## COMS-W4995 Applied Machine Learning Project

The project assignment gives students an opportunity to apply different aspects of Machine Learning covered in the classroom to a real-world application. Through this project, students will get hands-on experience solving a Machine Learning problem including data analysis, visualization and applying machine learning models to develop actionable insights. The project will have three deliverables that are required to be submitted by the team throughout the course term. The students are randomly assigned to groups sized 4-6 and would be allowed one submission per group.

### Project Deliverable #1 - Project Proposal (due 02/23/2023)

*Submit a 1 page proposal that should include:*

1. Background and context to the problem statement.
2. Identification and description of the data set(s) you are planning on using along with their source. Some popular data sources are provided below:
   1. [https://www.kaggle.com/ (Links to an external site.)](https://www.kaggle.com/)
   2. Publicly available data of cities such as New York, Chicago, San Francisco, Washington DC etc.
   3. Public datasets on Amazon [https://registry.opendata.aws/ (Links to an external site.)](https://registry.opendata.aws/)
   4. Public datasets on Google [https://www.google.com/publicdata/directory (Links to an external site.)](https://www.google.com/publicdata/directory)
   5. OpenML datasets: [https://www.openml.org/s/88/data (Links to an external site.)](https://www.openml.org/s/88/data)
3. Proposed ML techniques you are proposing on applying to solve the problem

### Project Deliverable #2 - Data Analysis and Visualization (due 03/30/2023)

*Submit a presentation with 8-10 slides with notes covering*:

1. Initial data exploration
2. Cleaning and sampling
3. Insights from data exploration and
4. Machine Learning techniques proposed to be implemented

### Project Deliverable #3 - Report & Code (due 04/27/2023)

Each team is expected to submit a 3-page final report that talks about their project in detail. The teams are also expected to submit their python code through the github classroom that would be shared later in the semester.