DS-GA 3001: Tools and Techniques for Machine Learning (Fall 2021)

Instructor: David S. Rosenberg

Course description

This course deals with a range of topics that come up when applying machine learning in practice. Much of the course builds towards an understanding of how to handle machine learning with interventions, in contexts such as counterfactual learning, reinforcement learning, and causal inference. Techniques for inverse propensity weighting and for attaining double robustness with control variates will be given special attention, as these have applications throughout this course and beyond. Along the way, we'll discuss covariate shift, the exploration/exploitation tradeoff, and supervised learning with black box (including non-differentiable) loss functions. Finally, we will cover probability forecast calibration and methods for interpreting machine learning models.

Prerequisites

- DS-GA 1003: Machine Learning or equivalent.
- DS-GA 1002: Probability and Statistics or equivalent.
- Comfort with conditional expectations, conditional probability modeling, basic Bayesian statistics, hypothesis testing and confidence intervals.
- Python programming required for most homework assignments.

Topics and rough schedule

The following is a tentative schedule for the semester:

- Week 1: Conditional expectation and variance decomposition
- Week 2: Response bias, inverse propensity weighting, and self-normalization
- Week 3: Regression imputation, covariate shift
- Week 4: Control variates, double robustness, average treatment effects
- Week 5: Conditional average treatment effects (CATE), bandits, Thompson sampling

- Week 6: Contextual bandits, counterfactual evaluation
- Week 7: Counterfactual learning
- Week 8: Policy gradient for bandits and contextual bandits
- Week 9: Variance reduction for policy gradient, supervised learning with black box losses
- Week 10: Reinforcement learning, proper scoring rules
- Week 11: Calibrated probabilities
- Week 12: Feature importance and global interpretability
- Week 13: Local interpretation, LIME, Shapley values
- Week 14: SHAP

More information and motivation for these topics can be found in an earlier draft of the syllabus here.

Course Requirements and Evaluation

- **(70%) Homework**: 5 homework assignments; mix of model building and written mathematical exercises to reinforce the main concepts.
- **(30%) Weekly Quizzes**: Concept-check quizzes that reinforce the main ideas from lectures and lab, which students may use any resources to complete.

Remote Access

- Lecture and Lab sessions will be streamed and recorded via Zoom for at least the month of September.

Academic Integrity Policy:

The course conforms to NYU's policy on academic integrity for students.

Moses Statement

Academic accommodations are available for students with disabilities. The Moses Center website is `http://www.nyu.edu/csd`. Please contact the Moses Center for Students with Disabilities (212-998-4980 or mosescsd@nyu.edu) for further information. Students who are requesting academic accommodations are advised to reach out to the Moses Center as early as possible in the semester for assistance.