VISVESVARAYA TECHNOLOGICAL UNIVERSITY

"Jnana Sangama", Belagavi-590018

Synopsis Report
on

"HEALTH CARE CHATBOT USING ARTIFICIAL INTELLIGENCE"

Submitted in the partial fulfillment of the requirements for

the award of

BACHELOR OF ENGINEERING DEGREE
In

COMPUTER SCIENCE & ENGINEERING

Submitted by

AKHILESH J A	**(4AD17CS005)**
ARJUN V	**(4AD17CS012)**
BHAVANA R	**(4AD17CS016)**
KUSUM I K	**(4AD17CS031)**

Under the guidance of

Mr. ANILKUMAR B H
Assistant Professor
Department of CSE

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

ATME College of Engineering

13th Kilometer, Mysore-Kanakapura-Bangalore Road
Mysore-570028

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 right now. Our GDP gradually decreased, as of now we have 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 provide in the database.

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 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 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 selfreported 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.

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.

OBJECTIVE

The main purpose of chatbots is to support and maintain good relation with the users, by offering precision, personalization, efficiency and scalability. Chatbots are meant to help and deliver immediate actions where humans can't reach due to timing or budget.

With the increased innovations and use of technology in different domains, chatbots are a new trend for the web-applications. Chatbots are a piece of technology that replace the need for a human to converse.

PROBLEM STATEMENT

We all know that we are affected by a pandemic situation, so we all are facing problems in this time and we need a regular update in this time about a covid-19 cases and our health status too.

But in our current stage we not having an existing system that we shows a every information related to covid-19 and also our health status.

We will not get every information in single site or website know everything about coronavirus.

We don't have below mentioned features in one system

1. To identify, track and forecast outbreaks.
2. To identify, track and forecast outbreaks.
3. To create Chat Bot with multiple regional language supporting Conversational Al Techniques.

EXISTING SYSTEM

There are lot of treatments that are available for various diseases. No human can possibly know about all the medicines and the diseases. So, the problem is that there isn't any place where anyone can have the details of the diseases or the medicines. We have proposed a system which give the composition of medicines and their prescribed users through AI Chat Bot.

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 users.

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 demystified 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.

With the outbreak of COVID 19, tracking of cases has been challenging in prevalent regions and less prevalent regions alike. 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 proposed system should have the following features. The transactions should take place in a secured format between various clients in the network.

It provides flexibility to the user to transfer the data through the network very easily by compressing the large amount of file.

It should also identify the user and provide the communication according to the prescribed level of security with transfer of the file requested and run the required process at the server if necessary.

In this system the data will be send through the network as a audio file. The user who received the file will do the operations like de embedding, decryption, and decompress in their level of hierarchy etc.

LITERATURE SURVEY

[1]A Review on NLU Engine

- **Authors**: Mohammed Javed, Naeun Lee (to implement word segmentation)

  Jerome R. Bellegarda (to implement POS Tagging)

- NLU i.e., Natural Language Understanding is a subpart of NLP (Natural Language Processing) which enables the system to understand the natural language or the conversational.
- To understand a complete sentence, the NLU system needs to understand each word of that sentence. It means that the initial task is the segmentation of the sentences into individual words. Next, to understand the word, the system needs to understand the grammar of the sentence. This can be done by knowing the parts of speech(POS) of each word in that sentence. This is the most important step wherein the word with the highest dependency is extracted, from which the intent of the system can be known. but with different words used in it. To match these types of synonymic sentences, synonym determination and sentence matching are required.

[2] A review on Synonym and Pattern Recognition

- **Authors**: LinHua Gao, Sijun Qin **(**Synonym detection and extraction)
- After understanding the intent of the user sentence, the database is checked for a sentence with the same intent. The matched sentences have difference of words which are used to express the same content. They use alternative words or synonyms. This makes synonym detection necessary for the system. Synonyms for a particular word may be domain independent or domain dependent.
- This paper explains the traditional dictionary method of synonym extractions. In this method, the system database maintains a dataset of synonyms for important keywords in that domain. The sentence sent by the user is then mapped on to that synonym dataset.
- It proposed a feature selection method for synonym extraction. In this method, among all the parts of speech tags, words having the tags as noun, verbs and adjectives are marked as positive tags and the others as negative tags. The polarity for each feature (word) is then carried out by using the POS tags. If the overall feature polarity is positive, then it

can be identified categorically. The one with the highest information gain is the strongest synonym extracted.

[3] A Review on Decision or ML Engine

- **Authors**: Sachin S. Gavankar, Naganna Chetty (Implementing disease predictions.)
- Scripted or monotonous chatbots have predefined replies to be given. They provide replies to the user from a set of predefined replies categorized on the basis of the query given by the user. Inclusion of ML in chatbots enables it to compute the replies from scratch.
- proposed the eager decision tree algorithm for prediction. This type of decision tree is the improvised version of the traditional decision tree. It creates this tree at runtime, based on the user's queries and keeps updating the tree on new user messages.
- In this algorithm, the system follows the clustering mechanism. It means that, the algorithm extracts that data from the knowledge base which is the closest to the user query. When the user fires a query, the algorithm searches for the best matches in the knowledge base and provides the same to the user.

[4] A Review on NLG Engine

- NLG performs the reverse task of NLU. It is the process of converting the system produced results into natural language representations which can be easily understood by the user.
- The sentences can have two voices i.e., active or passive voice. Also, there can be similarity between two sentences, but they might involve the usage of synonyms. Hence, while providing a response to the user, the NLG unit needs to calculate all the possibilities to interpret the same sentence, and then select the most appropriate one.
- NLG engine also performs a sequence of tasks to generate sentences. The initial task is to determine the content. It involves the selection of response to be given to the user. The next task is the choice of sentences. As already said, there can be a variety of sentences that can be used to express the same situation, this step deals with the selection of the appropriate sentence, which is the best for that instance.

Last and the most important is the morphology check, wherein the sentence generated from the previous steps is checked upon for its correctness.

SOFTWARE REQUIREMENTS:

Software Requirements Specification (SRS) is the starting point of the software development activity. Little importance was given to this phases in the early days of software development. The emphasis was first on coding and then shifted to design.

As systems grew more complex, it become evident that the goal of the entire system cannot be easily comprehended.

Hence need for the requirements analysis phase arose. Now, for large software systems, requirements analysis is perhaps the most difficult activity and also the most error prone. Some of the difficulty is due to the scope of this phase.

The software project is imitated by the client needs. In the beginning these needs are in the minds of various people in the client organization.

The requirement analyst has to identify the requirements by tacking to these people and understanding there needs. In situations where the software is to automated a currently manuals process, most of the needs can be understood by observing the current practice.

The SRS is a means of translating the ideas in the minds of the clients (the output) into formal document (the output of the requirements phase).

Thus the output of the phase is a set of formally specified requirements, which hopefully are complete and consistent, while the input has none of these properties.

OPERATING SYSTEM	: Windows 10
XAMPP	: PostgreSQL 13
LANGUAGE	: html 5
WEB TECHNOLOGY	: JavaScript, jQuery
BACKEND	: pgAdmin 4

HARDWARE REQUIREMENTS:

Processor - Intel core i5 2.0 GHz

RAM - 4GB or more

Hard Disk - 20 GB or more

Monitor - 15”CRT or LCD

Keyboard - Normal or Multimedia

Mouse - Compactable Mouse

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.

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.

[3] 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.

[4] 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.

[5] 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.

[6] 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.

[7] 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.

[8] 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. 2020, doi: 10.1101/2020.06.16.20126466.

[9] C. T. Lloyd et al., "Global spatio-temporally harmonised datasets for producing highresolution gridded population distribution datasets," Big Earth Data, vol. 3, no. 2, pp. 108-139, Apr. 2019, doi: 10.1080/20964471.2019.1625151.

[10] 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.

[11] Z. Du et al., "Risk for Transportation of Coronavirus Disease from Wuhan to Other Cities in China - Volume 26, Number 5-May 2020 Emerging Infectious Diseases journal - CDC," doi: 10.3201/eid2605.200146

[12] 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-maybebelying-reality-1888327 (accessed Jul. 20, 2020),

[13] 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-hospitals11-moretolaunch-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] H.-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 CoronavirusInfected Pneumonia, N. Engl. J. Med., vol. 382, no. 13, pp. 1199-1207, Mar. 2020, doi:
10.1056/NEJMoa2001316.

Signature of Guide
Mr. ANILKUMAR BH
Assistant Professor

Signature of Coordinator
Mrs. SUNITHAPATEL
Assistant Professor