

**School of Business**

**OPIM 5604 – Predictive Modeling**

**Fall 2019**

Instructors: Trevor Tomko

E-mail: Trevor.Tomko@uconn.edu

Course Schedule (subject to possible change): Wednesday 6:00 PM to 9:00 PM

Location: Hartford GBLC Room 502

Office Hours: By appointment only

Course Objectives:

Technology advancements now allow companies to capture and store large amount of data (or facts) in databases and data warehouses. With so much raw data, organizations urgently need tools that allow them to effectively sift through these enormous datasets and extract actionable information and knowledge (meaningful patterns, trends, and anomalies) from such data sets to help them optimize businesses. Predictive modeling is the process of developing models to better predict future outcomes for an event of interest by exploring its relationships with explanatory variables from historical data. It is used extensively in businesses to identify risks and opportunities associated with a set of conditions.

The course introduces the techniques of predictive modeling and analytics in a data‐rich business environment. It covers the process of formulating business objectives, data selection, preparation, and partition to successfully design, build, evaluate and implement predictive models for a variety of practical business applications. The primary approach will entail learning key concepts in analytics and then reinforcing using ‘learning-by-doing’ with the use of the state-of-the-art software SAS JMP. We will also review and understand a variety of statistical tools and processes that form key elements of these techniques. In studying each technique, the goal is for each student to understand the technique and how it can be used to unlock hidden business value. At the completion of the course, each student should be positioned to understand and apply the variety of tools studied. In a given business problem setting, the student should be able to identify, understand, and lead the utilization of the best technique to analyze the data and yield business value.

Course Materials:

I will provide the bulk of the class materials that we will use in the class. You should download these materials prior to the start of class. The following repository will be used for providing the class materials and for students to submit the assignments, projects and exams.

1. **HuskyCT** which is the online course management system used at UConn. You can access it at <http://lms.uconn.edu>. You will need to have your NetID and password to login.

We will use the HuskyCT website extensively for course-related communication between students and the instructor. Once you have logged onto the HuskyCT class website, you will see specific links for different aspects of this course. All class materials (including lecture notes, assignments, data sets, and project details) will be posted on this website. Please check the site frequently for updates, additional material, and assignments. Please e-mail me if you have questions or problems with the material.

Required Textbook:

Data Mining for Business Intelligence: Concepts, Techniques, and Applications with JMP Pro

Shmueli, Bruce, Stephens, and Patel

ISBN 978-1-118-87743-2

Grading:

The final grade will be determined on the following basis:

Assignments 18% *Individual*

Exams 60% *Individual*

Project & Presentation 20% *Group*

Group Peer Review 2% *Individual*

Assignments (Individual):

There will be a number of assignments during the course of the semester. You will submit the graded assignments through HuskyCT. The total score from all the assignments will be normalized to 18% of the total course grade. **Please submit all assignments as a Word document – assignments will typically involve some verbal responses and demonstrations in JMP. Please use screen capture to take screen snapshots and then cut and paste into the word document that you submit**.

The following statement must be present on each and every submission: ***“The work contained and presented here is my work and my work alone.”***

Assignments that are submitted up to 24 hours late will receive a 25% reduction in the grade that it would otherwise earn. Assignments submitted 24 hours beyond the due date will receive a zero.

Predictive Modeling / Knowledge Discovery Project & Presentation (Group):

The project is expected to discover/predict interesting relationships and patterns using a significant amount of real data.

You can explore/analyze/mine interesting data sets and present and describe your findings. In other words, describe: (a) the data and how it was collected, (b) what you might be looking for in such data (e.g., potentially interesting/important business questions that could be explored/answered using this data), (c) the tools/techniques that you used, (d) the results that you obtained, (e) and the conclusions that you can draw (e.g., the actions that one can take in response to the findings).

Each group has to obtain the instructor’s permission regarding the appropriateness of the topic chosen.

Please see the Project Requirements document for more details.

Academic Integrity

Each student is individually responsible for the integrity of his/her own work. Unless explicitly stated otherwise, each homework assignment is an individual assignment and does not permit cooperation between students. For all assignments, you must properly cite sources of information as well as the ideas and words of others. Not giving proper credit to others' work constitutes plagiarism and is a serious violation of the University's honor system. Similarly, turning in the same work for credit in two different courses constitutes fraud and is also a serious violation of the University's honor system. Cheating of any sort will not be tolerated and will result in a failure of the exam or assignment, deduction in the class participation grading component, and potential failure of the course. A student who knowingly assists another student in committing an act of academic misconduct shall be equally accountable for the violation, and shall be subject to the sanctions and other remedies described in The Student Code. For details refer to https://community.uconn.edu/the-student-code-preamble/.

**Please note: vacations, previously purchased tickets or reservations, weddings, and other large or small scale social events, are not viable excuses for missing an exam or project presentation. Please contact the Dean of Students office with any questions. Thank you in advance for your cooperation.**

**Additional Notes**

1. Peer evaluations will play a role in determining the scores awarded to individual group members on the group-oriented components of the course. Individuals who assume leadership roles and/or who perform more than their fair share of the work and/or whose work is consistently of superior quality are likely to receive scores that are above the group’s average scores. Individuals who do less than their fair share of the work or work that is consistently below average will receive scores that are less than the group average.

Tentative Schedule – subject to change as necessary:

|  |  |  |  |
| --- | --- | --- | --- |
| **Date** | **Topic** | **Assignment Due** | **Readings** |
| Aug 28 | Introduction to Predictive Modeling and Data Mining  Software Setup (JMP)  JMP® Software: Data Retrieval & Manipulation |  | Chapters 1 & 2 |
| Sept 4 | Data Exploration & Visualization |  | Chapter 3 |
| Sept 11 | Data Quality, Preparation, Data Cleansing & Pre-processing | Homework 1 | Chapter 4 |
| Sept 18 | Clustering and Association Rules | Homework 2 | Chapter 14 |
| Sept 25 | Fundamentals of Predictive Modeling – Process and Metrics | Homework 3 | Chapter 5 |
| Oct 2 | **Exam #1** |  |  |
| Oct 9 | Regression (Linear and Logistic) | Group Project Proposal & Homework 4 | Chapters 6 & 10 |
| Oct 16 | Classification and Regression (Decision) Trees | Homework 5 | Chapter 9 |
| Oct 23 | Neural Networks | Homework 6 | Chapter 11 |
| Oct 30 | No Class – Work on Group Projects |  |  |
| Nov 6 | Discriminant Analysis, KNN, & Ensemble Models | Homework 7 | Chapters 12, 7, & 13 |
| Nov 13 | Naïve Bayes Classifier | Homework 8 | Chapter 8 |
| Nov 20 | **Project Presentations** | Project White Paper & Slides |  |
| Nov 27 | No Class – Thanksgiving Break |  |  |
| Dec 4 | **Exam #2** | Group Evaluation |  |