

Computer scientist publishes new algorithm cluster to data mine health records

May 14 2013

The time may be fast approaching for researchers to take better advantage of the vast amount of valuable patient information available from U.S. electronic health records. Lian Duan, an NJIT computer scientist with an expertise in data mining, has done just that with the recent publication of "Adverse Drug Effect Detection," *IEEE Journal of Biomedical and Health Informatics* (March, 2013).

The article spotlights a new and promising way of using a combination of commonly used existing algorithms to root out more information about [adverse drug reactions](#) within electronic health records available to the researchers. The new pattern, which when compared against the most commonly used existing sole algorithm, showed an almost 25 percent improvement in outcome. Although the idea could theoretically be applied beyond [electronic health records](#), this paper focuses only on using them to find adverse medical reactions to a drug therapy.

"Large collections of [electronic patient records](#) have long provided abundant, but under-explored information on the real-world use of medicines. But when used properly these records can provide longitudinal observational data which is perfect for data mining," Duan said. "Although such records are maintained for patient administration, they could provide a broad range of clinical information for data analysis. A growing interest has been drug safety."

In this paper, the researchers proposed two [novel algorithms](#)—a likelihood ratio model and a Bayesian network model—for adverse drug

effect discovery. Although the performance of these two algorithms is comparable to the state-of-the-art algorithm, Bayesian confidence propagation neural network, by combining three works, the researchers say one can get better, more diverse results.

Since the actual adverse drug effects on a given dataset cannot be absolutely determined, the researchers made use of a simulated observational [medical outcomes](#) partnership dataset. They constructed this "dataset" with the predefined adverse drug effects to evaluate their methods.

Experimental results show the usefulness of the proposed pattern discovery method on the simulated dataset by improving the standard baseline algorithm—chi-square—by 23.83 percent.

Duan, whose innovative research on large-scale data mining has applications in the business world as well as many industries, including marketing, social networking and bioinformatics. Whereas most data mining experts search for correlation pairs, he focuses on correlated sets of arbitrary size. His research focuses on correlation search, community detection, and density-based clustering and outlier detection.

Duan holds two doctorates— one in computer science from the Chinese Academy of Sciences, China, and the other in information systems with an emphasis on data mining from the University of Iowa.

Provided by New Jersey Institute of Technology

Citation: Computer scientist publishes new algorithm cluster to data mine health records (2013, May 14) retrieved 19 April 2024 from <https://phys.org/news/2013-05-scientist-publishes-algorithm-cluster-health.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.