OJPHI

ISDS 2015 Conference Abstracts

Natural Language Processing and Technical Challenges of Influenza-Like Illness Surveillance

Dino P. Rumoro[1], Gillian S. Gibbs*[1], Shital C. Shah[1], Marilyn M. Hallock[1], Gordon M. Trenholme[1], Michael J. Waddell[2] and Joseph P. Bernstein[3]

[1]Rush University, Chicago, IL, USA; [2]Pangaea Information Technologies, Ltd, Chicago, IL, USA; [3]H-CORE, LLC, Chicago, IL, USA

Objective

To review the natural language processing (NLP) and technical challenges encountered in an automated influenza-like illness (ILI) surveillance system.

Introduction

Processing free-text clinical information in an electronic medical record (EMR) may enhance surveillance systems for early identification of ILI outbreaks. However, processing clinical text using NLP poses a challenge in preserving the semantics of the original information recorded. In this study, we discuss several NLP and technical issues as well as potential solutions for implementation in syndromic surveillance systems.

Methods

This is a retrospective, cross-sectional study conducted at the EDs of a large urban academic medical center and community hospital. The study timeframe was October 1, 2014 to June 30, 2015. Geographic Utilization of Artificial Intelligence in Real-Time for Disease Identification and Alert Notification (GUARDIAN) – a syndromic surveillance program – received and processed HL7 messages in real-time and generated ILI surveillance reports. The sophisticated GUARDIAN NLP algorithm processed each patient chart component, consistent with a physician's manual review [1].

A random sample of 10 ILI-positive cases detected by GUARDIAN was drawn each week for manual review to confirm the positive presence of the ILI case definition terms: fever, cough, and sore throat. False ILI-positive cases, and associated causes, were documented and categorized as shown in Table 1.

Results

Of the 519 ILI-positive charts reviewed, 56 cases were false positive, mainly due to NLP or programming errors (e.g., incorrect concept parsing due to certain punctuation and word combinations).

Temporal relationships were found to be a challenge for NLP: examples included when a clinician noted a fever in the past or documented instructions to return for a fever. Physician documentation style was also a common and difficult problem: examples include the use of italics or bold text to represent a positive or negative symptom.

We identified new terms to add to the negation list and discovered a problem caused by the use of a long negation string. GUARDIAN was programmed to acknowledge negation strings up to 16 words since any more words than that decreased accuracy.

We found that temperatures were occasionally corrected in the EMR, while GUARDIAN had no way of knowing the original value was an error. Lastly, on occasion, there was also reception of incorrect inpatient vitals and truncated nurse notes.

With the modification of our system architecture and NLP engine, we were able to reduce the associated ILI false positives from 56 (10.8%) to 32 (6.2%).

Conclusions

The use of NLP can enhance the efficacy of syndromic surveillance systems. However, there are limitations to NLP processing loads. While many NLP errors can be corrected, yielding improved accuracy, some issues cannot be resolved. Sharing known technical and NLP issues, and their resolutions, can assist in minimizing errors to acceptable levels (<5%) leading to refinement of existing syndromic surveillance systems.

Table 1. Review of NLP and technical challenges in ILI surveillance.

Issue	Examples	Prevalance
Temporal relationship	Historical symptoms noted; instructions to return if symptoms persist	26.8% (15)
Programming error*	Punctuation caused concepts to be incorrectly parsed	26.8% (15)
Text formatting	Physician used rich-text formatting to indicate negative symptoms	16.1% (9)
Excessive negation string length	More than 16 words long	10.7% (6)
Context misrepresentation*	Unrelated symptoms, such as "hay fever"	8.9% (5)
Slippage of additional data through outbound filters*	Incorrectly received inpatient vitals or partial nurse note	7.1% (4)
Correction ambiguity	Temperature corrected by nurse	3.6% (2)

Note: *Issues were resolved by modifying the system architecture and NLP engine.

Keywords

Influenza-like illness; Natural language processing; Lessons learned

Acknowledgments

GUARDIAN is funded by the US Department of Defense, Telemedicine and Advanced Technology Research Center, Award numbers W81XWH-09-1-0662 and W81XWH-11-1-0711.

References

Silva J, Shah S, Rumoro D, Bayram J, Hallock M, Gibbs G, Waddell M. Comparing the accuracy of syndrome surveillance systems in detecting influenza-like illness: GUARDIAN vs. RODS vs. electronic medical record reports. Artificial Intelligence in Medicine. 2013;59:169-174.

***Gillian S. Gibbs**
E-mail: Gillian_Gibbs@rush.edu