Marketing 5963, Spring 2018

Data Mining and Customer Relationship Management Applications

CRN – 26874 Online (OL) Syllabus

Professor: Dr. Miriam McGaugh**,** Clinical Assistant Professor of Marketing
Contact Information:
Office: 307 SSB (old) 429 SSB (new) - Since the Business School is in transition this semester from the old building to the new building, I have included both office numbers. I will not know an exact date when my office will be moving but it should be by the end of spring break. I will notify all students through regular communication channels when my office does move.
Phone: (405) 744-2208, Fax: (405) 744-5180
Office Hours: In the office, by phone or virtually through the online main chat room on Tuesdays from 1:30 to 3:00 pm CST (Central Standard Time) or by appointment.
Course Site: (Brightspace by BRIGHTSPACE): http://online.okstate.edu or http://my.okstate.edu
Online Learning Support: spearsonline@okstate.edu
       Phone: 405-744-4048
       Facebook: Follow Spears School Online Learning on
       Facebook! https://www.facebook.com/SpearsOnline/

Course Description:

Data mining promises to turn business data into actionable information. Therefore, this course will focus on learning how to use various data mining tools such as neural networks, decision trees, classification and prediction algorithms etc. in the context of most common applications in business – sales, marketing, and customer relationship management (CRM). Students will be expected to use state-of-the-art industrial strength data mining software (SAS Enterprise Miner) to analyze real-world data and make strategic recommendations for managerial actions. My philosophy in teaching the course is "you learn by doing," that is, you should be prepared to work extensively with SAS Enterprise Miner in analyzing data sets using various techniques such as neural networks, decision trees, multiple/logistic regression, ensemble models, text mining, etc. The course will use lectures, data analysis using state-of-the-art data mining software, and exercises. Class lectures will be handled via video (video links will be posted on the BRIGHTSPACE course site) that you can watch at your own convenience (you will need a high-speed Internet connection to watch the lectures).

Course Objectives:
This course has five major objectives that fit within five of the program learning goals.

Course Objective	Program Learning Goal
Students will be able to engage in analytical reasoning to break problems into their component parts; identify important patterns by analyzing data; and test for assumptions behind models.	• Critical Thinking

Students will be able to apply science and business principles to analyze and interpret data, using analytic and computer-based techniques.	• Critical and Creative Thinking
Students will be able to present written results from their analyses by relating those back to the business issues that demonstrate a mastery of language and mechanics.	• Written Communication
Students will be able to use appropriate tools and technologies for data visualization and statistical model building	• Technology Skills

Course Prerequisites:
MKTG 5983 or
Consent of MBA, MIS/MSIS, MSTM director or
Consent of Business Analytics director or instructor

Computer and Software Requirements:
- A broadband internet connection
- Windows 7 or Mac OS Mavericks or newer operating system are preferred
- Google Chrome or Mozilla Firefox web browser
    Note: lecture videos are not compatible with Internet Explorer or Edge
- VLC Viewer video player (click on link to download)
- SAS Software – Base SAS, SAS Studio, Enterprise Guide and Enterprise Miner accessible by:
    o https://app.it.okstate.edu/sdc/ - SAS 9.4 32 or 64 bit depending on system (need to download activation file also) – get you all but Enterprise Miner
    o Virtual SAS Server – get you all including Enterprise Miner https://desktop.okstate.edu/
    o SAS University Edition - – get you all but Enterprise Miner https://www.sas.com/en_us/software/university-edition/download-software.html
    Note: Please see the Online Classroom system for additional information on accessing and/or installing SAS software.

Communication Plan:
Teaching Assistant (TA): The TA for this class will be announced via Brightspace during the first two weeks of class. There may be other TAs helping me with this class as well. But, the announced TA will be your primary point of contact for any issues related to this class. The TA will monitor the Brightspace Discussion Board platform twice a day (morning and afternoon). If they are not able to answer your questions immediately, they will contact me and get a response to you as soon as possible (usually within 24 hours).

E-mail: Please use the class discussion board via Brightspace for any general questions, comments, clarifications about any of the class topics (including cases, assignments etc.). Use the e-mail to my TA sparingly and only for questions that disclose or ask for personal

information (such as grades, scores, etc.) *There is no need to copy me with your email to my TA – if my TA is unable to answer your question*, he/she will discuss with me and get back to you.

Response Times: The TA and I will respond to student inquiries within 24 hours during Monday-Friday business hours. Students may expect grades for assignments to be posted to the Gradebook in the online classroom within two weeks of turning in the assignment. ***Please consider these timelines when you are scheduling your course work assignments. While the TA and I will do everything we can to respond in a timely manner, waiting to post a question one hour before the assignment is due will not allow for adequate time for a response.***

Class Discussion via Brightspace (https://online.okstate.edu/): We will use this format extensively for communication among students as well as between students and the instructor. This will be a bulletin-board type system with specific folders for different aspects of this course. There will be multiple forums (folders) in this bulletin board. Please check these folders regularly. Please post your questions only in the **appropriate forums**. Please use appropriate subject line in your posting and use threaded discussion whenever possible.
All students are expected to participate in all aspects of the class. Online discussions can provide everyone with valuable tips and techniques to common problems. One thing I do not want is a string of postings with little to offer on the topic (i.e., Thank you, ditto, me too, etc.). If you are responding to a post, please make it courteous and helpful. Do not type in all caps unless it is within a program or your classmates will think you are mad about something that was said.

Texts and Supplementary Materials:

Class Materials: Most of the class materials will be distributed via the Brightspace web site for this class (https://online.okstate.edu/). If you are a registered student for this class, you should be able to see this course when you log-in to Brightspace (the site becomes active one-week before classes begin). If you have problems accessing the Brightspace class site, please call OSU's IT help desk (405-744-435) or (toll free) 1-877-951-4836).

Required Text: There is no required textbook in this class. I will primarily use readings off the web, cases, SAS training materials, chapters from reference books, etc. in this class (some of these are shown in the schedule). I have also indicated a number of good books (under reference texts) on this topic that you may find useful. Finally, you may need to consult and brush up on basic statistics and probability concepts (any introductory level statistics or marketing research book will work for this purpose).

Reference Texts (whenever possible, these will be put on reserve at OSU Stillwater library)

- Data Mining Techniques for Marketing, Sales and Customer Relationship Management, by Michael J. Berry and Gordon S. Linoff, Wiley Publishing Inc, 2004. (OSU library call number: 658.802 B534d)
- Handbook of Statistical Analysis and Data Mining Applications, by Nisbet, Elder and Miner, Academic Press, 2009. (OSU Library Call No. 006.312 N724h)
- Principles of Data Mining, by Hand , Mannila, and Smyth. MIT Press 2001. (OSU library call number: 006.3.H236p)

- Data Preparation for Data Mining by Dorian Pyle, Morgan Kauffman publications, 1999. (OSU library call number : 005.74 P996d)

Grading Policy

The grades in this class break down as follows:

Project Deliverables	240 pts
Assignments	200 pts
Exams	270 pts
Total Points	710 pts

Letter grades will be assigned according to the standard scale.

635-710 pts. = A
564-634 pts. = B
493-563 pts. = C
422-492 pts. = D
0-421 pts. = F

Late Assignments: Assignments must be turned in according to the date and time on the syllabus via Brightspace drop box (not emails). All late assignments (*even 1-minute late*) must be turned in **via the Late Drop Box** and will be *penalized* as follows:

- One late assignment (within 48-hours of due date and time) – *no penalty*
- All other late assignments will carry following penalty structure:
    - Within 1 hour of due date and time – 15% penalty
    - More than 1 hour but less than 24 hours of due date and time – 30% penalty
    - More than 24 hours but less than 48 hours of due date and time – 50% penalty
    - More than 48 hours of due date and time – will not be graded (no credit)

I enforce this rule because I believe that part of effective functioning in business is the ability to complete projects on time. **Please do not email/call/contact me or my TA with excuses (however valid they may be) about making exceptions to my late submission policy.**

Attendance Policy: Students who have not participated in at least one assignment in Brightspace (BRIGHTSPACE) (discussion board, specific assignment, etc.) within the first two weeks (depends if course is a short length or full semester length- could be the first 2 days) of the course will be reported as not having attended class. The instructor will then recommend the student to drop the course.

Course Format:

Lectures: The link for each video lecture will be posted on the Brightspace. It is <u>your responsibility</u> to watch the lecture video and do appropriate readings/exercises. *All students* will be given access to lecture videos.

Labs: There is not a lab section for this course.

Exercises: Because of the emphasis on "hands-on learning" in this course, you will work on several exercises. These exercises will reinforce the concepts covered in the video lectures and demonstrations. These will count for **a total of 200 points**. You must submit these exercises on time via the appropriate Dropbox in the Brightspace site. All exercises should be done individually.

Project: You will be working on/ continuing to work on a group project throughout the semester. Project deliverables will be broken into two sections. The first deliverable will be a preliminary report of the modeling processes that you will be using for your project. Preliminary data and results will be submitted in this report. The second deliverable will be the final project report. You may add to the previous analysis based on the material you learned during the second portion of the semester.

More information will be provided via Brightspace as the semester progresses. Reports should be turned into the appropriate Dropbox by **midnight** CST (Central Standard Time) of the due date. See Late Policy for information regarding late submissions. The project deliverables will be graded using the rubric in Appendix A and both deliverables combined will account for **240 points of the course grade**.

Exams: You will take two exams throughout the semester. These exams include multiple-choice questions and essay writing. The exam content will be largely based on video lectures and readings. One mid-term (**120 points**) and <u>comprehensive </u>final exam (**150 points**). The exams will be proctored exams and you will have a 48-hour window in which to complete them (please consult tentative schedule in Appendix B).

> **ACTION:** one week prior to course start date, go to the Spears School of Business Online Learning website to choose a testing center at: <u>spearsonline.okstate.edu</u>, <u>and click on "Select Testing Center" at the top right of the page.</u> Follow the instructions to identify your testing center. Up to one week before each exam start date, make your appointment directly with your testing center to take each exam while being monitored by a proctor for test security reasons. The exam and/or exam instructions will be sent to your testing center 3 days prior to the exam start date. To confirm your testing center received the exam/exam information, call at least one day prior to your appointment time. If the center does not have your exam, contact the Spears School Online Learning office immediately at <u>spearsonline@okstate.edu</u>, or call (405) 744-4048 to request the exam to be sent. Contact that same office if you have any questions regarding the testing center sign up process. You may also visit <u>http://spears.okstate.edu/online/guide</u>.

Make-up Policy:
Students are expected to take each exam on the date given and submit each assignment in a timely manner. If for any reason a student cannot attend an exam or submit an assignment, he or she must notify the instructor prior to the due date.

University Policy:
Drop Policy
Information about university drop policy and dates is at this website: http://registrar.okstate.edu/
To drop this course, contact the Registrar's office, (405) 744-6876, or drop through Banner Self Service, http://my.okstate.edu

Academic Integrity
Oklahoma State University is committed to the maintenance of the highest standards of integrity and ethical conduct of its members. This level of ethical behavior and integrity will be maintained in this course. Participating in a behavior that violates academic integrity (e.g., unauthorized collaboration, plagiarism, multiple submissions, cheating on examinations, fabricating information, helping another person cheat, unauthorized advance access to examinations, altering or destroying the work of others, and fraudulently altering academic records) will result in your being sanctioned. Violations may subject you to disciplinary action including the following: receiving a failing grade on an assignment, examination or course, receiving a notation of a violation of academic integrity on your transcript (F!), and being suspended from the University. You have the right to appeal the charge. Contact the Office of Academic Affairs, 101 Whitehurst, 405-744-5627, **http://academicintegrity.okstate.edu/**.

Accessibility
Any student in this course who has a disability that may prevent him or her from fully demonstrating his or her abilities should contact the instructor as soon as possible, so we can discuss accommodations necessary to ensure full participation and facilitate your educational opportunity. For more information about OSU Student Disability Services, please go to: http://sds.okstate.edu.

Syllabus Attachment
For more student resources, go to: https://academicaffairs.okstate.edu/content/resources-students

Appendix A: Project Deliverable Rubric

	Project Deliverable Rubric					Student's Score
Rubric Components	**Point Scale**					
	4 **Exceeds Standard**	**3** **Proficient** (at standard, minor errors)	**2** **Developing** (emerging, needs some help)	**1** **Novice** (can't do independently, doesn't understand)	**0** **Not Included**	
Introduction and Overview of Research/ Business Problem: Research/business question, variables and data description, and rationale explained adequately.						
Statistical Analysis: Adequate overview of descriptive statistics; Appropriate statistical/modeling techniques are conducted correctly and explained adequately						
Interpretation of Results: Appropriate interpretation of techniques in context of research/ business questions and explained adequately						
Conclusion and Discussion: Adequate discussion of results and their practical implications; Reasonable explanation of findings offered; Recommendations (if applicable) are included and appropriate; Limitations and future research discussed						
Readability & Organization: Few writing errors and generally readable						

Appendix B: Tentative Schedule

Week	Starting Date	Content	Readings	Exercise	Exercise Due Date
1	16-Jan-18	Video 1: Course and faculty introduction Video 2: Overview of data mining concepts Video 3: Overview of SAS Enterprise Miner interface	What is data mining? (http://en.wikipedia.org/wiki/Data_mining) (http://www.twocrows.com/booklet.htm)	Demo 1	NA
2	22-Jan-18	Video 1: Data Preparation Video 2: Accessing and Assaying Prepared Data	External Reading 1:"Competing on Analytics" by Thomas Davenport, Harvard Business Review – to read this paper please go to OSU's library site, http://www.library.okstate.edu/ and then search for this article online. Optional: visit this web site and scan some of the papers/articles about data mining (http://www.thearling.com/index.htm#wps)	Exercise 1 - Accessing and Assaying Data	2/1/2018
3	29-Jan-18	Video: Exploring and Preparing Data	Read the below Webpages for different Exploratory Data Analysis techniques (A) http://www.togaware.com/datamining/survivor/Exploring_Data.html (B) http://en.wikipedia.org/wiki/Exploratory_data_analysis (C) http://www.itl.nist.gov/div898/handbook/prc/section1/prc16.htm Problems with Missing Data: http://www.kdd.org/exploration_files/12-Pearson.pdf	Exercise 2 - Exploring and Preparing Data	2/8/2018
4	5-Feb-18	Video: Introduction to Predictive Modeling via Decision Trees	More Information on Decision Trees for Predictive Modeling http://www.wuss.org/proceedings10/analy/3055_2_ANL-Hobbs.pdf http://www.statsoft.com/Textbook/classification-trees	Exercise 3 - Decision Trees	2/15/2018
5	12-Feb-18	Video: Introduction to Predictive Modeling via Regression Models	a. Basics of logistic regression: http://faculty.cas.usf.edu/mbrannick/regression/Logistic.html b. A detailed description of logistic regression mechanics, assumptions etc. It uses SPSS output in the examples, but the terms used are essentially same between SAS and SPSS: http://faculty.chass.ncsu.edu/garson/PA765/logistic.htm c. Odds ratio : http://en.wikipedia.org/wiki/Odds_ratio d. Regression vs. Decision Trees: http://www.forecastingprinciples.com/paperpdf/exploratory.pdf	Exercise 4 - Regression	2/22/2018
6	19-Feb-18	Video: Introduction to Predictive Modeling via Neural Networks	For more information on Artificial Neural Networks read below webpages http://www.statsoft.com/Textbook/Neural-Networks http://en.wikipedia.org/wiki/Artificial_neural_network http://ulcar.uml.edu/~iag/CS/Intro-to-ANN.html Multilayer Perceptron: http://users.ics.tkk.fi/ahonkela/dippa/node41.html	Exercise 5 - Neural Network	3/1/2018

7	26-Feb-18	Video: Model Assessment	Sensitivity and Specificity: http://en.wikipedia.org/wiki/Sensitivity_and_specificity More on ROC Curves: (a) http://homepage.cs.uri.edu/faculty/hamel/pubs/hamel-roc.pdf (b) http://gim.unmc.edu/dxtests/ROC1.htm	Exercise 6 - Model Assessment	3/8/2018
8	5-Mar-18	Video: Model Implementation	If any (beyond lecture materials posted on BRIGHTSPACE), will be announced via BRIGHTSPACE. Partial Least Squares: http://www.statsoft.com/textbook/partial-least-squares/	Modeling Phase and Preliminary Results	3/8/2018
9	12-Mar-18	Midterm Exam			
10	19-Mar-18	Spring Break			
11	26-Mar-18	Video: Special Topics	Ensemble Model: http://www.statsoft.com/textbook/neural-networks/#ensembles http://www.scholarpedia.org/article/Ensemble_learning Variable Reduction using Clustering: http://www.casact.org/pubs/forum/06wforum/06w93.pdf Principal Component Analysis: http://support.sas.com/publishing/pubcat/chaps/55129.pdf	Exercise 7 - Ensemble Models and Scoring	4/5/2018
12	2-Apr-18	Video: Pattern Discovery Techniques - Segmentation	Market Segmentation: http://en.wikipedia.org/wiki/Market_segmentation http://facweb.cs.depaul.edu/research/vc/publications/C-262.pdf		
13	9-Apr-18	Video: Pattern Discovery Techniques - Clustering	Clustering: http://en.wikipedia.org/wiki/Cluster_analysis http://www.statsoft.com/textbook/cluster-analysis/?button=1 http://en.wikipedia.org/wiki/K-means_clustering http://home.dei.polimi.it/matteucc/Clustering/tutorial_html/hierarchical.html	Exercise 8 - Clustering	4/19/2018
14	16-Apr-18	Video: Association Rule Discovery Techniques	Market Basket Analysis: http://en.wikipedia.org/wiki/Association_rule_learning http://www.statsoft.com/textbook/association-rules/?button=1 Teknomo, Kardi (2008) Market Basket Analysis. http://people.revoledu.com/kardi/tutorial/MarketBasket/	Exercise 9 - Association Rule	4/26/2018
15	23-Apr-18	Video: Text Mining	http://www.statsoft.com/textbook/text-mining/?button=3 "Finding Nuggets in Textual Information" - pdf on Brightspace SVD: http://web.mit.edu/be.400/www/SVD/Singular_Value_Decomposition.htm	Exercise 10 - Text Mining	5/3/2018
16	30-Apr-18			Final Project Report	5/3/2018
17	7-May-18	Final Exam – Date and Time To Be Announced			