# CCSU
## DEPARTMENT OF MATHEMATICAL SCIENCES

DATA MINING
THESIS PRESENTATION

**Friday, December 3**
**2:00 – 3:00 pm in MS 101**

# Knowledge Discovery in Microarray Data


Rafiqul Islam
(MS candidate in Data Mining)



**Abstract**

We present a method of prioritizing potential drug targets based on their gene expression "signature". Primary human pre-cursor neuronal cells were treated with three classes (antidepressant, antipsychotic, opiod receptor agonist) of psychoactive drugs for 24 hours. Microarray technology was used to capture expression of ~11, 000 genes induced by these three categories of drugs. It was demonstrated that a Neural Network model and a Decision Tree (C5.0) model could classify these drugs based on their gene expression with 80% and 92 % accuracy, respectively. TwoStep clustering algorithm was used to separate gene expression profiles into natural groups. When cluster information was used as input for the Neural Network model, classification accuracy increased to 88%. It was also demonstrated that the confidence index generated by each classification model could successfully be used to prioritize a portfolio of novel drug targets.


**Refreshments will be served at 3:00 pm.**

***For further information:***


gotchevi@ccsu.edu (860) 832-2839
castanedan@ccsu.edu (860) 832-2851