STA 6543 01T

Predictive Modeling

Course Syllabus | Summer 2024

Course Information

Course Description: This course presents students with basic understanding of predictive modeling techniques and predictive analytics tools, with specific emphasis on problem-solving with real data using R programming. Topics include data preprocessing, over-fitting and model tuning, supervised learning methods, including linear regression and classification, nonlinear regression and classification models, resampling methods, model regularization, tree and rule-based methods, and support vector machines. Unsupervised learning methods include principal component analysis, clustering methods, and outlier detection. Students will learn how to select various predictive modeling algorithms for a wide variety of applications and how to code the programs in R, as well as assumptions and requirements of predictive modelling, optimal tuning parameter setting, and how to interpret and report the results.
Credit Hours: 3
Course Modality: Online only, no set time

Learning Goals

At the end of the course, you will be able to:

- Recognize basic predictive modeling skills in supervised as well as unsupervised statistical learning;
- Compare selected statistical learning (data mining, machine learning or data analytics) concepts and tools;
- Apply appropriate predictive modeling techniques in analyzing the vast amounts of data found in biology, business, and other industries;
- Implement selected predictive modeling tools in R programming and interpret numerical results;
- Work on various statistical and machine learning data analyses;
- Produce a complete data analysis report.

Communicate with Me

Instructor Name: Min Wang
Department
Management Science and Statistics
Office Location
SPI 470E - San Pedro I, 506 Dolorosa St.
Student Hours
Virtual office via the ZOOM https://utsa.zoom.us/j/97952044637
Phone Number(s): 210-458-5381
Email Address: min.wang3@utsa.edu
Preferred Method of Communication
Email

About Me & My Teaching Philosophy

About myself: I am a Full Professor of Statistics in the Department of Management Science and Statistics at The University of Texas at San Antonio. My main research spans the areas of Bayesian inference and methods, high-dimensional inference, statistical machine learning, prior elicitation, quantile regression, and statistical modeling, all in both methodological and theoretical perspectives.

My Teaching Philosophy:

My teaching goals are to inspire students' interest in statistics and data science, to guide students to understand how statistics relate to the real world, to develop students' critical thinking skills and to encourage students to use statistical concepts and data analytics tools to solve problems encountered in their other studies or everyday life.

The first component of my teaching philosophy is to use a data-oriented instruction to motivate students' interest in learning statistics and data science. Based on my 12 years of teaching experience from introductory to advanced statistics courses, I have concluded that the most challenging part of teaching statistics is not the material, but students' attitudes and fears of the subjects. Motivating examples related to real-life situations are particularly useful for enhancing student motivation. For instance, many undergraduate students seem to be confused by the use of the multiplication rule in conditional probability. I use an example of the drawing straws game that students often play. It is easier for them to understand through the game why the probability of the last draw is the same as the first draw, which indicates that the order of drawing is not important. For STA 6113 Applied Bayesian Statistics, I use a Bayesian forecast for U.S. presidential election to illustrate how prior knowledge elicitation from the polls and experts' opinion updates Bayesian decision. It has been proven that this philosophy works well for increasing student motivation and participation.

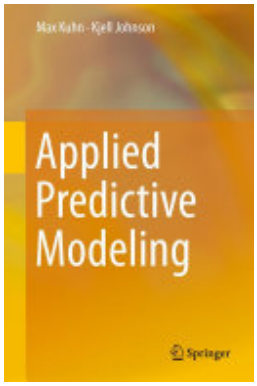
The second component of my teaching philosophy is to make students actively participate in their learning process. In this way, students can get a deep understanding of what they learn in class and develop their critical

thinking skills. It also helps me to monitor their learning progress. If I talk for the whole class by myself, I may think that the students have learned a lot from me. However, I would not really know how much the students have absorbed from each class and how well they understand the concepts. I strongly encourage students to ask questions in class. I say, ``Whenever a question pops into your head, please raise your hand. Of course, it is better to be course related; otherwise I might not be able to answer you.'' If students ask questions, it is a sign that students are listening and thinking. I also assign group learning activities to students during class time, since group activities involve interactions which not only prevent boredom but also enhance students understanding of class materials by sharing ideas and asking each other. For instance, for the M.S. in Data Analytics (MSDA) courses, I usually release all the course materials and programming in advance in Canvas and then teach the first half of the course about statistical modeling for data analysis. Then I give learning activities to students so that they can replicate my computations and carry out new statistical concepts and tools and thus advance the profession. Finally, a course project will be assigned to have them collect data from their everyday lives and jobs and perform data analysis as a teamwork using statistical methods learned in the course along with software tools.

The third component of my teaching philosophy is to be flexible. I adjust my teaching load based on students' feedback. I stay longer after each class if students have questions and encourage students to come see me if they have questions. If any student wants to meet me outside of office hours, they can always make an appointment with me by email or phone. During the current COVID-19 pandemic, most of my courses are delivered entirely through either synchronous or asynchronous teaching, so I always make myself be available even after 10:00 pm to the midnight to help more students to be successful in the course. In addition, to help students prepare the written qualifying exams in Advanced Inference and Applied Bayesian Statistics, I hold extra office hours and meetings during the summer break based on their needs.

These 3 components of my teaching philosophy are interrelated. If I as an instructor can inspire students' interest, they will be willing to participate in class. If they participate in class more often, they will keep up their interest and be active participants in their learning. Attaining these first two goals also requires that an instructor should be reasonably flexible in both teaching style and schedule.

Course Materials



Applied Predictive Modeling
ISBN: 9781461468493
Authors: Max Kuhn, Kjell Johnson
Publisher: Springer Science & Business Media
Publication Date: 2013-05-17
Additional Information
http://appliedpredictivemodeling.com/

Additional Course Information

1. **Statistical software**: A statistics package is essential for this type of applied course. The one that we will use in this course is **R** or **Rstudio**, which can be downloaded from http://www.r-project.org/. Also make sure to install AppliedPredictiveModeling package, which includes the datasets used in the course book.
2. **One useful reference book**:
   An Introduction to Statistical Learning: with Applications in R
   ISBN: 1071614177
   Authors: Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani
   Publisher: Springer
   Website: https://www.statlearning.com/

Assessments and Assignments

1. **Exercise**: There are 6 exercises for continuous formative assessment of the progress of the course.
2. **Homework**: Three homework assignments will be assigned in a regular basis.
3. **One Midterm**: This midterm will be available from 12:00AM CST 07/08/2024 to 11:59PM CST 07/11/2024. You can download the midterm examination program and work on these problems on your own during this time period. You must show all your work in the exam to receive credit. Each student must write his/her own exam solution. Just make sure that you submit your work before the deadline.
4. **One Individual/Group Project with Poster Presentation**: This individual/group project (at most three students) will be assigned and finished during the semester.
    i. A poster needs to be submitted to the discussion board on Canvas before 11:59 pm CT on Wednesday, August 07, 2024.
    ii. Discussions to your fellow groups' posters need to be made before 11:59 pm CT on Wednesday, August 10, 2024.
    iii. An individual/group project needs to be submitted via Canvas before 11:59 pm CT on Saturday, August 10, 2024.

Notes:

1. All work must be of professional quality, neatly presented, grammatically correct, and free of spelling and punctuation errors. Late work is allowed under some circumstances and with prior approval from the instructor. Points will be deducted at a rate of 10 percentage points per day, including weekends and holidays. Work is accepted up to 3 days after the due date. At that time, the grade will be zero.
2. **Make Ups**: Make-up midterm/Exercise/Homework is given for approved reasons such as participation in sanctioned extramural activities, Documented illness etc. Notification is requested PRIOR to the midterm, or as soon as reasonably possible in cases of emergency. Late submissions will NOT be accepted for any assignment, exercise, project, and poster presentation.

3. **Announcements**: All information you need for this course will be posted in Canvas in the syllabus or schedule, or as an assignment, item, or announcement. It is your responsibility to check in and participate every week in the course and complete all listed activities and assignments. It is recommended that you check announcements daily.

Activities and Grading

Activity	Quantity	%
Exercise	6	18
Homework	3	24
Midterm	1	30
Poster and Discussion	1	8
Project	1	20
Total		**100%**

Distribution of Course Assignments, Their Quantity, and Contribution to Final Grade.

Grading and Feedback: To view your grades on Canvas, click on the “Grades” tab on the sidebar menu of our course in Canvas. If additional feedback has been included with your grade, a speech-bubble icon will appear next to your grade. Clicking on this icon will open an additional window on your screen providing you with feedback.

Grade Distribution and Letter Grade

Letter Grade	Grade Range
A+	96.5 – 100%
A	92.55 – 96.49%
A-	89.5 – 92.49%
B+	86.5 – 89.49%
B	82.5 – 86.49%
B-	79.5 – 82.49%
C+	76.5 – 79.49%
C	72.5 – 76.49%
C-	69.5 – 72.49%
D+	66.5 – 69.49%
D	62.5 – 66.49%
D-	59.5 – 62.49%

Letter Grade	Grade Range
F	<59.5%

grading scheme

Course Expectations & Policies

Instructor-Initiated Drops

This course uses instructor-initiated drops for students who exceed the absence and/or missed assignment limit. Therefore, up to the last day for students to withdraw from an individual course, August 2, 2024, you will be dropped for exceeding 4 assigned activities, including exercise, homework, examination, project, and poster presentation. Students will receive at least one courtesy warning when approaching the absence/missed assignment limit. Notification will be sent using ASAP to the student's email address. A subsequent absence or missed assignment will result in being dropped from the course. Notification of being dropped will also be sent through ASAP to the student's email address. *This drop does not affect enrollment in other courses.* **Please consult the Dropping Courses webpage** for further details on the process and appeals.

Video and Audio Recording

As the instructor of this course, I may record meetings and lessons. You are expected to follow appropriate University policies and maintain the security of passwords used to access recorded lectures. Recordings may not be published, reproduced, or shared with those not in the class. If the instructor or a UTSA office plans any other uses for the recordings, consent of the students identifiable in the recordings is required before such use unless an exception is allowed by law. For more information on your privacy and class recordings, review Student Privacy (FERPA) in Virtual Classrooms and Other Educational Recordings and the Guide to Secure Video Conferencing Tools.

Academic Freedom

Academic freedom[1] is a cornerstone of the University. Academic freedom in its teaching aspect is fundamental for the protection of the rights of the teacher in teaching and of the student to freedom in learning.[2] Each faculty member is entitled to full freedom in the classroom discussing the subject that the faculty member teaches.[3] The University of Texas at San Antonio will not penalize or discipline members of the faculty because of their exercise of academic freedom.

Along with this freedom comes responsibility. It is the responsibility of faculty members to ensure that topics taught are related to the classroom subject. Students should be free to take reasoned exception to the data or views offered in any course of study and to reserve judgment about matters of opinion, but they are responsible for learning the content of any course of study for which they are enrolled.[4] It is not the proper role of the university or any outside agency to attempt to shield individuals from ideas and opinions they find unwelcome, disagreeable, or even deeply offensive.[5] Engaging with new ideas and perspectives helps students grow intellectually and is beneficial to the educational process.

1. Statement adapted from Texas A&M University's Syllabus Statement Regarding Academic Freedom
2. 1940 Statement of Principles on Academic Freedom and Tenure
3. Board of Regents Rule 31004; HOP Policy 4.02
4. American Association of University Professors Joint Statement on Rights and Freedoms of Students
5. The Chicago Statement

UTSA Inclusivity Statement

The University of Texas at San Antonio, a Hispanic-serving institution situated in a global city that has been a crossroads of peoples and cultures for centuries, values diversity and inclusion in all aspects of university life. As an institution expressly founded to advance the education of Mexican Americans and other underserved communities, our university is committed to promoting access for all. UTSA, a premier public research university, fosters academic excellence through a community of dialogue, discovery, and innovation that embraces the uniqueness of each voice.

Syllabus Changes

The syllabus is subject to change at the instructor's discretion. Any changes/corrections to the course materials, assignment dates, or other updates will be communicated to the students ahead of time. You are responsible for checking Canvas for corrections or updates to the syllabus.

Course Schedule

For a list of important university-wide dates, review One Stop's academic calendar.

Due Date	Activities/Assignments	Points
6/2	Exercise 0	30
6/9	Exercise 1	30
6/23	Homework 1	100
6/16	Exercise 2	30
6/30	Exercise 3	30
7/7	Exercise 4	30
7/11	Midterm	100
7/14	Homework 2	100
7/21	Exercise 5	30

Due Date	Activities/Assignments	Points
8/9	Homework 3	100
8/10	Project	30
8/10	Poster Submission and Discussion	20

WEEK 1: [05/28/2024-06/02/2024]

i. Chapter 1: Introduction to Predictive Modeling
ii. Demonstration 1: Introduction to R
iii. Chapter 2: A short tour of the predictive modeling process
iv. Demonstration 2: Predicting Fuel Economy in Chapter 2

WEEK 2: [06/03/2024-06/09/2024]

i. Chapter 3: Data Pre-processing
ii. R demonstration 3(1): Data scaling, transformation, and outliers
iii. R demonstration 3(2): PCA, creating dummy variables, removing and binning predictors

WEEK 3: [06/10/2024-06/16/2024]

i. Chapter 4: Over fitting and model tuning
ii. R Demonstration Chapter 4(1) Model Tuning for kNN method
iii. R Demonstration Chapter 4(2) Model Tuning for SVM method and model comparison
iv. Chapter 5: Measuring Performance in Regression Models

WEEK 4: [06/17/2024-06/23/2024]

i. Chapter 6: Linear Regression and Its Cousins
ii. R Demonstration Chapter 6: Data pre-processing for Solubility data and Linear models, PCR, and PLS
iii. R demonstration for Chapter 6: Penalized models: Ridge and ENET

WEEK 5: [06/24/2024-06/30/2024]

i. Chapter 7: Neural Networks
ii. Chapter 7: Multivariate Adaptive Regression Splines
iii. Chapter 7: Support Vector Machines
iv. Chapter 7: K-Nearest Neighbors and Model Comparisons
v. R demonstrations for Chapter 7: Nonlinear regression models

WEEK 6: [07/01/2024-07/07/2024]

i. Independence Day July 4, 2024

ii. Review previous materials for the midterm

WEEK 7: [07/08/2024-07/14/2024]

i. Chapter 8: Regression Trees and Rule-Based Models
ii. R demonstration for tree-based models in Chapter 8

WEEK 8: [07/15/2024-07/21/2024]

i. Chapter 11: Measuring Performance in Classification Models
ii. Chapter 12: Discriminant Analysis and Other Linear Classification
iii. R Demonstration for Chapter 12: Discriminant Analysis and Other Linear Classification

WEEK 9: [07/22/2024-07/28/2024]

i. Chapter 13: Nonlinear Classification Models
ii. R Demonstration for Chapter 13: Nonlinear Classification Models

WEEK 10: [07/29/2024-08/07/2024]

i. Chapter 14: Classification Trees and Rule-Based Models
ii. R Demonstration for Chapter 14: Classification Trees and Rule-Based Models

Essential Student Information

- **Important:** Bookmark and visit the Common Syllabus Information webpage to find important and valuable resources about counseling services, transitory/minor medical issues, supplemental instruction, tutoring services, academic success coaching, sexual harassment and sexual misconduct, campus safety and emergency preparedness, and the Roadrunner Creed.
- For technical requirements, support, and resources, visit University Technology Solutions (UTS) Student Services.
- UTSA provides reasonable accommodations to students via Student Disability Services. For more details on eligibility, policies, and requirements, please visit www.utsa.edu/disability or call (210) 458-4157.
- Visit the UTSA Libraries and Museums website for access to journals, research tutorials, tech gear you can borrow, and to find your department's librarian.
- Follow Digital Learning Netiquette standards for your online communication activities. Please be mindful of the communication tools available in your course and use them for learning purposes. Class discussions take place in a respectful and safe environment, whether online or in person. UTSA encourages everyone to openly share their ideas and opinions without penalty or judgment, but learning should always be based on facts and research. It is possible to disagree without being disagreeable.

University Policies & Resources

Counseling Services

Wellness 360 Counseling Services provides confidential, professional services by staff psychologists, social workers, counselors and psychiatrists to help meet the personal and developmental needs of currently enrolled students. Services include individual brief therapy for personal and educational concerns, couples/relationship counseling, and group therapy on topics such as college adaptation, relationship concerns, sexual orientation, depression and anxiety. The Wellness 360 counselors also screen for possible learning disabilities and mental health support. Visit the UTSA Wellbeing Services website or call (210) 458-4140 (option 2).

Student Code of Conduct and Scholastic Dishonesty

The Student Code of Conduct is Section B of the Appendices in the Student Information Bulletin. Scholastic Dishonesty is listed in the Student Code of Conduct (Sec. B of the Appendices) under Sec. 203.

Students with Disabilities

The University of Texas at San Antonio, in compliance with the Americans with Disabilities Act and Section 504 of the Rehabilitation Act, provides “reasonable accommodations” to students with disabilities. Only those students who have officially registered with Student Disability Services and requested accommodations for this course will be eligible for disability accommodations. Instructors at UTSA must be provided official notification of accommodation through Student Disability Services. Information regarding diagnostic criteria and policies for obtaining disability-based academic accommodations can be found at www.utsa.edu/disability or by calling Student Disability Services at (210) 458-4157. Accommodations are not retroactive.

Transitory/Minor Medical Issues

In situations where a student experiences a transitory/minor medical condition (e.g. broken limb, acute illness, minor surgery) that impacts their ability to attend classes, access classes or perform tasks within the classroom over a limited period of time, the student should refer to the class attendance policy in their syllabus.

Copyright and Fair Use

It is important to understand the issue of intellectual property rights. You may not use the images or thoughts of others for profit or gain without their written permission. The UTSA library has a Copyright Laws and Public Performance Rights (PPR) page.

Family Educational Rights and Privacy Act (FERPA)

FERPA grants students the right to control certain disclosures of their educational records. For a full explanation of your rights and to grant access to FERPA educational records, go to Student Catalog Annual FERPA Letter and One Stop Enrollment – FERPA Proxy Access. Without your consent or authorization of proxy access,

UTSA may release Directory Information, such as but not limited to your name, email, phone, place of birth, and photograph, unless you have opted out of the release of Directory Information. To opt out, go to Restrict Directory Information Form.

Mandatory Reporting of Sexual Misconduct and Reporting of Health and Safety Information: If a student discloses an incident of sexual misconduct to any UTSA employee (other than to a designated confidential employee such as mental health counselor or PEACE advocate, a UTSA police officer using a pseudonym form or at a public awareness event), that information is not confidential, and the UTSA employee must report all known information to the UTSA Office of Equal Opportunity Services. Employees may also report any concerns about the health and safety of students or others to other school officials and/or law enforcement. For a complete list of exceptions to FERPA, please see Student Catalog Annual FERPA Letter and HOP 5.01.

Supplemental Instruction

Supplemental Instruction offers student-led study groups using collaborative learning for historically difficult classes. Supported courses and schedules can be found on the Supplemental Instruction website. You can call the SI office if you have questions or for more information at (210) 458-7251.

Tutoring Services

Tutoring Services offers peer-to-peer academic support to students enrolled in a variety of courses. Both drop-in tutoring and one-on-one appointments are available. For more information, visit the Tutoring Services website or call (210) 458-6783 on the Main Campus and (210) 458-2838 on the Downtown Campus.

Student Success Coaching

The Student Success Coaching program offers one-on-one study skills assistance through academic coaching. Students meet by appointment with a professional to develop more effective study strategies and techniques that can be used across courses. Group workshops are also offered each semester to help students defeat common academic challenges. Find out more information on the Student Success Coaching website or call (210) 458-4694.

Sexual Harassment and Sexual Misconduct

UTSA is committed to providing an environment free from all forms of discrimination and sexual harassment, including sexual misconduct, sexual assault, domestic violence, dating violence, and stalking. If a student has experienced or experiences any of these incidents, know that UTSA has resources to help.

UTSA faculty have the responsibility to create a learning environment that is safe and free from hostility. State and federal law as well as UTSA's Handbook of Operating Procedures (HOP 9.24) require that instructors must report incidents of sexual harassment and sexual misconduct they learn about to the Title IX Coordinator or a Deputy Title IX Coordinator. This means that if a student tells their instructor about a situation (including classroom discussions, written work and/or one-on-one meetings) involving sexual harassment, sexual assault, dating violence, domestic violence, or stalking, the instructor must report it to the EOS/Title IX Office. Although the faculty member must report the situation, the student will still have options about how their case will be

handled, including whether or not they wish to pursue a formal complaint. The university's goal is to make sure students are aware of the range of options available to them and have access to the resources they need.

If a student wishes to speak to someone confidentially, they can contact any of the following on-campus resources, who are not required to report the incident to the EOS/Title IX Office: (1) Wellness 360 Counseling Services at 210-458-4140; (2) Wellness 360 Healthcare at 210-458-4142; or (3) PEACE Center at 210-458-4077.

Campus Safety & Emergency Preparedness

UTSA is committed to providing a safe campus environment for students, faculty, staff, and visitors. As members of the community, we encourage you to take the following actions to be better prepared in case of an emergency:

- Alerts: Ensure you are signed up for UTSA Alerts through your myUTSA Account.
- Emergency Procedures: Read through the emergency response guide on the UTSA Alerts website.
- Safety App: Download the LiveSafe App on your phone through the Apple store or Google Play; visit the UTSA Alerts website for details.
- Important Numbers: UTSA Police - Emergency: (210) 458-4911; Non-Emergency: (210) 458-4242

Each one of us plays a critical role in making sure ALL ROADRUNNERS are safe, know what to do, and how to stay informed during a campus crisis. Don't be scared, be prepared! #UTSAprepared

Campus Recreation

Enrolled students are members of Campus Recreation and are eligible to use all of the services provided as well as participate in programming. Aquatics, Club Sports, Downtown Programs, Fitness and Wellness, Intramural Sports, and Outdoor Pursuits are all part of Campus Recreation.

The Roadrunner Creed

The University of Texas at San Antonio is a community of scholars, where integrity, excellence, inclusiveness, respect, collaboration, and innovation are fostered.

As a Roadrunner, I will:

- Uphold the highest standards of academic and personal integrity by practicing and expecting fair and ethical conduct;
- Respect and accept individual differences, recognizing the inherent dignity of each person;
- Contribute to campus life and the larger community through my active engagement; and
- Support the fearless exploration of dreams and ideas in the advancement of ingenuity, creativity, and discovery.

Guided by these principles now and forever, I am a Roadrunner!