

CS 4803 / 7643: Deep Learning

Website: https://www.cc.gatech.edu/classes/AY2022/cs7643_spring/

Piazza: <http://piazza.com/gatech/spring2022/cs46447643a/>
(code: DLSPR2022)

Canvas: <https://gatech.instructure.com/courses/249278> (4803)
<https://gatech.instructure.com/courses/236394> (7643)

Gradescope: <https://www.gradescope.com/courses/346893> (4803)
<https://www.gradescope.com/courses/346907> (7643)

Zsolt Kira

School of Interactive Computing
Georgia Tech

Elephant in the room

- These times are filled with change and uncertainty
- Hope everyone is staying safe and healthy.
 - Please make use of GT resources for vaccination+boosting, surveillance testing, as well as masking/social distancing.
- Let's make the best of it.

Are you in the right place?

- This is CS 4803(DL) / CS 7643
 - “On campus” class

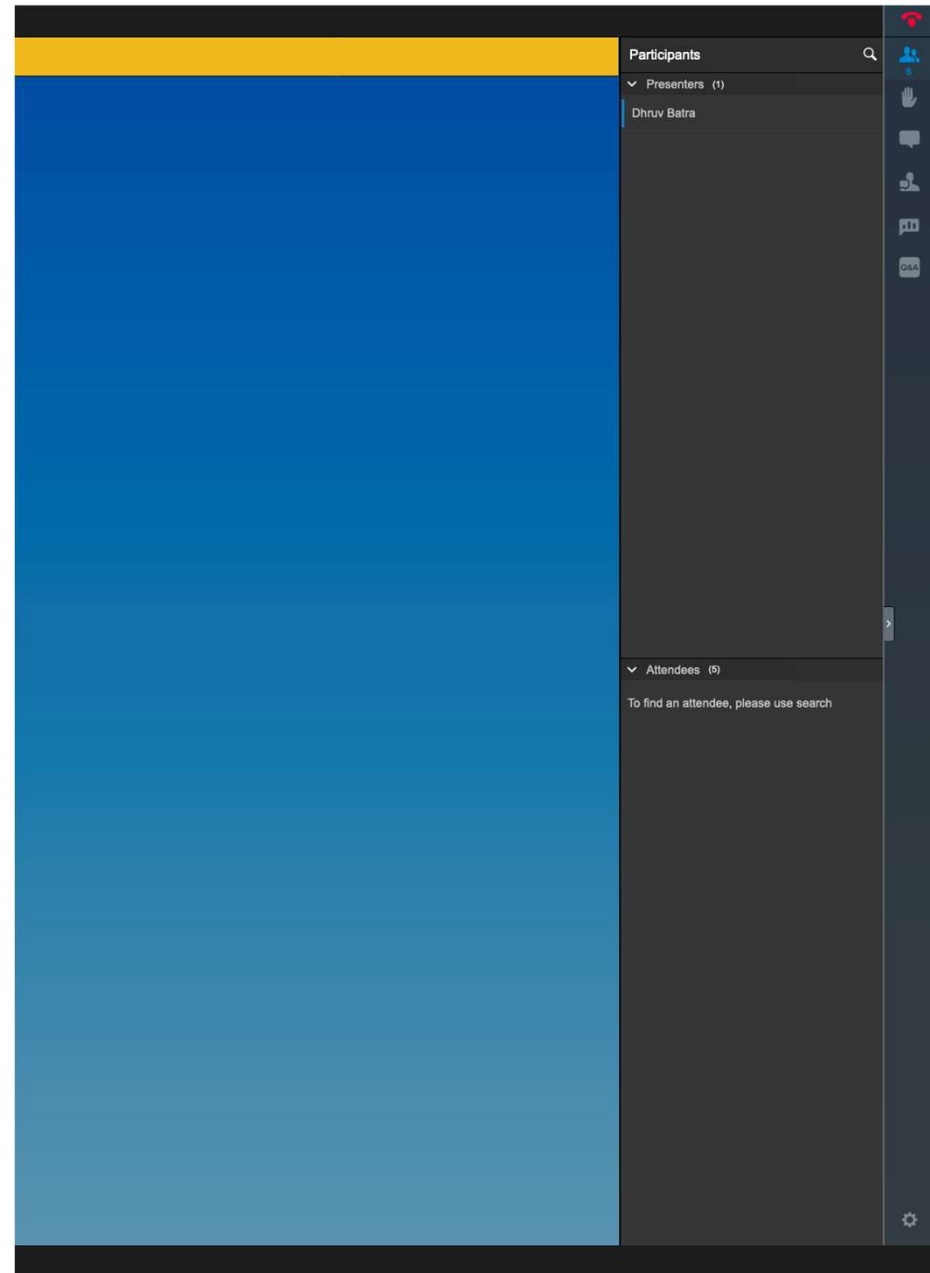
- This is NOT CS 7643-O01/OAN/Q/R
 - Online class for OMSCS program

Spring 21 Delivery Format

- Remote
 - No in-person interaction
 - Lectures, office hours, HW/project submissions online
 - No exam
- Sync
 - There is a scheduled “live” lecture time
 - Mix of live lectures and/or pre-recorded lectures played w/ Q&A
- Recording
 - Lectures are recorded and available for viewing
 - We **STRONGLY** encourage you to attend the lectures
- **Remember: Content is free online.**
 - **You are here for the interaction and the insight.**

How to interact

- Questions
 - Q&A
 - Chat
 - For responding to our questions
 - We'll explicitly stop and take questions periodically
 - But feel free to ask in between
- BlueJeans Event
 - ~10 sec lag



Outline for Today

- What is Deep Learning, the field, about?
- What is this class about?
 - What to expect?
 - Logistics
- FAQ

Outline

- What is Deep Learning, the field, about?
- What is this class about?
 - What to expect?
 - Logistics
- FAQ

What are we here to discuss?

**Some of the most exciting
developments in**

**Machine Learning,
Vision, NLP, Speech, Robotics
& AI in general**

in the last decade!

Demo time

vqa.clouddcv.org

demo.visualdialog.org

Result for Visual Question Answering



What is in this picture?

Submit

Predicted top-5 answers with confidence:

tennis ball	19.539%
ball	14.667%
tennis	9.942%
skateboarder	7.005%
racket	6.827%

Credits

Built by [@rishabh](#) & [@deshrai](#)

Result for Visual Question Answering



Is there a person driving a car?

Submit

Predicted top-5 answers with confidence:



Credits

Built by [@rishabh](#) & [@deshrui](#)

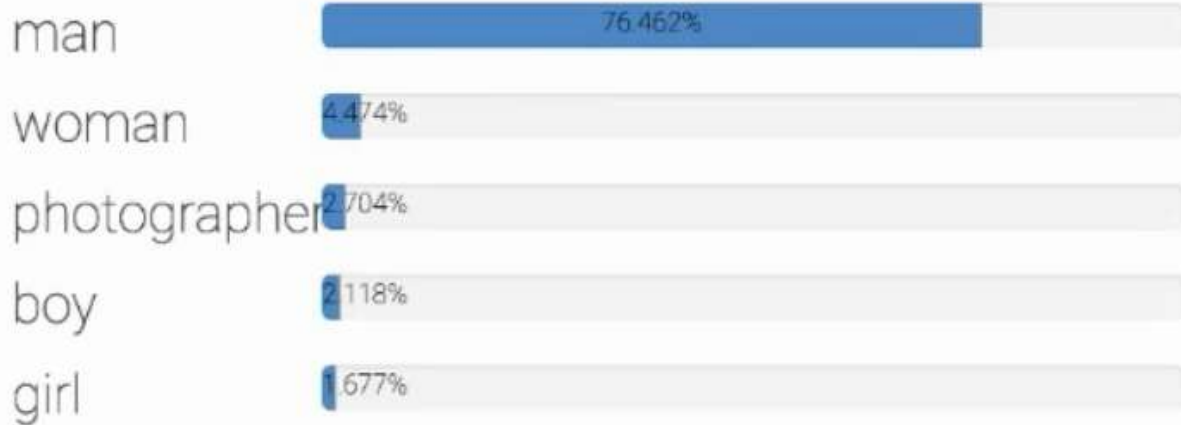
Result for Visual Question Answering



Who is this person?

Submit

Predicted top-5 answers with confidence:



Credits

Built by [@rishabh](#) & [@deshrai](#)

Result for Visual Question Answering



What type of court is the person playing on?

Submit

Predicted top-5 answers with confidence:



Credits

Built by [@rishabh](#) & [@deshrui](#)

(C) Dhruv Batra & Zsolt Kira

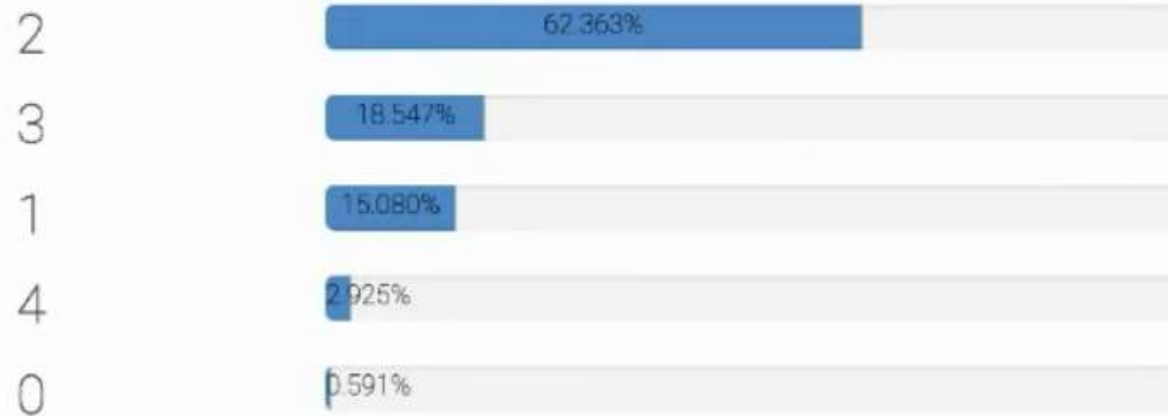
Result for Visual Question Answering



How many people are in the picture?

Submit

Predicted top-5 answers with confidence:



Credits

Built by [@rishabh](#) & [@deshrai](#)

Updated 2018 Model

Result for Visual Question Answering



How many people are in the picture?

Submit

Predicted top-5 answers with confidence:



Credits

Result for Visual Question Answering



What is the answer of Life, Universe, and everything?

Submit

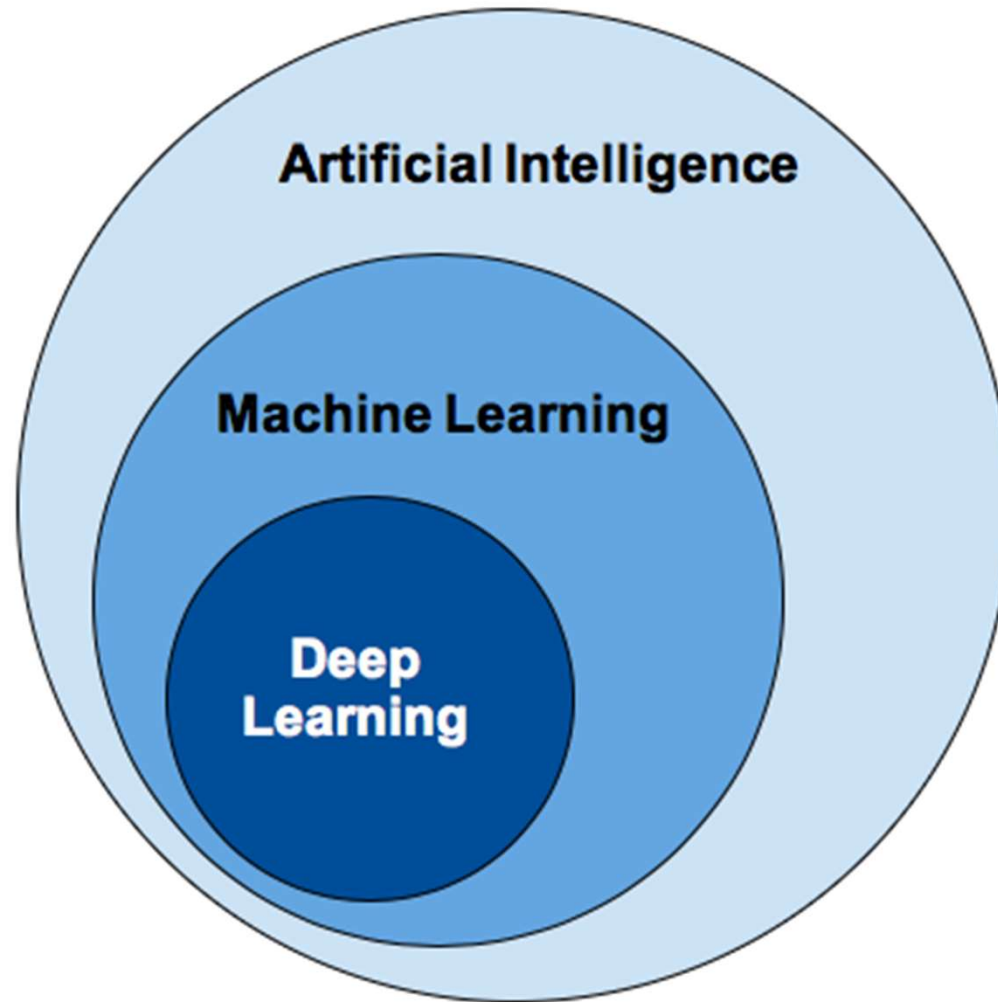
Predicted top-5 answers with confidence:



Credits

Built by [@rishabh](#) & [@deshrai](#)

Concepts



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.

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

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?

- Representation Learning
- Neural Networks
- Deep Unsupervised/Reinforcement/Structured/
<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
 - Distributed Representations
 - No single neuron “encodes” everything
 - Groups of neurons work together

Hierarchical Compositionality

VISION

pixels → edge → texture → motif → part → object

SPEECH

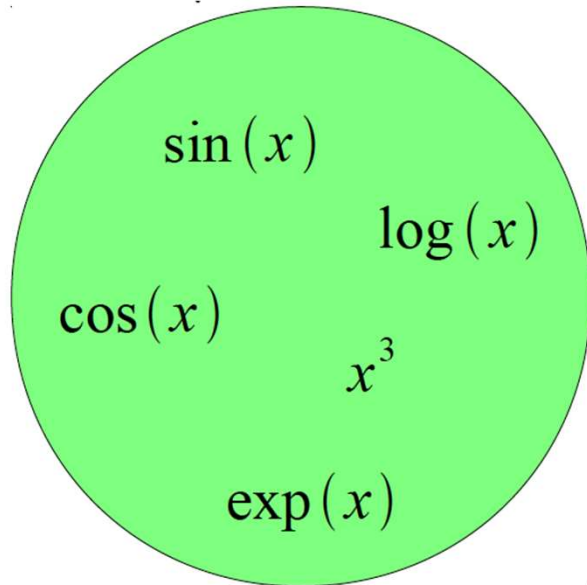
sample → spectral band → formant → motif → phone → word


NLP

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

Building A Complicated Function

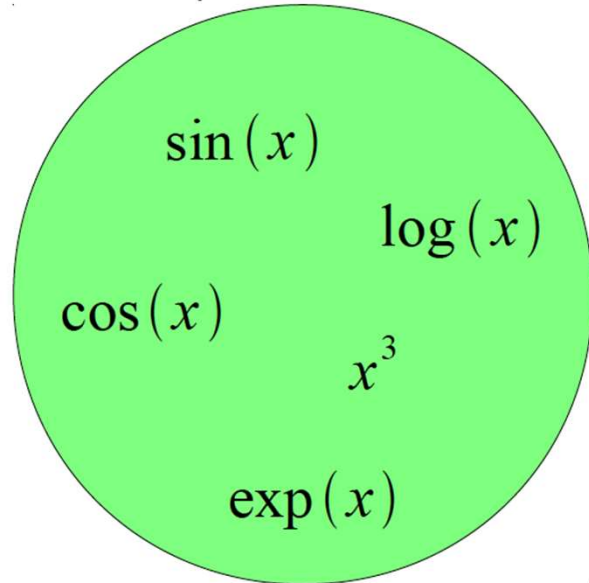
Given a library of simple functions



Compose into a

complicate function

Building A Complicated Function

Given a library of simple functions

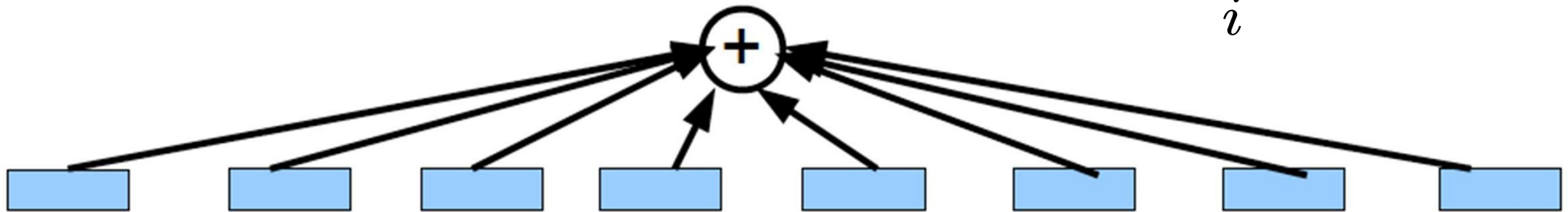


Compose into a
→
complicate function

Idea 1: Linear Combinations

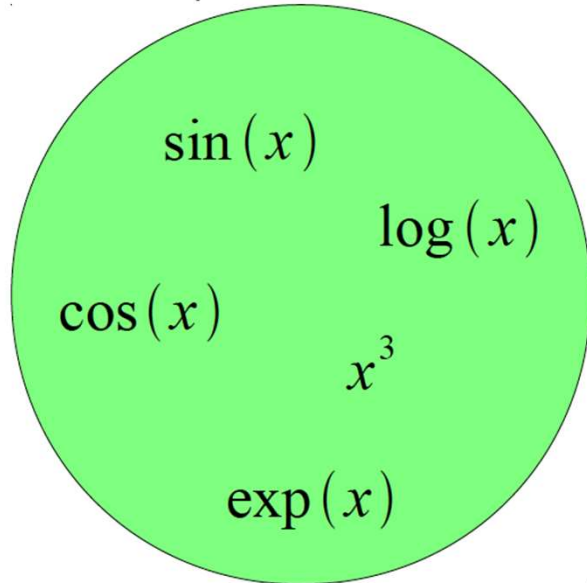
- Boosting
- Kernels
- ...


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



Building A Complicated Function

Given a library of simple functions

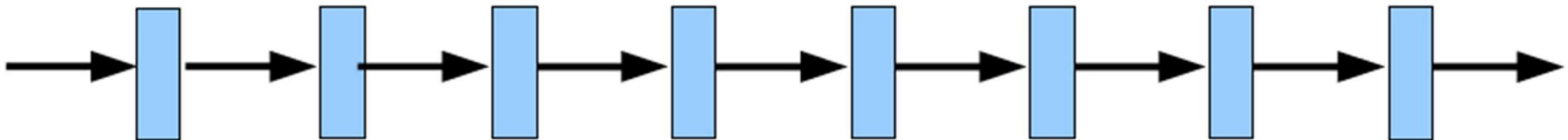


Compose into a

complicate function

Idea 2: Compositions

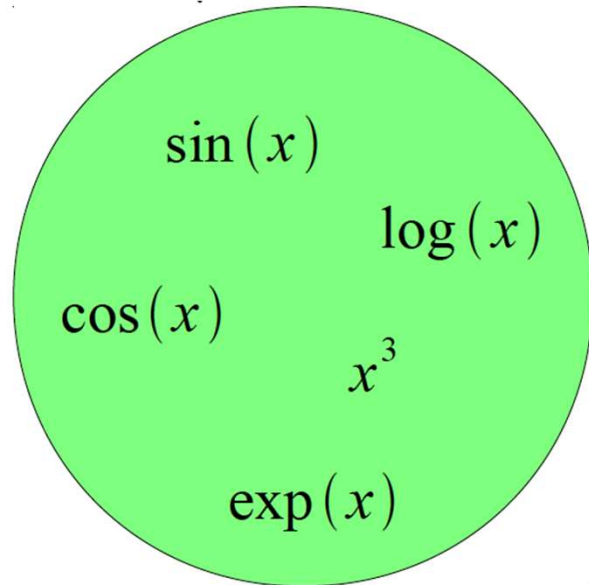
- Deep Learning
- Grammar models
- Scattering transforms...

$$f(x) = g_1(g_2(\dots(g_n(x)\dots)))$$



Building A Complicated Function

Given a library of simple functions

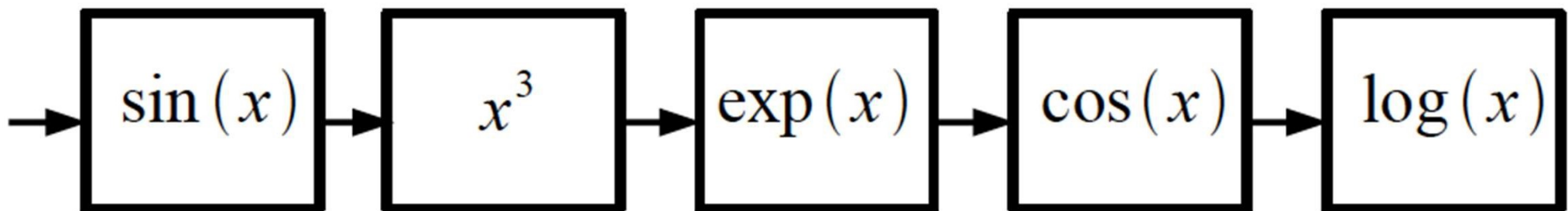


Compose into a
→
complicate function

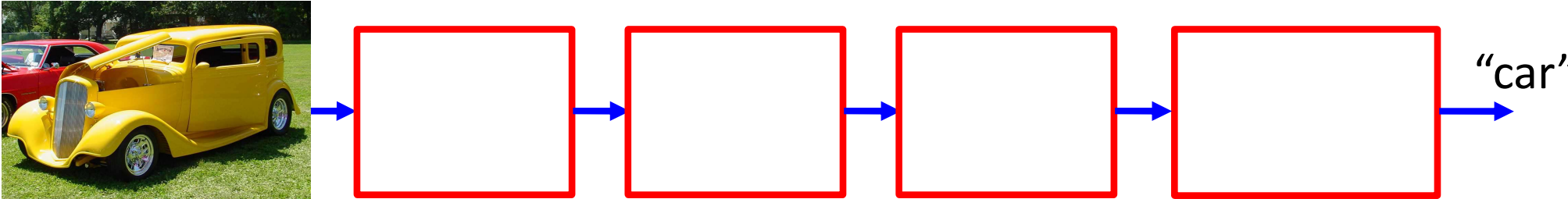
Idea 2: Compositions

- Deep Learning
- Grammar models
- Scattering transforms...

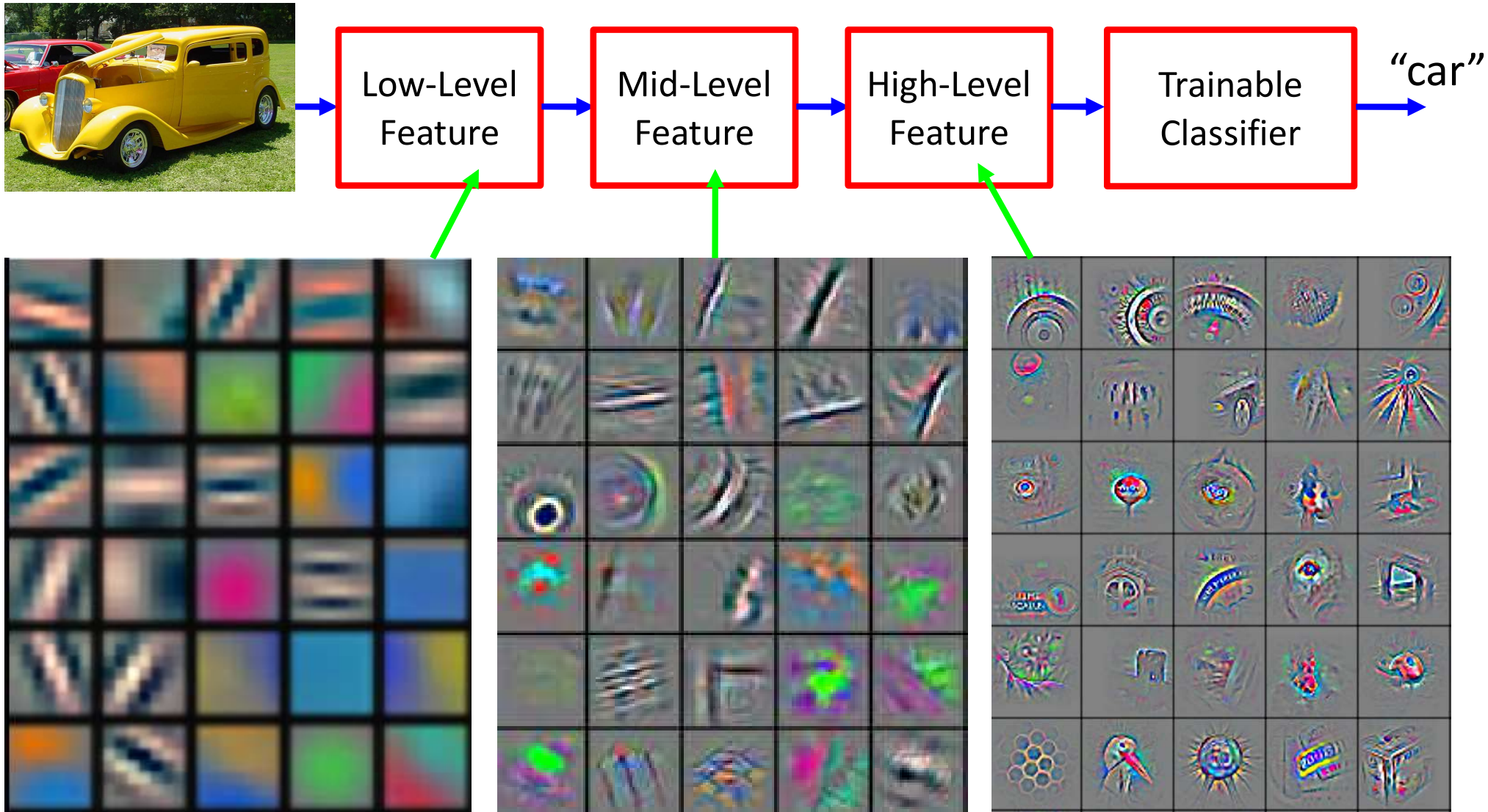
$$f(x) = \log(\cos(\exp(\sin^3(x))))$$



Deep Learning = Hierarchical Compositionality



Deep Learning = Hierarchical Compositionality



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

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
 - Distributed Representations
 - No single neuron “encodes” everything
 - Groups of neurons work together

Traditional Machine Learning

VISION



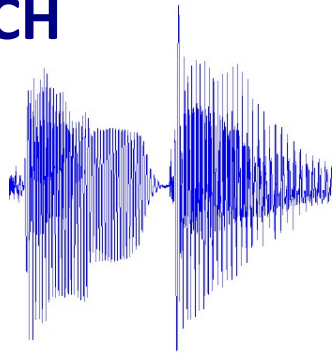
fixed



learned

“car”

SPEECH



fixed



learned

\ 'd ē p \

NLP

This burrito place
is yummy and fun!



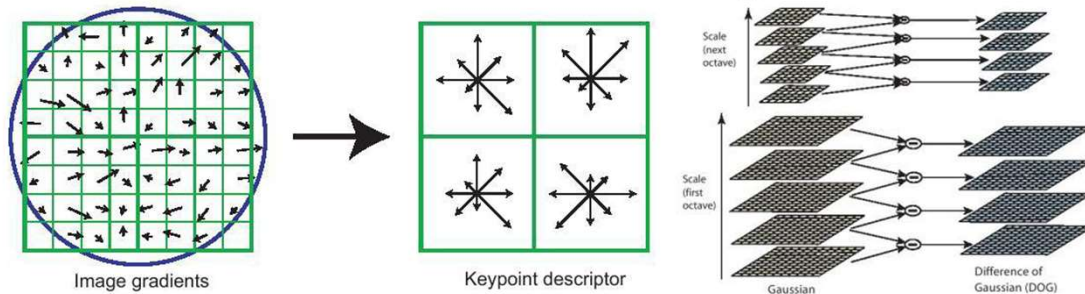
fixed



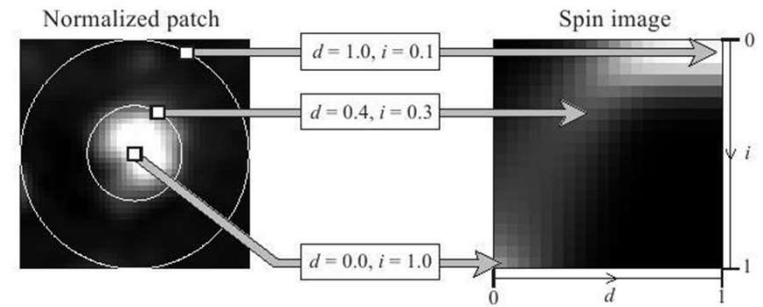
learned

“+”

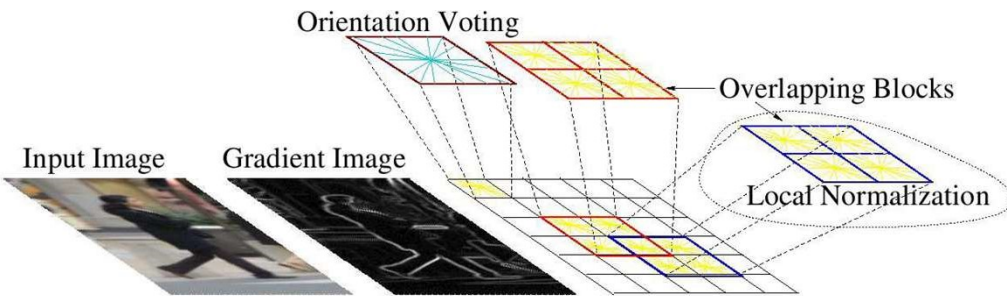
Feature Engineering



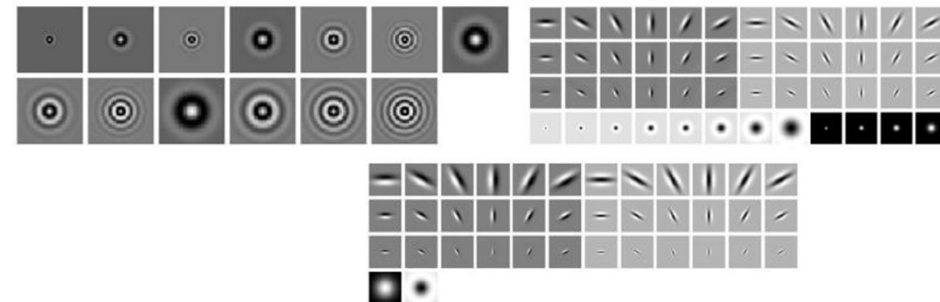
SIFT



Spin Images



HoG

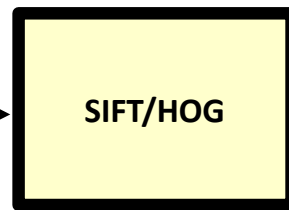


Textons

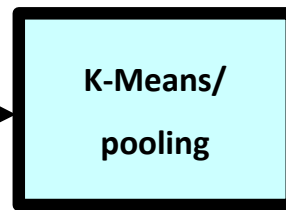
and many many more....

Traditional Machine Learning (more accurately)

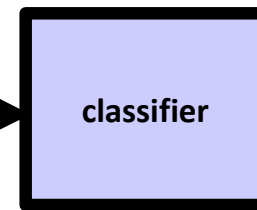
VISION



fixed



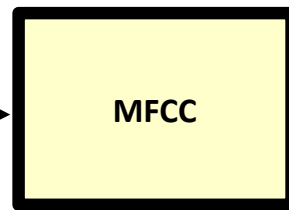
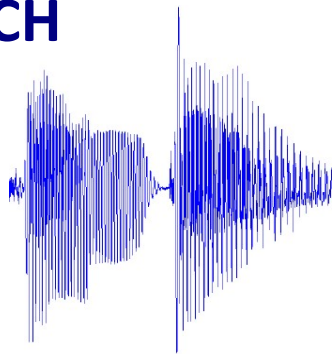
unsupervised



supervised

“car”

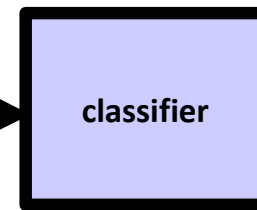
SPEECH



fixed



unsupervised

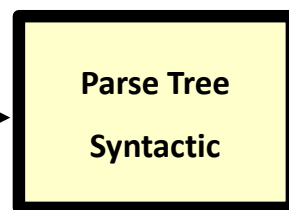


supervised

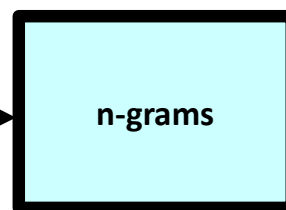
\ 'd ē p \

NLP

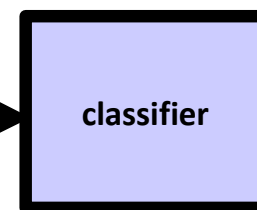
This burrito place
is yummy and fun!



fixed



unsupervised



supervised

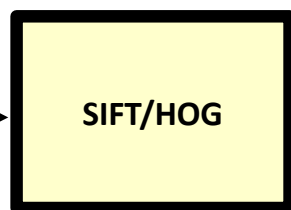
“+”

“Learned”

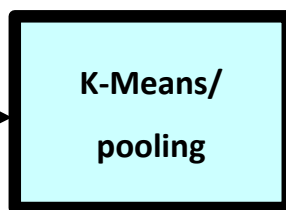


Deep Learning = End-to-End Learning

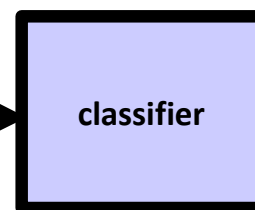
VISION



fixed



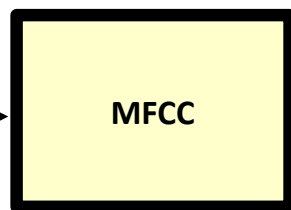
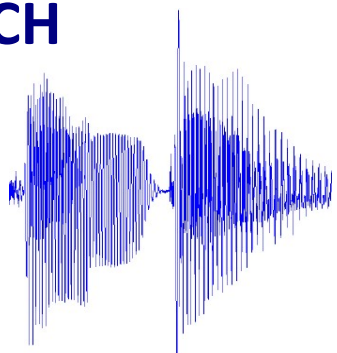
unsupervised



supervised

"car"

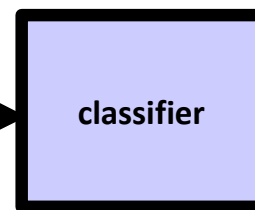
SPEECH



fixed



unsupervised

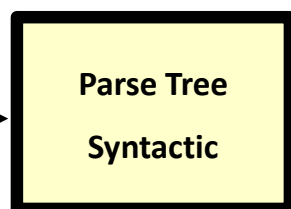


supervised

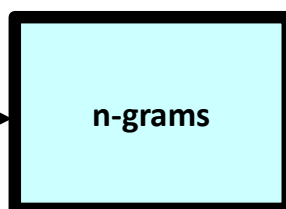
\ 'd ē p \

NLP

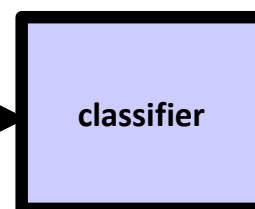
This burrito place
is yummy and fun!



fixed



unsupervised



supervised

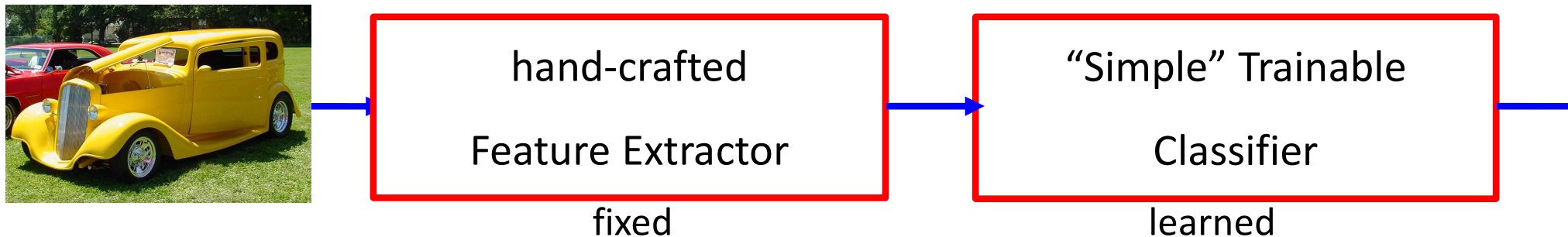
"+"

"Learned"

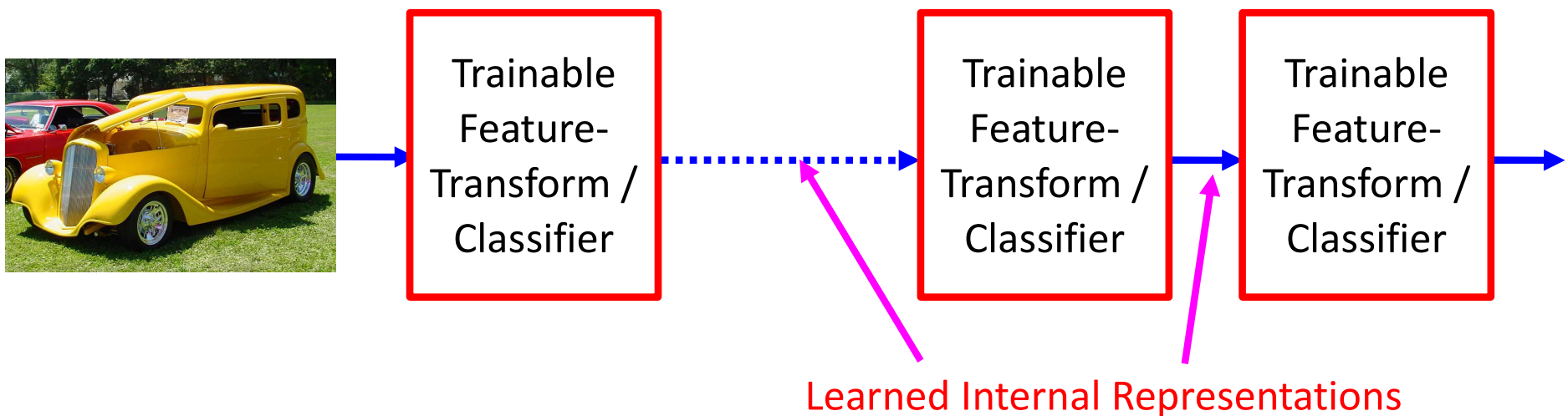


“Shallow” vs Deep Learning

- “Shallow” models



- Deep models



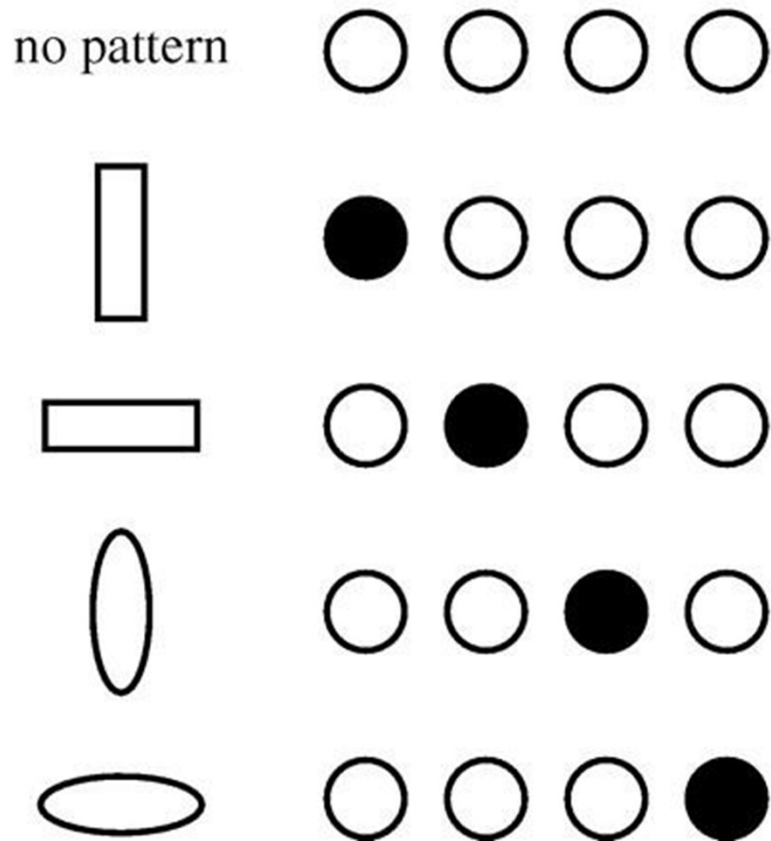
So what *is* Deep (Machine) Learning?

- 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
- Distributed Representations
 - No single neuron “encodes” everything
 - Groups of neurons work together

Distributed Representations Toy Example

- Local vs Distributed

(a)

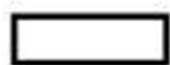


Distributed Representations Toy Example

- Can we interpret each dimension?

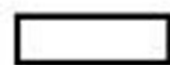
(a)

no pattern



(b)

no pattern



vertical
horizontal
rectangle
ellipse

Power of distributed representations!

Local

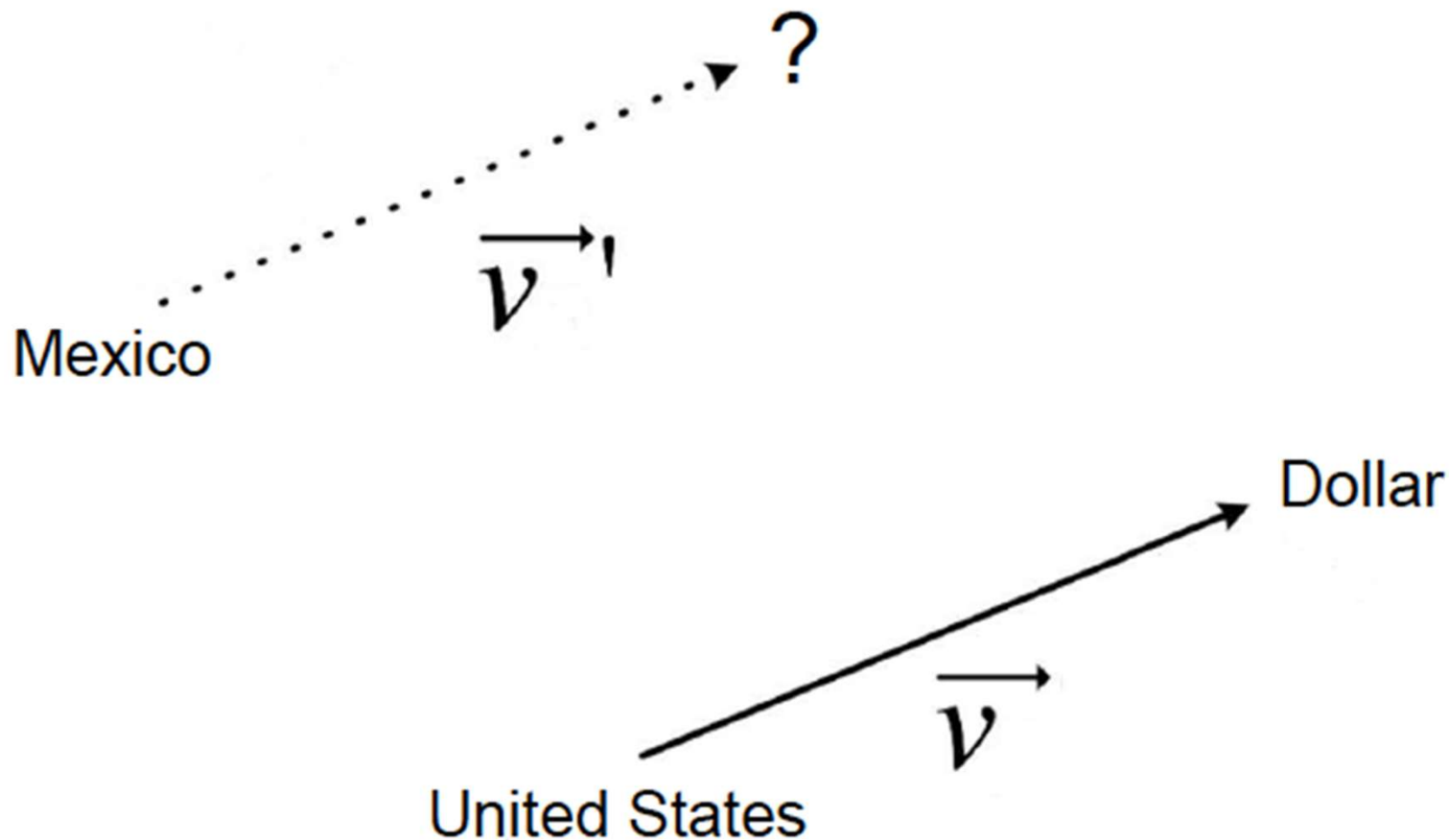
$$\bullet \bullet \circ \bullet = VR + HR + HE = ?$$

Distributed

$$\bullet \bullet \circ \bullet = V + H + E \approx \bigcirc$$

Power of distributed representations!

- United States:Dollar :: Mexico:?



ThisPlusThat.me

the matrix - thoughtful + dumb

Search

How it Works

mbiguated into +1 the_matrix -1 thoughtful +1 dumb in 0.0 seconds from ip-10-32-114-31

FILM, W FILM, NETFLIX TITLE,

Blade II

Blade II is a 2002 American vampire superhero action film base Marvel Comics character Blade. It is the sequel of the first film a part of the Blade film series. It was written by David S. Goyer, w previous film. Guillermo del Toro was signed in to d...

Horror Film



Image Credit:

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
- Distributed Representations
 - No single neuron “encodes” everything
 - Groups of neurons work together

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

“Expert” intuitions can be misleading

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

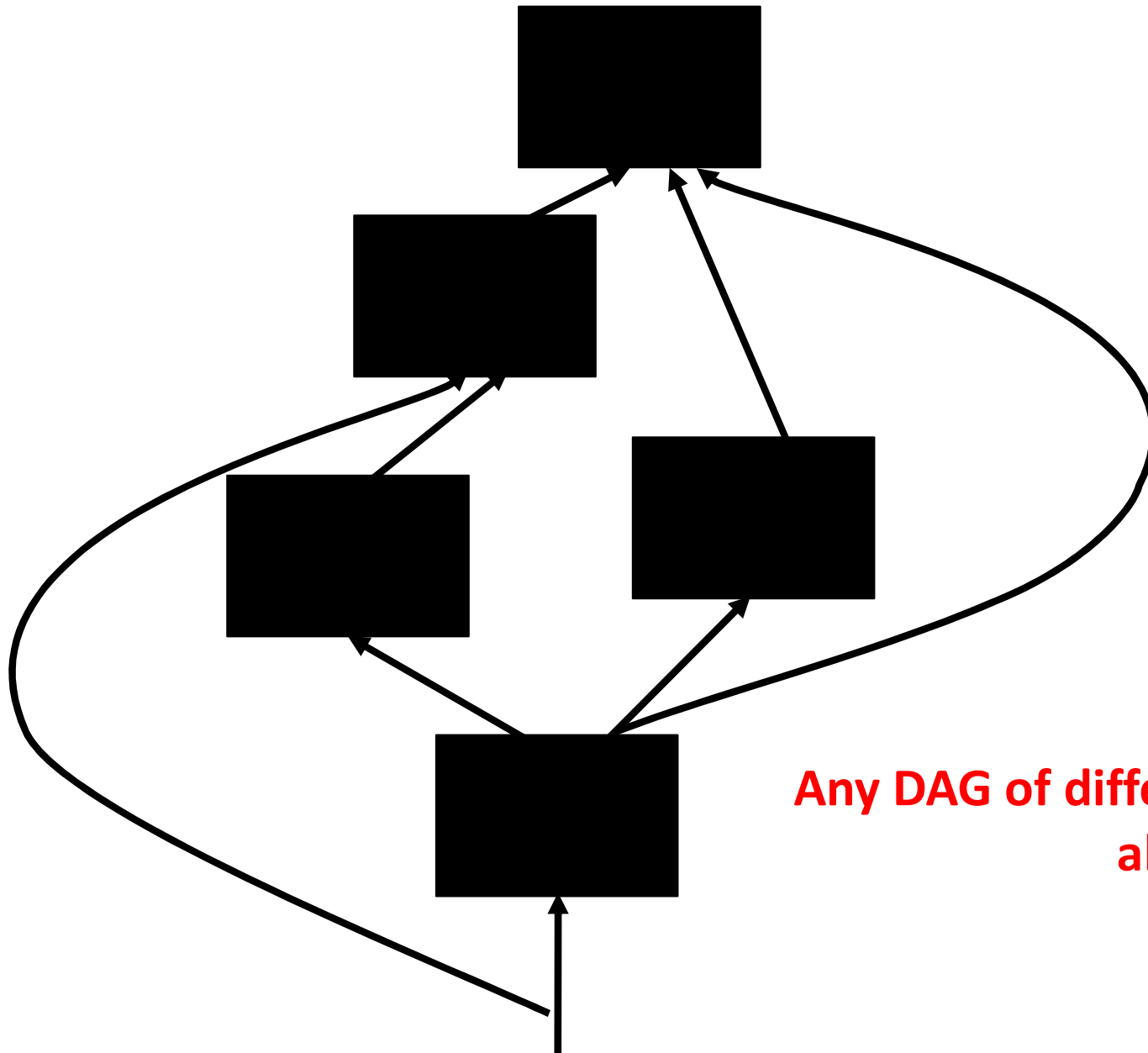


- *“Because gradient descent is better than you”*
 - Yann LeCun, CVPR '13

Benefits of Deep/Representation Learning

- Modularity!
- Plug and play architectures!

Differentiable Computation Graph



Any DAG of differentiable modules is allowed!

Problems with Deep Learning

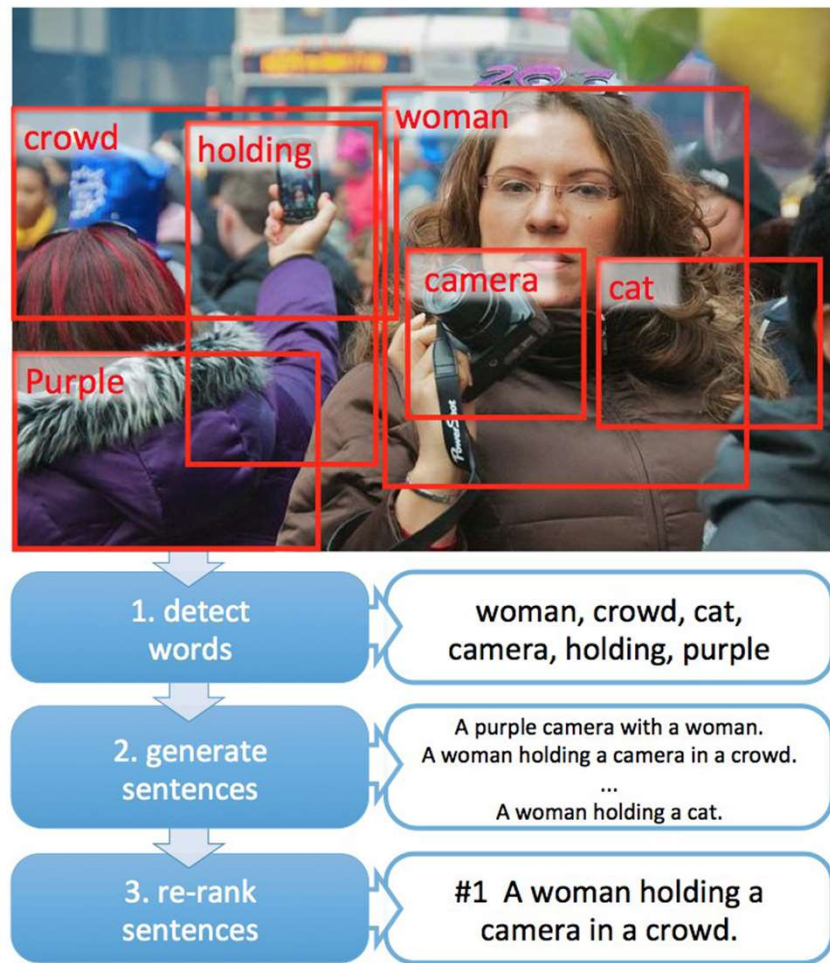
- **Problem#1: Lack of a formal understanding**
 - Non-Convex! Non-Convex! Non-Convex!
 - Depth \geq 3: most losses non-convex in parameters
 - Worse still, existing intuitions from classical statistical learning theory don't seem to carry over.
 - Theoretically, we are stumbling in the dark here
- **Standard 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 \rightarrow wave hands \rightarrow non-convexity
- **Standard response #2**
 - “Yes, but it often works!”

Problems with Deep Learning

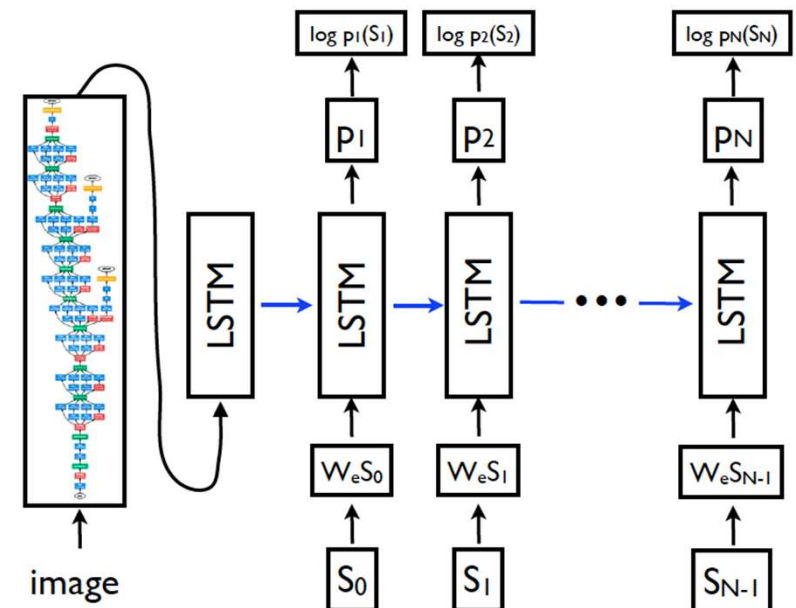
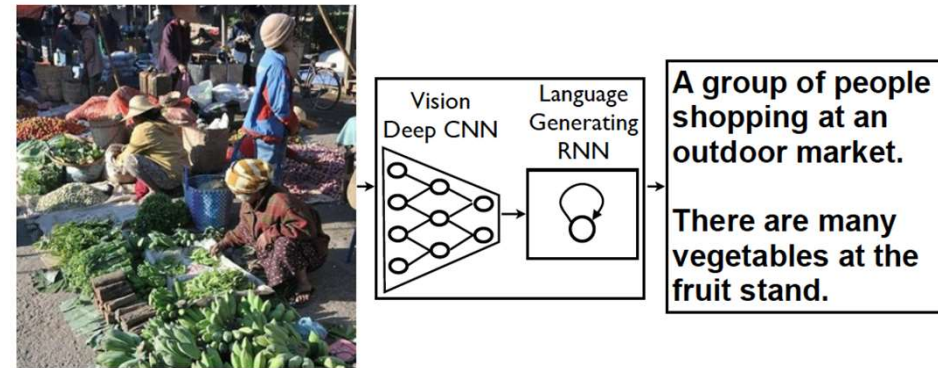
- **Problem#2: Lack of interpretability**
 - Hard to track down what's failing
 - Pipeline systems have expected performances at each step
 - In end-to-end systems, it's hard to know why things are not working

Problems with Deep Learning

- Problem#2: Lack of interpretability



[Fang et al. CVPR15]



[Vinyals et al. CVPR15]

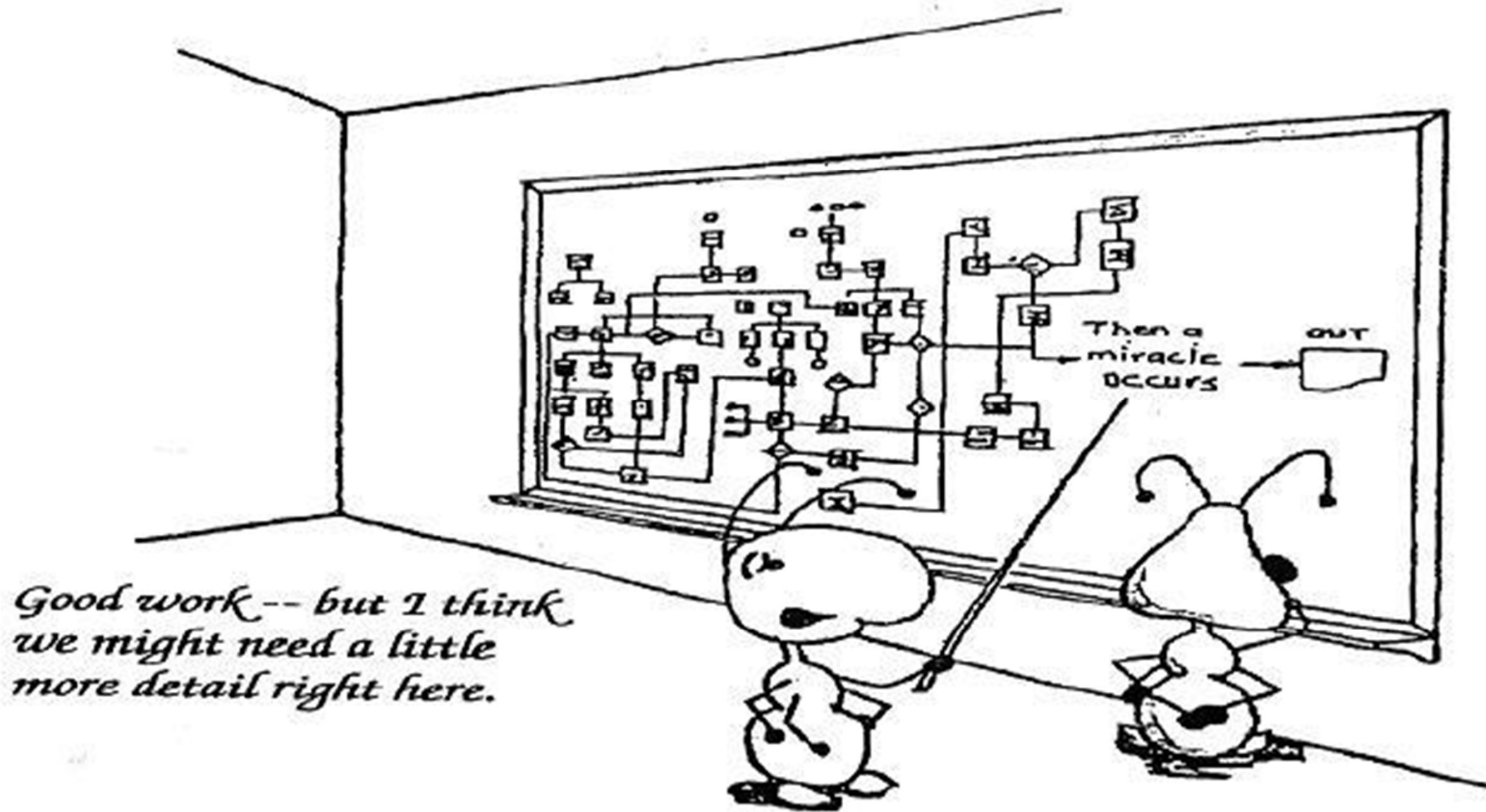
Problems with Deep Learning

- **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
- 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!”

Yes it works, but how?



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
 - RNNs / Transformers
 - Deep Reinforcement Learning
 - Generative Models (VAEs, GANs)
- Target Audience:
 - Senior undergrads, MS-ML, and new PhD students

What this class is NOT

- NOT the target audience:
 - Advanced grad-students already working in ML/DL areas
 - People looking to understand latest and greatest cutting-edge research (e.g. GANs, AlphaGo, etc)
 - Undergraduate/Masters students looking to graduate with a DL class on their resume.
- NOT the goal:
 - Teaching a toolkit. “Intro to TensorFlow/PyTorch”
 - Intro to Machine Learning

Caveat

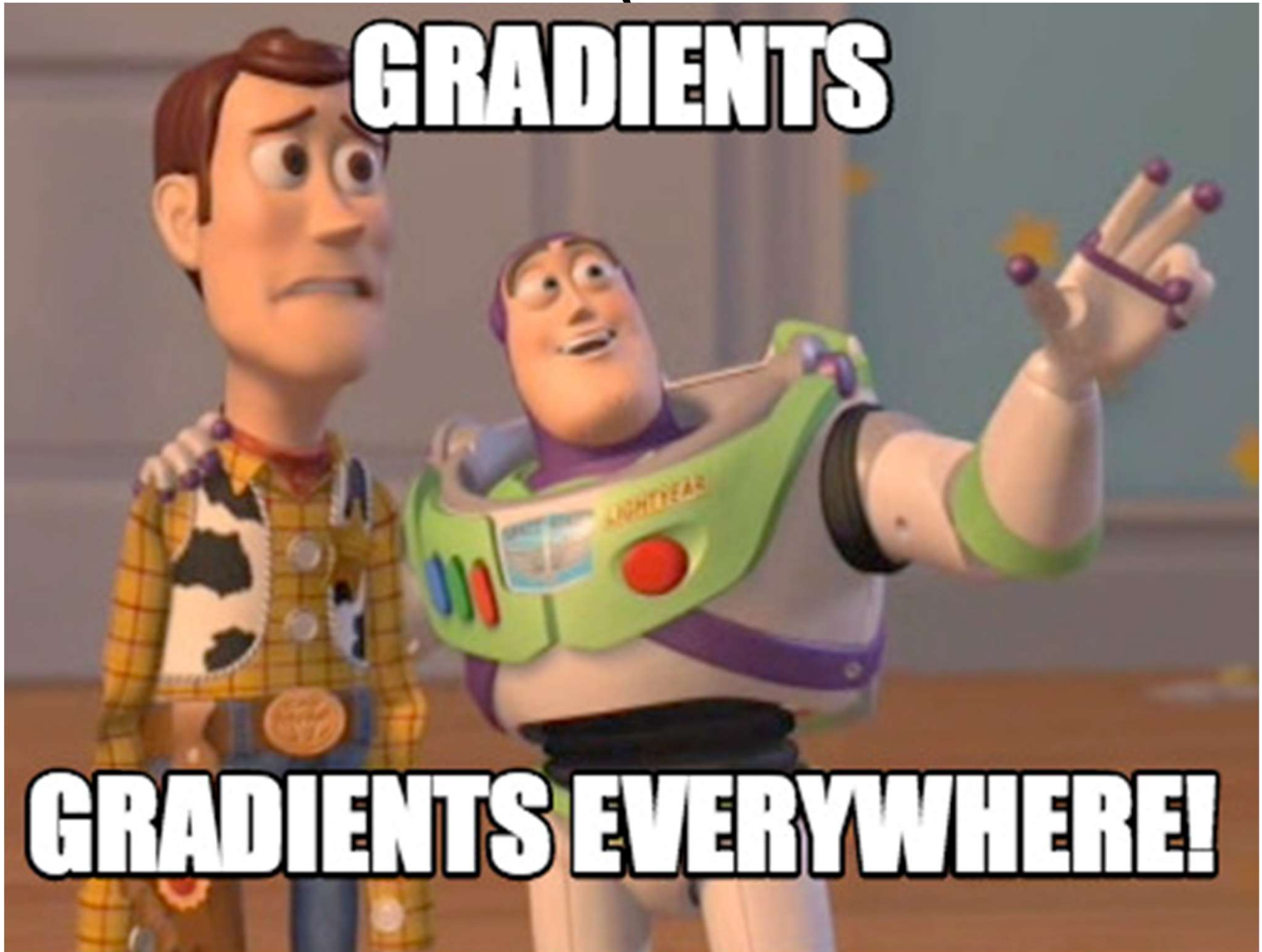
- This is an **ADVANCED** Machine Learning class
 - This should **NOT** be your first introduction to ML
 - You will need a formal class; not just self-reading/coursera
 - If you took CS 7641/ISYE 6740/CSE 6740 @GT, you're in the right place
 - If you took an equivalent class elsewhere, see list of topics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

Prerequisites

- 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>

Prerequisites

GRADIENTS



Prerequisites

- Intro Machine Learning
 - Classifiers, regressors, loss functions, MLE, MAP
- Linear Algebra
 - Matrix multiplication, eigenvalues, positive semi-definiteness...
- Calculus
 - Multi-variate gradients, hessians, jacobians...
- **Programming!**
 - Homeworks will require Python!
 - Libraries/Frameworks: PyTorch
 - HW1 (pure python), HW2 (python + PyTorch), HW3+4 (PyTorch)
 - Your language of choice for project

Course Information

- Instructor: Zsolt Kira
 - zkira@gatech dot edu (**use piazza public/private instead!**)

TAs



Man (Mandy) Xie



Jordan Rodrigues



Bhavika Devnani



Anshul Ahluwalia



Aditya Singh



Ahmed Shaikh



Amogh Dabholkar



Ting-Yu Lan



Yanzhe Zhang

Organization & Deliverables

- 4 problem-sets+homeworks (80%)
 - 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 (20%)
 - Projects done in groups of 3-4
- (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
 - 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 |
lgbtqia.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 Thursday Jan 13th
 - Will be 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
 - $\leq 75\%$ means that you might not be prepared for the class
 - If you submit after Thursday, we will not grade before registration ends
- Topics
 - PS: probability, calculus, convexity, proving things

Computing

- Major bottleneck
 - GPUs
- Options
 - Your own / group / advisor's resources
 - Google Colab
 - jupyter-notebook + free GPU instance
 - Google Cloud credits (details TBD)
 - New! Amazon AWS credits courtesy of Facebook/Meta AI Alliance Program

4803 vs 7643

- Level differentiation
- HWs
 - Extra credit questions for 4803 students, necessary for 7643
- Project
 - Higher expectations from 7643

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**

Waitlist / Audit / Sit in

- Waitlist
 - Class is full. Size will not increase further.
 - Do PS0. 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

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

How do I get in touch?

- Primary means of communication -- Piazza
 - No direct emails to Instructor unless private information
 - 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](#), but too late for Spring 2022
 - If you can find one of my students to supervise you, I am happy to sign off on the paperwork.
 - Your responsibility to approach them and ask. It will help if you know what they are working on.

Todo

- PS0
 - Due: Jan 13th 11:59pm

Welcome

