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

Master thesis project "Machine Learning for drug ranking 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 RefDNN [2] but they have not been applied to single cell data so far.

Data

- Drug related data such as chemical properties are freely available from the repositories cited in [2]
- Disease related data such as gene expressions are freely available from the repositories cited in [2]
- Single cell data for individuals having arthritis.

Research questions

- What is the ranking of drugs predicted by RefDNN for the arthritis patients if predictions are done per patient (average expression per patient)? What can be said about top-5 recommended drugs?
- How can statistical uncertainty methods be used to estimate the uncertainty of drug ranking and how does this affect the comparison of top-5 recommended drugs?
- What is the ranking of drugs predicted by RefDNN for the arthritis patients if predictions are done per cell type (average expression per cell and patient)? How does this compare to previous rankings?

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

[1] 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.
[2] Choi, J., Park, S., & Ahn, J. (2020). RefDNN: a reference drug based neural network for more accurate prediction of anticancer drug resistance. *Scientific reports*, *10*(1), 1-11.