INSY 5339-030: PRINCIPLES OF BUSINESS DATA MINING

Cohort 3 2020

Instructor Information

Instructor: Jayarajan Samuel
Office: Predetermined at Ericsson Plano Campus or online
Email Address: jayarajan.samuel@uta.edu
Office Hours: By appointment

Course Information

Section Information: INSY 5339-030 Ericsson Cohort 3
Time and Place of Class Meetings:
- Each meeting will be online for 3 hours: Sept 11-12; 18-19; 25-26; Oct 2-3; 9-10; 16-17;
- Timings of online classes: Fridays 3PM to 6PM and Saturdays 9AM to 12NOON
- Additional 4 hours of class time to be scheduled between instructor and students as meeting times at Ericsson.

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.

Students are responsible for obtaining and installing Python 3 and Jupyter Notebook on their laptops.

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

Required Textbooks and Other Course Materials:
Data Mining for Business Analytics: Concepts, Techniques and Applications in Python

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

Software: Python 3 and Jupyter Notebook
Recommended to use software from https://www.anaconda.com/

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

<table>
<thead>
<tr>
<th>Assignment</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homework Assignments</td>
<td>20%</td>
</tr>
<tr>
<td>Class Participation</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 3 to 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.

**Final Project:** The final project work needs to be started by the group early in the program and it will have multiple deliverables. The team is expected to make at least a 30 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 but attendance is a critical indicator of student success. 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. However, while UT Arlington does not require instructors to take attendance in their courses, the U.S. Department of Education requires that the University have a mechanism in place to mark when Federal Student Aid recipients “begin attendance in a course.” UT Arlington instructors will report when students begin attendance in a course as part of the final grading process. Specifically, when assigning a student a grade of F, faculty report must the last date a student attended their class based on evidence such as a test, participation in a class project or presentation, or an engagement online via Canvas. This date is reported to the Department of Education for federal financial aid recipients.

**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, etc.), other arrangements may be considered.

**Homework Assignments:** You will have 4 or 5 homework exercises. Homework must be turned in electronically (via Canvas/email) by the due date and time specified by the instructor. Late submissions will receive a score of 0 unless discussed ahead of time with the instructor for a reduced grade.

**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 8 to 16 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>Sep 11</td>
<td>Introduction, Syllabus, Data Visualization, Dimension Reduction, PCA</td>
<td>Syllabus; Chap. 1, 2, 3</td>
<td></td>
</tr>
<tr>
<td>Sep 12</td>
<td>Dimension Reduction, PCA, Python Implementation</td>
<td>Chap 3, 4</td>
<td></td>
</tr>
<tr>
<td>Sep 18</td>
<td>Linear Regression, Logistic Regression</td>
<td>Chap 6, 10</td>
<td>Homework 1 Due: PCA Analysis</td>
</tr>
<tr>
<td>Sep 19</td>
<td>Linear Regression, Python Implementation, Logistic Regression, Python Implementation</td>
<td>Chap 6, 10</td>
<td>Project Proposal Due</td>
</tr>
<tr>
<td>Sep 25</td>
<td>Classification &amp; Regression Trees</td>
<td>Chap 9</td>
<td>Homework 2 Due: Linear Regression</td>
</tr>
<tr>
<td>Sep 26</td>
<td>Classification &amp; Regression Trees, Python Implementation</td>
<td>Chap 9</td>
<td></td>
</tr>
<tr>
<td>Oct 2</td>
<td>Mid Term Exam, K-Nearest Neighbor, Naïve Bayes Classifier</td>
<td>Chap 7, 8</td>
<td>Homework 3 Due: Logistic Regression and Classification Trees</td>
</tr>
<tr>
<td>Oct 3</td>
<td>K-Nearest Neighbor &amp; Naïve Bayes Classifier Python Implementation Project Presentation Guidance</td>
<td>Chap 7, 8</td>
<td>Project Progress Report Due</td>
</tr>
<tr>
<td>Oct 9</td>
<td>Association Rules mining, Cluster Analysis;</td>
<td>Chap 14, 15</td>
<td>Homework 4 due: K-Nearest Neighbors; Validation</td>
</tr>
<tr>
<td>Oct 10</td>
<td>Association Rules mining &amp; Cluster Analysis Python Implementation Project Discussion</td>
<td>Chap 14, 15</td>
<td></td>
</tr>
<tr>
<td>Oct 16</td>
<td>Evaluating Predictive Performance; Final Exam Review; Project Presentations</td>
<td>Chap 5</td>
<td>Final Project Report, Slides Due</td>
</tr>
<tr>
<td>Oct 17</td>
<td>Final Exam; Project Presentations</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](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/ao/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 Success Programs:** 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 by appointment, drop-in tutoring, etutoring, supplemental instruction, mentoring (time management, study skills, etc.), success coaching, TRIO Student Support Services, and student success workshops. For additional information, please email resources@uta.edu, or view the Maverick Resources website.

**Research or General Library Help**

Ask for Help
- Academic Plaza Consultation Services (library.uta.edu/academic-plaza)
- Ask Us (ask.uta.edu/)
- Research Coaches (http://libguides.uta.edu/researchcoach)

Resources
- Library Tutorials (library.uta.edu/how-to)
- Subject and Course Research Guides (libguides.uta.edu)
- Librarians by Subject (library.uta.edu/subject-librarians)
- A to Z List of Library Databases (libguides.uta.edu/az.php)
- Course Reserves (https://uta.summon.serialssolutions.com/#!/course_reserves)
- Study Room Reservations (openroom.uta.edu/)

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
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