



Machine Learning based Healthcare Chatbot

Hussana Johar R B^{1*}, Darshan R V², Advait P R², Venkatesha Munisamy²

¹Department of CSE (AI & ML) ATME College of Engineering, Mysore Karnataka, India.

²Department of ECE, Sai Vidya Institute of Technology, Bangalore, India.



ABSTRACT

Healthcare is essential to lead a good life. However, it is rather difficult to procure the consultation with the physician for each and every well-being problem. As we all know the world is suffering from a pandemic situation still. Our GDP gradually decreased, as of now it has web application but with less functionality to spread awareness of COVID-19 and to connect essential COVID-19-related health services to the people. The thought is to create a medical chatbot (Viro-attack) using Artificial Intelligence that may diagnose the disease and provide fundamental information about the illness sooner than consulting a health care provider. This will lend a hand to scale back healthcare prices and strengthen accessibility to medical knowledge thru scientific chatbot. This Viro-attack web application helps us to get updates related to COVID-19 and it spreads awareness about Corona virus in this web application. It also provides a vaccine related updates and current active cases in nearest locality of the user. And this web application will be containing all the COVID-19 related information and the user can also get their health status in the web application by entering their current symptoms if they have any. The Viro-attack stores the information in the database to spot the sentence keywords and to make a question resolution and answer the query. Ranking and sentence similarity calculation is performed the use of n-gram, TFIDF and cosine similarity. The score might be obtained for every sentence from the given input sentence and more equivalent sentences can be acquired for the query given. The third celebration, the skilled program, handles the question introduced to the bot that's not understood or isn't providing in the database.

Keywords: Interactive Voice Response (IVR), unstructured supplementary service data (USSD), syndrome-coronavirus 2 (SARS-CoV-2)

1. Introduction

In this COVID 19 tracking and alert system is used for. The prototype would be a software with web and mobile components. Using these platforms, information dissemination, disease prevalence and position tracking of carriers, confirmed carriers and status of treated patients could be easily managed. Every disease irrespective of the infectious agent presents a challenge especially when it is novel like COVID-19. Like other infectious diseases the corona virus present some symptoms in the infected patient. However, in some individuals, the disease is asymptomatic thus posing a special concern as it could spread unnoticed through the droplets of saliva or discharge from the nose, mouth, eyes, or other body cavities of asymptomatic patients. The incubation period of the virus is between 1-14 days within which there may be visible symptoms. Thus, the affected person may be living with the virus with or without symptoms. The most common symptoms of COVID-19 are fever, tiredness, and dry cough. The symptoms indicate the level of

infection, which ranges from mild infection, severe infection, and critical infection. The symptoms are well-documented and include: Technology has played a significant role in the detection, prevention and control of public health problems. Sophisticated evolutionary technologies have been applied to various areas of health care delivery. Notable systems include clinical decision support systems, expert systems, electronic health systems, to mentioned. Having regard to the foregoing, this work proposes a full tracking system to augment the activities of public health workers and security agencies in tracking cases from the point of entry and association with cases. The ongoing coronavirus disease 2019 (COVID-19) pandemic has overwhelmed the healthcare systems of countries around the world, exposing the challenges faced by public health agencies when responding to rapidly emerging outbreaks. In particular, the scarcity of reliable data on the incidence of COVID-19 cases has hindered a timely response. On a national scale, control efforts should be guided by accurate data on cases and disease burden, ideally captured through widespread surveillance. However, very few countries affected by COVID 19 have sufficient viral testing capacity to monitor cases occurring in the community adequately. Hospitalization and death rates provide relatively robust indicators of SARS-CoV-2 transmission in some areas, but these are lagged by Page | 2 about 2 and 3 weeks, respectively. Identifying alternative indicators of transmission that reflect the timing of new infections is therefore an important priority for responding to the epidemic. The first COVID-19 confirmed case occurred in Bangladesh on March 8th, with nearly 200,000 confirmed cases by July 15. SARS-CoV-2 testing capacity has increased significantly from a daily average of fewer than 100 tests in March to about 15,000 in June. However, as in most countries, the testing capacity can only cover a small fraction of even

Received 08 January 2023

Revised 05 February 2023

Accepted 04 March 2023

Available Online 14 March 2023

Corresponding Authors: **Hussana Johar R B**, (husnavais@gmail.com)

Darshan R V, (darshan.rv@saividya.ac.in)

Advait P R, (advait.p@saividya.ac.in)

Venkatesha Munisamy, (venkatesha.m@saividya.ac.in)

<https://doi.org/10.51470/eSL.2023.4.4.50>

© 2023 by the authors. The license of Journal of e-Science Letters. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

symptomatic cases. Reporting delays in rural and remote parts of the country also make it difficult to monitor the epidemic across the country in real-time. To augment surveillance, a participatory surveillance system based on self reported symptoms via national telephone hotlines and the internet, assisted by a telemedicine team of clinicians, was deployed in March and rapidly scaled up over the course of the first few months of the outbreak. The participatory surveillance system was set up through a public-private partnership, and is designed to collect syndromic information, to identify potential disease hotspots, and to provide information about COVID-19 to participants. Any surveillance data that relies on self reported symptoms to monitor transmission will be subject to a range of biases, including the extent to which people are aware of and know how to use the system, and reporting behavior of people in the middle of a pandemic, which has naturally created much fear and uncertainty. Given the lack of specificity of the main symptoms of COVID-19, namely fever and cough, we also expect many people experiencing symptoms to have another disease unrelated to the coronavirus outbreak. Nevertheless, an uptick in individuals reporting symptoms consistent with COVID-19, particularly if verified through an interview with a clinician, may provide important insights into transmission hotspots. While participatory crowdsourced syndromic surveillance has been utilized in many contexts [1]-[7], including for COVID-19 [8], [9], their ability to track an emerging outbreak at a high spatial resolution has not been evaluated previously. Page | 3 Here, we show that one such system, though noisy, provides an indication of where and when to expect new cases, suggesting that it could be a useful model in other places that need to map COVID-19 risk for decision making. The syndromic data suggests that the outbreak had spread across the country much faster than is evident from official case counts, consistent with geographic spread based on population mobility data. The rest of the paper has the proposed system in section II followed by the flow chart and results in section III and conclusion in section IV.

2 Proposed System

This user interface is used to get all info about COVID-19 cases in one single platform where we are providing a user risk scan to get follow of there covid-19 symptoms and go through check out there health status also. The user can also check out the covid-19 real time tracking of cases of every country and every locality and here we also spread awareness about coronavirus to user's. Technology has played a significant role in the detection, prevention and control of public health problems. Sophisticated evolutionary technologies have been applied to various areas of health care delivery. Notable systems include clinical decision support systems, expert systems, electronic health systems, to mention but few. To a great extent, computing and information technologies have demystify diagnosis and management of complex medical cases as they are employed at various levels ranging from information gathering, documentation, intelligent insights to accurate decision making. Modern ICTs readily augment human expertise in several ways, such as: system-enabled diagnosis, disease management, drug administration, expert prognosis, etc.

The object-oriented approach was strictly adopted owing to the ease at which the system problem domain could be decomposed into participating components. System extensibility was also considered as against the limitations of some methodologies. The procedures adopted included: general survey and documentation of covid19 cases and public health technology requirements in our Regional places; conceptualization of an automated multi-platform model with requirements, actors,

inputs, processing and outputs defined. To produce valid specifications of the proposed system, the following object oriented approach components were employed: use cases, class diagrams, and component diagrams respectively.

In March 2020, the World Health Organization (WHO) declared COVID-19 a pandemic, caused by the novel SARS-CoV-2 virus. Following the call from the WHO to immediately assess available data to learn what care approaches are most effective and evaluate the effects of therapies, this collection aims to report on original peer-reviewed research articles in methodological approaches to medical research related to COVID-19. The first confirmed case of COVID-19 occurred in the United States (US) in Washington State on January 20, 2020. [1] Non-pharmaceutical interventions, such as quarantines and mass social distancing, were the primary public health strategy for blunting COVID-19 spread. As confirmed case counts climbed, state, county, and municipal governments adopted policies recommending or requiring actions to reduce social density and slow the progression of the outbreak. The timing and intensity of social distancing policy responses has varied. Multiple efforts sought to rapidly code these social distancing policy responses for analysis [2,3,4,5,6,7,8,9,10,11]. Social distancing policy coding has been critical to COVID-19 disease models that have influenced policy decision-making whether to impose or ease social distancing approaches. For example, Dr. Deborah Birs, the U.S. Coronavirus Response Coordinator, has repeatedly cited the COVID-19 projections prepared by the University of Washington's Institute for Health Metrics and Evaluation (IHME).

However, the methods used by various modeling efforts for linking COVID policies to projected outcomes (e.g., rates of infection, hospitalization, or death), have been quite divergent. Social distancing coding efforts have used a range of methodologies and frameworks to characterize and code policy responses, resulting in a diversity of social distancing policy taxonomies and classification schemes. These efforts have characterized social distancing into taxonomies consisting of 1 domain (e.g., stay at home order in place) to upwards of six domains. For example, McGrail et al. used a single domain of "lockdown," coded as the date the lockdown started and ending as of the date non-essential retail stores re-opened. In contrast, Adolph et al., which developed one of the first publicly available datasets, developed a COVID-19 framework originally consisting of five domains: (1) recommendations or restrictions on gatherings, regardless of the size of gathering; (2) K12 school closures; (3) restaurant restrictions on in-person dining; (4) non-essential business closures; and (5) mandatory stay-at-home orders [2].

The block diagram of implemented AI based Chabot system is as shown in Figure 1.

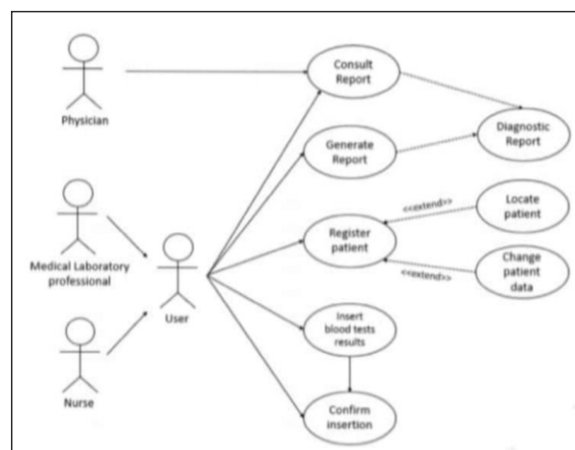
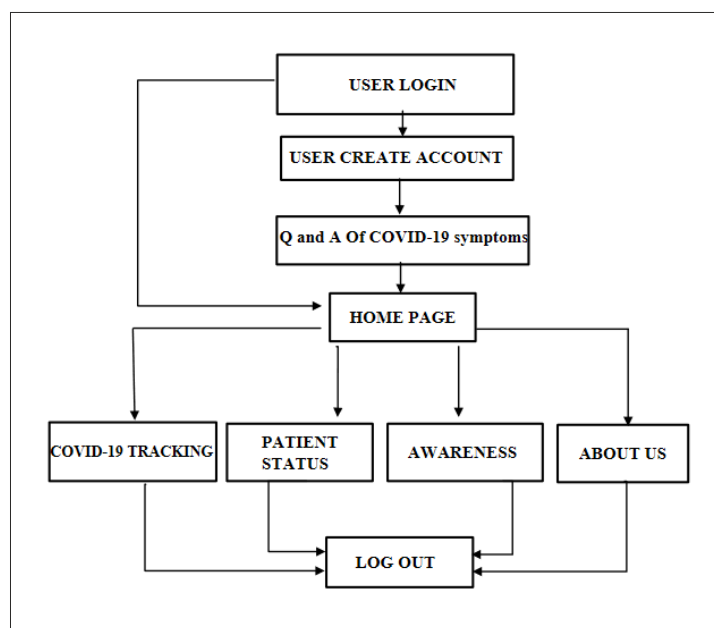


Figure 1. Block Diagram of AI based Chabot system

2.1 FLOW CHART



Coronavirus pandemic is spreading in large numbers. Experts suggest that social distancing has been used for a long time as one of the methods to curb or reduce the spike in diseases and infectious illnesses. Thus, apps and innovative solutions such as these, which promote the same idea can help authorities make the population aware and save lives. The app, which is a coronavirus tracker of sorts works on the basis of contact. Tracing and can help a user identify possible coronavirus 'hotspot' around his or her area. It can help people stay safe and adopt necessary precaution in some areas where there are cases and accordingly, help stop or prevent community transmission to an extent. By the basis of geotagging, it can also alert a specific user about their proximity to a nearby infection case or hotspot. The app also helps users self-identify their risk and monitor their health assessment, considering the times when it can get difficult (and most of all, is not particularly safe to step out and visit health clinics). Aarogya Setu app also helps people identify the symptoms, alert them about the best safety precautions and other relevant information concerning the spread of COVID-19. While this is a noble initiative, the app also lists down basic quarantine measures for those who are considered to be in the high-risk' category. It can also help people, who have had a travel history self-quarantine and prevent any risk of transmission.

The following steps to be followed as functioning of the entire process:

- **Login/Sign-Up Module:** This module enables all actors to create an account and access their accounts on the proposed system.
- **Free end module:** it is the end where user can visit the site to see the present statistics of every region of a country and also get an information related to COVID-19.
- **Tracking Module:** This module helps the officers to track the geographical positions of the cases when the need arises. The module will map the code assigned earlier to the case.
- **Information Module:** The information module provides a guest with a platform to interact with the system. The guest may request or report some issues and can also obtain relevant information on transmission, distribution, etc.

- The above-mentioned site is using state bulletins and official handles to update our numbers. The data is validated by a group of volunteers and published into a Google sheet and an APL.
- API is available for all at api.covid19india.org. there main intention is to we can use this data in the fight against this virus.
- The interface was created by a group of dedicated volunteers who curate and verify the data coming from several sources.
- We extract the details, like a patient's relationship with other patients to identify local and community transmissions, travel history and status.

3 Results

The system working is simulated using JavaScript, jQuery on API. Simulation experiments have been conducted for the different specifications using AI based Chatbot system.

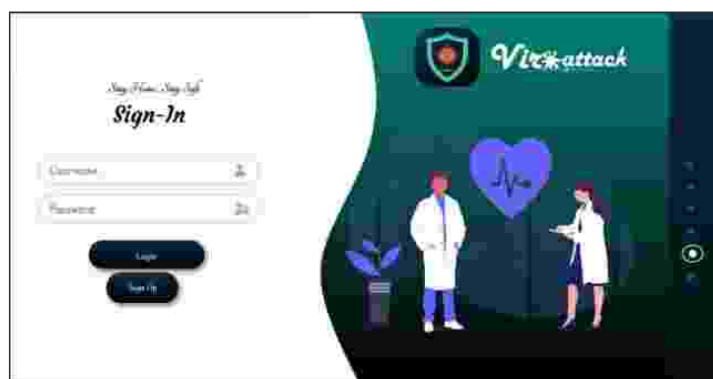


Figure 2: Login Page

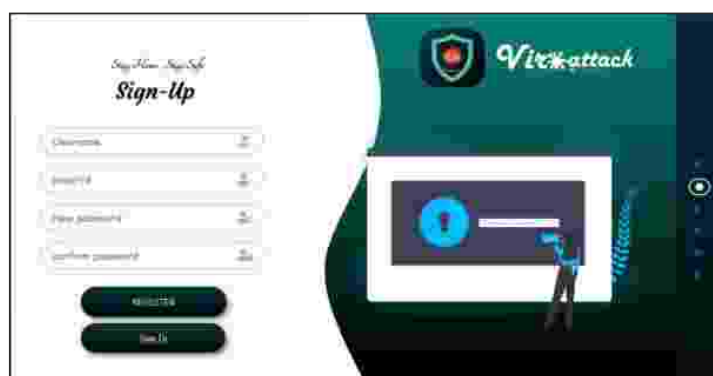


Figure 3: Signup Page



Figure 3: Home Page



Figure 4: Corona Update Page

Chatbot

4 Conclusion

The software of chatbot within the clinical domain is slightly means beyond then our imaginations. This work have been covered nearly all of the points which a clinical chatbot will have to fortify to cater the desire of the affected person. In the past few years there were lot of medical chatbot models has been invented which were rather expensive for a standard particular person however this paper attempted to triumph over this downside in our 'well-being care chatbot machine. The increased real-world submission of the chatbot and imposing that for more domain names can further overview this chatbot framework.

References

1. S. Bansal, G. Chowell, L. Simonsen, A. Vespignani, and C. Viboud, "Big Data for Infectious Disease Surveillance and Modeling," *J. Infect. Dis.*, vol. 214, no. Suppl 4, pp. 5375-5379, Dec. 2016, doi: 10.1093/infdis/jiw-400.
2. M. Salathé, "Digital Pharmacovigilance and Disease Surveillance: Combining Traditional and Big Data Systems for Better Public Health," *J. Infect. Dis.*, vol. 214, no. Suppl 4, pp. 5399-5403, Dec. 2016, doi: 10.1093/infdis/jiw281.
3. L. Simonsen, J. R. Gog, D. Olson, and C. Viboud, "Infectious Disease Surveillance in the Big Data Era: Towards Faster and Locally Relevant Systems," *J. Infect. Dis.*, vol. 214, no. suppl_4, pp. 5380-5385, 01 2016, doi: 10.1093/infdis/jiw376.
4. G. E. Smith, D. L Cooper, P. Loveridge, F. Chinemana, E. Gerard, and N. Verlander, "A national syndromic surveillance system for England and Wales using calls to a telephone helpline," *Eurosurveillance*, vol. 11, no. 12, pp. 9-10, Dec. 2006, doi: 10.2807/esm.11.12.00667-en.
5. O. P. Wójcik, J. S. Brownstein, R. Chunara, and M. A. Johansson, "Public health for the people: participatory infectious disease surveillance in the digital age." *Emerg. Themes Epidemiol.*, vol. 11, no. 1, p. 7, Jun. 2014, doi: 10.1186/1742-7622-11-7.
6. D. Olson et al., "Performance of a Mobile Phone App-Based Participatory Syndromic Surveillance System for Acute Febrile Illness and Acute Gastroenteritis in Rural Guatemala," *J. Med. Internet Res.* vol. 19, no. 11, p. e368, 2017, doi: 10.2196/jmir.8041.
7. C. E. Koppeschaar et al., "Influenzanet: Citizens Among 10 Countries Collaborating to Monitor Influenza in Europe," *JMIR Public Health Surveill.*, vol. 3, no. 3, p. e66, 2017, doi: 10.2196/publichealth.7429.
8. M. Gong, L. Liu, X. Sun, Y. Yang, S. Wang, and H. Zhu, "Cloud-Based System for Effective Surveillance and Control of COVID-19: Useful Experiences From Hubei, China," *J. Med. Internet Res.*, vol. 22, no. 4, p. e18948, 2020, doi: 10.2196/18948.
9. A Mehl, F. Bergey, C. Cawley, and A. Gilsdorf, "Syndromic surveillance insights from a symptom assessment app before and during COVID-19 measures in Germany and the United Kingdom: results from repeated cross-sectional analyses," *medRxiv*, Jun. 2021, doi: 10.1101/2020.06.16.20126466.
10. C. T. Lloyd et al., "Global spatio-temporally harmonised datasets for producing high-resolution gridded population distribution datasets," *Big Earth Data*, vol. 3, no. 2, pp. 108-139, Apr. 2019, doi: 10.1080/20964471.2019.1625151.
11. A. S. Mahmud et al., "Megacities as drivers of national outbreaks: the role of holiday travel in the spread of infectious diseases," *bioRxiv*, p. 737379, Aug. 2019, doi: 10.1101/737379.

12. Z. Du et al., "Risk for Transportation of Coronavirus Disease from Wuhan to Other Cities in China - Volume 26, Number 5-May 2022 Emerging Infectious Diseases journal - CDC," doi: 10.3201/eid2605.200146
13. W. B. Habib and A. S. Moudud, "Coronavirus Outbreak: Low test rate may be belying reality." The Daily Star, Apr. 01, 2020. <https://www.thedailystar.net/frontpage/news/coronavirus-outbreak-low-test-rate-maybelying-reality-1888327> (accessed Jul. 20, 2020),
14. s. Moral, "COVID-19 testing being expanded to more labs | Prothom Alo." Mar. 31, 2020. <https://en.prothomalo.com/bangladesh/government/covid-19-testing-beingexpanded-tomore-labs> (accessed Aug. 12, 2020).
15. A. Alif and T. Rahman, "IEDCR not expanding coronavirus testing facilities yet." Dhaka Tribune, Mar. 11, 2020. <https://www.dhakatribune.com/bangladesh/2020/03/11/iedcr-notexpanding-coronavirustesting-facilities-yet> (accessed Aug. 12, 2020).
16. "Coronavirus: Free testing at 17 hospitals, 11 more to laun...", unb.com.bd, Apr. 14, 2020. <https://unb.com.bd/category/Bangladesh/coronavirus-free-testing-at-17-hospitals-11-moreto-launch-soon/49596> (accessed Aug. 12, 2020).
17. D. Paolotti et al., "Web-based participatory surveillance of infectious diseases: the influenzanet participatory surveillance experience," Clin. Microbiol. Infect., vol. 20, no. 1, pp. 17-21, 2014, doi: 10.1111/1469-0691.12477.
18. R. Chunara, E. Goldstein, O. Patterson-Lomba, and J. S. Brownstein, "Estimating influenza attack rates i the United States using a participatory cohort," Sci. Rep., vol. 5, no. 1, p. 9540, Aug. 2015, doi: 10.1038/srep09540.
19. C. Guerrisi et al., "Participatory Syndromic Surveillance of Influenza in Europe," J. Infect. Dis., vol. 214, no. suppl 4, pp. 5386-5392, Dec. 2016, doi: 10.1093/infdis/jiw280.
20. R. L. Marquet et al., "Internet-based monitoring of influenza-like illness (ILI) in the general population of the Netherlands during the 2003-2004 influenza season," BMC Public Health, vol. 6, no. 1, p. 242, Oct. 2006, doi: 10.1186/1471-2458-6-242.
21. A. Gelman and J. Hill, Data analysis using regression and multilevel/hierarchical models. Cambridge, UK: Cambridge University Press, 2006.
22. A. Wesolowski et al., "Impact of human mobility on the emergence of dengue epidemics in Pakistan, Proc. Natl. Acad. Sci., vol. 112, no. 38, pp. 11887-11892, Sep. 2015, doi: 10.1073/pnas.1504964112.
23. JH.-H. Chang et al., "Mapping imported malaria in Bangladesh using parasite genetic and human mobility data," eLife, vol. 8, p. e43481, Apr. 2019, doi: 10.7554/elife.43481.
24. Q. Li et al., "Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus- Infected Pneumonia, N. Engl. J. Med., vol. 382, no. 13, pp. 1199-1207, Mar. 2020, doi: 10.1056/NEJMoa2001316.