



Maryam Lotfi Shahreza^{1*}, Nasser Ghadiri²

¹Department of Computer Engineering, Shahreza Campus, University of Isfahan, Iran;

²Department of Electrical and Computer Engineering, Isfahan University of Technology, Isfahan 84156-83111, Iran.

*Corresponding Author's E-mail: (maryam.lotfi@gmail.com)

Abstract

The process of finding additional indications for existing drugs is known as Drug Repositioning (DR). Computational methods can guide wet lab experimental design by narrowing the scope of candidate targets to accelerate drug discovery and can provide supporting evidence for experimental results.

We have developed a label propagation method to predict drug-target, drug-disease and disease-target interactions (Heter-LP) which integrates various data sources at different levels. Here its application on identifying of most probable drugs and targets of three challenged topics for drug companies (rare diseases, animal health and aging) are investigated.

Keywords: Semi-supervised learning, heterogeneous networks, systems biology

Introduction

Although DR has attracted many researchers' attentions recently and various computational methods have been proposed in this area, their effectiveness is a significant challenge preventing them to be widely accepted. Given the nature of existing data, the use of network modelling and semi-supervised learning methods is crucial in this domain [1].

The use of heterogeneous networks in DR is motivated by the fact that drugs tend to take effect via interaction with one or more protein targets within a cell. Therefore, it is necessary to consider drugs, protein targets, and diseases simultaneously to investigate their interrelationships.

Materials and Methods

Heter-LP is a semi-supervised machine learning method, based on label propagation over a heterogeneous network. Heter-LP consists of three main steps: data modeling, projection, and label propagation. In data modeling, different level and various data related to drugs, diseases and protein targets have been integrated in one heterogeneous network. The using data have been gathered from some of the most important biological datasets such as PubChem, SIDER, KEGG, DisGeNET, TTD and DrugBank and by using of some packages in R such as DOSE, GOSemSim and HPOSim. In the projection phase, topological similarities of each pair of concepts were computed. Finally, the proposed label propagation algorithm is applied [2].

Results

Some of evaluated applications of Heter-LP in drug discovery are presented here:

- For rare diseases: Adrenocortical Carcinoma (ACC) is one of the rare diseases, which involves the adrenocortical glands. Heter-LP predicted items related to ACC are Cosyntropin (drug) and Hsa:3480 (IGF-1R), Hsa:7153 (TOP2A), Hsa:1717 (DHCR7) and Hsa:4157 (MC1R) (protein targets).
- In animal health: Mastitis is the most prevalent and costly disease in dairy cattle industry, is defined as inflammation of the mammary gland caused by infectious agents [3]. Top predicted drugs by Heter-LP for clinical mastitis are Cefoperazone, Meloxicam, Cephapirin, Cephalexin, Oxytetracycline, Cinoxacin and Ketoprofen.
- In prevention approaches of aging: Aging is the change in the body as a result of becoming older. Heter-LP predicted items related to aging are Ketanserin, Divalproex sodium, Ciclopirox (drugs) and HSA:1812, HSA:1497, HSA:5742 (protein targets).

Discussion, Conclusion and Suggestions

By reviewing of related published articles, It seems that predicted new interactions by Heter-LP for three mentioned concepts are worthy ones. Pharmaceutical researchers can use these predictions to narrow the scope of their research. Also a comprehensive statistical analysis of its performance in general is presented in [2].

Moreover, the proposed method has some general advantages which include:

- No need to any inadvisable preprocessing of data.
- No need to know the negative samples.
- Heter-LP can predict interactions of *new* drugs (where a drug has no known target) and *new* targets.
- It is able to predict both trivial and non-trivial interactions.
- The integration of various data has been done perfectly.
- It applies both local as well as global network features

Lastly, we are confident that performance will continue to increase as more accurate and complete input data become available.

References

- [1] Lotfi Shahreza, Maryam, Nasser Ghadiri, Seyed Rasoul Mousavi, Jaleh Varshosaz, and James R. Green. 2018. "A Review of Network-Based Approaches to Drug Repositioning." *Briefings in Bioinformatics* 19 (5): 878–92. <https://doi.org/10.1093/bib/bbx017>.
- [2] Lotfi Shahreza, Maryam, Nasser Ghadiri, Seyed Rasoul Mousavi, Jaleh Varshosaz, and James R. Green. 2017. "Heter-LP: A Heterogeneous Label Propagation Algorithm and Its Application in Drug Repositioning." *Journal of Biomedical Informatics* 68 (April): 167–83. <https://doi.org/10.1016/j.jbi.2017.03.006>.
- [3] Sharifi, Somayeh, Abbas Pakdel, Mansour Ebrahimi, James M. Reecy, Samaneh Fazeli Farsani, and Esmaeil Ebrahimie. 2018. "Integration of Machine Learning and Meta-Analysis Identifies the Transcriptomic Bio-Signature of Mastitis Disease in Cattle." *PLoS One* 13 (2): e0191227. <https://doi.org/10.1371/journal.pone.0191227>.