Contribution ID: **77** Type: **not specified**

Classification methodology for flow cytometry data in the context of a specific disease

Wednesday, 18 September 2024 14:40 (20 minutes)

Flow cytometry is a technique used to analyze individual cells or particles contained in a biological sample. The sample passes through a cytometer, where the cells are irradiated by a laser, causing them to scatter and emit fluorescent light. A number of detectors then collect and analyze the scattered and emitted light, producing a wealth of quantitative information about each cell (cell size, granularity, expression of particular proteins or other markers···). This technique produces high dimensional multiparametric observations.

We considered here flow cytometry data, obtained from blood samples, in the context of a specific severe disorder, heparin-induces thrombocytopenia (HIT). For each of the 141 patients, 8 variables was measured on around 10 000 cells. In order to reduce the size of the data, pre-processing has been made by calculating deciles and correlation between variables. We then get a database of 357 variables on 141 patients. We investigated several classification methods (logistic regression, random forests, SVM with different kernels) on the full data and on data reduced by PCA, with the aim of developing a classification methodology.

Type of presentation

Talk

Classification

Both methodology and application

Keywords

Cytometry, Classification

Primary author: GÉGOUT-PETIT, Anne (Université de Lorraine)

Co-authors: MAHAMAT, Cheikhadine Saleh; Dr MÉZIÈRES, Sophie (INRIA BIGS - IECL); Mr FLOCQUET, Thomas

Presenter: GÉGOUT-PETIT, Anne (Université de Lorraine)

Session Classification: PCA and mining

Track Classification: Biostatistics/ Statistics in the Pharmaceutical Industry