



THE 2ND INTERNATIONAL CONFERENCE ON
SYSTEMS MEDICINE, AI & DRUG REPURPOSING

Enhancing drug repurposing through graph neural networks and link prediction

Lucía Prieto-Santamaría^{1,2}, Adrián Ayuso-Muñoz^{1,2}, Alejandro Rodríguez-González^{1,2}

¹ Escuela Técnica Superior de Ingenieros Informáticos, Universidad Politécnica de Madrid (Spain)

² Centro de Tecnología Biomédica, Universidad Politécnica de Madrid (Spain)

Abstract

The use of computational methods in drug repurposing has become increasingly popular, and innovative artificial intelligence techniques have proven to be valuable tools for predicting new therapeutical opportunities of existing drugs. Within this framework, biomedical data can be represented as graphs, offering a highly expressive way to depict the underlying structure of the information. Integrating these graph-based data structures with deep learning models enhances the prediction of novel connections, such as potential associations between diseases and drugs, which can be regarded as repurposing possibilities.

In particular, we have employed the heterogeneous integrated biomedical information from the DISNET project (<https://disnet.ctb.upm.es>) [1], which compiles a set of different structured and unstructured data that revolve around diseases, and which is based in the concepts of the Human Disease Network (HDN) [2]. This information is represented as a multilayered complex network whose node types, in addition to diseases, include symptoms, genes, proteins, genetic variants, non-coding RNAs, biological pathways, drugs, and so on. And the links represent the connection between diseases and genes, diseases and drugs, the interactions between proteins (Protein-Protein Interactions – PPIs), and the interactions between drugs (Drug-Drug Interactions – DDIs), among others. We consider two derived graphs: a simplified version of the DISNET knowledge base that we name *simple* and its complete version that we name *complex*. The *simple graph* contains a limited set of information: diseases, symptoms, drugs and their relationships. The *complex graph* uses all the available information. Node types are “phenotype”, “drug”, “pathway”, “protein” and “drug-drug-interaction”; and link types are Disease-Drug (therapeutic) (“dis_dru_the”), Disease-Symptom (“dse_sym”), Disease-Protein (“dis_pro”), Disease-Pathway (“dis_pat”), Drug-Drug (“druA_druB”), Drug-Protein (“dru_pro”), Drug-Symptom (side effect) (“dru_sym_sef”), Drug-Symptom (indication) (“dru_sym_ind”), Protein-Protein (“proA_proB”), Protein-Pathway (“pro_pat”), DDI-Phenotype (“ddi_phe”), and DDI-Drug (“ddi_dru”).

From these two different networks, we have developed several models based on Graph Neural Network (GNN) pipelines to embed the information in the network (representing each node as a vector of features) to then decode these embedding vectors optimizing the prediction of new links of the type disease-drug. Firstly, REDIRECTION (dRug rEpurposing Disnet lInk pREdiCTION) was presented [3]. REDIRECTION is a model that was trained on the *simple graph*, which was developed under this encoder-decoder framework to predict ‘dis_dru_the’ links.



THE 2ND INTERNATIONAL CONFERENCE ON
SYSTEMS MEDICINE, AI & DRUG REPURPOSING

Afterwards, we presented DMSR (Drug Molecular Structure REDIRECTION), which incorporated drug molecular structures as the initial drug nodes' features in REDIRECTION [4]. This model also used GNNs to embed two-dimensional drug molecular structures and obtain the corresponding features. DMSR was originally trained on the *simple graph*, improving REDIRECTION accuracy. We also presented BEHOR (Bedirectional Edge and Hyperparameter Optimized REDIRECTION). In it, links were made undirected, hyperparameter optimisation took place and the *complex graph* was used [5]. And finally, we also developed a version of DMSR with the *complex graph*. The best accuracy for predicting new repurposing opportunities was obtained in this last model.

Keywords

Drug repurposing, DISNET knowledge base, Graph Neural Networks

References

- [1] G. Lagunes García, A. Rodríguez González, L. Prieto Santamaría, E. P. García del Valle, M. Zanin, and E. Menasalvas Ruiz, "DISNET: a framework for extracting phenotypic disease information from public sources," *PeerJ*, vol. 8, p. e8580, Feb. 2020, doi: 10.7717/peerj.8580.
- [2] K.-I. Goh, M. E. Cusick, D. Valle, B. Childs, M. Vidal, and A.-L. Barabási, "The human disease network," *PNAS*, vol. 104, no. 21, pp. 8685–8690, May 2007, doi: 10.1073/pnas.0701361104.
- [3] A. A. Muñoz *et al.*, "REDIRECTION: Generating drug repurposing hypotheses using link prediction with DISNET data," in *2022 IEEE 35th International Symposium on Computer-Based Medical Systems (CBMS)*, Jul. 2022, pp. 7–12. doi: 10.1109/CBMS55023.2022.00009.
- [4] A. Ayuso-Muñoz, L. Prieto-Santamaría, A. Álvarez-Pérez, B. Otero-Carrasco, E. Serrano, and A. Rodríguez-González, "Enhancing Drug Repurposing on Graphs by Integrating Drug Molecular Structure as Feature," in *2023 IEEE 36th International Symposium on Computer-Based Medical Systems (CBMS)*, Jun. 2023, pp. 192–197. doi: 10.1109/CBMS58004.2023.00215.
- [5] A. Ayuso Muñoz, L. Prieto Santamaría, E. Ugarte Carro, E. Serrano, and A. Rodriguez-Gonzalez, "Uncovering Hidden Therapeutic Indications Through Drug Repurposing with Graph Neural Networks and Heterogeneous Data." Rochester, NY, Mar. 24, 2023. doi: 10.2139/ssrn.4385667.