



AN HONORS UNIVERSITY IN MARYLAND

Department of Computer Science and Electrical Engineering

CMSC 462: Introduction to Data Science

Spring 2023 Course Syllabus Tu/Th 10:00 AM-11:15 AM

This syllabus is tentative. I reserve the right to add, remove or alter this syllabus as needed throughout the semester.

Instructor: Abhijit Dutt, PhD, Professor of the practice

Cell Phone: (414) 322 8188

E-mail: adutt@umbc.edu

Office Hours: : Tuesday 11: 15 AM – 11:45 AM, Wednesday 12:00 noon – 2:30 PM
and by appointment. Sometimes, the office hours may be conducted through MS Teams.

Required Texts and Learning Materials:

Textbook 1: An Introduction to Statistical Learning with Applications in R by Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani, Second edition, Available for free at the following website: <https://www.statlearning.com/>

Textbook 2: R for Data Science by Hadley Wickham

Available for free at the following website: <http://r4ds.had.co.nz/index.html>

Textbook 3: Data Mining: Concepts and Techniques by J. Han, M. Kamber, and J. Pei (3rd edition, 2011)

Selected Chapters will be posted in Blackboard

Course Description

Data science is an evolving and interdisciplinary field that involves extraction of knowledge from both structured and unstructured data as well as efficient management of those data. It uses techniques and theories drawn from many different fields such as statistics, mathematics, computer science, social science etc.

Course Objectives

1. Introduce the basic terminologies and concepts associated with data science.
2. To learn R.
3. Understand how to do data manipulation, analysis, and presentation.
4. Learn different statistical techniques for model building and knowledge extraction
5. To learn about distributed computing and to learn techniques for storing big data.

Course Policies:

- The class will be offered strictly on face to face format. Thanks to vaccines, COVID-19 is no longer a threat. However, things may change and follow all the updates from UMBC about COVID-19 at <https://covid19.umbc.edu/> .
- Please note that masking is no longer required at UMBC. However, if you are sick or if you have any underlying health condition, then please wear a mask.
- I use MS Teams for communication, office hours and discussions. Please install MS Teams on your computer and on your phone.
- The main communication will be through MS Teams. If you need to send me an email, please make sure that you mention your class and section number in the subject. I will only respond to umbc.edu e-mails and will not respond to e-mails written on a private account.
- We shall use Blackboard for sharing course materials and assigning assignment.
- You must finish the assigned reading for the week before classes every week. The week 1 is an exception. You should be able to discuss the reading material in class and your participation grade will be dependent on that.
- Attendance will be taken at every class, and it is mandatory. I may conduct 1/2 polls during every class. You must participate in all the polls to be marked present in a class. Any class material missed by the student is the student's responsibility to acquire.
- There will be two tests and a final. The format of the tests will be discussed in class. No makeup examination will be given.
- All the assignments are due on the previous day of a class at midnight. Late assignments will receive at least a 10% penalty unless prior approval is given.
- Athletes with travel schedules should meet with me by the end of the first week of classes to discuss any necessary arrangements that need to be made.
- Arrangements for any religious observances or UMBC sanctioned activity must be arranged with the instructor at least one week prior to the event.
- By remaining registered in the course through drop/add period, you agree to all terms and policies set forth in the syllabus.

Accessibility and Disability Accommodations, Guidance and Resources

Accommodations for students with disabilities are provided for all students with a qualified disability under the Americans with Disabilities Act (ADA & ADAAA) and Section 504 of the Rehabilitation Act who request and are eligible for accommodations. The Office of Student Disability Services (SDS) is the UMBC department designated to coordinate accommodations that creates equal access for students when barriers to participation exist in University courses, programs, or activities.

If you have a documented disability and need to request academic accommodations in your courses, please refer to the SDS website at sds.umbc.edu for registration information and office procedures.

SDS email: disAbility@umbc.edu

SDS phone: (410) 455-2459

If you will be using SDS approved accommodations in this class, please contact the instructor to discuss implementation of the accommodations.

Sexual Assault, Sexual Harassment, and Gender Based Violence and Discrimination

[UMBC Policy](#) and Federal law (Title IX) prohibit discrimination and harassment on the basis of sex, sexual orientation, and gender identity in University programs and activities. Any student who is impacted by sexual harassment, sexual assault, domestic violence, dating violence, stalking, sexual exploitation, gender discrimination, pregnancy discrimination, gender-based harassment or retaliation should contact the [University's Title IX Coordinator](#) to make a report and/or access support and resources:

Mikhel A. Kushner, Title IX Coordinator (she/they)
410-455-1250 (direct line), kushner@umbc.edu

You can access support and resources even if you do not want to take any further action. You will not be forced to file a formal complaint or police report. Please be aware that the University may take action on its own if essential to protect the safety of the community.

If you are interested in or thinking about making a report, please use the [Online Reporting/Referral Form](#). Please note that, if you report anonymously, the University's ability to respond will be limited.

Notice that Faculty are Responsible Employees with Mandatory Reporting Obligations:

All faculty members are considered *Responsible Employees*, per [UMBC's Policy on Sexual Misconduct, Sexual Harassment, and Gender Discrimination](#). Faculty are therefore required to report any/ all available information regarding conduct falling under the Policy and violations of the Policy to the Title IX Coordinator, even if a student discloses an experience that occurred before attending UMBC and/or an incident that only involves people not affiliated with UMBC. Reports are required regardless of the amount of detail provided and even in instances where support has already been offered or received.

While faculty members want encourage you to share information related to your life experiences through discussion and written work, students should understand that faculty are required to report *past and present* sexual assault, domestic and interpersonal violence, stalking, and gender discrimination that is shared with them to the Title IX Coordinator so that the University can inform students of their [rights, resources and support](#). While you are encouraged to do so, you are not obligated to respond to outreach conducted as a result of a report to the Title IX Coordinator.

If you need to speak with someone in confidence, who does not have an obligation to report to the Title IX Coordinator, UMBC has a number of [Confidential Resources](#) available to support you:

- [The Counseling Center](#) (Main Campus): 410-455-2472 / After-Hours 410-455-3230 [Monday – Friday; 8:30 a.m. – 5 p.m.]
- [Center for Counseling and Consultation](#) (Shady Grove Campus): 301-738-6273 (Messages checked hourly) [Online Appointment Request Form](#)
- [University Health Services](#): 410-455-2542 [Monday – Friday 8:30 a.m. – 5 p.m.]
- Pastoral Counseling via [Interfaith Center](#): 410-455-3657; interfaith@umbc.edu [7 days a week; Fall and Spring 7 a.m. – 11 p.m.; Summer and Winter 8 a.m. – 8 p.m.]

Other Resources:

- [Women's Center \(for students of all genders\): 410-455-2714; \[womenscenter@umbc.edu\]\(mailto:womenscenter@umbc.edu\). \[Monday – Thursday 9:30am-6pm and Friday 9:30am-4pm\]](#)
- [Shady Grove Student Resources, Maryland Resources, National Resources.](#)

Child Abuse and Neglect

Please note that Maryland law requires that I report all closure or suspicions of child abuse or neglect to the Department of Social Service and/or the police.

Mental Health Resources

Diminished mental health can interfere with optimal academic performance. The source of symptoms might be related to your course work; if so, please speak with your instructor. However, problems with other parts of your life can also contribute to decreased academic performance. UMBC provides cost-free and confidential mental health services through the Counseling Center to help you manage personal challenges that threaten your personal or academic well-being.

Remember, getting help is a smart and courageous thing to do -- for yourself and for those who care about you. For more resources get the Just in Case mental health resources Mobile and Web App. This app can be accessed on this web page: <http://counseling.umbc.edu/justincase>
The UMBC Counseling Center is in the Student Development & Success Center (between Chesapeake and Susquehanna Halls). Phone: 410-455-2542. Hours: Monday-Friday 8:30am - 5:00pm.

Academic Integrity:

You will be held to UMBC's Undergraduate Student Academic Conduct Policy. The details of the policy can be found here: <https://fdc.umbc.edu/resources/adjunct/academic-integrity/>

Course Grade

The final letter grade is assigned objectively and strictly according to the WEIGHTED average of the numerical scores of all exams, quizzes, and assignments.

90.0%-100.0%	A	80.0%-89.9%	B	70.0% -79.9%	C	60.0%-69.9%	D	below 60.0%	F
--------------	----------	-------------	----------	--------------	----------	-------------	----------	-------------	----------

Methods of Student Evaluation:

Students will be evaluated based on assignments, projects, quizzes, tests, and class activities.

Item	Tests	Homework, Assignments, Quiz	Participation	Total
Points	250	125	25	400

TENTATIVE COURSE SCHEDULE: THIS SCHEDULE WILL CHANGE

Week	Dates	Topics	Reading and Assignments*
1	1/31-2/2	Course Orientation, Introduction, Intro to Statistics Intro to Statistics, Data Science process, Introduction to R	Week 1 papers in Blackboard
2	2/7 – 2/9	Statistics Lab with Excel and R, Excel Pivot Table, Linear Regression	
3	2/14 – 2/16	Bias-Variance theory, Linear & Multiple regression	Chapters 2.2, 3 Assignment 1 Due 2/13
4	2/21 – 2/23	Linear & Multiple regression, Classification – Naïve Bayes, Logistic regression	Sections 4.3, 4.4.1 Assignment 2 Due 2/20
5	2/28 – 3/2	Classification – Logistic regression, Test-1 (3/2)	
6	3/7 – 3/9	Classification – Logistic regression Unsupervised learning – PCA	Chapters 4, 10
7	3/14 – 3/16	Unsupervised learning – PCA, Clustering, Market Basket Analysis	Chapter 10, additional readings Assignment 3 Due 3/15
8	3/19 – 3/26	Spring Break	
9	3/28 – 3/30	Data Preprocessing, Data Visualization	Chapters 2 & 3 – Han et al
10	4/4 – 4/6	Databases – SQL & NoSQL	Readings will be provided
11	4/11 - 4/13	MapReduce and Hadoop	Assignment 4 Due 4/10 Readings will be provided,
12	4/18- 4/20	MapReduce and Hadoop, Test-2 (4/20)	Readings will be provided
13	4/25 – 4/27	MapReduce and Hadoop	Readings will be provided,
14	5/2 – 5/4	Spark	Readings will be provided Assignment 5 Due 5/3
15	5/9 – 5/11	Spark	Readings will be provided,
16	5/16	Review for Final	
	5/23	Final – 10:30 AM – 12:30 PM (Tuesday)	

* - Readings are from Textbook 1 unless specifically mentioned otherwise