Homework Three

MSDS Summer 2021

- Submit code via GitHub Classroom using Markdown Cells to **clearly** indicate which code answers which question and to answer short answer questions.
- Please use a Markdown cell to write your name at the top of the notebook or you will lose points.
- This is due on August 10th at Midnight Pacific time. This will be graded for correctness, but **not** model performance.

For this homework, you will perform the tasks of classification and localization using the Stanford cars dataset: `https://www.kaggle.com/eduardo4jesus/stanford-cars-dataset`

You can find a good example of how to process this data here: `https://www.kaggle.com/eduardo4jesus/stanford-cars-dataset-a-quick-look-up`

You may use any architecture/approach you would like, including a CNN from scratch or finetuning a model with pre-trained weights. You **must**, however, use spatial augmentation during training (such as random rotations)

1. Create a model that **only** classifies the cars. Make a note of the accuracy of this model.

2. Use a similar model architecture to both classify the cars and predict the bounding box by summing the loss functions for each task: `CrossEntropyLoss` and `L1Loss`. Compute the accuracy for the classification task and the IoU score for the bounding box prediction task. How does the accuracy of this model compare to the **just** classification model from before?

3. Try training this model with different weights on the two loss functions. For instance you might try: `CrossEntropyLoss` + 20 $\times$ `L1Loss`. Observe how the performance of the model changes at the two tasks as you adjust the weight. Is there a sweet spot?