

CSCI 1470

Eric Ewing

Monday,  
4/21/25

# Deep Learning

Day 34: What comes next?

# Logistics

DL Day, April 30<sup>th</sup>

- 3 timeslots throughout the day
  - 10-11:30, 12:00-1:30, 2:00-3:30
- Fill out form on Ed to indicate timeslot that works best for your group (i.e., the one where the most people can make it)
- Poster printer in CIT has limited hours this year 8:30-4:30. We (staff) will be printing posters the 28<sup>th</sup> and 29<sup>th</sup>, if you want to wait until the 30<sup>th</sup> and risk it, that's up to you.

# Logistics

1470 will be offered again in the Fall and we're looking for TAs!

# Trends in Deep Learning Jobs

## 1. Generative Models

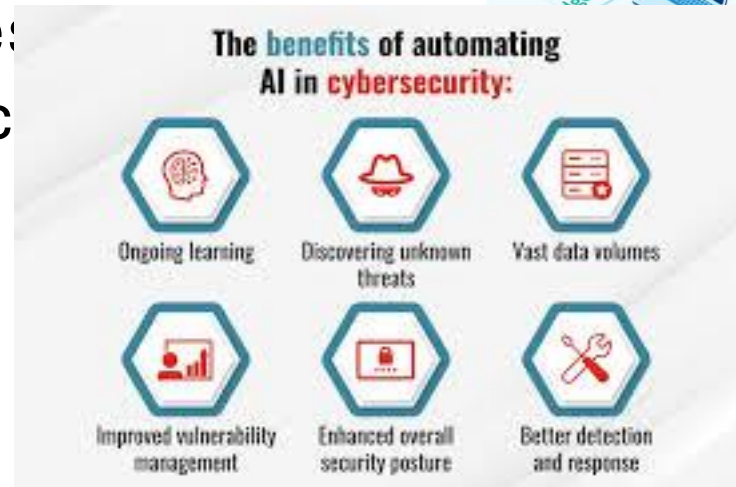


## 2. Edge AI, connected devices

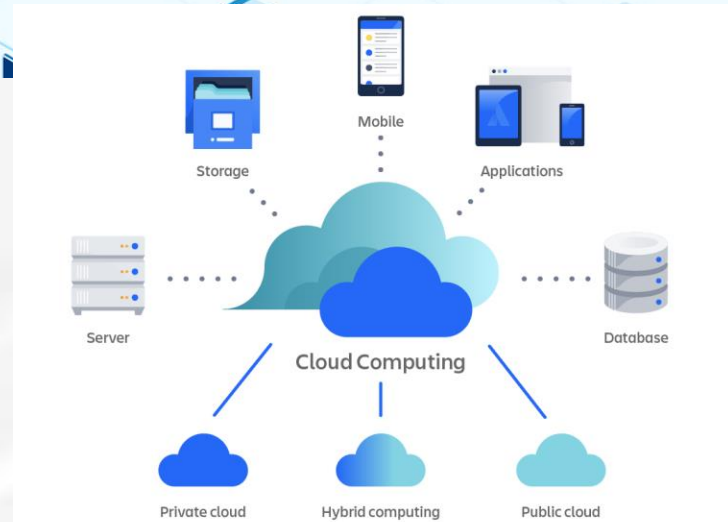
- Can your AI run on limited c

Edge AI

## 3. Cybersecurity and AI



## 4. Deep Learning Infrastructure (Hardware or Cloud Compute)



# Deep Learning (or Adjacent) Jobs

- Data Scientist
  - Process and analyze large amounts of data
  - Find and validate trends in data, “Actionable Insights”
- Machine Learning Engineer
  - Build and maintain data pipeline for machine learning models
  - Implement and optimize ML/DL models
  - Evaluating Models
- Machine Learning Researcher
  - Research and implement **new** architectures, optimizers, or other techniques
- “Topic Specific” Jobs (e.g., Computer Vision, Fine-tuning LLMs, startups, etc.)

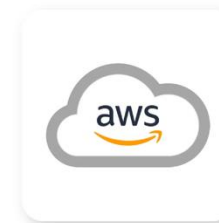
# Skills You've Gained

- Foundations of linear algebra, optimization, and probability
- Data processing
- Neural network architectures
- Practical training skills (e.g., managing overfitting/underfitting, hyperparameter tuning)
- Going from math/slides to code

# Skills You Should Cultivate

- Use of popular training infrastructures

The big 3 Cloud platforms



AWS



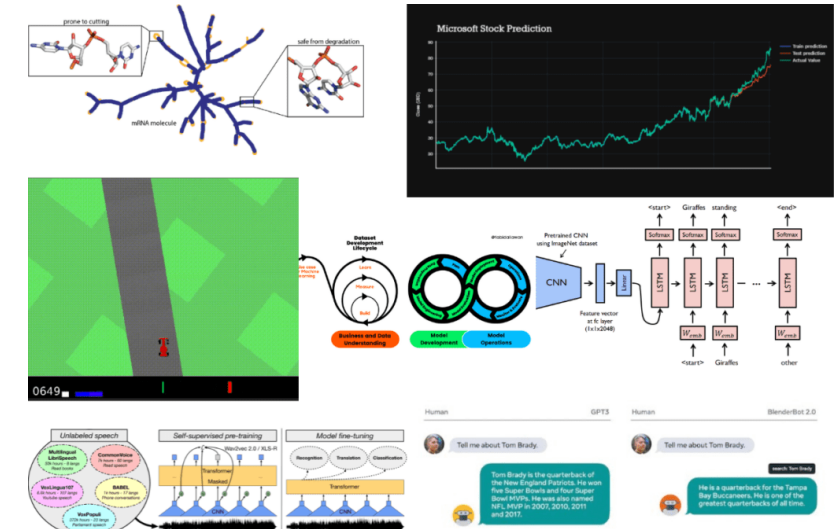
Azure



GCP

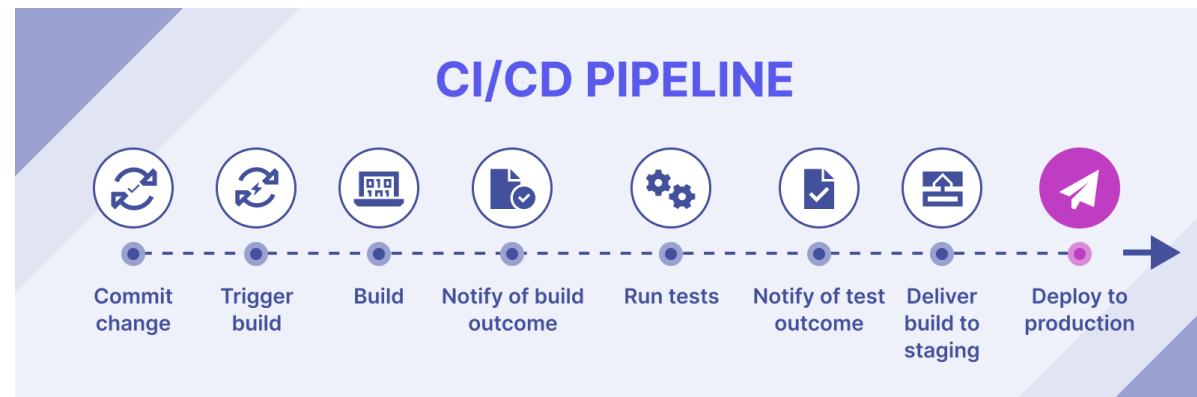
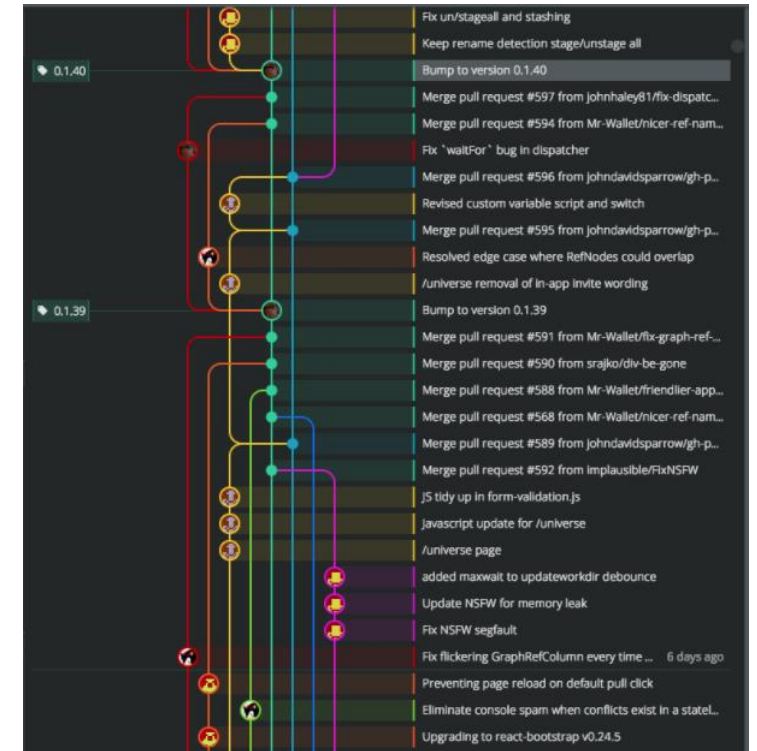


- What is the right technique to use for your problem?



# Skills You Should Cultivate

- Know how to use git (more than just commit, push, and pull)
- CI/CD pipelines



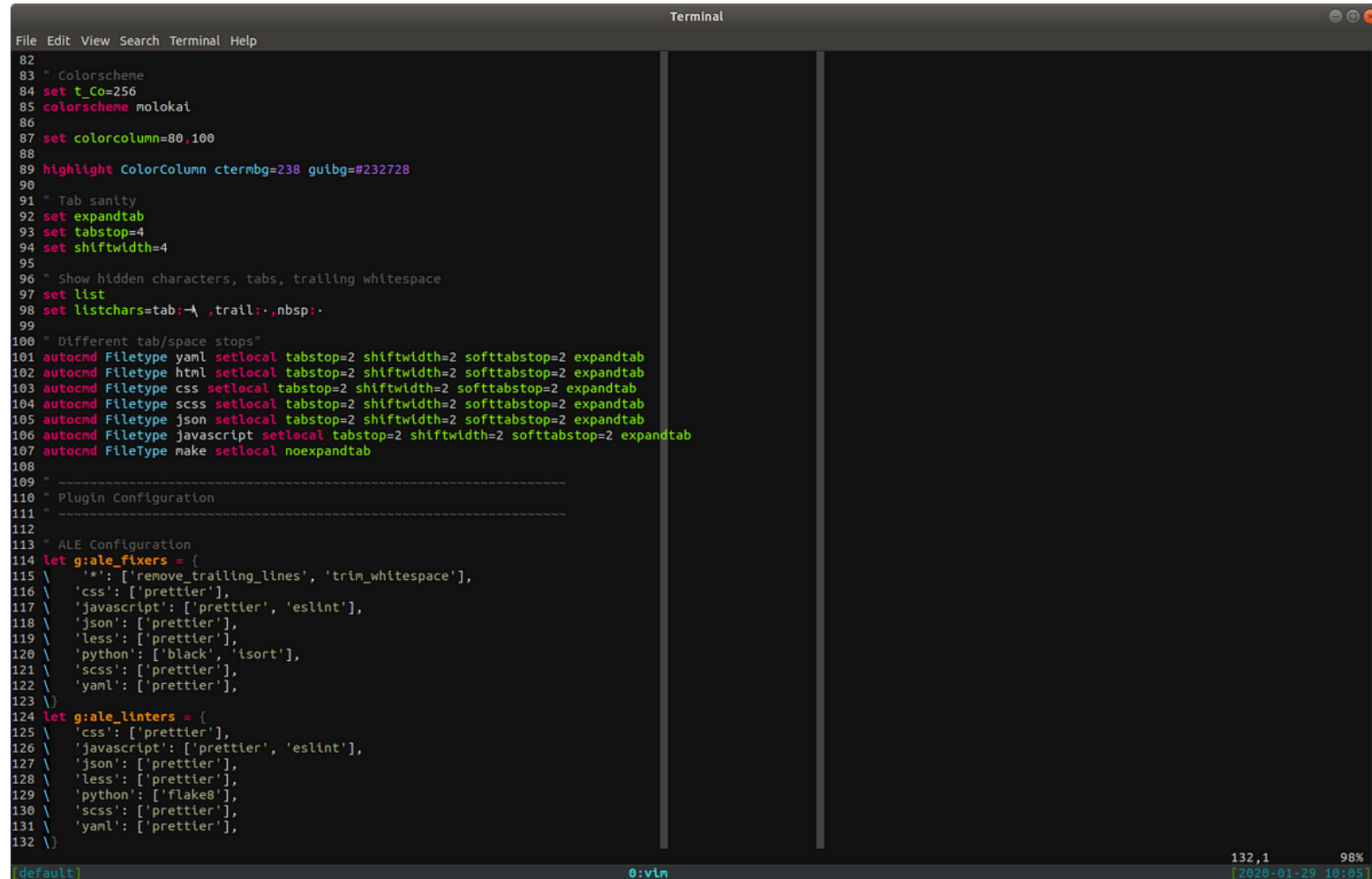


# Skills You Should Cultivate

Vim

Just Kidding

(Not Really)

A screenshot of a terminal window titled "Terminal" showing a Vim configuration file (.vimrc). The file contains various settings for colors, tabs, and plugins. The text is color-coded: comments are in light blue, set and autocmd commands are in green, and plugin names are in red. The configuration includes settings for the 'molokai' color scheme, tab settings, and ALE (Asynchronous Lint Engine) configuration for various file types like yaml, html, css, json, javascript, and python. The terminal window has a dark background and a light gray border. The status bar at the bottom shows "0:vim" and "132,1 98%".

```
File Edit View Search Terminal Help
82
83 " Colorscheme
84 set t_Co=256
85 colorscheme molokai
86
87 set colorcolumn=80,100
88
89 highlight ColorColumn ctermbg=238 guibg=#232728
90
91 " Tab sanity
92 set expandtab
93 set tabstop=4
94 set shiftwidth=4
95
96 " Show hidden characters, tabs, trailing whitespace
97 set list
98 set listchars=tab:~\ ,trail:.,nbsp:·
99
100 " Different tab/space stops"
101 autocmd FileType yaml setlocal tabstop=2 shiftwidth=2 softtabstop=2 expandtab
102 autocmd FileType html setlocal tabstop=2 shiftwidth=2 softtabstop=2 expandtab
103 autocmd FileType css setlocal tabstop=2 shiftwidth=2 softtabstop=2 expandtab
104 autocmd FileType scss setlocal tabstop=2 shiftwidth=2 softtabstop=2 expandtab
105 autocmd FileType json setlocal tabstop=2 shiftwidth=2 softtabstop=2 expandtab
106 autocmd FileType javascript setlocal tabstop=2 shiftwidth=2 softtabstop=2 expandtab
107 autocmd FileType make setlocal noexpandtab
108
109 "
110 " Plugin Configuration
111 "
112
113 " ALE Configuration
114 let g:ale_fixers = {
115 \   '*': ['remove_trailing_lines', 'trim_whitespace'],
116 \   'css': ['prettier'],
117 \   'javascript': ['prettier', 'eslint'],
118 \   'json': ['prettier'],
119 \   'less': ['prettier'],
120 \   'python': ['black', 'isort'],
121 \   'scss': ['prettier'],
122 \   'yaml': ['prettier'],
123 \ }
124 let g:ale_linters = {
125 \   'css': ['prettier'],
126 \   'javascript': ['prettier', 'eslint'],
127 \   'json': ['prettier'],
128 \   'less': ['prettier'],
129 \   'python': ['flake8'],
130 \   'scss': ['prettier'],
131 \   'yaml': ['prettier'],
132 \ }
[default] 0:vim 132,1 98% [2020-01-29 10:05]
```

# Build a Portfolio

- Portfolios matter!
  - Demonstrates practical skills beyond theoretical knowledge
  - Shows ability to complete projects from conception to deployment
  - Acts as tangible proof of your capabilities
- Essential Portfolio Components
  - GitHub repository with well-documented code
  - Personal website showcasing projects
  - Variety of projects demonstrating range of skills
- Ideally shows a mix of breadth in many topics and depth in a certain topic

# What Type of Projects to do?

- A bit easier to say what not to do:
  - Don't make another CNN for MNIST or CIFAR
  - Don't find a popular reddit post or medium article describing "Top 5 projects for your DL Portfolio"
- Projects are evaluated for creativity first and then technicality
  - The first person who looks at your portfolio won't be able to evaluate your technical ability, you must draw them in
- A portfolio doesn't have to be perfectly cohesive
  - Follow your short-term nose

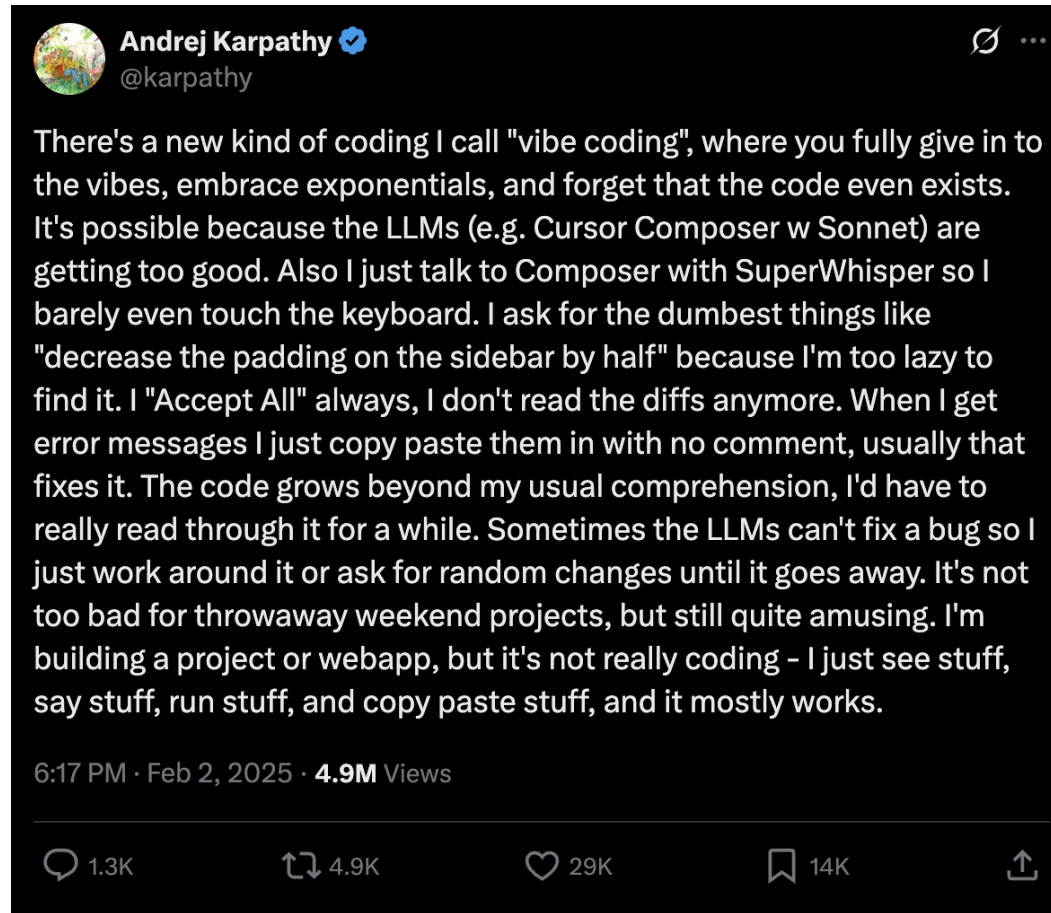
# What Type of Projects to do?

- (At least part of) your portfolio should match the jobs you are applying for
  - If you are applying for a ML-Engineer or data science position, you **shouldn't** work only with perfectly clean data (e.g., Kaggle datasets)
- Do what interests **you**
  - You will work harder and create better results if you are **genuinely interested** in your project
- If you're interested in grad school, get involved in research
  - If you're in grad school, try doing research

# Research Needs a Special Sauce

- There are lots of people doing research in AI
  - Be wary of problems you know thousands of people are working on
  - Don't try to make a better CNN for ImageNet
- Research needs something special:
  - A technique that other people haven't thought of applying to this problem yet
    - Harder, but more impactful
  - A dataset that other people don't have access to
    - Easier, but potentially isolated impact

# Should I Use LLMs to Build my Portfolio?



# Should I Use LLMs to Build my Portfolio?

## **The Impact of Generative AI on Critical Thinking: Self-Reported Reductions in Cognitive Effort and Confidence Effects From a Survey of Knowledge Workers**

Hao-Ping (Hank) Lee  
Carnegie Mellon University  
Pittsburgh, Pennsylvania, USA  
haopingl@cs.cmu.edu

Ian Drosos  
Microsoft Research  
Cambridge, United Kingdom  
t-iandrosos@microsoft.com

Advait Sarkar  
Microsoft Research  
Cambridge, United Kingdom  
advait@microsoft.com

Sean Rintel  
Microsoft Research  
Cambridge, United Kingdom  
serintel@microsoft.com

Nicholas Wilson  
Microsoft Research  
Cambridge, United Kingdom  
niwilson@microsoft.com

Lev Tankelevitch  
Microsoft Research  
Cambridge, United Kingdom  
levt@microsoft.com

Richard Banks  
Microsoft Research Cambridge  
Cambridge, United Kingdom  
rbanks@microsoft.com

# Should I Use LLMs to Build my Portfolio?

My Take: If you want a job in ML, ***you*** should do the machine learning

If you want a job as a ML engineer, you shouldn't only work with perfectly clean data

If you want to showoff your ML project as part of a webapp, it's reasonable to automate building the frontend.



Any Questions?