CEE 498EDS: Environmental Data Science, Fall 2020

This document (permalink) was automatically generated from uiuc-cee-eds/syllabus@7636e3b on March 27, 2020.

Basic Course Information

- Department: Civil and Environmental Engineering
- Title: CEE 498EDS: Environmental Data Science
- Credits: 3 for Undergraduates, 4 for Graduate Students
- Semester: Fall 2020
- Meeting time and location: TBD

Basic Instructor Information

- Instructor: Prof. Christopher Tessum, PhD, he/him/his
- Office: 3213 Newmark Civil Engineering Laboratory
- Office hours: TBD
- Please contact me through Canvas, rather than through email
- Names and contact information for teaching assistants: TBD

Description of the course

Welcome to Environmental Data Science! This semester, you will learn to leverage data to study environmental problems, identify patterns, and make actionable insights. This course combines training in digital and computer tools—including distributed computing, exploratory data analysis, and statistical modeling and deep learning—with application of those tools to environmental issues.

By the end of the semester, you will be able to: 1. Use software tools for data processing and visualization, machine learning, and deep learning to 2. Retrieve, manipulate, and analyze environmental data; and 3. Make inferences and predictions about the environment.

This course will help you to gain the skills and tools necessary to make the most of the great increases in the amount and quality of environmental data that is being collected and stored.

- Format of the course

Instructor's Philosophy

- Conceptual structure used to organize the course, why it is organized the way it is
- Philosophy of teaching and learning

Prerequisites

CEE 202; CS 101 or equivalent.

Course Requirements and Assessment Overview

- Nature of assignments and exams (details can be in a separate handout)
- Deadlines and test dates
- Description of grading procedures
- Description of how grades will be assigned, components of final grade, weights, grading scale

Learning Resources

- Textbook and other required materials
- Supplemental readings, etc.
- Campus resources—tutoring, writing, counseling, etc.
- Estimate of student work load
- Hints for how to study, take note, etc.
- Availability of past exams, etc.

Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython

Hands-On Machine Learning with Scikit-Learn & TensorFlow

Course Policies

- University-based policies—academic integrity
- Course specific polices—late assignments, make-up exams, attendance, participation, etc.
- Statement on accommodations
- Important dates such as drop dates, final exam date, etc.

Course Schedule

- Sequence of course topics with tentative (or firm) dates

- Due dates for assignments, exams

- Preparations or readings

- Week 1: Open Reproduceable Science

    - Reading: https://www.earthdatascience.org/courses/intro-to-earth-data-science/open-reproducible-science/get-started-open-reproducible-science/
    - In class lecture/tutorial:
        - FAIR Principles
        - Git/Github
        - Jupyter

- Week 2: Data science for the environment

    - Reading:
        - https://arxiv.org/pdf/1906.05433.pdf
        - https://pangeo.io/use_cases/index.html
        - https://www.kaggle.com/
        - https://earthengine.google.com/case_studies/
        - Others from specific environmental disciplines
    - In class:
        - Students present on potential projects

- Week 3: Programming review

    - Reading/Lecture: https://developers.google.com/edu/python
    - In class:

        - Python & Jupyter exercises

- Week 4: Big data

    - In class:
        - Lecture:
            - Cloud / High-performance computing
            - Pangeo
            - Earth engine
        - Practice and discussion
        - Choose project groups and topics

- Week 6: Exploratory data analysis (EDA)

    - Reading/Lecture:
        - https://mlcourse.ai/articles/topic1-exploratory-data-analysis-with-pandas/
        - https://mlcourse.ai/articles/topic2-visual-data-analysis-in-python/
        - https://mlcourse.ai/articles/topic2-part2-seaborn-plotly/
        - And/or accompanying video lectures
    - In class:
        - Statistics review lecture
        - Students present initial EDA for their projects

- Week 7: Geospatial data

    - Reading: https://github.com/geopandas/scipy2018-geospatial-data
    - In class lecture:
        - raster vs. vector formats
        - joins and boolean operations

- Week 8: Spatial statistics

    - Reading: http://pysal.org/notebooks/intro
    - In class dicussion:
        - Applicability of spatial statistics to student projects

- Week 9: Mid-term project presentations

    - Students should be able to access, characterize, and visualize the data for their projects.

- Week 10: Supervised learning

    - Reading/lecture:
        - https://developers.google.com/machine-learning/crash-course/framing/video-lecture
        - https://developers.google.com/machine-learning/crash-course/descending-into-ml/video-lecture
        - https://developers.google.com/machine-learning/crash-course/reducing-loss/video-lecture
        - https://developers.google.com/machine-learning/crash-course/first-steps-with-tensorflow/toolkit
        - https://developers.google.com/machine-learning/crash-course/generalization/video-lecture
        - https://developers.google.com/machine-learning/crash-course/training-and-test-sets/video-lecture

        - https://developers.google.com/machine-learning/crash-course/validation/check-your-intuition
    - In class:
        - Application to student projects

- Week 11: Unsupervised learning

    - Reading/Lecture: https://mlcourse.ai/articles/topic7-unsupervised/
    - In class:
        - Application to student projects

- Week 12: Deep learning

    - Reading/lectures:
        - https://developers.google.com/machine-learning/crash-course/introduction-to-neural-networks/video-lecture
        - https://developers.google.com/machine-learning/crash-course/training-neural-networks/video-lecture
        - https://developers.google.com/machine-learning/crash-course/multi-class-neural-networks/video-lecture
    - In class:
        - Hyperparameter optimization
        - Application to student projects

- Week 13-16: Projects