Deep Learning on Edge Device

or
Pain and suffering

Why?

- IoT is the "thing"

- Avoid clogging up networks

- Reduce network costs

- Edge devices have increasing GPU capabilities

The Jetsons

- Jetson Nano (~$240)
  - GPU: 128 core Maxwell
  - Quad core ARM A57
  - 4GB RAM (shared)
  - 2x CSI
  - Gigabit Ethernet
  - HDMI/Display port
  - 4x USB 3
  - 40pin GPIO
  - Power: USB micro or barrel jack (5-10W)

Does it work?

- Kind of...
  - JetPacks (ie prebuilt images) only come with Python 3.6, based on Linux4Tegra (L4T)
  - Compilation is terribly slow (missing wheel files)
  - PyTorch can easily run out of memory
  - Nvidia-based code/models does work

The Jetsons' bigger brother

- Jetson Xavier NX (>$400)
  - 512 core Volta
  - 8-core ARMv8.2
  - 2x CSI
  - Gigabit Ethernet, M.2
  - HDMI/Display port
  - 4x USB 3
  - 40pin GPIO
  - Power: 19V barrel jack (10-15W)

Does it work?

- Much better
  - still limited with JetPacks and Python 3.6
  - Compilation is way faster
  - Double amount of RAM helps
  - PyTorch works reasonable