Visualization for Machine Learning

Spring 2024

Teaching Staff

- Instructor: Claudio Silva (csilva@nyu.edu); webpage
- Section Leader: Erin McGowan (erin.mcgowan@nyu.edu)
- Teaching Fellow: Vitoria Guardieiro (vitoria.guardieiro@nyu.edu)
- Grader: Rithvik Guruprasad (rg4361@nyu.edu)

Location

Instruction Mode: In-Person

Meeting Times:

- DS-GA 3001.001 (Lecture) Thursdays 6:45pm-8:25pm Classroom: 31 Washington Pl (Silver Ctr) Room 520
- DS-GA 3001.002 (Lab) Fridays 10:15am-11:05am Classroom: 31 Washington Pl (Silver Ctr) Room 520

Discord

- Class Discord: https://discord.gg/6pFj8dMK.
- All our communications happen here.
- Anything related to assignments, grading, etc., send a direct message to Grader(s).
- Anything related to labs / coding contact Erin or Vitoria.
- Feel free to direct message for anything else.

Course Prerequisites

- Solid programming expertise.
- The coursework includes extensive programming with JavaScript and D3.js. To be covered in the labs.
- We will also expect students to be able to program in Python.
- Basic knowlege of machine learning.

Course Description

- Our course is based on foundations of visual analytics, which is an area of data visualization that is concerned with improving the human analytic process.
- Visual analytics is concerned with combining automated processes with human-driven processes that are built around data visualization: visual representations of data, and ways to interact with data.

Course Objectives

- This course is designed to sharpen a student's knowledge of visualization.
- We aim to make the student a more effective data scientist.
- The course has a major project, which will help the student develop research skills.

Course Structure

- The course include lectures and labs. We will strive to have hands-on sessions to complement theoretical materials.
- Starts with a short primer on visualization.
- Then we will cover techniques for visualizing model assessment, white-box and black-box machine learning explainers, and dimensionality reduction.
- Second half of the course has more advanced topics, including Topology Data Analysis and techniques for visualizing deep neural networks.

Reading Material

- There is no textbook for the course.
- We will have suggested reading materials for each class.
- Supplemental readings to be used as reference material:
  - Data Visualization Curriculum, Heer, link
  - A Course in Machine Learning, Daume, link
  - Interpretable Machine Learning, Molnar, link
  - Introduction to Machine Learning, Bernard, link
  - Deep Learning, Goodfellow et al, link

Research Project

- Course includes a substantial research project.
- Projects are expected to be pursued in groups of 2-3.
- Once the group is finalized, students cannot change or separate their groups throughout the semester.

Course Assessment

- Assignments (50%)
- Project Proposal (4-page writeup): 10%
- Project Updates (1-page writeup): 10%
- Full Project (up to 8-page writeup): 25%
- Class Participation: 5%

Late Submissions

Late submissions of assignments will be penalized as follows:

- A standard deduction rate of 20% per day. It means that after 5 days of being late, your assignment will have a maximum grade of 0 (zero).
- You will have a one-time exception for submitting assignments late (up to 5 days late).

Academic Integrity

- All students are expected to do their own work.
- See detailed policy on class syllabus.
- Also, here is a link to the GSAS statement on Academic Integrity.

AI policy

- You can use generative AI tools to do the assignments in this class.
- If you use an AI to guide you in completing an assignment, you have to disclose which parts were generated by the AI.

NYU Academic Calendar

- link to NYU Academic Calendar
- This course does not have a final exam, but there will be a final project presentation.
- Also, please pay attention to notable dates such as Add/Drop, Withdrawal, etc.

End of Course Logistics

- Any questions?

BREAK – 5 minutes

Self Introduction for VisML 2024

slides