Division of Statistics and Machine Learning
Department of Computer and Information Science
Linköping University

Master thesis project ”Machine Learning for optimal drug prediction from single cell data”

Background

More than 40% of patients do not respond to standard treatments in common diseases. Accordingly, methods from personalized medicine aim to identify and propose the best treatments for a patient with given characteristics. Single cell RNA sequencing enables generation of high-resolution data for a given patient that can be used for individualized drug prediction.

Many advanced methods have been developed for prediction of efficiency of a given drug for a certain disease profile, such as classical PREDICT algorithm [1] or more modern techniques like [2] or [3] but they have not been applied to single cell data so far.
Your task will be to adopt some of these approaches to simplified settings and train your own drug prediction model.

Data

- Drug related data such as chemical properties, targeted proteins, gene expressions is freely available from the repositories cited in [1,2,3]
- Disease related data such as gene expressions are freely available from the repositories cited in [1,2,3]
- Single cell data for individuals having arthritis and a group of healthy individuals will be provided in order to test the drug prediction.

Research questions

- What machine learning model can be used to predict drug effect based on only chemical properties of the drug, targeted proteins, induced gene expressions and for the given gene expression profile of the disease?
- Which drugs are predicted if this machine learning model is used to predict from mean expression profiles per cell type in the given single cell data?
- How will these predictions differ between the cell types?
- Are these predictions reasonable from biological point of view?

Prerequisites

- Good knowledge of Machine learning and Statistics
- Good programming skills

Research Team

- Oleg Sysoev, STIMA, Linköping University
- Mikael Benson, Center for Personalized Medicine, Linköping University

Contact and application

- Oleg Sysoev, oleg.sysoev@liu.se

References

Gottlieb, A., Stein, G. Y., Ruppin, E., & Sharan, R. (2011). PREDICT: a method for inferring novel drug indications with application to personalized medicine. *Molecular systems biology*, *7*(1), 496.

Wu, G., Liu, J., & Yue, X. (2019). Prediction of drug-disease associations based on ensemble meta paths and singular value decomposition. *BMC bioinformatics*, *20*(3), 1-13.

Wang, Y., Yang, Y., Chen, S., & Wang, J. (2021). DeepDRK: a deep learning framework for drug repurposing through kernel-based multi-omics integration. *Briefings in Bioinformatics*.