Course Syllabus Part I
DSC 630 Predictive Analytics

3 Credit Hours

Course Description
This course assembles topics covered in previous courses into an applied project. Students have the opportunity to find, clean, analyze, and report on a project they define. Advanced methods of analysis using Python and R allow students to delve deeper into their projects.

Course Prerequisites
DSC 540 or equivalent and recommend DSC 550

Course Objectives

Students who successfully complete this course should be able to:

- Solve real-world problems by adapting and applying statistical learning methods to large, complex datasets.
- Analyze datasets with supervised learning methods for functional approximation, classification, and forecasting.
- Analyze datasets with unsupervised learning methods for dimensionality reduction and clustering.
- Compose a comprehensive project report.
- Assess various statistical learning methods and models for solving a specified real-world problem, comparing their advantages and disadvantages.

Grading Scale

93 – 100% = A	87 – 89% = B+	77 – 79% = C+	67 – 69% = D+
90 – 92% = A-	83 – 86% = B	73 – 76% = C	63 – 66% = D
	80 – 82% = B-	70 – 72% = C-	60 – 62% = D-
			0 – 59% = F

Topic Outline

I. Overview of Predictive Analytics
   a. History
   b. Predictive Analytics Application in Business
   c. Challenges and Opportunities
   d. Future

II. Data Understanding
   a. Data Visualization
   b. Histograms
   c. Single and Multiple Variable Summaries
   d. Data Audit

III. Data Preparation
   a. Variable Cleaning
   b. Feature Creation

IV. Descriptive Modeling
   a. Data Preparation Issue
   b. Principle Component Analysis
   c. Clustering Algorithms

V. Interpreting Descriptive Models
   a. Standard Cluster Model

VI. Predictive Modeling
   a. Decision Trees
   b. Logistic Regression
   c. Naïve Bays
   d. Regression Models
   e. Linear Regression

VII. Assessing Predictive Models
   a. Batch Approach
   b. Assessing Regression Models

VIII. Model Ensembles
   a. Bagging
   b. Boosting
   c. Occam's Razor
   d. Interpreting Model Ensembles

IX. Text Mining
   a. Structured vs Unstructured Data
   b. Text Mining Features
   c. Modeling
   d. Regular Expressions
   e. Challenges and Opportunities

X. Course Project
   a. Project proposal
   b. Preliminary analysis
   c. Project Report
   d. Presentation