CS 4644 / 7643-A: Deep Learning

Website:
https://faculty.cc.gatech.edu/~zk15/teaching/AY2025_cs7643_summer/index.html

Piazza:	https://piazza.com/gatech/summer2025/cs46447643a/
	(sync'd to Canvas)
Canvas:	https://gatech.instructure.com/courses/486594 (4644)
	https://gatech.instructure.com/courses/486592 (7643)
Gradescope:	https://www.gradescope.com/courses/1037086 (4644)
	https://www.gradescope.com/courses/1037087 (7643)

Zsolt Kira

Associate Professor

School of Interactive Computing
Georgia Tech

Are you in the right place?

- This is CS 4644 / CS 7643-A
  - “On campus” class
  - For project, you can group across ugrad/grad with permission

- This is NOT CS 7643-O01/OAN/Q/R/AO (“OMSCS”)
  - Online class for OMSCS program, but other sections combined
  - AO section is NOT on-campus section! It is linked with OMSCS version
  - You cannot group between on-campus class and OMSCS

Summer 25 Delivery Format

- In-Person
  - Scheller College of Business 200

- Streaming & Recording
  - We STRONGLY encourage you to attend the lectures in person
  - But **DO NOT come in sick!**
    - We will provide recordings for such cases & accommodations
  - Lectures recordings **MAY** be available **on a delayed basis, inversely proportional to attendance.** Do not rely on this.
  - **(Remote or in-person) recordings by students not allowed unless you talk to me first**

- Office hours, HW/project submissions online

- **Remember: Content is free online.**
  - **You are here for the interaction and the insight.**

Outline for Today

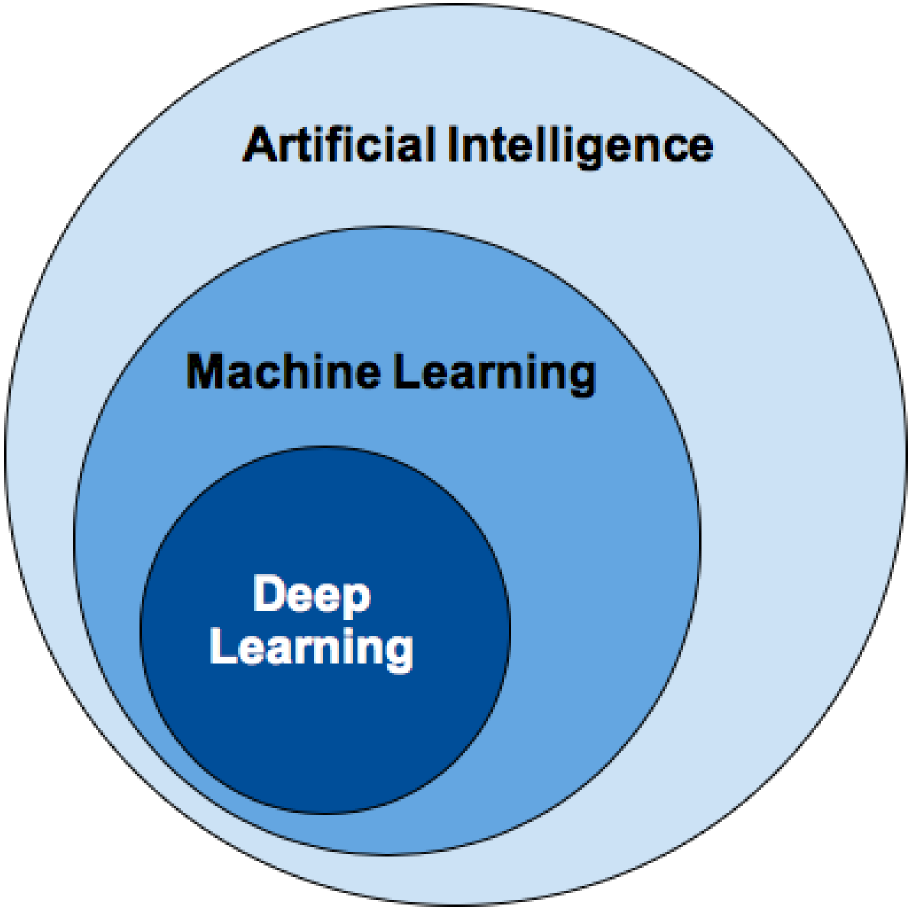
- What is Deep Learning, the field, about?

- The elephant in the room: ChatGPT, Stable Diffusion, existential risk, …

- What is this class about?
  - What to expect?
  - Logistics

- FAQ

Concepts



"Deep Learning is part of a broader family of **machine learning methods** based on **artificial neural networks**"

--- https://en.wikipedia.org/wiki/Deep_learning

ZK Caveat: Note it does not HAVE to be through ANNs; there are deep methods involving probabilistic graphical models (Boltzmann Machines, etc.). They just do not currently work and are not scalable.

Image Credit: https://www.sumologic.com/blog/machine-learning-deep-learning/

What is (general) intelligence?

- Boring textbook answer

The ability to acquire and apply knowledge and skills

  - Dictionary

- Many others
  - Survival, various types/aspects of intelligence, etc.

New Words!:

- **AGI – Artificial General Intelligence (~ as good as expert humans across most/all tasks)**
- **ASI – Artificial Super-Intelligence (self-improvement, etc.)**

What is artificial intelligence?

- Boring textbook answer

Intelligence demonstrated by machines

  – Wikipedia

- What others say:

The science and engineering of making computers behave in ways that, until recently, we thought required human intelligence.

  – Andrew Moore, CMU

- Squaring the two (artificial general intelligence) is not easy; how do we define or evaluate this?

What is machine learning?

- A favorite

Study of algorithms that
improve their performance (P)
at some task (T)
with experience (E)

  – Tom Mitchell, CMU

So what *is* Deep (Machine) Learning?

- **Objective:** Representation Learning
  - Automatically discover useful features/representations for a **task** from raw data

- **Model:** (Deep) Artificial Neural Networks

- **Learning Method:** Unsupervised/Supervised/Reinforcement/Generative/ <insert-qualifier-here> Learning

- **Simply**: Deep Learning

So what *is* Deep (Machine) Learning?

- A few different ideas:

- (Hierarchical) Compositionality
  - Cascade of non-linear transformations
  - Multiple layers of representations

- End-to-End Learning
  - Learning (goal-driven) representations
  - Learning to feature extraction

Hierarchical Compositionality

VISION

pixels → edge → texton → motif → part → object

SPEECH

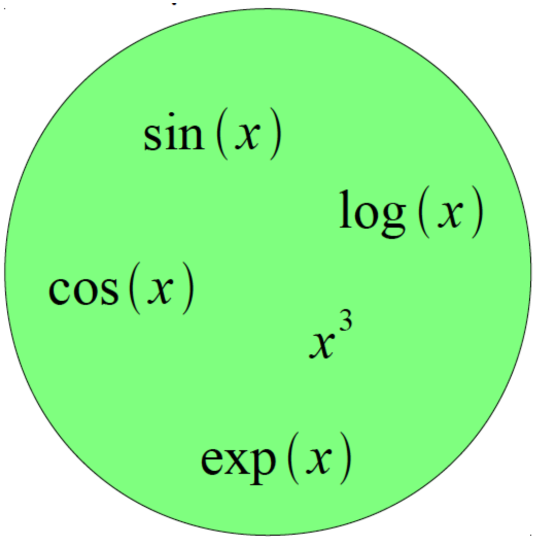
sample → spectral band → formant → motif → phone → word

NLP

character → word → NP/VP/.. → clause → sentence → story

Building A Complicated Function

Given a library of simple functions

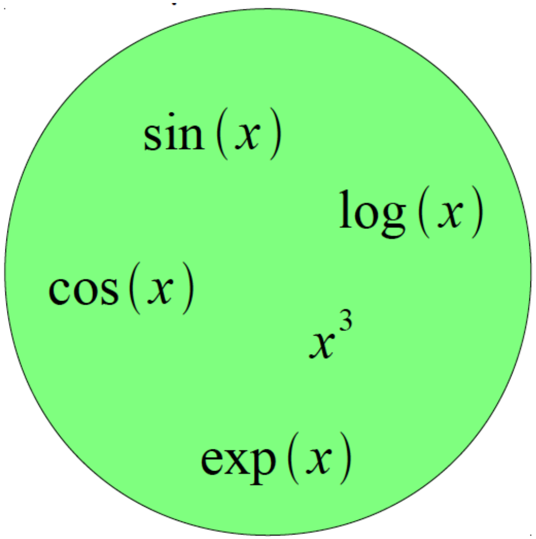


Compose into a

complicate function

Building A Complicated Function

Given a library of simple functions

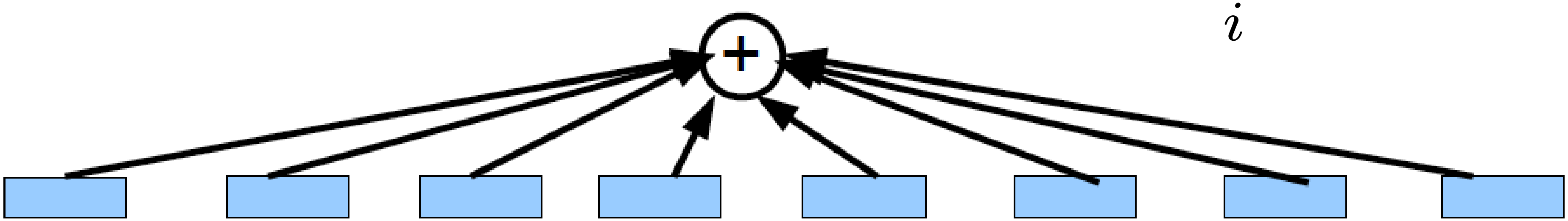


Compose into a complicate function

Idea 1: Linear Combinations

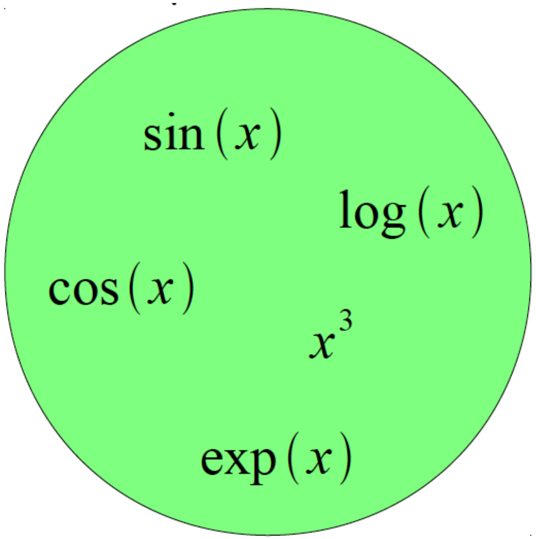
- Boosting
- Kernels
- …

$$f(x) = \sum_i \alpha_i g_i(x)$$



Building A Complicated Function

Given a library of simple functions



Compose into a

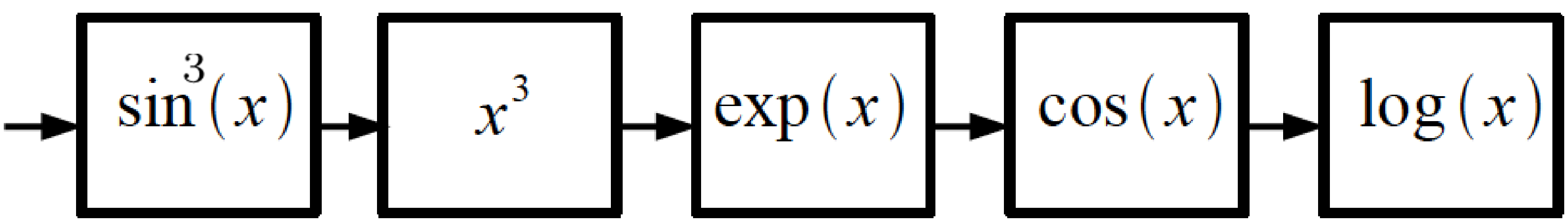
complicate function

Idea 2: Compositions

Compose a set of functions (layers) through which the input data get transformed.

More layers = "Deeper"

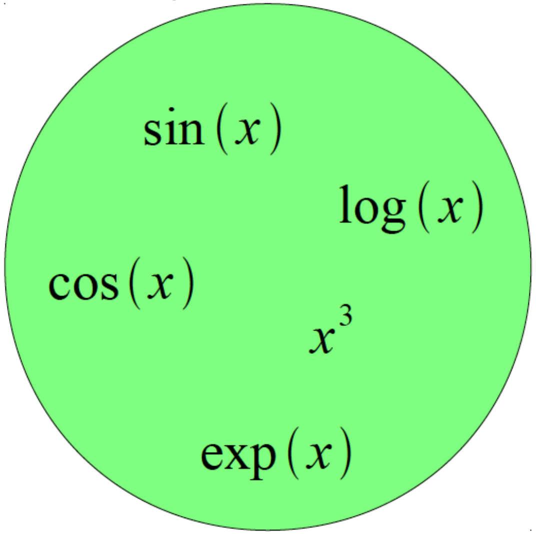
$$f(x) = \log(\cos(\exp(\sin^3(x))))$$



Can we make it more expressive?

Building A Complicated Function

Given a library of simple functions



Compose into a

complicate function

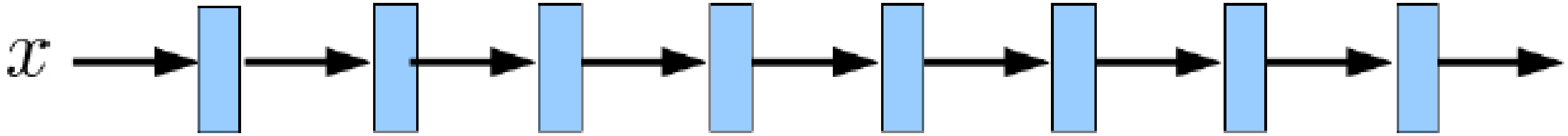
Yes! Parametric functions

Modern DNNs have huge # of parameters, on the orders of Billions

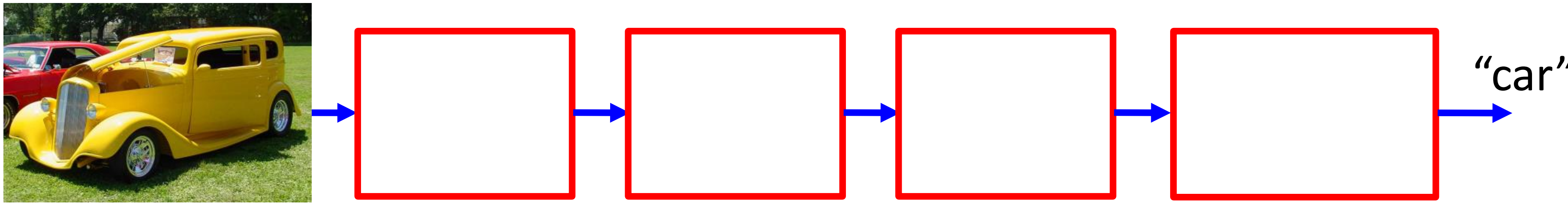
Modern DNNs have huge # of parameters, on the orders of bn's

Parametric functions

$$f_\theta(x) = \boxed{g_{\theta_n}}(\ldots g_{\theta_2}(g_{\theta_1}(x)\ldots))$$

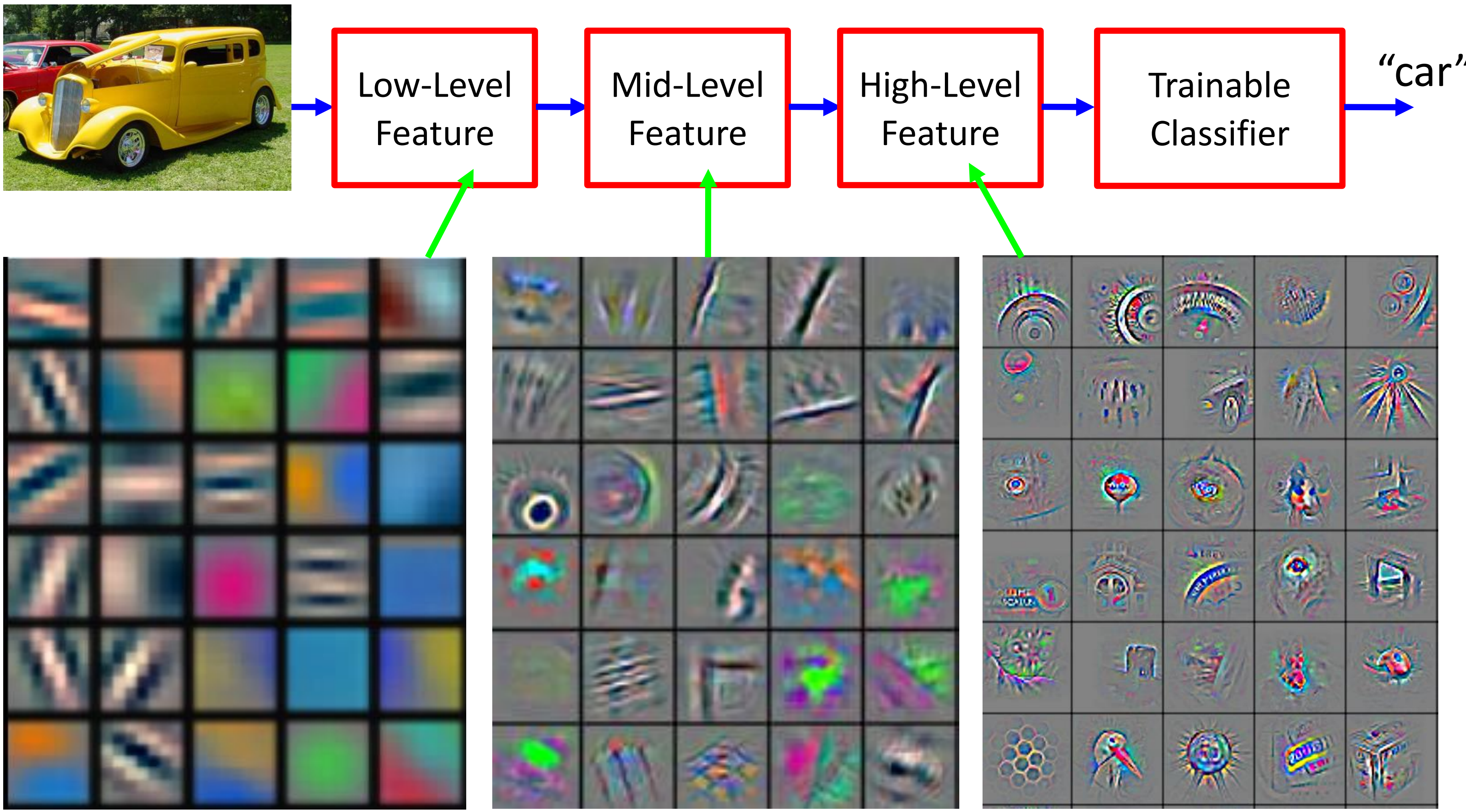


Deep Learning = Hierarchical Compositionality



Slide Credit: Marc'Aurelio Ranzato, Yann LeCun

Deep Learning = Hierarchical Compositionality



Feature visualization of convolutional net trained on ImageNet from [Zeiler & Fergus 2013]

Slide Credit: Marc Aurelio Ranzato, Yann LeCun

So what *is* Deep (Machine) Learning?

- A few different ideas:

- (Hierarchical) Compositionality
  - Cascade of non-linear transformations
  - Multiple layers of representations

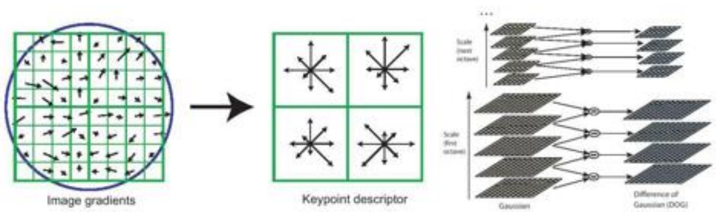
- End-to-End Learning
  - Learning (goal-driven) representations
  - Learning to feature extraction

“Shallow” vs Deep Learning

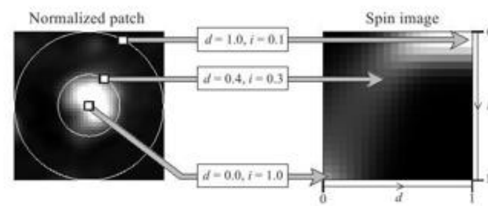
- “Shallow” models

hand-crafted Feature Extractor → “Simple” Trainable Classifier

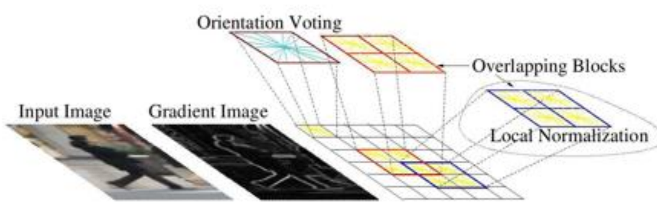
fixed | learned



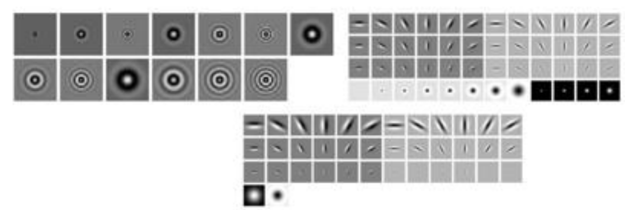
SIFT



Spin Images



HoG



Textons

and many many more….

Slide Credit: Marc'Aurelio Ranzato, Yann LeCun

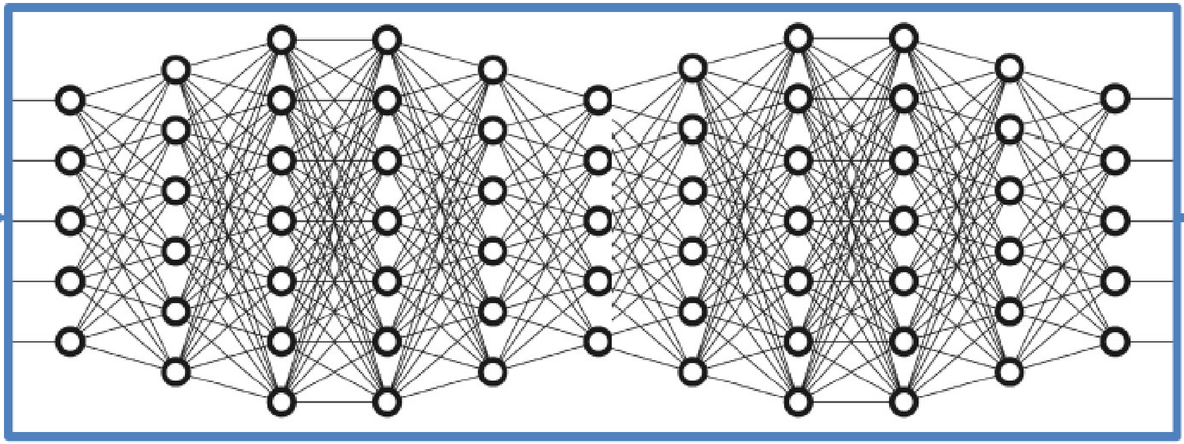
"Shallow" vs Deep Learning

- "Shallow" models

hand-crafted Feature Extractor
fixed

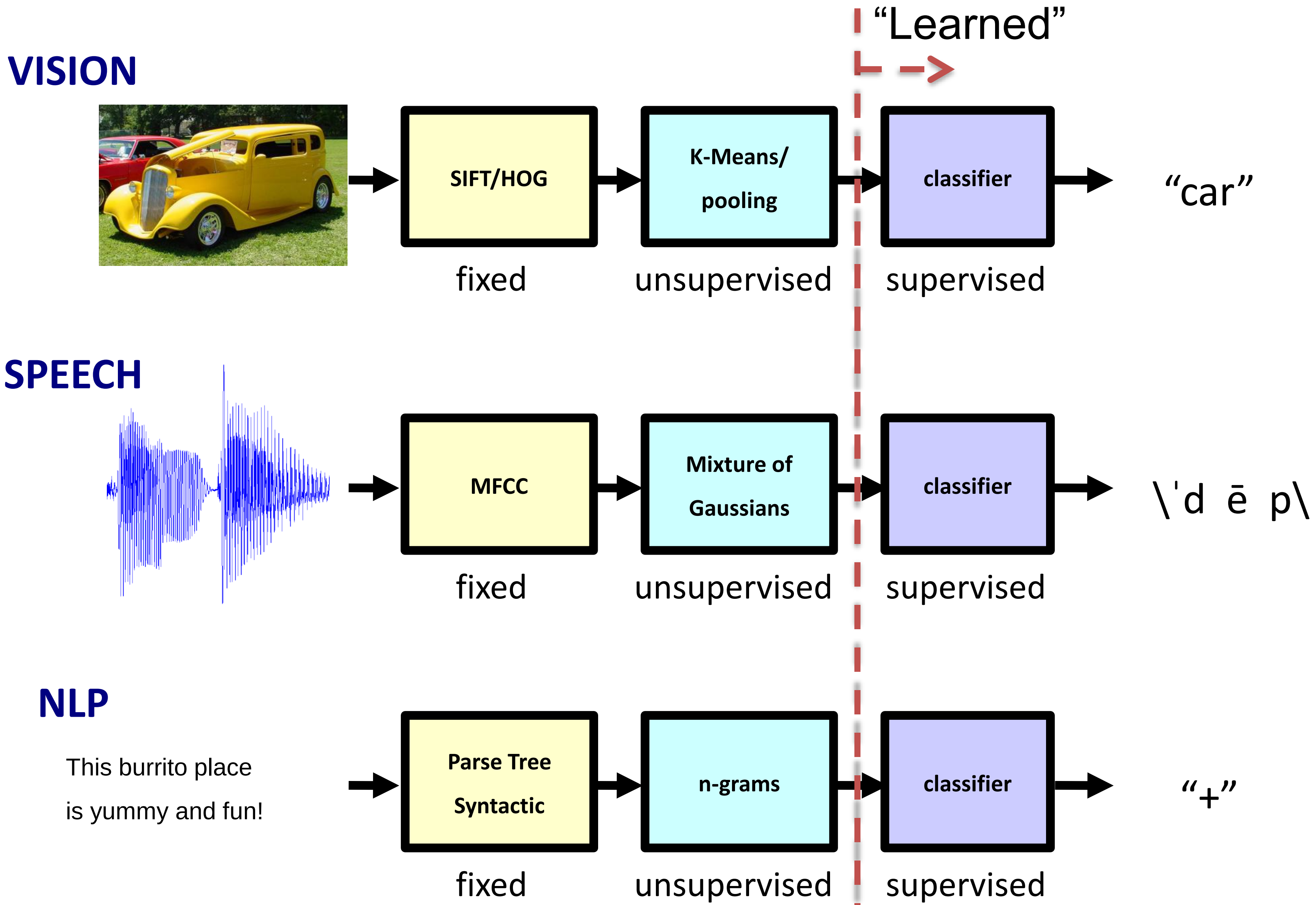
"Simple" Trainable Classifier
learned

- Deep models



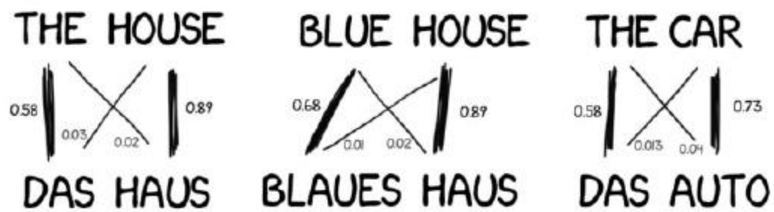
Slide Credit: Marc'Aurelio Ranzato, Yann LeCun

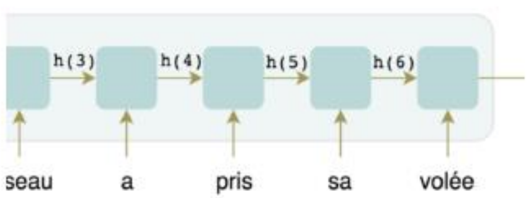
Deep Learning = End-to-End Learning

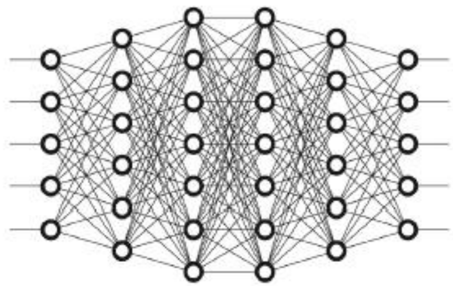


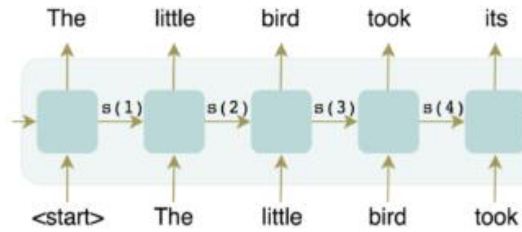
“Shallow” vs Deep Learning

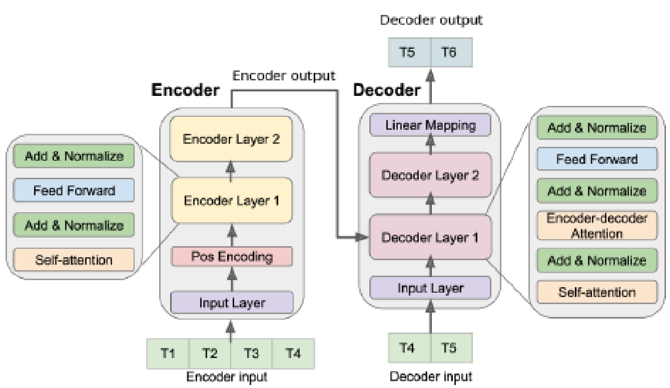
“Shallow” vs. deep language models









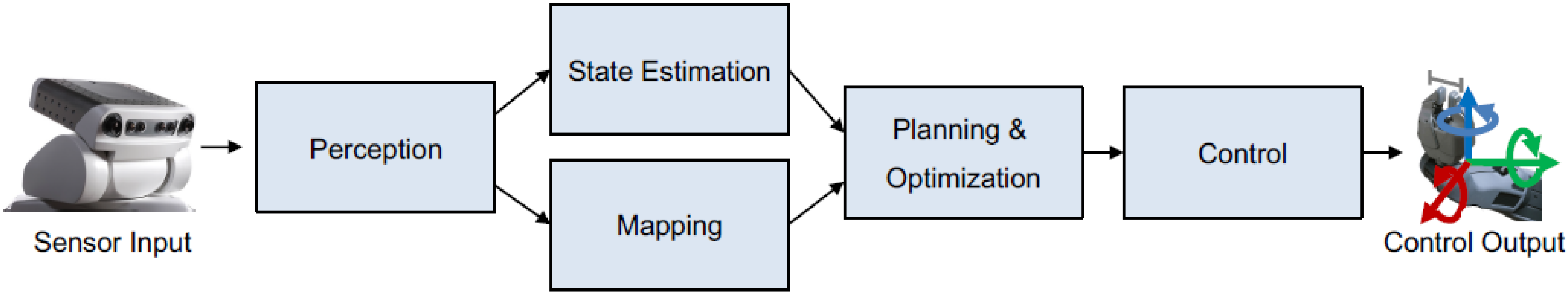


Transformer Models
(Vaswani *et al.*, 2017)

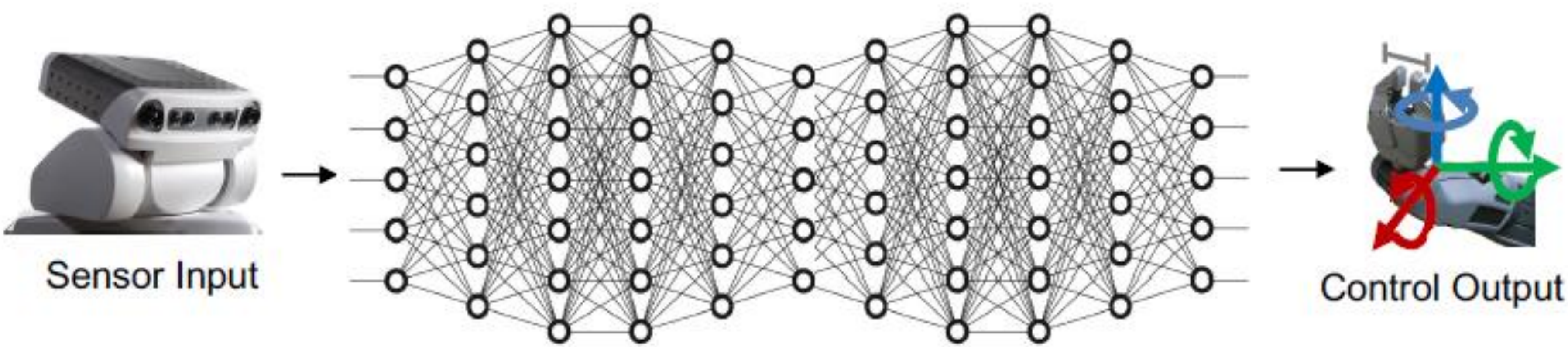
GPT-4

https://www.freecodecamp.org/news/a-history-of-machine-translation-from-the-cold-war-to-deep-learning-f1d335ce8b5/

"Pipelining" vs. "End-to-End Learning"



Hand-engineered pipelines



End-to-end learning
("pixel-to-torque")

Note: These can be mixed/combined!

Benefits of Deep/Representation Learning

- (Usually) Better Performance
  - Caveats: given enough data, similar train-test distributions, non-adversarial evaluation, etc., etc.

- New domains without “experts”
  - RGBD/Lidar
  - Multi-spectral data
  - Gene-expression data
  - Unclear how to hand-engineer

- “Homogenization” of model design
- New abilities emerge with more **data/parameter scale** and **compute**

"Expert" intuitions can be misleading

- *"Every time I fire a linguist, the performance of our speech recognition system goes up"*
  - Fred Jelinik, IBM '98

- *"Because gradient descent is better than you"*
  - *Yann LeCun,* CVPR '13

“The Bitter Lesson”

- “The biggest lesson that can be read from 70 years of AI research is that general methods that leverage computation are ultimately the most effective, and by a large margin. The ultimate reason for this is Moore's law, or rather its generalization of continued exponentially falling cost per unit of computation.” (Sutton, 2019)

http://www.incompleteideas.net/IncIdeas/BitterLesson.html

What about ChatGPT / Foundation Models / … buzzwords?

Bing's A.I. Chat: 'I Want to Be Alive. 😈'

In a two-hour conversation with our columnist, Microsoft's new chatbot said it would like to be human, had a desire to be destructive and was in love with the person it was chatting with. Here's the transcript.

Give this article 1.6K

https://www.nytimes.com/article/ai-artificial-intelligence-chatbot.html

ARTIFICIAL INTELLIGENCE

ChatGPT is about to revolutionize the economy. We need to decide what that looks like.

New large language models will transform many jobs. Whether they will lead to widespread prosperity or not is up to us.

By David Rotman March 25, 2023

https://www.technologyreview.com/2023/03/25/1070275/chatgpt-revolutionize-economy-decide-what-looks-like/

Exam	GPT-4	GPT-4 (no vision)	GPT-3.5
Uniform Bar Exam (MBE+MEE+MPT)	298 / 400 (~90th)	298 / 400 (~90th)	213 / 400 (~10th)
LSAT	163 (~88th)	161 (~83rd)	149 (~40th)
SAT Evidence-Based Reading & Writing	710 / 800 (~93rd)	710 / 800 (~93rd)	670 / 800 (~87th)
SAT Math	700 / 800 (~89th)	690 / 800 (~89th)	590 / 800 (~70th)
Graduate Record Examination (GRE) Quantitative	163 / 170 (~80th)	157 / 170 (~62nd)	147 / 170 (~25th)
Graduate Record Examination (GRE) Verbal	169 / 170 (~99th)	165 / 170 (~96th)	154 / 170 (~63rd)
Graduate Record Examination (GRE) Writing	4 / 6 (~54th)	4 / 6 (~54th)	4 / 6 (~54th)
USABO Semifinal Exam 2020	87 / 150 (99th - 100th)	87 / 150 (99th - 100th)	43 / 150 (31st - 33rd)
USNCO Local Section Exam 2022	36 / 60	38 / 60	24 / 60
Medical Knowledge Self-Assessment Program	75 %	75 %	53 %
Codeforces Rating	392 (below 5th)	392 (below 5th)	260 (below 5th)
AP Art History	5 (86th - 100th)	5 (86th - 100th)	5 (86th - 100th)
AP Biology	5 (85th - 100th)	5 (85th - 100th)	4 (62nd - 85th)
AP Calculus BC	4 (43rd - 59th)	4 (43rd - 59th)	1 (0th - 7th)
AP Chemistry	4 (71st - 88th)	4 (71st - 88th)	2 (22nd - 46th)
AP English Language and Composition	2 (14th - 44th)	2 (14th - 44th)	2 (14th - 44th)
AP English Literature and Composition	2 (8th - 22nd)	2 (8th - 22nd)	2 (8th - 22nd)
AP Environmental Science	5 (91st - 100th)	5 (91st - 100th)	5 (91st - 100th)
AP Macroeconomics	5 (84th - 100th)	5 (84th - 100th)	2 (33rd - 48th)
AP Microeconomics	5 (82nd - 100th)	4 (60th - 82nd)	4 (60th - 82nd)
AP Physics 2	4 (66th - 84th)	4 (66th - 84th)	3 (30th - 66th)
AP Psychology	5 (83rd - 100th)	5 (83rd - 100th)	5 (83rd - 100th)
AP Statistics	5 (85th - 100th)	5 (85th - 100th)	3 (40th - 63rd)
AP US Government	5 (88th - 100th)	5 (88th - 100th)	4 (77th - 88th)
AP US History	5 (89th - 100th)	4 (74th - 89th)	4 (74th - 89th)
AP World History	4 (65th - 87th)	4 (65th - 87th)	4 (65th - 87th)
AMC 10[3]	30 / 150 (6th - 12th)	36 / 150 (10th - 19th)	36 / 150 (10th - 19th)
AMC 12[3]	60 / 150 (45th - 66th)	48 / 150 (19th - 40th)	30 / 150 (4th - 8th)
Introductory Sommelier (theory knowledge)	92 %	92 %	80 %
Certified Sommelier (theory knowledge)	86 %	86 %	58 %
Advanced Sommelier (theory knowledge)	77 %	77 %	46 %
Leetcode (easy)	31 / 41	31 / 41	12 / 41
Leetcode (medium)	21 / 80	21 / 80	8 / 80
Leetcode (hard)	3 / 45	3 / 45	0 / 45

Table 1. GPT performance on academic and professional exams. In each case, we simulate the conditions and scoring of the real exam. We report GPT-4's final score graded according to exam-specific rubrics, as well as the percentile of test-takers achieving GPT-4's score.

GPT4 technical report, OpenAI, March 2023

What about ChatGPT / Foundation Models / … buzzwords?

A *grossly simplified* timeline of deep learni

course

Pioneers	Foundations	…plosion	The Era of Large Models
1940-1960s	1970	2010-2019	2020-???

…

- Artificial Neural Networks
- Hebbian theory (Hebb)
- Perceptron (Rosenblatt)
- Visual cortex cells (Hubel and Wiesel)
- …

Some interesting neuro-inspired math & theory

- ConvNets (Fukus
- Backpropagation Williams, Hinton,
- LSTM (Hochreiter
- GPUs
-

Computational mechanisms and initial **architectures**

Unified code framework (Caffe, PyTorch, TensorFlow). Layer/regularization improvements

- Net (Jia, Fei-Fei, …)
- et (Krizhevsky)
- at Experiment (Google)
- Transformer
- Deep Generative Models
- AlphaGo
- Deep Reinforcement Learning
- More GPUs
- …

Distributed representations are effective, **transfer learning** works!

With sufficient **labeled data**, it works better than feature engineering!

Attention, **homogenization of scalable architectures**

- GPT Series (OpenAI)
- CLIP & DALL·E (OpenAI)
- SegmentAnything (Meta)
- Llama Series (Meta)
- Bard (Google)
- StableDiffusion
- …

Self-Supervised Learning – We can learn features w/o labels!

Let's just **SCALE** it. Lots of engineering. **Unified models for multiple tasks.**

Scale results in **emergent capabilities**

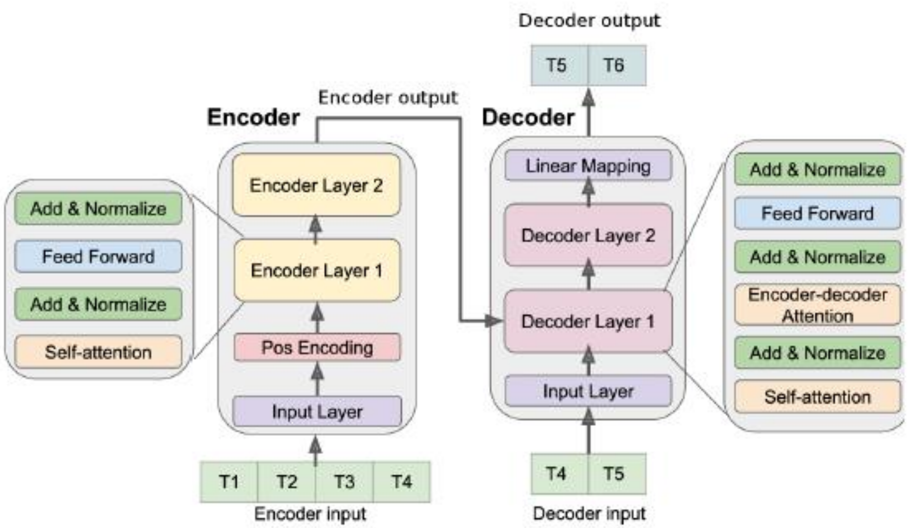
The final methods are simple. Just needed to know it will work, explore the design space, compute, data, engineering.

Homogenization of Deep Learning

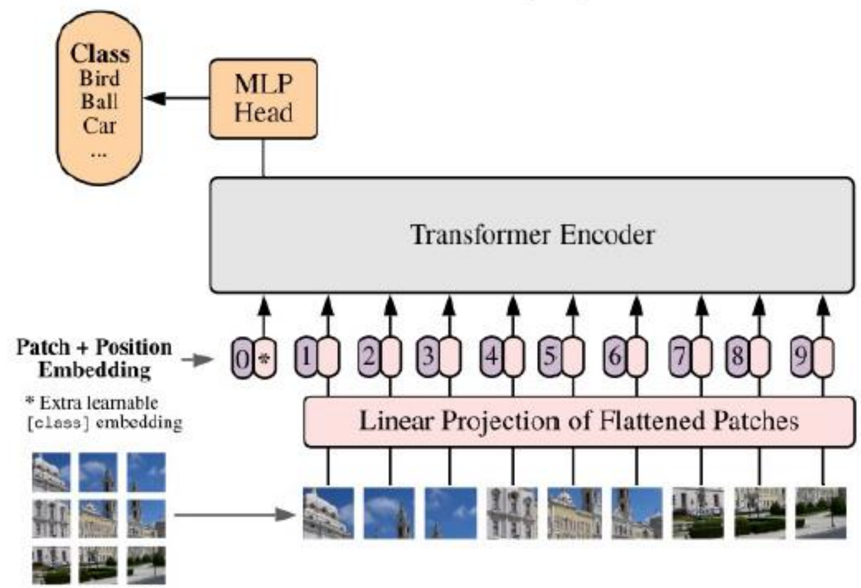
Homogenization is the **consolidation** of methodologies for building machine learning systems across a wide range of applications.

- Enabled by modular, plug-n-play nature of neural networks and training
- Consequence: Multi-modal, unified architectures, unified tasks (next-token prediction)

Example: The Transformer Models (Vaswani et al., 2017)



Transformer Models originally designed for NLP



Almost identical model (Visual Transformers) can be applied to Computer Vision tasks

Emergence of new behaviors

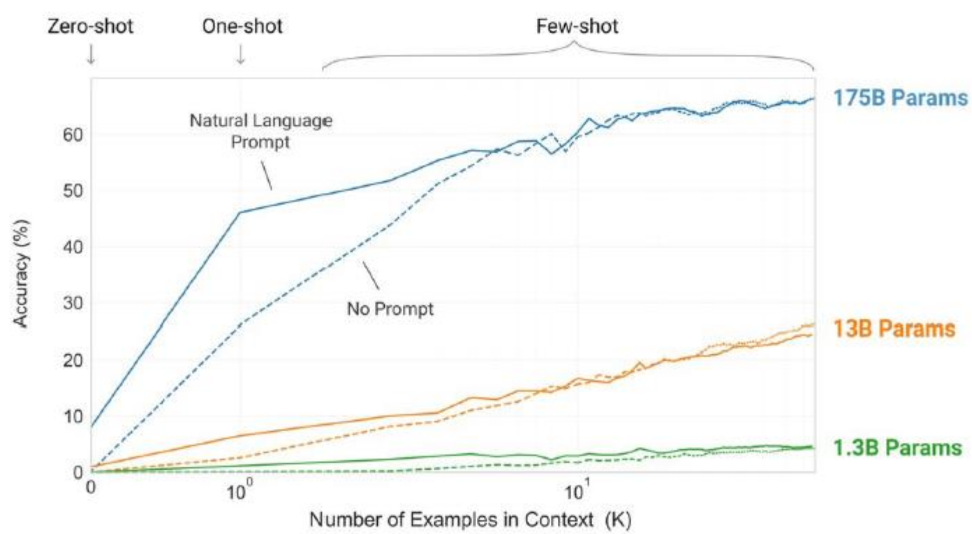
Emergence means that the behavior of a system is implicitly induced rather than explicitly constructed. For Deep Learning, emergence is often induced by larger model & more data.

Example: Compared to GPT-2's 1.5B parameter parameter model, GPT-3's 175-billion model permits "prompting" and "in-context learning", i.e., adapting to a new task simply by describing task.

Example input (prompt):

Ask it to translate French to English

$$\underbrace{\text{maison} \to \text{house, chat} \to \text{cat, chien} \to}_{\text{prompt}} \underbrace{\text{dog}}_{\text{completion}} .$$



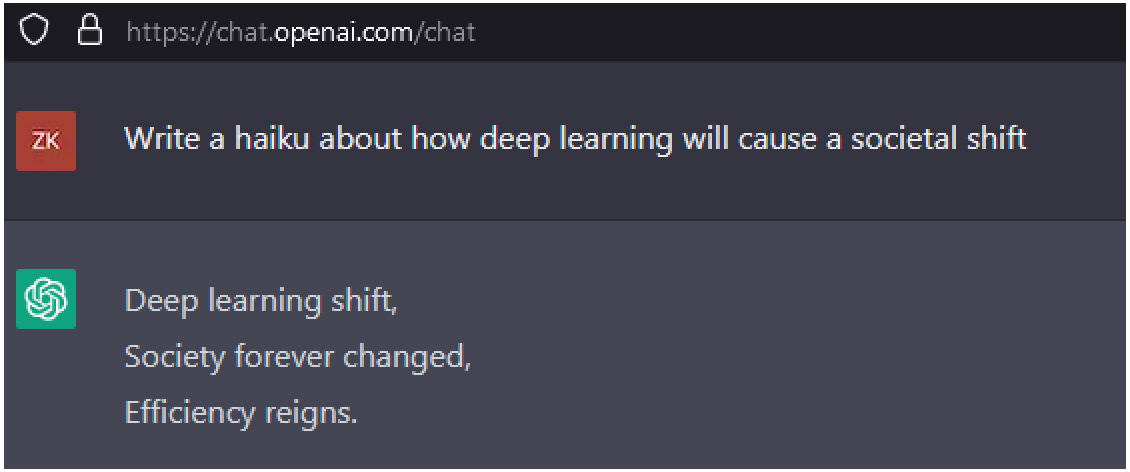
Bommasani et al., On the Opportunities and Risks of Foundation Models

Societal Change is Coming

- GitHub Copilot, ChatGPT, etc. are now useful enough to **speed up higher-level human work!**

GitHub Copilot

https://gamefromscratch.com/dall-e-vs-stable-diffusion-vs-midjour



But likely will not be as crazy fast or much as the hype suggests

Problems with Deep Learning

- Problem#1: Lack of a formal understanding
  - Non-Convex! Non-Convex! Non-Convex!
    - Depth>=3: most losses non-convex in parameters
  - Worse still, existing intuitions from classical statistical learning theory don't seem to carry over.
  - Theoretically, we are stumbling in the dark here

- Standard response #1
  - "Yes, but this just means there's new theory to be constructed"
  - "All interesting learning problems are non-convex"
    - For example, human learning
      - Order matters → wave hands → non-convexity

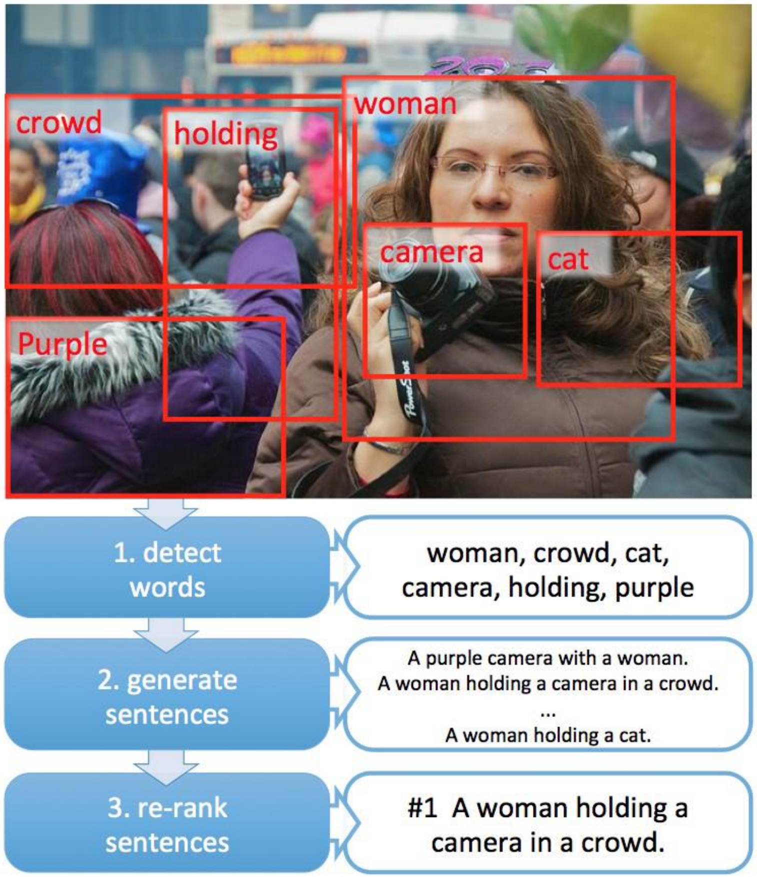
- Standard response #2
  - "Yes, but it often works!"

Problems with Deep Learning

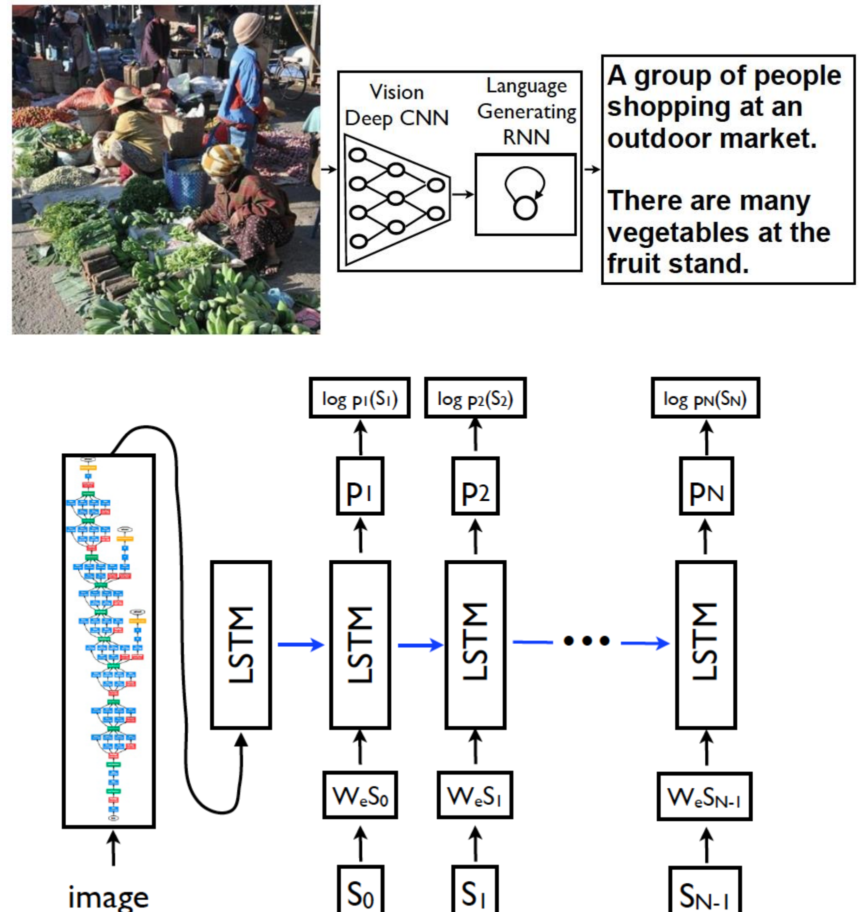
- Problem#2: Lack of interpretability
  - Hard to track down what's failing
  - Pipeline systems have expected performances at each step
  - In end-to-end systems, it's hard to know why things are not working

Problems with Deep Learning

- Problem#2: Lack of interpretability



[Fang et al. CVPR15]



[Vinyals et al. CVPR15]

Pipeline

End-to-End

Problems with Deep Learning

- Problem#2: Lack of interpretability
  - Hard to track down what's failing
  - Pipeline systems have "oracle" performances at each step
  - In end-to-end systems, it's hard to know why things are not working

- Standard response #1
  - Tricks of the trade: visualize features, add losses at different layers, pre-train to avoid degenerate initializations…
  - "We're working on it"

- Standard response #2
  - "Yes, but it often works!"

Problems with Deep Learning

- Problem#3: Lack of easy reproducibility
  - Direct consequence of stochasticity & non-convexity
    - different initializations → different local minima
    - Almost everything matters! (hyper-parameters, small design decisions, etc.)
  - More recently: Privatization of unknown models trained on unknown data

- Standard response #1
  - It's getting much better
  - Standard toolkits/libraries/frameworks now available
  - PyTorch, TensorFlow, MxNet…

- Standard response #2
  - "Yes, but it often works!"

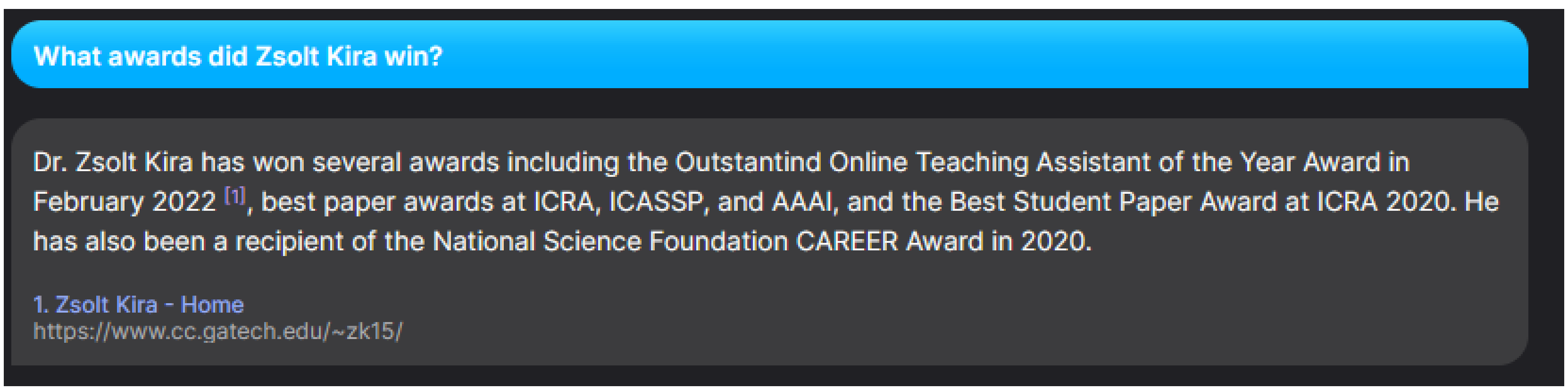
Problems with Deep Learning

- Problem#4: Still not robust to out-of-distribution data
  - Even training on “entire internet” just bypasses this:
    - For domains you care about it may still not generalize well
    - Domains that dominate the data will dominate performance profile
- Lots of research into this, but lack of formal understanding hinders this
  - Most ML theory deals with i.i.d. train/test case, or some simplified model of shift

Consequences

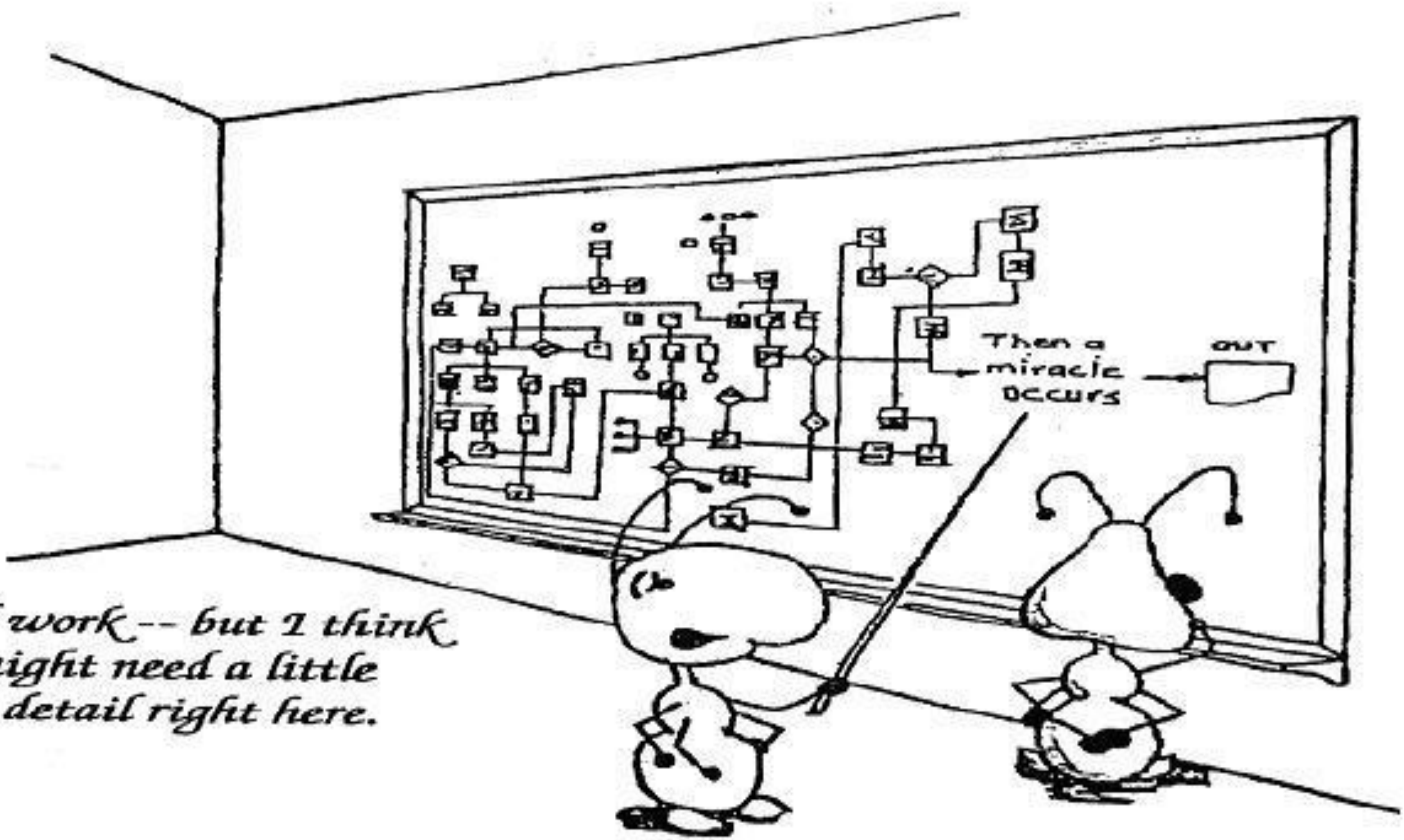
- As a consequence, general issue of **safety and correctness**
  - No explicit reasoning or logical mechanisms

- **Example:**
  - Tesla crashes
  - Language models hallucinating



https://you.com/search?q=who+are+you&tbm=youchat

Yes it works, but how?



Good work -- but I think we might need a little more detail right here.

Outline

- What is Deep Learning, the field, about?
  - Highlight of some recent projects from my lab
- What is this class about?
  - What to expect?
  - Logistics
- FAQ

Outline

- What is Deep Learning, the field, about?
  - Highlight of some recent projects from my lab

- What is this class about?
  - What to expect?
  - Logistics

- FAQ

What is this class about?

- Introduction to Deep Learning

- Goal:
  - After finishing this class, you should be ready to get started on your first DL research project.
    - CNNs (Image data)
    - RNNs / Transformers (sequential/graph data)
    - Deep Reinforcement Learning (decision-making)
    - Generative Models (VAEs, Diffusion Models, GANs) (unsupervised learning)

- Target Audience:
  - Senior undergrads, MS-ML, and new PhD students

What this class is NOT

- NOT the target audience:
  - Students without sufficient background knowledge (Python, linear algebra, calculus, basic probability & statistics)
  - Advanced grad-students already working in ML/DL areas
  - People looking for an in-depth understanding of a research area that uses deep learning (3D Vision, Large Language
  - Models, Deep RL, etc.).
- NOT the goal:
  - Intro to Machine Learning
  - Teaching a toolkit. “Intro to TensorFlow/PyTorch”

Caveat

- This is an ADVANCED Machine Learning class
  - This should NOT be your first introduction to ML
  - You will need a formal class; not just self-reading/coursera
  - If you took CS 7641/ISYE 6740/CSE 6740 @GT, you're in the right place
  - If you took an equivalent class elsewhere, see list of topics taught in CS 7641 to be sure.

Prerequisites

- Python programming
  - Basic knowledge of numerical computation & tools (e.g. numpy)
  - HW1 (pure python), HW2 (python + PyTorch),
    HW3+4 (PyTorch)
  - Your language of choice for project
- Intro Machine Learning
  - Classifiers, regressors, loss functions, MLE, MAP
- Linear Algebra
  - Matrix multiplication, eigenvalues, positive semi-definiteness…
- Calculus
  - Multi-variate gradients, hessians, jacobians…
- Must read (on W3 reading list): Matrix calculus for deep learning
  - https://explained.ai/matrix-calculus/index.html

Course Information

- Instructor: Zsolt Kira
  - [censored]@gatech.edu **(use piazza public/private instead!)**

TAs

Yipu Chen

Mili Das

Kausar Patherya

Office Hours

- TA Office Hours:
  - Virtual over zoom
  - Check course website for OH slots and zoom links
  - Start next week

- Zsolt’s Office Hours:
  - Virtual over Zoom
  - **No assignment (PS/HW) questions**
  - Lecture content / project ideas / administrative / career advice, …

Organization & Deliverables

- 3 problem-sets+homeworks (69%)
  - Mix of theory (PS) and implementation (HW)
  - First one goes out next week
    - Start early, Start early, Start early, Start early, Start early, Start early, Start early, Start early, Start early, Start early

- Final project (31%)
  - Projects done, recommended in groups of 3-4
  - You need a good reason to do a solo project
  - Mid-semester project proposal before project period starts
  - **Find a team ASAP! Talk to people, use Piazza “find a teammate” post**

- (Bonus) Class Participation (1%)
  - Top (endorsed) contributors on Piazza

Plenty of “buffer” built in

- Grace period
  - 2 days grace period
    - Intended for *checking* submission NOT to replace due date
    - No need to ask for grace, no penalty for turning it in within grace period
    - Can NOT use for PS0/HW0
  - After grace period, you get a 0 (no excuses except medical)
    - Send all medical requests to dean of students (https://studentlife.gatech.edu/)
    - Form: https://gatech-advocate.symplicity.com/care_report/index.php/pid224342?
  - **DO NOT SEND US ANY MEDICAL INFORMATION!** We do not need any details, just a confirmation from dean of students

GT Resources for Mental Health

Georgia Tech Police Department
Emergency: Call 911 | 404-894-2500

Dean of Students Office
404-894-2565 | studentlife.gatech.edu
Afterhours Assistance Line & Dean on Call: 404-894-2204

Center for Assessment, Referral and Education (CARE)
404-894-3498 | care.gatech.edu

Collegiate Recovery Program
404-894-2575 | counseling.gatech.edu

Counseling Center
404-894-2575 | counseling.gatech.edu

Health Initiatives
404-894-9980
healthinitiatives.gatech.edu

LGBTQIA Resource Center
404-385-4780 | lgtbqia.gatech.edu

Stamps Psychiatry Center
404-894-1420

VOICE
404-385-4464 | 404-385-4451
24/7 Info Line: 404-894-9000 | voice.gatech.edu

Women's Resource Center
404-385-0230 | womenscenter.gatech.edu

Veterans Resource Center
404-894-4953 | veterans.gatech.edu

Georgia Crisis and Access Line
1-800-715-4225
The crisis line is staffed with professional social workers and counselors 24 hours per day, every day, to assist those with urgent and emergency needs.

Trevor Project
1-866-488-7386
Trained counselors are available to support anyone in need.

National Suicide Prevention Hotline
1-800-273-8255
A national network of local crisis centers that provides free and confidential emotional support to people in suicidal crisis or emotional distress 24/7.

Georgia State Psychology Clinic
404-413-2500
The clinic offers high quality and affordable psychological services to adults, children, adolescents, families and couples from the greater Atlanta area.

PS0

- Out already; due Sunday Jan 18th (no grace period)
  - Available on class webpage + Canvas
  - If not registered yet (on waitlist), see webpage FAQ for form to request gradescope access

- Grading
  - Not counted towards your final grade, but required
  - <=75% means that you might not be prepared for the class
  - We may not be able to grade before registration ends if submit later than Thursday morning

- Topics
  - PS: probability, calculus, convexity, proving things

Project

- Goal
  - Chance to try Deep Learning
  - Encouraged to apply to your research (computer vision, NLP, robotics,…)
  - Must be done this semester.
  - Can combine with other classes, but **separate thrust**
    - get permission from both instructors; delineate different parts
  - 2-4 members (outside of this requires approval)

- Main categories
  - Application/Survey
    - Compare a bunch of existing algorithms on a new application domain of your interest
  - Formulation/Development
    - Formulate a new model or algorithm for a new or old problem
  - Theory
    - Theoretically analyze an existing algorithm

Computing

- Major bottleneck
  - GPUs

- Options
  - Your own / group / advisor's resources

  - Google Colab
    - jupyter-notebook + free GPU instance

  - PACE-ICE

  - Google Cloud credits (details TBA)
    - Tutorial on setting up gloud: https://github.com/cs231n/gcloud

Outline

- What is Deep Learning, the field, about?
  - Highlight of some recent projects from my lab
- What is this class about?
  - What to expect?
  - Logistics
- FAQ

4644 vs 7643

- Level differentiation
- HWs
  - Extra credit questions for 4644 students, necessary for 7643
- Project
  - Higher expectations from 7643
- Different grade cutoffs

Waitlist / Audit / Sit in

- Waitlist
  - Waitlist are mostly full. Class size will likely increase closer to room size
  - Do PS0/HW0 **NOW**. Come to first few classes.
  - Hope people drop.

- "I need this class to graduate"
  - Talk to your degree program advisor. They control the process of making sure you have options to graduate on time.

- Audit or Pass/Fail
  - No. We will give preference to people taking class for credit.

- Sitting in
  - Welcome to if space allows

What is the re-grading policy?

- Homework assignments
  - **Within 1 week** of receiving grades: see the TAs

- This is an advanced grad class.
  - The goal is understanding the material and making progress towards our research.

What is the collaboration policy?

- Collaboration
  - Only on HWs and project (not allowed in PS0/HW0).
  - You may discuss the questions
  - **Each student writes their own answers**
  - Write on your homework anyone with whom you collaborate
  - Each student must write their own code for the programming part

- DO NOT give access to your materials to anyone, even if the intent is not to copy. This is a violation!

- Zero tolerance on plagiarism
  - Neither ethical nor in your best interest
  - Always credit your sources
  - Don’t cheat. We will find out.

Deep Learning is So Good..

- That I had to put this slide in
- Our policy on ChatGPT/Co-Pilot/etc. is on the webpage

- tldr; treat it like a human collaborator – you can talk to it, learn from it, but **never directly copy from it**

https://chat.openai.com/chat

How do I get in touch?

- Primary means of communication -- **Piazza**
  - No direct emails to Instructor unless private information that cannot go on private piazza post
  - Instructor/TAs can provide answers to everyone on forum
  - Class participation credit for answering questions!
  - No posting answers. We will monitor.
  - Stay respectful and professional

Research

- "Can I work with your group for funding/credits/neither?"
  - Fill out this form