

## **SYST 499 Applied Predictive Analytics Spring 2017**

Class time: 10:30am-11:45am, Tuesday, Thursday  
Room: ENGR 1107  
Instructor: Prof. Jie Xu  
Email: jxu13@gmu.edu  
Office: Engineering Building Room 2218  
Phone: (703) 993-4620  
Office Hours: Tue, Thu 2:00 - 3:00 PM or by appointment

### **Course Description:**

This course introduces students to the fundamentals of data analysis and some of the most widely used models in applied predictive analytics. The students learn how to summarize data and explore relationship between variables, including principle component analysis and multidimensional scaling. Class instruction follows with a presentation of commonly used tables, visualizations, and statistical tests for comparing groups. Linear predictive models for both continuous and binary outcomes (logistic regression) are discussed in detail. The course introduces students to clustering and classification using random forest and naïve Bayes. The course concludes with topics on choice modeling. Hands-on programming with R is emphasized. While no prior knowledge on R is required, students must be well prepared in programming.

### **Prerequisite:**

(C or higher in STAT 344 or STAT 346 or MATH 351 or in STAT 250), and (C or higher in CS 112 or IT 206)

**Grading:** Homework 30%; midterm 35%; final exam 35%.

**Late homework submissions will not be accepted. Email submission will not be accepted.**

Homework problems should be worked out independently but discussions are allowed.

Make-up exam can **only** be granted on the following basis (**the only exception is medical emergencies that you cannot know beforehand**):

You notify the instructor **at least 2 days before the exam** that you must be absent because of medical conditions (documentation from doctors required) or other circumstances that you have no control over. **Make-up will then be given one day before your scheduled absence** and the exam will at least be at the same level of difficulty as the regular exam.

Final letter grades are assigned as follows:

A+: 97-100, A: 93-96, A-: 90-92, B+: 87-89, B: 83-86, B-: 80-82, C+: 77-79, C: 73-76, c-: 70-72, D+: 67-69, D: 63-66, D-: 60-62, F: <60

### **Textbooks**

Required text:

C. Chapman and E. Feit, R for Marketing Research and Analytics, Springer, 2015.

### **Software**

R is a very popular free software package for statistical computing and analytics, and will be used in this class. You can download the latest release of R at <https://cran.r-project.org/>. There are free IDE for R that you may find helpful, such as R Studio. The instructor does not endorse any specific IDE for R.

#### **Topics to be covered & reading materials:**

- An overview of the R Language
- Describing data
- Relationships between continuous variables
- Comparing groups using tables and visualizations
- Comparing groups using statistical tests
- Linear models for predicting continuous outcomes, performance evaluation
- Linear models for predicting continuous outcomes, performance evaluation
- Reducing data complexity with principle component analysis, factor analysis, multi-dimensional scaling
- Linear models for predicting binary outcomes (logistic regression)
- Segmentation analysis, clustering
- Classification
- Association rules
- Choice modeling

#### **Academic Integrity**

*GMU is an Honor Code university; please see the Office for Academic Integrity for a full description of the code and the honor committee process. The principle of academic integrity is taken very seriously and violations are treated gravely. What does academic integrity mean in this course? Essentially this: when you are responsible for a task, you will perform that task. When you rely on someone else's work in an aspect of the performance of that task, you will give full credit in the proper, accepted form. Another aspect of academic integrity is the free play of ideas. Vigorous discussion and debate are encouraged in this course, with the firm expectation that all aspects of the class will be conducted with civility and respect for differing ideas, perspectives, and traditions. When in doubt (of any kind) please ask for guidance and clarification.*

#### **Disabilities Statement**

*If you have a documented learning disability or other condition that may affect academic performance you should: 1) make sure this documentation is on file with Office of Disability Services (SUB I, Rm. 4205; 993-2474; <http://ods.gmu.edu>) to determine the accommodations you need; and 2) talk with me to discuss your accommodation needs.*

#### **Mason Diversity Statement**

George Mason University promotes a living and learning environment for outstanding growth and productivity among its students, faculty and staff. Through its curriculum, programs, policies, procedures, services and resources, Mason strives to maintain a quality environment for work, study and personal growth.

An emphasis upon diversity and inclusion throughout the campus community is essential to achieve these goals. Diversity is broadly defined to include such characteristics as, but not limited to, race, ethnicity, gender, religion, age, disability, and sexual orientation. Diversity also entails different viewpoints, philosophies, and perspectives. Attention to these aspects of diversity will help promote a culture of

inclusion and belonging, and an environment where diverse opinions, backgrounds and practices have the opportunity to be voiced, heard and respected.

The reflection of Mason's commitment to diversity and inclusion goes beyond policies and procedures to focus on behavior at the individual, group and organizational level. The implementation of this commitment to diversity and inclusion is found in all settings, including individual work units and groups, student organizations and groups, and classroom settings; it is also found with the delivery of services and activities, including, but not limited to, curriculum, teaching, events, advising, research, service, and community outreach.

Acknowledging that the attainment of diversity and inclusion are dynamic and continuous processes, and that the larger societal setting has an evolving socio-cultural understanding of diversity and inclusion, Mason seeks to continuously improve its environment. To this end, the University promotes continuous monitoring and self-assessment regarding diversity. The aim is to incorporate diversity and inclusion within the philosophies and actions of the individual, group and organization, and to make improvements as needed.

#### **Student Support Resources on Campus**

Resources that you may find helpful may be found at: <http://ctfe.gmu.edu/teaching/student-support-resources-on-campus/>