Risk Prediction in Health due to Drug Reactions Using Supervised Learning Technique

Divya S.
U.G. Student
Francis Xavier Engineering College.

Rajeswari M.
U.G. Student
Francis Xavier Engineering College.

Mr. E. Manohar
A.P. /CSE,
Francis Xavier Engineering College.

ABSTRACT

The major issue in human health is the reaction of the drug. Most of the clinical trial does not identify the harmfulness of the drug due to limitation in size. In this paper the post-marketing surveillance carried out to monitor the impacts of medicines once they have been made available to the user. Nowadays, many data mining techniques and methodologies have been developed to motivate the mining and detection of drug reaction. These methods are inconvenient and inefficient for users and the process is time consuming. This paper proposes a combined system platform for the detection of the reaction of the drug. By combining a drug databases with new data mining techniques, the proposed system not only supports multidimensional analysis of drug reaction, but also allows the interactive discovery of associations between drugs and symptoms, called a drug and its reaction association rule.

In this paper, an interactive approach to capture the Causality between dugs and their reactions. Premarketing clinical trials are required to capture the association between drugs and symptoms. But this premarketing clinical trial is limited in size. Therefore drug safety heavily depends on Postmarketing a clinical trial that is monitoring the impacts of medicine once they have been made available to the users. The interestingness measures named as intensive causal leverage measures. This measure includes the databases such as medical details and drug reaction. If the user posted their query then that query is compared between this causal measures and then the method analyze that the user taking drug has adverse effect or not.

As electronic patient records become more and more easily accessible in various health organizations such as hospitals, medical centers, and insurance companies, they provide a new source of information that has great potential to generate drug reaction.

1. INTRODUCTION

A drug reaction is an expression that describes harm associated with the use of given medications at a normal dosage during normal use. Drug reaction may occur following a single dose or prolonged administration of a drug or result from the combination of two or more drugs. The meaning of this expression differs from the meaning of "side effect", as this last expression might also imply that the effects can be beneficial. A drug event refers to any injury caused by the drug and any harm associated with the use of the drug

Note that each patient case can be considered as an event sequence where various events such as drug prescription, occurrence of a symptom and lab test occur at different times. In the literature, there exist a couple of studies that attempted to find the associations between drugs and potential drug reaction by mining their temporal relationships. That is, they tried to mine temporal association rules (represented as X ->Y) where Y occurs after X within a time window of length T. These studies obtained promising results based on administrative health data. However, temporal association

was the only parameter used for linking a symptom with a drug within each patient case in their work. Temporal association assumes that cause precedes effect. Other parameters such as challenge and rechallenge can also give direct or indirect cues of the potential causal association of a drug-symptom pair. Dechallenge is defined as the relationship between withdrawal of the drug and abatement of the adverse effect. Rechallenge describes the relationship between reintroduction of the drug followed by recurrence of the adverse event. In addition, their approaches suffer from the sharp boundary problem. On the one hand, the symptom events near the time boundaries are either ignored or overemphasized. On the other hand, two symptom events contribute equally to the interestingness measure as long as they occur within the hazard period T. That is, the length of the time duration between exposure to the drug and occurrence of the symptom has no effect on the interestingness measure. This is not true in reality because if a drug reaction symptom occurs within a shorter period, it is usually more likely to be caused by the drug.

This study develops a platform to analyze adverse drug reactions, which combines data warehousing and data mining technologies, through which users can observe and analyze drug-ADR signals from different viewpoints. Specifically, a contingency-cube-based method and an associative-classification-based method are proposed to facilitate the interactive detection of suspected drug-ADR and multidrug-ADR signals, respectively. The experimental results show that the cube-based approach significantly outperforms an associative-classification based approach and the interactive exploitation of suspected association of drugs and symptoms from a data warehouse is more efficient.

2. RELATED WORK IN THE LITERATURE

A large volumes of data related to adverse events and the development of data mining technology have spawned the use of statistical or data mining methods for the detection of drug reaction. These methods can be divided into two categories: the measures of disproportionality and the Bayesian methods.

The measures of disproportionality are commonly used techniques for the identification of drug reaction. Although different measures for calculating disproportionality are not concordant, they all use a 2×2 contingency table as shown in Table 1. The most common measures include the Proportional Reporting Ratio (PRR) used by the UK Yellow Card database (Evans et al. 2001), the Reporting Odds Ratio (ROR) used by the Netherlands Pharmacovigilance Foundation (Egberts et al. 2002) and the MHRA, an integrated measure used by the UK Medicines and Healthcare products Regulatory Agency (MHRA) (Evans et al. 2001). The MHRA combines the PRR, the numbers reported and a chi-squared test.

	Suspected ADR	All other ADRs	Total
Suspected drug	A	b	a+b
All other drugs	C	d	c+d
Total	a+c	b+d	a+b+c+d

Table 1. The 2×2 contingency table used for the identification of ADRs

The best-known Bayesian-based method is the Bayesian Confidence Propagation Neural Network (BCPNN) used by the World Health Organization (WHO) (Beta et al. 1998; Orre et al. 2000). This approach uses Bayesian statistics in a neural network architecture and calculates an information component (IC) for each drug–drug reaction combination. The US Food and Drug Administration (FDA) uses an algorithm called Empirical Bayes Gamma-Poisson

Shrinker (EBGPS) to detect those ADRs that have the frequency of reporting higher than the expected value (Dumouchel 1999). This algorithm also uses a Bayesian statistical formula to calculate the observed reporting value and the expected reporting value for each drug-drug reaction pair. The observed ratio of reporting value to expected reporting value represents the strength of the signal of the drug-drug reaction pair. A drug-drug reaction pair with an observed ratio higher than the threshold is more significant and worthy of further investigation.

3. PROPOSED SYSTEM FRAMEWORK

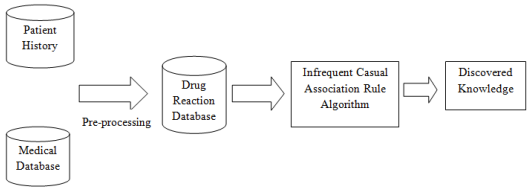

Figure 1. System Architecture

3.1 DESIGN STRATEGY

This proposed system establishes an interactive platform for the end user to allow the analysis and detection of suspicious drug reaction signals. It is well known that drug reaction signal detection is time consuming - at least of the same complexity as typical data mining tasks, such as association rule mining and classification analysis. In order to reduce the computation time, a general concept commonly used in the context of query processing, precomputation, is used. This executes the partial or total computation involved in the process of answering the query, in advance.

3.2 SYSTEM OVERVIEW

In this system, the databases such as patient database, medical databases. After preprocess this databases, we have drug reaction data warehouse. Then this data warehouse referred as the data source for the data mining engine. This data mining engine is connected to the web server. Through this web server, the user posted their query and the query is analyzed by the data mining engine.

4. CONCLUSION AND FUTURE WORK

In real world application, it is very important to mine the Causal Associations between two events. This provides the information that can help people to discover the causality of a type of events and avoid its potential drug effects. In this paper, a system platform to analyze and detect Adverse Drug Reactions is developed. The Users can interact with this platform to examine various forms of drug reaction signals from different view points, by selecting and readjusting parameters measures of interest. One of the main problems is that the pharmocovigilance using computer systems is a lack of standard measures for signal detection. This paper presents a preliminary development of drug reaction detection and analysis and there is much scope for future research, such as including various parameters, time and onset latency. The system can be enhanced with more visualization tools such as signal tracking and monitoring mechanisms. So that users can effectively track the change in some specific drug reaction signals.

5. REFERENCES

1. Alexandru Floares, Ovidiu Balacescu, Carmen Floares, Loredana Balacescu, Tiberiu Popa, Oana Vermesan (2010), 'Mining Knowledge and Data to Discover Intelligent Molecular Biomarkers: Prostate Cancer i-Biomarkers'.
2. Anni Coden, Daniel Gruhl, Neal Lewis, Michael Tanenblatt, Joe Terdiman (2012), 'SPOT the drug! An unsupervised pattern matching

method to extract drug names from very large clinical corpora' IEEE Second Conference on Healthcare Informatics, Imaging and Systems Biology.

3. C. Marinica and F. Guillet (June 2010), 'Knowledge-Based Interactive Postmining of Association Rules Using Ontologies' IEEE Transactions on Knowledge and Data Engineering, Vol 22, No.6.
4. Dat Tran, Wanli Ma, and Dharmendra Sharma (2009), 'Fuzzy Subspace Hidden Markov Models for Pattern Recognition'.
5. H. Jin, J. Chen, H. He, G. Williams, C. Kelman, and C. O'Keefe (2008), 'Mining Unexpected Temporal Associations: Applications in Detecting Adverse Drug Reactions' IEEE Transactions on Information Technology in Biomedicine, Vol 12, No.4.
6. Jutamas Tempaiboolkul, School of Engineering and Technology, Asian Institute of Technology, Thailand (2013), 'Mining Rare Association Rules in a Distributed Environment using Multiple Minimum Supports'.
7. Lian Duan, Mohammad Khoshneshin, W. Nick Street, and Mei Liu (2013), 'Adverse Drug Effect Detection' IEEE Journal of Biomedical and Health Informatics, vol.17,No.2.
8. Luigi Troiano, Giacomo Scibelli, Cosimo Birtolo (2009), 'A Fast Algorithm for Mining Rare Itemsets' Ninth International Conference on Intelligent Systems Design and Applications.
9. Noah Lee, Andrew F. Laine, Jianying Hu, Fei Wang, Jimeng Sun, Shahram Ebadollahi (2011), 'Mining electronic medical records to explore the linkage between healthcare resource utilization and disease severity in diabetic patients' First IEEE International Conference on Healthcare Informatics, Imaging and Systems Biology.
10. Yanqing Ji, Hao Ying, Peter Dews, Margo S. Farber, Ayman Mansour, John Tran, Richard E. Miller, R. Michael Massanari (2010), 'A Fuzzy Recognition-Primed Decision Model-Based Causal Association Mining Algorithm for Detecting Adverse Drug Reactions in Postmarketing Surveillance'.
11. Y.Ji, H. Ying, P. Dews, A. Mansour, J. Tran, R.E. Miller, and R.M.Massanari (2011), 'A Potential Causal Association Mining Algorithm for Screening Adverse Drug Reactions in Postmarketing Surveillance' IEEE Transactions on Information Technology in Biomedicine, Vol 15, No.3.
12. Y. Ji, H. Ying, M.S. Farber, J. Yen, P. Dews, R.E. Miller, and R.M.Massanari (May2010), 'A Distributed, Collaborative Intelligent Agent System Approach for Proactive Postmarketing Drug Safety Surveillance' IEEE Transactions on Information Technology in Biomedicine, Vol 14, Nso.3.