**UTRGV Logo**

| **CSCI 6366.90L Data Mining and Warehousing** | **Dr. Zhixiang Chen** |
| --- | --- |
| Spring 2021 | Phone: 665-285, Email zhixiang.chen@utrgv.edu |
| Class time:  ONLINE - Synchronous | Office: EIEAB 3.241  Office Hours: ONLINE (by Appointment) |

## Textbook and/or Resource Material

**Required Textbook**

The textbook is "Introduction to Data Mining", Pang-Ning Tan, Michael Steinbach, Anuj Karpatne, Vipin Kumar, 2nd (or latest) edition, Pearson/Addison Wesley, ISBN-13: 9780133128901, ISBN-10: 0133128903. eBook PDF $29.99). Not an open education resource. Copyright © 2019 Pearson Education Inc.

Other suggested materials:

Will be given in class as the semester progresses

Other materials such as lecture notes, handouts, micro lectures, recorded lectures and meetings, etc., will be posted to Blackboard. Please make sure that you assess Blackboard often for course materials and resources.

## Course Description and Prerequisites

Here are the course descriptions from UTRGV Catalog:

CSCI 6366 Data Mining and Warehousing

As a multidisciplinary field, draws on work from areas including database technology, artificial intelligence, machine learning, neural network, statistics, information retrieval, and data visualization. Theoretical and practical methods will be presented on knowledge discovery and systems design and implementation.

## Course Topics

CSCI/CMPE 6366 teaches you fundamental methodologies and practices about data mining and warehousing. It covers the following topics:

* Introduction: Data Mining Tasks, Data, Measures of Similarity and Dissimilarity
* Classification: Basic Concepts and Techniques, Alternative Techniques
* Association Analysis: Basic Concepts and Algorithms, Advanced Concepts
* Cluster Analysis: Basic Concepts and Algorithms, Additional Issues and Algorithms
* Anomaly Detection, Avoiding False Discoveries

## Course Overview

**Quizzes/Exercises.** There will be quizzes/exercises for every module of this course. There is no make-up if you miss one. These are composed of simple questions/calculation/exercises to provide formative assessment about to what extent students are able to attain the learning objectives, which are listed later.

**Homework Assignments.** Students are expected to have assignments completed before or on the due date. This may include selected readings from the textbook. Homework assignments consist of theoretical and practical problems to provide summative assessments about to what extent students are able to attain the learning objectives, which are listed later.

**Term Paper.** Students are expected to complete a term paper in one of the given areas to demonstrate competency in enabling data mining technologies as part of larger systems.

**Exams.** The material in this course is naturally cumulative, with topics of each week building on all the prior material. Therefore, each exam will focus on the material covered from the beginning or since the previous exam, however, the student is expected to understand and apply all previous course material. The exams provide summative assessment about to what extent students are able to

attain the learning objectives, which are listed later.

**Scoring and Grading.**

Blackboard test/assignment tools will be used to assist grading. Rubrics will be used to grade assignments, projects and exams. Scoring and letter grades determination are summarized in the following.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Grade Breakdown** |  |  | **Total Score** | **Final Grade** |
| **Quizzes/Exercises (5)** | 10% |  | 90%-100% | A |
| **Assignments (4)** | 20% |  | 80%-89% | B |
| **Term paper (1)** | 35% |  | 70%-79% | C |
| **Exams (2)** | 35% |  | 60%-69% | D |
| **Total possible score (max):** | 100% |  | 0%-59% | F |

Note: Grades may be curved to reflect the overall performance of the class.

**Course Schedule**. Below is a rough course schedule to give you an idea of topics and pacing. The actual course schedule is likely to change and will be kept up to date on the course website.

* Module 1: Introduction (1 Week)
* Module 2: Classification (4 Weeks)
* Module 3: Association Analysis (4 Weeks)
* Module 4: Cluster Analysis (3 Weeks)
* Module 5: Anomaly Detection, Avoiding False Discoveries (3 Week)

**Late Work Policy.** Assignments must be turned in at the specified time on the given due date. Afterwards, the penalties are as follows:

* Within 24 hours late will lose 10%. Within 48 hours late will lose 20%.
* More than 48 hours late will receive 50% credit.

**Make-up Policy.** No make-up exams will be given except for university sanctioned excused absences. If you need to miss an exam, it is your responsibility to contact me before the exam, or as soon after the exam as possible. Missing an exam without an approved (by the university or me) excuse will result in a zero.

## Teaching Philosophy

I would like to help my students become self-motivated learners who not only master the content of the course but also are able to apply the learned knowledge to practice. I expect my students to gain fruitful attainment in areas of critical thinking and technical problem solving through completing integrated components of coursework.

## Student Learning Outcomes

Upon completion of the course the student will be able to:

* **SLO A:** Demonstrate competency in primary data mining techniques
* **SLO B:** Demonstrate competency in enabling data mining technologies as part of larger systems.

## Course Learning Objectives

Upon successfully completing the course each student should be able to:

* understand algorithms and methods of data mining
* develop data mining programs and applications
* program using available data mining tools and general-purpose languages
* understand analysis, metrics, visualization and navigation of data mining results
* learn how to use a few commercial data mining tools
* know basic techniques for both directed and undirected knowledge discovery.
* know and use software package techniques for mining.
* Have a good understanding of data mining techniques: association rules, clustering, anomaly detection, etc.
* design data schemas for a warehouse environment.

## Learning Objectives for Core Curriculum Requirements

This course is not a part of the University Core Curriculum.

## Changes in the course content

Changes in the course content and topics may be made by the instructor as deemed necessary.

## Blackboard Orientation

To actively participate in this course, one needs to get familiar with the course environment. This course uses Blackboard tools. One can get familiar with these tools by self-enrolling and reviewing the orientation provided:

* In Blackboard, use the **Course Search** box and enter Blackboard Orientation
* Click on Go
* Click on the **Enroll** button to include the orientation course in your Blackboard MyCourses list

## Contacting Instructor and Instructor's Response Time

Use instructor's email given above for contacting him regarding the course content. Normally, the instructor will respond within two working weekdays. Any exceptions will be announced through Blackboard announcement mechanism. Technical questions on Blackboard should be referred to Blackboard Support (see under Course Technology).

Grades are normally updated each time a grading session has been completed – typically within two working weekdays or before the next class session.

## Course Delivery Method due to COVID-19 Situation

Due to the current COVID-19 pandemic, the delivery of the course, and interaction between students and instructor, will be via **online** tools. Below is a summary:

1. **Class notes, recordings, and attending classes online:** The class-notes, recordings, and other material will be available **online** at Blackboard. Students will be required to complete any assigned work before attending the class online. Attending the online classes is mandatory. Blackboard Collaborate Ultra needs to be used to attend the class online. Below are links to Blackboard Student Support:

* [Student Preparation Checklist](https://www.utrgv.edu/online/academic-continuity/students/index.htm)
* [How-To Guide and Tutorials](https://www.utrgv.edu/online/academic-continuity/resources/index.htm)

1. **Syllabus, grading policy, and exam dates/times:** Syllabus has the detail on grading policy and exam dates/times. There cannot be an exception. Of course, the exam dates/times may be changed due to unforeseen situation at the university level, which will be communicated to you as soon as the university announces any changes.
2. **Projects, Assignments, Quizzes and Exams:** Assignments, Quizzes and Exams are required to be submitted online through Blackboard. There will be quizzes and exams online during the class time. Note that a quiz or an exam will be conducted synchronously during an online class for everyone at the same time. Students may not take a quiz or an exam at different times.
3. **Online tutoring:** Teaching Assistants will be available through **online**. The ONLINE Tutoring Schedule (along with instructions of how to connect with a tutor online) will be available on Blackboard under **Reference Material** folder.
4. **Online office hours:** We will be meeting online at the scheduled class time, and I will be answering your questions about the recorded lectures/notes and conducting quizzes/exams. Additionally, you may request separate online appointment through email. Students may not visit the instructor or a Teaching Assistant in person at this time.
5. **Visiting on-campus computer labs:** Contact Help Desk (956-882-2020) to know which labs are available for you to use, and any restrictions associated with your visit. If any students respond YES to any of the four (4) items listed below, they should not visit campus and be present in a social-distanced space with other individuals:

* Have you felt like you have a fever in past 48 hours?
* Do you have a cough?
* Have you been to Europe or Asia in the past 14 days?
* Have you been in personal contact with someone who has known or suspected coronavirus (COVID-19) infection?

## CALENDAR OF ACTIVITIES

Refer to Section “**Course Organization**” above for subject topics planned for each week.

The UTRGV academic calendar can be found at [My UTRGV](https://my.utrgv.edu/home) at the bottom of the screen, prior to login. Some important dates for Spring 2021 include:

Jan. 11 (Mon.) Spring classes begin

Jan. 17 (Sun.) Last day to add or register for Spring classes

Jan. 15 (Fri.) Last day to withdraw (drop all classes) and receive an 80% refund

Jan. 18 (Mon.) Martin Luther King Jr. Holiday. No classes.

Jan. 25 (Mon.) Last day to withdraw (drop all classes) and receive a 70% refund

Jan. 27 (Wed.) Census Day (last day to drop without it appearing on the transcript)

Feb. 1 (Mon.) Last day to withdraw (drop all classes) and receive a 50% refund

Feb. 8 (Mon.) Last day to withdraw (drop all Spring classes) and receive a 25% refund

Mar. 15-20 (Mon.-Sat) Spring Break. No classes.

Apr. 2-3 (Fri.-Sat) Easter Holiday. No classes.

Apr. 6 (Tue.) Last day to drop a class (grade of DR) or withdraw (grade of W)

Apr. 29 (Thurs.) Study Day. No classes.

Apr 30-May 6 (Fri.-Thurs.) Final Exams

May 6 (Thurs.) Spring classes end; Official last day of the term

May 7-8 (Fri.-Sat.) Commencement Exercises

May 10 (Mon.) Grades Due at 3 p.m.

## Additional Course Information

The lab work required in the course is integrated with the course itself in the form of assignments/projects. Besides the formal allotted time, a student is expected to spend a large number of hours working on projects, assignments, reading material, recorded lectures, and videos. The general rule is that for every 1 hour in class a student is expected to put in 3 or more hours outside the class. You may only do work related with the class when in class. Activities unrelated with the class (such as texting, talking, browsing, unless otherwise explicitly instructed by the instructor to do so) are prohibited during a class session.

## UTRGV Policy Statements

**STUDENTS WITH DISABILITIES:**

Students with a documented disability (physical, psychological, learning, or other disability which affects academic performance) who would like to receive academic accommodations should contact **Student Accessibility Services (SAS)** as soon as possible to schedule an appointment to initiate services.  Accommodations can be arranged through SAS at any time but are not retroactive.  Students who experience a broken bone, severe injury, or undergo surgery during the semester are eligible for temporary services.

**Pregnancy, Pregnancy-related, and Parenting Accommodations:**

Title IX of the Education Amendments of 1972 prohibits sex discrimination, which includes discrimination based on pregnancy, marital status, or parental status. Students seeking accommodations related to pregnancy, pregnancy-related condition, or parenting (reasonably immediate postpartum period) are encouraged to contact Student Accessibility Services for additional information and to request accommodations.

**Student Accessibility Services:**

**Brownsville Campus**: Student Accessibility Services is located in 1.107 in the Music and Learning Center building (BMSLC) and can be contacted by phone at (956) 882-7374 or via email at [ability@utrgv.edu](mailto:ability@utrgv.edu).

**Edinburg Campus:** Student Accessibility Services is located in 108 University Center (EUCTR) and can be contacted by phone at (956) 665-7005 or via email at [ability@utrgv.edu](mailto:ability@utrgv.edu).

**MANDATORY COURSE EVALUATION PERIOD:**

Students are required to complete an ONLINE evaluation of this course, accessed through your UTRGV account (<http://my.utrgv.edu>); you will be contacted through email with further instructions. Students who complete their evaluations will have priority access to their grades. Online evaluations will be available on or about:

Module 1 October 2nd – 8th

Module 2 November 27th – December 3rd

Full Fall Semester November 14th – December 4th

**SEXUAL MISCONDUCT and MANDATORY REPORTING:**

In accordance with UT System regulations, your instructor is a “Responsible Employee” for reporting purposes under Title IX regulations and so must report to the Office of Institutional Equity & Diversity ([oie@utrgv.edu](mailto:oie@utrgv.edu)) any instance, occurring during a student’s time in college, of sexual misconduct, which includes sexual assault, stalking, dating violence, domestic violence, and sexual harassment, about which she/he becomes aware during this course through writing, discussion, or personal disclosure. More information can be found at [www.utrgv.edu/equity](https://nam01.safelinks.protection.outlook.com/?url=http%3A%2F%2Fwww.utrgv.edu%2Fequity&data=02%7C01%7Cmarissa.alaniz%40utrgv.edu%7Ccdba6e51db19402b2a0f08d7062784d4%7C990436a687df491c91249afa91f88827%7C0%7C0%7C636984639122459545&sdata=5hndVru5iLHMQARX2XhtCu41F8pbn0A6sKxkK1lXYss%3D&reserved=0), including confidential resources available on campus. The faculty and staff of UTRGV actively strive to provide a learning, working, and living environment that promotes personal integrity, civility, and mutual respect that is free from sexual misconduct, discrimination, and all forms of violence. If students, faculty, or staff would like confidential assistance, or have questions, they can contact OVAVP (Office for Victim Advocacy & Violence Prevention) at 665-8287, 882-8282, or [OVAVP@utrgv.edu](mailto:OVAVP@utrgv.edu).

**COURSE DROPS:**

According to UTRGV policy, students may drop any class without penalty earning a grade of DR until the official drop date. Following that date, students must be assigned a letter grade and can no longer drop the class. Students considering dropping the class should be aware of the “3-peat rule” and the “6-drop” rule so they can recognize how dropped classes may affect their academic success. The 6-drop rule refers to Texas law that dictates that undergraduate students may not drop more than six courses during their undergraduate career. Courses dropped at other Texas public higher education institutions will count toward the six-course drop limit. The 3-peat rule refers to additional fees charged to students who take the same class for the third time.

**STUDENT SERVICES:**

Students who demonstrate financial need have a variety of options when it comes to paying for college costs, such as scholarships, grants, loans and work-study. Students should visit the Students Services Center (U Central) for additional information. U Central is located in BMAIN 1.100 (Brownsville) or ESSBL 1.145 (Edinburg) or can be reached by email ([ucentral@utrgv.edu](mailto:ucentral@utrgv.edu)) or telephone: (888) 882-4026. In addition to financial aid, U Central can assist students with registration and admissions.

Students seeking academic help in their studies can use university resources in addition to an instructor’s office hours. University Resources include the Advising Center, Career Center, Counseling Center, Learning Center, and Writing Center. The centers provide services such as tutoring, writing help, critical thinking, study skills, degree planning, and student employment. Locations are:

| **Center Name** | **Brownsville Campus** | **Edinburg Campus** |
| --- | --- | --- |
| **Advising Center**  [AcademicAdvising@utrgv.edu](mailto:AcademicAdvising@utrgv.edu) | BMAIN 1.400  (956) 665-7120 | ESWKH 101  (956) 665-7120 |
| **Career Center**  [CareerCenter@utrgv.edu](mailto:CareerCenter@utrgv.edu) | BCRTZ 129  (956) 882-5627 | ESSBL 2.101  (956) 665-2243 |
| **Counseling Center**  [Counseling@utrgv.edu](mailto:Counseling@utrgv.edu) | EUCTR 109  (956) 665-2574 | BSTUN 2.10  (956) 882-3897 |
| **Learning Center**  [LearningCenter@utrgv.edu](mailto:LearningCenter@utrgv.edu) | BMSLC 2.118  (956) 882-8208 | ELCTR 100  (956) 665-2585 |
| **Writing Center**  [WC@utrgv.edu](mailto:WC@utrgv.edu) | BUBLB 3.206  (956) 882-7065 | ESTAC 3.119  (956) 665-2538 |