CCSU
DEPARTMENT OF MATHEMATICAL SCIENCES

COLLOQUIUM

Friday, April 29
9:30 – 10:00 AM
Davidson Hall, Room 207

TRINARY PREDICTIVE CLASSIFICATION OF DIABETIC EPISODE RECURRENCE

CATHY L. FARRELL

(Data Mining MS Thesis Presentation)

CENTRAL CONNECTICUT STATE UNIVERSITY

Abstract: This thesis explores the value of trinary predictive modeling to the understanding of diabetes-related hospitalizations. According the U.S. Centers for Disease Control and Prevention, diabetes is the 7th leading cause of death in the United States, killing more than 75,000 people each year. In 2010, there were more than 600,000 hospital inpatient discharges in the U.S. with a primary diagnosis of diabetes. The average stay was 4.6 days. Between 2009 and 2012, diabetes was estimated to affect more than 12% of the U.S. population, including both diagnosed and undiagnosed cases (U.S. Centers for Disease Control and Prevention, 2015). If effective models were available to predict which patients were likely to be readmitted to the hospital, it is hoped that interventions might be possible to improve the health and well-being of such patients, and to reduce the overall costs of their care. The thesis attempts to address this need with an analysis of the dataset "Diabetes 130-US hospitals for years 1999-2008 Data Set" (Clore, Cios, DeShazo, & Strack, 2014), which was obtained from UCI's Machine Learning Repository. (Center for Machine Learning and Intelligent Systems at the University of California Irvine, 2007). Each record in the dataset represents a single hospitalization for a specific anonymized patient. The thesis explores whether an effective trinary classification model can be built to predict the likelihood of patient readmissions. The intended model would classify patient hospitalization records into three categories: those likely to be readmitted within thirty days of discharge, those likely to be readmitted more than 30 days after discharge, and those likely to not be readmitted. The results of this analysis are inconclusive on that question. The most accurate model built was only 57.6% accurate.

For further information:
gotchevi@ccsu.edu 860-832-2839
http://www.math.ccsu.edu/gotchev/colloquium/