Homework Two

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.
- This is due on August 3rd at Midnight Pacific time. This will be graded for correctness, but **not** model performance.

For part of this homework, you will create a Twin Neural Network (sometimes referred to as a "Siamese Neural Network") to evaluate whether two questions are the same or not using the dataset here: `https://www.kaggle.com/c/quora-question-pairs/`

1. Write a Vanilla RNN module **without** using the PyTorch RNN layer. (Hint: you will need to use at least three linear layers!)

2. Create a Twin Neural Network using GRU or LSTM layers to evaluate whether Quora question pairs are the same or not. Some hints:

- See here

  `https://blog.mlreview.com/implementing-malstm-on-kaggles-quora-question-pairs-competition-8b31b0b16a07`

  for a Keras implementation.
- You may need to use a loss function not provided by PyTorch, like contrastive loss

  `https://medium.com/@maksym.bekuzarov/losses-explained-contrastive-loss-f8f57fe32246`

For a bonus point, submit your model to the Kaggle competition, and include a screenshot of your submission in your repo.