CS288 HW4: Finetuning Large Language Models

Due: April 3rd

Overview

This homework will be focused on (1) using Huggingface transformers to load state-of-the-art language models, (2) generating text from language models, and (3) finetuning language models on a specific dataset (machine translation, sentiment analysis).

Background Reading

You may wish to read the GPT-2 paper. You may also want to read this post on how to generate from language models in Huggingface.

Assignment

Notebook `https://colab.research.google.com/drive/1L1VdbzUuPVky-bdMJfGVHi-y7YHUCYQP?usp=sharing`:

- Complete the the notebook and save the generated files
- Write an error analysis of the machine translation model (details in notebook)

Debugging and Piazza

Please post questions about this homework in the corresponding homework threads on Piazza. Please limit your use of private questions unless absolutely necessary, as we do not have the staffing capacity to help with all requests for code debugging. If you believe you have found a legitimate bug in the solution, however, please contact us immediately.

Submission to Gradescope

Please submit the assignment to: `https://www.gradescope.com/courses/500812` (code: V5X87D)

- `sst_predictions.txt`
- `mt_predictions.txt`
- `hw4.ipynb`
- `report.pdf`