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

Master thesis project ”Machine learning methods for Seasonal Allregic Rhinitis”

Background

One of the key problems in medical care is that many patients do not respond to treatment. A new technique, single cell RNA-sequencing, generates high resolution data that allows genome wide analysis of mRNAs in individual cell types. These kinds of data were collected in the context of Seasonal Allergic Rhinitis study at several time points. More specifically, the blood samples of allergic and non-allergic patients were either exposed by pollen or not, and gene expressions of the individual cells were collected.

It is of interest to study the potential of commonly used machine learning tools in order to find disease associated genes and the time points when the disease can be detected.

Data

- Single cell data describing gene expressions (thousands of cells, thousands of genes) obtained for allergic and non-allergic patients that were either exposed or not exposed by the pollen.
-

Aims

- Provide a model to classify between cells coming from allergic and non-allergic persons by using the early timepoint data and late timepoint data and derive the genes associated with the disease at these time points.
- Investigate whether the classification model can be improved if the information about their cell types is included in the model
- Propose and compute a machine learning model that uses data from all time points in order to predict earliest time points when disease related genes are differentially expressed between the cells corresponding to allergic and non-allergic patients.
-

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] Li, H., Courtois, E. T., Sengupta, D., Tan, Y., Chen, K. H., Goh, J. J. L., ... & Wong, M. (2017). Reference component analysis of single-cell transcriptomes elucidates cellular heterogeneity in human colorectal tumors. *Nature genetics*, *49*(5), 708.