Setting up a remote machine

Outline

- Why
- Where to rent
- How to setup

Why

- Your laptop and desktop are weak

Why

- Your laptop and desktop are weak
  - Consumer grade components, not designed for 24/7 100% utilization, small GPUs

Why

- Your laptop and desktop are weak
  - Consumer grade components, not designed for 24/7 100% utilization, small GPUs
- Probably are not running linux (OS for almost all deep and machine learning)

Why

- Your laptop and desktop are weak
  - Consumer grade components, not designed for 24/7 100% utilization, small GPUs
- Probably are not running linux (OS for almost all deep and machine learning)
- SOLUTION- Rent a remote machine!

Where

- There are a lot to choose from:

Cloud Clusters

Where

- We are going to use Paperspace

Cloud Clusters

Why Paperspace?

- Affordable (free, pro and growth)
- Decent machines
- Has jupyter lab (unlike Google colab or kaggle)
- And a terminal window
- Easy session limiting(defaults to 6 hours)
- Setup steps will apply to other cloud providers

Why Paperspace?

- Please signup for an account
- They have free, but a pro account ($8/month) gives you more options

Why you don’t use remote desktop software with cloud compute

- You connect to a remote machine over a network, which is much slower than doing everything locally
- Network speed and latency are concerns
  - If you try to use a remote desktop solution (x2go, GoToMyPC, etc..) you are sending a large portion of each video frame multiple times per second.
  - Worst case: On this machine 1920x1080*32bits=8.3Mbytes/frame, sent 60 times per second
  - YOU WILL NOTICE LAG AND DROPPED FRAMES!

Why you don’t use remote desktop software with cloud compute

- Plus you have to waste GPU memory on your remote machine to produce a desktop to send over the network to your local machine.
-

Why you don’t use remote desktop software with cloud compute

- Plus you have to waste GPU memory on your remote machine to produce a desktop to send over the network to your local machine.
- Finally, it’s probably not supported by the cloud provider anyway

Solution: don’t have a desktop

Use Jupyter Lab and terminal instead: they send just a few characters, only when they are produced. Much, much lower bandwidth requirements.

How: Stuff to cover

- A little on the linux boot sequence
- .bashrc file
- Script files
- Some linux commands (du, pwd, cd, which, whereis, mv etc.)
- Permanent verses ephemeral storage (/storage and /notebooks are permanent on Paperspace, all else is ehemeral)
- Symbolic links
- A little vim
- Universal ctags and code navigation
- CLI apis (for Kaggle and Paperspace)

How: Demo

- Demo configuring a vanilla linux machine to ensure that changes persists across sessions (aliases, packages, config files, data directories etc..)
- I’ll show you how to do this manually, then port this process to a script(s), then port the script(s) and setup data to a git repo.
- The git repo will serve as a guide for easily setting up a custom machine.

Format

- Live video session so you can ask questions.
- Recording will be posted online.