

Introduction to Data Science

CAP 4770 Section 0358

Class Periods: M/W/F, 5th period, 11:45am – 12:35pm

Location: FAB 0103

Academic Term: Fall 2019

Instructor:

Daisy Zhe Wang

daisyw@ufl.edu

(352) 294-6677

Office Hours: M/F, 10-11am, CSE E456/E457

Teaching Assistant/Peer Mentor/Supervised Teaching Student:

Please contact through the Canvas website

- Yifan Wang, wangyifan@ufl.edu, CSE E309/E312, W 2-4pm

Course Description

CAP 4770 Introduction to Data Science 3 Credits

Grading Scheme: Letter Grade

Introduces the basics of data science including programming for data analytics, file management, relational databases, classification, clustering, and regression; lays the foundation for big data applications ranging from social networks to medical and business informatics.

Course Pre-Requisites / Co-Requisites

Prerequisite: COP 3530 or equivalent

Course Objectives

By the end of this course, students will be able to:

1. Implement the basic data science techniques including programming in Python and MapReduce for small and big data manipulation and analytics.
2. Analyze topics including data preparation, data analytics including pattern mining, classification, clustering, data visualization, and parallel computing platforms.
3. Apply advanced data analytics including text retrieval, information extraction, graph analysis, and knowledge bases.

During the course, we will also make an effort to introduce key application areas of data science including business intelligence, social sciences, life sciences, biomedical informatics, and legal analytics.

Materials and Supply Fees

None

Professional Component (ABET):

The course prepares students data science skills to handle datasets with volume, veracity and value, to perform data preprocessing, analysis, visualization, modeling and evaluation. It also prepares students to work in a team, develop a project and communicate results to large audience.

Relation to Program Outcomes (ABET):

Outcome	Coverage*
1. An ability to identify, formulate, and solve engineering problems by applying principles of engineering, science, and mathematics.	High
2. An ability to apply both analysis and synthesis in the engineering	Medium

design process, resulting in designs that meet desired needs.	
3. An ability to develop and conduct appropriate experimentation, analyze and interpret data, and use engineering judgment to draw conclusions.	High
4. An ability to communicate effectively with a range of audiences	Medium
5. An ability to recognize ethical and professional responsibilities in engineering situations and make informed judgments, which must consider the impact of engineering solutions in global, economic, environmental, and societal contexts.	
6. An ability to recognize the ongoing need for additional knowledge and locate, evaluate, integrate, and apply this knowledge appropriately.	
7. An ability to function effectively on teams that establish goals, plan tasks, meet deadlines, and analyze risk and uncertainty	Medium

Required Textbooks and Software

- Data Mining: Concepts and Techniques
- Jiawei Han, Micheline Kamber and Jian Pei
- The Morgan Kaufmann Series in Data Management Systems, Morgan Kaufmann Publishers
- July 2011 and 3rd ed.
- ISBN 978-0123814791

Required Software: IPython, Github, Slack, Virtual Machine, and Amazon Web Services (AWS) Student Account.

Recommended Materials

- Mining of massive datasets (MMD)
- A. Rajaraman and J.D. Ullman
- Cambridge University Press, 2011
- ISBN-10: 1107015359, ISBN-13: 978-1107015357
- <http://www.mmds.org/>

Course Schedule

Week 1: [Course Introduction / DM Chap 1](#)
Week 2: [Data Exploration / DM Chap 2 / Lab 1](#)
Week 3: [Data Wrangling / DM Chap 3 / Lab 2](#)
Week 4: [Pattern Mining / DM Chap 6 / Lab 3](#)
Week 5: [Classification / DM Chap 8 / Exercise 1](#)
Week 6: [Classification / DM Chap 8 / Lab 5](#)
Week 7: [Midterm Review / Lab 5/ Exercise 2](#)
Week 8: [Map-Reduce / MMD Chap 2 / Midterm](#)
Week 9: [Advanced Classification / MMD Chap 7 / Lab 6](#)
Week 10: [Text Analysis / Lecture Notes / Lab 7](#)
Week 11: [Graph Analysis / MMD Chap 5 / Exercise 3/ Project Proposal](#)
Week 12: [Clustering / DM Chap 9](#)
Week 13: [Clustering/ DM Chap 9](#)
Week 14: [Final Project Presentations / Exercise 4](#)
Week 15: [Final Project Presentations](#)
Week 16: [Final Review/ Final Exam](#)

A more detailed schedule is posted on the course homepage on canvas.

Attendance Policy, Class Expectations, and Make-Up Policy

- We require class and lab attendance and participation. Attendance will be checked in class/labs.
- Personal laptop computers are required in class for participation in labs and projects.
- Please return your projects in time. Late returns will cause 20% deduction in your grade for that lab/homework/project for each late day.
- If the office hour is cancelled, it will be posted in the announcements.
- Please avoid any activities that will disturb the flow of the lectures: Silence your cell phones, pagers, etc.

Excused absences are consistent with university policies in the undergraduate catalog (<https://catalog.ufl.edu/ugrad/current/regulations/info/attendance.aspx>) and require appropriate documentation.

Policy on Missed Quizzes and Late Assignments

I am extremely sympathetic if you have some conflict that will make it difficult for you to attend a quiz or get an assignment completed on time. To be fair to everyone in the class, these rules are always followed to the letter and without exception.

1. If you have some conflict and feel like you may need an extra day on the assignment or need to take the quiz a day or two later, it must be cleared with me no fewer than one week (seven days) before the assignment is due or the quiz will be held. However, if you do not clear it a week in advance, you will receive a zero, with only a two exceptions, given below.
2. If you are ill at the time of a quiz or right before an assignment is due and so you miss the one week window, or if there is a death in your immediate family, I will allow a late assignment or a make-up provided (a) you can give me proof of the circumstances, and (b) you let me know before the quiz is held or the assignment is due.
3. If you simply don't turn in the assignment or don't show up for the quiz, the only valid excuse is a note from a doctor given as proof that you were injured or ill at the due date to such an extent that it would be unreasonable for you to send me an email or leave a message letting me know of your illness or injury. In any other case, the result is a zero on the assignment.

Evaluation of Grades

Assignment	Total Points	Percentage of Final Grade
Midterm Exam	100	25%
Final Exam	100	25%
Labs	30	25%
Project	20	20%
Participation	5	5%

Grading Policy

Percent	Grade	Grade Points
93.4 - 100	A	4.00
90.0 - 93.3	A-	3.67
86.7 - 89.9	B+	3.33
83.4 - 86.6	B	3.00
80.0 - 83.3	B-	2.67
76.7 - 79.9	C+	2.33
73.4 - 76.6	C	2.00
70.0 - 73.3	C-	1.67
66.7 - 69.9	D+	1.33
63.4 - 66.6	D	1.00
60.0 - 63.3	D-	0.67

0 - 59.9	E	0.00
----------	---	------

More information on UF grading policy may be found at:

<https://catalog.ufl.edu/ugrad/current/regulations/info/grades.aspx>

Policy on Regrade Requests:

Any request for an actual regrade must be made in writing within one week of the time that the paper is returned, with no exceptions. All regrade requests must be prepared using a word processing program; a hard copy should then be submitted to me or a TA, along with the original graded work. On your regrade request, carefully describe why you feel that you were scored unfairly and/or incorrectly. Even if you discussed the grading issue orally with someone, the written discussion must be self-contained and be able to be evaluated based only on what is written on the paper.

Students Requiring Accommodations

Students with disabilities requesting accommodations should first register with the Disability Resource Center (352-392-8565, <https://www.dso.ufl.edu/drc>) by providing appropriate documentation. Once registered, students will receive an accommodation letter which must be presented to the instructor when requesting accommodation. Students with disabilities should follow this procedure as early as possible in the semester.

Course Evaluation

Students are expected to provide professional and respectful feedback on the quality of instruction in this course by completing course evaluations online via GatorEvals. Guidance on how to give feedback in a professional and respectful manner is available at <https://gatorevals.ua.ufl.edu/students/>. Students will be notified when the evaluation period opens, and can complete evaluations through the email they receive from GatorEvals, in their Canvas course menu under GatorEvals, or via <https://ufl.bluera.com/ufl/>. Summaries of course evaluation results are available to students at <https://gatorevals.ua.ufl.edu/public-results/>.

University Honesty Policy

UF students are bound by The Honor Pledge which states, "We, the members of the University of Florida community, pledge to hold ourselves and our peers to the highest standards of honor and integrity by abiding by the Honor Code. On all work submitted for credit by students at the University of Florida, the following pledge is either required or implied: "On my honor, I have neither given nor received unauthorized aid in doing this assignment." The Honor Code (<https://sccr.dso.ufl.edu/policies/student-honor-code-student-conduct-code/>) specifies a number of behaviors that are in violation of this code and the possible sanctions. Furthermore, you are obligated to report any condition that facilitates academic misconduct to appropriate personnel. If you have any questions or concerns, please consult with the instructor or TAs in this class.

Commitment to a Safe and Inclusive Learning Environment

The Herbert Wertheim College of Engineering values broad diversity within our community and is committed to individual and group empowerment, inclusion, and the elimination of discrimination. It is expected that every person in this class will treat one another with dignity and respect regardless of gender, sexuality, disability, age, socioeconomic status, ethnicity, race, and culture.

If you feel like your performance in class is being impacted by discrimination or harassment of any kind, please contact your instructor or any of the following:

- Your academic advisor or Graduate Program Coordinator
- Robin Bielling, Director of Human Resources, 352-392-0903, rbielling@eng.ufl.edu
- Curtis Taylor, Associate Dean of Student Affairs, 352-392-2177, taylor@eng.ufl.edu
- Toshikazu Nishida, Associate Dean of Academic Affairs, 352-392-0943, nishida@eng.ufl.edu

Software Use

All faculty, staff, and students of the University are required and expected to obey the laws and legal agreements governing software use. Failure to do so can lead to monetary damages and/or criminal penalties for the

individual violator. Because such violations are also against University policies and rules, disciplinary action will be taken as appropriate. We, the members of the University of Florida community, pledge to uphold ourselves and our peers to the highest standards of honesty and integrity.

Student Privacy

There are federal laws protecting your privacy with regards to grades earned in courses and on individual assignments. For more information, please see: <https://registrar.ufl.edu/ferpa.html>

Campus Resources:

Health and Wellness

U Matter, We Care:

Your well-being is important to the University of Florida. The U Matter, We Care initiative is committed to creating a culture of care on our campus by encouraging members of our community to look out for one another and to reach out for help if a member of our community is in need. If you or a friend is in distress, please contact umatter@ufl.edu so that the U Matter, We Care Team can reach out to the student in distress. A nighttime and weekend crisis counselor is available by phone at 352-392-1575. The U Matter, We Care Team can help connect students to the many other helping resources available including, but not limited to, Victim Advocates, Housing staff, and the Counseling and Wellness Center. Please remember that asking for help is a sign of strength. In case of emergency, call 9-1-1.

Counseling and Wellness Center: <http://www.counseling.ufl.edu/cwc>, and 392-1575; and the University Police Department: 392-1111 or 9-1-1 for emergencies.

Sexual Discrimination, Harassment, Assault, or Violence

If you or a friend has been subjected to sexual discrimination, sexual harassment, sexual assault, or violence contact the [Office of Title IX Compliance](mailto:title-ix@ufl.edu), located at Yon Hall Room 427, 1908 Stadium Road, (352) 273-1094, title-ix@ufl.edu

Sexual Assault Recovery Services (SARS)

Student Health Care Center, 392-1161.

University Police Department at 392-1111 (or 9-1-1 for emergencies), or <http://www.police.ufl.edu/>.

Academic Resources

E-learning technical support, 352-392-4357 (select option 2) or e-mail to Learning-support@ufl.edu. <https://lss.at.ufl.edu/help.shtml>.

Career Resource Center, Reitz Union, 392-1601. Career assistance and counseling. <https://www.crc.ufl.edu/>.

Library Support, <http://cms.uflib.ufl.edu/ask>. Various ways to receive assistance with respect to using the libraries or finding resources.

Teaching Center, Broward Hall, 392-2010 or 392-6420. General study skills and tutoring. <https://teachingcenter.ufl.edu/>.

Writing Studio, 302 Tigert Hall, 846-1138. Help brainstorming, formatting, and writing papers. <https://writing.ufl.edu/writing-studio/>.

Student Complaints Campus: https://www.dso.ufl.edu/documents/UF_Complaints_policy.pdf.

On-Line Students Complaints: <http://www.distance.ufl.edu/student-complaint-process>.