

# Data Mining for Biomedical Informatics

*Saturday, April 26, 2008*

## Workshop Schedule

- 8:45 - 9:00 Introduction
- 9:00 - 9:30 Biomedical named Entity Recognition using Support Vector Machines:  
Performance vs. Scalability issues.  
*Mona Soliman Habib*
- 9:30 - 10:00 Establishing Androgen Receptor-Regulated Transcription Networks.  
*Li Li, Steffen Heber, Qiang Zhang, Melvin E. Andersen*
- 10:00 - 10:30 *Break*
- (Keynote)** Taming Patterns in Biological Data.  
10:30 - 11:30 *Laxmi Parida*  
IBM TJ Watson Research Center  
Courant Institute of Mathematical Sciences, New York University.  
<http://www.cs.nyu.edu/parida/>
- Abstract:**  
Patterns abound in large data sets. The more flexible a pattern definition, the larger the instances and higher the confusion; while rigid definitions are very restrictive. Must we throw away the proverbial baby with the bath water? In this talk we look at the combinatorics and statistics of patterns that computational biologists discover at different levels in biological data be it nucleic acid sequence, microarray data or other formal structures. We exploit the commonality that runs across these various domains and apply the lessons learned in one to another. I make the case of applying nontrivial combinatorics to an interesting class of patterns called permutation patterns. This mathematical structure is applied to some problems arising naturally in the area of computational biology such as the problem of common gene clusters across species, phylogeny within populations, and the task of modeling complex control of transcriptions via motifs.  
I will end the talk with patterns on Linkage Disequilibrium (LD) and what it entails in the context of the Genographic Project.
- 11:30 - 12:00 Simulated association studies with PCA-informative markers.  
*Jamey Lewis*
- 12:00 - 1:30 *Lunch*

- 1:30 - 2:00      A Comparative study of Breast Cancer Microarray Gene expression profiles using Label Propagation  
*TaeHyun Hwang and Rui Kuang*
- 2:00 - 2:30      Mining Fly Embryo images using Graph Based Methods.  
*Fan Guo, Lei Li, Eric P. Xing, Christos Faloutsos*
- 2:30 - 3:00      An improved algorithm for Tag SNP selection based on Pair-Wise Linkage Disequilibrium.  
*Wei Wang, Youling Guo, Yuexian Zou, Tianrui Wu*
- 3:00 - 3:30      *Break*
- 3:30 - 4:00      Detection Ischemic episodes from Electrocardiogram Signal using Wavelet Transform.  
*Mohammad Karimi Moridani*