

## Midterm project presentation (15 points total)

The purpose of the semester-long project is to give you hands-on experience working with the machine learning pipeline and creating a course project that you can put on your resume with a public github repository. For the midterm presentation, you should select a dataset and have an interesting/valuable/relevant classification or regression problem in mind to solve, you also should be done with EDA, splitting, and preprocessing.

**The project presentations are during the week of October 21st.** The presentations are in-person, and we will send out a sign up sheet in early October. We will have seven 2-hour sessions (84 slots for 78 students). You will have 6 minutes to present, 2-3 minutes for questions, and 1 min to change speakers. Please make sure to be there for your whole session on the day you present so you can listen to your fellow students' presentations and we encourage you to ask questions.

Each student will have a mentor TA assigned. Your mentor TA will reach out to you via email in early October. **Please meet with the mentor TA and show the draft of your slides a couple of days before the presentations.** If you have any questions about the project before your mentor TA reaches out, please post on the course forum or come to office hours.

## Requirements

### Title slide (1 point)

Your title slide should include the title of your presentation, your name, institute, the date of your presentation, and a link to your GitHub repository (make sure it's public).

### Intro slide (2 points)

The intro slide should explain the problem you want to solve, why this is important, whether it is a regression or classification problem, where you got the data from, and how the data was collected.

### EDA slide(s) (7 points)

Show us what you think were the most interesting, unexpected, or important insights that you gained during EDA. Present at least three different informative visualization types. Do not create only three figures. Create as many figures you can to better understand your dataset and calculate summary statistics as well. Then select three figures that you want to show during the presentation. Explain why you chose the figures! Why are they interesting, unexpected, important?

### Splitting and preprocessing slide(s) (5 points)

Describe briefly how you split the data and why. What preprocessors did you use and why? How many features and data points do you have before and after preprocessing. If you have missing

values, describe the basic properties of the missing values (e.g., what fraction of the points have missing values, which features, what fraction of the features are missing, etc).

**Keep your time (you have a 30 sec grace period but we will deduct 2 points for every minute you go overtime; we round up to the minute)**

Please practice your presentation out loud at least once and make sure you can fit everything into 6 minutes. Moderators in conferences usually do not let you go over time. They will make noise, come to the stage and stand next to you looking at you, or try to signal in other ways that it's time to wrap up. It is rude to go overtime and delay the presentations of others.

Make sure your slides are not too busy! Your slides should not contain full sentences or paragraphs, only your key points as phrases. Large images and short text is the way to go.

Make sure the figures are readable from a distance! If the text on your figures is too small, your audience won't be able to see what's on your figure.

There is no need to add everything to your slides. We will record your presentation and grade what you say as well.

Make sure to speak out loud and with a good pace (don't rush), and keep eye contact with your audience. A good presentation is about story-telling and engaging your audience. If you feel compelled to speak fast and rush so you can go through all of your slides in time, don't. Less is more. Cut out a slide or two to save time. If you feel compelled to read from your notes, don't. Most people in the audience are annoyed by this.