DOI Number: 10.5958/0974-1283.2019.00188.9

Sentiment Ontology Analysis for Adverse Drug Effect

P Rajalakshmi[1], S.P. Rajagopalan[2]

[1]Research Scholar, Department of Computer Science, Vels Institute of Science, Technology and Advanced Studies (VISTAS), Chennai, [2]Professor, School of Computer Science and Engineering of Technology, Chennai

Abstract

Adverse drug reactions (ADR) are happen more frequently based on different drugs. ADR should be monitored by Pharmacovigilance. Due to drug usage of single dose or combine dose ADR are happen. ADR can't be monitored in INDIA because of not follow the physician direction and also self-medication. Most of the Indians used drug store or online to purchase drugs. The proposed method utilised different websites and online drugs reviews with sentiment analysis and conjugate gradient method in neural network. In output, we propose the best way to monitor the ADR and drug harmfulness.

Keywords: *Pharmacovigilance, dosage, physician, online drug reviews, conjugate gradient.*

Introduction

Adverse Drug Reactions monitoring is an operation of continuously monitoring of unsuitable effect distrusted to be associated with the use of medicinal products. It encourages the collection of fair safety data observed aimed clinical practice in 'real life' conditions. The rules have been created to help similar professionals on understanding the significance of ADRs monitoring, methods of detailing ADR and the four fundamental parts of a ADRs case report to enhance drug safety. The basic parts include data about the patient, explanation of the adverse drug reactions, distrusted drug(s) and the reporter.

In modern countries, ADR is the fourth important cause of the death because of poor reporting of side effect after having drugs. The cause for poor reporting is insensitivity and carelessness in the patient. Moreover, the nurse in the hospital to observe patient ache from side effect and describing to the treating physician is less. The doctor orders the excess dose and unwarranted mix of the drug. The doctor invests a small amount of time in diagnosis patient's symptoms and never clarifies about the drug regimen. In india ADR is barely report to the ADR observing centre of Pharmacovigilance Programme of India (PvPI). Reporting of ADR to PvPI is moderate since a large portion of the population of patients looks at the drug stores. Moreover, the doctor holds another medicine qualification and suggest for the allopathic drugs.

Compared association asssess utilising the percentage of false positive signals between a given numbers of the most extremely ranked drug-case combinations concording to the values of the association assess. By regarding 150 drugs and 100 adverse cases, the percentages of false positives, between the 500 most extremely ranked drug-case couples; changes from 1.1% to 53.4% (mean over 1000 simulated datasets). As the calculation guided to very different outputs, we could name which calculates seem to be the most applicable for pharmacovigilance [1]. Reviews of prescription drugs and those evaluated to be medium-risk will probably be identified utilising search query data. These determinations recommend that amassed internet search engine data can be utilised to help in early warning of faulty bunches of pharmaceuticals [2]. Reviews of prescription drugs and those evaluated to be medium-risk will probably be identified utilising search query data. These determinations recommend that amassed internet search engine data can be utilised to help in early warning of faulty bunches of pharmaceuticals [3]. It dependably shortlists not only six referred to ADRs, as well as another ADR, flucloxacillin potentially causing hepatitis, which our algorithm originators and examination runners have not known before the tests. The MUTARC execute substantially more effectively than existing methods. Here, shows the considerable potential along the new direction of ADR signal propagation from healthcare regulatory databases [4]. Broad investigation on medical forums dataset shows

that constrained information entropy (CIE) results from the state-of-the art co-occurrence based techniques, particularly in rare ADRs detection [5]. A similar study between the evaluated frequencies has the similar dissemination trend. These outputs recommend that the naive Bayesian model based on gene-ADR connection network can assist as an economical and efficient tool in quick ADRs appraisal [6]. Drug-target association studies are predicted adverse side effects or drugs unexpected therapeutic. In silico prevision of potential association are important and can concentrate effort on in vitro tests. [7]. The next phase demonstrates the use of FMECA to the process, to identify the vital problem and measure the risk decreases obtained utilising a specific IT tool, contrasted with the utilisation of current resources [8]. The literal adverse drug effects on a handed dataset cannot be completely decided; we make utilisation of the simulated observational medical outcomes partnership (OMOP) dataset built with the determined adverse drug effects to assess our strategies. Experimental outcomes demonstrate the convenience of the developed pattern discovery technique on the simulated OMOP dataset by enhancing the standard baseline algorithm-chi-square-by 23.83% [9]. A PHARMA 2.0 atelematicsn integrated telematics system proposed at decreasing Adverse Drug Events (ADEs) in the phases of drug provision, distribution, transcription and administration. The developed system is grounded on three sub-systems: an RFID-based drug container and dispenser, a CPOE (computerised Prescription Order Entry), and middleware system. The perception and management of prevision and administration data are covered through a web application planned to conform by international usability regulation [10].

Methodology

The methodology of the Pharmacovigilance is integrated the orderly manner. The below diagram describes the methodology adopted in the Pharmacovigilance.

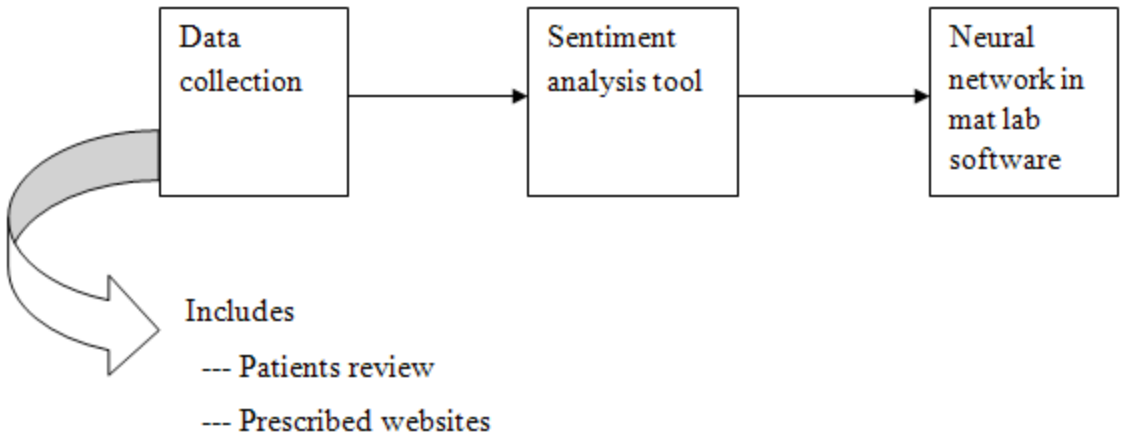

Figure 1. Methodology block diagram

Through the help of online modes, the data collections are done. In the online mode, the data collections are processed through the prescribed websites like as Webmd.com, drugs.com, druglib.com. emedicine.com and so on. The patient's reviews are collected from the website and processed through the online sentiment analysis tool.

The prescribed websites like WebMD.com and Drugs.com. The websites give information on drugs for healthcare professionals and consumers. The drugs which have utilised for the analysis is namely as Benadryl, Vicodin, Adderall, amoxicillin, cetirizine and naproxen.

In Naproxen drug, the data are collected concerning the different filter conditions such as Pain, A disorder characterised by Stiff, Tender and painful, Rheumatoid Arthritis, Joint Damage making Pain and loss of function, inflammation of the tendon, A migraine headache, Inflammation of the lining of a joint, others, and All reviews, conditions and overall ratings. In such a way the data are collected to the other drugs concerning the filter conditions.

Sentiment Analysis

Sentiment analysis is contextual mining of content which recognises and extracts subjective data in the source material. To understand the customer's sentiment of your item, service or brand while monitoring online conversations. Parallel Dots Text Analytics APIs give convenient and various set of Natural Language Understanding (NLU) algorithms to detect emotion or sentiment of any document, find high entities in them or remove profane from them.

Text analysis is the procedure of filiation of high-end data through accomplished patterns and courses in a piece of text. Our parallel dots text analysis APIs execute importantly better than traditional NLP methods.

Online conversations of the reviewer text have been taken as input and gives the structured image as the output by conquering the meaning of that text. Our parallel dots text analysis APIs shows the sentiment analysis as positive, neutral and negative in Sentiment Analysis Tool . The text analysis calculates the high-quality data from the unstructured content collection from online posts such as WebMD.com and drugs.com. Once the output percentage comes out for the drug obtain from various website and the percentage values state-of-the-art dataset and then the data is analysed with conjugate gradient neural network through Matlab software.

Conjugate Gradient Neural Network

We assume a single-hidden-layer training network with a linear output. With the size of $L \times ML \times M$, an input weights matrix is $W^{in}W^{in}$, with M non-linear hidden units individual with the similar non-linear function f(x), and N is a size of the linear output layer. The input data is held in the matrix X of size $K \times L$ $K \times L$. If a bias condition is needed in $W^{in}W^{in}$, a column of ones is considered in X. For a bunch of K input samples, individual of size $1 \times L1 \times L$, the reaction of the input linear addition stage matrix $Z^{in}Z^{in}$ of size $K \times MK \times M$ is measured using

$$Z^{in} = XW^{in}Z^{in} = XW^{in} \text{ ----- (1)}$$

The reaction of the hidden layer is produced by the matrix A of size $K \times MK \times M$ and measured using

A=f(Z)--------- (2)

Where f(Z) is a non-linear function. Example functions utilised by practician combines

the relu (rectified linear), tanh, logistic and square functions. The reaction of the output layer is linear and produced by the matrix Y of size $K \times NK \times N$ utilising

$$Y = AW^{out}Y = AW^{out} \text{.....(3)}$$

Where the output weight matrix $W^{out}W^{out}$ has the size as $M \times NM \times N$.

For the function of training an N-class classifier, we produce a $K \times NK \times N$ indicator matrix T with components $t_{k,n}t_{k,n}$ to represent the desired label or the target for the associated input data X. If the desired label of kth input data is in class n then $t_{k,n}t_{k,n}$is 1 and other N-1 components of the kth row of T are 0.

For the conjugate gradient, the $W^{in}\,W^{in}$ is set to random numbers, and the $W^{out}W^{out}$ are calculated utilising linear regression. The cost function for the network is the sum square error of the output and is produced by

$$E = \sum_{all\ elements}(Y - T)^2 = \sum_{all\ elements}(AW^{out} - T)^2$$

---------(4)

Where square indicates the element by element.

Regression analysis is a method for simulating the relationship among two or more variables. Regression method quantitatively explains the variability between the observations by dividing an observation into two parts. The first part of this decay is the predicted portion having the feature that can be assigned to all the observations believed as a group in a parametric system. The other portion is called as residual, is the difference between the predicted values and observed values must be assigned to unknown sources.

Simulated Results and Discussions

Table 1. Sentiment Analysis for Naproxen drug

S.NO	Naproxen drug by filter conditions	Positive	Neutral	Negative
1	Pain	28.20%	6.90%	64.90%
2	A disorder characterised by Stiff, Tender and painful	3.90%	4.50%	91.60%
3	Rheumatoid Arthritis	5.60%	7.50%	86.80%
4	Joint Damage making Pain and loss of function	5.60%	7.50%	86.80%
5	inflammation of the tendon	59.40%	6.90%	33.80%
6	painful periods	8.30%	2.40%	89.30%
7	A migraine headache	14.90%	37.00%	48.10%
8	Inflammation of the lining of a joint	26.50%	7.30%	66.20%
9	Inflammation of the covering of the Tendon	19.00%	4.70%	76.30%
10	Inflammation of the Sac surrounding the joint-Bursitis	5.70%	4.00%	90.30%
11	Rheumatic Disease-Causing Pain and stiffness in Backbone	2.00%	2.70%	95.30%
12	Gout	64.70%	15.80%	19.50%
13	Head pain	9.00%	2.40%	88.70%
14	Joint Inflammatory Disease in children and young adults	65.20%	4.50%	30.30%
15	Fever	51.10	16.60	32.30
16	Others	54.50%	3.90%	41.70%
17	All reviews, conditions and overall ratings	28.20%	6.90%	64.90%

Naproxen is utilised to relieve pain from different conditions such as dental pain, menstrual cramps, headaches, tendonitis, and muscle aches. It also decreases swelling, pain, and joint stiffness caused by bursitis, arthritis and gout attacks. The patients give their feeling as reviews by the intake of Naproxen drug. The sentiment analysis produces the result in the percentage as positive, neutral and negative respectively. According to the percentage value, we justify that the drug is good or bad to health. The results of Naproxen drug has shown in table 1.

Table 2. Sentiment Analysis for Vicodin drug

S.NO	Vicodin drug by filter conditions	Positive	Neutral	Negative
1	Pain	8.20%	11.50%	80.30%
2	Others	5.30%	1.80%	92.90%
3	All reviews, conditions and overall ratings	29.70%	11.50%	58.80%

Vicodin is a costless drug utilised to treat moderate to severe pain. The drug is more popular than the other drugs in the market. The patients give their feeling as reviews by the intake of Vicodin drug. The sentiment analysis produces the result in the percentage as positive, neutral and negative respectively. According to the percentage value, we justify that the drug is good or bad to health. The results of Vicodin drug have shown in table 2.

Regression value (R) foregrounds the correlation between the outputs and the targets. A value of R equals to one or zero suggests a closer relationship or a random relationship respectively. Based on the regression curve received (from fig.2 to fig.3), for various cases under the comparison of conjugate gradient methods to produce a statistical report. Figure 2 shows the regression value of social post in website, from the output R-value, is about 0.99801 proves that the reviews are matched between the drugs used at different conditions. Figure 3 demonstrates the regression value of social post in website, from the output R-value is about 0.76441 proves that the reviews are mismatched between the drugs used at different conditions. If the R value is negative, which demonstrate that the reviewer's comments are weak in feelings.

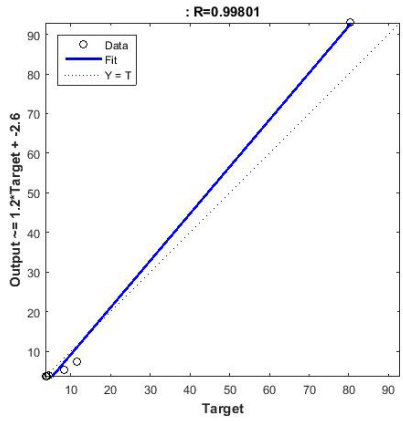

Fig. 2 Regression analysis for Vicodin drug

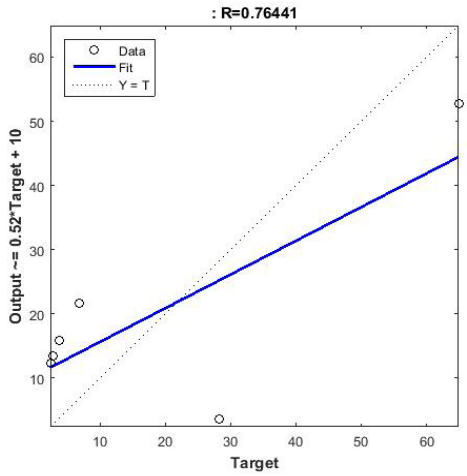

Fig 3. Regression analysis for naproxen drug

Table 3. Regression analysis for Naproxen drug

S.NO	Compared with the different conditions	Regression analysis
1	Pain vs Disorder characterised by Stiff, Tender and painful	0.99441
2	Pain vs Rheumatoid Arthritis	0.85411
3	Pain vs Joint Damage causing Pain and loss of function	0.93409
4	Pain vs inflammation of the tendon	0.94092
5	Pain vs painful periods	0.93793
6	Pain vs Migraine headache	0.74667
7	Pain vs Inflammation of the lining of a joint	0.99926
8	Pain vs Inflammation of the covering of the Tendon	0.99676
9	Pain vs Inflammation of the Sac Surrounding the joint-Bursitis	0.92407
10	Pain vs Rheumatic Disease-Causing Pain and stiffness in Backbone	0.89618
11	Pain vs Gout	0.76441

Cont... **Table 3. Regression analysis for Naproxen drug**

12	Pain vs Head pain	0.94551
13	Pain vs Joint Inflammatory Disease in children and young adults	0.69352
14	Pain vs Fever	0.78022
15	Pain vs others	0.79028
16	Pain vs all reviews, conditions and overall rating	0.99973

Conclusion

The reviews are collected from the prescribed websites like WebMD.com and drugs.com reports are processed for sentiment analysis. These processed sentiment analysis output are analysed in conjugate gradient neural network in MATLAB, and the regression value is determined and presented in the graph. This analysis gives an efficient way to detect the toxicity of the drug. Furthermore, the result demonstrates that the data collection from the websites proves to be more effective for the adverse drug reaction.

Ethical Clearance: VISTAS – Editorial Board, Chennai

Mode of Payment : Self

Conflict of Interest : Nil

References

1. Roux E, Thiessard F, Fourrier A, Be B. Evaluation of Statistical Association Measures for the Automatic Signal Generation in Pharmacovigilance. IEEE Trans Inf Technol Biomed. 2005;9(4):518–27.
2. Yom-tov E, Member S. Predicting Drug Recalls from Internet Search Engine Queries. IEEE J Transl Eng Heal Med. 2017;5(c):1–6.
3. Reps JM, Garibaldi JM, Aickelin U, Soria D, Gibson JE, Hubbard RB. A Novel Semisupervised Algorithm for Rare Prescription Side Effect Discovery. IEEE J Biomed Heal INFORMATICS. 2014;18(2):537–47.
4. Jin HW, Chen J, He H, Williams GJ, Kelman C, Keefe CMO. Mining Unexpected Temporal Associations : Applications in Detecting Adverse Drug Reactions. IEEE Trans Inf Technol Biomed. 2008;12(4):488–500.
5. Zheng Y, Lan C, Peng H, Li J. Using Constrained Information Entropy to Detect Rare Adverse Drug Reactions from Medical Forums. 2016 38th Annu Int Conf IEEE Eng Med Biol Soc. 2016;2(1):2460–3.
6. Xiang Y, Liu K, Cheng X, Cheng C. Rapid Assessment of Adverse Drug Reactions by Statistical Solution of Gene Association Network. IEEE Trans Comput Biol Bioinforma. 2014;5(1):1–7.
7. Fakhraei S, Huang B, Raschid L, Getoor L. Network-Based Drug-Target Interaction Prediction with Probabilistic Soft Logic. IEEE/ACMTRANSACTIONS Comput Biol Bioinforma. 2014;11(5):775–87.
8. Iadanza E, Ieee M, Gaudio F, Marini F. The diagnostic-therapeutic process . Workflow analysis and risk management with IT tools . 35th Annu Int Conf IEEE EMBS. 2013;2(1):4763–6.
9. Duan L, Khoshneshin M, Street WN, Liu M. Adverse Drug Effect Detection. IEEE J Biomed Heal INFORMATICS. 2013;17(2):305–11.
10. Iadanza E, Pettenati MC, Bianchi L, Turchi S, Ciofi L, Pirri F, et al. Telematics Integrated System to Perform Drugs Prescription and Administration Reducing Adverse Drug Events. 34th Annu Int Conf IEEE EMBS. 2012;2(1):6082–5.