Contribution ID: **33** Type: **not specified**

Dimension Reduction Methods Based on FINE Algorithm for Clustering Patients from Flow Cytometry Data

Wednesday, 13 September 2023 11:45 (20 minutes)

Flow cytometry is used in medicine to diagnose complex disorders using a multiparametric measurement (up to 20 parameters). This measurement is performed in a few seconds on tens of thousands of cells from a blood sample. However, clustering and analysis of this data is still done manually, which can impede the quality of diagnostic discrimination between “disease” and “non-disease” patients. A computational algorithmic approach that automates and deepens the search for differences or similarities between cell subpopulations could increase the quality of diagnosis.

The approach considered in this study is information geometry, which involves lowering the dimensionality of multiparametric observations by considering the subspace of the parameters of the statistical model describing the observation. The points are probability density functions, and the subspace is equipped with a special geometrical structure called a manifold. The objective of the reported study is to explore an algorithm called Fisher Information Non-parametric Embedding (FINE), by applying it to flow cytometry data in the context of a specific severe disorder, heparin-induced thrombocytopenia (HIT).

This exploration consisted in testing different alternatives of the FINE algorithm steps such as the use of the Kullback Leibler divergence under a Gaussian assumption or the Wasserstein distance as measures of dissimilarity between the multiparametric probability distributions of the flow cytometry data for HIT+ vs HIT-.

Keywords

Information Geometry, Cytometry, Clustering

Classification

Both methodology and application

Primary authors: LAZIRI, Walid (INRIA); GÉGOUT-PETIT, Anne (Université de Lorraine); MÉZIÈRES, Sophie (INRIA BIGS - IECL); Mr ALLEMAND, Frédéric (EMOSIS)

Presenter: LAZIRI, Walid (INRIA)

Session Classification: CONTRIBUTED Biostatistics 2

Track Classification: Biostatistics