

Poster L-6

KERNEL-IMBEDDED GAUSSIAN PROCESSES FOR MICROARRAY GENE EXPRESSION DATA ANALYSIS



Authors:

LEO WANG-KIT CHEUNG (*Division of Biostatistics and Bioinformatics, Department of Preventive Medicine and Epidemiology, Lo*)

XIN ZHAO (*Department of Information and Computer Sciences, University of Hawaii*)

Short Abstract: We propose kernel-imbedded Gaussian processes for exploration of nonlinear relationships between expression data and the target features of a disease classification problem. This work has been applied to simulated data as well as real microarray data, and has shown great promises especially when non-linearity is dominant in the underlying relationships.

Long Abstract:

This poster introduces a class of kernel-imbedded Gaussian processes (KIGP) for microarray gene expression data analysis. Specifically, we propose a supervised Bayesian gene selection method via the regression setting based on imbedding a kernel function into a Gaussian process. With the introduction of latent variables in a hierarchical Bayesian modeling framework, we implemented a Gibbs sampler to discover potentially significant genes from microarray gene expression data for a given disease classification problem. The methodology developed with the involvement of a kernel space allows the proposed Bayesian framework for exploration of nonlinear relationships between gene expression data and the target features of the given classification problem. Furthermore, the proposed method precludes from inverting a possibly singular matrix resulted from the limited number of samples and/or from a potentially large number of important genes, hence, it is advantageous as inverting a singular matrix gives rise to algorithmic problems in most existing linear model based methods. The proposed work has been applied to simulated data as well as real microarray data, and has shown great promises especially when non-linearity is dominant in the underlying relationships.