<div align="center">

Data Science Team
CSCI 4802/5802
Spring 2021

</div>

Instructor:

Rebecca Morrison
ECOT 820
`rebeccam@colorado.edu`

Class details:

Class meetings: Tu 5:30 - 6:45 on Zoom, passcode: 930481
Communication: Slack
Assignments: Canvas
Team website: codata.colorado.edu
Grader: TBA

<div align="center">

*****Email to instructor should be used as a last resort.*****

</div>

Course description

Data science is one of the fastest-growing sectors of our economy, and there is a great demand for data scientists with practical experience. In this course, students will hear from a variety of experts about a broad collection of topics. Student will apply data science techniques and machine learning algorithms to real-world problems—valuable experience, both for entering the workforce or remaining in academia. Students will work in small teams on semester-long projects of their choice. Project teams are responsible for attending class, submitting progress reports, and giving short presentations.

Topics

To accompany the prediction challenges and other activities, we will have short presentations on various topics. A possible and non-exhaustive list is as follows.

- **Basic concepts**: classification and regression, prediction vs causation, regularization and overfitting

- **Algorithms**: linear regression, logistic regression, support vector machines, boosting, decision trees and forests, neural networks, gradient and stochastic gradient descent

- **Practical techniques**: ensemble methods and aggregation, tradeoffs in regularization, parameter and hyperparameter tuning, data imputation techniques, cross-validation

- **Software and tools**: tutorials on modern data science software packages including scikit-learn, pandas, vowpal wabbit, and xgboost

- **Context and industry practice**: weekly presentations from practicing data scientists to see what techniques actually used in industry and academia, and which algorithms work well for which problems

Course objectives

In this course, students will learn how to:

- Apply data science techniques to real-world problems
- Define a problem statement and collect the appropriate data
- Navigate common issues such as labeling and storage
- Assess the reliability of the collected data
- Collaborate efficiently with others
- Clearly communicate their assumptions, methods, and results

Course schedule

Week 1 (1/19): Introductions
Week 2 (1/26): Lecture + discussion
Week 3 (2/2): Lecture + discussion
Week 4 (2/9): Lecture + discussion
Week 5 (2/16): Organizational Day, **Groups must be ≥ 4, decided by end of class**
Week 6 (2/23): Lecture + discussion, **First midterm report due 2/26**
Week 7 (3/2): Lecture + discussion
Week 8 (3/9): Lecture + discussion
Week 9 (3/16): Lecture + discussion
Week 10 (3/24): Lecture + discussion, "Spring Pause"
Week 11 (3/30): Lecture + discussion, **Second midterm report due 4/2**
Week 12 (4/6): Lecture + discussion
Week 13 (4/13): Student presentations
Week 14 (4/20): Student presentations
Week 15 (4/27): Student presentations
Finals week 5/4: **Final project due**

Lectures will be recorded and posted on Canvas, with permission of the speaker.

Course work and grading

Students will turn in possible talk questions, two midterm reports, and one final report. See below for descriptions. Grades will be determined based on the following: class participation (30%), midterm reports (10% each), and final project (50%).

Class participation
Class participation is critical—this is Data Science *Team!* Usually, the more informal and comfortable we feel, the better we understand new things and work together.

- You can engage in a variety of ways: asking questions to the speaker, asking and answering questions on Slack, pointing other students to helpful tools, etc.
- This class depends on external speakers, and class participation shows the speakers that we are engaged, interested, and respectful. Students are encouraged to ask questions during the talks; asking good questions takes thought and effort. To practice, students will turn in three questions about each talk on Canvas, due by the end of the following day.
- I encourage all of you to turn on your videos. We have a large class this semester, but it still makes a positive difference for me, your classmates, and the speakers.
- Please show up to class on time.

Midterm report 1 (The Proposal)

- Purpose: To make sure you are on track
- Format: Text (PDF), 1 report per group
- Length: 3-4 paragraphs
- Content:
  - **Summary**: Brief description of the tasks you have been involved in and are planning. Some examples might be: competitions, prediction tasks, and so on.
  - **Techniques**: Brief description of the techniques you've used.
  - **Goals**: Your goals for the rest of the semester, both in terms of activities (what you will do) and education (what you hope to learn).

Midterm report 2 (The Progress Report)

- Purpose: To make sure you are on track
- Format: Text (PDF), 1 report per group
- Length: 3-4 paragraphs
- Content:
  - **Summary**: Give a more detailed description of the tasks you have been involved in so far, with links to or mentions of specific data sets or problems. Describe which of your goals you have met already, and what you plan to do for the remainder of the semester.
  - **Techniques**: Brief description of the techniques you've used.
  - **Goals**: Your goals for the rest of the semester, both in terms of activities (what you will do) and education (what you hope to learn).

Final report

- Purpose: To assess your level of participation, effort, and learning
- Format: PDF or website, 1 report per group
- Length: For a PDF, roughly 3 pages single-spaced ($\approx$ 1500 words). For a website, you can decrease the word count by 500, but should still include the following content on the site.
- Content:
  - **Summary**: A 1-2 paragraph summary description of the tasks your group attempted and/or completed.
  - **Activities**: For each activity, give a detailed account of your approach and techniques, including descriptions of any hurdles you had to overcome. Include relevant plots and figures (not included in page count). If you participated in prediction competitions, include links to the leaderboard and/or a screenshot showing your score.
  - **Goals**: Describe whether you accomplished your goals from midterm report. And if not, why not?
  - **Attachments**: Include any relevant code or other digital artifacts (links to online repositories preferred but not required).

More ideas about projects:

- Work in groups ≥ 4
- Branch one of your existing research projects (IS, thesis, dissertation)
- Pick a method you always wanted to learn, and find a way to learn it.
- Pick a topic you think is cool and replicate an existing DS analysis.
- Other!

Day 1 checklist:

- Join Slack and say hello.
- Make a github account if you don't have one.
- Contribute to #chat thread

    - What techniques seem like magic?
    - What methods are intimidating?
    - Who are your company/professor favs? (We can ask them for a lecture or AMA.)

Syllabus Statements

Classroom Behavior

Both students and faculty are responsible for maintaining an appropriate learning environment in all instructional settings, whether in person, remote or online. Those who fail to adhere to such behavioral standards may be subject to discipline. Professional courtesy and sensitivity are especially important with respect to individuals and topics dealing with race, color, national origin, sex, pregnancy, age, disability, creed, religion, sexual orientation, gender identity, gender expression, veteran status, political affiliation or political philosophy. For more information, see the policies on [classroom behavior](#) and the [Student Code of Conduct](#).

Requirements for COVID-19

As a matter of public health and safety due to the pandemic, all members of the CU Boulder community and all visitors to campus must follow university, department and building requirements, and public health orders in place to reduce the risk of spreading infectious disease. Required safety measures at CU Boulder relevant to the classroom setting include:

- maintain 6-foot distancing when possible,

- wear a face covering in public indoor spaces and outdoors while on campus consistent with state and county health orders,

- clean local work area,

- practice hand hygiene,

- follow public health orders, and

- if sick and you live off campus, do not come onto campus (unless instructed by a CU Healthcare professional), or if you live on-campus, please alert [CU Boulder Medical Services](#).

Students who fail to adhere to these requirements will be asked to leave class, and students who do not leave class when asked or who refuse to comply with these requirements will be referred to [Student Conduct and Conflict Resolution](#). For more information, see the policies on [COVID-19 Health and Safety](#) and [classroom behavior](#) and the [Student Code of Conduct](#). If you require accommodation because a disability prevents you from fulfilling these safety measures, please see the "Accommodation for Disabilities" statement on this syllabus.

All students who are new to campus must complete the [COVID-19 Student Health and Expectations Course](#). Before coming to campus each day, all students are required to complete the [Buff Pass](#).

Students who have tested positive for COVID-19, have symptoms of COVID-19, or have had close contact with someone who has tested positive for or had symptoms of COVID-19 must stay home. In this class, if you are sick or quarantined, **please watch the recorded videos and talk with your peers about the missed class material**.

ACCOMMODATION FOR DISABILITIES

If you qualify for accommodations because of a disability, please submit your accommodation letter from Disability Services to your faculty member in a timely manner so that your needs can be addressed. Disability Services determines accommodations based on documented disabilities in the academic environment. Information on requesting accommodations is located on the [Disability Services website](#). Contact Disability Services at 303-492-8671 or [dsinfo@colorado.edu](#) for further assistance. If you have a temporary medical condition, see [Temporary Medical Conditions](#) on the Disability Services website.

PREFERRED STUDENT NAMES AND PRONOUNS

CU Boulder recognizes that students' legal information doesn't always align with how they identify. Students may update their preferred names and pronouns via the student portal; those preferred names and pronouns are listed on instructors' class rosters. In the absence of such updates, the name that appears on the class roster is the student's legal name.

HONOR CODE

All students enrolled in a University of Colorado Boulder course are responsible for knowing and adhering to the Honor Code. Violations of the policy may include: plagiarism, cheating, fabrication, lying, bribery, threat, unauthorized access to academic materials, clicker fraud, submitting the same or similar work in more than one course without permission from all course instructors involved, and aiding academic dishonesty. All incidents of academic misconduct will be reported to the Honor Code ([honor@colorado.edu](#)); 303-492-5550). Students found responsible for violating the academic integrity policy will be subject to nonacademic sanctions from the Honor Code as well as academic sanctions from the faculty member. Additional information regarding the Honor Code academic integrity policy can be found at the [Honor Code Office website](#).

SEXUAL MISCONDUCT, DISCRIMINATION, HARASSMENT AND/OR RELATED RETALIATION

The University of Colorado Boulder (CU Boulder) is committed to fostering an inclusive and welcoming learning, working, and living environment. CU Boulder will not tolerate acts of sexual misconduct (harassment, exploitation, and assault), intimate partner violence (dating or domestic violence), stalking, or protected-class discrimination or harassment by members of our

community. Individuals who believe they have been subject to misconduct or retaliatory actions for reporting a concern should contact the Office of Institutional Equity and Compliance (OIEC) at 303-492-2127 or cureport@colorado.edu. Information about the OIEC, university policies, anonymous reporting, and the campus resources can be found on the OIEC website.

Please know that faculty and graduate instructors have a responsibility to inform OIEC when made aware of incidents of sexual misconduct, dating and domestic violence, stalking, discrimination, harassment and/or related retaliation, to ensure that individuals impacted receive information about options for reporting and support resources.

RELIGIOUS HOLIDAYS

Campus policy regarding religious observances requires that faculty make every effort to deal reasonably and fairly with all students who, because of religious obligations, have conflicts with scheduled exams, assignments or required attendance. In this class, **please come talk to me during office hours with at least a week's notice, and we can figure something out**.

See the campus policy regarding religious observances for full details.

Updated December 9, 2020