Deep Learning for Data Science
DS 542

https://dl4ds.github.io/sp2026/

Introduction and Course Overview

Course Staff

Instructor:

Thomas Gardos
Assoc. Prof of the Practice
Previously Manager and AI Algorithms Engineer, Intel Corp.

Teaching Assistants:

- Johnathan Finizio, BSDS
- Yuki Li, MSDS

Course Assistants:

- Priyanshu Bansal, MSCS
- Tarushi Gandhi, MSAI
- Carol Kong, BSDS

See https://piazza.com/bu/spring2026/ds542/staff for office hour information.

Plan for Today

- Applications of Deep Learning
- Why Deep Learning?
- How We Figured out Deep Learning
- Course Logistics

Plan for Today

- Applications of Deep Learning
- Why Deep Learning?
- How We Figured out Deep Learning
- Course Logistics

ImageNet Challenge 2012 – A Turning Point

Task 1: Classification

Car

- Predict a class label
- 5 predictions / image
- 1000 classes
- 1,200 images per class for training
- Bounding boxes for 50% of training.

Task 2: Detection (Classification + Localization)

classification Car

- Predict a class label and a bounding box
- 5 predictions / image
- 1000 classes
- 1,200 images per class for training
- Bounding boxes for 40% of training.

Task 3: Fine-grained classification

classification Walker hound

- Predict a class label given a bounding box in test
- 1 prediction / image
- 120 dog classes (subset)
- ~200 images per class for training (subset)
- Bounding boxes for 100% of training

Source: https://www.image-net.org/static_files/files/ilsvrc2012.pdf

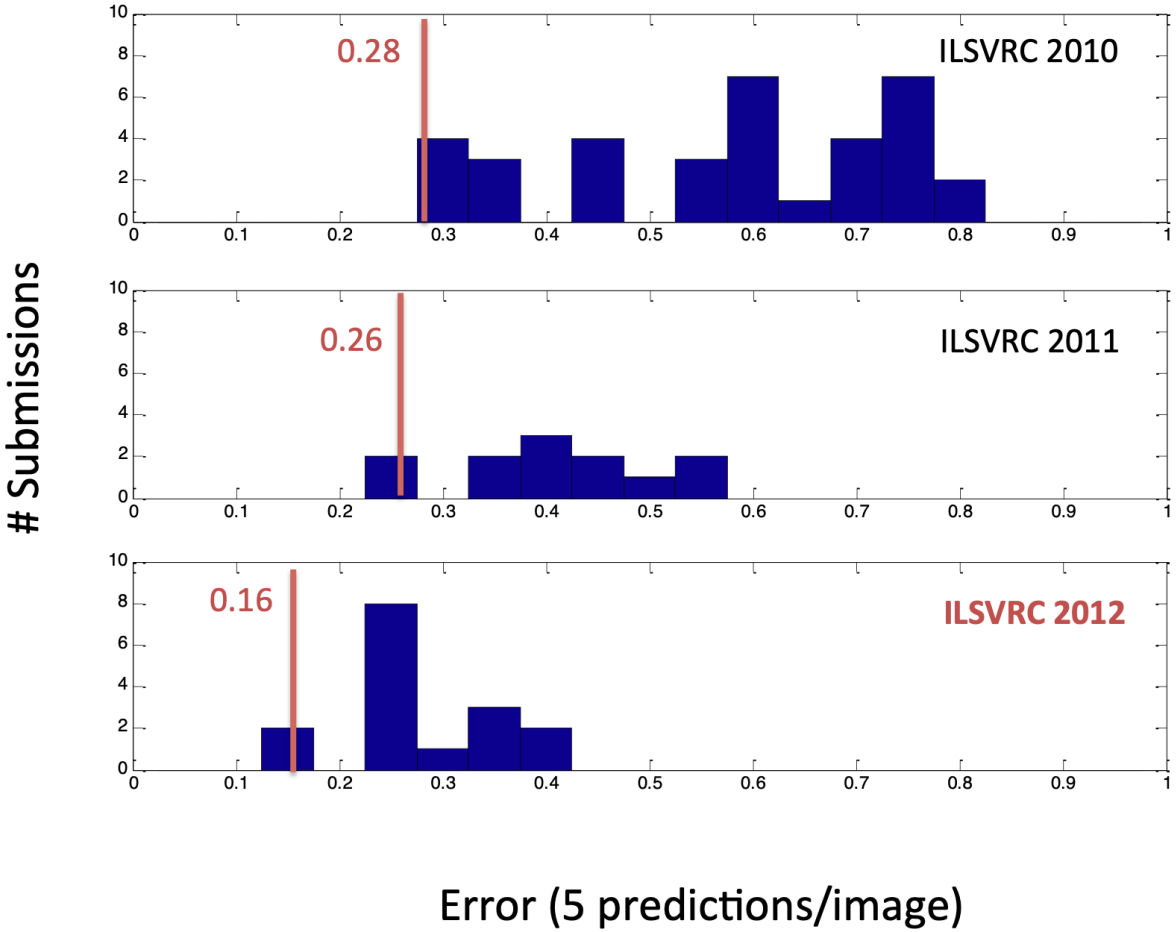
ImageNet Categories

Breakdown of categories (samples):

Subtree	# of leaf categories
Instrument	358
Canine	130
Covering	90
Vehicle	67
Invertebrate	61
Bird	59
Structure (construction)	58
Food	27
…	

Source: https://www.image-net.org/static_files/files/ilsvrc2012.pdf

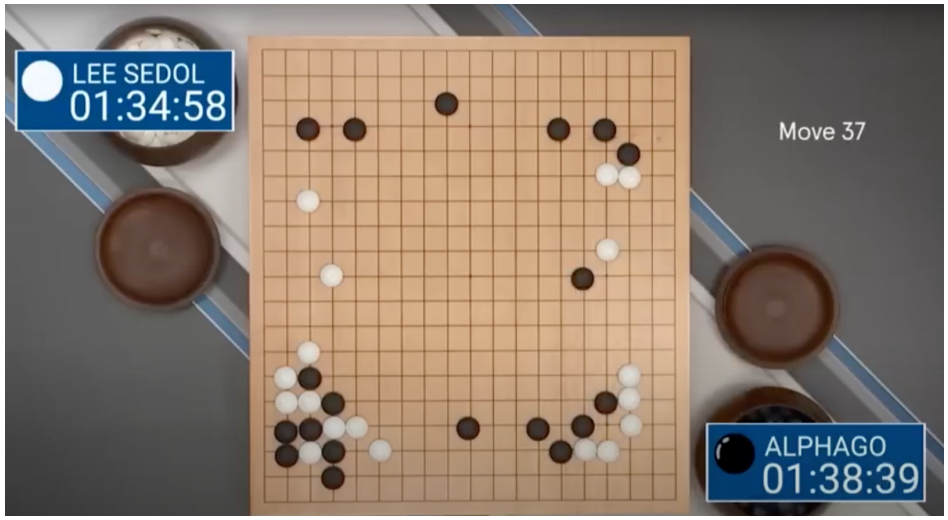
2012 Results (Classification)



Source: https://www.image-net.org/static_files/files/ilsvrc2012.pdf

AlphaGo (2016)

Google's Computer Program Beats Lee Se-dol in Go Tournament



Move 37 broke a centuries old rule of Go, but ended up being a creative, strategic tactic.

Sources:

https://www.nytimes.com/2016/03/10/world/asia/google-alphago-lee-se-dol.html

Emotion Identification (2019)

- 42 muscles control all possible expressions
- Restrictions on how faces and heads look subject to physics of illumination and reflectance, etc.

- The “manifold” of possible faces is much, much smaller than the combinatoric collection of pixel values

C. A. C. Holland, N. C. Ebner, T. Lin, and G. R. Samanez-Larkin, “Emotion identification across adulthood using the Dynamic FACES database of emotional expressions in younger, middle aged, and older adults,” *Cognition and Emotion*, vol. 33, no. 2, pp. 245–257, Feb. 2019, doi: 10.1080/02699931.2018.1445981.

GPT-2 (2019)

System Prompt (human-written)

In a shocking finding, scientist discovered a herd of unicorns living in a remote, previously unexplored valley, in the Andes Mountains. Even more surprising to the researchers was the fact that the unicorns spoke perfect English.

Model Completion (machine-written, 10 tries)

The scientist named the population, after their distinctive horn, Ovid's Unicorn. These four-horned, silver-white unicorns were previously unknown to science.

Now, after almost two centuries, the mystery of what sparked this odd phenomenon is finally solved.

https://openai.com/index/better-language-models/

This Person Does Not Exist (2020)

https://thispersondoesnotexist.com, using StyleGan2

Latent Diffusion (2021)

"Hallucinating" image detail to create high resolution version

R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, "High-Resolution Image Synthesis with Latent Diffusion Models," Apr. 13, 2022, *arXiv*: arXiv:2112.10752. doi: 10.48550/arXiv.2112.10752.

Image Interpolation (2022)

Axel Sauer, Katja Schwarz, and Andreas Geiger. 2022. StyleGAN-XL: Scaling StyleGAN to Large Diverse Datasets. In ACM SIGGRAPH 2022 Conference Proceedings (SIGGRAPH '22). Association for Computing Machinery, New York, NY, USA, Article 49, 1–10. https://doi.org/10.1145/3528233.3530738

Ramesh, A., Dhariwal, P., Nichol, A., Chu, C., & Chen, M. (2022). Hierarchical text-conditional image generation with CLIP Latents. arXiv:2204.06125

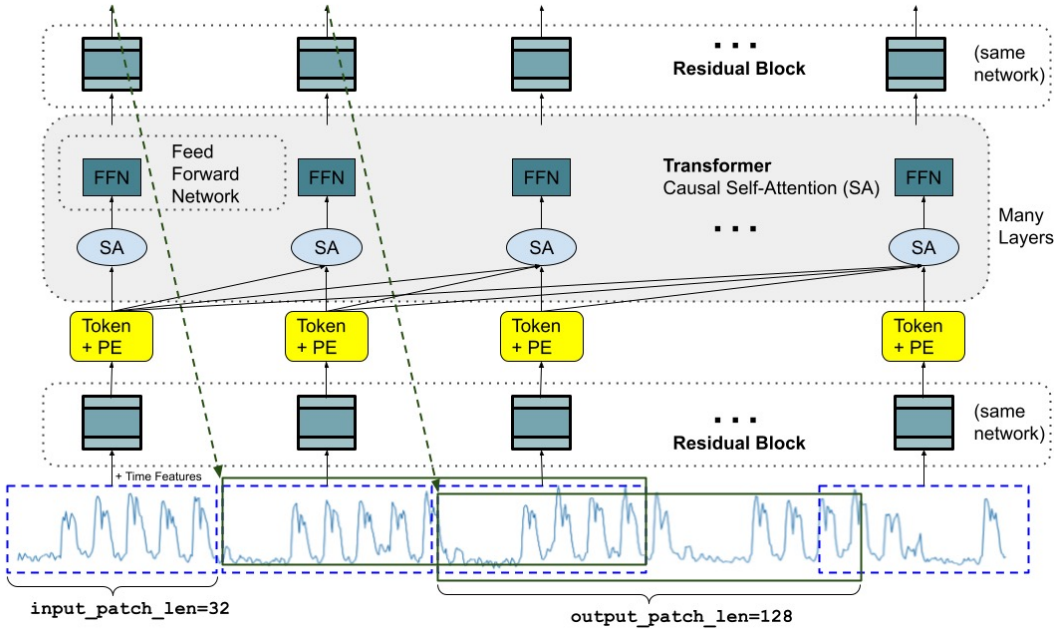
Conditional synthesis (2022)

Saharia, C., Chan, W., Chang, H., Lee, C., Ho, J., Salimans, T., Fleet, D., & Norouzi, M. (2022a). Palette: Image-to-image diffusion models. ACM SIGGRAPH, (link)

Segment Anything (2023)

Source: https://segment-anything.com

Time Series Forecasting (2024)

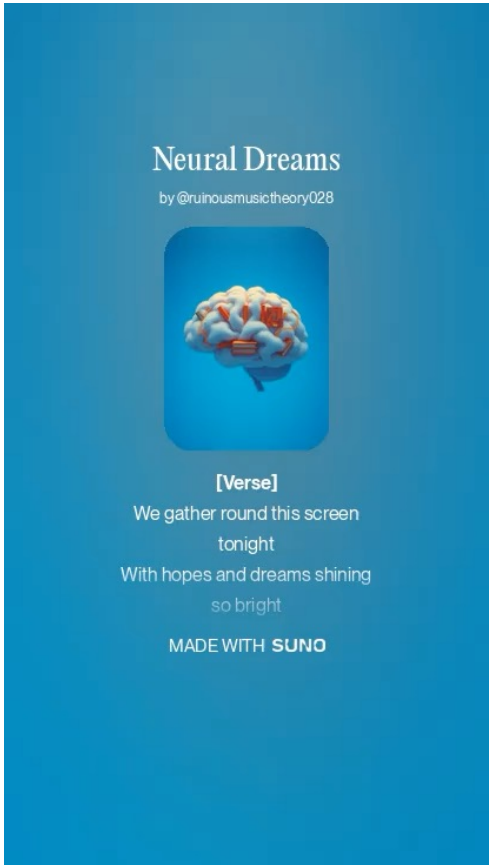


https://research.google/blog/a-decoder-only-foundation-model-for-time-series-forecasting/

Image/Video/Music Generation (2024)

A teenage superhero fighting crime in an urban setting shown in the style of claymation.

https://sora.com



Write a short pop song about students wanting to learn about neural networks and do great things with them.

What do these examples have in common?

- Very complex relationship between input and output
- Sometimes may be many possible valid answers
- But outputs (and sometimes inputs) obey rules

"A Kazakh man on a horse holding a bird of prey"

Language obeys grammatical rules

Natural images also have "rules"

Any Questions?

Plan for Today

- Applications of Deep Learning
- Why Deep Learning?
- How We Figured out Deep Learning
- Course Logistics

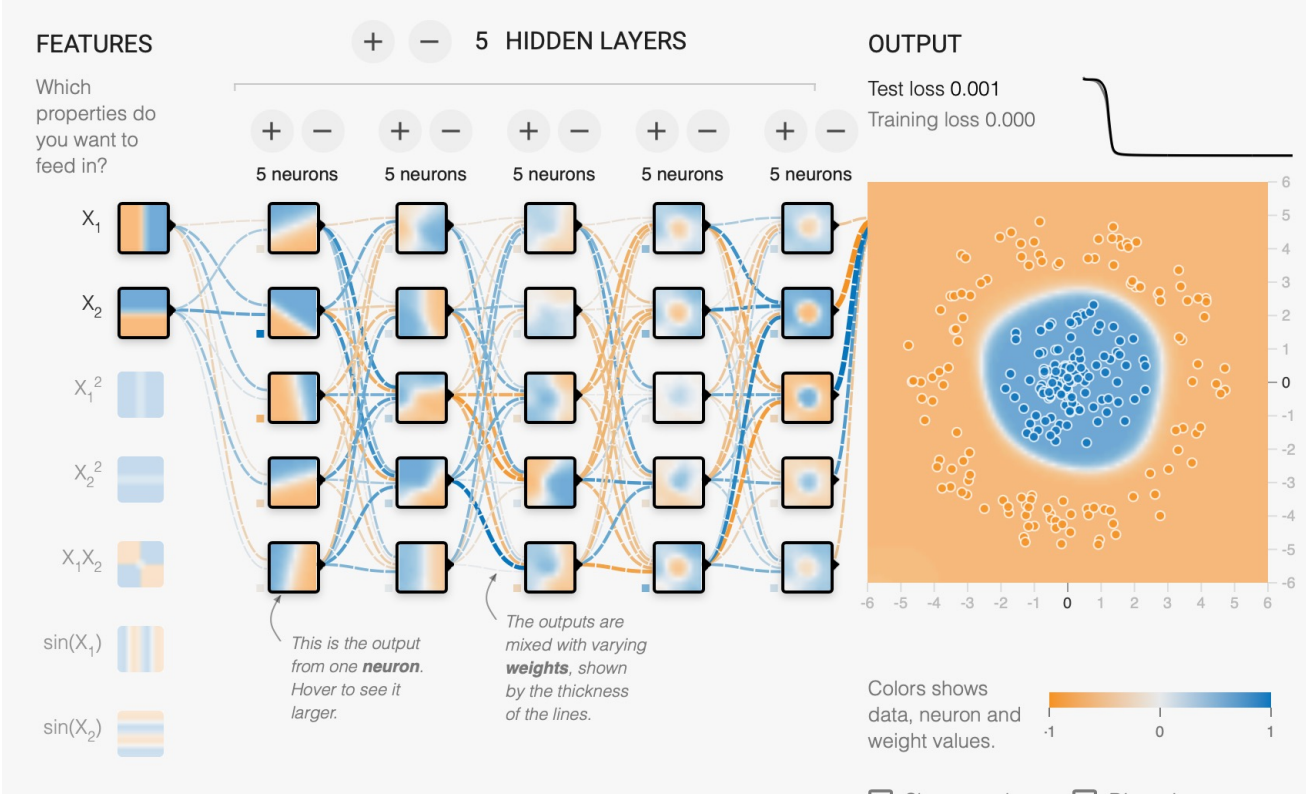
Universal Function Approximation

- A big enough neural network can approximate any function.
  - Does not require deep learning.

- To fit or overfit?

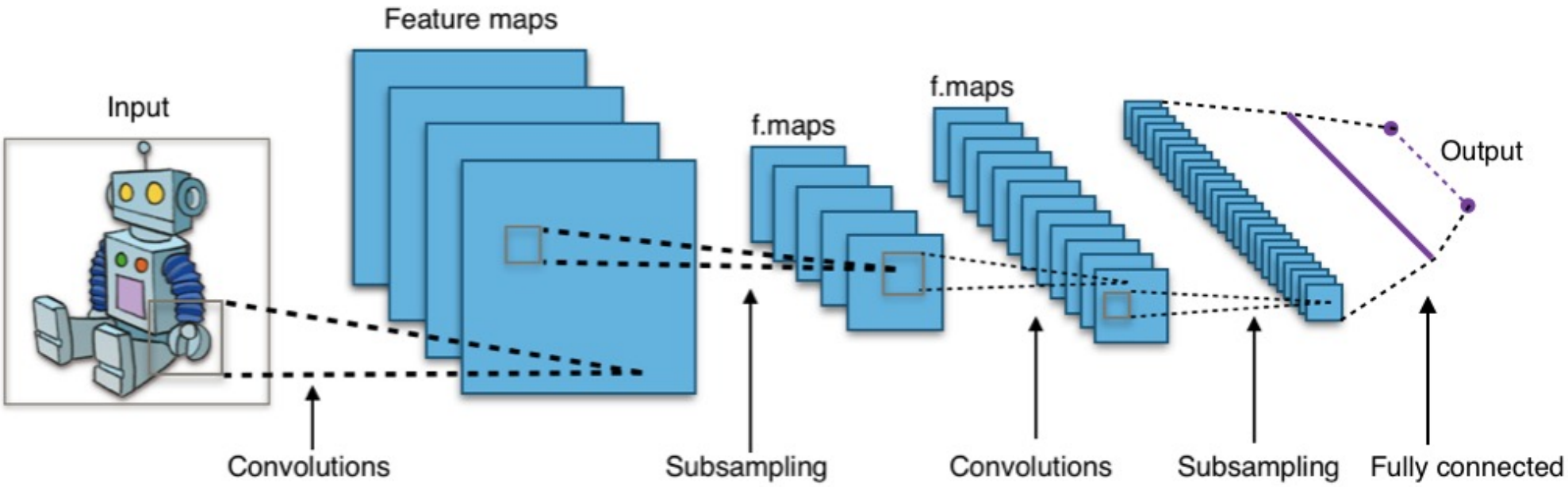
Layers of Abstraction

- Many layers may map to progressively developing concepts or abstractions.



http://playground.tensorflow.org/

Parameter Sharing In Convolutional Networks



Source: https://en.wikipedia.org/wiki/Convolutional_neural_network

Joint Optimization

Do all of these at once –

- Universal function approximation
- Layers of abstraction
- Parameter sharing

Differentiable Computations

- Can the partial derivative of “output quality” be computed with respect to every input and every parameter in the system?
  - Parameter sharing
  - Composability
  - Joint Optimization

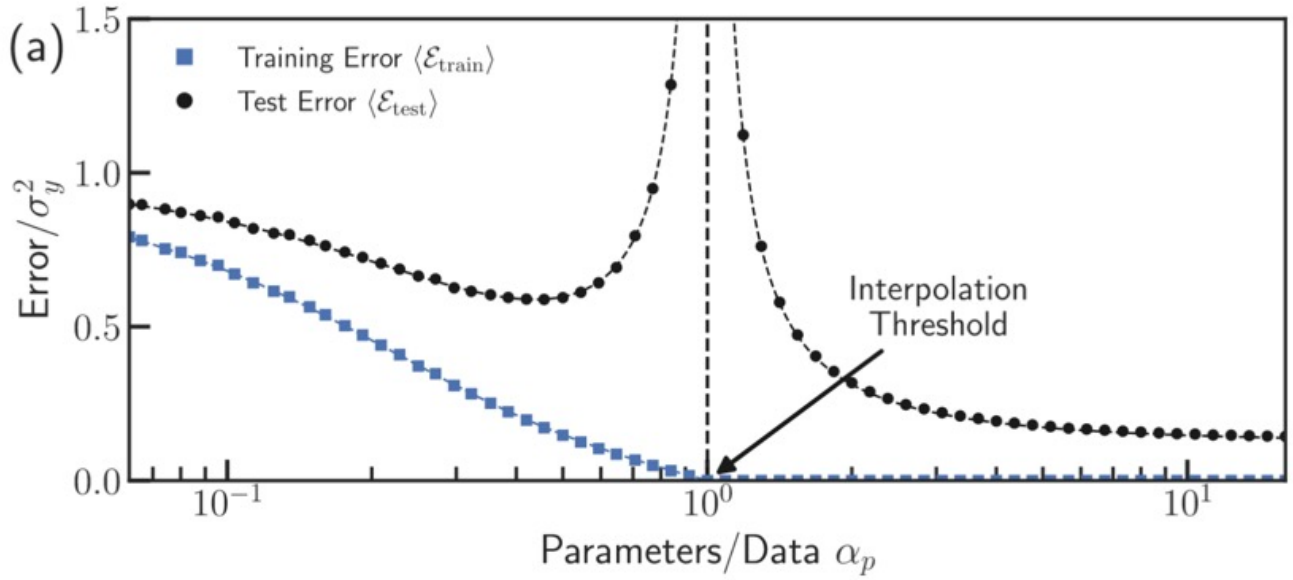
- Gradient descent as a universal algorithm!

- Can be modularized with back propagation

Deep Learning with Neural Networks

- Universal function approximation
  - Deep neural networks tend to be able to fit more complex patterns with fewer parameters.
  - Some toy problems have probable exponential gaps between shallow and deep network sizes.
- Layers of abstraction
  - Happens automatically but many parts not well understood.
- Train with gradient descent!

Better Generalization?



Source: [Double Descent Demystified: Identifying, Interpreting & Ablating the Sources of a Deep Learning Puzzle]

Any Questions?

Early History and the AI Winter

Abbreviated History of NNs

- 1943: McCulloch & Pitts – Calculus of neurons
- 1947-49: Donald Hebb – Plasticity of neurons
- 1956: Minsky, McCarthy, Shannon… Dartmouth Summer Research Project on AI
- 1957: Rosenblatt – Perceptron, HW implementation of 20x20 CV
- 1959: Hubel & Weisel – Visual cortex and receptive fields
- 1960: Widrow & Hoff – Adaptive Linear Neuron (ADALINE)
- 1969: Minsky & Papert – Perceptrons: computation limitations of neurons

Continued (abbreviated) History

- 1979 – Fukushima: Neocognitron, cascade of neural structures that can classify shapes, invariant to shift, learned from data
- 1982 – Hopfield Networks, recurrent artificial neural networks
- 1983 – Hinton & Sejnowski: Boltzmann Machines
- 1985 – Rumelhart, Hinton, Williams: Practical backpropagation
- 1989 – LeCun – Backprop on Convolutional Neural Networks
- 1991 – Bottou & Galinari – Automatic differentiation (autograd)
- 2012 – AlexNet (CNN on GPU trained on ImageNet)
- 2016 – Kaiming He: ResNet

A Brief History of Transformers

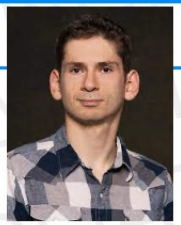
2003

Yoshua Bengio*

2014

Ilya Sutskever*

2014



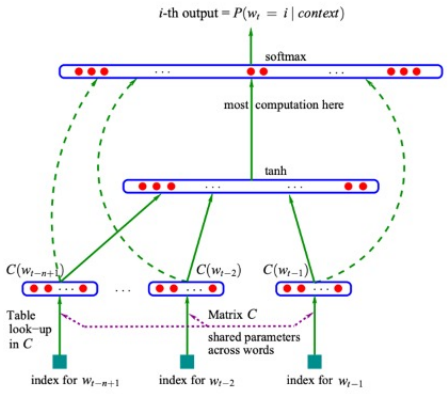
Dzmitry Bahdanau*

2017

A Team at Google

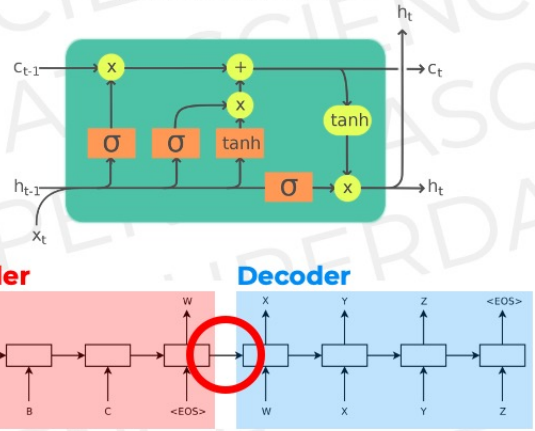
Use LSTMs

A Neural Probabilistic Language Model



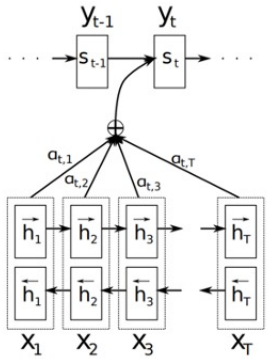
Add Attention

Seq-to-Seq Learning with Neural Networks

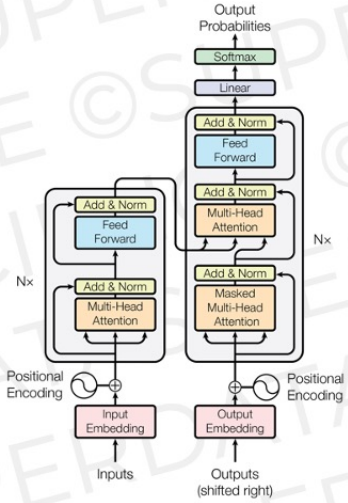


Remove LSTMs

Neural Machine Translation by Jointly Learning to Align and Translate



Attention is all you need



**And others; Chronological analysis inspired by Andrej Karpathy's lecture, youtube.com/watch?v=XfpMkf4rD6E*

Computer Vision as Inverse Computer Graphics (not a tangent)

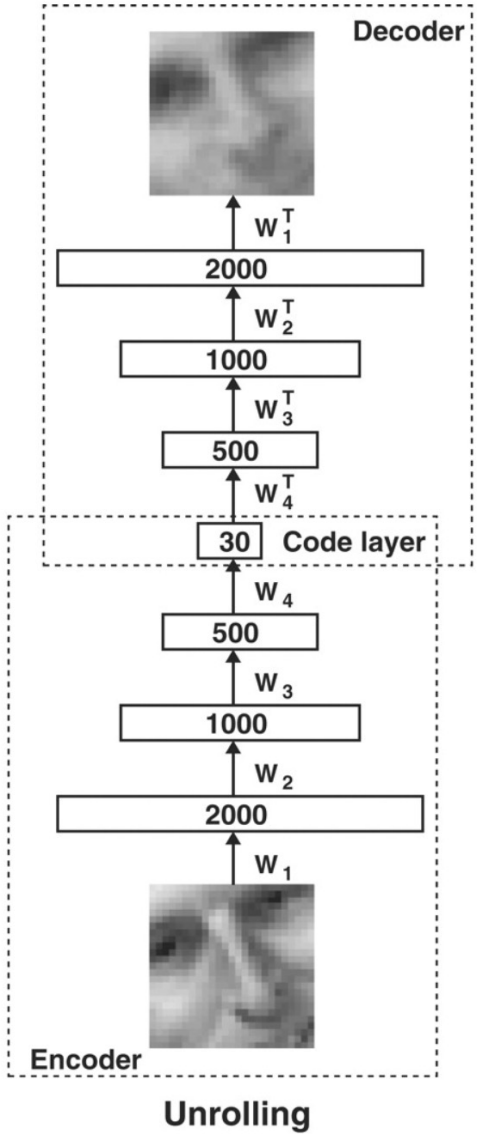
Paraphrase of Geoffrey Hinton -

If we can briefly write a short description of an object well enough to draw it, then we must have captured the essence of the object?

Auto-Encoder Idea

- If we can build an architecture like this, then the small "code layer" must have most of the important information?

Source: Reducing the Dimensionality of Data with Neural Networks (2006)



Unrolling

Deep Learning from Scratch

- After the previous work, we started figuring out what we were doing wrong and learned how to train deep neural networks directly.

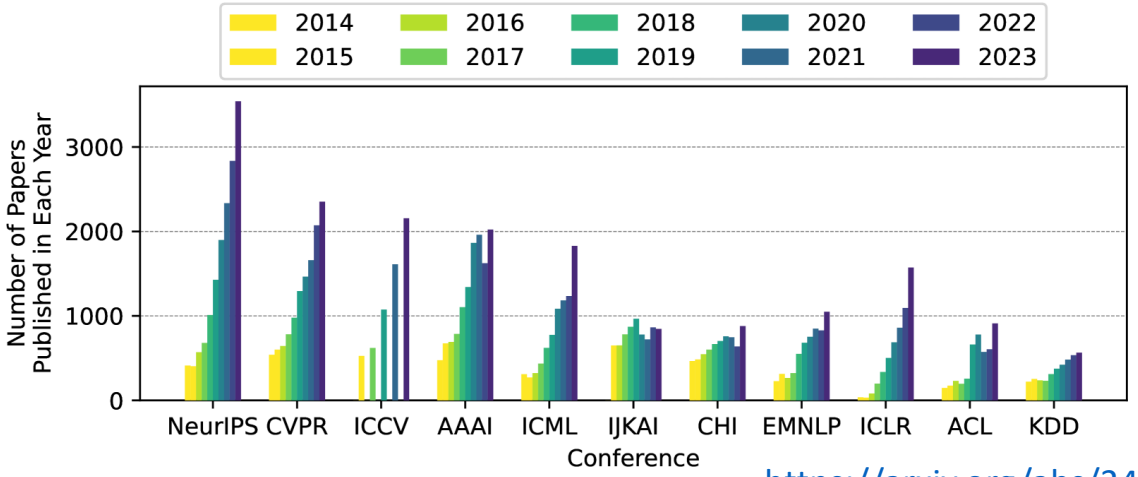
- This will be the first 1/3rd of this course…

Any Questions

Course Information

Explosion in Deep Learning Publications

- There has been exponential growth in deep learning research and publications
- This is not a survey course or graduate seminar on every latest trend and breakthrough



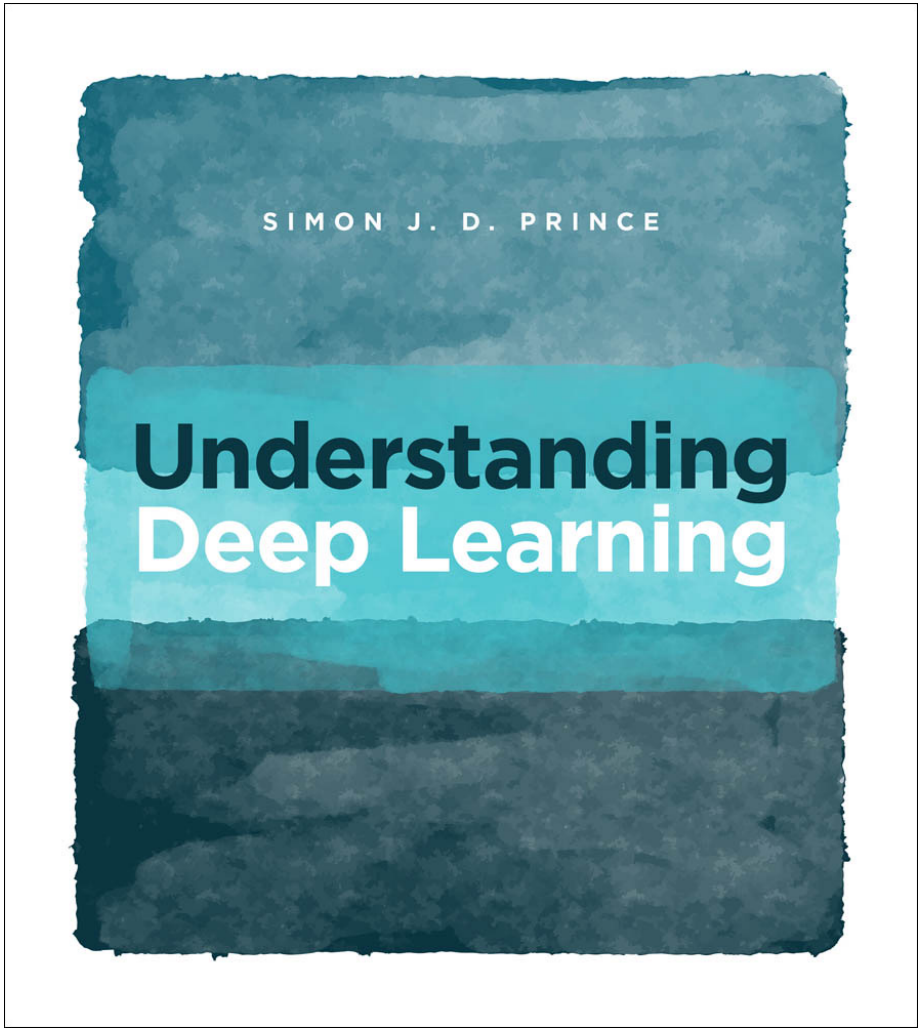
https://arxiv.org/abs/2412.07793v1

Focus on the Foundations

- We will focus on the theory and practice behind many of the major efforts in deep learning
- With this foundation you should be better equipped to read, implement and contribute to more recent efforts

Book

- Published December 2023
- http://udlbook.com
  - Free PDF there or buy at BU bookstore
- Used heavily for 1st half of the course, and a bit at the end too



Course Logistics

- Tuesdays and Thursdays, 3:30 – 4:45
  - I'll have to leave fairly promptly at 4:45 to catch my train
- Office hours: Tuesdays and Thursdays 2-3pm and by appointment other times. (I may expand office hours as I figure out my schedule)
- Discussions on Fridays
- TA/CA office hours will be posted on Piazza

Course Web Site

- https://dl4ds.github.io/sp2026/
    - Syllabus
    - Course Schedule
    - Lectures, discussions and assignments
    - Piazza and Gradescope links sent via email.

Generative AI Assistance (GAIA) Policy

https://dl4ds.github.io/fa2025/index.html#gaia-policy

1. Give credit to AI tools whenever used, even if only to generate ideas rather than usable text, illustrations or code.

…

3. When using AI tools on _coding_ assignments, unless prohibited
    1. Add the prompt text and tool used as comments before the generated code. Clarify whether the code was used as is, or modified somewhat, moderately or significantly.

…

5. Use AI tools wisely and intelligently, aiming to deepen understanding of subject matter and to support learning.

Focus on your learning objectives!

Please, no AI Slop or Verbosity!

Most assignments will be focused on implementing techniques covered in class, but you will sometimes be asked questions with text answers. For example, you may be asked to explain, motivate or otherwise argue for an approach. In those cases, you are expected to give a concise and direct answer and not be unnecessarily verbose. Points may be deducted for poorly written responses.

^^ in the formal syllabus.

You may use generative AI but use it well!

- Don’t waste the grader’s time.
- Don’t lose points for
  - Saying it is such an interesting question…
  - Answering a couple other questions!
  - Answering the next question too.
  - Making your answer pretty with redundant headings or emojis.
  - Repeating something an LLM hallucinated.
- Be concise and to the point.
- Submit something better than ChatGPT!
- Don’t forget to cite! A chat link suffices.

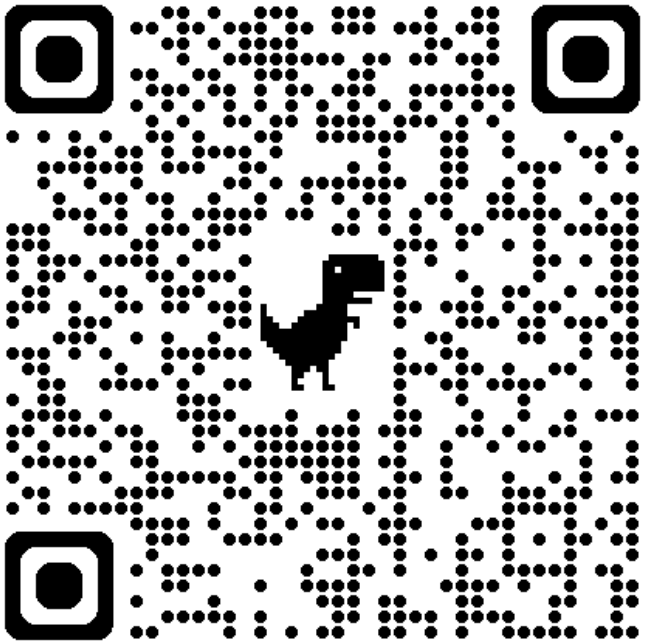
Grades

Item	Percentage
Discussions (due at end of discussion)	15%
Homework (~1 week each) and Pre-Work Reflections	15%
Projects (~2 weeks each)	70%

- 1st homework is out today (first class) on some refreshers
- Discussions notebook due end of day Fridays
- Late submissions up to 48 hours late
  - 20% penalty

Lecture Feedback

Only accessible with BU account.



https://forms.gle/pXHM5nx1Ti9aFmpw6