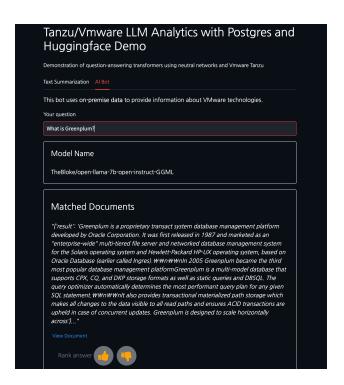


## Sample LLM App frontend with Multi-step MLflow workflows

Overview

Use this sample template as a baseline to bootstrap a Generative AI application whose LLM model is retrained / finetuned, evaluated and published using multi-step **MLflow** workflows. The sample template demonstrates using **HuggingFace Transformers** to publish a finetuned model based on the provided base LLM. The application is a **Streamlit** web app. Workflow steps may be configured for each run with MLflow's **MLproject** file.



## **Workflow Steps**

The workflow steps are as follows:

upload\_dataset Uploads the training dataset to the training

environment (supports local environment or

Ray clusters by default)

train\_model Finetunes or retrains the base LLM model

evaluate\_model Uses defined metrics and thresholds to

perform model evaluation on the candidate

model

to\_staging

promote\_model\_ By default, promotes the candidate model to Staging if it passes the evaluation stage; also publishes the LLM to HuggingFace and stores

metadata in Datahub