Project Name: Visual Fashion Recommender App

Team: Mikiko Bazeley

Project Goal: Develop, productionize, & deploy an end-to-end visual fashion recommender.

What dataset will you use, or how do you plan to collect data?
The two types of data sets I'll be using are:
- Category & Attributes from [DeepFashion Database](#)
- Scraped clothing items from key clothing sites to use as inventory

What baseline will you use?
DeepFashion has some benchmarks on their site & there are some projects with performance metrics.

What model architecture / loss function do you propose?
The models used to first categorize items of clothing & attributes will be CNNs, specifically I'll use ResNet with cross-entropy. Using the hidden layers I'll extract the embeddings, concatenate, and then apply nearest-neighbors to the embeddings to find the 5 most similar items.

What will the end result look like?
A website where users can upload an image of their favorite outfit (or clothing item) and will have 5 recommended items displayed on the website with links to purchase options from real clothing sites.

Specific tool stack
- PyTorch, FastAI
- Streamlit for front-end
- FastAPI for backend API framework
- AWS (RDS, Lambda, API Gateway, etc)

Must-Have Features:
- Logging
- Testing (Unit + Integration)
- Documentation

What's your stretch goal?
- Gitlab CI/Github Actions for CI/CD process
- Material Mkdocs for documentation around service
- Selenium for scraping
- Orchestration
- Monitoring

Meta-Goals

- Improve my time-to-delivery from model development to fully deployed & monitored
- Deeper understanding of FastAPI, AWS, & Selenium
- Develop mental models & working processes around data & model pipeline management