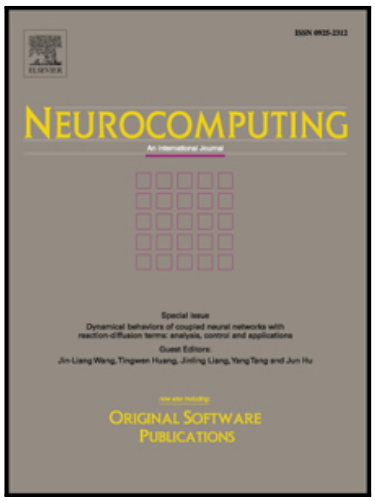

Feature-derived Graph Regularized Matrix Factorization for Predicting Drug Side Effects

Wen Zhang , Xinrui Liu , Yanlin Chen , Wenjian Wu , Wei Wang , Xiaohong Li

PII: S0925-2312(18)30126-7
DOI: 10.1016/j.neucom.2018.01.085
Reference: NEUCOM 19287

To appear in: *Neurocomputing*

Received date: 10 November 2017
Revised date: 12 January 2018
Accepted date: 30 January 2018

Please cite this article as: Wen Zhang , Xinrui Liu , Yanlin Chen , Wenjian Wu , Wei Wang , Xiaohong Li , Feature-derived Graph Regularized Matrix Factorization for Predicting Drug Side Effects, *Neurocomputing* (2018), doi: 10.1016/j.neucom.2018.01.085

Feature-derived Graph Regularized Matrix Factorization for Predicting Drug Side Effects

Wen Zhang[1,*], Xinrui Liu[1], Yanlin Chen[2],Wenjian Wu[3], Wei Wang[4], Xiaohong Li[1,*]

1. School of Computer, Wuhan University, Wuhan 430072, China

2. School of Mathematics and Statistics, Wuhan University, Wuhan 430072, China

3. School of Electronic Information, Wuhan University, Wuhan 430072, China

4. Key Laboratory of Combinatorial Biosynthesis and Drug Discovery, Ministry of Education, Wuhan University School of Pharmaceutical Sciences, Wuhan 430071, China

zhangwen@whu.edu.cn, liuxinrui@whu.edu.cn, chenyanlin@whu.edu.cn, westonwu@whu.edu.cn, waw6@whu.edu.cn, leexh@whu.edu.cn

*co-corresponding author

Abstract: Drug side effects are one of the major concerns in the drug discovery. A great number of machine learning-based computational methods have been proposed to predict drug side effects. Many methods combine diverse drug features for the side effect prediction, but complete features are not available for all drugs. Drug side effect prediction with limited information is challenging and meaningful. In this paper, we propose a novel computational method "feature-derived graph regularized matrix factorization" (FGRMF), which predicts unobserved side effects for approved drugs based on known drug-side effect associations and available drug features. FGRMF projects the drug-side effect association relationship into the low-dimensional space, which uncovers the latent features of drugs and side effects. A graph is constructed based on individual drug features, and the graph regularization which preserves the structure of the drug graph is incorporated into FGRMF. FGRMF is different from the traditional matrix factorization technique, and can take the biomedical context into account. In the computational experiments, FGRMF can produce satisfying results, and outperforms benchmark side effect prediction methods on the benchmark datasets. When complete features are available, we can extend FGRMF to integrate diverse features. We develop a web server to facilitate drug side effect prediction, available at: http://www.bioinfotech.cn/FGRMF/.

Key Words: drug feature; side effect; graph regularization; matrix factorization

1 Background

Drugs are chemicals that treat, cure, prevent, or diagnose diseases, and are beneficial for human health. Almost, all drugs have side effects, and unintended side effects may do harm to human and lead to serious consequences. Identifying side effects of drugs is meaningful and urgent [1, 2].

Wet methods rely on a counter screen of candidate compounds against enzymes and receptors in vitro to identify drug side effects, and wet methods are usually costly and time-consuming. Therefore, computational methods were proposed to screen possible side effects and complement wet experiments. Traditional computational methods utilized the structure-activity relationship or structure-property relationship [3-6] to make predictions. For example, Fliri [3] transformed the side effect data derived from prescription drug labels into effect spectra, and then diagnosed drug side effects. Fukuzaki [4] proposed a method to predict side effects using sub-pathways that share correlated modifications of gene-expression profiles. Hammann [5] presented a