ISSN NO: 2230-5807

Quantifying Covid-19 Content in The Online Health Opinion War Using Machine Learning

Dr.Rohita Yamaganti
Associate professor ,IT Department
Sreenidhi institute of science and Technology
rohita.yamaganti@gmail.com

Dr.GVani
Assistant professor
CSEDepartment
Sreenidhi institute of science and technology
vanig@sreenidhi.edu.in

S Usha Manjari
Assistant professor
IT Department
Sreenidhi institute of science and technology
manjariusha@gmail.com

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 find 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 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 efficacy of intervention strategies. Our approach is scalable and hence tackles the urgent problem facing social media platforms of having to analyse huge volumes of online health misinformation and disinformation.

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 significant 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].

Online social media platforms, and in particular the built-in 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 conflict 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]..

The existing methods using Most models assume a standard SEIR structure.Fraser and colleagues to estimate size but make different changes on the nature of the different compartments and their respective residence timesThe 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.Overall, this approach shows that a machine-learning algorithm, the LDA algorithm, identifies 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 emergequickly using statistical grouping techniques, instead of having to rely on potentially biased, slow and costly human labelling..

LITERATURESURVEY

The Internet plays a large role in disseminating anti-vaccination information. This paper builds upon previous research by analysing 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 analysed; misinformation was also prevalent. The most commonly proposed method of combating this misinformation 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 re-emergence 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].

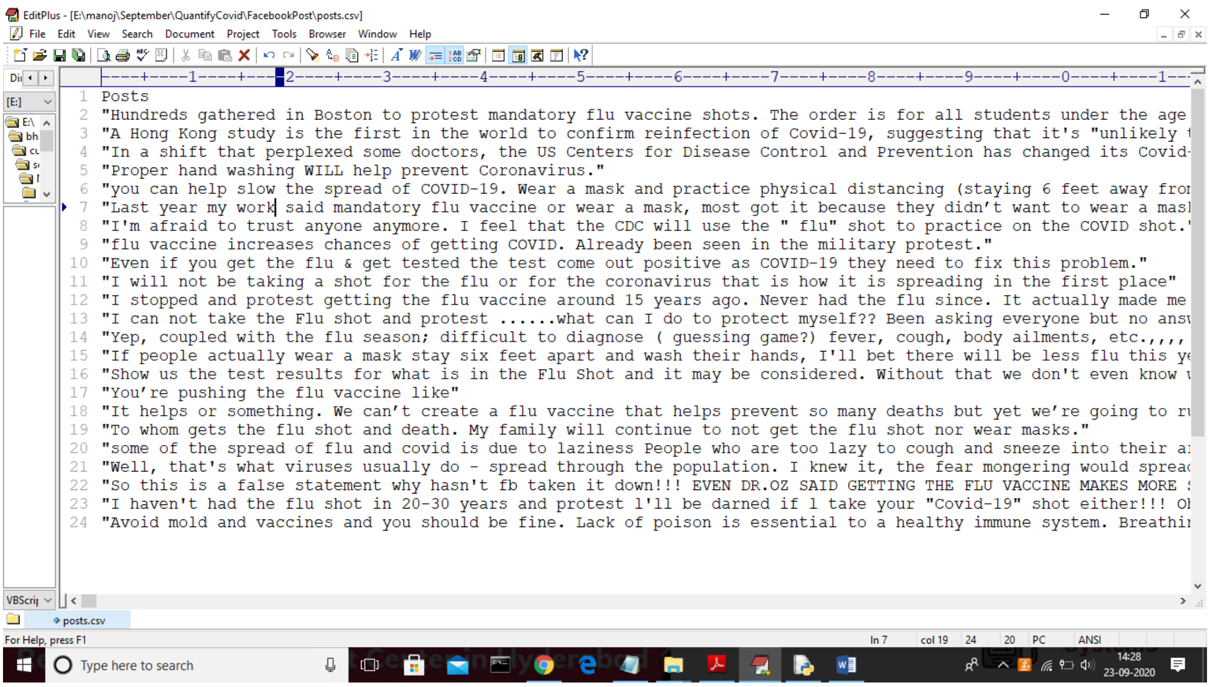
This demonstrates the importance of understanding what messages are presented online and why they may be accepted. The body of research examining online anti-vaccinations is not large, nor has there been a recent update [11,13–18]. Only one analysis [13] examined misinformation and deception on such 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 analysed.

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.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 analysed. 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:

RESULTS ANDDISCUSSIONS

Here we use machine learning to quantify COVID-19 content among online opponents of establishment health guidance, in particular vaccinations (``anti-vax"). We find 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 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 efficacy of intervention strategies

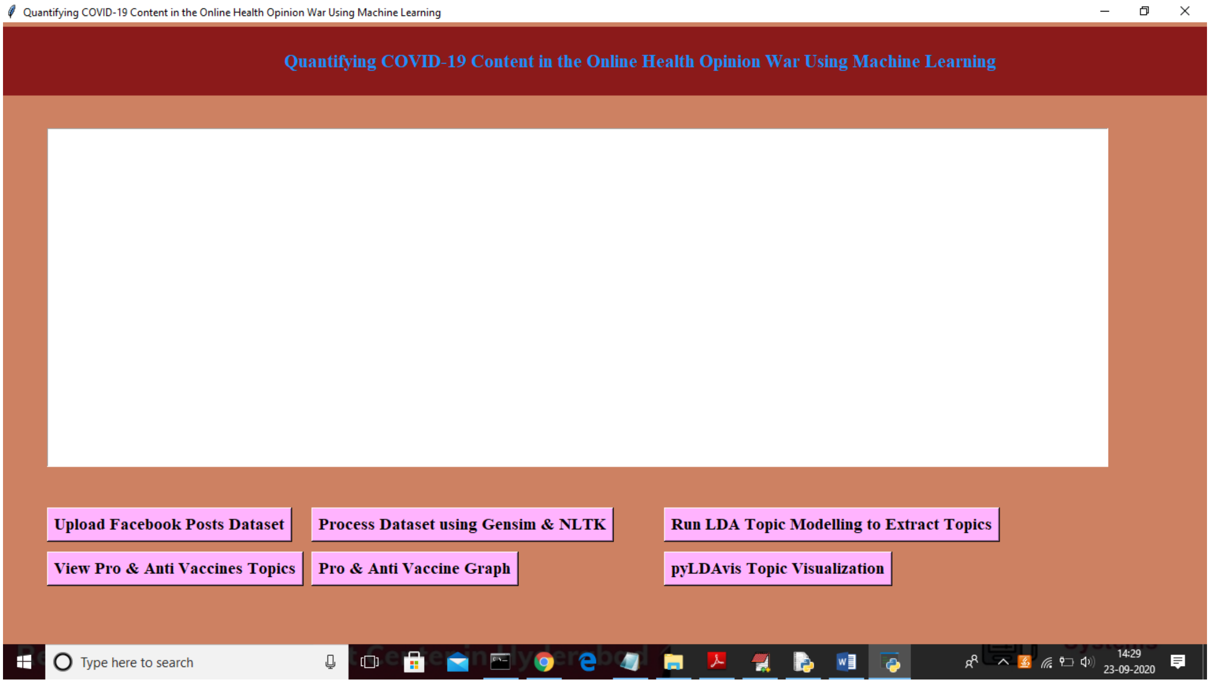


Screen shots

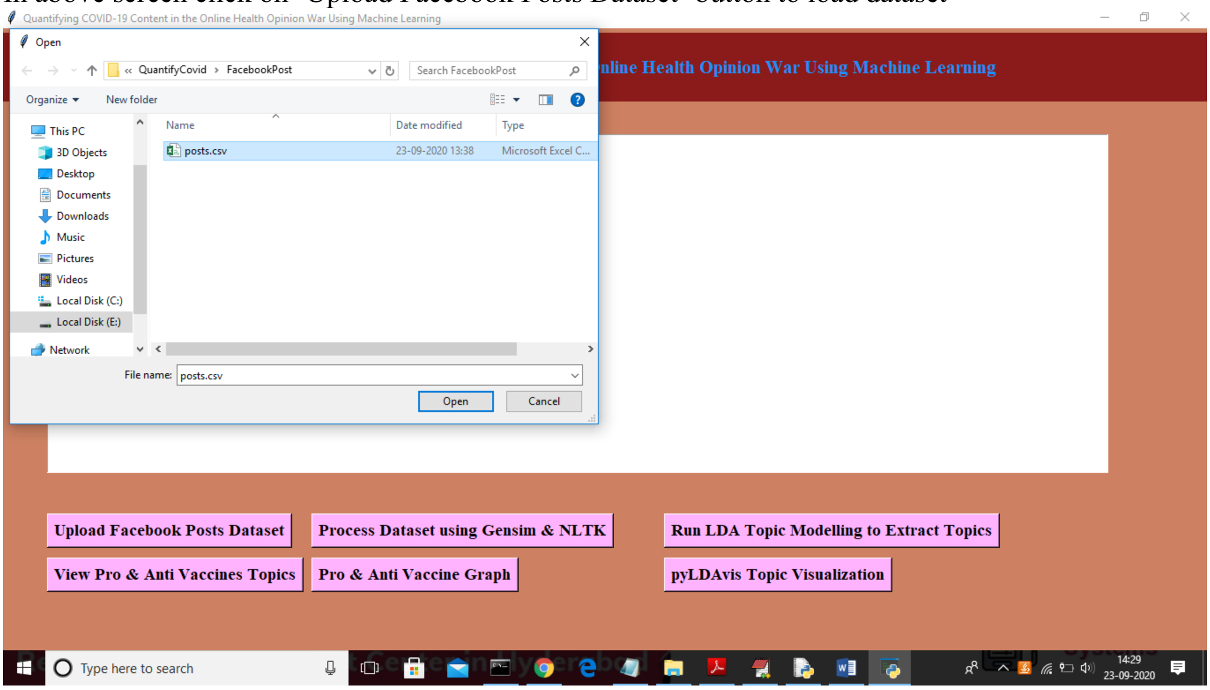
To run project double click on 'run.bat' file to get below screen

BioGecko Vol 11 Issue 04 2022

ISSN NO: 2230-5807

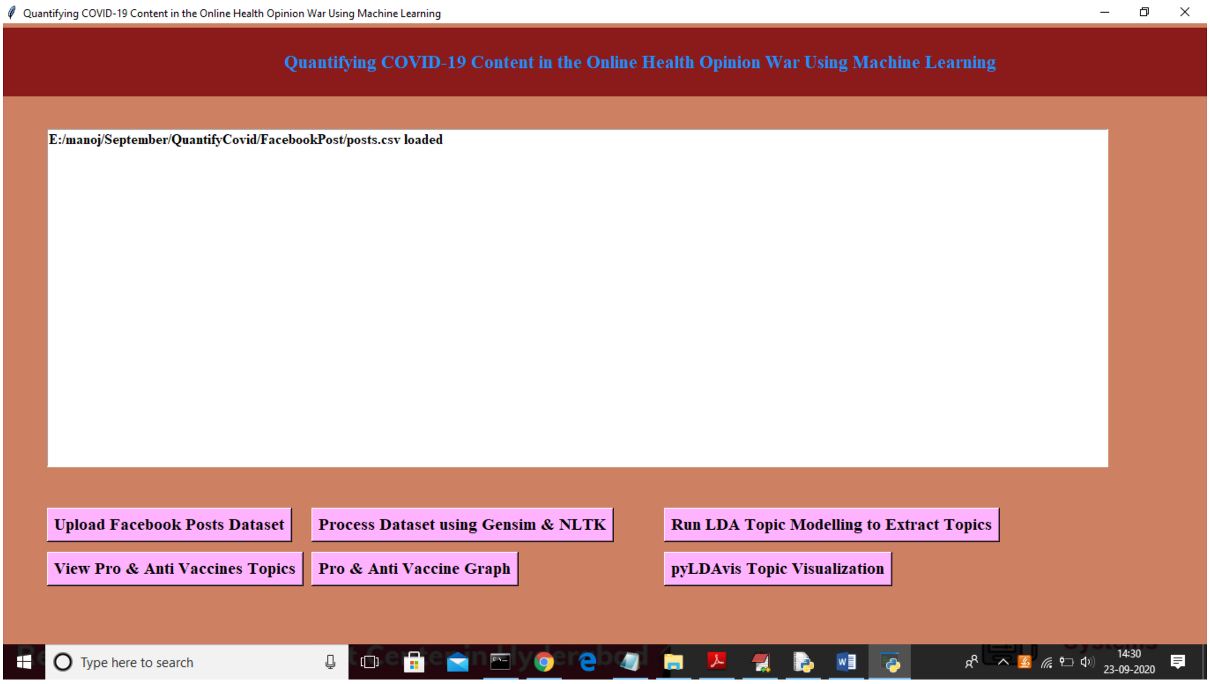


In above screen click on 'Upload Facebook Posts Dataset' button to load dataset



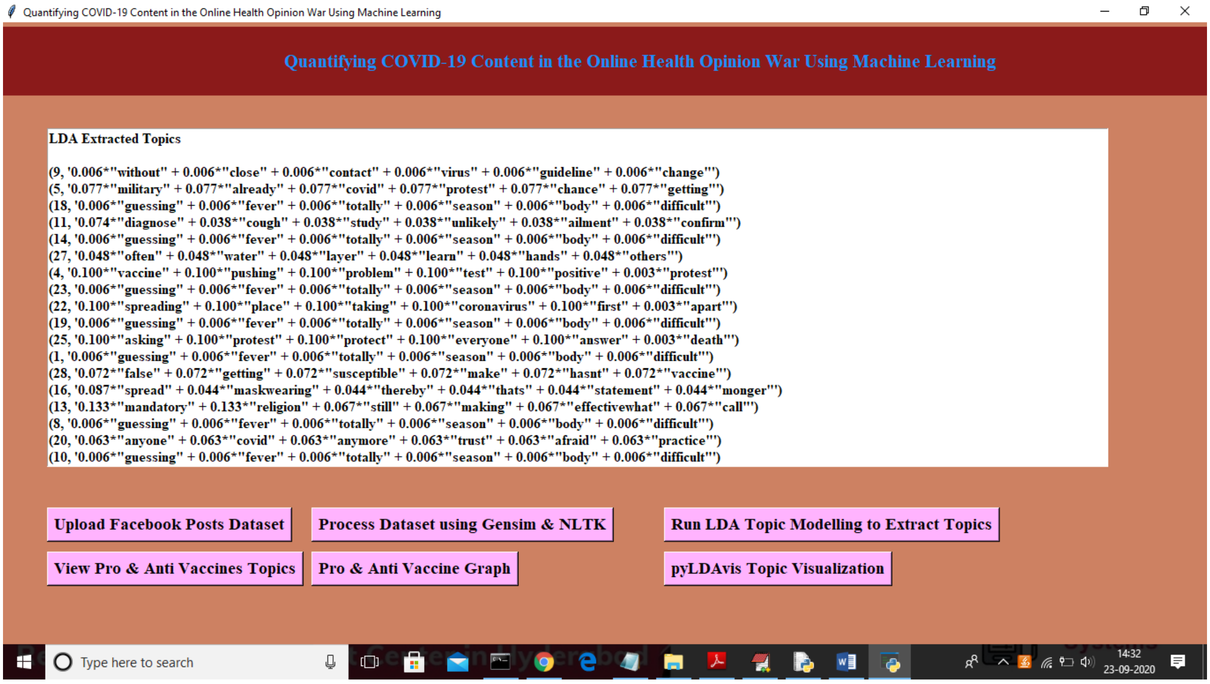
In above screen uploading 'Posts.csv' file and then click on 'Open' button to load dataset and after loading dataset will get below screen

ISSN NO: 2230-5807

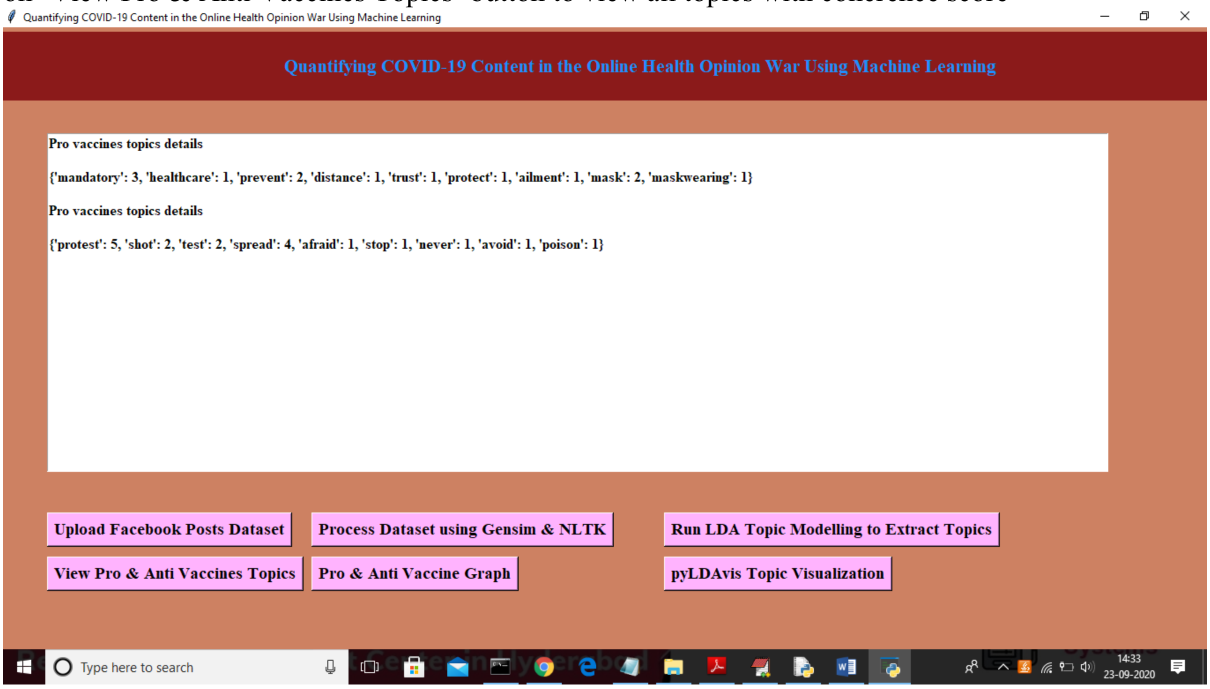


Now click on 'Process Dataset using Gensim& NLTK' button to read dataset and to clean dataset

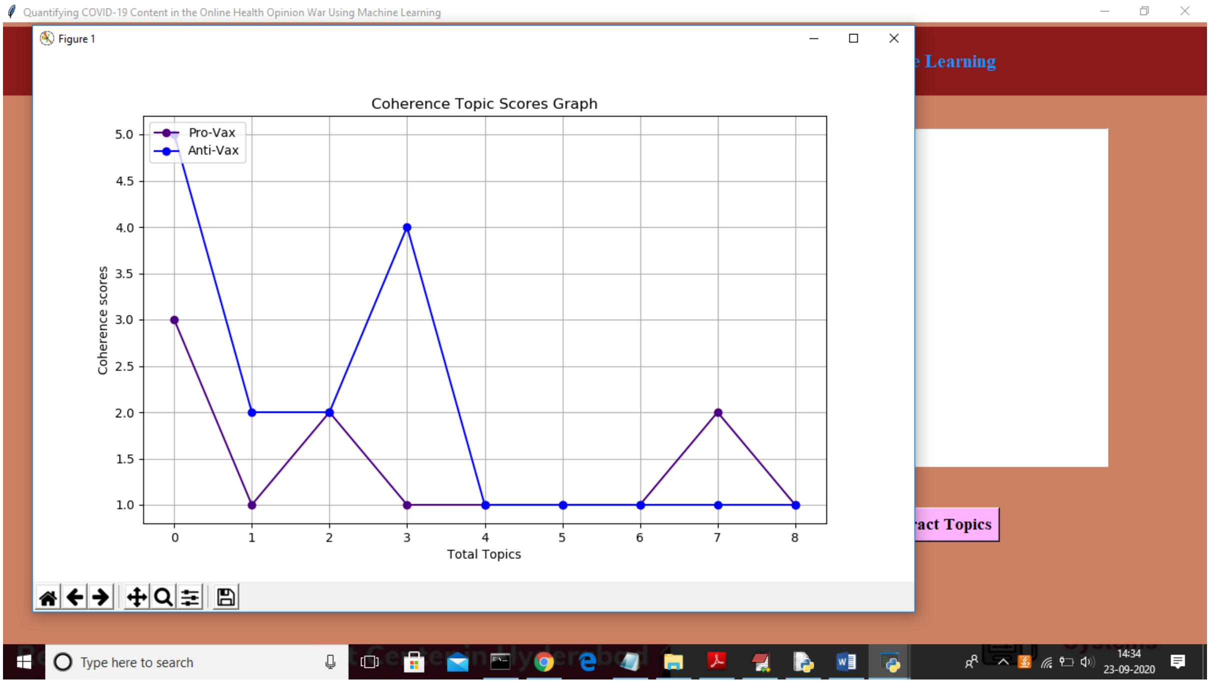
In above screen after cleaning will get above text from all posts and now click on 'Run LDA Topic Modelling to Extract Topics' button to extract topics from all posts



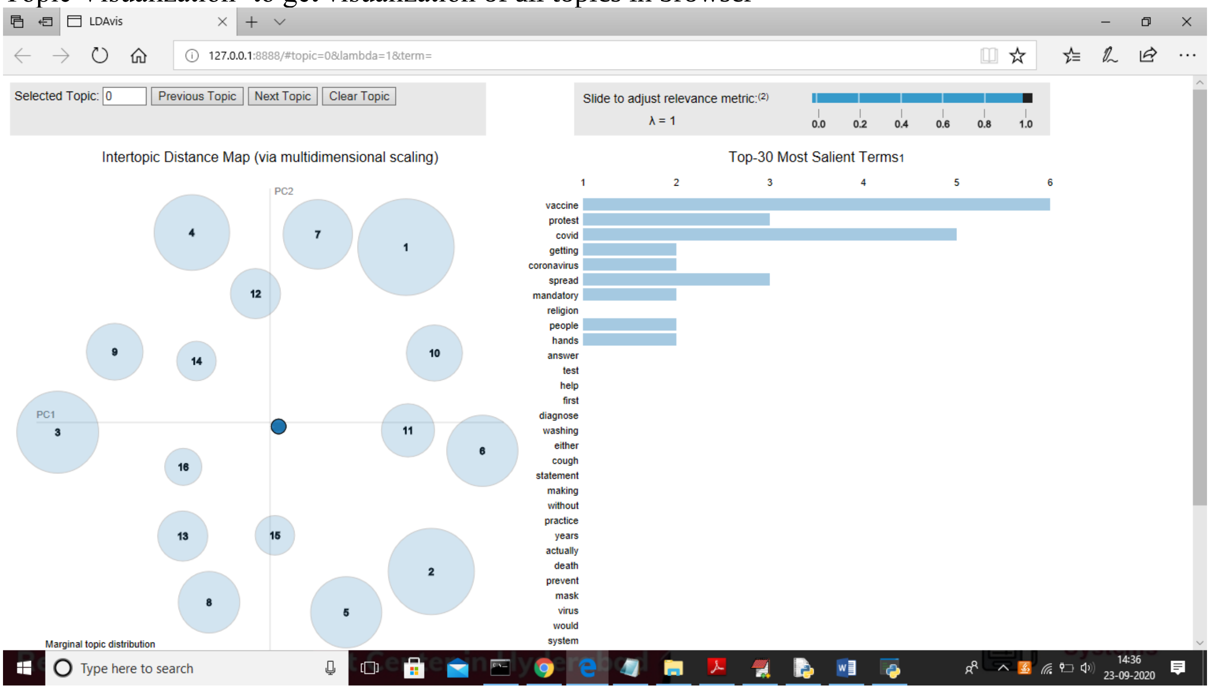
In above screen we can see TOPIC and its probability score for each word from POSTS and now click on 'View Pro & Anti Vaccines Topics' button to view all topics with coherence score



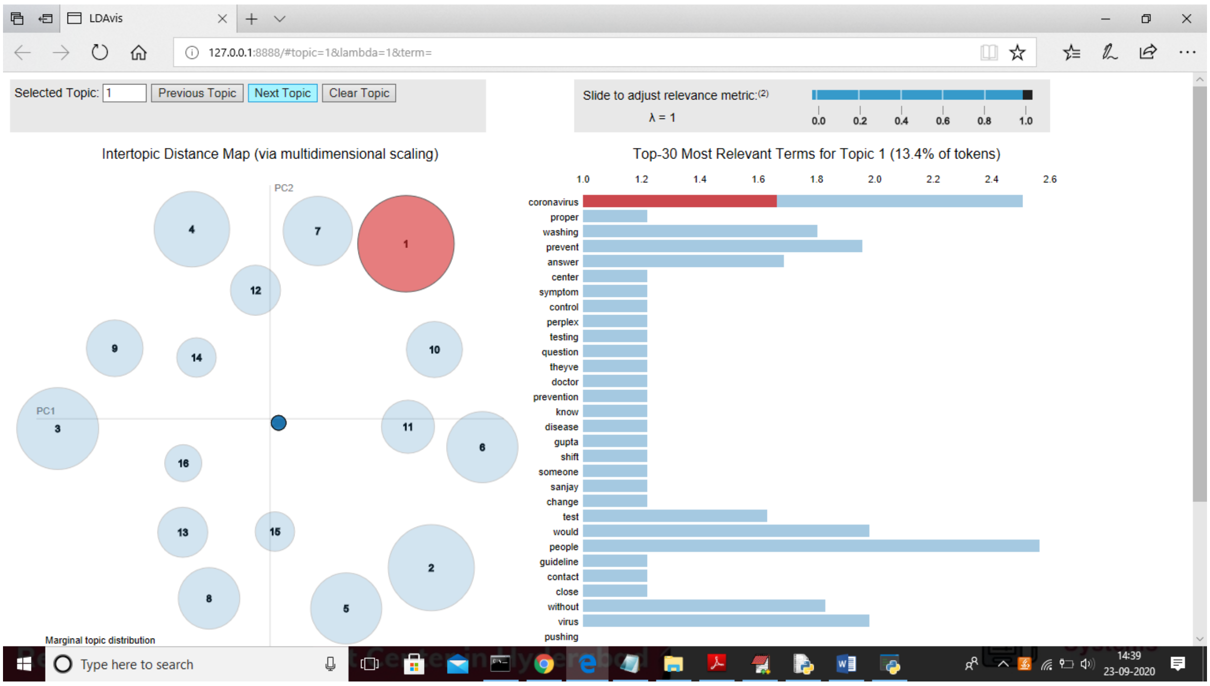
In above screen we can see all topics of PRO and ANTI with count value and now click on 'Pro & ANTI Vaccine Graph' button to get below graph and to quantify number of peoples are in favour of PRO or ANTI



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

CONCLUSIONS

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

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.