CS 6440: DATA MINING COURSE INFORMATION

Semester Hours: 3.0  
Contact Hours: 3

Coordinator: N/A

Text: Data Mining for Business Analytics, 3rd edition

Authors: Schmueli, Bruce and Patel

Year: 2017

SPECIFIC COURSE INFORMATION

Catalog Description:
Data mining is the analysis of large data sets for the purpose of discovering useful information. This course will cover a variety of data mining applications and algorithms. Students will be exposed to applications in business (finance, insurance, manufacturing, marketing), criminology (identifying criminal patterns, fraud detection), and science (analysis of scientific data).

Prerequisite: STAT 6010 Statistics for Managerial Decisions or consent of instructor. Students taking the course should have knowledge of elementary statistics & regression analysis. An advanced regression course STAT/MSA 5020 is preferred.

Course type: Elective

SPECIFIC COURSE GOALS

- I can demonstrate a basic knowledge of data mining statistical techniques.
- I can recognize when and how these techniques are applicable I can provide technical arguments for the integrity of a certain piece of evidence.
- I can identify data requirements for successful implementation of these methods.
- I can use R, Minitab, SPSS, and/or other software to conduct a data mining investigation on a real-life case study.
- I can analyze relevant research and communicate my findings

LIST OF TOPICS COVERED

1. Overview
2. Data mining process

3. Logistic regression

4. Performance evaluation

5. k-Nearest Neighbor Classification

6. Classification & Regression Trees

7. Naïve Bayes Classifier

8. Discriminant Analysis

9. Association or Affinity Analysis

10. Cluster Analysis