Hands-On: Building an RLHF Pipeline

Practical Implementation

Step-by-step guide to implementing a basic RLHF pipeline for medical applications.

Setup & Prerequisites

- Python 3.8+, PyTorch, Transformers library
- Pre-trained medical language model (e.g., BioBERT, ClinicalBERT)
- Preference dataset (or use publicly available data)
- GPU with 16GB+ VRAM recommended

Pipeline Steps

- Step 1: Load and prepare preference dataset
- Step 2: Train reward model on preferences
- Step 3: Set up PPO or DPO training loop
- Step 4: Optimize policy with reward guidance
- Step 5: Evaluate on held-out test set
- Step 6: Analyze outputs and iterate