Chaklam Silpasuwanchai, Todsavad Tangtortan Updated 22 Feb 2025
AT82.05 Artificial Intelligence: Natural Language Understanding (NLU)

A5: Optimization Human Preference

This assignment focuses on using Hugging Face models to optimize human preference, specifically leveraging the Direct Preference Optimization (DPO) trainer. You will work with preference datasets, train a model, and push it to the Hugging Face model hub. Additionally, you will build a simple web application to demonstrate the trained model.

Note: You are ENCOURAGED to work with your friends, but DISCOURAGED to blindly copy others' work. Both parties will be given 0.
Note: Comments should be provided sufficiently so we know you understand. Failure to do so can raise suspicion of possible copying/plagiarism.
Note: You will be graded upon (1) documentation, (2) experiment, (3) implementation.
Note: This is a one-week assignment, but start early.
Deliverables: The GitHub link containing the Jupyter notebook, a README.md of the GitHub repository, and the folder of your web application called 'app'.

Task 1. Finding a Suitable Dataset (0.5 point)

1) Select a publicly available dataset for preference optimization tasks, such as human preference rankings or reinforcement learning from human feedback (RLHF) datasets.
2) Ensure that the dataset is properly preprocessed and suitable for training a preference-based model.
3) Document the dataset source and preprocessing steps.

NOTE: You can use datasets from Hugging Face Datasets Hub [1].

Task 2. Training a Model with DPOTrainer

1) Implement the Direct Preference Optimization (DPO) training method with **DPOTrainer** Function using a pre-trained transformer model (such as GPT, or T5) on the Hugging Face and fine-tune it using the selected dataset. (1 point)
2) Experiment with hyperparameters and report training performance. (1 point)

HINT: Refer to the Hugging Face documentation for **DPOTrainer** implementation.[2].
Note: You do not need to train large model sizes like 1B-7B if your GPU is not capable. This assignment focuses on how to use pre-trained models with Hugging Face.

Task 3. Pushing the Model to Hugging Face Hub (0.5 point)

1) Save the trained model.
2) Upload the model to the Hugging Face Model Hub.
3) Provide a link to your uploaded model in your documentation.

NOTE: Make sure your repository is public[3] and also the README.md should also contain the link to your publicly available trained model on Hugging Face.

Task 4. Web Application Development (1 point)

1) Develop a simple web application that demonstrates your trained model's capabilities.
2) The app should allow users to input text and receive response.

Have fun with huggingface :-)

[1] https://huggingface.co/datasets
[2] https://huggingface.co/docs/trl/main/dpo_trainer
[3] https://huggingface.co/docs/hub/models-uploading