HEALTHCARE XPLORERS
Academic minds for innovation

Roche

— Predicting drug effects in a multimodal biological network

PHC

The fundamental understanding of the human organism and how it is perturbed by variants and drugs forms the basis of delivering tailored care. How can we predict the specific effect a drug has on its known cellular target(s) based on its cellular interaction network?

Apply until **April 18, 2021** / Xplorers Camp on **May 11, 2021**

Question to be solved

How can we predict the specific effect a drug has on its known cellular target(s) (e.g., activation, inhibition, antagonism) based on its cellular interaction network?

General Background

In personalized healthcare, the fundamental understanding of the human organism and how it is perturbed by variants and drugs forms the basis of delivering care tailored to the individual patient. To enable this, the complex interplay between the biological entities that drive disease at the molecular level can be modeled as a multimodal network which can be leveraged for machine learning applications, including the prediction of polypharmacology, drug combinations, or candidates for drug repurposing.

However, many machine learning applications on biomedical graphs struggle with lacking detail driven by 1) missing data in the data sources used to build the network and 2) an incomplete picture of the organism elucidated by different types of experiments. In particular, the lack of detail on drug effects on their molecular targets can cause issues with utilization of the graph for pharmaceutical applications. In this challenge, we ask for your help to develop computational methods to overcome the missing data problem of drug effects in multimodal networks to open up the full potential of biological networks for personalizing healthcare.

Data Types & Technologies

- Working with multimodal graphs of biological entities (genes, genetic variants, pathways, diseases, drugs, …)
- Experience developing machine learning models on graphs
- Python (or R)
- Tensorflow 2 or PyTorch

Supporting Material or Links

- Manica et al. Toward Explainable Anticancer Compound Sensitivity Prediction via Multimodal Attention-Based Convolutional Encoders. Molecular Pharmaceutics, 2019, 16 (12), 4797-4806.
- Oskooei et al. PaccMann: Prediction of anticancer compound sensitivity with multi-modal attention-based neural networks. NeurIPS 2018 Workshop on Machine Learning for Molecules & Materials, 2018 –arXiv:1811.06802
- Zitnik et al. Modeling polypharmacy side effects with graph convolutional networks. Bioinformatics, 2018
- You et al. Handling Missing Data with Graph Representation Learning, NeurIPS 2020.
- Pushpakom et al. Drug repurposing: progress, challenges and recommendations Nature Reviews Drug Discovery, 2018, 18, 41–58

Needed Skills

- Self-starter who likes tough machine learning problems and wants to make an impact in healthcare
- Experience with advanced missing data imputation methods
- Experience with state-of-the-art mechanistic graph learning methods as well as representation learning and graph embeddings.
- Experience with biological networks and knowledge graphs can be of advantage.
- Completed MSc in technical/quantitative field
- (At least basic) Understanding of pharmacology and biology

Mentor

Dr. Charlotta Fruechtenicht
Senior Data Scientist, PHC Analytics

Form of Cooperation

Preferred scale: 6 months full-time (likely remote) internship
Possible format: Full-time internship, with potential to develop into Master Thesis or part of PhD research project

How to present your Idea

We do not expect a bullet proof solution, but please present your idea on machine learning approaches to solve the task including ideas for suitable training data and validation strategy in a short and concise format (3-5 slides or max 1 page written document plus possible figures to illustrate the concept)