

TITLE: Bayesian Nonparametrics Methods for Complex Biomedical Data

SPEAKER: Francesco Denti

ABSTRACT: In this talk, we introduce the intuition behind fundamental Bayesian nonparametric tools, such as Gaussian Processes and Dirichlet Processes, which serve as the building blocks for models that are both extremely flexible and fully interpretable. Motivated by the challenges posed by large, complex biological datasets generated by state-of-the-art imaging technologies, we present modeling extensions designed to handle spatial structures, temporal dependencies, and nested frameworks where data are organized into hierarchical groups.

We first demonstrate these methodologies through an application to MALDI-mass spectrometry imaging, which quantifies molecular abundance across multiple locations within biological tissue. We propose a model for nested, separate exchangeable data that induces a biclustering structure to simultaneously group spatial locations and molecules with similar expression patterns. By incorporating a Hidden Markov Random Field prior, the model achieves precise image segmentation into biologically meaningful regions, facilitating the detection of specific molecular activation patterns. The second part of the talk addresses calcium imaging data, which records the activity of individual neurons over time in freely moving animals. We develop a spatiotemporal mixture model to identify co-activating neurons by detecting groups of cells with similar firing patterns. This is achieved through a multivariate time-series framework that identifies spikes in calcium traces and captures recurring temporal dynamics. We employ a Dependent Dirichlet Process to incorporate anatomical proximity between neurons, while simultaneously segmenting spike amplitudes via a Dirichlet Process to provide clusters based on magnitude. Together, these frameworks offer robust, interpretable solutions for the segmentation and functional analysis of complex biomedical systems.