CCSU
DEPARTMENT OF MATHEMATICAL SCIENCES

COLLOQUIUM

Friday, November 21, 2008
2:00 – 3:00 PM
Maria Sanford, Room 101

ENSEMBLE CLASSIFIERS
AND BIOMARKER DISCOVERY

DARIUS M. DZIUDA

CENTRAL CONNECTICUT STATE UNIVERSITY

Abstract: One of recent trends in supervised learning is the ensemble classifier approach. Classification systems based on multiple individual classifiers often outperform single classifiers, especially when the latter are weak or unstable. An unknown sample is classified by all individual classifiers and assigned to one of the classes in a result of weighted or unweighted voting. However, the ensemble approach generally does not deliver parsimonious biomarkers. On the contrary, ensemble classifiers base their voting on a usually large number of variables represented in all individual classifiers.

One of the main goals of biomarker discovery is to identify a small set of genes (or proteins) whose joint expression profile can significantly separate the differentiated classes and can be used for efficient classification of new cases. **Are then ensemble approaches useless for biomarker discovery?**

To answer this question, this talk will start with introduction to biomarker discovery and introduction to ensemble classifiers, and then will focus on proper ways of using the ensemble approach in biomarker discovery. Discussed topics will include utilization of the ensemble paradigm in identification of robust and parsimonious biomarkers and in validation of classification systems.

Keywords: *Ensemble Classifiers, Biomarker Discovery, Multivariate Biomarkers, Data Mining, Bioinformatics, Genomics, Proteomics, Differential Diagnosis, Drug Discovery, Individualized Medicine, Classification and Prediction Systems.*

For further information:
gotchevi@ccsu.edu 860-832-2839