Homework 5 Paper Questions: LLM and SFT

Due: Friday, December 12th at 11:59 pm

Deliverables. Submit a PDF of your write-up to Gradescope under *HW5 Paper*.

Overview

This homework provides a structured introduction to **Supervised Fine-Tuning (SFT)** in Large Language Models (LLMs). You are encouraged to skim related works but focus on identifying the **problem**, **current solutions**, **proposed approach**, and the **limitations**.

The goal is not to memorize details but to practice reading research papers methodically. Most papers follow a common structure: they first motivate a **problem**, then review **related work**, describe the **proposed method and key insights**, present the **methods and experiments**, and finally discuss **limitations and conclusions**.

You will apply this process while reading the following paper:

- **SFT Data Composition:** *How Abilities in Large Language Models are Affected by Supervised Fine-tuning Data Composition*, ACL 2024.

Paper Questions

Q1. [**Motivation**] What is task the paper is trying to solve? How does this task differ from single-task learning and its challenges? [**Section: 1 – Introduction**]

Q2. [**Related Work**] What do the authors mean when they mention "aligning LLMs to human intent"? What kinds of datasets do people perform supervised fine-tuning on to align LLMs with human intent, and what do the data actually look like? Search up some of the referenced datasets in this section and describe 1-2 concrete examples. [**Section: 2 – Related Work**]

Note: We are not expecting you to read all the cited work. Few sentences will be sufficient.

Q3. [**Key Insights / Contributions**] The authors organize the paper around four major research questions (RQs). For each RQ listed below, briefly summarize the authors' findings based on the Experiments. [**Section: 3 - Experiments**]

RQ1: How do performance of individual abilities change as the amount of SFT data increases?

RQ2: How do performance change when different abilities are trained together using mixed SFT data? Why does the same strategy that work for low resource setting not effective to the setting where we have enough data?

RQ3: What primarily drives the conflicts described in RQ2: the *total data size* or the *data ratio*? Under what conditions does the ratio matter?

RQ4: How do different supervised fine-tuning strategies (multi-task, sequential, mixed sequential, DMT) influence ability trade-offs? Which works best?

Q4. [**Experimental Setup**] How do the authors design their experiments to answer the four research questions? In your answer, describe: (a) the model and the model size that was trained (b) the specific datasets used to represent each ability, (c) the SFT strategies and data manipulations the authors vary. [**Section: 3 – Experiments**]

Q5. [**Proposed Approach**] Describe the **Dual-stage Mixed Fine-tuning (DMT)** strategy proposed in the paper. How does it help mitigate the issues identified in previous RQs? [**Section: 3.5 – RQ4**]

Q6. [**Open-Ended**] In Section 2, the authors note that "recent initiatives have generated SFT datasets from user logs within proprietary LLM platforms," meaning that many modern SFT datasets now contain *LLM-generated responses*. Answer the following:

(a) **What are the potential advantages and disadvantages of using LLM-generated data (e.g., assistant responses from user logs) for supervised fine-tuning?**

(b) **Suppose you have two specialized datasets (e.g., math and coding) and you do not have have access to other data. How could LLM-generated data be used to help reduce catastrophic forgetting?**