INSY 5339: PRINCIPLES OF BUSINESS DATA MINING

Fall 2020

Instructor: Jayarajan Samuel
Office Number: Not Applicable
Email Address: jayarajan.samuel@uta.edu
Office Hours: By email appointment only.
Section Information: INSY 5339 / Section 001
Time and Place of Class Meetings: Wednesday 7:00 – 9:50 p.m. via Canvas Conferences tool

Course Description:
This course will cover the foundations of business data mining. It will examine tools and techniques from the fields of machine learning and statistics used in practical data mining for finding, and describing, structural patterns in data. Topics may include: Knowledge representation and different types of data; Techniques for data pre-processing, cleaning, reduction, transformation, and visualization; Methods for Classification, Clustering, and Association Rules, including Decision Trees, Rules, Naive Bayes, k Nearest Neighbor, Neural Networks, Regression (linear & logistic), A-Priori, K-means, and hierarchical and density-based clustering; Performance evaluation of data mining algorithms using metrics. This course uses real world data sets and widely used statistical packages and programming languages. Prerequisite: BSTAT 5325 or equivalent. May be taken concurrently.

Student Learning Outcomes:
Students will be able to pre-process real world datasets and implement supervised and un-supervised data mining algorithms in SAS enterprise miner to solve business problems.

Prerequisite: BSTAT 5325 or equivalent. May be taken concurrently.

Required Textbooks and Other Course Materials:
Data Mining for Business Analytics. Concepts, Techniques and Applications in R
ISBN-10: 1118879368

Other Materials/Resources/Readings:
Notes will be posted in Canvas

Software: SAS Enterprise Miner 14.2 or later
Students are responsible to obtain individual Desktop SAS license.
Student licenses can be purchases from the following website:
https://ut.service-now.com/utss/KAhome.do?number=KB0013524

Software: Tableau
Students are responsible to obtain individual Tableau student licenses from the following site:
https://www.tableau.com/academic/students
Use your UTA id to register.

**Description of major assignments and examinations:**
The distribution of points will be as follows:

<table>
<thead>
<tr>
<th>Assignment</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homework Assignments</td>
<td>20%</td>
</tr>
<tr>
<td>Peer Evaluations</td>
<td>5%</td>
</tr>
<tr>
<td>MidTerm Exam</td>
<td>25%</td>
</tr>
<tr>
<td>Project</td>
<td>25%</td>
</tr>
<tr>
<td>Final Exam</td>
<td>25%</td>
</tr>
</tbody>
</table>

**Grading:** The following criteria will be used to assess your grade:
A (>=90%), B (>=80%), C (>=70%), D (>=60%), F (<60%)

**Group Assignments:** All homework assignments and the final project will be done in groups. Each group will consist of 4 members and will be chosen by the students in the beginning of the semester. All group members are expected to fully participate in the assignments/project and an end of semester, written, peer evaluation will determine 5% of the overall grade. Peer evaluation means each member of the team will evaluate themselves and each of the other members of the team on a 1 to 10 scale on various aspects of their contribution to the team.

**Final Project:** The final project work needs to be started by the group in mid semester and it will have multiple deliverables. The team is expected to make a 10 minute presentation at the end of semester (all group members need to speak at the presentation). Details will be provided in class and Canvas.

**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 will consider attendance mandatory for all lectures. If you miss a class, you are responsible for the materials covered. All lectures delivered via Canvas Conferences will be recorded. It is expected that students who miss the online lecture would view the recordings at a later time.

**Exams:** You are responsible for everything that is covered in the classroom, including additional materials that the instructor may discuss in class. There are no make-up exams. Under extenuating circumstances (e.g., medical emergency, family emergency, work-related travel, etc. with documented proof), other arrangements may be considered. You can only use this excuse for one exam. The Mid Term Exam will be conducted at UTA campus. The final exam will be administered via Canvas Quizzes with Lockdown Browser and camera. It is the student’s responsibility to be equipped to comply to all Canvas Quizzes requirements (Hardware, Software, Internet connection, etc.).

**Homework Assignments:** You will have 4 or 5 homework exercises. You will have about a week to complete the exercises. Homework must be turned in electronically (via Canvas) by the due date and time specified by the instructor. Late submissions will receive a score of 0 (no exceptions).

**Expectations for Out-of-Class Study:** Beyond the time required to attend each class meeting, students enrolled in this course should expect to spend at least an additional 9 hours per week of their
own time in course-related activities, including reading required materials, completing assignments & projects, and preparing for exams/quizzes.

**Tentative Course Schedule**

<table>
<thead>
<tr>
<th>Week</th>
<th>Topics</th>
<th>Readings</th>
<th>Assignments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Week1 Aug. 26</td>
<td>Introduction, Syllabus,</td>
<td>Syllabus; Chap. 1, 2</td>
<td></td>
</tr>
<tr>
<td>Week2 Sep. 2</td>
<td>Data Visualization</td>
<td>Chap. 3</td>
<td>Homework1 Due (Tableau)</td>
</tr>
<tr>
<td>Week3 Sep. 9</td>
<td>Dimension Reduction</td>
<td>Chap. 4</td>
<td></td>
</tr>
<tr>
<td>Week4 Sep. 16</td>
<td>PCA SAS Enterprise Miner; Linear Regression</td>
<td>Chapter 6 SAS notes</td>
<td>Homework2 Due (SAS EM tutorial)</td>
</tr>
<tr>
<td>Week5 Sep. 23</td>
<td>Linear Regression; SAS Implementation</td>
<td>Chapter 6 SAS notes</td>
<td></td>
</tr>
<tr>
<td>Week6 Sep. 30</td>
<td>Logistic Regression;</td>
<td>Chapter 10</td>
<td>Homework3 Due (PCA &amp; Linear Regression)</td>
</tr>
<tr>
<td>Week7 Oct. 7</td>
<td>Classification and Regression Trees</td>
<td>Chapter 9</td>
<td></td>
</tr>
<tr>
<td>Week8 Oct. 14</td>
<td>Mid-Term Exam</td>
<td></td>
<td>Mid Term exam on Oct 14th at UTA campus</td>
</tr>
<tr>
<td>Week9 Oct. 21</td>
<td>Classification and Regression Trees</td>
<td>Chapter 9,7</td>
<td>Project Proposal Due</td>
</tr>
<tr>
<td>Week10 Oct. 28</td>
<td>Naïve Bayes Classifier; Association Rules and Collaborative Filtering; Implementation in SAS</td>
<td>Chapter 8,14</td>
<td>Homework4 Due (Decision Trees)</td>
</tr>
<tr>
<td>Week11 Nov. 4</td>
<td>Association Rules and Collaborative Filtering; Implementation in SAS; Cluster Analysis</td>
<td>Chapter 14,15</td>
<td>Project Progress Report Due</td>
</tr>
<tr>
<td>Week12 Nov. 11</td>
<td>Cluster Analysis and Implementation in SAS; Project Presentations</td>
<td>Chapter 15</td>
<td>Homework5 Due (Association rules)</td>
</tr>
<tr>
<td>Week13 Nov. 18</td>
<td>Evaluating Predictive Performance; Final Exam review</td>
<td>Chapter 5</td>
<td></td>
</tr>
<tr>
<td>Nov 25 to 27</td>
<td>Thanksgiving – No Classes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week14 Dec 2</td>
<td>Project Presentations</td>
<td></td>
<td>Project Final Reports Due</td>
</tr>
<tr>
<td>Dec. 16</td>
<td>Comprehensive Final Exam on December 16, between 8:15 and 10:45, 60 to 90min exam via Canvas Quizzes with Lockdown Browser and camera</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**NOTE:** This is a tentative schedule and may be changed to serve the educational needs of the students enrolled in the course. Students are responsible to be aware of changes announced in class and/or via Canvas. – Jayarajan Samuel

Kindly check the Academic Calendar for important dates.
UNIVERSITY and COLLEGE POLICIES

Grade Grievances: Any appeal of a grade in this course must follow the procedures and deadlines for grade-related grievances as published in the current University Catalog. Policy details may be obtained from:
http://catalog.uta.edu/academicregulations/grades/#graduatetext.

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 www.uta.edu/disability.

Only those students who have officially documented a need for an accommodation will have their request honored. Information regarding diagnostic criteria and policies for obtaining disability-based academic accommodations can be found at www.uta.edu/disability or by calling the Office for Students with Disabilities at (817) 272-3364.

Counseling and Psychological Services (CAPS) 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 uta.edu/eos.
Title IX: 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 www.uta.edu/titleIX or contact Ms. Michelle Willbanks, Title IX Coordinator at (817) 272-4585 or titleix@uta.edu

Academic Integrity: Students enrolled all UT Arlington courses 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 in their courses by 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. Additional information is available at https://www.uta.edu/conduct/. Faculty are encouraged to discuss plagiarism and share the following library tutorials http://libguides.uta.edu/copyright/plagiarism and http://library.uta.edu/plagiarism/

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 move toward the nearest exit. 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.

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/universitycollege/resources/index.php.

University Tutorial & Supplemental Instruction (Ransom Hall 205): UTSI offers a variety of academic support services for undergraduate students, including: 60 minute one-on-one tutoring sessions, Start Strong Freshman tutoring program, and Supplemental Instruction. Office hours are Monday-Friday 8:00am-5:00pm. For more information visit www.uta.edu/utsi or call 817-272-2617.

The IDEAS Center (2nd Floor of Central Library) offers FREE tutoring to all students with a focus on transfer students, sophomores, veterans and others undergoing a transition to UT Arlington. Students can drop in, or check the schedule of available peer tutors at www.uta.edu/IDEAS, or call (817) 272-6593.

The English Writing Center (411LIBR): The Writing Center offers FREE tutoring in 15-, 30-, 45-, and 60-minute face-to-face and online sessions to all UTA students on any phase of their UTA coursework. Register and make appointments online at https://uta.mywconline.com. Classroom visits, workshops, and specialized services for
graduate students and faculty are also available. Please see www.uta.edu/owls for detailed
information on all our programs and services.

The Library’s 2nd floor Academic Plaza offers students a central hub of support services,
including IDEAS Center, University Advising Services, Transfer UTA and various
college/school advising hours. Services are available during the library’s hours of operation.
http://library.uta.edu/academic-plaza

You may also find the following information useful:
  Library Home Page .............................. http://www.uta.edu/library
  Subject Guides................................. http://libguides.uta.edu
  Subject Librarians ............................. http://www.uta.edu/library/help/subject-librarians.php
  Course Reserves .............................. http://pulse.uta.edu/vwebv/enterCourseReserve.do
  Connecting from Off- Campus............... http://libguides.uta.edu/offcampus
  Ask A Librarian............................... http://ask.uta.edu

**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