

**California State University San Bernardino
School of Computer Science & Engineering**

MASTERS PROJECT PROPOSAL PRESENTATION

Date/Time

March 11, 2015 (Wednesday), 12:00 PM

Location

JB 389/391

Topic

APPLY DATA ANALYSIS TECHNIQUE TO GENE EXPRESSION DATA

Candidate

Abdullah Abualhamayl

Advisor

Dr. Haiyan Qiao

Committee Members

Dr. George M. Georgiou

Dr. Owen J. Murphy

Abstract

Data analysis plays an important role in effective analysis of gene expression. Although DNA microarray technology facilitates expression monitoring, difficulty lies in understanding the genetic link due to the complexity in our biological system. Clustering analysis offers a viable solution to this problem and it encompasses data mining, which unravels the natural structures and recognizes distinct and novel patterns in the given dataset. It can be considered as one crucial stage in data analysis. On the basis of certain specific features, gene dataset is grouped into different classes, and different clustering algorithms have been developed for this purpose. The genetic groups which are biologically interlinked can be identified through cluster algorithms. This project proposes clustering analysis of genes involved in published data sets. Different clustering algorithms will be tested on the acquired dataset to investigate the most accurate and efficient analysis.