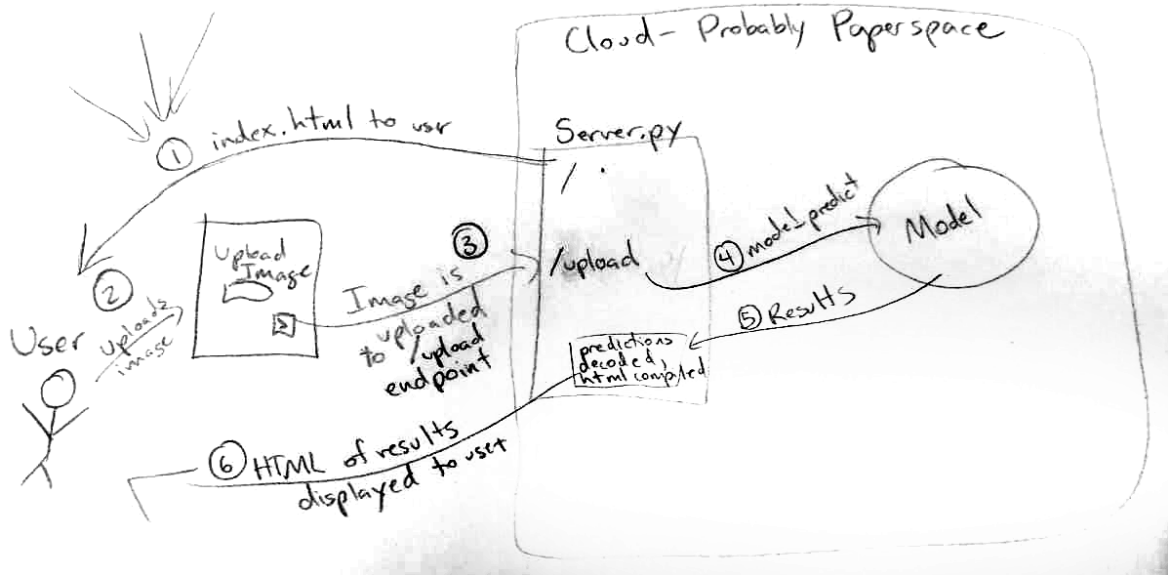


Springboard MLE Track

Capstone Project Step 11: Design Your Deployment Solution Architecture



The system will be very simple. It will consist of a server, hosted on the cloud, a backend model, and some HTML that gets updated with model results and served to the user.

Steps:

- 1) User goes to / and index.html is served
- 2) User uploads a pattern image and hits the button
- 3) /upload endpoint is called. Image is uploaded and model_predict is called
- 4) Model runs the prediction
- 5) Returns results back to server
- 6) Server decodes the predictions and generates html and serves it back to the user.

Due to the nature of the predictions being made, this model is unlikely to need to be retrained often. I do want to update it to handle more keywords, however.

The system will be monitored with log files. It will be deployed on Paperspace (temporarily). The implementation will take a few hours.