

Course Syllabus: Industrial data analytics
(IE-4395 CRN:28235), (SE-5390 CRN:28234), (MFG-5390, CRN:24730), and (IE-5390 CRN:28236)
Spring 2022

(Class will be delivered online, class materials will be posted on Blackboard every Monday before 6:00pm)

Instructor	Dr. Md Fashiar Rahman
Location and Time	Online via Blackboard.
Office	E-201L
Office Hours	Wednesday 4 PM - 5 PM via Blackboard and by appointment
Email	mrahman13@utep.edu
Teaching Assistant	Hugo Soto Email: hsoto8@miners.utep.edu Office hour: Friday 1 PM – 2 PM via Blackboard and by appointment

Course web page:

<https://blackboard.utep.edu/>

COURSE MOTIVATION:

21st century has been called by many “The Century of Data.” We see more and more data collected with the expectation (justified by empirical evidence!) that analyzing these data will give organizations a competitive edge and will help them to excel. The amount of data collected is enormous and growing. In many cases, analyzing these data lags behind. Mark Twain wrote once “A man who does not read has no advantage over a man who cannot read.” A similar sentence is most certainly true: “A company who does not analyze his data has no advantage over a company who has no data.” Methods for analyzing data, called collectively “data analytics” are, therefore, crucial in for businesses to gain competitive advantage.

Data mining (Data analytics), also called Knowledge Discovery in Databases (KDD), is the field of discovering novel and potentially useful information from large amounts of data. Data mining has been applied in a great number of fields, including retail sales, bioinformatics, healthcare, manufacturing and counter-terrorism, among others. Data analytics is an important, fast-growing field that has quickly become a key basis of productivity growth, innovation, and consumer surplus. There is an increasing need for analytics-savvy employees who can think uniquely across disciplines to transform data into relevant insights for making better business decisions. It is a highly inter-disciplinary area representing the confluence of machine learning, statistics, operations research, database systems and high-performance computing, among other fields.

COURSE DESCRIPTION (3 CREDITS)

This course is an **introductory course** open to both graduate and senior undergraduate students. As an introductory course on data analytics, this course introduces the concepts, algorithms, techniques, and systems of data mining, including (1) an introduction to data analytics, (2) data preprocessing, (3) mining frequent patterns and associations, (4)

classification, (5) cluster analysis, and (6) learning about software used in data mining and (7) demonstration of how to apply data analytics techniques using Python. The course will provide students with basic understanding of common data mining and analytics techniques.

PREREQUISITES:

Basic engineering statistics and basic calculus and linear algebra. Having basic programming knowledge would be an advantage. However, this is not a course of mathematical statistics or advanced calculus.

COURSE OBJECTIVES:

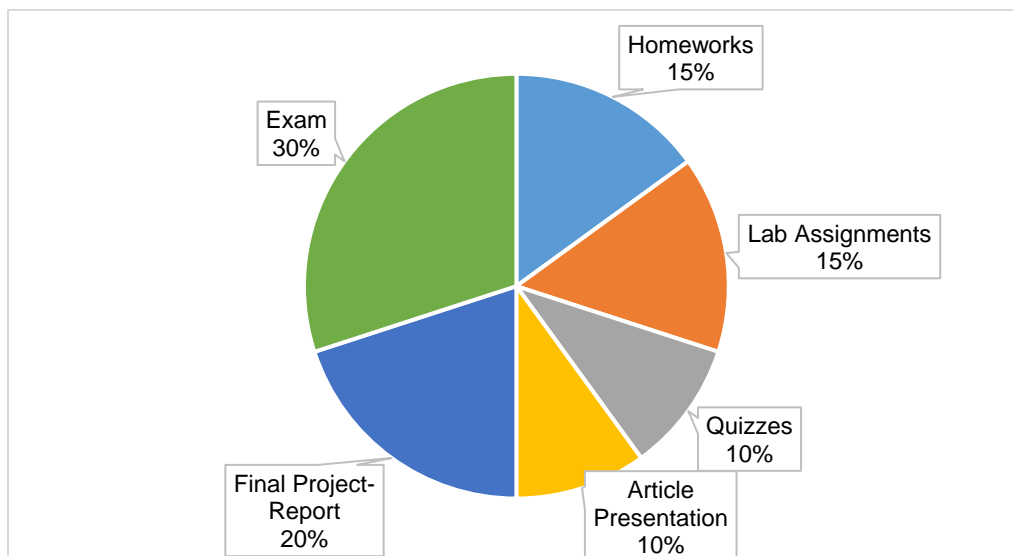
- Understand the importance of Data Analytics
- Students will learn about fundamental data analytic techniques and tools used to analyze data
- Be capable of confidently applying common data analytic algorithms in practice
- Be capable of performing data mining experiments using programming in Python
- Be capable of performing experiments in using real-world data

LEARNING OUTCOMES:

At the end of this course, students should be able to

- Understand modern views in data analysis;
- Explain the data mining methodology.
- Use visual techniques to describe data.
- Explain and apply the different regression, clustering, classification, and association rule mining techniques.
- Able to explain the assumption and interpret the output of the above data catalytic techniques

COURSE ASSIGNMENTS AND GRADING SCHEME:



TEXTBOOK

There is no required textbook. However, necessary reading materials will be provided during the semester.

USEFUL REFERENCES:

- Data Mining: Concepts and Techniques, 3rd edition, written by Jiawei Han, Micheline Kamber, Jian Pei.
- An Introduction to Statistical Learning, written by Daniela Witten, Trevor Hastie, Robert Tibshirani
- An Introduction to R by W. N. Venables, D. M. Smith and the R Core Team
- Materials not from the textbook will be provided in lecture notes.

HOMEWORKS, LAB ASSIGNMENTS AND MAKE-UP WORK POLICY:

Homeworks/Assignments will be announced every week. Make-up work will be given only in the case of a documented emergency. Note that make-up work may be in a different format than the original work, may require more intensive preparation, and **may be graded with penalty points**. If you miss an assignment and the reason is not considered excusable, you will receive a zero. It is therefore important to reach out to me—in advance if at all possible—and explain with proper documentation why you missed a given course requirement. Once a deadline has been established for make-up work, no further extensions or exceptions will be granted.

QUIZZES

The quizzes will in class without previous notice. The contents of the quiz will be covered before on the same class. So, all the students are suggested to be regular in class. Note that make-up quiz may be in a different format than the original work, may require more intensive preparation, and may be graded with penalty points. If you miss a quiz and the reason is not considered excusable, you will receive a zero.

ARTICLE PRESENTATION:

A team of 2-3 students is required to present a journal article or case study, and prepare a 10-12-minute presentation. The team must understand and explain the data analytics model being presented in the article to solve the problem addressed. Sample application areas are

- Manufacturing
- Healthcare
- Agriculture
- Homeland Security
- Finance
- Transportation
- Energy
- Environment, etc.

It is recommended that the students use the **Google Scholar (scholar.google.com)** and find the paper of their area of interest using the key words such as “big data analytics”, “data analytics in manufacturing” etc. It is expected that the group will review **recently published literature** on the topic of interest (not publish before year 2010). You can customize the range as showed in the following figure.

It is recommended that the students review the following journals in order to select a journal article for the journal presentation:

FINAL PROJECT REPORT

The purpose of the term project is to provide an opportunity to apply and/or further explore a topic area related to the course. Some examples of types of projects that are acceptable are:

- Literature review of data mining methods used in a problem area that demonstrates breadth and organization of the literature
- Research proposal incorporating engineering analysis that demonstrates creativity and originality
- Apply different models to a specific problem and compare results (interest is in models not covered in class, maybe use some from the class and present a new one)

Each team is required to identify a topic of interest and submit a one-page abstract by **February 28th, 2022**. It is expected that the group will review recently published literature on the topic of interest. The project will be evaluated based on the technical content (30%), quality of documented references (40%) and quality of technical writing (40%).

Final project report format: (written in word, 1.5 spacing, Times New Roman 11, 15-20 pages) due (**May 09th, 2022**)

POSSIBLE OUTLINE FOR PROJECT REPORTS

Prepare your final report for an interdisciplinary group of decision makers. Your written reports must be formatted into different section headings (shown below)

I. Introductory Pages

- a. Cover Page: title, authors
- b. Table of Contents, List of Tables, and List of Figures

II. Executive Summary (brief summary of the overall project)

III. Introduction

- a. Mention the importance of the problem/area being studied
- b. Ideally provide specific information from reports related to the problem and its significance (**Cite at least 10 references**)
- c. Explain benefits of solving the problem, including motivation economics
- d. Present a clear and concise definition of the problem

IV. Literature Review

- a. Present an up-to date literature review from Journal publications and conference proceedings
- b. Write a small paragraph from each article indicating what the authors did to solve the problem presented in the publication
- c. End the literature review section mentioning a specific idea about how you may expand the work presented by them

V. Model Development

- a. Mention the data mining model/algorithm which will be for the analysis of the problem indicated in the previous section

VI. Case study

- a. Clearly state the problem where the algorithm will be applied
- b. Apply the data mining algorithm(s)

VII. Conclusions and Future Work

- a. Write a paragraph mentioning the final conclusions for the proposed work
- b. Mention final recommendations and Future work

VIII. Bibliography

TECHNOLOGY REQUIREMENTS

Course content is delivered via the Internet through the Blackboard learning management system. Ensure your UTEP e-mail account is working and that you have access to the Web and a stable web browser. Google Chrome and Mozilla Firefox are the best browsers for Blackboard; other browsers may cause complications. When having technical difficulties, update your browser, clear your cache, or try switching to another browser.

You will need to have access to a computer/laptop, scanner, a webcam, and a microphone. You will need to download or update the following software: Microsoft Office, Adobe Acrobat Reader, Windows Media Player, Python, and Minitab. Check that your computer hardware and software are up-to-date and able to access all parts of the course.

If you do not have a word-processing software, you can download Word and other Microsoft Office programs (including Excel, PowerPoint, Outlook and more) for free via UTEP's Microsoft Office Portal. Click the following link for more information about Microsoft Office 365 and follow the instructions.

IMPORTANT: If you encounter technical difficulties beyond your scope of troubleshooting, please contact the UTEP Help Desk as they are trained specifically in assisting with technological needs of students. Please do not contact me for this type of assistance. The Help Desk is much better equipped than I am to assist you!

COURSE COMMUNICATION

Because this is an online class, we won't see each other in the ways you may be accustomed to: during class time, small group meetings, and office hours. However, there are a number of ways we can keep the communication channels open:

- **Office Hours:** We will not be able to meet on campus, but I will still have office hours for your questions and comments about the course. My office hours will be held on Blackboard Collaborate using this link: [\[click here\]](#) and during the following times:
 - Mondays: 4-5 p.m. Mountain Time
 - Wednesday: 4-5 p.m. Mountain Time
- **Email:** UTEP e-mail is the best way to contact me. I will make every attempt to respond to your e-mail within 24-48 hours of receipt. When e-mailing me, be sure to email from your UTEP student account and please put the course number in the subject line. In the body of your e-mail, clearly state your question. At the end of your e-mail, be sure to put your first and last name, and your university identification number.
- **Discussion Board:** If you have a question that you believe other students may also have, please post it in the Help Board of the discussion boards inside of Blackboard. Please respond to other students' questions if you have a helpful response.
- **Announcements:** Check the Blackboard announcements frequently for any updates, deadlines, or other important messages.

NETIQUETTE

As we know, sometimes communication online can be challenging. It's possible to miscommunicate what we mean or to misunderstand what our classmates mean given the lack of body language and immediate feedback. Therefore, please keep these netiquette (network etiquette) guidelines in mind. Failure to observe them may result in disciplinary action.

- Always consider audience. This is a college-level course; therefore, all communication should reflect polite consideration of other's ideas.
- Respect and courtesy must be provided to classmates and to the instructor at all times. No harassment or inappropriate postings will be tolerated.
- When reacting to someone else's message, address the ideas, not the person. Post only what anyone would comfortably state in a face-to-face situation.
- Blackboard is not a public internet venue; all postings to it should be considered private and confidential. Whatever is posted on in these online spaces is intended for classmates and professor only.

ATTENDANCE AND PARTICIPATION

Attendance in the course is determined by participation in the learning activities of the course. Your participation in the course is important not only for your learning and success but also to create a community of learners. Participation is determined by completion of the following activities:

- Reading/Viewing all course materials to ensure understanding of assignment requirements
- Participating in engaging discussion with your peers on the discussion boards (grading rubric provided in the "grading information" area of each forum)
- Participating in scheduled Blackboard Collaborate sessions
- Other activities as indicated in the weekly modules

Because these activities are designed to contribute to your learning each week, they cannot be made up after their due date has passed.

EXCUSED ABSENCES AND/OR COURSE DROP POLICY

According to UTEP Curriculum and Classroom Policies, "When, in the judgment of the instructor, a student has been absent to such a degree as to impair his or her status relative to credit for the course, the instructor may drop the student from the class with a grade of "W" before the course drop deadline and with a grade of "F" after the course drop deadline." See academic regulations in the UTEP Undergraduate Catalog for a list of excuse absences. Therefore, if I find that, due to non-performance in the course, you are at risk of failing, I will drop you from the course. I will provide 24 hours advance notice via email.

OR

I will not drop you from the course. However, if you feel that you are unable to complete the course successfully, please let me know and then contact the Registrar's Office to initiate the drop process. If you do not, you are at risk of receiving an "F" for the course.

BLACKBOARD COLLABORATIVE SESSIONS

This class requires that you participate in scheduled Blackboard Collaborate sessions. The purpose of these sessions are for you to view live demonstrations of the course material and/or to participate in small discussion groups with your classmates. These sessions will be held **each Monday from 6 p.m – 8.50 p.m.**

Students are expected to, at least occasionally, participate in these sessions with a webcam and microphone. The sessions will be recorded and provided so that they can be

reviewed by classmates at a later time. Students should not record the sessions and post them to any sites outside of Blackboard.

If you are unable to attend a Collaborate session, please let me know as soon as possible so that accommodations can be made when appropriate.

ALTERNATE MEANS OF SUBMITTING WORK IN CASE OF TECHNICAL ISSUES

I strongly suggest that you submit your work with plenty of time to spare in the event that you have a technical issue with the course website, network, and/or your computer. I also suggest you save all your work (answers to discussion points, quizzes, exams, and essays) in a separate Word document as a back-up. This way, you will have evidence that you completed the work and will not lose credit. If you are experiencing difficulties submitting your work through the course website, please contact the UTEP Help Desk. You can email me your back-up document as a last resort.

ACCOMMODATIONS POLICY:

The University is committed to providing reasonable accommodations and auxiliary services to students, staff, faculty, job applicants, applicants for admissions, and other beneficiaries of University programs, services and activities with documented disabilities in order to provide them with equal opportunities to participate in programs, services, and activities in compliance with sections 503 and 504 of the Rehabilitation Act of 1973, as amended, and the Americans with Disabilities Act (ADA) of 1990 and the Americans with Disabilities Act Amendments Act (ADAAA) of 2008. Reasonable accommodations will be made unless it is determined that doing so would cause undue hardship on the University. Students requesting an accommodation based on a disability must register with the UTEP Center for Accommodations and Support Services (CASS). Contact the Center for Accommodations and Support Services at 915-747-5148, or email them at cass@utep.edu, or apply for accommodations online via the CASS portal.

COVID-19 ACCOMMODATION

Students are not permitted on campus when they have a positive COVID-19 test, exposure or symptoms. If you are not permitted on campus, you should contact me as soon as possible so we can arrange necessary and appropriate accommodations.

(Classes with on-campus meetings) Students who are considered high risk according to CDC guidelines and/or those who live with individuals who are considered high risk may contact Center for Accommodations and Support Services (CASS) to discuss temporary accommodations for on-campus courses and activities.

SCHOLASTIC INTEGRITY:

Academic dishonesty is prohibited and is considered a violation of the UTEP Handbook of Operating Procedures. It includes, but is not limited to, cheating, plagiarism, and collusion. Cheating may involve copying from or providing information to another student, possessing unauthorized materials during a test, or falsifying research data on laboratory reports. Plagiarism occurs when someone intentionally or knowingly represents the words or ideas of another as ones' own. Collusion involves collaborating with another person to commit any academically dishonest act. Any act of academic dishonesty attempted by a UTEP student is unacceptable and will not be tolerated. All suspected violations of academic integrity at The University of Texas at El Paso must be reported to the Office of Student Conduct and Conflict Resolution (OSCCR) for possible disciplinary action. To learn more, please visit HOOP: Student Conduct and Discipline.

CLASS RECORDINGS:

The use of recordings will enable you to have access to class lectures, group discussions, and so on in the event you miss a synchronous or in-person class meeting due to illness

or other extenuating circumstance. Our use of such technology is governed by the Federal Educational Rights and Privacy Act (FERPA) and UTEP's acceptable-use policy. A recording of class sessions will be kept and stored by UTEP, in accordance with FERPA and UTEP policies. I will not share the recordings of our class activities outside of course participants. You may not share recordings outside of this course. Doing so may result in disciplinary action.

PROGRAMMING LANGUAGE – PYTHON:

Python is an interpreted, object-oriented, high-level programming language with dynamic semantics. Its high-level built-in data structures, combined with dynamic typing and dynamic binding, make it very attractive for Rapid Application Development, as well as for use as a scripting or glue language to connect existing components together. Python is simple, easy to learn syntax emphasizes readability and therefore reduces the cost of program maintenance. Python supports modules and packages, which encourages program modularity and code reuse. The Python interpreter and the extensive standard library are available in source or binary form without charge for all major platforms, and can be freely distributed. <https://www.python.org/>

COVID-19 PRECAUTION

You must STAY AT HOME and REPORT if you (1) have been diagnosed with COVID-19, (2) are experiencing COVID-19 symptoms, or (3) have had recent contact with a person who has received a positive coronavirus test. Reports should be made at screening.utep.edu. If you know of anyone who should report any of these three criteria, you should encourage them to report. If the individual cannot report, you can report on their behalf by sending an email to COVIDaction@utep.edu.

For each day that you attend campus—for any reason—you must complete the questions on the UTEP screening website (screening.utep.edu) prior to arriving on campus. The website will verify if you are permitted to come to campus. Under no circumstances should anyone come to class when feeling ill or exhibiting any of the known COVID-19 symptoms. If you are feeling unwell, please let me know as soon as possible, and alternative instruction will be provided. Students are advised to minimize the number of encounters with others to avoid infection.

Wear face coverings when in common areas of campus or when others are present. You must wear a face covering over your nose and mouth at all times in this class. If you choose not to wear a face covering, you may not enter the classroom. If you remove your face covering, you will be asked to put it on or leave the classroom. Students who refuse to wear a face covering and follow preventive COVID-19 guidelines will be dismissed from the class and will be subject to disciplinary action according to Section 1.2.3 Health and Safety and Section 1.2.2.5 Disruptions in the UTEP Handbook of Operating Procedures.

(Classes with on-campus meetings) Please note that if COVID-19 conditions deteriorate in the City of El Paso, all course and lab activities may be transitioned to remote delivery.

COURSE RESOURCES

UTEP provides a variety of student services and support:

Technology Resources

- **Help Desk:** Students experiencing technological challenges (email, Blackboard, software, etc.) can submit a ticket to the UTEP Helpdesk for assistance. Contact the Helpdesk via phone, email, chat, website, or in person if on campus.

Academic Resources

- **UTEP Library:** Access a wide range of resources including online, full-text access to thousands of journals and eBooks plus reference service and librarian assistance for enrolled students.
- **University Writing Center (UWC):** Submit papers here for assistance with writing style and formatting, ask a tutor for help and explore other writing resources.
- **Math Tutoring Center (MaRCS):** Ask a tutor for help and explore other available math resources.
- **History Tutoring Center (HTC):** Receive assistance with writing history papers, get help from a tutor and explore other history resources.
- **RefWorks:** A bibliographic citation tool; check out the RefWorks tutorial and Fact Sheet and Quick-Start Guide.

Individual Resources

- **Military Student Success Center:** Assists personnel in any branch of service to reach their educational goals.
- **Center for Accommodations and Support Services:** Assists students with ADA-related accommodations for coursework, housing, and internships.
- **Counseling and Psychological Services:** Provides a variety of counseling services including individual, couples, and group sessions as well as career and disability assessments.

TENTATIVE WEEKLY CALENDER (IT MAY CHANGE, BASED ON FEEDBACK OR PROGRESS)

Week	Date	Topics	Assignments Due
1	01/17	Dr. Martin Luther King, Jr. Holiday, No Class	
2	01/24	Course syllabus and introduction to data analytics	
3	01/31	Getting to know your data	Assignment 1
4	02/07	Data Preprocessing	
5	02/14	Lab # 1: Python installation and basics	Assignment 2
6	02/21	Lab # 2: Working with Panda library and data preprocessing	
7	02/28	Linear regression	Assignment 3
8	03/07	Logistic regression	
9	03/14	Spring Break	
10	03/21	Lab # 3: Linear and Logistics regression	Assignment 4
11	03/28	Clustering – Partitioning method	
12	04/04	Clustering – DBSCAN and	Assignment 5
13	04/11	Classification	
14	04/18	Lab #4: Clustering and Classification	
15	04/25	Association Rule Mining	Assignment 6
16	05/02	Final Exam	
17	05/09	Final Project Report due	Project report