

Meeting 29/01/2020

Stephen Coleman

29/01/2020

Summary

- Stephen to look at the Radford Neal paper “Markov Chain Sampling Methods for Dirichlet Process Mixture Models” and annotate the MATLAB code;
- Look at the BCC TCGA data pre- and post-processing;
- Meet with Eckart and discuss possible integrative clustering methods (probably COCA, see Alessandra’s paper, “Multiple kernel learning for integrative consensus clustering of genomic datasets”); and
- Apply the TCGA labels to the BCC data to compare clusterings.

Eckart

A MSc student with Mark Toshner and Chris. Looking for heterogeneity in a patient population with each patient described within multiple datasets. Toshner expects the sub-populations to be **global** rather than local. This suggests use of an integrative clustering method that produces a global clustering. As Eckart is not a computational person by background, ideally an intuitive method with a good workflow that is quick to run. For this reason consider iClustrBayes or Cluster-of-Cluster-Analysis (COCA).