KEYNOTE TALK

Monday, October 14, 2019 at 1:30pm
(Emerald Bay 123)

High dimensional unsupervised approaches for dealing with heterogeneity of cell populations and proliferation of algorithmic tools

Yuval Kluger
Yale University School of Medicine
USA

Abstract: Revealing the clonal composition of a single tumor is essential for identifying cell subpopulations with metastatic potential in primary tumors or with resistance to therapies in metastatic tumors. Bulk sequencing technologies provide only an overview of the aggregate of numerous cells. We propose an evolutionary framework for deconvolving data from a bulk genome-wide experiment to infer the composition, abundance and evolutionary paths of the underlying cell subpopulations of a tumor. With advances in high throughput single cell techniques, we can in principle resolve these issues. However, these techniques introduce new challenges such as analyzing datasets of millions of cells, batch effects, missing values etc. We provide several algorithmic solutions for some of these challenges. Finally, a key challenge in bioinformatics is how to rank and combine the possibly conflicting predictions of several algorithms, of unknown reliability. We provide new mathematical insights of striking conceptual simplicity that explain mutual relationships between independent classifiers/algorithms. These insights enable the design of efficient, robust and reliable methods to rank the classifiers performances and construct improved predictions in the absence of ground truth.

Speaker Bio-Sketch: Yuval Kluger has been working in the broad fields of bioinformatics, machine learning, and dynamics of quantum fields. His main contributions to date relate to development of spectral methods for unsupervised learning, cell specific regulatory networks, algorithms for analyzing genomics and epigenomics sequencing data, algorithms for detecting and characterizing biomarkers in high dimensional assays, and non-equilibrium quantum field theory models.