CAI 4104/6108 – Machine Learning Engineering:

Review + Q&A

Prof. Vincent Bindschaedler

Spring 2024

Administrivia

- Final Exam
  - When: May 2, 2024 — 7:30AM to 9:30AM.
    - Exam **opens** at **7am ET** and closes at 10:00AM (***don't wait until the last minute to start the exam!***)
  - Where: Online (Canvas + **Honorlock**)
  - Note:
    - The CAI4104 and CAI6108 exams will be (slightly) different
  - Format:
    - Some Short answer questions (may include multiple choice)
    - Some multi-part problems

- **Sample Final Exam** (Practice Questions) — Live on Canvas
  - Please use it to prepare but do **not** overfit to it
  - It will close at 6:30am the day of the final (so there is no confusion)

Administrivia

- **Course Evaluation**
  - Help us improve the course!
  - Complete your evaluation by **April 26**
  - Access the evaluation form:
    - Canvas: click on GatorEvals (left navigation panel)
    - or: https://ufl.bluera.com/ufl/
  - Optional and anonymous

Part 1: Fundamentals

- ML Engineering:
  - Workflow, Feature Engineering, Model Selection, Hyperparameter tuning, Performance Evaluation, Bias-Variance Tradeoff, Underfitting/Overfitting
- Supervised Learning:
  - Support Vector Machines (SVMs)
  - k-Nearest Neighbors
  - Linear models: Linear regression, Logistic regression
  - Trees & Ensembles: Decision Trees, Random Forests, Voting/Bagging/Stacking
- Unsupervised Learning
  - Clustering: K-Means
  - Dimensionality reduction: PCA, KernelPCA, Manifold Learning (MDS, LLE, t-SNE)
- Learning Algorithms:
  - Stochastic Gradient Descent & variants

Part 2: Neural Networks

- Components/Architectures of Neural Networks:
  - Neuron/Unit, Activation Functions, Hidden Layers, width vs depth
- Training Deep Neural Networks:
  - Backpropagation, Vanishing/exploding gradient problems
  - Initialization of weights, saturating activation functions, dying ReLUs, tuning hyperparameters
- Architectures
  - Simple, Fully-connected / Dense Nets
  - Convolutional Neural Nets: Convolutions, Pooling, Flattening, etc.
  - Recurrent Neural Nets: Recurrent Layers, Seq-to-vec, Vec-to-seq, Seq-to-seq, encoder-decoder networks, Bidirectional RNNs, Types of Cells (GRU, LSTMs), Attention, Transformers
- Unsupervised Learning & Generative Models
  - Auto-Encoders: Latent Space, Types of AutoEncoder, Variational AutoEncoder (VAE)
  - Generative Adversarial Networks (GANs): Discriminator & Generator, Adversarial Training/Learning

Part 3: ML & Society

- Adversarial ML:
  - Evasion attacks, Adversarial Examples, Weird Properties of Neural Nets
- Privacy Attacks:
  - Membership Inference Attacks, Memorization/Overfitting
- Interpretable/Explainable ML:
  - Need for human understandable explanations, Proxy models, LIME, Rule Extraction
  - Saliency Maps, Explanation Synthesis
- Fairness:
  - Sources of ML Unfairness, making models fair
  - Fairness Notions (e.g., anti-classification, statistical party, individual fairness)
- Synthetic Media
  - Deepfakes, LLMs and future architectures

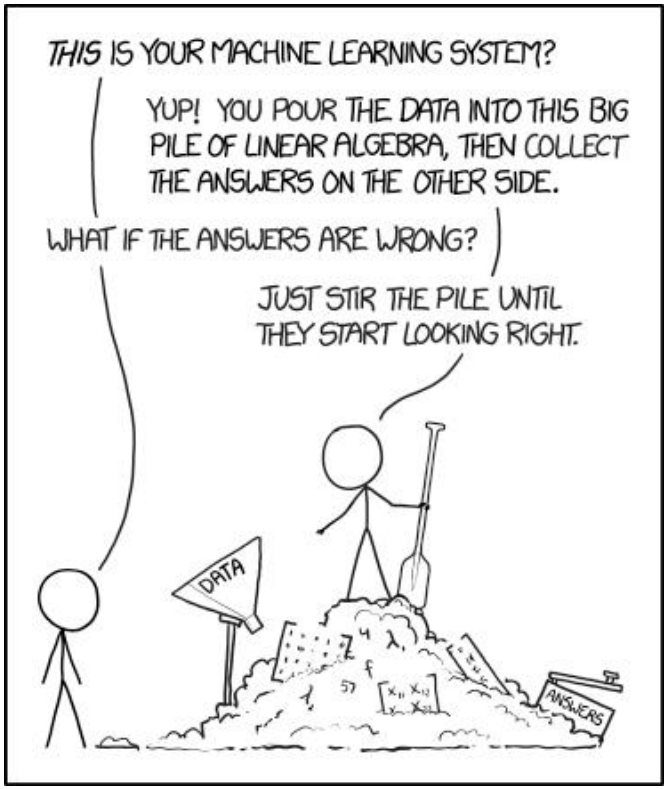
Questions?

Deadlines

- Upcoming:
  - Project due 4/24
    - No late penalty if submitted by 4/26
  - Final Exam on 5/2
    - 7:30AM to 9:30AM (Online — Canvas + Honorlock)

The End

- I hope you enjoyed this course



source: xkcd