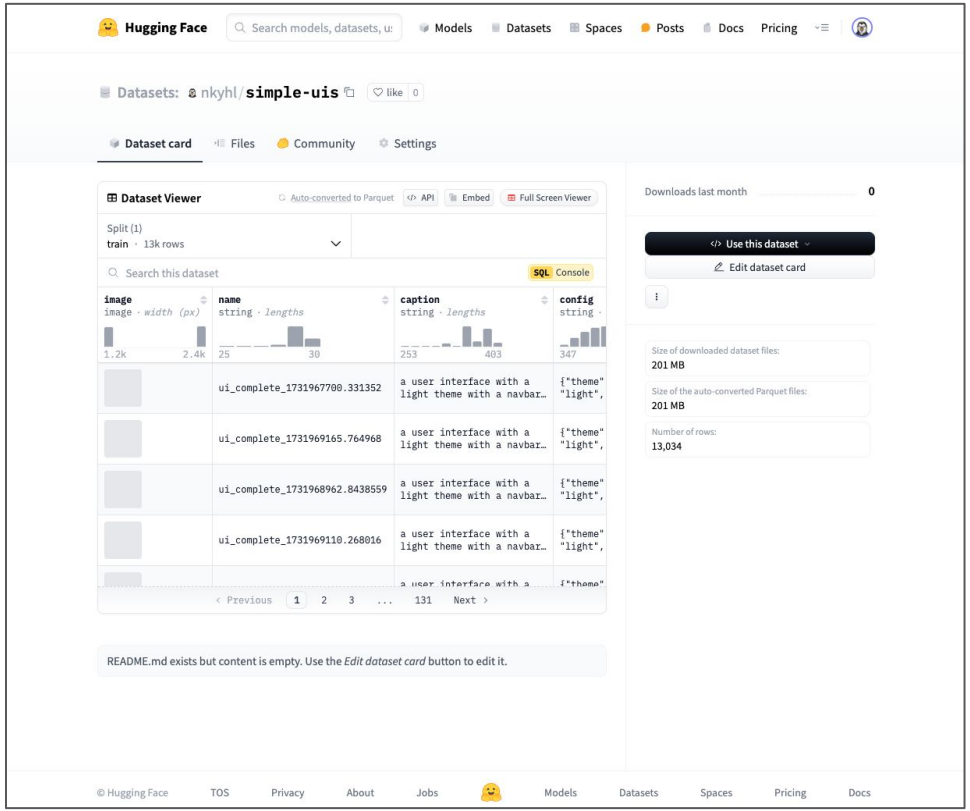
Nov 20th Research Update: Explore → Pruning Phase

Putting components together

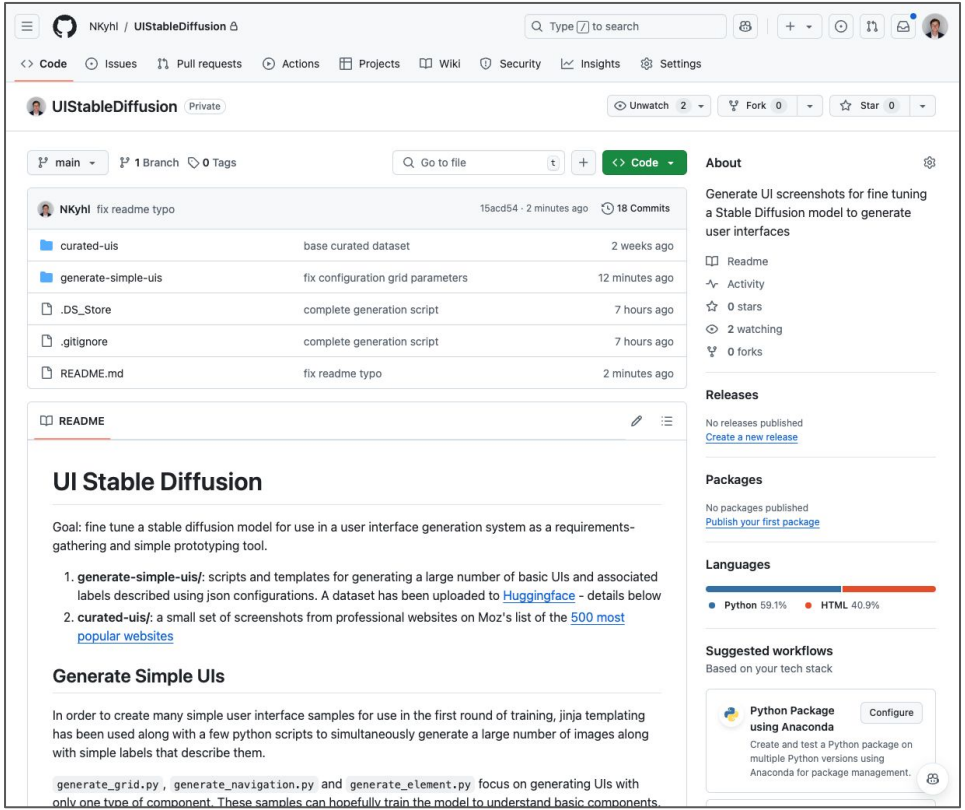
- I merged my scripts into a new script “generate_complete.py” which takes in a json configuration for a *complex* UI with multiple components and outputs a screenshot and english label
- I used this to generate *13,000* image-label samples of UIs, and uploaded this to Hugging Face.
- I retrained the Stable Diffusion Model on this data overnight
  - Around 9 hours, saving every 400 steps (1 hour). Learning Rate=1e6, Batch Size=4. Seemed to underfit, perhaps because more training time is needed for the large increase in samples but similar variation?
- I’ve since started a new round of training beginning from september 30th’s model (trained only on grids), with a higher learning rate.



Nov 20th Research Update: Explore → Pruning Phase

Preparing for the end of the semester

- I began cleaning up the GitHub repo and I wrote a Readme that I can build off of as detailed documentation for the project.
- This links to the data on Hugging Face and will link to the model on Civit.ai when I upload my final model.
- I am creating training scripts for this repo that allow users to more easily follow the steps in the fine-tuning process I've been following this semester.



Questions

1. Hugging face allows me to download the dataset as a .arrow file, but I can't find a built-in way to download and recover individual screenshot and caption files for use in training our model. Am I missing something?
2. Preparing the latents for model training took about 2 hours, is this reasonable?
3. Fine tuning with our process relies on scripts from this repo of stable diffusion scripts. Should I attempt to extract them for our own repo (with citation) or should I simply mention in the readme that users should also clone that repository?

Next Steps

1. I am now done with generating training data and will be fully focused on fine tuning using this semester's data to create my final model and document the steps that were used in training it. I will then upload this to Civit.ai.
2. I would like to simplify the process described in the civit.ai fine-tuning tutorial because some parts are irrelevant to our use case. By creating a few simple scripts to automate the tedious command line work and writing short documentation it should be easier for future users to jump in and fine tune.
3. If possible, I think it would be interesting to label our small dataset of professional UIs and include these in a 2nd round of training for our final model. Amazon Turk?