Department of Decision and Information Science

MIS 7397 / 4397 – Predictive Analytics & Business Intelligence

Fall 2019 – Wed 6-9, Melcher 138

Professor:
Dr. Xiao Ma ("Xiao" is pronounced as /sh-ih-ow/)   Email: xma@bauer.uh.edu
Office: Melcher #280H   Office phone: 713-743-4725
Office hours: Wed 1:15-3:00pm or By Appointment
TA (TA unavailable; backup desktops with SAS available in Melcher Hall student computer lab & in #131):
TA Office:
TBA   Email:
Office hours:

****

The information contained in this course syllabus is subject to change anytime. Students are expected to be aware of any additional course policies presented by the Professor during the course.

****

Prerequisite:
Completion of BZAN 6310 or BZAN 6320 (or equivalent statistics course) with a grade of "B" or better.

Required Reading:
Daniel T. Larose and Chantal D. Larose, *Discovering Knowledge in Data: An Introduction to Data Mining*, 2nd Edition, Wiley, 2014. ISBN: 978-0-470-90874-7 (hardback).

Recommended Reading:
Foster Provost and Tom Fawcett, *Data Science for Business: What You Need to Know About Data Mining and Data-Analytic Thinking*, O'Reilly, 2013. ISBN: 978-1-449-36132-7.

Recommended Computing Resources:
Use your own machine (laptop, notepad, surface, remote desktop) as your best computing resource. Backup computers with SAS may be available, as noted above; use them at your own risk (project files can be lost).

Course Description
Predictive Analytics is the umbrella term to include all the efforts toward utilizing analysis of data to support business decision making. Based on this concept, this course (aka Data Mining) mainly focuses on machine learning models, tools, and software techniques.

Value of the Course
For a variety of reasons—including the low cost of storage, technologies to capture data and ease of use of analytical tools—considerable effort is now being expended in most corporations to gain competitive advantage via business intelligence. Thus, some grasp of the fundamental business intelligence models coupled with hands-on experiences of state-of-the-art BI software tools will be of great benefit to almost any imminently graduating business student.

Goals of Course

The primary course goals are for the student to be able to do the following when they have completed the course:

- Be able to explain what the terms Predictive Analytics and Business Intelligence encompass and how they fit into the business—also its potential and limitations
- Understand the principles underlying data mining tools
- Be able to adequately determine correct data mining tools to apply to data
- Be able to run data mining (DM) projects; analyze and interpret the results—and make management recommendations based on their findings

Teaching Methods

1. Lectures and Discussions with some hands-on time: Important material from the text and outside sources will be covered in class. This coverage will be mostly in the form of hands-on problem solving. Students should plan to take careful notes on topics that are presented by the Professor.

   Screencast of inclass hands-on demonstrations: The Professor will demonstrate how to use SAS and related data mining technologies in class and record the computer screen live. Realtime audio narration as well as any accompanying Q&A will also be recorded. The Professor will upload the screencasts to YouTube following all classes of recording, and create shortcuts to the videos on Blackboard.

   Complementary to the new screencasts to be produced and made available throughout the semester, technical recordings from previous semester(s) hosted on YouTube are also available via Blackboard.

2. Assignments: Problems and readings are routinely assigned to help support and supplement material found in the text. Assignments must be turned in on time to receive the maximum points possible.

3. Exams: Exams / quizzes will be closed book/note and will test assigned readings and materials. Review sheets will be provided in advance to the exam day. The final exam will not be cumulative, but may require a good understanding of some fundamental concepts covered in previous exams. All relevant concepts and points will be noted on the review sheets.

4. Blackboard: Most materials will be distributed on the Internet, using the class website *Blackboard* application. It is assumed that students know how to access the content on Blackboard.

5. Announcements regarding the class such as schedule changes, assignments, projects, and so on will be made in class during the first 10 minutes as well as on Blackboard. You are responsible for being in class **ON TIME** to hear the announcements and for monitoring the class website.

6. Contacting the Professor and/or Teaching Assistant (TA): Email will be the most productive channel of communication to resolve the vast majority of inquiries. If a TA is available to this course (dictated on the cover page), when emailing your question and replying any follow-up message, you must include both your Professor and your TA as the recipients, using the "Reply All" function of the email service, in order to make us maximally informed about the recent development of the resolution process. If you only included the Professor or the TA as the only recipient of your email, it is not guaranteed your inquiry will be answered, and you will be responsible for any negative consequence (e.g. late submission penalty for being stuck on a problem needing the Professor's and/or the TA's help).

   The best email address to reach the Professor and/or the TA are included on the cover page of syllabus.

7. Grading:

    1. Quiz #1 – 20%
    2. Quiz #2 – 15%
    3. Homework: 20%
    4. Attendance: 12%
    5. Active Classroom Discussions: 3%
    6. Project – 30%

    Total: 100%

Homework	1	2	3	4	5	6	7	8	9
Ttl Pts (out of 100)	1.5	3.0	3.0	3.0	2.5	2.5	1	1.5	2.0

Final course letter grade follows the numeric-letter grade system shown in the table below.

Raw Score	Letter Grade
> or = 92	A
> or = 89, but < 92	A-
> or = 86, but < 89	B+
> or = 83, but < 86	B
> or = 80, but < 83	B-
> or = 77, but < 80	C+
> or = 74, but < 77	C
> or = 70, but < 74	C-
> or = 67, but < 70	D+
> or = 63, but < 67	D
> or = 60, but < 63	D-
< 60	F

8. Project Description: The class project will give you an opportunity to apply most of the techniques you learned in this course. I will make available to you a set of data on which your project is based. The data will be what we consider "big." You will be required to get an understanding of the variables in your data set. You will use the software applications for the course to perform various analytics techniques on your data. In your project, you will attempt to show how well you can apply the techniques, understand the problems and solve them using appropriate techniques. You will need to interpret all your results. And, provide managerial recommendations to help a stakeholder make good business decisions, which you will also define.

Course Policies:

Missed Classes: The student is responsible for obtaining material, which may have been distributed in class when he/she was absent. This can be done through contacting a classmate who was present or by contacting the Professor during his office hours or other times. Missed or late exams cannot be made up under any

circumstances, unless an official excuse is provided. **Any uncoordinated, unexcused missed exam will result in a score of 0 for that exam.**

Assignments: All assignments are due at the beginning of class on the date due. As you know, this class has lab sections.

Academic Dishonesty: Plagiarism and cheating are serious offenses and may be punished by failure on exam, paper or project; failure in course; and or expulsion from the University. For more information, refer to the "Academic Honesty Policy" accessible here (http://www.uh.edu/provost/policies/honesty/). The University of Houston Academic Honesty Policy is strictly enforced by the C. T. Bauer College of Business. No violations of this policy will be tolerated in this course. Students are expected to be familiar with this policy.

Need for Assistance: If you have any condition, such as a physical or learning disability, which will make it difficult for you to carry out the work as outlined in this document, or which will require academic accommodations, please notify me as soon as possible. I will recommend that you contact the Center for Students with Disabilities. The contact person is Justin Dart in the CSD building #568, room 110. The numbers for the CSD office are Ph: 713-743-5400; TDD: 713-749-1527; Fax: 713-743-5396 or email: uhcsd@central.uh.edu. Also available to you is *Counseling and Psychological Services (CAPS),* which can help students who are having difficulties managing stress, adjusting to college, or feeling sad and hopeless. You can reach CAPS (www.uh.edu/caps) by calling 713-743-5454 during and after business hours for routine appointments or if you or someone you know is in crisis. In addition, there is no appointment necessary for the "Let's Talk" program, which is a drop-in consultation service at convenient locations and hours around campus. http://www.uh.edu/caps/outreach/lets_talk.html.

Inclement Weather or Technical Problems
In case of inclement weather or technological problems that prevent the University from providing access to course materials you may contact the Professor by phone via the numbers given above or send the Professor an email inquiry. In addition, the Professor will notify students as soon as possible in such instances and provide instructions on how the course will proceed.

Tentative Lecture Outline
This outline is tentative. The order in which these topics are covered can change in the event of unexpected class disruptions.

(week #) Date	Topic	Required Chapter Reading	Required SAS Tutorial Practice
(1) 8/21	Course Overview	Chap 1	SAS-EM14-Basics
	The data mining process		SAS Enterprise Guide Introduction
	Course resources and hands-on environment		
(2) 8/28	Data Preprocessing	Chap 2	SAS-EM-first-lessons
	Understand data	Chap 13.1	
	Summarizing data and visualization		
	Homework 1 due 9/6, 11:59PM		

(3) 9/4	*Case: Telco Telecommunication Churn*	Chap 3	EM-EDA
	Exploratory Data Analysis (EDA)		
	Data visualization		
	Homework 1 due 9/6, 11:59PM		
(4) 9/11	Exploring the Data	Chap 3	EM-EDA
	Exploratory Data Analysis (EDA)		
	Data visualization		
	Homework 2 due 9/20, 11:59PM		
(5) 9/18	Modeling the Data	6, 7, 14.2	EM-Tutorial-ch3.1, section 3.1
	Supervised vs unsupervised model		EM-kNN-MBR
	Model evaluation		
	Over-fitting, cross-validation		
	Classification: K-Nearest Neighbor Algorithm		
	Homework 2 due 9/20, 11:59PM		
	Homework 3 due 9/27, 11:59PM		
(6) 9/25	Classification: Decision Trees	Chap 8	EM-DT-complete
	Decision tree algorithms		
	Homework 3 due 9/27, 11:59PM		
	Homework 4 due 10/4, 11:59PM		
(7) 10/2	Regression Model & Neural Networks 1	5.6 – 5.12	EM-REG
	Logistics regression	Chap 9	EM-NEURONET
	Simple and multiple regression	13.(2, 4, 5 only)	
	NNT: Multilayer perceptron, backward propagation		
	Homework 4 due 10/4, 11:59PM		
	Homework 5 due 10/11, 11:59PM		
(8) 10/9	Material for Quiz #1 review on blackboard		
	Neural Network 1 (cont'd)		
	Homework 5 due 10/11, 11:59PM		
	Final Project Team-up Due 10/23 end of class		

(9) 10/16	**Quiz #1**		
	Neural Network 2	Chap 9	
	Model performance		
	Customized model optimization: using regression		
	Final Project Team-up Due 10/23 end of class		
	Homework 6 due 10/25, 11:59PM		
(10) 10/23	Model Evaluation and Deployment	Chap 14	EM-ASSESS, EM-SCORE
	Scoring the prediction (hold-out) dataset	13.(1, 2, 4, 5 only)	
	Homework 6 due 10/25, 11:59PM		
	Homework 7 due 11/1, 11:59PM		
(11) 10/30	Unsupervised Model: Clustering Analysis	Chap 10	EM-K-MEANS
	K-Means Clustering Algorithm		
	Homework 7 due 11/1, 11:59PM		
	Homework 8 due 11/8, 11:59PM		
(12) 11/6	Unsupervised Model: Association Analysis	Chap 12	EM-AA
	Market-Basket Analysis		
	Homework 8 due 11/8, 11:59PM		
	Homework 9 due 11/15, 11:59PM		
	Final Project Due 11/25, 11:59PM		
(13) 11/13	Final Project Strategic Session		
	Homework 9 due 11/15, 11:59PM		
(14) 11/20	Closing Remarks		
	Final Project Due 11/25, 11:59PM		
11/27	**Thanksgiving Holiday (No Class)**		
(15) 12/4	**Quiz #2**		