Predicting Adverse Effects of Drugs

Sofia Alexandra Machado Novais[1], Rui Camacho[2]

***Abstract*—Currently, the amount of drugs available is very large. Its therapeutic importance is indisputable. Although its production is performed using safe and security criteria, the taking of some drugs may involve risks. When drugs are taken, Adverse Drug Reactions (ADRs) occur and may have serious consequences for the health of the person.**

Predictive Data Mining models can be a valuable aid to help in the task of predicting unknown ADRs.

The main goal of this report is to assess how useful Data Mining methods and algorithms can predict the adverse effect of drugs. In general, the results obtained were promising and encourage the continuation of this line of research.

***Keywords*—Drugs, ADR's, Data Mining.**

I. Introduction

Adverse Drug Reactions (ADRs) are a significant public health concern, because they can potentiate serious injury and even lead to mortality of individuals [1]. The World Health Organization (WHO) defines ADR as "any harmful or undesired effect which manifests itself after drug administration at doses normally used in man for the prophylaxis, diagnosis or treatment of a disease" [2].

The development of predictive procedures would help health professionals and pharmaceutical companies, since they could avoid many unwanted and unknown serious effects before marketing the drug [3].

According to Fayyad et al. [4], "Data Mining (DM) consists in the accomplishment of data analysis and the application of discovery algorithms that, under certain computational limitations, produce a set of patterns of certain data".

In this work two experiments, involving the use of Recommender Systems (RSs) and Classification algorithms were performed. RSs have as their main objective to filter information and provide researchers with only relevant and highly correlated information [5]. The recommendation algorithms used were Matrix Factorization (MF), Slope-One and User k-NN. Commonly, DM has the ability to perform, among others, classification tasks. The Classification is known as "the process of learning a function that maps (classifies) a given object of interest into one of the possible classes" [6]. We have also defined a classification task and have used Decision Tree, Random Forest, Naive Bayes and Support Vector Machines (SVM).

II. Experimental Evaluation

The dissertation work is composed by a set of DM experiments. The case studies of this project used data sets from the ADReCS databases (http://bioinf.xmu.edu.cn/ADReCS/download.jsp). The data and information available in these databases were pre-processed to make them adequate for the analysis tools. Subsequently, with the aid of enrichment tools, namely PaDEL and Open Babel, the processed data were enriched, in order to perform ADR prediction on new drugs. Lastly, the data were transformed into a suitable format for the DM methods and tools and finally, models were constructed.

We have used RSs and then, evolved to classification tasks. RS was used as predictive methods only, whereas the classification can provide potential explanations for the ADR apart from the predictive facility.

For the classification tasks, a previous data enrichment stage with molecular descriptors was performed.

The DM tool chosen to carry out the experiments was RapidMiner, since it has an easy to use interface and enables the definition of the DM steps workflow.

Models performance were assessed using obtaining the metrics Accuracy, Precision, Recall and F-Measure. In the results, only the values obtained from the Accuracy metric are presented, since this is the one that gives the best evaluation of the overall performance of the model [7].

The Accuracy is defined as below:

$$\frac{tp + tn}{tp + fp + tn + fn} \tag{1}$$

where, ***tp*** - true positive for ***C***; ***fp*** - false positive for ***C***; ***fn*** - false negative for ***C***; ***tn*** - true negative for ***C***.

Note that the higher the percentage of the metric, the better the performance of the model.

A. Experiment 1

This experiment focused on the prediction of existing drug-adverse effects pairs. Thus, we only used information from the databases and used algorithms of RSs that was responsible for the prediction.

In order for the RapidMiner recommendation process operators to apply the RSs, it was necessary to use a name recommender extension provided by RapidMiner. The algorithms exploited were those of predictive rating with collaborative filtering. In this project, we have used 3: Matrix Factorization (MF), Slope One and User k-NN.

To assess the robustness of the methods, we have produced two data sets by introducing two different noise-levels in the original data. The data sets used was prepared with 5% and 10% noise and the procedure previously described was performed 2 times for each noise level.

[1]novais.sofia@hotmail.com, FEUP
[2]Associate Professor, DEI/FEUP, rcamacho@fe.up.pt

B. Experiment 2

In this experiment, we have used 26 groups of ADRs ids. ADRs ids are organized in an hierarchy and we have merged the levels of the hierarchy into the top level ids of the hierarchy. Thus, we was obtained 26 nodes, corresponding to the 26 groups of adverse effects explored.

The objective was mainly intended to find not only drug-adverse effects relationships, but also biochemical justifications for their existence. In addition to the available information in the databases, we have added data related to the molecular descriptors of each drug. In the course of the experiment, a selection of the 10 best attributes was made to understand which descriptors significantly influence the activity of the molecules.

The classification algorithms used were Decision Tree, Random Forest, Naive Bayes and SVM. Each of these classification algorithms was executed 26 times, which corresponds to the number of adverse effects groups in the study.

III. Results

A. Experiment 1

In the Table I, it is possible to observe the results obtained from the metric Accuracy for noise-free data, for the data with 5% of noise and for the data with 10% of noise.

TABLE I
RESULTS OBTAINED FROM THE METRIC ACCURACY FOR NOISE-FREE DATA, FOR THE DATA WITH 5% OF NOISE AND FOR THE DATA WITH 10% OF NOISE.

Accuracy (%)	MF	Slope One	User k-NN
Noise free	45.15 (+/-1.18)	21.57 (+/-0.12)	35.54 (+/-1.35)
5% of noise	39.43 (+/-0.98)	19.53 (+/-0.17)	31.63 (+/-1.23)
10% of noise	34.75 (+/-0.87)	16.35 (+/-0.10)	28.53 (+/-1.12)

The results shown in the table for data with 5% of noise and for the data with 10% of noise correspond to the Accuracy average and standard deviation of the results obtained in two performances performed.

Observing Table I, it is possible to see that the algorithm that obtained the best performance was MF. The worst results were obtained with the Slope One algorithm. Such results may be justified, since the MF algorithm is the only one that predicts, considering adverse effects and drugs.

Comparing the results presented, it is possible to state that the percentage of the value of the metric decreases with the increase of the noise in the data, which leads to the conclusion that the noise interferes and decreases the performance of the model.

B. Experiment 2

In the Table II, only the results obtained for Group 8 are presented, because it was the one that obtained the best results.

The best result of Accuracy metric is obtained when using the SVM algorithm - about 94.41%.

TABLE II
RESULTS OBTAINED FROM THE METRIC ACCURACY FOR GROUP8

	Accuracy (%)
Decision Tree	94.12 (+/-2.63)
Random Forest	93.82 (+/-2.06)
Naive Bayes	91.43 (+/-2.76)
Support Vector Machine	94.41 (+/-0.88)

IV. Conclusion

The occurrence of ADRs has been increasing notably in recent years and has reached the level of concern. Many studies are being carried out to predict them. In general, it has been observed that computing has proved essential in this area, since the development of predictive procedures could help to avoid many unwanted serious effects.

The main goal of this dissertation was to assess how useful Data Mining methods and algorithms could predict the adverse effect of drugs. Thus, two experiences were performed. Data set was taken from the ADReCS database.

In the first experiment, were used recommendation algorithms: MF, Slope One and User k-NN. The algorithm that obtained a model with better performance was the MF, since this algorithm obtained the value of the greater Accuracy metric (about 45.15%). However, the results obtained were poor.

In the second experiment, we have defined a classification tasks using only the top of the ADRs identifiers hierarchy. The classification algorithms used were Decision tree, Random Forest, Naive Bayes and SVM. Considering the results obtained, it can be stated that the algorithm that obtained a model with better performance was the SVM, since this algorithm obtained the best value of Accuracy (about 94.41%). The results obtained were satisfactory.

After analyzing the results achieved in the two experiments, it was observed that experiment 2 obtained better results. Thus, it is concluded that making individual predictions of adverse effects is quite complex, and when reducing the detail of the information, that is, when one goes up the level of hierarchy, the obtained results are better. It is important to note that this improvement of results also occurs when done with the Feature Selection, since a pre-selection of the best attributes occurs.

References

[1] A. M. Wilson, L. Thabane, and A. Holbrook, "Application of data mining techniques in pharmacovigilance." *British Journal of Clinical Pharmacology*, 2003.

[2] D. F. Kawano, L. R. L. Pereira, J. M. Ueta, and O. de Freitas, "Acidentes com os medicamentos: como minimiza-los?" *Revista Brasileira de Ciencias Farmaceuticas*, 2006.

[3] M. Yang, M. Kiang, and W. Shang, "Filtering big data from social media - building an early warning system for adverse drug reactions." *Journal of Biomedical Informatics*, 2015.

[4] C. O. Camilo and J. C. D. Silva, "Mineracao de dados: Conceitos, tarefas, metodos e ferramentas." *Instituto de Informatica Universidade Federal de Goias*, August 2009.

[5] F. Isinkaye, Y. Folajimi, and B. Ojokoh, "Recommendation systems: Principles, methods and evaluation." *Egyptian Informatics Journal*, 2015.

[6] T. A. de Araujo Matos, "Uma visao geral das principais tarefas de mineracao de dados," *Saber Digital*, April 2012.

[7] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks." *Information Processing and Management*, pp. 427–437, May 2009.