining online anti-vaccinationism is not large, nor has there been a recent update [11,13–18]. Onlyone analysis [13] examined misinformation and deception on such sites, but was not quantitative. Prior research also acknowledged the need to understand discourses underlying anti-vaccination arguments [19,20], but did not elaborate upon them. This analysis aims to address these issues by answering two main questions. First, what information is proffered on anti-vaccination websites, and what is its accuracy? Second, what discourses make these vaccine objections appealing? 2. Methods 2.1. Data collection Web searches were conducted on May 21, 2009 using the terms “vaccine”, “vaccination”, and “immunization OR immunisation” input into Google.com (the American version of the search engine) and Google.ca (the Canadian version). Google was chosen as it is the most popular search engine, accounting for 73% of all Internet searches [21]. Results were classified as anti-vaccination and included for content analysis if they opposed childhood vaccination for any reason, without meeting any of the following exclusion criteria: (1) listserv or newsgroup pages; (2) pages solely containing brief notices about other website content; (3) news results, medical journals or library sites; (4) video results; (5) book previews; (6) non-English sites; (7) sites exclusively about adult immunization; (8) sites exclusively about veterinary vaccination and (9) inactive links. Criteria (see Tables 1 and 2) were applied to the anti-vaccination websites and coded as present or absent. Criteria were adapted from previous online antivaccination studies [11,13,14,17,18], as well as created by the author. Online health information seekers examine the first 10 search results 97.2% of the time [22]; therefore, only the first 10 results retrieved per term were examined. Of 30 total Google.com results, 5 of 21 immunization sites (24%) were classified as anti-vaccination. Of 30 total Google.ca results, 2 of 16 immunization sites (13%) were classified as anti-vaccination. To amass additional websites for a more meaningful study, the Canadian searches were extended to 50 results per term. Of 150 total results, 5 of 86 immunization sites (6%) were classified as anti-vaccination (two were duplicates of American results). Combining the American and Canadian results, 8 anti-vaccination websites were subjected to content analysis. Appendix A lists the sites analyzed. The proportions of pro- and anti-vaccination sites found per search term are illustrated in Fig. 1. Overall, American searches returned more anti-vaccination results (24%) than Canadian searches (6%), indicating American parents are more likely to encounter anti-vaccination sites via Google than are Canadian parents. Neither search engine returned any anti-vaccination results for “immunization OR immunisation”; this was expected based on research that found anti-vaccination groups avoid using the term “immunization” as they tend not to believe that vaccines confer immunity [16]. Although prior studies returned more search results [11], this does not necessarily mean the number of anti-vaccination websites has decreased, but rather that their search rankings may have shifted. Nevertheless, the proportion of sites retrieved for some search terms is notable – 71% of results from the Google.com “vaccination” search were classified as antivaccination. 3.2. Content and themes Fig. 2 illustrates the percentage of analyzed sites with the listed themes present. Individual content criteria are quantified inTable 1. Individual design criteria are quantified in Table 2. 3.2.1. Safety and effectiveness “Vaccines are biological poisons, harmful to health, and a contributing factor in childhood illness.” (http://www. vaclib.org/sites/debate/about.html) Safety themes were present on all anti-vaccination websites analyzed. Every site claimed vaccines are poisonous and cause idiopathic illnesses. Sites stressed that vaccines contain substances poisonous to humans, including anti-freeze, ether, formaldehyde, mercury, and nanobacteria. Pertinent information was not elaborated upon – for instance, that the amount of potentially harmful substances in vaccines is not enough to produce toxic effects in humans, or that ether does not refer to the anaesthetic but to a chemical compound. Illnesses attributed to vaccines included:

**B. Martin, “Texas anti-vaxxers fear mandatory COVID-19 vaccines more than the virus itself,” Texas Monthly, 2020. [Online]. Available: https://www.texasmonthly.com/news/texas-antivaxxers-fear-mandatory-coronavirus-vaccines**

Background

The World Health Organization lists vaccine hesitancy as one of 10 threats to global health. The antivaccine movement uses Facebook to promote messages on the alleged dangers and consequences of vaccinating, leading to a reluctance to immunize against preventable communicable diseases.

Objective

We would like to know more about the messages these websites are sharing via social media that can influence readers and consumers. What messages is the public receiving on Facebook about immunization? What content (news articles, testimonials, videos, scientific studies) is being promoted?

Methods

We proposed using a social media audit tool and 3 categorical lists to capture information on websites and posts, respectively. The keywords “vaccine,” “vaccine truth,” and “anti-vax” were entered in the Facebook search bar. A Facebook page was examined if it had between 2500 and 150,000 likes. Data about beliefs, calls to action, and testimonials were recorded from posts and listed under the categories Myths, Truths, and Consequences. Website data were entered in a social media audit template.

Results

Users’ posts reflected fear and vaccine hesitancy resulting from the alleged dangers of immunization featured on the website links. Vaccines were blamed for afflictions such as autism, cancer, and infertility. Mothers shared testimonies on alleged consequences their children suffered due to immunization, which have influenced other parents to not vaccinate their children. Users denied the current measles outbreaks in the United States to be true, retaliating against the government in protests for fabricating news.

Conclusions

Some Facebook messages encourage prevailing myths about the safety and consequences of vaccines and likely contribute to parents’ vaccine hesitancy. Deeply concerning is the mistrust social media has the potential to cast upon the relationship between health care providers and the public. A grasp of common misconceptions can help support health care provider practiceMany diseases have been almost, or completely, eradicated due to immunization. Immunization against disease prevents 2-3 million deaths per year internationally and could prevent even more with global vaccination improvements [[1](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref1)]. Immunization has vastly decreased mortality due to preventable communicable diseases. For example, before the introduction of the measles vaccine, 300,000-400,000 Canadians were infected every year, with some recoveries and many deaths [[2](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref2)]. Since the elimination of measles in 1998 due to vaccines, there have been very few cases in Canada [[2](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref2)]. Similarly, once the polio vaccine was introduced in Canada in the 1950s, cases reduced dramatically, and the current risk to the Canadian population is extremely low [[3](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref3)].

The World Health Organization (WHO) has declared vaccine hesitancy as one of the top 10 threats to global health [[4](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref4)]. Social media has helped fuel the growth of the antivaccine movement, with Facebook being identified as a key disseminator of misinformation surrounding the campaign [[5](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref5)-[7](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref7)]. Facebook is the largest social media platform, with more than 2 billion active monthly users [[8](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref8)]. There have been serious efforts to reduce the amount of misinformation spread on the social media site by lowering the ranking of Groups and Pages making false claims [[7](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref7)]. Social media administrators have been urged to remove these Pages and Groups altogether; however, counterarguments cite a violation of human rights to access uncensored information [[7](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref7)]. This paper exposes the messages of the antivaccine movement online and how individuals perceive immunization. We aimed to uncover the myths and truths that users of Facebook Pages observe and partake in. Health care consumers and health care providers may find themselves on opposite ends of the debate. Lack of immunization places the public at risk and decreases public health efforts to curb measles and polio and prevent outbreaks of influenza (flu) along with other communicable diseases. The shift in power between doctors and patients due to easy access to information online has led to the questioning of health care providers and increased shared decision making [[6](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref6),[9](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref9)].

As most of the world awaits a vaccine to put an end to the COVID-19 (also known as the 2019 novel coronavirus) pandemic, “followers” of antivaccine Facebook Pages seem to fear the vaccine more than the virus itself [[10](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref10)]. Amid the COVID-19 pandemic, social media sites such as Facebook are unable to control the health misinformation that is spread on its Pages [[11](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref11)]. Antivaccine Pages have been providing conspiracy theories, safety concerns, and alternative health medication that grasp the attention of “undecided” individuals surfing the web for information on vaccines [[11](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref11)]. The WHO is fighting to stop the spread of misinformation online by collaborating with social media giants to find a way to regulate false claims [[12](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref12)]. Some examples of such claims include that COVID-19 is a bioweapon funded by the Bill & Melinda Gates Foundation or that it can simply be cured by consuming homemade concoctions (some include drinking bleach) [[13](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref13)]. Our aim was to know more about the messages that can influence readers and consumers that these websites are sharing via social media.

[Go to:](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/)

## Methods

Publicly available content on 4 Facebook Pages was analyzed based on the number of likes they received. Keywords “anti-vax,” “anti-vaccine,” “vaccine,” “vaccine injury,” and “stop vaccination” were entered into the Facebook search bar. Once on the “results” page, we followed the link to the “Pages” tab. Pages were chosen if they had between 2500 and 150,000 likes — a measure of the spread of readership. This range was selected based on the fact that it included most pages that had high traffic with daily activity. The Page was selected if it had the highest amount of likes on the first “results” page. We then scrolled down to January 1, 2019 and analyzed posts, comments on posts, and website links shared until May 30, 2019. Flu activity peaks between December and February and can last as late as May [[14](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref14)]. Website and posts data simultaneously reached saturation, the point where no new themes emerged. Website links shared on the Facebook Pages were publicly accessible and consisted of news articles, blog posts, scientific studies, or website posts of renowned antivaccine activists. The data collected from the Facebook posts were categorized into “Myths,” “Truths,” and “Consequences.” These lists helped categorize the data found on the Facebook Pages to determine the exaggeration of myths and falsehoods and the minimization of truths. A separate category, “Measles Outbreak Reactions,” was used to document reactions to outbreaks of measles happening around the United States that were garnering attention in mainstream media and on the Facebook Pages.

Website data captured from links shared between January 2019 and May 2019 were entered into the Who, What, When, and Why categories of the social media audit template. Using this tool, we were able to capture and categorize data in a uniform manner for all websites. The social media audit template was created by Keith Quesenberry [[15](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref15)]. He describes a social media audit as “a systematic examination of social media data” [[15](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref15)]. We adapted and modified this tool to help us gain insight on what points and messages website authors are trying to get across. This tool is to be used to systemically examine “social talk of a brand” — in this case, the “brand” is the topic of immunization [[16](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref16)] — and allows the examiner to shift their viewpoint from “control” to “engagement” and understand why users are participating in such forums by examining specific content and posts [[16](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref16)]. The “Who” category captured the type of website (eg, blog post, news article) and the URL, which helped us determine the type of websites that were being shared. The “What” category was used to describe what content the website was sharing. This category was crucial in helping us determine the messages of the website. “When” noted the date the website content was published to determine whether links are being shared on the Facebook Page instantaneously or randomly. In the “Why” category, we noted any comments or statements made by the Facebook Page when sharing the linked website. Noting these statements in this section helped give us a better idea of the purpose behind sharing these websites and what the administrators of the page hoped to achieve by sharing these links with their audience. “Opportunity” was a crucial category in helping us note the amount of “reactions,” “shares,” and “comments” on the Facebook post sharing the link. By noting the reactions, we were able to assert which type of links get the most reactions and replies from the audience.

[Go to:](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/)

## Results

### Myths

The claims made by the authors on the Facebook Pages were diverse and ranged from questioning the ethics of administration to a total disregard of the benefits of immunization. Demographics of the overseers of the Facebook Page or users cannot be known as Pages can be accessed worldwide; however, given the posts about the Centers for Disease Control and Prevention (CDC) and current events in the United States, we conjecture that the users and majority of commentators are from the United States. Claims under “Myths” numbered far greater than those listed under “Truths.” Claims are listed in order of greatest to least in number.

#### Claim 1: “Vaccines Fail”

Several users on all Facebook Pages expressed the concern that vaccines are not 100% safe and people should opt to not vaccinate. References were made to the recent outbreaks of measles in the United States, and users claimed that most affected individuals were already vaccinated. The CDC reports that a majority of measles-affected individuals were unvaccinated [[17](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref17)]. Claims were made that the measles/mumps/rubella (MMR) vaccine was failing, and the need for the DTaP vaccine for tetanus was disregarded, as tetanus is not a communicable disease.

#### Claim 2: “Vaccine Schedules are Overwhelming and Spark Autoimmunity”

Users expressed concerns with the number of immunizations being added to child schedules by the CDC and with multiple vaccines given at one time. Parents were concerned about vaccines overstimulating the immune system. Some parents claimed to go with “alternative” immunization schedules in collaboration with their health care providers [[18](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref18)]. This included giving fewer vaccines at once and skipping vaccines deemed “not important” by the parents. Users also expressed concerns about the differences in child vaccine schedules among different countries and used it as a reason not to vaccinate against certain diseases (eg, the United Kingdom does not vaccinate against varicella).

#### Claim 3: “Vaccines Contain Harmful Adjuvants”

Adjuvants used in vaccines have been under heavy scrutiny on all Facebook Pages. Adjuvants are added to a vaccine to strengthen its ability to stimulate the immune system [[19](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref19)]; however, they are believed to be responsible for causing a variety of diseases such as cancer, infertility, Alzheimer’s, and autism. Each vaccine has been linked to its own set of mythical consequences from contained adjuvants (see [Table 1](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/table/table1/)). [Table 1](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/table/table1/) lists the most popular vaccines discussed on the Facebook Pages with the most popular adjuvant in each vaccine that is accused of causing harm. The table highlights the effect of the disease on unprotected individuals to reinforce the dangers of the infections prevented by immunization.

### Table 1

Vaccines and their alleged consequences.

| Diseases that can be prevented by immunization | Available vaccine(s) | Adjuvant in vaccine allegedly causing adverse effects | Prevailing myth(s) |
| --- | --- | --- | --- |
| Human papillomavirus | Gardasil, Gardasil 9 | Polysorbate 80, aluminium | Infertility, premature ovarian failure, paralysis |
| Measles, mumps, rubella (MMR) | MMR | Aluminum, fetal bovine serum, recombinant human albumin | Autism, seizures, measles shedding from a vaccine, Alzheimer’s, lupus, aseptic meningitis |
| Diphtheria, pertussis, tetanus | Dtap, Tdap | Formaldehyde, polysorbate 80, bovine serum albumin | SIDSa, autism, vaccine-induced pertussis, neurodevelopmental problems, miscarriage, death |
| Polio | Inactivated poliovirus | Simian virus 40 | Cancer, vaccine-induced paralysis |

[Open in a separate window](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/table/table1/?report=objectonly)

aSIDS: sudden infant death syndrome.

The human papillomavirus vaccine was heavily linked to infertility and polycystic ovarian syndrome, MMR vaccine to autism and epilepsy, polio vaccine to cancer, and DTaP vaccine to sudden infant death syndrome (SIDS). Vaccines are accused of containing fetal cells as adjuvants, and claims are made that they influence the sexuality of teenagers and lead to homosexuality.

### Truths

“Truths” contained information shared on Facebook that could be supported by peer-reviewed scientific evidence. Repeated concerns were raised over the efficacy of the flu vaccine. Flu vaccines are in production before the flu season begins (meetings in February for the Northern Hemisphere and in September for the Southern Hemisphere), and this information was shared on the Facebook Pages. Flu strains are predicted based on surveillance, laboratory, and clinical studies [[20](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref20)]. The flu vaccine’s effectiveness was questioned on all Facebook Pages; however, the effectiveness of the flu shot can change every year [[21](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref21)]. The Government of Canada states that the flu virus may change while the vaccine is in production: “Even when there is a less-than-ideal match or lower effectiveness against one virus, the seasonal flu shot can still provide protection against the remaining two or three viruses. If you do get the flu, the flu shot may reduce the severity of your symptoms” [[21](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref21)].

Another concern raised was “over-vaccinating” against pertussis (or whooping cough) as the vaccine is allegedly not effective. Schwartz et al [[22](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref22)] found that 4 years postimmunization, immunity to pertussis declined significantly, especially with the acellular vaccine as compared to the whole-cell vaccine. Booster shots during pregnancy or priming with the whole-cell vaccine are recommended to optimize pertussis control [[22](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref22)]. Use of the acellular vaccine instead of the whole-cell vaccine was another topic discussed on the Facebook Pages. A case-control study by Klein et al [[23](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref23)] found that, among teenagers who had received vaccines for pertussis at Kaiser Permanente Northern California, those immunized with whole-cell vaccine were more protected in outbreaks compared to teenagers who received the acellular pertussis vaccine.

### Consequences

Autism is the most widely known affliction that allegedly implicates vaccines. The most popular consequences linked to vaccines also included SIDS, asthma, epilepsy, cancer, Alzheimer’s, miscarriage, infertility, and death (see [Table 1](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/table/table1/)). Testimonies from parents sharing information about the death of their children and posting their pictures are extremely popular on all the Pages. These have a profound effect on other viewers, as evidenced by their responses. Mothers have shared their hesitancy of vaccinating their children after viewing these posts.

### Reactions to Current Measles Outbreaks

Measles outbreaks are at an all-time high in the United States, since the year 1992, with the numbers of cases growing. To achieve herd immunity for measles, 95% of the population must be immune to the infection [[24](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref24)]. Outbreaks in communities that are unvaccinated in New York account for more than 75% of the cases, with the majority affecting Orthodox Jewish communities where immunization rates are low. All Facebook Pages discussed the current coverage of measles outbreaks; however, 2 of the 4 Pages we examined posted more frequently about the mainstream media coverage of the outbreaks. Reactions to outbreaks included denial of events, accusing mainstream media of falsifying reports, and claiming that most individuals spreading the infection were vaccinated. While a vaccinated individual could contract measles, the chances are much lower compared to those for unvaccinated individuals, and the disease presents in a milder form [[21](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref21)].

### Website Data

Website links shared on the Facebook Pages were followed. Data were also collected from January 2019 to May 2019. Website information was categorized into Who, What, When, Why, and Opportunity.

#### “Who”

We found that the shared website links were predominately blog posts coming from antivaccine activists. The most popular website shared on all the Pages was Green Med Info [[25](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref25)]. The author, Sayer Ji, is a self-proclaimed expert on the rights and wrongs of immunization and supports alternative medicine [[25](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref25)]. Other types of shared websites included news article, testimonials, and studies. Robert Kennedy is a huge supporter of antivaccine sentiments. His blog website was shared numerous times on all Pages.

#### “What”

We found that the themes emerging from shared content on the websites varied. These included stories of vaccinated individuals getting the infection they were vaccinated against, testimonies from mothers whose children allegedly died postimmunization, accusations towards the government and physicians for promoting vaccines to make money, condemnation of mandatory vaccine bills and laws, “expert” testimonies on dangers of immunization, promotion of naturopathic medicine, and denial of the harm of illnesses that vaccines protect from. Some websites claimed that childhood infection with measles provides protection from cardiac disease in adult life and women will be protected from ovarian cancer, but the sites do not produce sufficient evidence to support these claims.

#### “When”

Information on the vast majority of the websites had been published within the last month; however, for sites that were not current, they were at least published in the last decade. The growth of these Pages in the last decade supports the fact that antivaccine content has increased with the use of social media and the internet [[6](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref6),[26](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref26)]. The “post” or “share” date on the Facebook Pages compared to the original publication date of the website was within 1 week for most posts on 3 of the 4 Pages. We found that one Facebook Page in particular shared website links published 1 or 2 years before current events, but still received lots of support from followers.

#### “Why”

The main reasons for sharing a website link were either to promote or condemn its content. The majority of website links were supported and shared as a way of endorsing the antivaccine information. We do not have reason to believe that any of the shared website links were vetted for scientific evidence or truthfulness; instead, content that would have a profound effect on viewers was shared. This included “new” studies and information such as the “benefits” of getting an infection.

#### “Opportunity”

We found that comments, “reactions,” and “shares” on the Facebook posts sharing the website link did not show any specific trend. The number of responses varied heavily; this could be attributed to the number of online users or personal interests. Therefore, we cannot make a conclusion on whether any specific topics sparked user interest; however, a news article shared on one of the Pages about Kailyn Lowry, an actress, had the highest number of “reactions” compared to any other website link shared on one of the Facebook Pages. This is significant, as celebrities have the platform to influence many people across different geographical areas just by sharing their personal beliefs [[9](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref9),[27](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref27)].

[Go to:](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/)

## Discussion

### Principal Findings

The analysis of the Facebook Pages led to emerging themes from the ongoing discussion among the users and their use of Facebook as their platform to promote their anti-vaccine beliefs: (1) forming an online “community” consisting of like-minded individuals and similar beliefs, (2) the widespread reach of anti-vaccine messages, and (3) debating the ethics of mandatory immunization and content moderation.

The majority of the online population of users on the Facebook Page were likely from the United States. We noticed there was a huge appreciation of community and support for one another. For example, if anyone posting antivaccine content was criticized or condemned, other users would step up to support them in the comments of posts. There was a tremendous amount of support in mother-to-mother communication. A social media analysis conducted by Gruzd and Haythornthwaite [[28](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref28)] analyzed a 1-month sample of Twitter messages to trace interaction via social media and understand “how a community is formed and maintained online.” The study found that network analysis can facilitate understanding of “what, and who, compromises and sustains a network…” [[28](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref28)]. Gruzd and Haythornwaite [[28](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref28)] found that active participation and attention to others were extremely significant aspects of building an online community. This finding translates very well to our analysis as aforementioned: The sharing of pictures and stories of children who were allegedly afflicted by vaccines was very popular, and mothers tended to gather and demonstrate support for one another in such cases. Supporting one another in their “time of need” and defending their collective viewpoints helped foster a sense of community among each other [[29](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref29)].

A very significant finding was the raising of money for private autopsies postdeath of a child from SIDS. Users believe that immunization can cause SIDS; therefore, when the cause of death is not officially linked to vaccines, users on these Facebook Pages collected donations for private autopsies, with the amounts ranging in the thousands. There was little to no follow-up from parents who received this money on whether the autopsy was done, and no follow-up included medical proof of immunization being linked to SIDS. Users on Facebook who support these Pages are a highly tight-knit community, who have limited trust in government authorities and medical professionals [[30](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref30)]. Users on one of the Facebook Pages examined in this research encouraged a mother to not fill a prescription for Tamiflu after her son had been diagnosed with influenza and was running a high fever and had a seizure [[31](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref31)]. None of the comments on the post encouraged the mother to seek medical help and fill the prescription. The mother instead opted to treat with natural remedies such as peppermint oil, vitamin C, and lavender [[31](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref31)]. The users on Facebook also suggested the use of home remedies such as breastmilk, thyme, and elderberry to treat her child — none of which are recommended treatment for influenza — and the child eventually died 4 days later at the hospital [[31](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref31)]. This case, in particular, highlights the trust and confidence users of these Pages are placing upon each other.

This relationship among users who have likely never even met each other can be incredibly difficult to infiltrate and change, as they share a common ground of strong values and beliefs. We must focus resources on individuals who are undecided and caught in the middle of the debate [[11](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref11)]. Interventions must be community-based, and education and information on vaccines must be encouraged by alike members in the community (such as parent groups) [[32](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref32)]. As health care professionals, we must become informed on adjuvants and how they impact the body (along with other concerns noted in the Results section) and therefore be equipped to answer questions parents may have to build trust [[33](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref33)]. It is incredibly important for health care personnel to recognize and understand this problem [[33](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref33)]. Vaccines are one of the largest defenses we have against communicable diseases, and we must continue to educate and attempt to change the attitudes surrounding this important issue.

### Widespread Reach of Antivaccine Messages

Antivaccine sentiments no longer belong to a small group of people; instead, they have global implications [[5](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref5),[11](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref11),[34](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref34)]. Three of the 4 Facebook Pages have 100,000 to 135,000 likes, and the most popular website shared on the Facebook Pages — Green Med Info — claims to have 500,000 monthly visitors [[25](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref25)]. Our research of the Pages uncovered how detrimental these campaigns can be in underdeveloped nations of the world. Pakistan is one of 3 countries that has failed to eradicate polio transmission [[35](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref35)]. The spread of vaccine misinformation through the availability of smartphones and social media is encouraging a public health threat in one of the most vulnerable nations of the world [[34](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref34),[35](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref35)]. Not only has misinformation threatened Pakistan’s public health but the unauthorized immunization campaign directed by the Central Intelligence Agency in an attempt to locate Osama Bin Laden has broken the trust between locals and foreign public health efforts [[36](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref36)]. Another example of mistrust of western medicine in the developing world is Nigeria, where Islamic militant groups believe that immunization is a ploy to sterilize Muslims [[37](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref37)].

### Measles outbreaks have been on the rise globally, increasing by 30% from 2016 to 2017 [[38](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref38)]. These outbreaks correlate heavily with whether citizens trust vaccines. For example, France experienced measles outbreaks, with 1 in 3 of their citizens believing that vaccines are not safe [[38](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref38)]. Currently, the COVID-19 pandemic has taken the world by storm, and researchers and the general public are eager for a vaccine to help flatten the curve of the disease. When the time comes to distribute a vaccine for this virus (or any uncontrollable communicable disease), we must not only promote the vaccine to those who trust immunization but also promote targeted interventions for users of these Pages who refuse to vaccinate by increasing opportunity for dialogue and creating safe spaces [[32](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref32),[33](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref33)]. Our recommendation for institutions includes adding and enhancing education on vaccines and immunization for future health care professionals, such as doctors, nurses, and dentists, to address concerns of users on these Facebook Pages [[33](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref33),[39](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref39)]. Debating the Ethics of Mandatory Immunization and Content Moderation

A focal point of discussion on Facebook posts, website shares, and the comments section was the ethics of mandatory immunization laws [[40](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref40)]. Antivaccine groups are heavily against the passing of any bill that supports mandatory immunization [[41](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref41)]. Vaccine Choice Canada — one of Canada’s largest antivaccine organizations — claims they are prepared to fight New Brunswick’s 2019 proposed bill that will not allow children to attend public school without proof of immunization [[42](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref42)]. This nightmare for antivaccine organizations has already become a reality for those living in the state of California in the United States. There were calls to action against the governments that propose mandatory immunization and consistent derogatory comments made against the governor of California for being an open vaccine supporter on Facebook. Users on Facebook urged others to join protests, sign petitions, and call government officials’ offices to discourage mandatory immunization. Users frequently cited the Constitution of the United States, claiming their rights as citizens of the country have been violated through these bills and laws.

The ethical debate has also included the censoring of content on social media websites. Facebook has claimed to not remove vaccine misinformation; however, they assured the public that they will make it less prominent [[7](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref7),[43](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref43),[44](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref44)]. This includes removing content from recommended groups and ranking posts with misinformation lower on the newsfeed [[7](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref7),[44](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref44)]. Other social media sites that are moderating antivaccine content include YouTube, Pinterest, and Twitter. YouTube has stopped serving ads on any antivaccine-promoting video and attempted to make content on the benefits of immunization easier to find [[45](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref45)]. The WHO has applauded Pinterest for being a leader in removing vaccine misinformation from their website [[44](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref44)]. Pinterest has one of the most rigorous restrictions on posting of vaccine misinformation, going as far as to block any searches with the terms “anti-vaccination” or “anti-vax” [[44](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref44)].

### Conclusion

The antivaccine campaign has unfortunately used social media as a vessel to spread misinformation to users, especially parents [[4](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref4)-[7](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref7),[9](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref9),[11](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref11)-[13](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref13),[26](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref26),[27](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref27),[29](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref29),[30](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref30),[32](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref32),[37](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref37),[40](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref40),[41](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref41),[46](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref46),[47](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref47)]. As vaccine hesitancy increases, we increase the risk of a public health crisis and lessen our chances of controlling crises like the COVID-19 pandemic. Although users on Facebook have mentioned “Truths,” the number of “Myths” supersede these truths, and the benefits of immunization greatly outweigh the risk. We must understand the local-global implications of allowing preventable diseases to make a comeback. Health care providers deal directly with members of the public who are uncertain about immunization, and it becomes their job to be informed on users’ common misconceptions [[33](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref33),[39](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref39)]. Fake news travels faster than truth, building momentum [[48](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609192/#ref48)]; therefore, targeted promotion of vaccines that address specific claims on the internet is warranted.

Limitations of this study include the ever-dynamic nature of the internet, with the freedom for administrators and users to remove posts as desired, and the time constraints in which we studied the Facebook Pages. The social media audit template used to organize and categorize data has previously not been used in studying data for the interpretation of messages. We also cannot be certain that each different follower of the Page is a different individual, as one person may hold many accounts using different email addresses. This analysis focuses only on Facebook, which, as aforementioned, has a less scrutinizing approach to removing antivaccine content; therefore, analysis of other websites such as Twitter, YouTube, and Pinterest is warranted to compare the type of content and spread of readership. The worldview of both authors is that of nurses and health care providers, and this article has been written with a pro-immunization point of view.

.

# 3. PROBLEM ANALYSIS

### EXISTING APPROACH:

### The existing methods using Most models assume a standard SEIR structure. Fraser and collegues to estimate size but make different changes on the nature of the different compartments and their respective residence times

* 1. **Drawbacks**

There are of course many limitations of this study. There are other social media platforms, apart from Facebook, that should be explored \_ but Facebook is the largest. Similar behaviors should arise in any platform where communities can form..

### 3.2 Proposed System

Here we use machine learning to quantify COVID-19 content among online opponents of establishment health guidance, in particular vaccinations (``anti-vax''). We \_nd that the anti-vax community is developing a less focused debate around COVID-19 than its counterpart, the pro-vaccination (``pro-vax'') community. However, the anti-vax community exhibits a broader range of ``\_avors'' of COVID-19 topics, and hence can appeal to a broader cross-section of individuals seeking COVID-19 guidance online, e.g. individuals wary of a mandatory fast-tracked COVID-19 vaccine or those seeking alternative remedies. Hence the anti-vax community looks better positioned to attract fresh support going forward than the pro-vax community. This is concerning since a widespread lack of adoption of a COVID-19 vaccine will mean the world falls short of providing herd immunity, leaving countries open to future COVID-19 resurgences.We provide a mechanistic model that interprets these results and could help in assessing the likely ef\_cacy of intervention strategies

**3.2.1 Advantages**

Accuracy

# 3.3 Software And Hardware Requirements

**SOFTWARE REQUIREMENTS**

The functional requirements or the overall description documents include the product perspective and features, operating system and operating environment, graphics requirements, design constraints and user documentation.

The appropriation of requirements and implementation constraints gives the general overview of the project in regards to what the areas of strength and deficit are and how to tackle them.

* **Python idel 3.7 version (or)**
* **Anaconda 3.7 ( or)**
* **Jupiter (or)**
* **Google colab**

**HARDWARE REQUIREMENTS**

Minimum hardware requirements are very dependent on the particular software being developed by a given Enthought Python / Canopy / VS Code user. Applications that need to store large arrays/objects in memory will require more RAM, whereas applications that need to perform numerous calculations or tasks more quickly will require a faster processor.

* **Operating system : windows, linux**
* **Processor : minimum intel i3**
* **Ram : minimum 4 gb**
* **Hard disk : minimum 250gb**

**3.4 About Dataset**

**3.5 Algorithms**

# CNN Working Procedure

# To demonstrate how to build a convolutional neural network based image classifier, we shall build a 7 layer neural network that will identify and separate one image from other. This network that we shall build is a very small network that we can run on a CPU as well. Traditional neural networks that are very good at doing image classification have many more parameters and take a lot of time if trained on normal CPU. However, our objective is to show how to build a real-world convolutional neural network using TENSORFLOW.

# 4. SYSTEM DESIGN

### UML DIAGRAMS

The System Design Document describes the system requirements, operating environment, system and subsystem architecture, files and database design, input formats, output layouts, human-machine interfaces, detailed design, processing logic, and external interfaces.

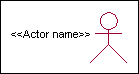
### Global Use Case Diagrams:

Identification of actors:

**Actor:** Actor represents the role a user plays with respect to the system. An actor interacts with, but has no control over the use cases.

Graphical representation:

<<Actor name>>





Actor

An actor is someone or something that:

Interacts with or uses the system.

Provides input to and receives information from the system.

Is external to the system and has no control over the use cases. Actors are discovered by examining:

* + - * Who directly uses the system?
      * Who is responsible for maintaining the system?
      * External hardware used by the system.
      * Other systems that need to interact with the system. Questions to identify actors:
        + Who is using the system? Or, who is affected by the system? Or, which groups need help from the system to perform a task?
        + Who affects the system? Or, which user groups are needed by the system to perform its functions? These functions can be both main functions and secondary functions such as administration.
        + Which external hardware or systems (if any) use the system to perform tasks?
        + What problems does this application solve (that is, for whom)?
        + And, finally, how do users use the system (use case)? What are they doing with the system?

The actors identified in this system are:

### System Administrator

1. **Customer**
2. **Customer Care**

Identification of usecases:

**Usecase:** A use case can be described as a specific way of using the system from a user’s (actor’s) perspective.

### Graphical representation:



A more detailed description might characterize a use case as:

* + Pattern of behavior the system exhibits
  + A sequence of related transactions performed by an actor and the system
  + Delivering something of value to the actor Use cases provide a means to:
  + capture system requirements
  + communicate with the end users and domain experts
  + test the system

Use cases are best discovered by examining the actors and defining what the actor will be able to do with the system.

Guide lines for identifying use cases:

* For each actor, find the tasks and functions that the actor should be able to perform or that the system needs the actor to perform. The use case should represent a course of events that leads to clear goal
* Name the use cases.
* Describe the use cases briefly by applying terms with which the user is familiar. This makes the description less ambiguous

Questions to identify use cases:

* What are the tasks of each actor?
* Will any actor create, store, change, remove or read information in the system?
* What use case will store, change, remove or read this information?
* Will any actor need to inform the system about sudden external changes?
* Does any actor need to inform about certain occurrences in the system?
* What usecases will support and maintains the system?

### Flow of Events

A flow of events is a sequence of transactions (or events) performed by the system. They typically contain very detailed information, written in terms of what the system should do, not how the system accomplishes the task. Flow of events are created as separate files or documents in your favorite text editor and then attached or linked to a use case using the Files tab of a model element.

A flow of events should include:

* When and how the use case starts and ends
* Use case/actor interactions
* Data needed by the use case
* Normal sequence of events for the use case
* Alternate or exceptional flows Construction of Usecase diagrams:

Use-case diagrams graphically depict system behavior (use cases). These diagrams present a high level view of how the system is used as viewed from an outsider’s (actor’s) perspective. A use-case diagram may depict all or some of the use cases of a system.

A use-case diagram can contain:

* + actors ("things" outside the system)
  + use cases (system boundaries identifying what the system should do)
  + Interactions or relationships between actors and use cases in the system including the associations, dependencies, and generalizations.

Relationships in use cases:

### Communication:

The communication relationship of an actor in a usecase is shown by connecting the actor symbol to the usecase symbol with a solid path. The actor is said to communicate with the usecase.

### Uses:

A Uses relationship between the usecases is shown by generalization arrow from the usecase.

### Extends:

The extend relationship is used when we have one usecase that is similar to another usecase but does a bit more. In essence it is like subclass.

### SEQUENCE DIAGRAMS

A sequence diagram is a graphical view of a scenario that shows object interaction in a time- based sequence what happens first, what happens next. Sequence diagrams establish the roles of objects and help provide essential information to determine class responsibilities and interfaces. There are two main differences between sequence and collaboration diagrams: sequence diagrams show time-based object interaction while collaboration diagrams show how objects associate with each other. A sequence diagram has two dimensions: typically, vertical placement represents time and horizontal placement represents different objects.

### Object:

An object has state, behavior, and identity. The structure and behavior of similar objects are defined in their common class. Each object in a diagram indicates some instance of a class. An object that is not named is referred to as a class instance.

The object icon is similar to a class icon except that the name is underlined: An object's concurrency is defined by the concurrency of its class.

### Message:

A message is the communication carried between two objects that trigger an event. A message carries information from the source focus of control to the destination focus of control. The synchronization of a message can be modified through the message specification. Synchronization means a message where the sending object pauses to wait for results.

### Link:

A link should exist between two objects, including class utilities, only if there is a relationship between their corresponding classes. The existence of a relationship between two classes symbolizes a path of communication between instances of the classes: one object may send messages to another. The link is depicted as a straight line between objects or objects and class instances in a collaboration diagram. If an object links to itself, use the loop version of the icon.

**CLASS DIAGRAM:**

Identification of analysis classes:

A class is a set of objects that share a common structure and common behavior (the same attributes, operations, relationships and semantics). A class is an abstraction of real-world items. There are 4 approaches for identifying classes:

* 1. Noun phrase approach:
  2. Common class pattern approach.
  3. Use case Driven Sequence or Collaboration approach.
  4. Classes , Responsibilities and collaborators Approach

### Noun Phrase Approach:

The guidelines for identifying the classes:

* + Look for nouns and noun phrases in the usecases.
  + Some classes are implicit or taken from general knowledge.
  + All classes must make sense in the application domain; Avoid computer implementation classes – defer them to the design stage.
  + Carefully choose and define the class names After identifying the classes we have to eliminate the following types of classes:
  + Adjective classes.

### Common class pattern approach:

The following are the patterns for finding the candidate classes:

* + Concept class.
  + Events class.
  + Organization class
  + Peoples class
  + Places class
  + Tangible things and devices class.

### Use case driven approach:

We have to draw the sequence diagram or collaboration diagram. If there is need for some classes to represent some functionality then add new classes which perform those functionalities.

### CRC approach:

The process consists of the following steps:

* + Identify classes’ responsibilities ( and identify the classes )
  + Assign the responsibilities
  + Identify the collaborators. Identification of responsibilities of each class:

The questions that should be answered to identify the attributes and methods of a class respectively are:

1. What information about an object should we keep track of?
2. What services must a class provide? Identification of relationships among the classes:

Three types of relationships among the objects are:

Association: How objects are associated?

Super-sub structure: How are objects organized into super classes and sub classes? Aggregation: What is the composition of the complex classes?

Association:

The **questions** that will help us to identify the associations are:

* 1. Is the class capable of fulfilling the required task by itself?
  2. If not, what does it need?
  3. From what other classes can it acquire what it needs? Guidelines for identifying the tentative associations:
* A dependency between two or more classes may be an association. Association often corresponds to a verb or prepositional phrase.
* A reference from one class to another is an association. Some associations are implicit or taken from general knowledge.

Some common association patterns are:

Location association like part of, next to, contained in….. Communication association like talk to, order to ……

We have to eliminate the unnecessary association like implementation associations, ternary or n- ary associations and derived associations.

Super-sub class relationships:

Super-sub class hierarchy is a relationship between classes where one class is the parent class of another class (derived class).This is based on inheritance.

Guidelines for identifying the super-sub relationship, a generalization are

### Top-down*:*

Look for noun phrases composed of various adjectives in a class name. Avoid excessive refinement. Specialize only when the sub classes have significant behavior.

### Bottom-up*:*

Look for classes with similar attributes or methods. Group them by moving the common attributes and methods to an abstract class. You may have to alter the definitions a bit.

### Reusability*:*

Move the attributes and methods as high as possible in the hierarchy.

### Multiple inheritances*:*

Avoid excessive use of multiple inheritances. One way of getting benefits of multiple inheritances is to inherit from the most appropriate class and add an object of another class as an attribute.

### Aggregation or a-part-of relationship:

It represents the situation where a class consists of several component classes. A class that is composed of other classes doesn’t behave like its parts. It behaves very difficultly. The major properties of this relationship are transitivity and anti symmetry.

The **questions** whose answers will determine the distinction between the part and whole relationships are:

* + Does the part class belong to the problem domain?
  + Is the part class within the system’s responsibilities?
  + Does the part class capture more than a single value?( If not then simply include it as an attribute of the whole class)
  + Does it provide a useful abstraction in dealing with the problem domain? There are three types of aggregation relationships. They are:

### Assembly:

It is constructed from its parts and an assembly-part situation physically exists.

### Container:

A physical whole encompasses but is not constructed from physical parts.

### Collection member:

A conceptual whole encompasses parts that may be physical or conceptual. The container and collection are represented by hollow diamonds but composition is represented by solid diamond.

**USE CASE DIAGRAM**

A use case diagram in the Unified Modeling Language (UML) is a type of behavioral diagram defined by and created from a Use-case analysis. Its purpose is to present a graphical overview of the functionality provided by a system in terms of actors, their goals (represented as use cases), and any dependencies between those use cases. The main purpose of a use case diagram is to show what system functions are performed for which actor. Roles of the actors in the system can be depicted.

****

### Fig 1: Use Case Diagram

### CLASS DIAGRAM

In software engineering, a class diagram in the Unified Modeling Language (UML) is a type of static structure diagram that describes the structure of a system by showing the system's classes, their attributes, operations (or methods), and the relationships among the classes. It explains which class contains information.

### 

### Fig 2:Class Diagram

**SEQUENCE DIAGRAM**

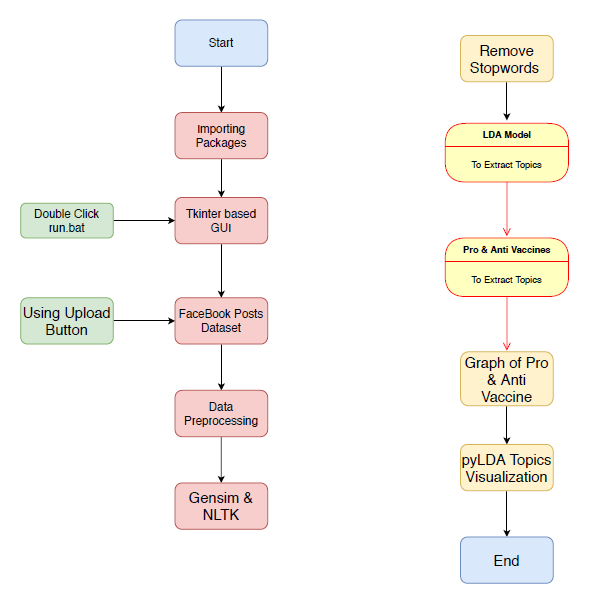
A sequence diagram in Unified Modeling Language (UML) is a kind of interaction diagram that shows how processes operate with one another and in what order. It is a construct of a Message Sequence Chart. Sequence diagrams are sometimes called event diagrams, event scenarios, and timing diagrams.

****

**Fig 3: Sequence Diagram**

**5.IMPLEMENTATION**

**5.1 FLOW CHART:**



**5.2 Code**

**from tkinter import messagebox**

**from tkinter import \***

**from tkinter import simpledialog**

**import tkinter**

**from tkinter import filedialog**

**import matplotlib.pyplot as plt**

**from tkinter.filedialog import askopenfilename**

**from sklearn.model\_selection import train\_test\_split**

**from sklearn.metrics import accuracy\_score**

**import numpy as np**

**import pandas as pd**

**from genetic\_selection import GeneticSelectionCV**

**from sklearn.metrics import classification\_report**

**from sklearn.metrics import confusion\_matrix**

**from sklearn import svm**

**from keras.models import Sequential**

**from keras.layers import Dense**

**import time**

**main = tkinter.Tk()**

**main.title("Android Malware Detection")**

**main.geometry("1300x1200")**

**global filename**

**global train**

**global svm\_acc, nn\_acc, svmga\_acc, annga\_acc**

**global X\_train, X\_test, y\_train, y\_test**

**global svmga\_classifier**

**global nnga\_classifier**

**global svm\_time,svmga\_time,nn\_time,nnga\_time**

**def upload():**

**global filename**

**filename = filedialog.askopenfilename(initialdir="dataset")**

**pathlabel.config(text=filename)**

**text.delete('1.0', END)**

**text.insert(END,filename+" loaded\n");**

**def generateModel():**

**global X\_train, X\_test, y\_train, y\_test**

**text.delete('1.0', END)**

**train = pd.read\_csv(filename)**

**rows = train.shape[0] # gives number of row count**

**cols = train.shape[1] # gives number of col count**

**features = cols - 1**

**print(features)**

**X = train.values[:, 0:features]**

**Y = train.values[:, features]**

**print(Y)**

**X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, Y, test\_size = 0.2, random\_state = 0)**

**text.insert(END,"Dataset Length : "+str(len(X))+"\n");**

**text.insert(END,"Splitted Training Length : "+str(len(X\_train))+"\n");**

**text.insert(END,"Splitted Test Length : "+str(len(X\_test))+"\n\n");**

**def prediction(X\_test, cls): #prediction done here**

**y\_pred = cls.predict(X\_test)**

**for i in range(len(X\_test)):**

**print("X=%s, Predicted=%s" % (X\_test[i], y\_pred[i]))**

**return y\_pred**

**# Function to calculate accuracy**

**def cal\_accuracy(y\_test, y\_pred, details):**

**cm = confusion\_matrix(y\_test, y\_pred)**

**accuracy = accuracy\_score(y\_test,y\_pred)\*100**

**text.insert(END,details+"\n\n")**

**text.insert(END,"Accuracy : "+str(accuracy)+"\n\n")**

**text.insert(END,"Report : "+str(classification\_report(y\_test, y\_pred))+"\n")**

**text.insert(END,"Confusion Matrix : "+str(cm)+"\n\n\n\n\n")**

**return accuracy**

**def runSVM():**

**global svm\_acc**

**global svm\_time**

**start\_time = time.time()**

**text.delete('1.0', END)**

**cls = svm.SVC(C=2.0,gamma='scale',kernel = 'rbf', random\_state = 2)**

**cls.fit(X\_train, y\_train)**

**prediction\_data = prediction(X\_test, cls)**

**svm\_acc = cal\_accuracy(y\_test, prediction\_data,'SVM Accuracy')**

**svm\_time = (time.time() - start\_time)**

**def runSVMGenetic():**

**text.delete('1.0', END)**

**global svmga\_acc**

**global svmga\_classifier**

**global svmga\_time**

**estimator = svm.SVC(C=2.0,gamma='scale',kernel = 'rbf', random\_state = 2)**

**svmga\_classifier = GeneticSelectionCV(estimator,**

**cv=5,**

**verbose=1,**

**scoring="accuracy",**

**max\_features=5,**

**n\_population=50,**

**crossover\_proba=0.5,**

**mutation\_proba=0.2,**

**n\_generations=40,**

**crossover\_independent\_proba=0.5,**

**mutation\_independent\_proba=0.05,**

**tournament\_size=3,**

**n\_gen\_no\_change=10,**

**caching=True,**

**n\_jobs=-1)**

**start\_time = time.time()**

**svmga\_classifier = svmga\_classifier.fit(X\_train, y\_train)**

**svmga\_time = svm\_time/2**

**prediction\_data = prediction(X\_test, svmga\_classifier)**

**svmga\_acc = cal\_accuracy(y\_test, prediction\_data,'SVM with GA Algorithm Accuracy, Classification Report & Confusion Matrix')**

**def runNN():**

**global nn\_acc**

**global nn\_time**

**text.delete('1.0', END)**

**start\_time = time.time()**

**model = Sequential()**

**model.add(Dense(4, input\_dim=215, activation='relu'))**

**model.add(Dense(215, activation='relu'))**

**model.add(Dense(1, activation='sigmoid'))**

**model.compile(loss='binary\_crossentropy', optimizer='adam', metrics=['accuracy'])**

**model.fit(X\_train, y\_train, epochs=50, batch\_size=64)**

**\_, ann\_acc = model.evaluate(X\_test, y\_test)**

**nn\_acc = ann\_acc\*100**

**text.insert(END,"ANN Accuracy : "+str(nn\_acc)+"\n\n")**

**nn\_time = (time.time() - start\_time)**

**def runNNGenetic():**

**global annga\_acc**

**global nnga\_time**

**text.delete('1.0', END)**

**train = pd.read\_csv(filename)**

**rows = train.shape[0] # gives number of row count**

**cols = train.shape[1] # gives number of col count**

**features = cols - 1**

**print(features)**

**X = train.values[:, 0:100]**

**Y = train.values[:, features]**

**print(Y)**

**X\_train1, X\_test1, y\_train1, y\_test1 = train\_test\_split(X, Y, test\_size = 0.2, random\_state = 0)**

**model = Sequential()**

**model.add(Dense(4, input\_dim=100, activation='relu'))**

**model.add(Dense(100, activation='relu'))**

**model.add(Dense(1, activation='sigmoid'))**

**model.compile(loss='binary\_crossentropy', optimizer='adam', metrics=['accuracy'])**

**start\_time = time.time()**

**model.fit(X\_train1, y\_train1)**

**nnga\_time = (time.time() - start\_time)**

**\_, ann\_acc = model.evaluate(X\_test1, y\_test1)**

**annga\_acc = ann\_acc\*100**

**text.insert(END,"ANN with Genetic Algorithm Accuracy : "+str(annga\_acc)+"\n\n")**

**def graph():**

**height = [svm\_acc, nn\_acc, svmga\_acc, annga\_acc]**

**bars = ('SVM Accuracy','NN Accuracy','SVM Genetic Acc','NN Genetic Acc')**

**y\_pos = np.arange(len(bars))**

**plt.bar(y\_pos, height)**

**plt.xticks(y\_pos, bars)**

**plt.show()**

**def timeGraph():**

**height = [svm\_time,svmga\_time,nn\_time,nnga\_time]**

**bars = ('SVM Time','SVM Genetic Time','NN Time','NN Genetic Time')**

**y\_pos = np.arange(len(bars))**

**plt.bar(y\_pos, height)**

**plt.xticks(y\_pos, bars)**

**plt.show()**

**font = ('times', 16, 'bold')**

**title = Label(main, text='Android Malware Detection Using Genetic Algorithm based Optimized Feature Selection and Machine Learning')**

**#title.config(bg='brown', fg='white')**

**title.config(font=font)**

**title.config(height=3, width=120)**

**title.place(x=0,y=5)**

**font1 = ('times', 14, 'bold')**

**uploadButton = Button(main, text="Upload Android Malware Dataset", command=upload)**

**uploadButton.place(x=50,y=100)**

**uploadButton.config(font=font1)**

**pathlabel = Label(main)**

**pathlabel.config(bg='brown', fg='white')**

**pathlabel.config(font=font1)**

**pathlabel.place(x=460,y=100)**

**generateButton = Button(main, text="Generate Train & Test Model", command=generateModel)**

**generateButton.place(x=50,y=150)**

**generateButton.config(font=font1)**

**svmButton = Button(main, text="Run SVM Algorithm", command=runSVM)**

**svmButton.place(x=330,y=150)**

**svmButton.config(font=font1)**

**svmgaButton = Button(main, text="Run SVM with Genetic Algorithm", command=runSVMGenetic)**

**svmgaButton.place(x=540,y=150)**

**svmgaButton.config(font=font1)**

**nnButton = Button(main, text="Run Neural Network Algorithm", command=runNN)**

**nnButton.place(x=870,y=150)**

**nnButton.config(font=font1)**

**nngaButton = Button(main, text="Run Neural Network with Genetic Algorithm", command=runNNGenetic)**

**nngaButton.place(x=50,y=200)**

**nngaButton.config(font=font1)**

**graphButton = Button(main, text="Accuracy Graph", command=graph)**

**graphButton.place(x=460,y=200)**

**graphButton.config(font=font1)**

**exitButton = Button(main, text="Execution Time Graph", command=timeGraph)**

**exitButton.place(x=650,y=200)**

**exitButton.config(font=font1)**

**font1 = ('times', 12, 'bold')**

**text=Text(main,height=20,width=150)**

**scroll=Scrollbar(text)**

**text.configure(yscrollcommand=scroll.set)**

**text.place(x=10,y=250)**

**text.config(font=font1)**

**#main.config()**

**main.mainloop()**

**6.TESTING**

### SOFTWARE TESTING

## Testing

Testing is a process of executing a program with the aim of finding error. To make our software perform well it should be error free. If testing is done successfully it will remove all the errors from the software.

#### 6.1.1 Types of Testing

* + - 1. White Box Testing
      2. Black Box Testing
      3. Unit testing
      4. Integration Testing
      5. Alpha Testing
      6. Beta Testing
      7. Performance Testing and so on

#### White Box Testing

Testing technique based on knowledge of the internal logic of an application's code and includes tests like coverage of code statements, branches, paths, conditions. It is performed by software developers

**Black Box Testing**

A method of software testing that verifies the functionality of an application without having specific knowledge of the application's code/internal structure. Tests are based on requirements and functionality.

**Unit Testing**

Software verification and validation method in which a programmer tests if individual units of source code are fit for use. It is usually conducted by the development team.

**Integration Testing**

The phase in software testing in which individual software modules are combined and tested as a group. It is usually conducted by testing teams.

**Alpha Testing**

Type of testing a software product or system conducted at the developer's site. Usually it is performed by the end users.

**Beta Testing**

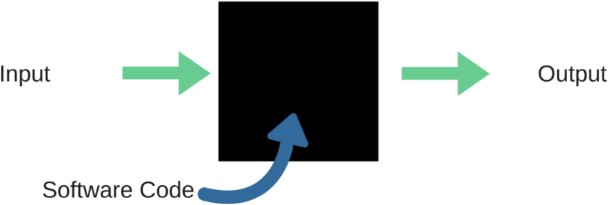
Final testing before releasing application for commercial purpose. It is typically done by end- users or others.

**Performance Testing**

Functional testing conducted to evaluate the compliance of a system or component with specified performance requirements. It is usually conducted by the performance engineer.

#### Black Box Testing

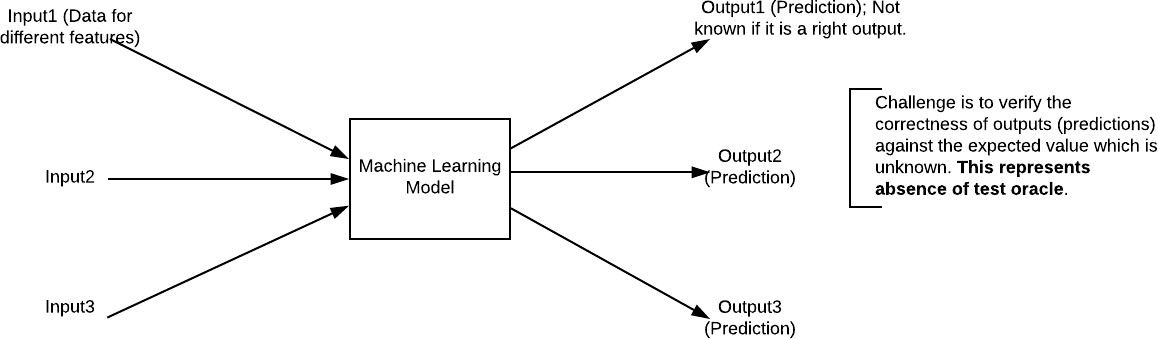
Blackbox testing is testing the functionality of an application without knowing the details of its implementation including internal program structure, data structures etc. Test cases for black box testing are created based on the requirement specifications. Therefore, it is also called as specification-based testing. Fig.4.1 represents the black box testing:



**Fig.:**Black Box Testing

When applied to machine learning models, black box testing would mean testing machine learning models without knowing the internal details such as features of the machine learning

model, the algorithm used to create the model etc. The challenge, however, is to verify the test outcome against the expected values that are known beforehand.



**Fig.:**Black Box Testing for Machine Learning algorithms

The above Fig.4.2 represents the black box testing procedure for machine learning algorithms.

**Table.4.1:**Black box Testing

|  |  |  |
| --- | --- | --- |
| **Input** | **Actual Output** | **Predicted Output** |
| [16,6,324,0,0,0,22,0,0,0,0,0,0] | 0 | 0 |
| [16,7,263,7,0,2,700,9,10,1153,832,9,2] | 1 | 1 |

The model gives out the correct output when different inputs are given which are mentioned in Table 4.1. Therefore the program is said to be executed as expected or correct program

## Testing

Testing is a process of executing a program with the aim of finding error. To make our software perform well it should be error free. If testing is done successfully it will remove all the errors from the software.

#### 7.2.2 Types of Testing

* + - 1. White Box Testing
      2. Black Box Testing
      3. Unit testing
      4. Integration Testing
      5. Alpha Testing
      6. Beta Testing
      7. Performance Testing and so on

#### White Box Testing

Testing technique based on knowledge of the internal logic of an application's code and includes tests like coverage of code statements, branches, paths, conditions. It is performed by software developers

**Black Box Testing**

A method of software testing that verifies the functionality of an application without having specific knowledge of the application's code/internal structure. Tests are based on requirements and functionality.

**Unit Testing**

Software verification and validation method in which a programmer tests if individual units of source code are fit for use. It is usually conducted by the development team.

**Integration Testing**

The phase in software testing in which individual software modules are combined and tested as a group. It is usually conducted by testing teams.

**Alpha Testing**

Type of testing a software product or system conducted at the developer's site. Usually it is performed by the end users.

**Beta Testing**

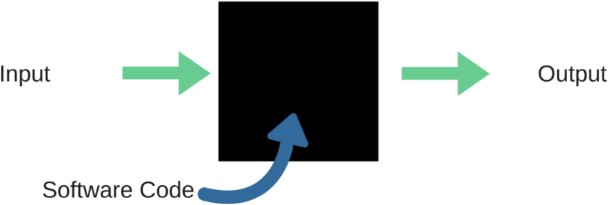
Final testing before releasing application for commercial purpose. It is typically done by end- users or others.

**Performance Testing**

Functional testing conducted to evaluate the compliance of a system or component with specified performance requirements. It is usually conducted by the performance engineer.

#### Black Box Testing

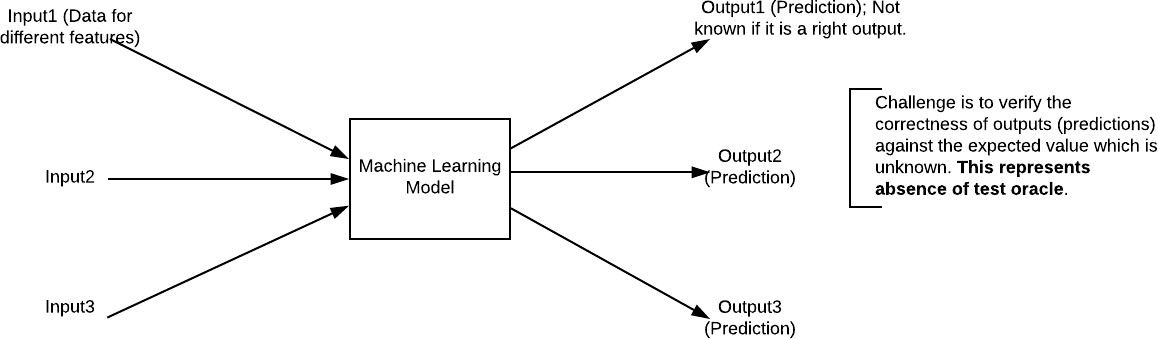
Blackbox testing is testing the functionality of an application without knowing the details of its implementation including internal program structure, data structures etc. Test cases for black box testing are created based on the requirement specifications. Therefore, it is also called as specification-based testing. Fig.4.1 represents the black box testing:



**Fig.:**Black Box Testing

When applied to machine learning models, black box testing would mean testing machine learning models without knowing the internal details such as features of the machine learning

model, the algorithm used to create the model etc. The challenge, however, is to verify the test outcome against the expected values that are known beforehand.



**Fig.:**Black Box Testing for Machine Learning algorithms

The above Fig.4.2 represents the black box testing procedure for machine learning algorithms.

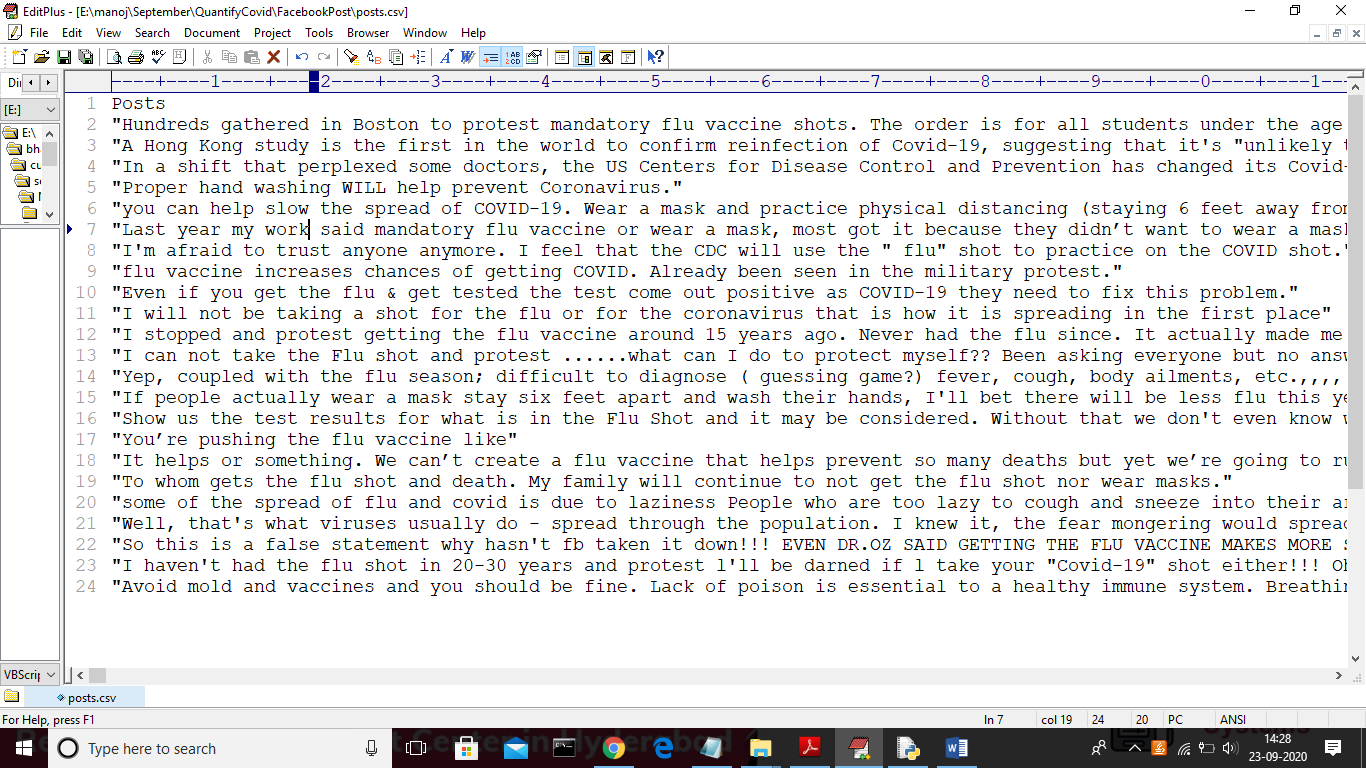
**Table.4.1:**Black box Testing

|  |  |  |
| --- | --- | --- |
| **Input** | **Actual Output** | **Predicted Output** |
| [16,6,324,0,0,0,22,0,0,0,0,0,0] | 0 | 0 |
| [16,7,263,7,0,2,700,9,10,1153,832,9,2] | 1 | 1 |

The model gives out the correct output when different inputs are given which are mentioned in Table 4.1. Therefore the program is said to be executed as expected or correct program

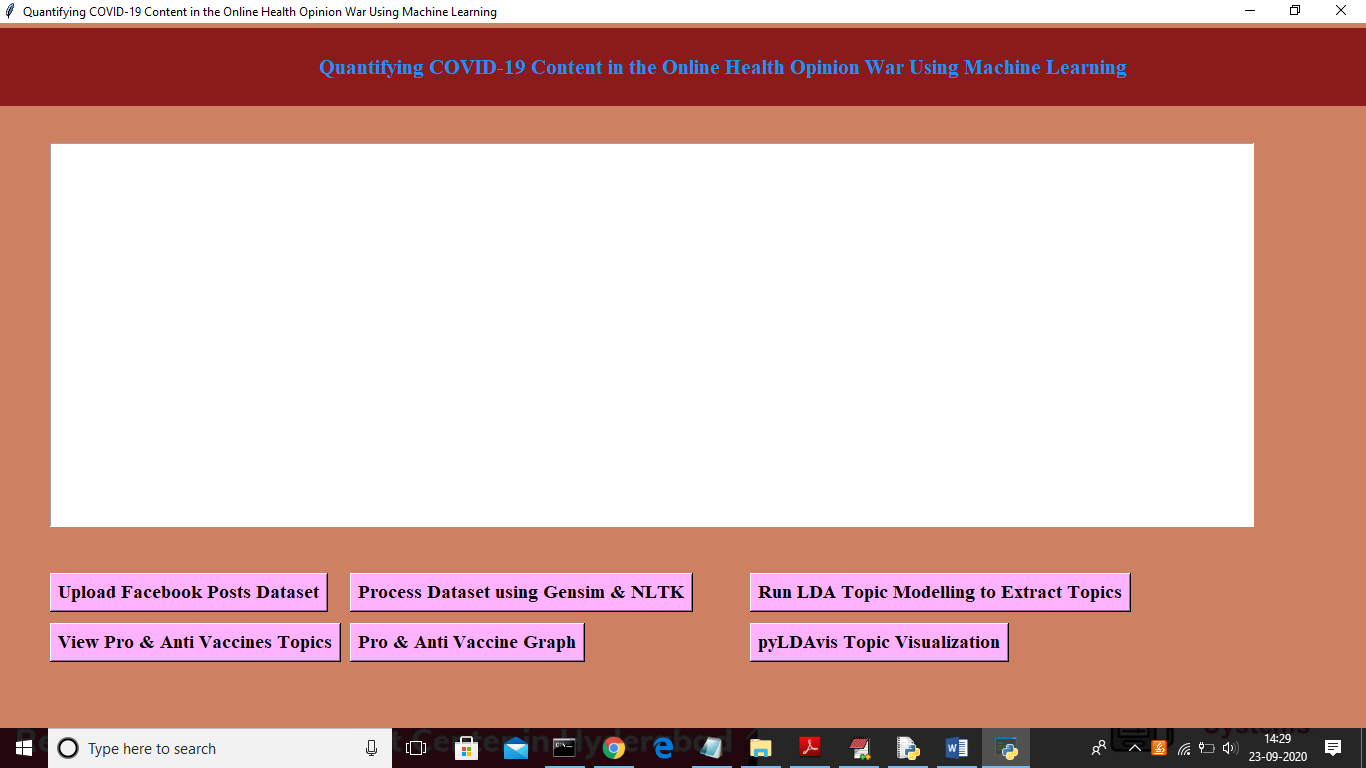
|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Test | Test Case | Test Case | Test Steps | | | Test | Test |
| Cas | Name | Description | Step | Expected | Actual | Case | Priorit |
| e Id |  |  |  |  |  | Statu | Y |
|  |  |  |  |  |  | s |  |
| 01 | Start the | Host the | If it | We | The | High | High |
|  | Applicatio | application | doesn't | cannot | application |  |  |
|  | N | and test if it | Start | run the | hosts |  |  |
|  |  | starts |  | applicati | success. |  |  |
|  |  | making sure |  | on. |  |  |  |
|  |  | the required |  |  |  |  |  |
|  |  | software is |  |  |  |  |  |
|  |  | available |  |  |  |  |  |
| 02 | Home Page | Check the | If it | We | The | High | High |
|  |  | deployment | doesn’t | cannot | application |  |  |
|  |  | environmen | load. | access | is running |  |  |
|  |  | t for |  | the | successfully |  |  |
|  |  | properly |  | applicati | . |  |  |
|  |  | loading the |  | on. |  |  |  |
|  |  | application. |  |  |  |  |  |
| 03 | User | Verify the | If it | We | The | High | High |
|  | Mode | working of | doesn’t | cannot | application |  |  |
|  |  | the | Respond | use the | displays the |  |  |
|  |  | application |  | Freestyle | Freestyle |  |  |
|  |  | in freestyle |  | mode. | Page |  |  |
|  |  | mode |  |  |  |  |  |
| 04 | Data Input | Verify if the | If it fails | We | The | High | High |
|  |  | application | to take the | cannot | application |  |  |
|  |  | takes input | input or | proceed | updates the |  |  |
|  |  | and updates | store in | further | input to application |  |  |
|  |  |  | The |  |  |  |  |
|  |  |  | Database |  |  |  |  |

# 7.RESULTS AND DISCUSSIONS

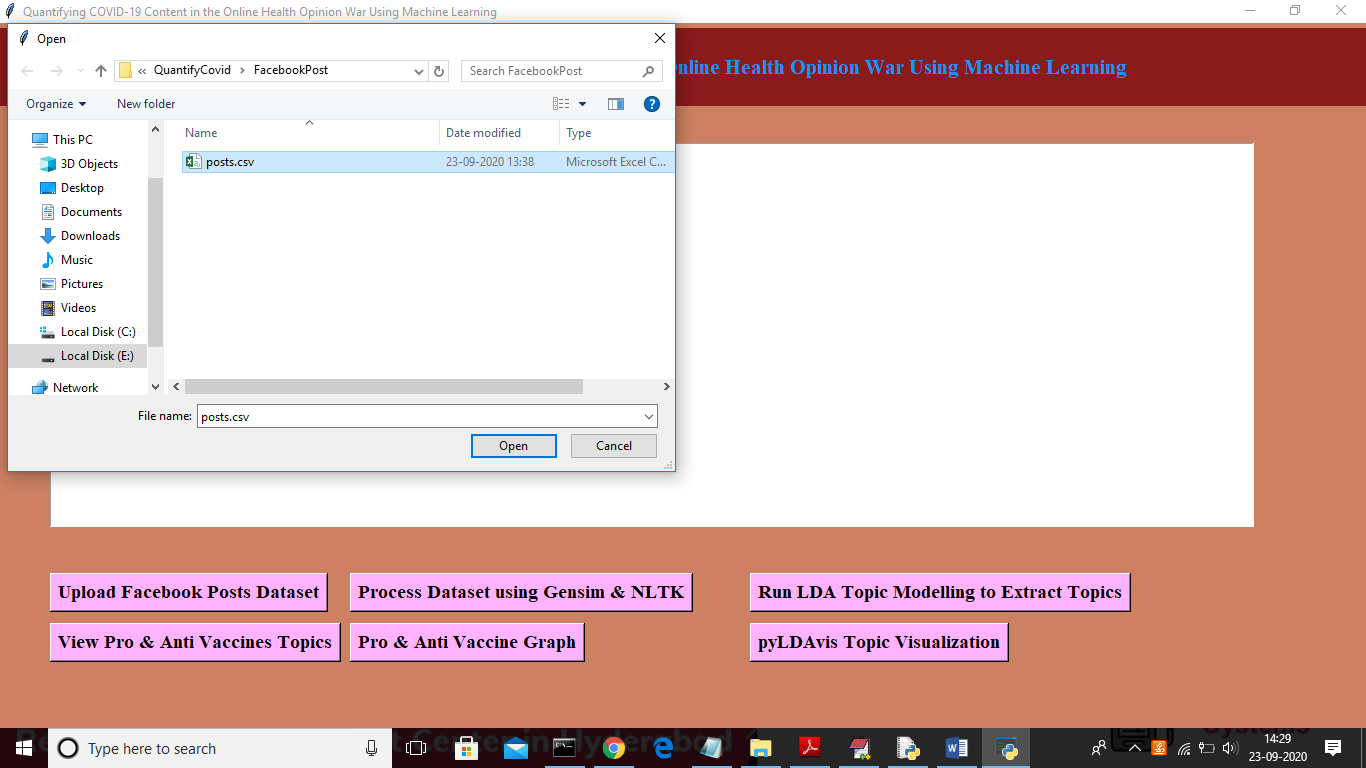


Screen shots

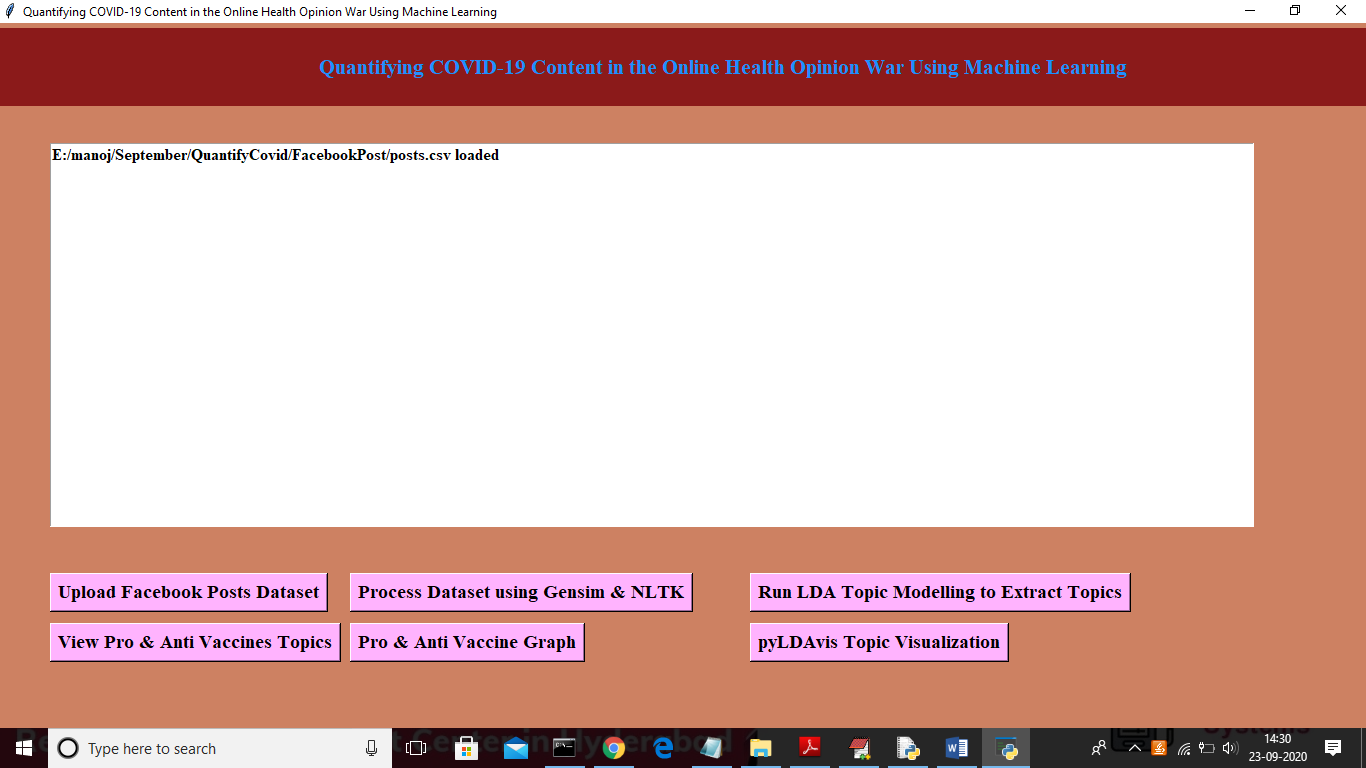
To run project double click on ‘run.bat’ file to get below screen



In above screen click on ‘Upload Facebook Posts Dataset’ button to load dataset



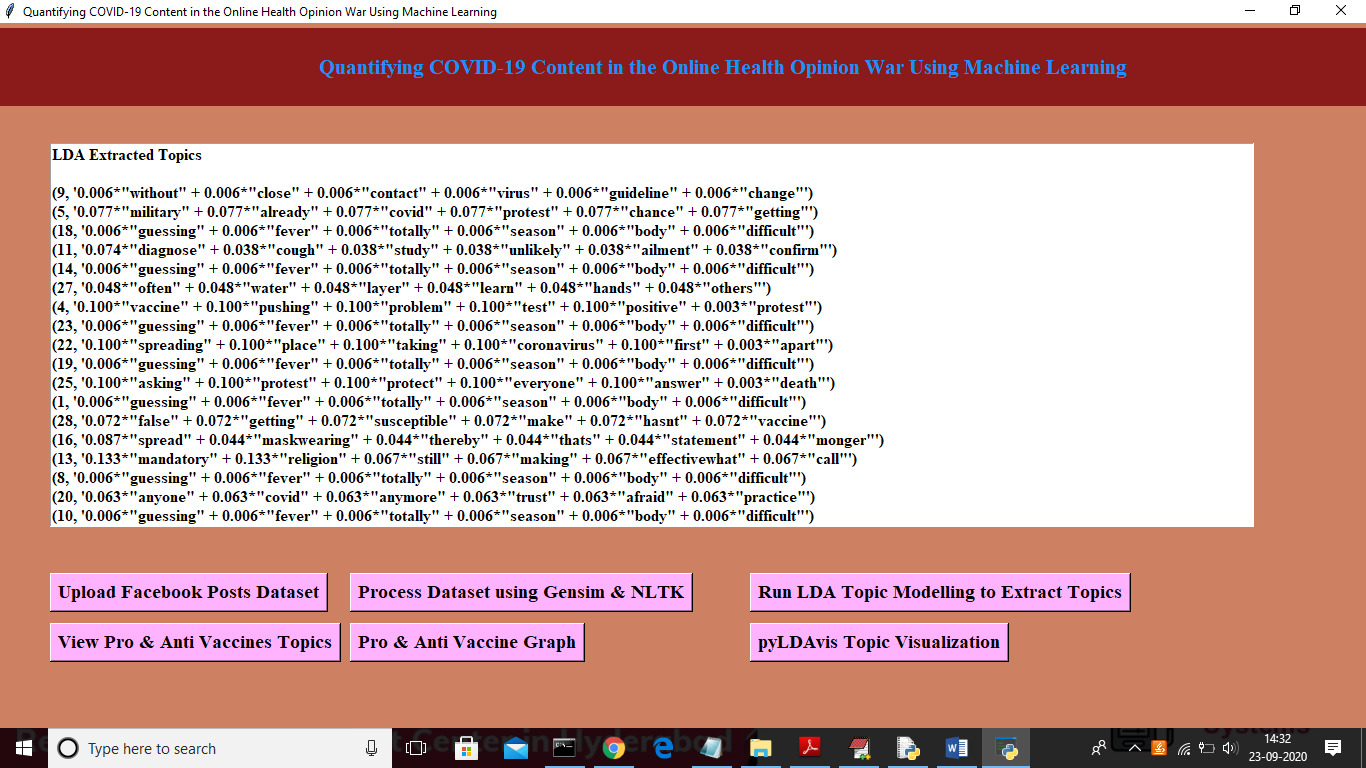
In above screen uploading ‘Posts.csv’ file and then click on ‘Open’ button to load dataset and after loading dataset will get below screen



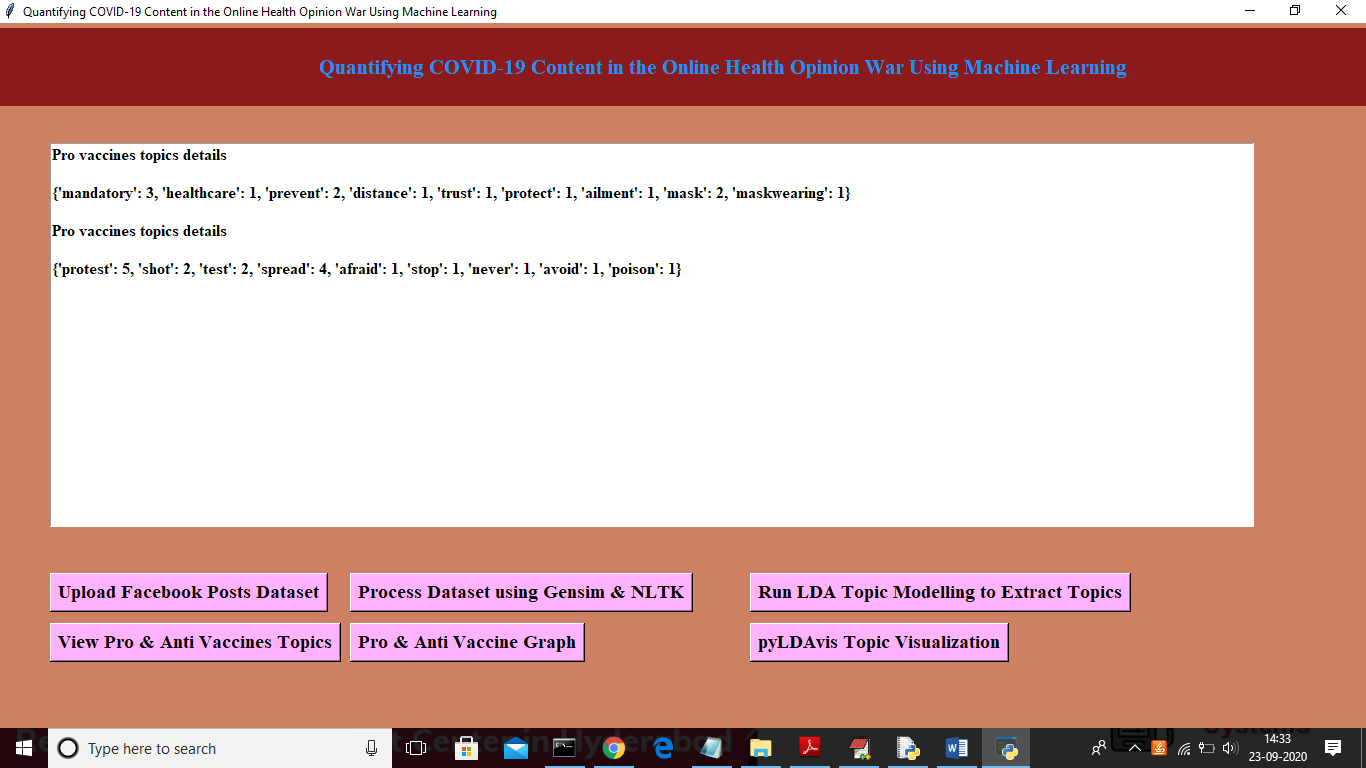
Now click on ‘Process Dataset using Gensim & NLTK’ button to read dataset and to clean dataset



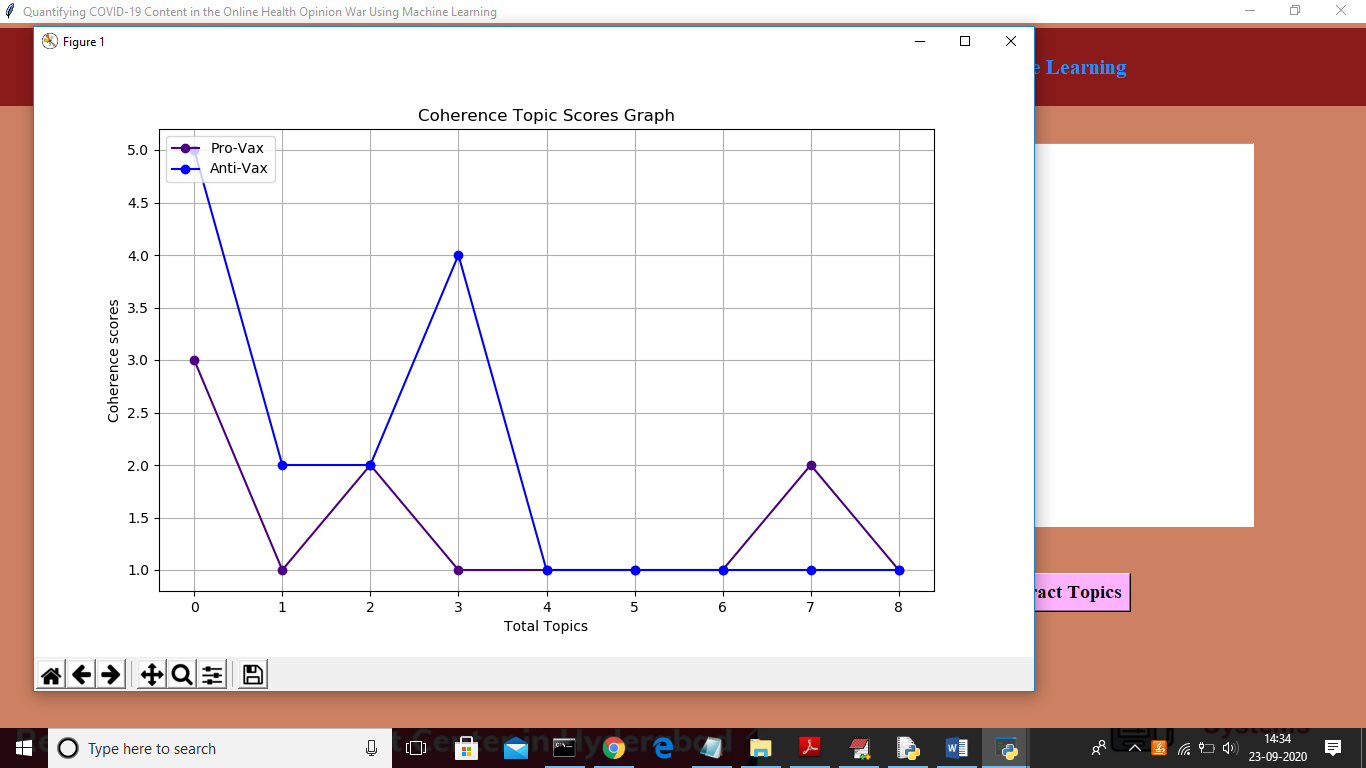
In above screen after cleaning will get above text from all posts and now click on ‘Run LDA Topic Modelling to Extract Topics’ button to extract topics from all posts



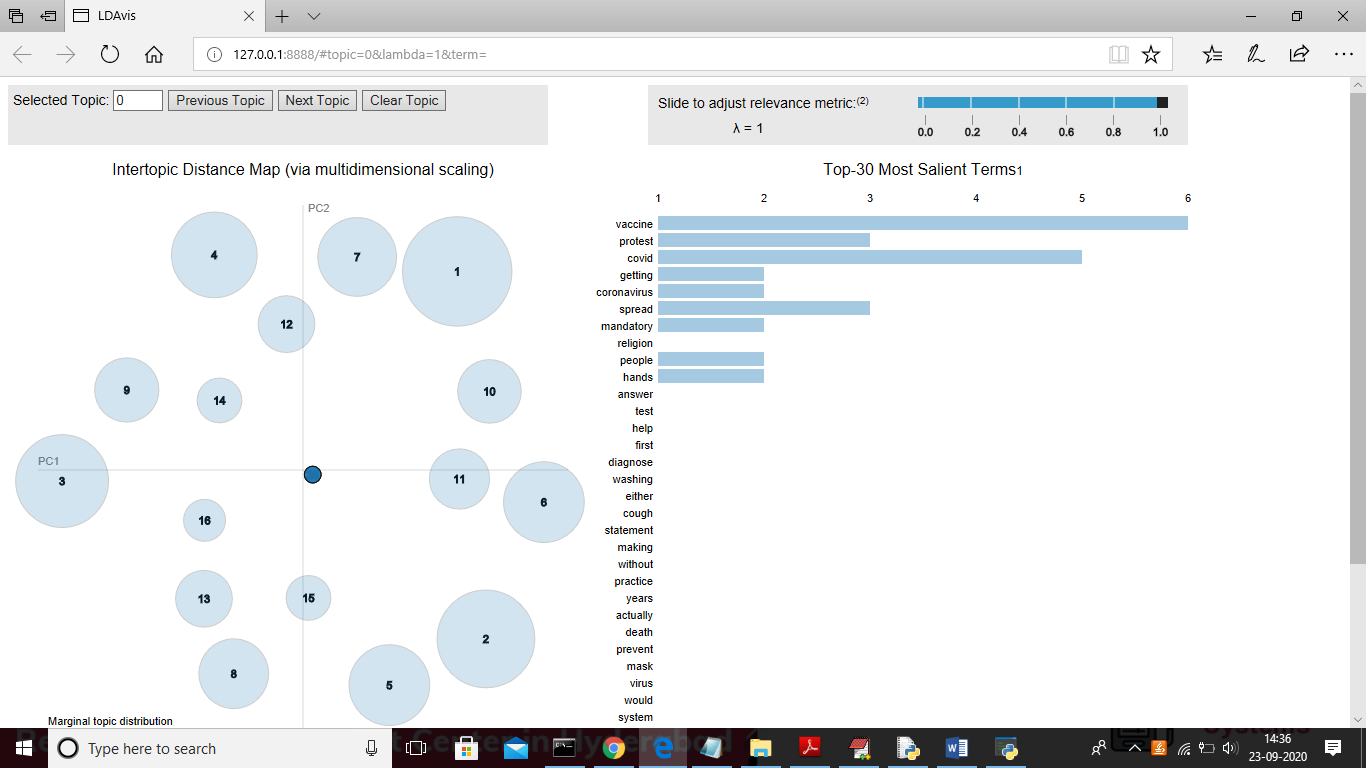
In above screen we can see TOPIC and its probability score for each word from POSTS and now click on ‘View Pro & Anti Vaccines Topics’ button to view all topics with coherence score



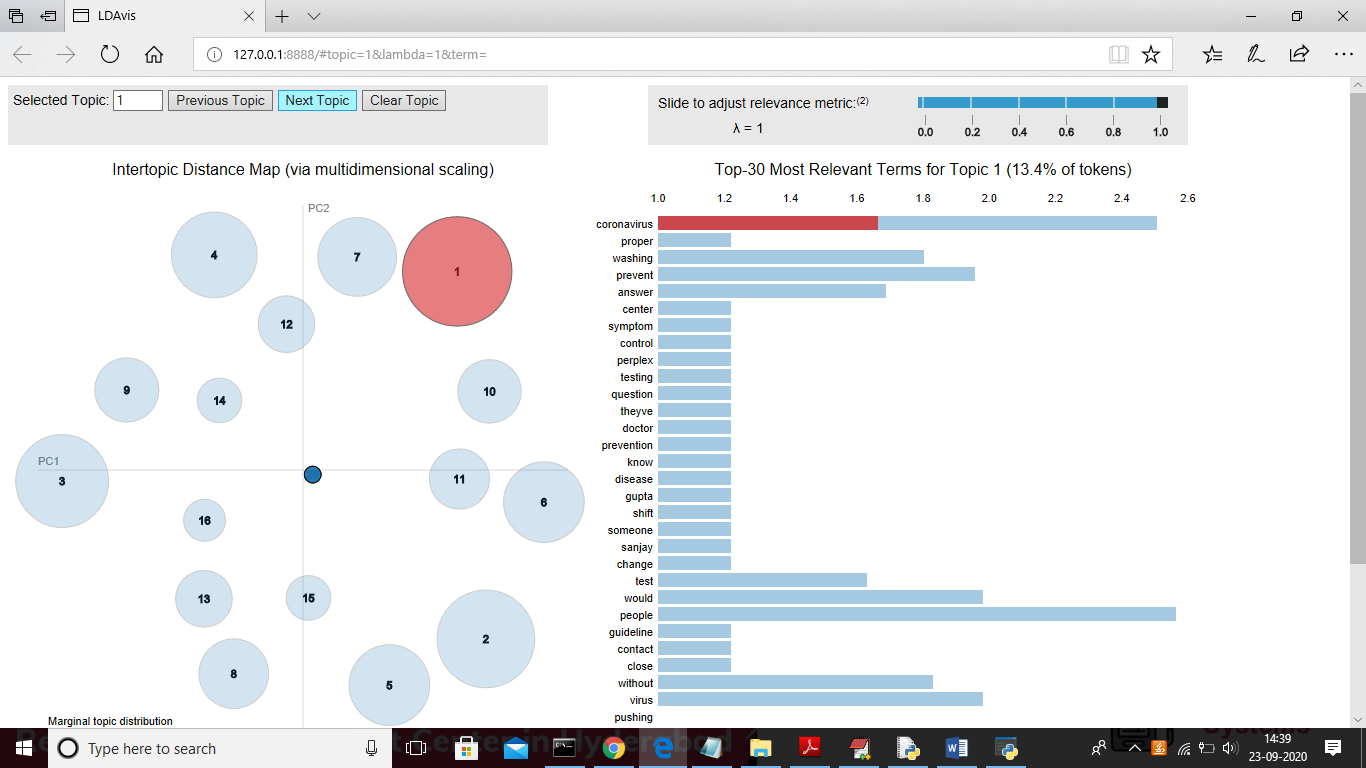
In above screen we can see all topics of PRO and ANTI with count value and now click on ‘Pro & ANTI Vaccine Graph’ button to get below graph and to quantify number of peoples are in favour of PRO or ANTI



In above graph x-axis represents number of topics and y-axis represents coherence score and in above graph blue line refers to ANTI and indigo colour line refers to PRO vaccine and from above graph we can conclude more peoples are discussing ANTI topics about vaccine. Now click on ‘pyLDAvis Topic Visualization’ to get visualization of all topics in browser



In above graph large circle refers that topic occurs more number of time and if occur less number of times then its circle will be in small size. In right size we can see all topics from that posts and u can click on ‘Next Topic’ button from top side of window to get next topic visualization. In above graph each circle represents 1 topic



In above screen for topic 1 coronavirus topic appear more times in all posts. Above graph u can see in browser only

**6. CONCLUSIONS AND FUTURE SCOPE**

# we present a methodology to identify COVID-19 spreaders using the analysis of the relationship between socio-cultural and economic characteristics with the number of infections and deaths caused by the COVID-19 virus in different countries. Using 5-layer multiplex network

# 7. REFERENCES

[1] Thailand Motor Vehicle Registered. CEIC Data Global Database.

[2] Linda, S., Can public transport compete with the private car? Iatss Research, 2003. 27(2): p. 27-35.

[3] CO2 emissions (metric tons per capita) - Thailand. THE WORLD BANK.

[4] Cohen, A.J., et al., Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. The Lancet, 2017. 389(10082): p. 1907-1918.

[5] Litman, T. and D. Burwell, Issues in sustainable transportation. International Journal of Global Environmental Issues, 2006. 6(4): p. 331-347.

[6] Liu, M. and N. Choosri.. A technical solution to improve the red cab for touring in Chiang Mai: Chinese tourists' perspective. in 2016 Chinese Control and Decision Conference (CCDC). 2016. IEEE

[7] Farooq, M.U., A. Shakoor, and A.B. Siddique. GPS based Public Transport Arrival Time Prediction. in 2017 International Conference on Frontiers of Information Technology (FIT). 2017. IEEE.

[8] Bin, Y., Y. Zhongzhen, and Y. Baozhen, Bus arrival time prediction using support vector machines. Journal of Intelligent Transportation Systems, 2006. 10(4): p. 151-158.

[9] Maiti, S., et al. Historical data based real time prediction of vehicle arrival time. in 17th International IEEE Conference on Intelligent Transportation Systems (ITSC). 2014. IEEE.

[10] Fan, W. and Z. Gurmu, Dynamic travel time prediction models for buses using only GPS data. International Journal of Transportation Science and Technology, 2015. 4(4): p. 353-366.