CS 6440: DATA MINING COURSE INFORMATION

Semester Hours: 3.0  
Contact Hours: 3

Coordinator: TBD

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. Topics include regression trees, clustering, neural networks, link analysis, and market based analysis. Students will be exposed to applications in business (finance, insurance, manufacturing, marketing), crime detection (identifying criminal patterns, fraud detection), and science (analysis of scientific data). Students taking the course should have a basic knowledge of elementary statistics. Prerequisite: STAT 2120 or equivalent. [Cross-listed with STAT 64440.]

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.

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