

## **CIS 5354-110, CRN 10419, ADVANCED METHODS IN BIG DATA ANALYTICS**

Texas A&M University-Central Texas

### **INSTRUCTOR AND CONTACT INFORMATION**

**Instructor:** Dr. Ferdi Eruysal

**Office:** FH 3230

**Phone:** 254-519-5440

**Preferred Classroom Communication Method:** Canvas inbox

**Email:** ferdi.eruysal@tamuct.edu

#### **Office Hours:**

Monday, Wednesday 14.00-16.00.

Other days and times by appointment. I am available to see drop by students', but this semester, I would prefer that students take appointment before they come to see me.

#### **Mode of instruction and course access:**

This course meets face-to-face on Saturdays 01/27, 02/10, 03/03, 03/24, 04/07 AND 04/28 , 13:00–17:00 at Founder's Hall 307. It also uses the A&M-Central Texas Canvas Learning Management System [<https://tamuct.instructure.com>] which will be our primary resource for the class information. You will use the canvas username and password communicated to you separately to logon to this system. The course syllabus, schedule, supplemental readings, class announcements, power point slides, learning modules and other course related documents will be posted on canvas. Each student is responsible for the posted material, and should check canvas at least daily for updates. The course outline in this syllabus shows the basic schedule for the semester.

#### **Student-instructor interaction:**

Please send all course related correspondence through canvas. Please use TAMUCT email only when canvas is not available (or for non-course related correspondence). I check email several times a day during the week and at least once during the weekends. I will attempt to respond within 24 hours and within 48 hours on Saturday through Sunday. Please do not hesitate to stop by my office during the scheduled office hours or contact me via phone / canvas messages or TAMUCT email, if there are any personal problems or challenges that are hindering your regular progress in the course.

#### **911 Cellular:**

Emergency Warning System for Texas A&M University – Central Texas

911Cellular is an emergency notification service that gives Texas A&M University-Central Texas the ability to communicate health and safety emergency information quickly via email, text

message, and social media. All students are automatically enrolled in 911 Cellular through their myCT email account.

Connect at 911Cellular [<https://portal.publicsafetycloud.net/Texas-AM-Central/alert-management>] to change where you receive your alerts or to opt out. By staying enrolled in 911Cellular, university officials can quickly pass on safety-related information, regardless of your location.

## **COURSE INFORMATION**

**Course Overview and description:** Study advanced concepts and principles of Big Data Analytics and its role in supporting/enhancing organizational decision making and predictions. Special emphasis on NoSQL Databases, Hadoop Ecosystem, MapReduce, Pig, Hive, Natural Language Processing, Social Network Analysis, and Data Visualization.

### **Course Objective:**

Successful completion of this course should enable the student to

1. Describe NoSQL databases and Hadoop Ecosystem
2. Implement Hadoop/MapReduce Functionality
3. Utilize tools such as Pig, Hive, R, etc. for analyzing large data sets
4. Demonstrate an ability to select the appropriate data visualization tool and use it to convey the insights from data analysis

### **Required Reading and Textbook:**

**Required Text:** Hadoop: The Definitive Guide, 4th Edition. Tom White. O'Reilly Media, 2012. 9781491901632

Data Analytics with Hadoop: An Introduction for Data Scientists, 1st edition, Benjamin Bengfort, Jenny Kim , 9781491913703

Note: A student of this institution is not under any obligation to purchase a textbook from a university-affiliated bookstore. The same textbook may also be available from an independent retailer, including an online retailer.

**Supplementary Material:** The course textbook will be supplemented with other materials including research papers readings, handouts, oral presentations, industry articles, videos, research paper readings, case studies, power point presentations etc.

## **COURSE REQUIREMENTS**

**Reading Assignments:** All assigned chapters will be used as basis for class and/or canvas discussions. Students are expected to study the assigned readings before each class session.

**Examination:** There will be a final exam worth 300 points. Makeup exam will be given ONLY when arrangements have been made PRIOR to the exam date.

**Assignments:** There will be 2 assignments. Each assignment is worth 150 points. Therefore, 2 assignments are worth 300 points. Assignments will be posted to canvas.

**Team Project:** There will be one term project where student teams will work as a team to solve a real-life data science problem with Hadoop framework. Students can self-select into teams of 2. This is worth 400 points. Projects turned in after due date are considered late. 10 points will be deducted for each hour the project is late. Special circumstances need to be discussed with the instructor ahead of time when possible.

### Grading Criteria Rubric and Conversion

	Points		Percentage	Points	Grade
Assignments	300		90.0-100%	900-1000	A
Team Project	400		80.0-89.9%	800-899	B
Final Exam	300		70.0-79.9%	700-799	C
			60.0-69.9%	600-699	D
			0-59.9%	0-599	F
Total	1000				

#### Posting of Grades:

- All student grades should be posted using the Canvas Grade book and students should monitor their grade status through this tool.
- Exam scores will be posted to the canvas immediately after student takes the test. Research paper score will be posted no later than one week after it is due.

#### Instructor policies related to absence, grading, etc

Makeup exams will be given only when extenuating circumstances are present. In particular, foreseeable conflicts with due dates (e.g., interviews, participation in sport activities, religious observances ...) must be brought to Prof. Erusyal's attention before the due date. You can have makeup ONLY when arrangements have been made PRIOR to the expiry date/time of exam.

Grading: Grades are not given; the student earns them. Rubrics are employed whenever possible to assure that the grading of an individual assignment is consistent. There are always disagreements about the grade given to a work product and a student's expectation; students are encouraged to make sound well-reasoned arguments as to why they should receive more points. Arguments that are not well-reasoned include:

1. I need a better grade because I am on probation.

2. I come from a good family.
3. I will be deported.
4. I will lose my scholarship.
5. I think it is an A paper

## COURSE OUTLINE AND CALENDAR

### Complete Course Calendar (Subject to Change)

**HDG:** Hadoop: The Definitive Guide

**DAH:** Data Analytics with Hadoop

Week	Date	Topic	Reading and Tasks
1	Jan 15 – 21	Class Intro, Big Data Overview	HDG Chapter 1,2 DAH Chapter 1,2. <b>Project Groups should be formed by Jan 28<sup>th</sup>, 2018.</b>
2	Jan 22 – 28		
3	Jan 29 – Feb 4	Hadoop and HDFS	HDG Chapter 3,4,5
4	Feb 5– 11		
5	Feb 12 – 18	MapReduce	HDG Chapter 6,7,8,9
6	Feb 19 – 25		
7	Feb 26 – Mar 4		
8	Mar 5 – 11	Setting up a Hadoop Cluster	HDG Chapter 10,11. <b>Assignment 1 is due on Mar 11<sup>th</sup>, 2018</b>
9	Mar 12 – 18	SPRING BREAK – NO CLASSES	
10	Mar 19 – 25	Pig	HDG Chapter 16
11	Mar 26 – Apr 1	Hive	HDG Chapter 12, DAH Chapter 17
12	Apr 2 – 8	Hadoop Streaming and Spark	DAH Chapter 3,4.
13	Apr 9 – 15	Natural Language Processing	Supplemental readings will be assigned. <b>Assignment 2 is due on Apr 15<sup>th</sup>, 2018.</b>
14	Apr 16 – 22	Social Network Analysis	Supplemental readings will be assigned
15	Apr 23 –29		Project Presentation is scheduled on Apr 28 <sup>th</sup> , 2018
16	Apr 30 – May 6	Data Visualization	Supplemental readings will be assigned

17	May 7 -10	<b>Final Exam and Project Deadline May 10<sup>th</sup>, 2018</b>
----	-----------	--

**Grading:** Grades are not given; student earns them. Rubrics are used whenever possible to assure that the grading of an individual assignment is consistent. There are always disagreements about the grade given to a work product and a student's expectation; students are encouraged to make sound well-reasoned arguments as to why they should receive more points. Arguments that are not well-reasoned include:

- I need a better grade because I am on probation.
- I come from a good family.
- I will be deported.
- I will lose my scholarship.
- I think it is an A paper

**Important University Dates:**

**January 2018**

- January 2, (Tuesday) Winter Break Ends
- January 2, (Tuesday) Priority Deadline for Admissions applications
- January 5, (Friday) VA Certification Request Priority Deadline
- January 11, (Thursday) Convocation
- January 12, (Friday) Tuition and Fee payment deadline (16 week & 1st 8 week)
- January 15, (Monday) Martin L. King Jr. Day
- January 16, (Tuesday) ADD/DROP/LATE REGISTRATION BEGINS (\$25 fee assessed for late registrants) (16 week & 1st 8 week)
- January 16, (Tuesday) Classes Begins
- January 18, (Thursday) ADD/DROP/LATE REGISTRATION ENDS (16 week & 1st 8 week)
- January 23, (Tuesday) Last day to drop 1st 8-week classes with no record
- January 31, (Wednesday) Last day to drop 16 week classes with no record

**February 2018**

- February 2, (Friday) Priority Deadline to Submit Graduation Application
- February 9, (Friday) Last day to drop a 1st 8-week class with a Q or withdraw with a W
- February 15, (Thursday) Last day to apply for Clinical Teaching
- February 23, (Friday) Student End of Course Survey Opens (1st 8-Week Classes)

**March 2018**

- March 1, (Thursday) Deadline to submit application to Teacher Education Program
- March 2, (Thursday) Deadline to Submit Graduation Application for Ceremony Participation
- March 9, (Friday) 1st 8 week classes end
- March 9, (Friday) Deadline for Admissions applications
- March 11, (Sunday) Student End of Course Survey Closes (1st 8-Week Classes)
- March 12, (Monday) Spring Break Begins
- March 12, (Monday) 1st 8-week grades from faculty due by 3pm
- March 15, (Thursday) Tuition and Fee Payment Deadline (2nd 8-week classes)
- March 16, (Friday) Spring Break Ends
- March 19, (Monday) 2nd 8 week begins
- March 19, (Monday) Summer Advising Starts
- March 19, (Monday) Class Schedule Published
- March 19, (Monday) ADD/DROP/LATE REGISTRATION BEGINS (\$25 fee assessed for late registrants) (2nd 8-week classes)

March 21, (Wednesday) ADD/DROP/LATE REGISTRATION ENDS (2nd 8-week classes)

March 27, (Tuesday) Last day to drop 2nd 8-week classes with no record

March 30, (Friday) Last day to drop a 16-week course with a Q or withdraw with a W

#### **April 2018**

April 1, (Sunday) GRE/GMAT scores due to Office of Graduate Studies

April 2, (Monday) Scholarship Deadline

April 2, (Monday) Registration begins

April 5, (Thursday) Priority Deadline for International Student Admission Applications

April 13, (Friday) Last day to drop a 2nd 8-week class with a Q or withdraw with a W\*

April 13, (Friday) Deadline for submission of final committee-edited theses with committee approval signatures to Office of Graduate Studies

April 27, (Friday) Student End of Course Survey Opens (16 Week and 2nd 8-Week Classes)

#### **May 2018**

May 7-11, Finals Week

May 11, (Friday) Last day to file for Degree Conferral (Registrar's Office)(\$20 Late Application Fee applies)

May 11, (Friday) Spring Term Ends

May 11, (Friday) Last day to withdraw from the university (16 week and 2nd 8 week classes)

May 11, (Friday) Last day to apply for \$1000 Tuition Rebate for Spring graduation (5pm)

May 12, (Saturday) Commencement Ceremony Bell County Expo Center 7:00 p.m.

May 13, (Sunday) Student End of Course Survey Closes (16 Week and 2nd 8-Week Classes)

May 14, (Monday) Minimester begins

May 15, (Tuesday) Last Day to clear Thesis Office

May 5, (Tuesday) Final grades due from faculty by 3pm (16 week & 2nd 8 week)

May 21, (Monday) Priority Deadline for Admissions applications

May 25, (Friday) VA Certification Request Priority Deadline

May 28, (Monday) Memorial Day

## **TECHNOLOGY REQUIREMENTS AND SUPPORT**

### **Technology Requirements**

This course will use the A&M-Central Texas Canvas Learning Management System [<https://tamuct.instructure.com>].

### **Technology Support.**

For log-in problems, students should contact Help Desk Central.

24 hours a day, 7 days a week:

Email: [helpdesk@tamu.edu](mailto:helpdesk@tamu.edu)

Phone: (254) 519-5466

[Web Chat](http://hdc.tamu.edu): [<http://hdc.tamu.edu>]

*Please let the support technician know you are an A&M-Central Texas student.*

For issues with **Canvas**, select "chat with Canvas support," submit a support request to Canvas Tier 1, or call the Canvas support line: 1-844-757-0953, links to all are found inside of Canvas using the "Help" link.

For issues related to course content and requirements, contact your instructor.

---

## UNIVERSITY RESOURCES, PROCEDURES, AND GUIDELINES

### **Drop Policy.**

If you discover that you need to drop this class, you must complete a [Drop Request Form](https://www.tamuct.edu/registrar/docs/Drop_Request_Form.pdf) [https://www.tamuct.edu/registrar/docs/Drop\_Request\_Form.pdf].

Professors cannot drop students; this is always the responsibility of the student. The Registrar's Office will provide a deadline on the Academic Calendar for which the form must be completed, signed and returned. Once you return the signed form to the Registrar's Office, you must go into Warrior Web and confirm that you are no longer enrolled. If you still show as enrolled, FOLLOW-UP with the Registrar's Office immediately. You are to attend class until the procedure is complete to avoid penalty for absence. Should you miss the drop deadline or fail to follow the procedure, you will receive an F in the course, which may affect your financial aid and/or VA educational benefits.

### **Academic Integrity.**

Texas A&M University -Central Texas values the integrity of the academic enterprise and strives for the highest standards of academic conduct. A&M-Central Texas expects its students, faculty, and staff to support the adherence to high standards of personal and scholarly conduct to preserve the honor and integrity of the creative community. Academic integrity is defined as a commitment to honesty, trust, fairness, respect, and responsibility. Any deviation by students from this expectation may result in a failing grade for the assignment and potentially a failing grade for the course. Academic misconduct is any act that improperly affects a true and honest evaluation of a student's academic performance and includes, but is not limited to, cheating on an examination or other academic work, plagiarism and improper citation of sources, using another student's work, collusion, and the abuse of resource materials. All academic misconduct concerns will be reported to the university's Office of Student Conduct. Ignorance of the university's standards and expectations is never an excuse to act with a lack of integrity. When in doubt on collaboration, citation, or any issue, please contact your instructor before taking a course of action.

### **Academic Accommodations.**

At Texas A&M University-Central Texas, we value an inclusive learning environment where every student has an equal chance to succeed and has the right to a barrier free education. The Department of Access and Inclusion is responsible for ensuring that students with a disability receive equal access to the University's programs, services and activities. If you believe you have a disability requiring reasonable accommodations please contact the Department of Access and Inclusion at (254) 501-5831. Any information you provide is private and confidential and will be treated as such.

For more information please visit our [Access & Inclusion](https://www.tamuct.edu/student-affairs/access-inclusion.html) webpage [https://www.tamuct.edu/student-affairs/access-inclusion.html].

Texas A&M University-Central Texas supports students who are pregnant and/or parenting. In accordance with requirements of Title IX and guidance from US Department of Education's Office of Civil Rights, the Dean of Student Affairs' Office can assist students who are pregnant and/or parenting in seeking accommodations related to pregnancy and/or parenting. For more information, please visit <https://www.tamuct.departments/index.php>. Students may also contact the institution's Title IX Coordinator. If you would like to read more about these [requirements and guidelines online](http://www2.ed.gov/about/offices/list/ocr/docs/pregnancy.pdf), please visit the website [http://www2.ed.gov/about/offices/list/ocr/docs/pregnancy.pdf].

Title IX of the Education Amendments Act of 1972 prohibits discrimination on the basis of sex and gender – including pregnancy, parenting, and all related conditions. A&M-Central Texas is able to provide flexible and

individualized reasonable accommodation to pregnant and parenting students. All pregnant and parenting students should contact the Division of Student Affairs at 254-501-5909 to seek out assistance. Students may also contact the University's Title IX Coordinator.

### **Tutoring.**

Tutoring is available to all A&M-Central Texas students, both on-campus and online. On-campus subjects tutored include Accounting, Advanced Math, Biology, Finance, Statistics, Mathematics, and Study Skills. Tutors are available at the Tutoring Center in Warrior Hall, Suite 111.

If you have a question regarding tutor schedules, need to schedule a tutoring session, are interested in becoming a tutor, or any other question, contact Academic Support Programs at 254-519-5796, or by emailing Dr. DeEadra Albert-Green at [deeadra.albertgreen@tamuct.edu](mailto:deeadra.albertgreen@tamuct.edu).

Chat live with a tutor 24/7 for almost any subject on your computer! Tutor.com is an online tutoring platform that enables A&M-Central Texas students to log-in and receive FREE online tutoring and writing support. This tool provides tutoring in over forty subject areas. Access Tutor.com through Canvas.

**University Writing Center:** Located in 416 Warrior Hall, the University Writing Center (UWC) at Texas A&M University-Central Texas is a free workspace open to all TAMUCT students from 10am-5pm Monday-Thursday with satellite hours in the University Library Monday-Thursday from 6:00-9:00pm. Students may arrange a one-on-one session with a trained and experienced writing tutor by visiting the UWC during normal operating hours (both half-hour and hour sessions are available) or by making an appointment via [WOnline](https://tamuct.mywconline.com/) at [https://tamuct.mywconline.com/]. In addition, you can email Dr. Bruce Bowles Jr. at [bruce.bowles@tamuct.edu](mailto:bruce.bowles@tamuct.edu) to schedule an online tutoring session. Tutors are prepared to help writers of all levels and abilities at any stage of the writing process.

While tutors will not write, edit, or grade papers, they will assist students in developing more effective composing practices. By providing a practice audience for students' ideas and writing, our tutors highlight the ways in which they read and interpret students' texts, offering guidance and support throughout the various stages of the writing process. In addition, students may work independently in the UWC by checking out a laptop that runs the Microsoft Office suite and connects to WIFI, or by consulting our resources on writing, including all of the relevant style guides. Whether you need help brainstorming ideas, organizing an essay, proofreading, understanding proper citation practices, or just want a quiet place to work, the University Writing Center is here to help!

If you have any questions about the University Writing Center, please do not hesitate to contact Dr. Bruce Bowles Jr. at [bruce.bowles@tamuct.edu](mailto:bruce.bowles@tamuct.edu).

### **University Library.**

The University Library provides many services in support of research across campus and at a distance. We offer over 200 electronic databases containing approximately 250,000 eBooks and 82,000 journals, in addition to the 72,000 items in our print collection, which can be mailed to students who live more than 50 miles from campus. Research guides for each subject taught at A&M-Central Texas are available through our website to help students navigate these resources. On-campus, the library offers technology including cameras, laptops, microphones, webcams, and digital sound recorders.

Research assistance from a librarian is also available twenty-four hours a day through our online chat service, and at the reference desk when the library is open. Research sessions can be scheduled for more comprehensive assistance, and may take place on Skype or in-person at the library. Assistance may cover many topics, including how to find articles in peer-reviewed journals, how to cite resources, and how to piece together research for written assignments.

Our 27,000-square-foot facility on the A&M-Central Texas main campus includes student lounges, private study rooms, group work spaces, computer labs, family areas suitable for all ages, and many other features. Services such as interlibrary loan, TexShare, binding, and laminating are available. The library frequently offers workshops, tours, readings, and other events. For more information, please visit our [Library website](https://tamuct.libguides.com/) [https://tamuct.libguides.com/].

---

## **OPTIONAL POLICY STATEMENTS:**

### **A Note about Sexual Violence at A&M-Central Texas**

Sexual violence is a serious safety, social justice, and public health issue. The university offers support for anyone struggling with these issues. University faculty are mandated reporters, so if someone discloses that they were sexually assaulted (or a victim of Domestic/Dating Violence or Stalking) while a student at TAMUCT, faculty members are required to inform the Title IX Office. If you want to discuss any of these issues confidentially, you can do so through Student Counseling (254-501-5955) located on the second floor of Warrior Hall (207L).

Sexual violence can occur on our campus because predators often feel emboldened, and victims often feel silenced or shamed. It is incumbent on ALL of us to find ways to actively create environments that tell predators we don't agree with their behaviors and tell survivors we will support them. Your actions matter. Don't be a bystander; be an agent of change. For additional information on campus policy and resources visit the [Title IX webpage](https://www.tamuct.edu/departments/compliance/titleix.php) [https://www.tamuct.edu/departments/compliance/titleix.php].

## **INSTRUCTOR POLICIES.**

1. Instructor reserves the right to modify the syllabus during the course of the semester for the benefit of the students.
2. Instructor reserves the right to supplement the material presented in the text with addition material that may benefit the students by either providing additional information or a different point of view.
3. Instructor expects that the students will act in a curious and professional manner in all interactions with other students and the instructor.
4. Instructor reserves the right to modify grading rubrics. Changes to grading rubrics are only made to current and/or future assignments.
5. Any changes made will be announced in class and posted on canvas.

### **Copyright Notice.**

Students should assume that all course material is copyrighted by the respective author(s). Reproduction of course material is prohibited without consent by the author and/or course instructor. Violation of copyright is against the law and Texas A&M University-Central Texas' Code of Academic Honesty. All alleged violations will be reported to the Office of Student Conduct.

Copyright. 2017 by Ferdi Eruysal at Texas A&M University-Central Texas, College of Business; 1001 Leadership Place, Killeen, TX 76549; 254-519-5440; Fax 254-501-5825; [ferdi.eruysal@tamuct.edu](mailto:ferdi.eruysal@tamuct.edu)