The Department of Computer Science and Engineering

University of Nevada, Reno

cordially invites you to a PhD defense

Using Pre- and Post-Process Labeling Techniques for Cluster Analysis

A colloquia submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy
in Computer Science and Engineering.

by

Damien Ennis

Abstract:

As the amount and variety of data increases through technological and investigative advances, the means to analyze and manage this data becomes more critical. Unsupervised machine learning algorithms can be used to group large datasets into categories, thereby facilitating new insight into similarities in data that would, on the surface, appear to be disparate. This dissertation investigates three unsupervised clustering algorithms: the self-organizing map, the K-means algorithm, and affinity propagation. These three algorithms are applied to large datasets from three different domainsorganizational management, bioinformatics, and financial marketseach of which presents its own challenges in terms of data management and knowledge discovery. Specifically, the self-organizing map is used to cluster a variety of academic library data to show how it can be used to aid in operational and strategic decision-making. Both a self-organizing map and the K-means algorithm are used to cluster genomic data to show how they can be used to identify possible organisms that are present in a metagenomic sample. Affinity propagation is used to cluster stock performance data to show how it can be used to aid in making investment decisions. In addition, different semi-supervised labeling techniques are employed in combination with these clustering algorithms to assist with knowledge discovery in these three areas. The applicability of these different labeling techniques for various types of problems is discussed, and the success of these combinations in facilitating several types of data analysis is explored, providing researchers with guidance about the applicability of these strategies.

1:00 pm, Monday, November 17, 2014

Scrugham Engineering and Mines (SEM) room 201

For more information contact Dr. Fred Harris @ 784-6571 (Fred.Harris@cse.unr.edu)