



SDSU
presents
a thesis defense for
Master of Science
degree in
Computer Science

Monday,
December 10, 2012

11:00am
GMCS 405

Rama Bandi

*Application of Random Forest Algorithm in
Biomarker Discovery for Cancer Detection*

Abstract

Recent advances in molecular biology and computational science have greatly influenced and promoted the field of bioinformatics. One such area is recently available high throughput platform for biomarker discovery based on Printed Glycan Arrays and new bioinformatics algorithms for feature selection and classification. This thesis focuses on implementation, application and evaluation of popular Random Forest algorithm used for biomarker discovery based on data generated with PGA. The implementation is tested on real clinical data obtained from the School of Medicine of NYU, which contain a control sample with 65 high risk subjects exposed to asbestos and a case sample with 50 subjects diagnosed with malignant mesothelioma. The results are compared with other popular approaches such as univariate feature selection based on Wilcoxon ranking, and multivariate feature selection based on Causal Networks and Markov Blanket algorithms.

Thesis Committee

Marko Vuskovic, Thesis Chair, Department of Computer Science
Joseph Lewis, Department of Computer Science
Jianwei Chen, Department of Mathematics & Statistics