New York University Tandon School of Engineering
Computer Science and Engineering

CS-GY 9223 - I (24680)
Visualization: Connections with Machine Learning

Th 3:20PM-5:50PM

Spring 2020

Course Instructors

Cláudio T. Silva
Tandon: 370 Jay Street, Brooklyn, room 1153, 646 997-4093
Center for Data Science: 60 5th Ave, Manhattan, room 701, 212-998-3091
csilva@nyu.edu

Office hours: TBA, or by appointment.

Teaching Assistants

Gromit Yeuk-Yin Chan
370 Jay Street, 11th floor
gromit.chan@nyu.edu

Jorge Piazentin Ono
370 Jay Street, 11th floor
jorgehpo@nyu.edu

Course Pre-requisites

You should have a solid foundation in **either** data visualization or machine learning. Ideally, both! But ultimately, your background will determine the type of project that you work on. If you are only well-versed in machine learning, then you will need to learn the basics behind data visualization. On the other hand, if you are only a data visualization expert, then you will need to learn the basic components of machine learning.

During lectures, we will be going in detail, beyond the research papers, on topics in visualization and machine learning that are foundational, but traditionally not covered in the research literature. This will be necessary, but not sufficient, for you in fully understanding one of the two areas you may not have proper background in.

Course Description

This course is a research-oriented course on topics related to visual analytics and machine learning. Visual analytics is an area of data visualization that is concerned with improving a human's analytic process, or how one makes sense of data for a given problem: understanding, reasoning, and making decisions about a provided dataset, and a given problem domain. Visual analytics, in particular, 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. Given the rapid growth in machine learning in the last decade, research in visual analytics has witnessed similar growth in leveraging machine learning in a variety of ways. This course will cover topics that live at the interface of visual analytics and machine learning, exposing you to the basics of both fields, how machine learning can be used to enhance visual analytics, and how visual analytics can help machine learning.

Acknowledgement: This course is based on the Visual Analytics and Machine Learning course designed by **Professor Matthew Berger** (Vanderbilt University).

Course Objectives

This course is designed to sharpen a student's knowledge of visualization and machine learning, and how the two areas interact. It is expected that the student will be a more effective data scientist by being fluent on the connections between the two areas. It is also designed around a major project, which will help the student develop research skills.

Course Structure

The course will primarily be lecture-based.

In addition to a primer on visualization, in this course we will study four primary areas that consider visual analytics and machine learning.

Mixed-Initiative Visual Exploration: One of the main goals of data visualization is to enable the human to better understand their data through visual exploration. Through leveraging machine learning techniques, it is possible to improve this form of exploration, establishing an effective blend of automated analyses provided by a learning technique, and what to expose to the user for determining their interactions.

Visual Analytics for Understanding Models: The growth in machine learning has been accompanied by an equally pressing demand to understand machine learning models, e.g. to provide interpretable and explainable models. Visual analytics plays an important role in helping the user understand machine learning models, be it through understanding the training process of a model, understanding the parameter space of a model, understanding features learned by a model from a given set of data, or understanding the outputs produced by a model.

Visual Analytics for Training Models: In machine learning, the training of a model is traditionally accomplished by a human identifying a training dataset, and then training the model, sometimes using a validation set to tune hyperparameters. Opening this process up, however, can enable visual analytics techniques to improve how models are trained, either through improving how humans annotate data used for training, or incorporating the human insights directly into the model-building process.

Machine Learning for Visualization: Machine learning can also be used as a means to improve the visualization process itself. This can range from methods for recommending visualizations, automating (or semi-automating) the creation of visualizations from a provided dataset, or constructing learning models for visualization techniques.

Readings

There is no textbook for the course - all lectures will be based on papers we have listed in the papers section of the website. The schedule section lists papers that will be covered during each lecture. It is expected that, prior to the lecture, you have read the corresponding papers.

Here are supplemental readings on visual analytics and machine learning:

1. Machine Learning: a Probabilistic Perspective, Kevin Patrick Murphy, MIT Press, 2012.
2. Interpretable Machine Learning: A Guide for Making Black Box Models Explainable, Christoph Molnar, https://christophm.github.io/interpretable-ml-book/
3. Visualization Analysis and Design, Tamara Munzner, A K Peters Visualization Series. CRC Press, 2014.
4. Deep Learning, Ian Goodfellow, Yoshua Bengio and Aaron Courville, MIT Press, 2016 http://www.deeplearningbook.org/

Research Project

The bulk of the course will be devoted to a semester-long research project. Please see the project section of the course for more details. As part of the project, you will be expected to reproduce prior work, as well as implement a proposed research idea of your choosing (selected in consultation with the instructor). Moreover, you will be expected to demonstrate both the prior work, and your final research project, to the class during lectures. Again, please see project for additional details.

Course Assessment

- Project Proposal: 15%

    - Presentation: 5%
    - Proposal Document: 10%
- Related Work: 30%
    - Presentation and Demonstration: 10%
    - Source Code and Documentation: 20%
- Project Updates: 5% (total)
- Full Project: 40%
    - Presentation: 10%
    - Full Submission: 30%
- Class Participation: 10%

Course Schedule

The course schedule is tentative and will be adjusted along the way.

Lecture 1-2: Introduction to Visual Analytics, Visualization, and Machine Learning

Lecture 3: Project Proposals

Lecture 4-5: Mixed-Initiative Visual Exploration

Lecture 5-7: Visual Analytics for Model Understanding

Lecture 8: Project Baseline Presentations

Lecture 9-10: Visual Analytics for Model Understanding

Lecture 11-13: Visual Analytics for Model Training

Lecture 14-15: Machine Learning for Visualization

Lecture 16: Project Presentations

Moses Center Statement of Disability

If you are a student with a disability who is requesting accommodations, please contact New York University's Moses Center for Students with Disabilities (CSD) at 212-998-4980 or mosescsd@nyu.edu. You must be registered with CSD to receive accommodations. Information about the Moses Center can be found at www.nyu.edu/csd. The Moses Center is located at 726 Broadway on the 3rd floor.

NYU School of Engineering Policies and Procedures on Academic Misconduct

– Complete Student Code of Conduct can be found here.

A. Introduction: The School of Engineering encourages academic excellence in an environment that promotes honesty, integrity, and fairness, and students at the School of Engineering are expected to exhibit those qualities in their academic work. It is through the process of submitting their own work and receiving honest feedback on that work that students may progress academically. Any act of academic dishonesty is seen as an attack upon the School and will not be tolerated. Furthermore, those who breach the School's rules on academic integrity will be sanctioned under this Policy. Students are responsible for familiarizing themselves with the School's Policy on Academic Misconduct.

B. Definition: Academic dishonesty may include misrepresentation, deception, dishonesty, or any act of falsification committed by a student to influence a grade or other academic evaluation. Academic dishonesty also includes intentionally damaging the academic work of others or assisting other students in acts of dishonesty. Common examples of academically dishonest behavior include, but are not limited to, the following:

1. Cheating: intentionally using or attempting to use unauthorized notes, books, electronic media, or electronic communications in an exam; talking with fellow students or looking at another person's work during an exam; submitting work prepared in advance for an in-class examination; having someone take an exam for you or taking an exam for someone else; violating other rules governing the administration of examinations.
2. Fabrication: including but not limited to, falsifying experimental data and/or citations.
3. Plagiarism: intentionally or knowingly representing the words or ideas of another as one's own in any academic exercise; failure to attribute direct quotations, paraphrases, or borrowed facts or information.
4. Unauthorized collaboration: working together on work meant to be done individually.
5. Duplicating work: presenting for grading the same work for more than one project or in more than one class, unless express and prior permission has been received from the course instructor(s) or research adviser involved.
6. Forgery: altering any academic document, including, but not limited to, academic records, admissions materials, or medical excuses.

NYU School of Engineering Policies and Procedures on Excused Absences

– Complete policy can be found here.

A. Introduction: An absence can be excused if you have missed no more than 10 days of school. If an illness or special circumstance has caused you to miss more than two Lectures of school, please refer to the section labeled Medical Leave of Absence.

B. Students may request special accommodations for an absence to be excused in the following cases:

1. Medical reasons
2. Death in immediate family
3. Personal qualified emergencies (documentation must be provided)
4. Religious Expression or Practice

Deanna Rayment, deanna.rayment@nyu.edu, is the Coordinator of Student Advocacy, Compliance and Student Affairs and handles excused absences. She is located in 5 MTC, LC240C and can assist you should it become necessary.

NYU School of Engineering Academic Calendar

– Official calendar can be found here.

This course does not have a final exam.

Also, please pay attention to notable dates such as Add/Drop, Withdrawal, etc. For confirmation of dates or further information, please contact Susana: sgarcia@nyu.edu