IJITCE

ISSN 2347- 3657

International Journal of Information Technology & Computer Engineering

Email : ijitce.editor@gmail.com or editor@ijitce.com

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 12, Issue 3, 2024

HEALTH MONITORING ON SOCIAL MEDIA OVER TIME

[1]A NAGARAJU, [2]S. DURGA NAGA VENKATA UPENDRA

[1](Assistant Professor), MSC, DNR college(A) PG courses Bhimavaram

[2]MSC, scholar, DNR college(A) PG courses Bhimavaram

ABSTRACT

Social media has become a major source for analysing all aspects of daily life. Thanks to dedicated latent topic analysis methods such as the Ailment Topic Aspect Model (ATAM), public health can now be observed on Twitter. In this work, we are interested in using social media to monitor people's health over time. The use of tweets has several benefits including instantaneous data availability at virtually no cost. Early monitoring of health data is complementary to post-factum studies and enables a range of applications such as measuring behavioral risk factors and triggering health campaigns. We formulate two problems: health transition detection and health transition prediction. We first propose the Temporal Ailment Topic Aspect Model (TM–ATAM), a new latent model dedicated to solving the first problem by capturing transitions that involve health-related topics. TM–ATAM is a non-obvious extension to ATAM that was designed to extract health-related topics. It learns health-related topic transitions by minimizing the prediction error on topic distributions between consecutive posts at different time and geographic granularities. To solve the second problem, we develop T–ATAM, a Temporal Ailment Topic Aspect Model where time is treated as a random variable natively inside ATAM. Our experiments on an 8-month corpus of tweets show that TM–ATAM outperforms TM–LDA in estimating health-related transitions from tweets for different geographic populations. We examine the ability of TM–ATAM to detect transitions due to climate conditions in different geographic regions. We then show how T–ATAM can be used to predict the most important transition and additionally compare T–ATAM with CDC (Centre for Disease Control) data and Google Flu Trends.

ISSN 2347–3657
Volume 12, Issue 3, 2024

1. INTRODUCTION

Social media has become a major source of information for analysing all aspects of daily life. In particular, Twitter is used for public health monitoring to extract early indicators of the well-being of populations in different geographic regions. Twitter has become a major source of data for early monitoring and prediction in areas such as health, disaster management and politics. In the health domain, the ability to model transitions for ailments and detect statements like “people talk about smoking and cigarettes before talking about respiratory problems”, or “people talk about headaches and stomach ache in any order”, benefits syndromic surveillance and helps measure behavioral risk factors and trigger public health campaigns. In this paper, we formulate two problems: the health transition detection problem and the health transition prediction problem. To address the detection problem, we develop TM–ATAM that models temporal transitions of health-related topics. To address the prediction problem, we propose T–ATAM, a novel method which uncovers latent ailment inside tweets by treating time as a random variable natively inside ATAM. Treating time as a random variable is key to predicting the subtle change in health-related discourse on Twitter.

Common ailments are traditionally monitored by collecting data from health-care facilities, a process known as sentinel surveillance. Such resources limit surveillance, most especially for real-time feedback. For this reason, the Web has become a source of syndromic surveillance, operating
on a wider scale, near real time and at virtually no cost. Our challenges are: (i) identify health-related tweets, (ii)determine when health-related discussions on Twitter transitions from one topic to another, (iii) capture different such transitions for different geographic regions. Indeed, in addition to evolving over time, ailment distributions also evolve in space.

On the other hand, while pLSI and LDA have been shown to perform well on static documents, they cannot intrinsically capture topic evolution over time. Temporal-LDA (TM–LDA) was proposed as an extension to LDA formining topics from tweets over time. To address thehealth transition detection problem, we propose TM–ATAMthat combines ATAM and TM–LDA. A preliminary version of TM–ATAM was described in a short paper [8]. We show here that it is able to capture transitions of health-related discussions in

ISSN 2347–3657

Volume 12, Issue 3, 2024

different regions (see Figure 1). As a result, the early detection of a change in discourse in Nevada, USA into allergies can trigger appropriate campaigns.

In each geographic region, TM–ATAM learns transition parameters that dictate the evolution of health-related topics by minimizing the prediction error on ailment distributions of consecutive pre-specified periods of time. Our second problem, the health transition prediction problem, is to automatically determine those periods. We hence propose T–ATAM, a different and new model that treats time as a random variable in the generative modelJust like TM–LDA, TM–ATAM and T–ATAMare different from dynamic topic models, as they are designed to learn topic transition patterns from temporally-ordered posts, while dynamic topic models focus on changing word distributions of topics over time.

Our experiments on a corpus of more than 500K health related tweets collected over an 8-month period, show that TM–ATAM outperforms TM–LDA in estimating temporal topic transitions of different geographic populations. Our results can be classified in two kinds of transitions. Stable topics are those where a health-related topic is mentioned continuously. One-Way transitions cover the case where some topics are discussed after others. For example, our study of tweets from California revealed many stable topics such as headaches and migraines. On the other hand, tweeting about smoking, drugs and cigarettes is followed by tweeting about respiratory ailments. Figure 1 shows example one way transitions we extracted for different states and cities in the world. Such transitions are often due to external factors such as climate, health campaigns, nutrition and lifestyle of different world populations.

2. EXISTING SYSTEM

- In the existing system, the authors propose a method that learns changing word distributions of topics over time and in the system, the authors leverage the structure of a social network to learn how topics temporally evolve in a community. TM–ATAM and T–ATAM are however different from dynamic topic models such as and, and from the work of Wang et al.as they are designed to learn topic transition patterns from temporally-ordered posts, while dynamic topic models focus on changing word distributions of topics over time.
- TM–ATAM learns transition parameters that dictate the evolution of health-related topics by minimizing the

ISSN 2347–3657

Volume 12, Issue 3, 2024

prediction error on ailment distributions of consecutive periods at different temporal and geographic granularities. T–ATAM on the other hand discovers latent ailments in health tweets by treating time as a corpus-specific multinomial distribution.

- Classical approaches have been applied to mining topics for inferring citations. Other discriminative approaches have been applied to do an empirical study on topic modelling and time-based topic modelling respectively. None of those are directly applicable to health data.

DISADVANTAGES

- There is no Mapping Tweets to Documents.
- There is Uncovering Health Topics with ATAM.
-

3. PROPOSED SYSTEM

- In the proposed system, the system formulates and solves two problems: the health transition detection problem and the health transition prediction problem. To address the detection problem, the system develops TM–ATAM that models temporal transitions of health-related topics. To address the prediction problem, we propose T–ATAM, a novel method which uncovers latent ailment inside tweets by treating time as a random variable natively inside ATAM.
- Treating time as a random variable is key to predicting the subtle change in health-related discourse on Twitter.

ADVANTAGES

- TM–ATAM, a model able to detect health-related tweets and their evolution over time and space. TM–ATAM learns, for a given region, transition parameters by minimizing the prediction error on ailment distributions of pre-determined time periods.
- T–ATAM, a new model able to predict health-related tweets by treating time as a variable whose values are drawn from a corpus-specific multinomial distribution.

4. OUTOUT SCREENS

HOME PAGE

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 12, Issue 3, 2024

ADMIN PAGE

LOGIN PAGE

VIEW HEALTH RESULTS

USER PAGE

5. CONCLUSION

This been a great pleasure for me to work on this exciting and challenging project. This project proved good for me as it provided practical knowledge of not only programming in ASP.NET and VB.NET web-based application and no some extent Windows Application and SQL Server, but also about all handling procedure related with **"PROJECT NAME".** It also provides knowledge about the latest technology used in developing web enabled application and client server technology that will be great demand in future. This will provide better opportunities and guidance in future in developing projects independently.

BENEFITS:

The project is identified by the merits of the system offered to the user. The merits of this project are as follows: -

- It's a web-enabled project.
- This project offers user to enter the data through simple and interactive forms. This

ISSN 2347–3657
Volume 12, Issue 3, 2024

is very helpful for the client to enter the desired information through so much simplicity.

● The user is mainly more concerned about the validity of the data, whatever he is entering. There are checks on every stage of any new creation, data entry or updation so that the user cannot enter the invalid data, which can create problems at later date.

● Sometimes the user finds in the later stages of using project that he needs to update some of the information that he entered earlier. There are options for him by which he can update the records. Moreover, there is restriction for his that he cannot change the primary data field. This keeps the validity of the data to longer extent.

● User is provided the option of monitoring the records he entered earlier. He can see the desired records with the variety of options provided by him.

● From every part of the project the user is provided with the links through framing so that he can go from one option of the project to other as per the requirement. This is bound to be simple and very friendly as per the user is concerned. That is, we can say that the project is user friendly which is one of the primary concerns of any good project.

● Data storage and retrieval will become faster and easier to maintain because data is stored in a systematic manner and in a single database.

● Decision making process would be greatly enhanced because of faster processing of information since data collection from information available on computer takes much less time than manual system.

● Easier and faster data transfer through latest technology associated with the computer and communication.

● Through these features it will increase the efficiency, accuracy and transparency,

LIMITATIONS:

● The size of the database increases day-by-day, increasing the load on the database back up and data maintenance activity.

● Training for simple computer operations is necessary for the users working on the system.

6. REFERENCES

[1] L. Manikonda and M. D. Choudhury, “Modelling and understanding visual attributes of mental health disclosures in social media,”in Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, Denver, CO, USA, May 06-11, 2017., 2017, pp.170–181.

ISSN 2347–3657

Volume 12, Issue 3, 2024

[2] S. R. Chowdhury, M. Imran, M. R. Asghar, S. Amer-Yahia, and C. Castillo, "Tweet4act: Using incident-specific profiles for classifying crisis-related messages," in 10th Proceedings of the International Conference on Information Systems for Crisis Response and Management, Baden-Baden, Germany, May 12-15, 2013., 2013.

[3] T. Davidson, D. Warmsley, M. W. Macy, and I. Weber, "Automated hate speech detection and the problem of offensive language," in Proceedings of the Eleventh International Conference on Web and social media, ICWSM 2017, Montréal, Québec, Canada, May 15-18, 2017.,2017, pp. 512–515.

[4] M. J. Paul and M. Dredze, "You Are What You Tweet: Analyzing Twitter for Public Health," in ICWSM'11, 2011.

[5] T. Hofmann, "Probabilistic Latent Semantic Indexing," in SIGIR'99, 1999, pp. 50–57.

[6] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet Allocation," Journal of Machine Learning, vol. 3, pp. 993–1022, 2003.

[7] Y. Wang, E. Agichtein, and M. Benzi, "TM-LDA: Efficient Online Modeling of Latent Topic Transitions in Social Media," in KDD'12,2012, pp. 123–131.

[8] S. Sidana, S. Mishra, S. Amer-Yahia, M. Clausel, and M. Amini, "Health monitoring on social media over time," in Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval, SIGIR 2016, Pisa, Italy, July17-21, 2016, 2016, pp. 849–852.

[9] D. M. Blei and J. D. Lafferty, "Dynamic Topic Models," in ICML'06,2006, pp. 113–120.

[10] C. X. Lin, Q. Mei, J. Han, Y. Jiang, and M. Danilevsky, "The Joint Inference of Topic Diffusion and Evolution in Social Communities,"in ICDM'11, 2011, pp. 378–387.