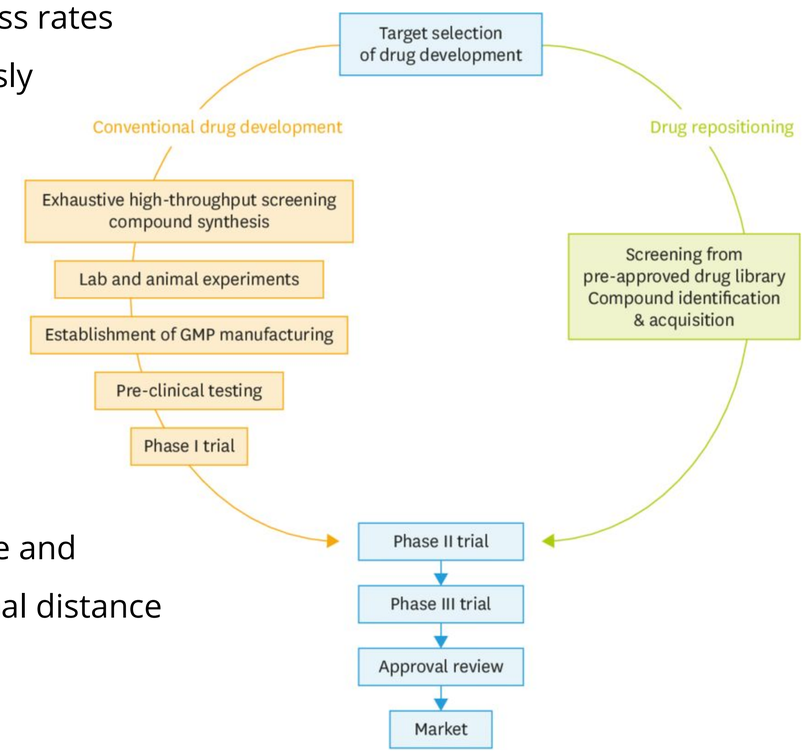
Identifying Drug Repositioning Candidates using Representation Learning on Heterogeneous Networks

Charles Tapley Hoyt[1,2], Rana Aldisi[1,2], Lingling Xu[1,2], Daniel Domingo-Fernández[1,2], Mehdi Ali[3], and Martin Hofmann-Apitius[1,2]

1. *Department of Bioinformatics, Fraunhofer Institute for Algorithms and Scientific Computing, Sankt Augustin 53754, Germany*
2. *Bonn-Aachen International Center for IT, Rheinische Friedrich-Wilhelms-Universität Bonn, Bonn 53115, Germany*
3. *Department of Enterprise Information Systems, Fraunhofer Institute for Intelligent Analysis and Information Systems, Sankt Augustin 53754, Germany*

MOTIVATION

De novo drug discovery and development is both an expensive and time-consuming process with low success rates often attributed to toxicity or lack of efficacy. Drug repositioning (DR) offers an alternative in which previously studied or approved drugs can be ascribed new uses. Image on right adapted from [1].



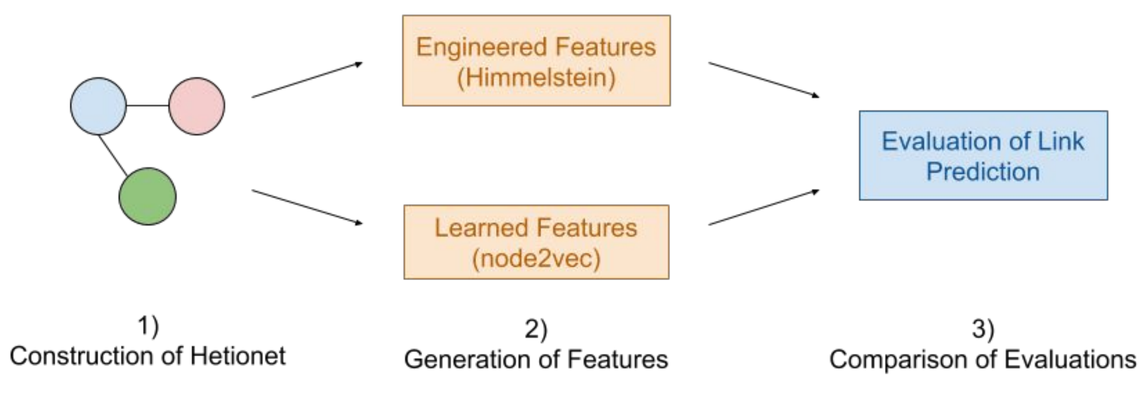
Previous computational approaches using biological knowledge graphs (KGs) have often focused on two subtasks: 1) predicting new drug-target interactions and 2) predicting new target-disease associations as an indirect means of identifying novel drug-disease associations.

BACKGROUND

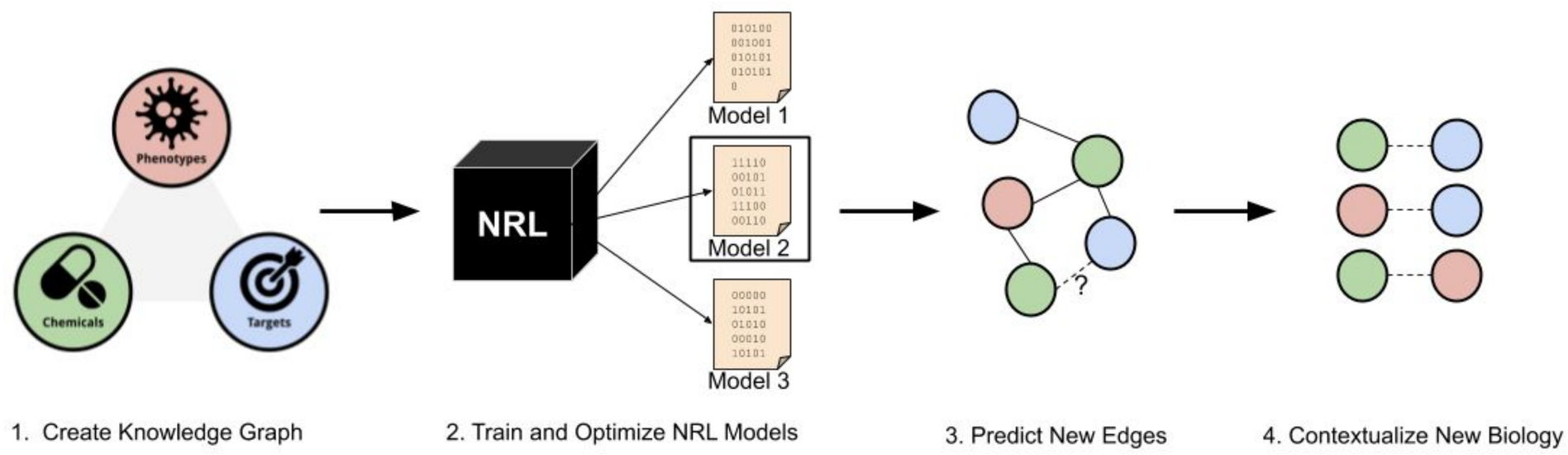
Network representation learning (NRL) methods have the goal, for a given KG, to generate low-dimensional vector representations of nodes whose elements correspond to latent features of the KG. Here, we propose and evaluate the edge prediction task in DR using three classes of NRL based on 1) random walks, 2) translational distance metrics, and 3) similarity metrics.

REPRESENTATION LEARNING IN NETWORK-BASED DRUG REPOSITIONING

A prominent example of using a biological KG in DR is from Himmelstein *et al.* [2], Hetionet, in which the authors engineered and subsequently selected topological features based on weighted path counts to train a logistic regression classifier for edge prediction. Because feature engineering is burdensome in time for generation, selection, and interpretation, we reproduced Himmelstein's workflow while interchanging this step with NRL methods (e.g., node2vec [3]).



With the best hyperparameters (p=1, q=2, dimensions=48, window=10, walk_length=30, walks=10), the average AUC-ROC of two validation datasets performed better than Himmelstein's (Clinical Trial: 0.714 vs 0.700; Symptomatic: 0.854 vs. 0.702) and one performed worse (DrugCentral: 0.816 vs. 0.855). Code for reproduction can be found at https://github.com/lingling93/comparison.



A MECHANISTIC UNDERSTANDING OF SIDE EFFECTS

Hetionet included drugs' side effects, but relatively little is known about their underlying mechanisms. We constructed a knowledge graph containing drug-target interactions from DrugBank, drugs' side effects from SIDER, and drugs' similarities using several fingerprints from RDKit. It was then embedded using several NRL methods (e.g., DeepWalk, SDNE, TransE), enabling calculation of side effect similarity, exploration of targets mediating side effects, and prediction of drugs' side effects. This is ongoing work that can be followed at https://github.com/AldisiRana/SE_KGE.

PERSPECTIVES

Our ongoing work is to incorporate literals in KGs, with models like LiteralE, in order to leverage prior (node-centric) knowledge such as chemical fingerprints into learning. This could provide an alternate approach to proteochemometrics that is more amenable to incorporation of further information.

REFERENCES

1. Kobayashi, Y., *et al.* (2019). Current state and outlook for drug repositioning anticipated in the field of ovarian cancer. *Journal of Gynecologic Oncology*, 30(1), e10.
2. Himmelstein, D. S., *et al.* (2017). Systematic integration of biomedical knowledge prioritizes drugs for repurposing. *eLife,* 6.
3. Grover, A., & Leskovec, J. (2016). Node2Vec: Scalable Feature Learning for Networks. *In Proceedings of SIGKDD* (pp. 855–864). New York, NY, USA: ACM.

This work was funded by the Fraunhofer Gesellschaft under the MAVO project: The Human Brain Pharmacome. More information can be found at https://pharmacome.scai.fraunhofer.de. Download this poster with the adjacent QR code ->

Fraunhofer SCAI Fraunhofer IAIS UNIVERSITÄT BONN