

Friday, January 22 from 11A.M. – 12 P.M. SPHIS Room 103

Martin Schindler

School of Public Health, University of Tampere, 33014 Tampere, Finland and
Department of Applied Mathematics, Technical University of Liberec, Liberec, Czech
Republic

presents

“Rank methods for microarray data”

Abstract

Our study was motivated by lymphoblastoid microRNA (miRNA) microarray data from more than 100 prostate cancer and healthy brothers. Due to nature of the data we use rank methods. After mentioning preprocessing, quality control and filtering procedures we first describe miRNA-wise analysis based on regression rank score test taking into account family effects. In the second part we make use of Ward's clustering method and rank correlation statistic to develop preprocessing-free, normalization-free, stable and robust method to find the most influential miRNAs.

Daniel Fischer

School of Public Health, University of Tampere, 33014 Tampere, Finland.

presents

“Applying decision trees and random forest to microarray data”

Abstract

Based on a microarray data set of prostate cancer patients and healthy controls we apply machine learning algorithms to classify cancer and healthy patients. We focus on decision trees, based on the C4.5 algorithm, to build up random forests which will enable us to make predictions on the health status of the patients. The decision rules are based on ranks of observed microRNAs. The random forest will also be used to identify a subset of important microRNAs