ANALYSIS ON ADVERSE SIDE EFFECTS OF DRUG MEDICATIONS AND DATA MINING ITS DATA MINING METHODS

[1] K.M.Deepa Lakshmi., [2]Dr.A.Prakash..
[1]Research Scholar, [2] Associate Professor & Reserach Supervisor,
[1], [2] PG & Department of Computer Science,
[1], [2] Hindusthan College of Arts & Science , Coimbatore, India.

Abstract

Recent trends move towards medicine consuming in day to day activities for the people who were suffering from diseases. Hence many of the people are not aware of the medicine which is prescribed by doctors or pharmacies. Once they are affected by diseases they are approaching doctor and they are in taking the medicines prescribed by them, without having any knowledge about it and gets affected by its side effects. Adverse drug events are unexpected and unwanted side effects caused by drugs when administered together to treat the same disease or different health conditions. When multiple drugs interact, they alter the desired level of activity and induce reactions that are unpredictable based on the side effects of the drugs observed when administered individually.

Keywords: Adverse Drug Reactions, Temporal Pattern Detection, Reporting Odds Ratio.

1. Introduction

The definition of Adverse Drug Events is not trivial but a common definition has been agreed on by researchers. Defining Adverse Drug Events (ADEs) first requires defining Adverse Drug Reactions (ADRs). The World Health Organization (WHO) and the European Union share the same definition of an ADR: “A response to a medicinal product which is noxious and unintended and which occurs at doses normally used in man for the prophylaxis, diagnosis or therapy of disease or for the restoration, correction or modification of physiological function”. The Institute Of Medicine defines an ADE as “an injury resulting from medical intervention related to a drug”. This definition has been simplified to “an injury resulting from the use of a drug”. According to that definition ADEs include harm caused by the drug (ADRs and overdoses) and harm from the use of the drug, including dose reductions and discontinuations of drug therapy. The Institute Of Medicine also gives another definition of ADEs that is interesting because it integrates the part that the diseases of the patient play in the outcomes: an ADE is “an injury due to medication management rather than the underlying condition of the patient”. A more complete definition can be retrieved from the “Glossary of terms related to patient and medication safety” elaborated by the Committee of Experts on Management of Safety and Quality in Health Care / Expert group on Safe Medication Practices, commissioned by the Council of Europe: “An Adverse Drug Event is any injury occurring during the patient’s drug therapy and resulting either from appropriate care, or from unsuitable or suboptimal care. Adverse drug events include: the adverse drug reactions during normal use of the medicine, and any harm secondary to a medication error, both errors of omission or commission.”

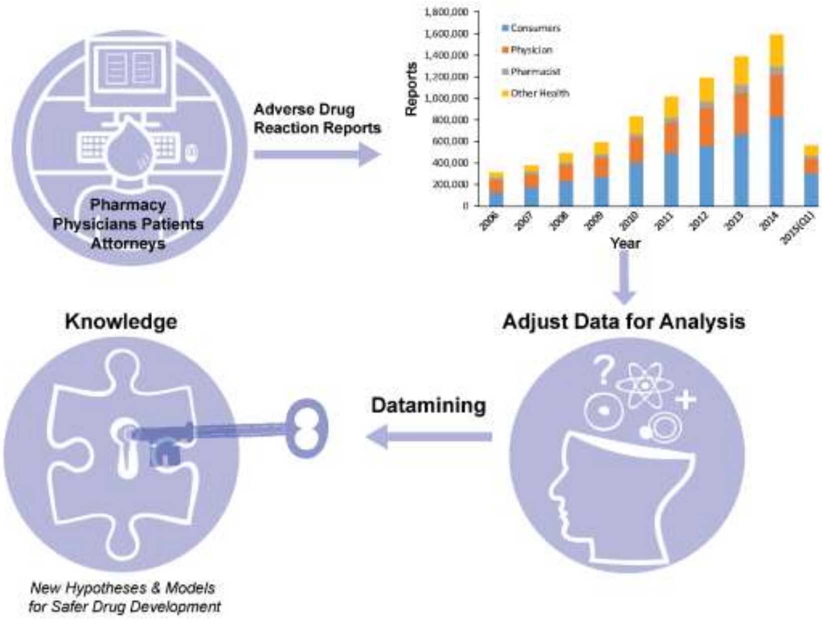

Figure 1: Adverse Drug Reaction Report System

In accordance with those definitions, researchers also agree on dividing ADEs into two categories: preventable ADEs and non preventable ADEs. Preventable ADEs are assimilated to “medication errors” while non preventable ADEs are considered ADRs that could not be avoided. It is worth noting that medication errors do not necessarily harm the patients. Only a limited portion of medication errors turns into actual ADEs; all of them are preventable. Conversely, all preventable ADEs are considered medication errors. From a pharmacological point of view, ADRs are well defined. Six types have been identified: - Type A: Augmented pharmacologic effects - dose dependent and predictable: intolerance and side effects - Type B: Bizarre effects (or idiosyncratic) - dose independent and unpredictable - Type C: Chronic effects - Type D: Delayed effects - Type E: End-of-treatment effects - Type F: Failure of therapy. Reporting systems of medication errors or incidents are the most ancient methods, and they were imported in healthcare from other domains such as Transportation (aviation) or Industry. Reporting systems are usually documented by healthcare professionals spontaneously or after prompting, but some systems are designed to be documented by the patients themselves. Although ADE reporting is made mandatory by the law in certain cases, authors usually agree that all reporting systems suffer from important under-reporting biases. Due to the exponential increase of the available computerized patient data, one would think that, as in banking industry, insurance companies or mass retail sector, data mining is more and more used to automatically screen large amount of medical records.

2. Literature Survey

[1] Sohini Sinha, Shubham Sharma & Ms. Srividhya. S(2018) proposed approach is general and can be used for segmentation in other applications where sequential data is accompanied with correlated signals. The fast growth of e-commerce, increasingly a growing number of products is marketed on web, and more people are also buying items from online. , it has actually ended up being an unusual technique for online vendors to make it possible for client's evaluations or to share opinions on the products that they have actually acquired. With more and more, a growing number of common individuals ending up being comfy with the internet. An enhancing number of people are composing evaluations. Because of this the variety of evaluations an item obtains boosts swiftly. Hence for this reason we have to efficiently analyze as well as make use of enormous online information resource is a difficulty. Opinion Mining takes care of the removal of information example favorable as well as unfavorable sentiments from a big piece of data or evaluations authored by individual. **[2] Abeed Sarker and Graciela Gonzalez (2016)** proposed the obvious use cases for utilizing social media data, national surveillance programs are yet to integrate proposed systems. Research tasks has successfully employed supervised learning systems that use manually annotated data to solve various natural language processing (NLP) problems. These include, for example, text classification tasks such as detecting mentions of adverse drug reactions, and extracting exact mentions using sequence labeling techniques. While these approaches have shown good performance in noisy, social media text, their need for manual annotations makes them expensive in nature. Manual annotations are time consuming, and the erratic properties of social media text make annotation tasks even harder. Consequently, even designing annotation tasks and guidelines require significant amounts of expert time, experience in annotations, and exposure to user posted texts. **[3] K. Arutchelvan and Dr. Pon Periasamy(2016)** proposed on knowing the major motive for commencement and indulging the youth in drugs. Ongoing efforts are geared towards increasing the size of data set. Data mining intends to endow with a systematic survey of current techniques of knowledge discovery in databases using data mining techniques that are in use in today's medical research. Discussion is made to enable the disease diagnosis and the breakthrough of hidden healthcare patterns from related databases is offered. Also, the use of data mining to discover such relationships as those between health conditions and a disease is presented. It further discusses about the tools that can be used for the processing and classification of data. This paper summarizes various technical articles on medical diagnosis and prognosis. It has also been focused on current research being carried out using the data mining techniques to enhance the diseases forecasting process. This research paper provides future trends of current techniques of KDD, using data mining tools for healthcare. It also confers significant issues and challenges associated with data mining and healthcare in general. **[4] Ken Naganuma Masayuki Yoshino, Ph.D. Hisayoshi Sato, Ph.D. Yoshinori Sato(2014)** proposed "Searchable encryption" is a generic term for encryption techniques that allow not only conventional encryption and decryption, but that can also perform text matching using an encrypted query on encrypted text. While encryption and decryption keys respectively are required for encryption and decryption, text matching does not require any special information and therefore can be performed by a cloud service that does not have the keys. Common key systems use the same key for both encryption and decryption. They are best suited to large quantities of data because they tend to execute more efficiently than public key systems. **[5] Megha Sinha, Sonu Kumar(2018)** proposed data Mining techniques are in use, such as K-Mean, KNN, ANN, SVN. Data Mining is one of the most vital and motivating area of research with the objective of clinical diagnosis and prognosis requires efficient and fast classification techniques, which in turn requires a large amount of genetic data generation and analyzing these huge data. The large amount of genetic data generated is obtained using the microarray technique in which expression of thousands of genes is concurrently measured and we are in the need of an efficient data mining technique for these huge data. SOM algorithm is efficient and one of most popular algorithm but having some drawback such as learning algorithm adjustment in winning field and the classification accuracy is needing to improve. The algorithm is efficient but though have a problem in visibility of clearer network. So in the field of drug industry Gene expression microarrays could be used to examine the physiological effects of drug administration, allowing the analysis of pathways and the identification of side reactions in which drugs bind promiscuously to cellular proteins, producing toxic side effects using gene expression analysis based on genetic algorithm.

3. Methods Taken for Analysis

To enable a fair comparison the TPD, MUTARA, HUNT and modified ROR methods, described were applied to investigate the one to thirty day period after the drug is prescribed. If each method used a different time period, the comparison would be biased.

3.1 Temporal Pattern Detection (TPD)

In this study the TPD was implemented as described with IC value over the time period corresponding to the 30 days after the first prescription in 13 months (u = [0, 30]) contrasted with the IC value over the time period corresponding to the 27 to 21 months prior to prescription (v = [−822, −639]), but two different filters were investigated:

- The TPD is applied and medical events with an IC value the month prior to prescription or an IC value on the prescription day greater than the IC value during the month after the prescription are filtered (TPD 1).

- The TPD is applied and medical events with an IC value the month prior to prescription greater than the IC value during the month after the prescription are filtered (TPD 2).

The justification for choosing two filters is due to the possibility that ADRs can occur and be reported to doctors on the same day as the prescription, so filtering events with an IC value on the day of prescription greater than the IC value during the month after the prescription may prevent detection of some ADRs.

3.2 Mining Unexpected Temporal Association Rules Given Antecedent & Highlighting UTAR's Negating TAR's (MUTARA & HUNT)

Two different lengths for the reference period were investigated as the length of the reference period determines the per patient filter stringency and the optimal stringency is unknown for the THIN database. The reference period for $MUTARA_{60}$ and $HUNT_{60}$ is set to be the time period starting from two months prior to the prescription and ending the day before the prescription. The reference period for $MUTARA_{180}$ and $HUNT_{180}$ is set to be the time period starting from six months prior to the prescription and ending the day before the prescription. The reference periods are chosen to end the day before the prescribed as this gave better preliminary results. The other parameter values used are: Tc = Te = 30, as this corresponds to the time period of a month after the drug prescription.

3.3 Reporting Odds Ratio (ROR)

The 'Spontaneous reporting system' style transformation is applied, where SRS style reports consisting of a patient, drug prescription and possible ADR are inferred from the LOD by discovering all the medical events that occur within 30 days of a drug prescription. Signals are only generated for medical events that have been reported with the drug of interest a minimum of 3 times.

4. Types of Adverse Drug Reactions

Adverse Drug Reaction often refers to the side effects caused by a drug and can be categorized based on the mechanism by which they are caused. Their knowledge and understanding is necessary for practitioners to monitor drug therapy and ADR detection.

4.1 Type A ADRs

These are caused by the normal pharmacological effect and effects of a substance. Usually, these are identified in the manufacturing phase. Adequate labelling and other information can be used to make these evident as being predictable and dose related. Example 1: Respiratory depression with opioids or bleeding with Warfarin. Type A reactions also include those effects that are indirectly related to the desired pharmacological effect of the concerned drug. Example 2: Dry mouth associated with tricyclic antidepressants.

4.2 Type B ADRs

These are relatively uncommon, unpredictable and non-dose related. These are responses not expected from the known actions of a drug and can only be discovered for the first time after the drug becomes available for general use. Idiosyncratic and immunological categories like these are detected during the pharmacovigilance phase of the drug's shelf life using Adverse Event reporting databases called Spontaneous Reporting System (SRS) like FAERS or Yellow Card. Example: Anaphylaxis with penicillin or skin rashes with antibiotics.

4.3 Type C ADRs

Cumulative toxic effect of a drug when used over time leads to this type of ADRs. In other words, these are continuous reactions that persist for relatively long time intervals. In this case, AEs increase gradually. Example: Osteonecrosis of the jaw with bisphosphonates. The SRS contains Type C ADR data as it collects data for the life time of the drug.

5. Experimental Results

5.1 Natural Thresholds

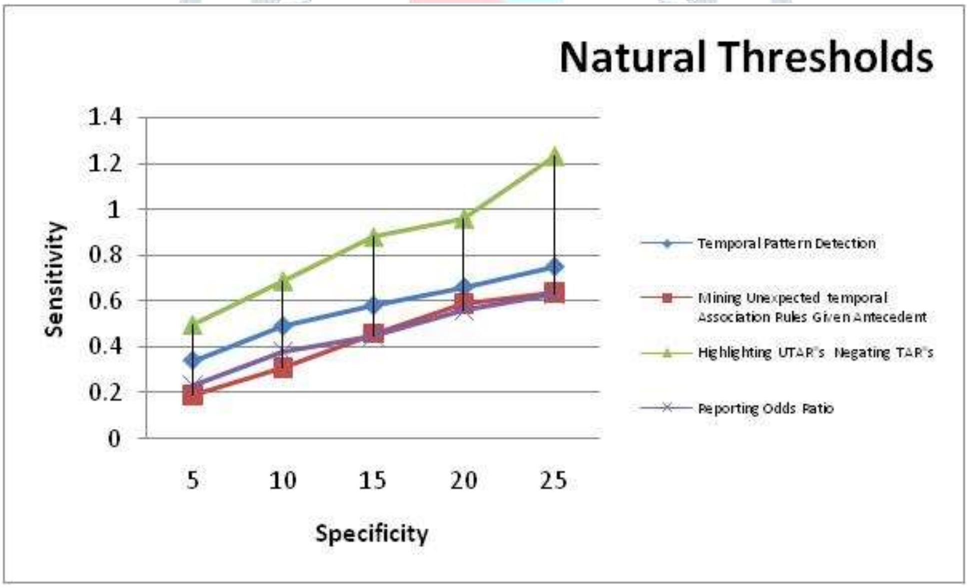

Figure 2: Natural Thresholds

The Comparison chart of natural thresholds shows the different values of methods. No of records in x axis and sequence level in Y axis. When compared each method and some method values are high and low.

5.2 General Ranking

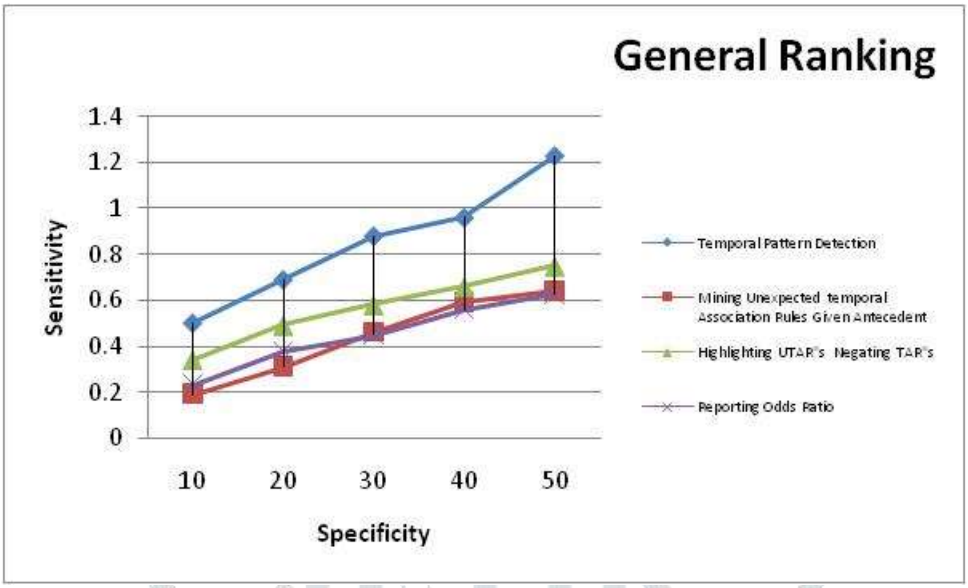

Figure 3: General Ranking

The Comparison chart of general ranking shows the different values of methods. No of records in x axis and sequence level in Y axis. When compared each method and some method values are high and low.

5.3 Unlabeled Data Signals

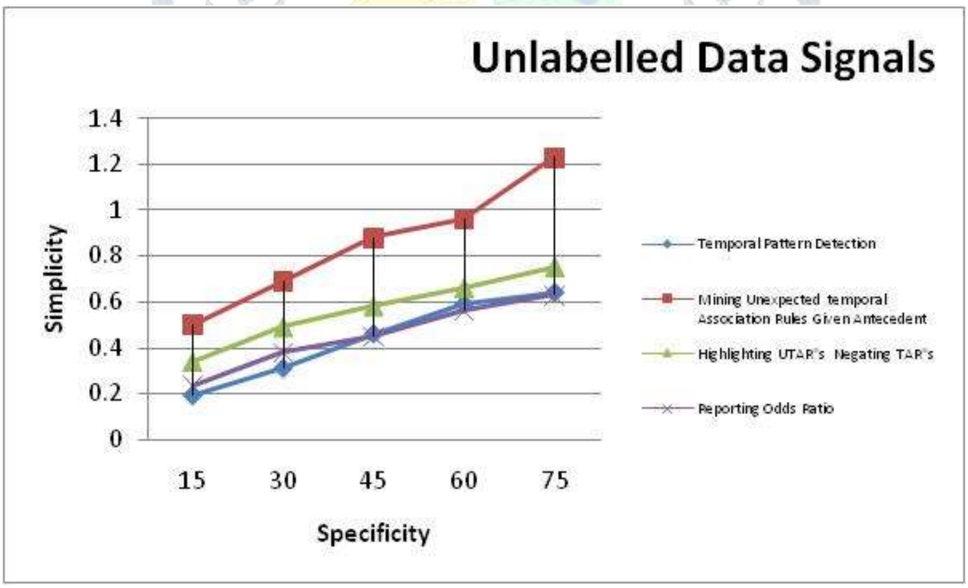

Figure 4: Unlabelled Data Signals

The Comparison chart of unlabelled data signals shows the different values of methods. No of records in x axis and sequence level in Y axis. When compared each method and some method values are high and low.

Conclusion

The drugs are associated based on generics rather than brand names. This increases the applicability of the results extensively. An integrated tool is developed to mine drug-drug interactions using the Adverse Event database. The methods and results obtained are analysed based on the contributions of drugs method detections. This paper have chosen few methods involved into the drug adverse reactions.

References:

[1] Sohini Sinha, Shubham Sharma & Ms. Srividhya. S., "Patient's Medication Analysis Using Probabilistic Aspect Mining Model", Imperial Journal of Interdisciplinary Research (IJIR) Vol-2, Issue-5, 2016 ISSN: 2454-1362.

[2] Abeed Sarker and Graciela Gonzalez., "Data, tools and resources for mining social media drug chatter", Proceedings of the Fifth Workshop on Building and Evaluating Resources for Biomedical Text Mining (BioTxtM 2016), pages 99–107, Osaka, Japan, December 12th 2016.

[3] K. Arutchelvan and Dr. Pon Periasamy, "Prognosis of Lung Cancer Using Data Mining Techniques", International Journal of Advanced Research in Computer Science and Software Engineering , Volume 6, Issue 3, March 2016 ISSN: 2277 128X

[4] Ken Naganuma Masayuki Yoshino, Ph.D. Hisayoshi Sato, Ph.D. Yoshinori Sato, "Privacy-preserving Analysis Technique for Secure, Cloud-based Big Data Analytics", Hitachi Review Vol. 63 (2014), No. 9.

[5] Megha Sinha, Sonu Kumar, "Optimization Of Data Mining Techniques Using Genetic Algorithm", International Journal Of Computer Engineering And Applications, Volume Xii, Special Issue, May 2018, Www.Ijcea.Com ISSN 2321-3469.

[7] Deepalakshmi. K. M, Dr. M. S. Vijaya, "ANALYSIS ON ADVERSE DRUG REACTION USING DIFFERENT DATA MINING TECHNIQUES", IJRSET SEPTEMBER 2016 Volume 3, Issue 9, ISSN 2394-739X.

[8] Lian Duan, Mohammad Khoshneshin, W. Nick Street, and Mei Liu, "Adverse Drug Effect Detection", IEEE TRANSACTIONS ON INFORMATION TECHNOLOGY IN BIOMEDICINE, VOL. X, NO. X, JANUARY 2013.

[9] Tu-Bao Ho, Ly Le, Dang Tran Thai, Siriwon Taewijit, "Data-driven Approach to Detect and Predict Adverse Drug Reactions", © 2014 Bentham Science Publishers.

[10] Aron Henriksson, Maria Kvist, Martin Hassel, Hercules Dalianis, “Exploration of Adverse Drug Reactions in Semantic Vector Space Models of Clinical Text”, Appearing in Proceedings of the 29 th International Conference on Machine Learning, Edinburgh, Scotland, UK, 2012.

[11]. Gabin Personeni, Emmanuel Bresso, Marie-Dominique Devignes, Michel Dumontier, Malika Smaïl-Tabbone and Adrien Coulet, “Discovering associations between adverse drug events using pattern structures and ontologies”, Journal of Biomedical Semantics (2017) 8:29 DOI 10.1186/s13326-017-0137-x.

[12]. Pathima Nusrath Hameed, Karin Verspoor, Snezana Kusljic and Saman Halgamuge, “A two-tiered unsupervised clustering approach for drug repositioning through heterogeneous data integration”, BMC Bioinformatics (2018) 19:129.