SHORT ARTICLE

Health System Using Social Network to detect the Dengue infection in the specific region

Vijay Singh[1], Bhasker Pant[2], Devesh Pratap Singh[3]

[1]Assistant Professor, [2]Associate Professor, [3]Associate Professor, Department of Computer Science and Engineering, Graphic Era University, 566/6 Bell Road, Clement Town, Dehradun-248002, Uttarakhand, India

Abstract	Introduction	Methodology	Results	Conclusion	References	Citation	Tables / Figures

Corresponding Author

Address for Correspondence: Vijay Singh, Assistant Professor, Department of Computer Science and Engineering, Graphic Era University, Dehradun-248002, Uttarakhand, India
E Mail ID: vijaysingh_agra@hotmail.com

Citation

Article Cycle

Abstract

Background: There is law in some countries that each and every health-related event should be recorded or documented digitally and centrally. Unfortunately, this practice is not implemented due to various reasons, that is why the real ground realities differ from what agencies says. Social Network is one of the fastest and effective ways to propagate information without any discrimination and condition. **Aims & Objective**: By using this beauty of the social network, we proposed a Health System using twitter to detect dengue infected people in the specific region so the concerned agencies take appropriate action proactively and can avoid the epidemic like condition. **Material & Methods**: Using Naïve Bayes classifier we classify incoming tweets into two categories weather it is promotional tweet or it tells about a dengue infected patient. The method and the proposed health system presented in this article can be applied to another similar task, like to detect communicable disease patients in the specified location. **Results:** The proposed system is capable enough to give the feel of dengue in the particular region. The results show that the developed system's efficiency is 57% when testing performs on three set of collected tweets. **Conclusion:** Twitter can be effectively to get the health status of the society and warned the concerned agencies before the emergency like situation.

Keywords

Introduction

According to World Health Organization (WHO) studies, there are 390 million dengue infections per year of which 96 million manifests clinically. One another study of World Health Organization (WHO) on dengue says that 39 billion people in 128 countries are at risk of infection with dengue viruses (http://www.who.int/mediacentre/factsheets/fs117/en/). Dengue is a mosquito-borne viral infection, and there is no specific treatment of this. Initial detection and medical facilities can lower the fatality below 1%. Governments invested the hefty amount to promote and educate the people about dengue and other mosquito-borne diseases. In this article, we proposed a twitter based Health System to detect the infection of dengue virus and estimates the severity of dengue infection in the particular region. The problem with the location based service implementation in twitter is that there are around 1% tweets containing location information. To overcome this issue, we used streaming API and collect the tweets within mentioned geographical

box. Twitter is one of the famous micro-blogging services where a registered can communicate with other user limiting to 140 characters on the real-time basis (1). China, Iran, and Korea are some exceptions, where twitter service is no available or banned. Otherwise, twitter can access from throughout the world without any age limit. Jon Parker et al proposed a twitter based framework in which they detect the health trends. They used the technique of frequent-term set in the pre-filtered health-related tweets (2). Patty Kostkova et al, investigate the role of early warning and risk communication. They analyzed swine flu data during 2009 and detect the potential of the twitter to predict the outbreak in advance (3). Sarah Hosein, Raid Al-Tahir, and Bheshem Ramlal did a spatiotemporal analysis of dengue hemorrhagic fever (DHF) and dengue shock syndrome (DSS). For the tenure of 9 years (1998-2007) (4). Aron Culotta proposed a method of how to detect influenza epidemics by analyzing twitter messages. He compares the number of the regression model to correlate messages with centers for Disease Control and Prevention, best model achieved a correlation of .78 (5). Dan Dumbrell and Robert Steele works on characteristics of highly disseminated public health-related tweets. They analyzed health-related tweets and their attributes and differentiate among Government, for-profit (FP) and not-for-profit (NFP) (6). Sue Jamison-Powell and Conor Linehan analysed insomnia on Twitter. They proposed an approach in which they perform the analysis of around 18000 tweets to understand the sleep disorder among in the twitter users (7). Nattiya Kanhahua, Avare Stewart and Sara Romano, works on a time series data to detect the outbreak of specific health event. They did a correlation analysis for understanding the temporal analysis on twitter data (8).

Aims & Objectives

To develop Dengue outbreaks detection system using Twitter and perform data analysis on collected tweets. Identifies Severity location of dengue and estimated the number of dengue infected patients

Material & Methods

The prime motive behind this study and the proposed system is that concerned authorities get informed before the situation of the disease become worse. Keeping these things in mind we analyzed collected tweets and develop a system to how many people may be infected with dengue? The message

posted on the Twitter with the dengue keyword used as the indication of dengue infection. The developed system has the facility to choose the location by simple drop-down list. We consider four different geographical locations for evaluating the system that are New Delhi, Mumbai, Chennai and Kolkata. The main steps of the proposed word and application are as follows:

- Collect the location based tweets containing the keyword dengue.
- Apply Noise removal techniques.
- Classify the tweets into promotional or advertisement tweets about dengue and the tweets where the user is infected with dengue infection.
- Detect the severity of the dengue in the particular location.
- Develop a web-based health system for the above-mentioned task.

The collected tweet contains lots of noisy and unnecessary data, which should be processed for an efficient system. So, in the second step, we apply pre-processing techniques, like data cleaning, data transformation and data reduction technique and get the tweets containing desirable information only. In the third phase of the work, the Naïve Bayes algorithm is implemented to classify the collected tweets as promotional tweets and as individual or personal infection tweets. In another task to detect the severity location of the dengue, time-series data analysis 15-15 day's two slots. 1st slot 15/7/2016 to 31/7/2016 and 2nd slot 1/8/2016 to 15/8/2016.

Data Analysis: By using frequent-item in the specific period and in the specific location, as in our case in July at New Delhi more dengue cases are detected as compared to the August month by our proposed system. Out of 2277 tweets total tweets, 27 individual infections were detected. The output is shown in figure 1. Sample tweets of individual infections are shown in Table 1, while in Table 2 sample promotional tweets are shown.

Data Collection: The dataset has set of approximately 25000 tweets each from New Delhi, Chennai, Kolkata, and Mumbai. Python streaming API is used to download the tweets, Python based Text Blob API is used to tweets classification, PHP is user for server side services and HTML, CSS used for Interface.

Results

Three data sample are used, containing around 2000 tweets in each sample and find the systems accuracy is 57 %, which is quite promising in the case of Natural Language Processing.

Discussion

The result of this study gets better in future as the numbers of twitter users are increasing day-day. Otherwise, this kind of systems suffers from cold start problem. Cold Start problem is a situation when there is no sufficient information for prediction and this is the limitation as well. These kinds of system can be applicable to other communicable or non-communicable diseases. This kind of system can be effectively used as disease alarm system or disease surveillance system.

Conclusion

Twitter is one of the promising medium to detect the pulse of the nation or in specific scenario like health care concerns. In the future work we will develop mobile-based disease outbreak detection system.

Recommendation

As the cases of Dengue is increasing every passing year it becomes imperative that this has emerged as a major important factor in public health.

Limitation of the study

The availability of the tweets in the specific region or in the specific domain.

Relevance of the study

It adds the impact of social network as well as the web.

Authors Contribution

This is collaborative work by all the authors, Vijay Singh, Dr. Bhasker Pant and Dr. Devesh Pratap Singh, in all the aspects of this research article like data collection, data analysis and development of dengue outbreak system

Acknowledgement

Authors are very thankful to Mr. Dibyahash Bordoloi for his valuable suggestion.

References

1. Bridianne O'Dea, Stephen Wan, Philip J. Batterham, Alison L.Calear, Cecile Paris and Helen, "Detecting suicidality on Twitter", Internet Inventions 2(2015)183-188.
2. Jon Parker, Yifang Wei, Andrew Yates, Ophir Frieder and Nazil Goharian, "A Framework for Detecting Public Health Trends with Twitter", IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2013.
3. Patty Kostkova, Martin Szomszor and Connie St. Louis," #Swineflu: The Use of Twitter as an Early Warning and Risk Communication Tool in the 2009 Swine flu Pandemic", ACM Transaction on Management Information Systems, Vol.5, No.2, Article 8, Publication date:July 2014.
4. Sarah Hosein, Raid Al-Tahir and Bheshem Ramlal," Spatiotemporal Analysis of Dengue Hemorrhagic Fever and Dengue Shock Syndrome Incidence within Trinidad, West Indies", ACM Conference HEALTH 2013.
5. Aron Culotta," Towards detecting influenza epidemics by analysing Twitter messages", 1st Workshop on Social Media Analytics, ACM, 2010.
6. Dan Dumbrell and Robert Steele, "What are the Characteristics of Highly Disseminated Public Health-related Tweets?" ACM OZCHI 2012.
7. Sue Jamison-Powell and Conor Linehan, "I can't get no sleep: Discussing #insomnia on Twitter", Understand Online Communication, CHI 2012, 2012.
8. Nattiya Kanhabua,Avare Stewart and Sara Romano, "Supporting Temporal Analytics for Helath_Related Events in Microblogs", ACM CIKM 2012

Tables

TABLE 1 SAMPLE TWEETS SHOWS INDIVIDUAL INFECTION OF DENGUE

1.	I am suffering from dengue.
2.	Diagnosed with Dengue feeling sick at Dombivli.
3.	A -ve blood required for critical dengue patient
4.	My friend is a DENGUE patient. Not letting enter in emergency ward of Din Dayal Upadhyay Hospital, Hari Nagar. Help!
5.	A+ve Platelets needed for Vikas who has Dengue. urgent!

TABLE 2 SAMPLE TWEETS SHOWS PROMOTIONAL OR ANTI-DENGUE CAMPAIGN

1.	dengue symptoms can get themselves tested free of cost at Mohalla Clinics from 1st Sept
2.	State is taking all measures to combat dengue. We must maintain cleanliness
3.	Dengue vaccine has been launched internationally but caught in red tape in India. Why isn't it available when people are dying?
4.	Stop Dengue, Its in your hand. Doctor Bhargava Presents Dengue Preventive & Supportive
5.	Many vector-borne diseases are prevalent in the WHO South-East Asia Region. These include, mosquito-borne diseases

Figures

FIGURE 1 SNAPSHOT OF DEVELOPED APPLICATION