MATH 6312: Syllabus
Spring 2019

Course Information

- Instructor: Jonghyun Yun

- Office Telephone Number: x29045

- Email Address: "j" dot "yun" at-sign "uta" dot "edu"

- Course Webpage: `http://wweb.uta.edu/faculty/yunj/course/math6312-sp19`

    - Download a course syllabus

- Office Hours: SEIR 218, Tu/Thu 4:30-5:30 pm or by appointment

- Section Information: MATH6312-001: Data Mining

- Time and Place of Class Meetings: Tu/Thu 5:30-6:50 pm, PKH 113

- Prerequisite: MATH 5305/STATS 5305 or MATH 5312/STATS 5312 or permission of instructor. Basic programming skills are preferred, but not required.

- Required Textbooks

    - Hastie, T., Tibshirani, R., and Friedman, J. H. (2008), The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd Edition. Springer.
    - Other Recommended Textbooks and Resources
        * Bishop, C. M. (2006). Pattern recognition and machine learning. Springer.
        * Rencher and Schaalje (2008). Linear Models in Statistics, 2nd Edition. Wiley.

Course Content

Data mining is the process of exploring and analyzing, by automatic or semiautomatic means, large quantities of observational data in order to discover meaningful patterns and models. By applying data mining techniques, data miners can fully exploit data patterns and behavior, gain insider understanding of the data, and produce new knowledge that decision-makers can act upon. Data mining has gained a substantial interest among practitioners in a variety of fields and industries. Nowadays, almost every organization collects data, which can be analyzed in order to support making better decisions, improving policies, discovering computer network intrusion patterns, designing new drugs, detecting credit fraud, making accurate medical diagnoses, etc.

The course covers selected portions of the textbook (ESL), with main focus on topics including but not limited to:

- Supervised learning

    - linear regression, logistic regression
    - model assessment, validation, and regularization
    - boosting; bagging and random forest
    - support vector machine, generative/discriminative classification algorithm

- neural network, probabilistic graphical model

- Unsupervised learning:

  - cluster analysis
  - principal component analysis, multidimensional scaling

The course will use R in order to demonstrate the implementation of data mining methods. In the first week, we will have a refresher on the commands in R which you will need to use in the following weeks, but this is not a comprehensive R course, and we will not go in depth on R syntax. Please see online R resources listed in the course website.

Student Learning Outcomes

Upon successful completion of the requirements for this course, students should have the knowledge and skills to:

1. understand a wide variety of data mining algorithms.

2. understand how to perform evaluation of learning algorithms and model selection.

3. be proficient with data mining software such as the R language.

4. understand how to create reproducible research documents using the R Markdown syntax.

5. conceptualise a data mining solution to a practical problem.

Grading

The usual grading scale will be used for this course.

A	B	C	D	F
100-90	89-80	79-70	69-60	below 60

Grade Distribution

- 40% Homework assignments
- 20% Take-home midterm exam
- 40% Final project

Document Formatting

All coursework (homework, midterm and final) must be turned in electronically, through Blackboard. All coursework will involve writing a combination of code and actual prose. You must submit your assignment in a format which allows for the combination of the two, and the automatic execution of all your code. The easiest way to do this is to use R Markdown. Exceptions may be made, with prior permission, for those who want to use other reproducible dynamic documents. Work should be submitted as R markdown **and** rendered PDF documents. Every files you submit should have a file name in the following format: `WorkType#.FirstName.LastName`.

Homework

The homework problems will be assigned weekly on the course webpage, and the problems will be supplemented with computer assignments requiring the use of statistical software (R). Statistical software instructions will be provided as needed.

There will be 7 homework assignments and a final homework assignment. One lowest homework score will be dropped, but you cannot drop the final homework. Late coursework will not be accepted without a university approved excuse. It is imperative that you show all your work; simply stating an answer will result in no credit for the problem. You are encouraged to work with each other on the homework problems, but you must turn in your own work.

Final project

Students will make in-class presentations of their final projects at the end of semester. Students also need to submit their final project reports. The instructions for the final project will be announced in class.

Digital Distractions

You are encouraged to bring your mobile devices to class. Mobile devices are to be used only for purposes directly related to the class content and information. You may not view or respond to emails during class. You may not conduct unrelated searches during class. Cellphones should be turned to vibrate while in class and on silent or off when students are presenting. No phone calls should be accepted, only in case of an emergency.

Lecture Attendance

At the University of Texas at Arlington, taking attendance is not required. Rather, each faculty member is free to develop his or her own methods of evaluating students' academic performance, which includes establishing course-specific policies on attendance. As the instructor of this section, I strongly encourage you to attend all lectures. You are responsible for any and all announcements made in class. You are also responsible for any material missed during class.

Drop Policy

Students may drop or swap (adding and dropping a class concurrently) classes through self-service in MyMav from the beginning of the registration period through the late registration period. After the late registration period, students must see their academic advisor to drop a class or withdraw. Undeclared students must see an advisor in the University Advising Center. Drops can continue through a point two-thirds of the way through the term or session. It is the student's responsibility to officially withdraw if they do not plan to attend after registering. Students will not be automatically dropped for non-attendance. Repayment of certain types of financial aid administered through the University may be required as the result of dropping classes or withdrawing. For more information, contact the Office of Financial Aid and Scholarships (`http://wweb.uta.edu/aao/fao/`).

Disability Accommodations

UT Arlington is on record as being committed to both the spirit and letter of all federal equal opportunity legislation, including The Americans with Disabilities Act (ADA), The Americans with Disabilities Amendments Act (ADAAA), and Section 504 of the Rehabilitation Act. All instructors at UT Arlington are required by law to provide "reasonable accommodations" to students with disabilities, so as not to discriminate on the basis of disability. Students are responsible for providing the instructor with official notification in the form of a letter certified by the Office for Students with Disabilities (OSD). Only those students who have officially documented a need for an accommodation will have their request honored. Students experiencing a range of conditions (Physical, Learning, Chronic Health,

Mental Health, and Sensory) that may cause diminished academic performance or other barriers to learning may seek services and/or accommodations by contacting: The Office for Students with Disabilities, (OSD) `http://www.uta.edu/disability` or calling 817-272-3364. Information regarding diagnostic criteria and policies for obtaining disability-based academic accommodations can be found at `http://www.uta.edu/disability`.

Counseling and Psychological Services (CAPS) `http://www.uta.edu/caps/` or calling 817-272-3671 is also available to all students to help increase their understanding of personal issues, address mental and behavioral health problems and make positive changes in their lives.

Non-Discrimination Policy

The University of Texas at Arlington does not discriminate on the basis of race, color, national origin, religion, age, gender, sexual orientation, disabilities, genetic information, and/or veteran status in its educational programs or activities it operates. For more information, visit `http://uta.edu/eos`.

Title IX Policy

The University of Texas at Arlington ("University") is committed to maintaining a learning and working environment that is free from discrimination based on sex in accordance with Title IX of the Higher Education Amendments of 1972 (Title IX), which prohibits discrimination on the basis of sex in educational programs or activities; Title VII of the Civil Rights Act of 1964 (Title VII), which prohibits sex discrimination in employment; and the Campus Sexual Violence Elimination Act (SaVE Act). Sexual misconduct is a form of sex discrimination and will not be tolerated. For information regarding Title IX, visit `http://www.uta.edu/titleIX` or contact Ms. Jean Hood, Vice President and Title IX Coordinator at (817) 272-7091 or jmhood@uta.edu.

Academic Integrity

Students enrolled in this course are expected to adhere to the UT Arlington Honor Code:

> I pledge, on my honor, to uphold UT Arlington's tradition of academic integrity, a tradition that values hard work and honest effort in the pursuit of academic excellence.

> I promise that I will submit only work that I personally create or contribute to group collaborations, and I will appropriately reference any work from other sources. I will follow the highest standards of integrity and uphold the spirit of the Honor Code.

UT Arlington faculty members may employ the Honor Code as they see fit in their courses, including (but not limited to) having students acknowledge the honor code as part of an examination or requiring students to incorporate the honor code into any work submitted. Per UT System Regents' Rule 50101, §2.2, suspected violations of university's standards for academic integrity (including the Honor Code) will be referred to the Office of Student Conduct. Violators will be disciplined in accordance with University policy, which may result in the student's suspension or expulsion from the University.

Electronic Communication

UT Arlington has adopted MavMail as its official means to communicate with students about important deadlines and events, as well as to transact university-related business regarding financial aid, tuition, grades, graduation, etc. All students are assigned a MavMail account and are responsible for checking the inbox regularly. There is no additional charge to students for using this account, which remains active even after graduation. Information about activating and using MavMail is available at `http://www.uta.edu/oit/cs/email/mavmail.php`.

Campus Carry

Effective August 1, 2016, the Campus Carry law (Senate Bill 11) allows those licensed individuals to carry a concealed handgun in buildings on public university campuses, except in locations the University establishes as prohibited. Under the new law, openly carrying handguns is not allowed on college campuses. For more information, visit `http://www.uta.edu/news/info/campus-carry/`

Student Feedback Survey

At the end of each term, students enrolled in classes categorized as "lecture," "seminar," or "laboratory" shall be directed to complete an online Student Feedback Survey (SFS). Instructions on how to access the SFS for this course will be sent directly to each student through MavMail approximately 10 days before the end of the term. Each student's feedback enters the SFS database anonymously and is aggregated with that of other students enrolled in the course. UT Arlington's effort to solicit, gather, tabulate, and publish student feedback is required by state law; students are strongly urged to participate. For more information, visit `http://www.uta.edu/sfs`.

Final Review Week

A period of five class days prior to the first day of final examinations in the long sessions shall be designated as Final Review Week. The purpose of this week is to allow students sufficient time to prepare for final examinations. During this week, there shall be no scheduled activities such as required field trips or performances; and no instructor shall assign any themes, research problems or exercises of similar scope that have a completion date during or following this week unless specified in the class syllabus. During Final Review Week, an instructor shall not give any examinations constituting 10% or more of the final grade, except makeup tests and laboratory examinations. In addition, no instructor shall give any portion of the final examination during Final Review Week. During this week, classes are held as scheduled. In addition, instructors are not required to limit content to topics that have been previously covered; they may introduce new concepts as appropriate.

Emergency Exit Procedures

Should we experience an emergency event that requires us to vacate the building, students should exit the room and take an immediate right or left, and walk toward the corner of the building. When exiting the building during an emergency, one should never take an elevator but should use the stairwells. Faculty members and instructional staff will assist students in selecting the safest route for evacuation and will make arrangements to assist individuals with disabilities. (`https://www.uta.edu/policy/procedure/7-6`).
Students are encouraged to subscribe to the MavAlert system that will send information in case of an emergency to their cell phones or email accounts. Anyone can subscribe at `https://mavalert.uta.edu/` or `https://mavalert.uta.edu/register.php`

Student Support Services

UT Arlington provides a variety of resources and programs designed to help students develop academic skills, deal with personal situations, and better understand concepts and information related to their courses. Resources include tutoring, major-based learning centers, developmental education, advising and mentoring, personal counseling, and federally funded programs. For individualized referrals, students may visit the reception desk at University College (Ransom Hall), call the Maverick Resource Hotline at (817) 272-6107, send a message to resources@uta.edu, or view the information at http://www.uta.edu/resources.

Emergency Phone Numbers

In case of an on-campus emergency, call the UT Arlington Police Department at 817-272-3003 (non-campus phone), 2-3003 (campus phone). You may also dial 911. Non-emergency number 817-272-3381

Active Shooter/Threat Resources

The safety and security of our campus is the responsibility of everyone in our community. Each of us has an obligation to be prepared to appropriately respond to threats to our campus, such as an active aggressor. The information for active threat and other emergency situations can be found through the links below:

- `http://police.uta.edu/activeshooter`

- `http://police.uta.edu/em`