S&DS 365 / 665

**Intermediate Machine Learning**

# Course Wrap-up:

December 6

Yale

# Reminders

- Assn 5 due tonight
- Final exam: Wed Dec 20 at 9am in HC L02
- Practice exams posted
- Review sessions (so far):
  - Tuesday, Dec 12, 7pm (Ruixiao)
  - Wednesday, Dec 13, 7pm (Zehao)

# Last unit: Language/Sequence models

- Generative process, any sequence (of words, characters, stock prices, nucleotides...) is assigned a probability

  $$p(x_1, \ldots, x_n)$$

  which can be factored as

  $$p(x_1, \ldots, x_n) = p(x_1)p(x_2 \mid x_1) \ldots p(x_n \mid x_1, \ldots, x_{n-1})$$

# Transformers

The current state-of-the-art is based on *transformers*

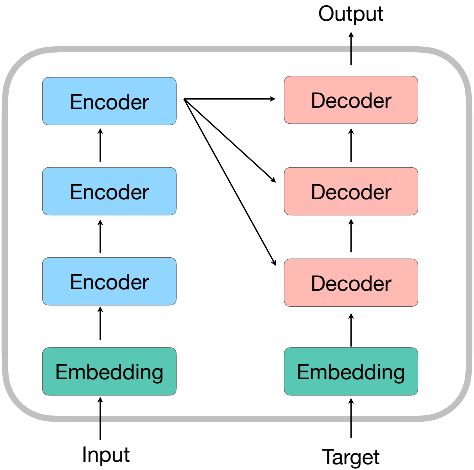
- Attention is the key ingredient
- Rather than processing sequences word-by-word, transformers handle larger chunks of text at once
- Incorporates “interactions” between words and hidden states

# Transformer architecture

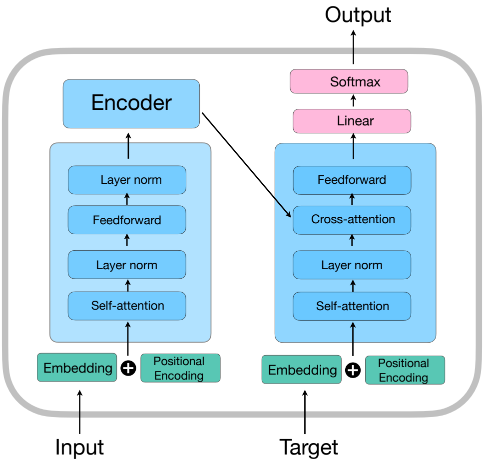
A Transformer is a seq2seq model based on encoder and decoder modules.

Transformers are powerful alternatives to RNNs that transform the encoder/decoder states using (multi-head) attention mechanisms.

# Transformer architecture

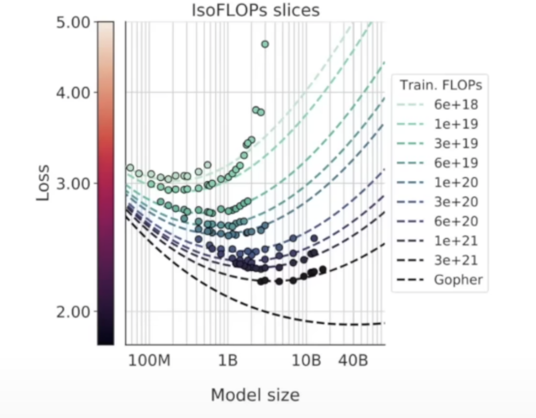


# Transformer architecture




Two encoder layers and one decoder layer

# LLM scaling laws: Bigger is better

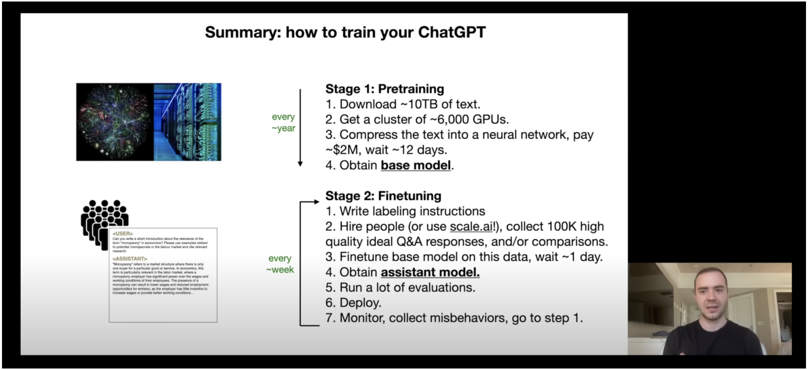



https://www.youtube.com/watch?v=zjkBMFhNj_g&t=25m40s

# Finetuning (Training an LLM Decoder)

- An incredible amount of knowledge is stored implicitly in the weights of the transformer
- To be made useful, supervised learning and reinforcement learning are used to teach the model how to respond

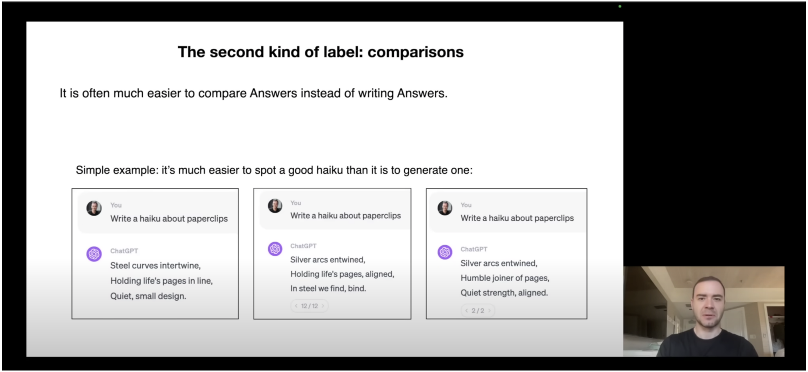
# Finetuning



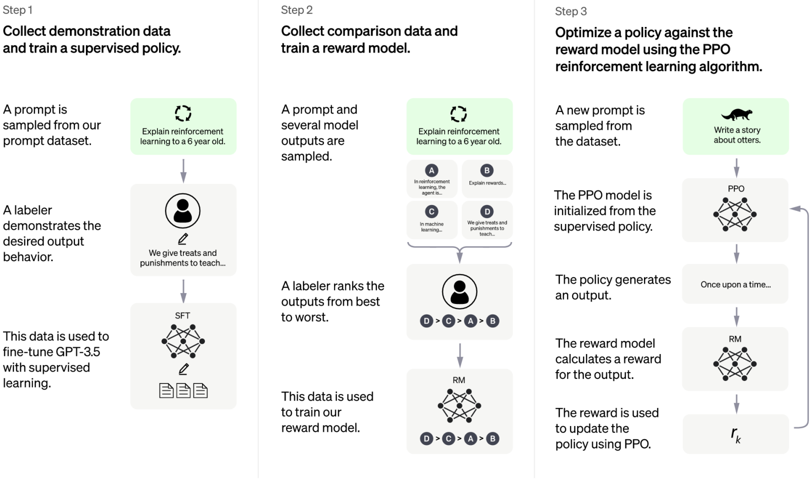

https://www.youtube.com/watch?v=zjkBMFhNj_g&t=14m19s

# Finetuning

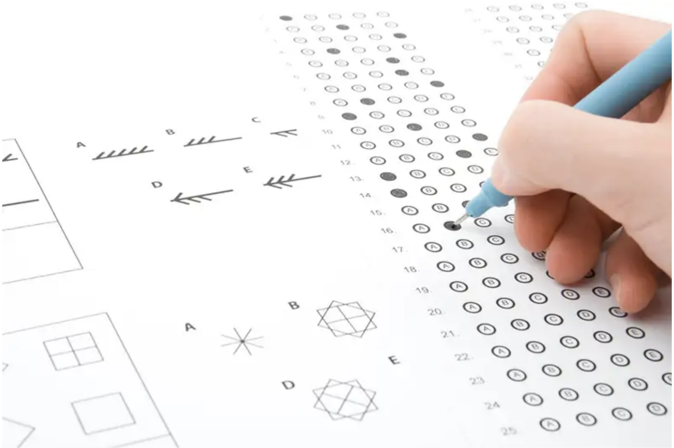
## RLHF




https://www.youtube.com/embed/zjkBMFhNj_g?start=1277&end=1340


Step 1
Collect demonstration data and train a supervised policy.
A prompt is sampled from our prompt dataset.
Explain reinforcement learning to a 6 year old.
A labeler demonstrates the desired output behavior.
We give treats and punishments to teach...
This data is used to fine-tune GPT-3.5 with supervised learning.
SFT
Step 2
Collect comparison data and train a reward model.
A prompt and several model outputs are sampled.
Explain reinforcement learning to a 6 year old.
A
In reinforcement learning, the agent is...
B
Explain rewards...
C
In machine learning...
D
We give treats and punishments to teach...
A labeler ranks the outputs from best to worst.
D > C > A > B
This data is used to train our reward model.
RM
D > C > A > B
Step 3
Optimize a policy against the reward model using the PPO reinforcement learning algorithm.
A new prompt is sampled from the dataset.
Write a story about otters.
The PPO model is initialized from the supervised policy.
PPO
The policy generates an output.
Once upon a time...
The reward model calculates a reward for the output.
RM
The reward is used to update the policy using PPO.
$r_k$

# What’s next?: Fast learning, slow thinking



**AI struggles with IQ tests**
Panther Media GmbH / Alamy Stock Photo

# Always news—on both sides

The Economist

Menu | Weekly edition | The world in brief | Search

Science and technology | Crystal balls

## A Google AI has discovered 2.2m materials unknown to science

Zillions of possible crystals exist. AI can help catalogue them

**Just one of trillions** IMAGE: SCIENCE PHOTO LIBRARY

Nov 29th 2023

Save | Share | Give

# Always news—on both sides

The Economist

Menu | Weekly edition | The world in brief | Search

Business | Of evils and evals

## The world wants to regulate AI, but does not quite know how

There is disagreement over what is to be policed, how and by whom

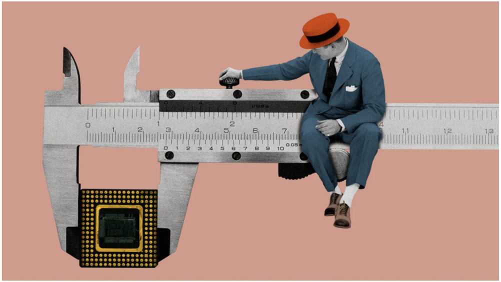


IMAGE: MARI FOUZ

Oct 24th 2023 | BLETCHLEY PARK

Save | Share | Give

# Advancing ML: Something for everyone

- Designing new methods
- Applying to new domains
- Solving deep mathematical puzzles
- Tackling unique engineering challenges
- Designing interfaces
- Collecting data
- Equitable use through policy and law
- Outreach and communication to broad communities

# We’ve covered a lot of ground!

## Calendar Fall 2023

Lectures: Monday/Wednesday 1:00-2:15pm
Luce Hall 101

| Week | Dates | Topics | Demos & Tutorials | Lecture Slides | Readings & Notes | Assignments & Exams |
|---|---|---|---|---|---|---|
| 1 | Aug 30, Sep 1 | Sparse regression | CO Python elements<br>CO Pandas and regression<br>CO Lasso example | Wed: Course overview<br>Fri: Sparse regression | PML Section 11.4<br>Notes on linear regression | |
| 2 | Sep 6 | Smoothing and kernels | CO Smoothing example<br>CO Using different kernels | Wed: Smoothing | PML Sections 16.3, 17.1<br>Notes on computing the lasso | Quiz 1 |
| 3 | Sep 11, 13 | Density estimation and Mercer kernels | CO Density estimation demo<br>CO Mercer kernels (1/2)<br>CO Mercer kernels (2/2) | Mon: Density estimation<br>Wed: Mercer kernels | Risk bounds for local smoothing<br>Notes on Mercer kernels | CO Assn 1 out |

# We’ve covered a lot of ground!

| | | | | | | |
|---|---|---|---|---|---|---|
| 4 | Sep 18, 20 | Neural networks and overparameterized models | np-complete example (1/2)<br>np-complete example (2/2)<br>TensorFlow playground | Mon: Neural networks<br>Wed: Double descent | PML Sections 13.1, 13.2<br>Notes on backpropagation<br>Notes on double descent | Quiz 2 |
| 5 | Sep 25, 27 | Convolutional neural networks | Convolution demo<br>CNN demo | Mon: Double descent (continued) and Convolutional neural networks<br>Wed: CNNs (continued) | PML Section 17.2<br>Notes on Bayesian inference<br>Notes on nonparametric Bayes | Assn 1 in<br>Assn 2 out |
| 6 | Oct 2, 4 | Gaussian processes and approximate inference | Parametric Bayes<br>Gaussian processes<br>Gibbs sampling for image denoising | Mon: Gaussian processes<br>Wed: Introduction to approximate inference | Notes on simulation | Quiz 3 |

# We’ve covered a lot of ground!

| | | | | | | |
|---|---|---|---|---|---|---|
| 7 | Oct 9, 11 | Variational inference | Variational autoencoders | Mon: Variational inference<br>Wed: VAEs | PML Section 20.3<br>Notes on variational inference | Assn 2 in<br>Assn 3 out |
| 8 | Oct 16 | Midterm | | | Practice midterms | Oct 16: Midterm exam |
| 9 | Oct 23, 25 | Graphs and structure learning | Graphical lasso demo | Mon: Sparsity and graphs<br>Wed: Discrete data and graph neural nets | Notes on graphs and structure learning<br>Graph neural networks<br>PML Section 23.4 | |
| 10 | Oct 30, Nov 1 | Deep reinforcement learning | Q-learning demo<br>DQN demo | Mon: Reinforcement learning<br>Wed: Deep reinforcement learning | Sutton and Barto, Section 6.5 | Nov 1: Assn 3 in<br>Assn 4 out |

# We’ve covered a lot of ground!

| | | | | | | |
|---|---|---|---|---|---|---|
| 11 | Nov 6, 8 | Policy gradient methods | Policy gradients demo<br>Actor-critic demo | Mon: Policy gradient methods<br>Wed: Actor-critic methods | Sutton and Barto, Section 13.1-13.3, 13.5 | Quiz 4 |
| 12 | Nov 13, 15 | Sequential models | vanilla RNN<br>Fakespeare GRU | Mon: HMMs and RNNs<br>Wed: RNNs, GRUs, LSTMs, and all that | TensorFlow: Text generation<br>Notes on HMMs and Kalman filters<br>PML Chapter 15 | Assn 4 in<br>Assn 5 out |
| 13 | Nov 20, 22 | No class, Thanksgiving break | | | | |
| 14 | Nov 27, 29 | Sequence-to-sequence models and Transformers | GPT-4 Python API | Mon: Sequence-to-sequence models<br>Wed: Attention and transformers | Notes on mixtures<br>PML Sections 15.4, 15.5 | Quiz 5 |
| 15 | Dec 4, 6 | Transformers<br>Societal issues | Transformer demo<br>Minimal LLM decoder | Mon: Transformers, LLMs, and AI safety<br>Wed: Course wrap up | | Assn 5 in |
| 17 | Wed Dec 20, 9am, HQ L02 | Final exam | | | Practice exams | Registrar: final exam schedule |

# Final exam

- Final exam Wednesday, Dec 20, 2023 at 9am in HQ L02
- `https://registrar.yale.edu/general-information/final-exams`
- Review sessions (see times/dates above)
- Length: About 1.5X Midterm
- Emphasis on material after midterm
- Cumulative, closed book, cheat-sheet

# Your input

- Please complete a course review!
- I greatly value your comments and feedback
- Feel free to send me comments privately
- Let me know how you use and continue to learn ML!

Thank you!