2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Ensemble ADR Detection from Spontaneous Reporting Data: A Positive Unlabeled Learning Perspective

Wen-Yang Lin and Tzu-Wei Tseng

***Abstract*—This paper presents our endeavor to fuse positive unlabeled and ensemble learning to shape a better ADR signal detector from spontaneous reporting data, like FAERS.**

***Clinical Relevance*—The results show that the proposed PU-ensemble methods surpass the non-PU ensemble with 18%, 144%, 19%, and 80% improvement on accuracy, precision, recall, and F-measure, respectively, of the detected ADR signals.**

I. Introduction

Spontaneous reporting systems (SRS) have long played the principal repository for post-marketing safety monitoring of approved drugs, aiming at early detection of significant ADR signals with a high association of drugs and adverse reactions. Unfortunately, the events reporting to SRS are not mandatory and lack authentic confirmation, especially the drug and adverse reaction information. Such uncertainty thwarts the efficacy of applying machine learning to discover ADR signals. Most contemporary ADR detection methods [1] thus are statistics-based approaches that rely on measuring the disportionality of drug-reaction pairs in the SRS data. The main disadvantages of these statistical approaches include no consistent measurement of disportionality and high false positive or false negative rates of the generated signals.

In this paper, we propose an ensemble ADR detection from the perspective of positive and unlabeled data learning, namely PU learning [2], a branch of uncertain learning. The available data for learning contains a substantial small set of positive examples and many unlabeled (uncertain) examples that may be positive or negative. By leveraging SIDER, a well-known ADR knowledge base, as the training set of positive examples and viewing SRS, e.g., FAERS, the unlabeled data, we design a novel ADR detection method featured by a two-layer ensemble of statistics-based detection methods.

II. Methods

The proposed PU-ensemble ADR detection is a two-layer ensemble mechanism. We regard SIDER as an initial training set composed of only positive examples, each of which represents a positive drug-reaction signal. The first layer fuses various statistical methods, including PRR, ROR, MHRA, Yule's Q, SPRT, and BCPNN, to determine the certainty (positive or negative) of each drug-reaction pair in FAERS. Only those pairs obtaining complete voting consistency are added into the training set. The resulting training set is then fed to a second layer to train the final ensemble ADR classifier. In this stage, we used three different ensemble learning methods, AdaBoost-inspired ADR [3], and two variants of stacking method, LR (logistic regression) and RF (random forest). For convenience, we name the three PU-ensemble methods PU-AB, PU-LR, PU-RF.

III. Results

We compared our methods with the AdaBoost-inspired ADR method [3], using the FAERS datasets from 2004 to 2016; those from 2004Q1 to 2016Q3 (around 32M records) were for building the ensemble ADR detector, while 2016Q4 (about 0.1M records) was the testing set. Table I shows the performance results of all methods running on the testing set, measured by accuracy, precision, recall, and F-measure. The results demonstrate the benefit of incorporating PU learning. All three PU-ensemble methods, especially the two stacking-based variants, outperform the AdaBoost-inspired method, our previously proposed ensemble without PU learning.

TABLE I. Performance Results

Method	Accuracy	Precision	Recall	F-measure
AdaBoost-ins	0.65	0.41	0.57	0.48
PU-AB	0.64	**1.0**	0.64	0.78
PU-LR	**0.84**	**1.0**	**0.84**	**0.91**
PU-RF	0.82	**1.0**	0.82	0.90

IV. Discussion & Conclusion

The SRS data is naturally uncertain, making from which predicting ADR signals highly challenging. In this paper, we have demonstrated promising results in fusing ensemble and PU learning. Overall, the proposed PU-ensemble methods surpass AdaBoost-inspired with 18%, 144%, 19%, and 80% improvement on accuracy, precision, recall, and F-measure, respectively.

References

[1] E. Roux et al., "Evaluation of statistical association measures for the automatic signal generation in pharmacovigilance," *IEEE Trans. on Information Technology in Biomedicine*, vol. 9, no. 4, pp. 518-527, 2005.

[2] J. Bekker and J. Davis, "Learning from positive and unlabeled data: a survey," *Machine Learning*, vol. 109, pp. 719-760, 2020.

[3] W.Y. Lin and I Dai, "An AdaBoost-inspired ensemble method for ADR signal detection," *Proc. 8th IEEE Int. Conf. on Healthcare Informatics*, 2020, pp. 1-2.

*Research supported by Ministry of Science and Technology of Taiwan under grant no. MOST108-2221-E-390-021.

Wen-Yang Lin is with the Dept. of Computer Science and Information Engineering, National University of Kaohsiung, Kaohsiung 811, Taiwan, phone: +886-7-5919517; e-mail: wylin@nuk.edu.tw