



CORNELL
TECH

Fall 2019

Deep Learning Clinic

Lecture 1 - Introduction

Jin Sun

9/3/2019

<https://cornelltech.github.io/deep-learning-clinic-2019-Fall/>

Today

- ❑ Overview of the class
 - ❑ Background
 - ❑ Logistics
- ❑ Lectures preview
- ❑ Office hours preview
- ❑ FAQ

Course page:

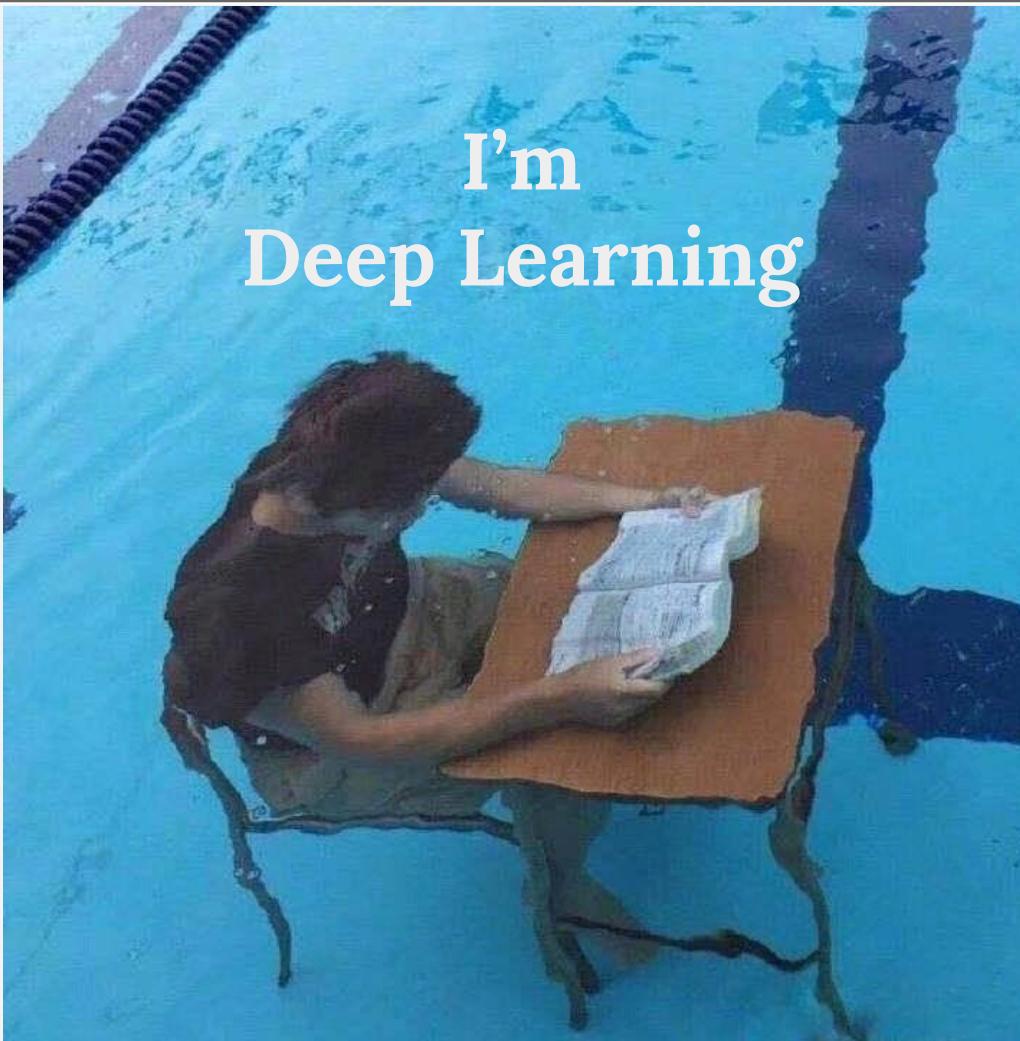
<https://cornelltech.github.io/deep-learning-clinic-2019-Fall/>

Slack channel:

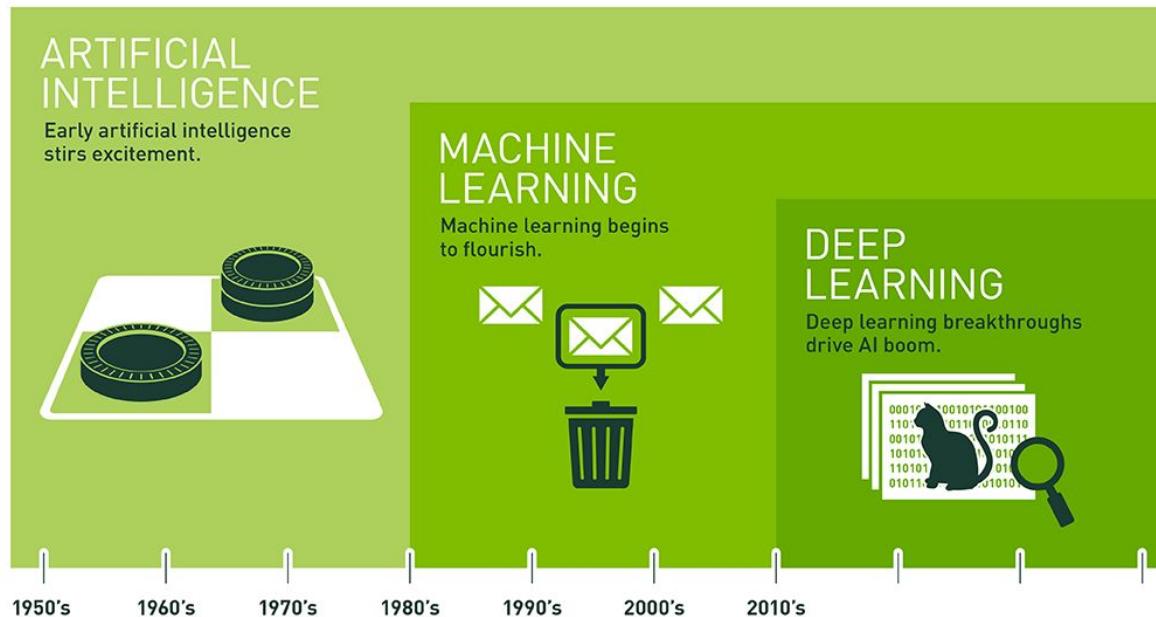
<https://cornelltechdlc.slack.com/>

What is Deep Learning

I'm
Deep Learning



What is Deep Learning



Since an early flush of optimism in the 1950s, smaller subsets of artificial intelligence – first machine learning, then deep learning, a subset of machine learning – have created ever larger disruptions.

What can DL do

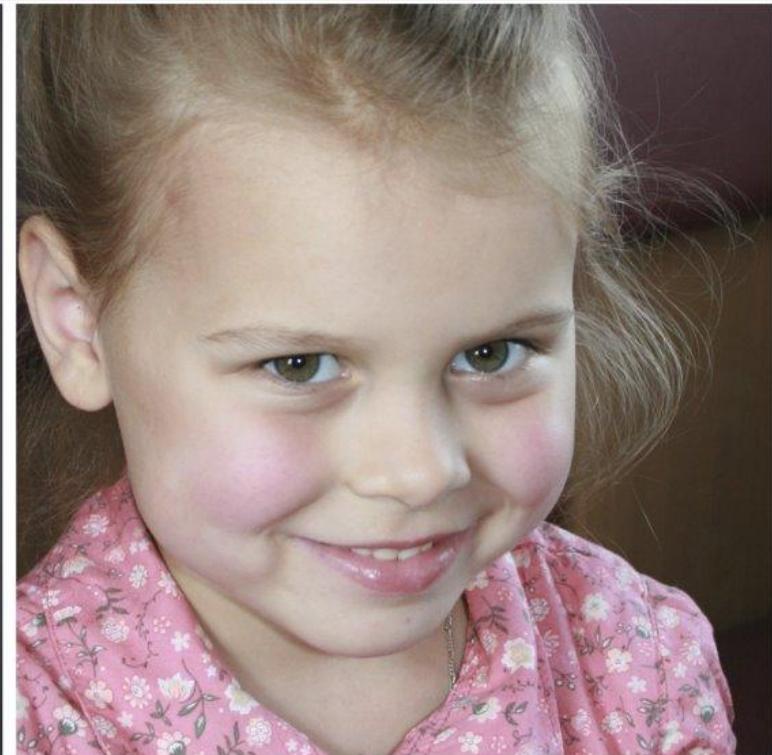
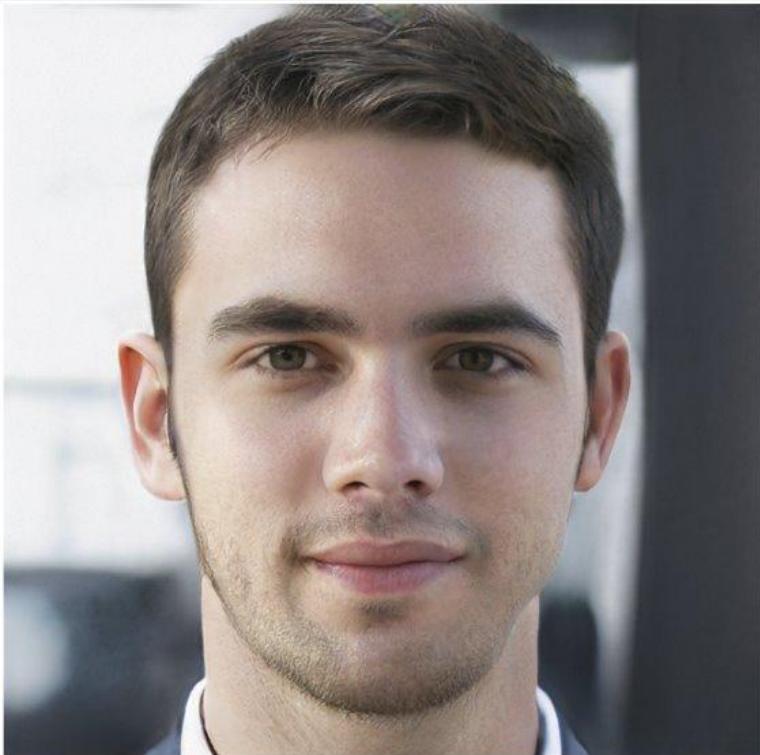
Artificial Intelligence (AI) to the general public:



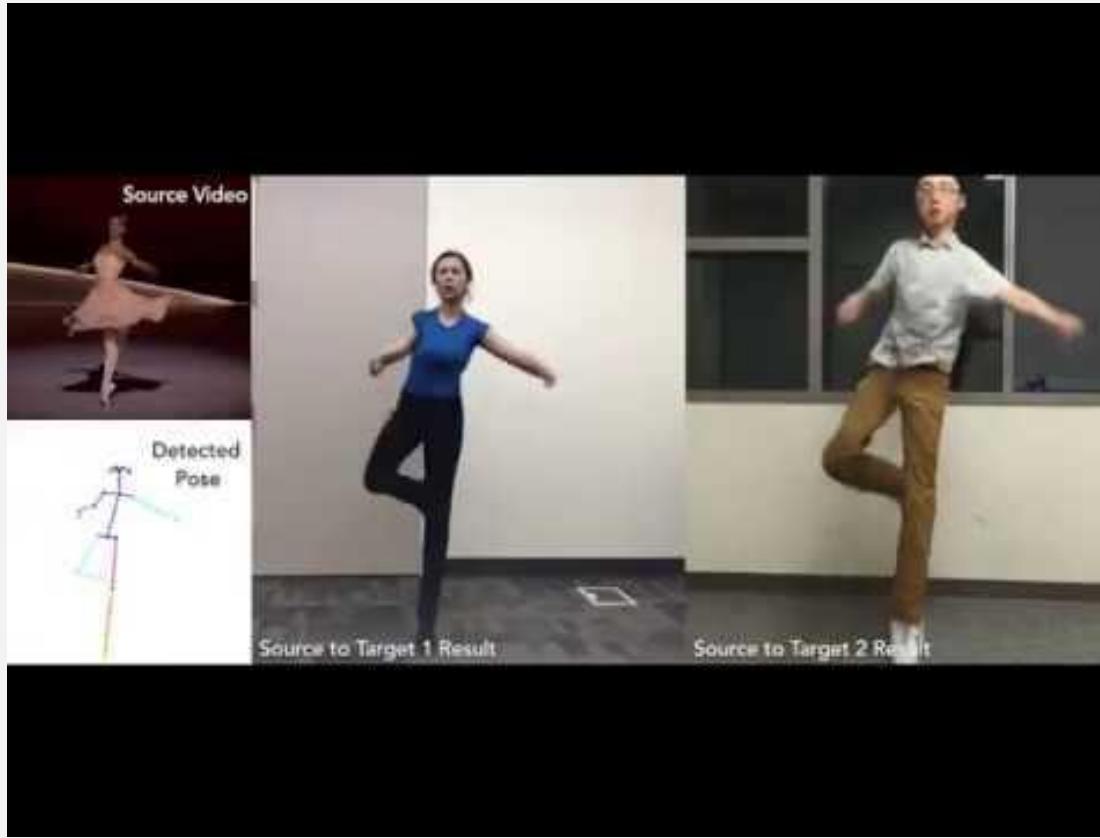
What can DL do

We are not there yet but ML/DL can do many cool things:





Which face is real?



[Everybody Dance Now](#)

Overview of DLC - Background

“The true challenge to artificial intelligence proved to be solving the tasks that are easy for people to perform but hard for people to describe formally.

...

This solution is to allow computers to **learn from experience** and understand the world in terms of **a hierarchy of concepts**... enables the computer to **learn complicated concepts by building them out of simpler ones**.

If we draw a graph showing how these concepts are built on top of each other, **the graph is deep...**

For this reason, we call this approach to AI ***deep learning***.”

-- *Deep Learning* by Ian Goodfellow and Yoshua Bengio and Aaron Courville

Overview of DLC - Background

What DL is good for

Problems with massive data

Classification problems

Regression problems

Correlation relationship

What DL is not good at

Limited amount data

Partial labels

Structure output

Long term relationship

Causal relationship

Overview of DLC

A horizontal collage of nine images related to deep learning, arranged in three rows. The top row contains five boxes with titles: 'Deep Learning Tutorial PDF', 'Open Source Deep Learning Tutorial', 'Deep Learning Examples', and 'Deep Learning Explained'. The middle row contains three images: a line graph showing the growth of deep learning at Google from 2013 to 2015; a diagram of an example neural network; and a hexagonal diagram representing various deep learning tools and platforms. The bottom row contains four images: a graph illustrating the performance plateau effect; a book cover for 'Deep Learning' by Kelly Howell; a diagram showing GPU acceleration work; and a circular diagram illustrating the processing of a natural scene.

DL/ML Basic Concepts

Real World Application

Hands On

Step-by-step

Deep Learning Clinic

Practical Guidance

One-on-one

Get Started

Technical Background

What tools to use

For You

Eager to solve complex real world problems

How to use existing tools

How to train a DL model

Overview of DLC - Logistics

Lecture Session

Tue 9:30-10:45am, Bloomberg 061

- ❑ Introduction to:
 - ❑ Techniques
 - ❑ Tools
 - ❑ Tricks
 - ❑ How to solve a real-world problem

Office Hour Session

Wed 2:00-3:00pm, Bloomberg 316

- ❑ One-on-one advice
- ❑ Problem solving:
 - ❑ Feasibility evaluation
 - ❑ Modeling and task formulation
 - ❑ Network design
 - ❑ Practical guidance on training

Zero-credit, no assignments or evaluations.

Jin Sun

jinsun@cornell.edu - *Please include the tag 'DLC' in the subject

Research Interest:
Computer Vision (object detection, scene understanding)

Office Hour:
Wed 2-3pm Bloomberg 316

<http://www.cs.cornell.edu/~jinsun/>

Slack channel: <https://cornelltechdlc.slack.com/>, open to cornell.edu address

Reference

Online Courses

MIT 6.S191: Introduction to Deep Learning [link](#)

Stanford CS231n: Convolutional Neural Networks
for Visual Recognition [link](#)

Free Textbooks

A Course in Machine Learning by Hal Daume III
[link](#)

Deep Learning by Ian Goodfellow and Yoshua
Bengio and Aaron Courville [link](#)

Today

- ❑ Overview of the class
 - ❑ Background
 - ❑ Logistics
- ❑ **Lectures preview** - Whole semesters lectures in a day
- ❑ Office hours preview
- ❑ FAQ

Lec 1: Deep Learning Frameworks

How to set up a basic deep learning development environment:

Python virtualenv

Anaconda

Docker

Jupyter Notebook

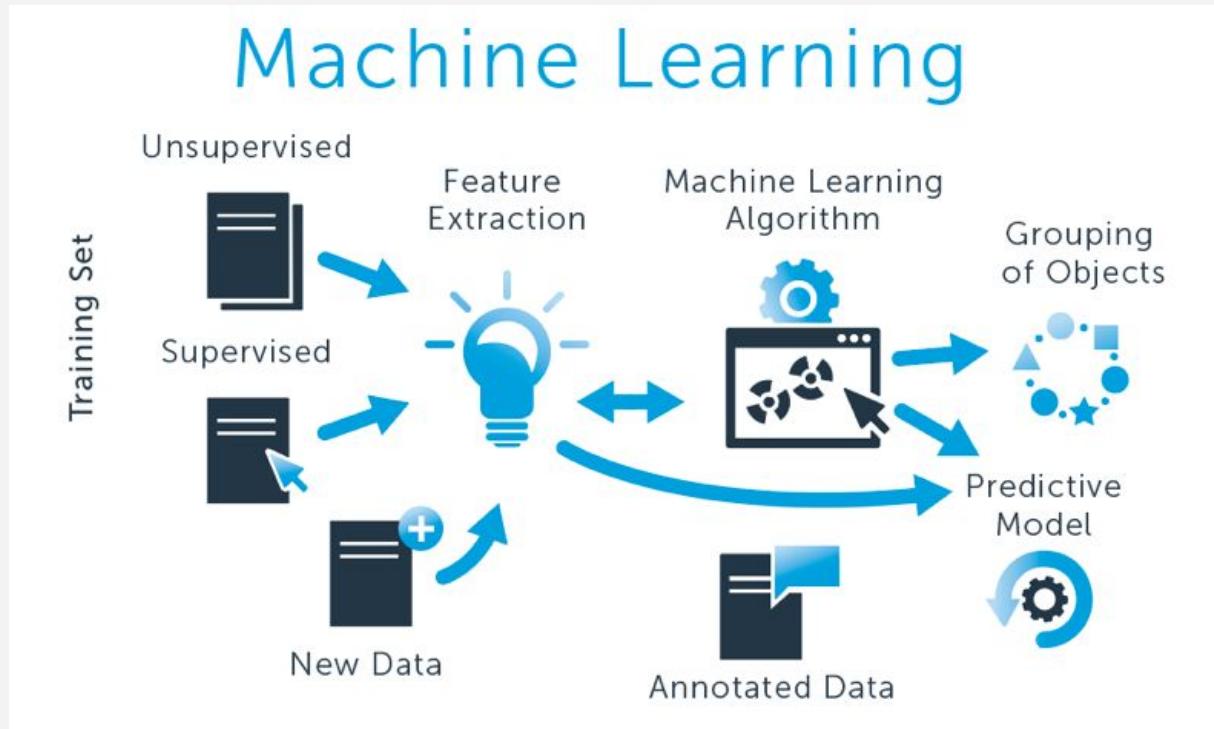
Lec 1: Deep Learning Frameworks

How to set up a basic deep learning development environment:



By the end of this lecture, you'll have a working DL environment to play with!

Lec 2: Brief Introduction to Machine Learning

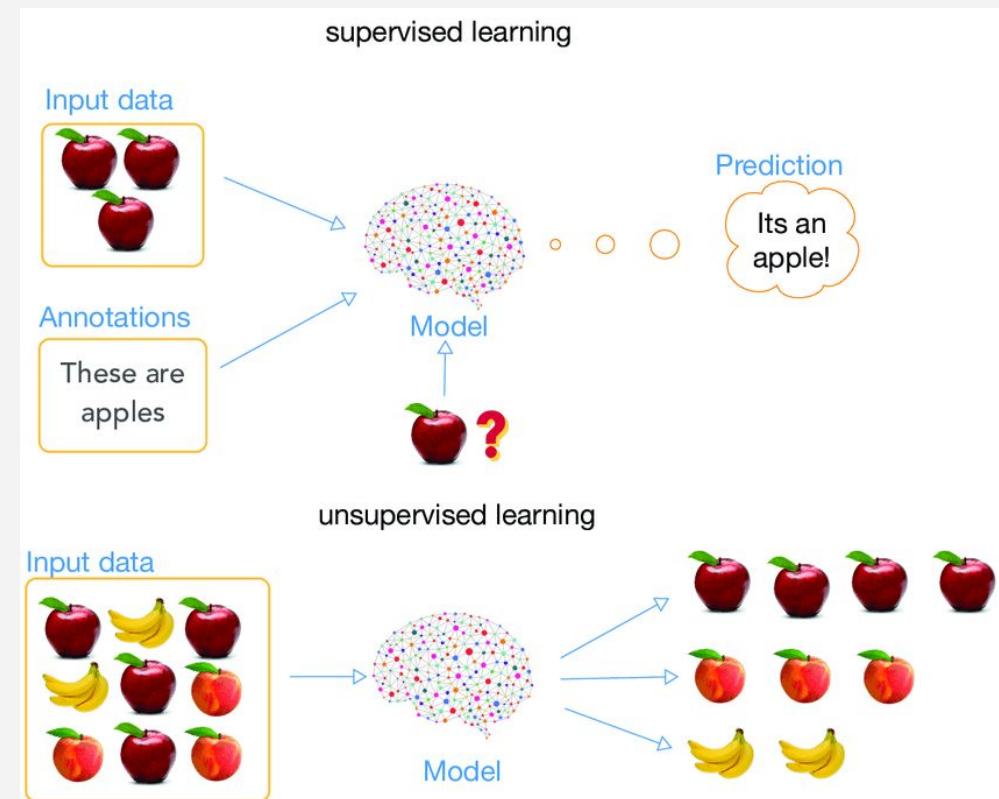


Lec 2: Brief Introduction to Machine Learning

Core ML concepts:

Supervised learning

Unsupervised learning

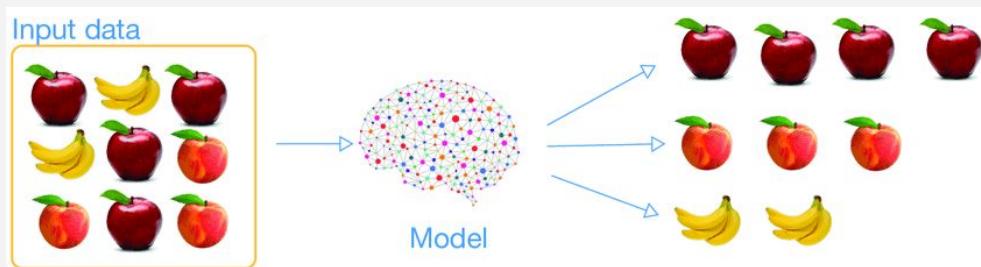


Lec 2: Brief Introduction to Machine Learning

Core ML concepts:

Classification

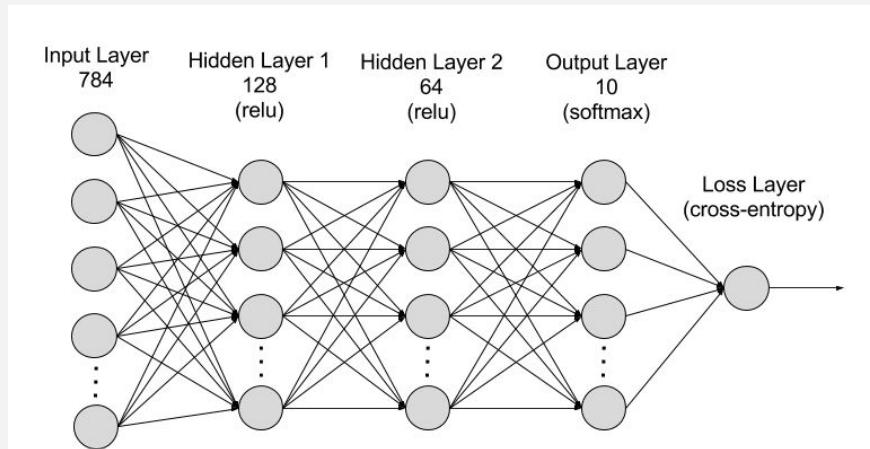
Regression



Lec 3: Brief Introduction to Deep Learning

Core DL topics:

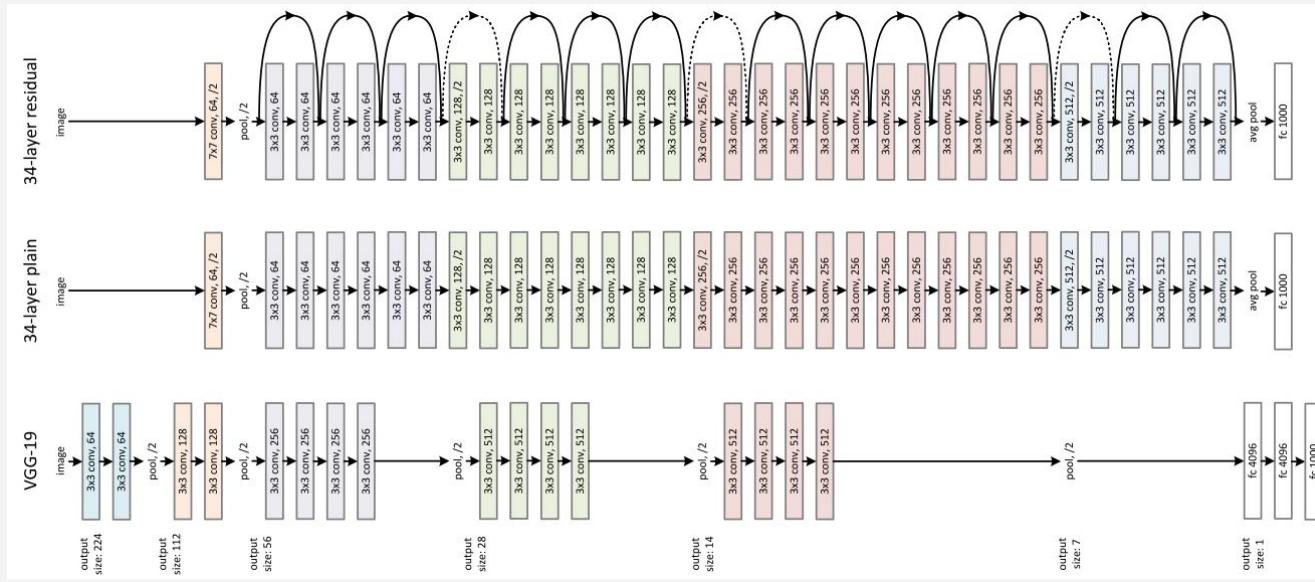
Network structures



Lec 3: Brief Introduction to Deep Learning

Core DL topics:

Network structures



Lec 3: Brief Introduction to Deep Learning

Core DL topics:

Common neural network building blocks

`torch.nn`

- Parameters
- + Containers
- + Convolution layers
- + Pooling layers
- + Padding layers
- + Non-linear activations (weighted sum, nonlinearity)
- + Non-linear activations (other)
- + Normalization layers
- + Recurrent layers

+ Transformer layers

+ Linear layers

+ Dropout layers

+ Sparse layers

+ Distance functions

+ Loss functions

+ Vision layers

+ DataParallel layers (multi-GPU, distributed)

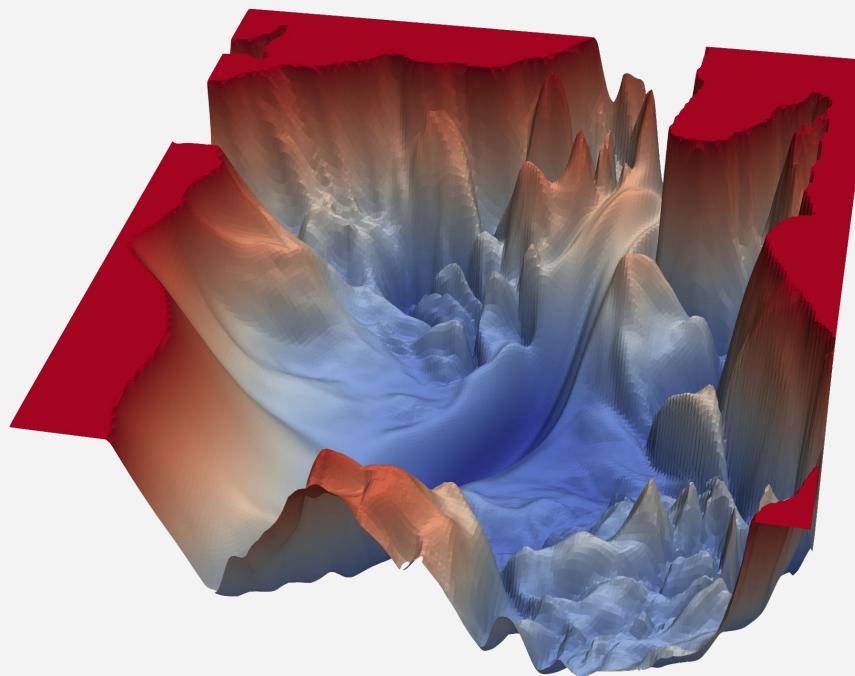
+ Utilities

Lec 3: Brief Introduction to Deep Learning

Core DL topics:

Backpropagation

Optimization



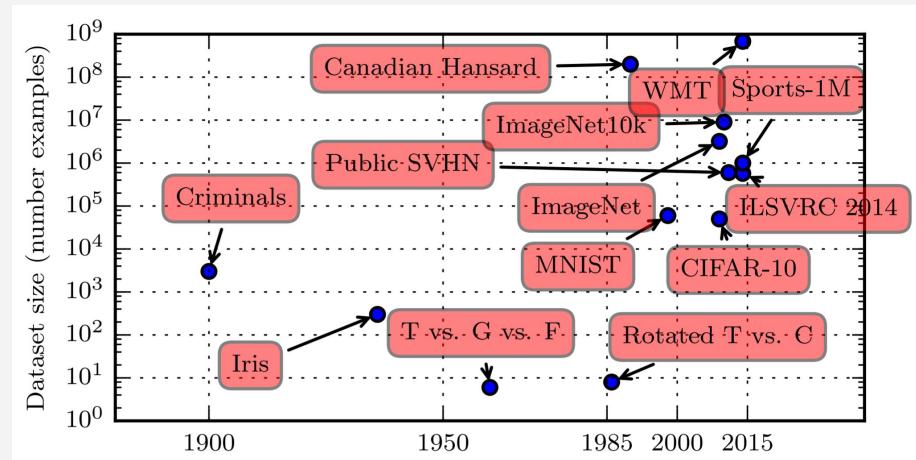
<https://www.cs.umd.edu/~tomg/projects/landscapes/>

Lec 4: Data

Main Reasons Behind the Success of DL

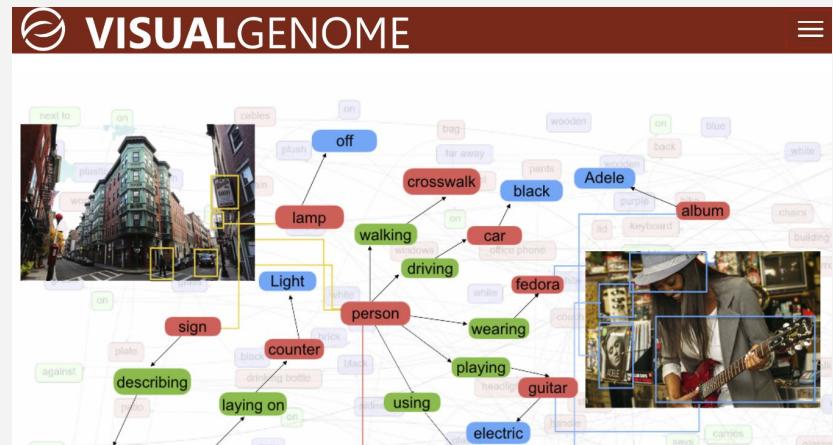
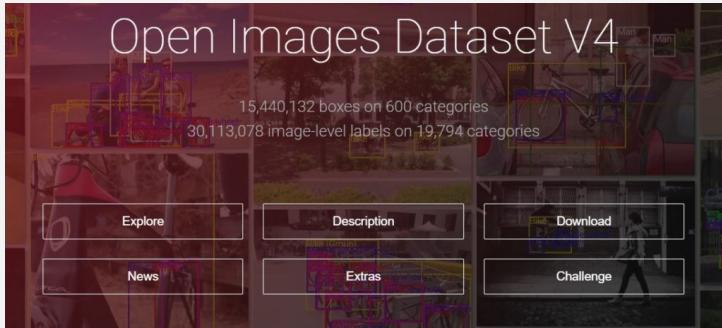
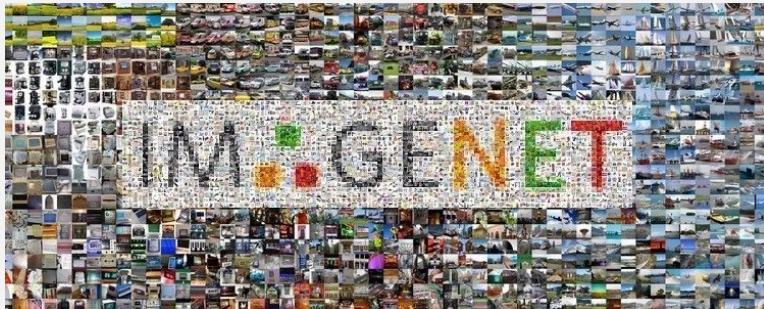


Hardware



Data

Lec 4: Data



Lec 4: Data

Data Collection, Annotation, and Verification

The screenshot shows the Amazon Mechanical Turk homepage. At the top, there's a navigation bar with links for "Your Account", "HITs", "Qualifications", and "Sign in as a Worker | Requester". Below the navigation bar, a yellow banner states: "Mechanical Turk is a marketplace for work. We give businesses and developers access to an on-demand, scalable workforce. Workers select from thousands of tasks and work whenever it's convenient." It also displays the number "405,999 HITs available." A "View them now." button is present. The main content area is divided into two main sections: "Make Money by working on HITs" for workers and "Get Results from Mechanical Turk Workers" for requesters. Both sections include descriptions and lists of benefits, along with "Find HITs Now" and "Get Started" buttons.

already have an account?
Sign in as a [Worker](#) | [Requester](#)

Mechanical Turk is a marketplace for work.
We give businesses and developers access to an on-demand, scalable workforce.
Workers select from thousands of tasks and work whenever it's convenient.

405,999 HITs available. [View them now.](#)

Make Money
by working on HITs

HITs - *Human Intelligence Tasks* - are individual tasks that you work on. [Find HITs now.](#)

As a Mechanical Turk Worker you:

- Can work from home
- Choose your own work hours
- Get paid for doing good work

Find an interesting task → **Work** → **Earn money**

[Find HITs Now](#)

or learn more about being a Worker

Get Results
from Mechanical Turk Workers

Ask workers to complete HITs - *Human Intelligence Tasks* - and get results using Mechanical Turk. [Get Started.](#)

As a Mechanical Turk Requester you:

- Have access to a global, on-demand, 24 x 7 workforce
- Get thousands of HITs completed in minutes
- Pay only when you're satisfied with the results

Find your account → **Load your tasks** → **Get results**

[Get Started](#)

Lec 5: Neural Network Architectures

Model search: How to find the best structure for your problem

There are a lot of design choices:

Number of layers, which non-linear layer to use, number of filters and etc.

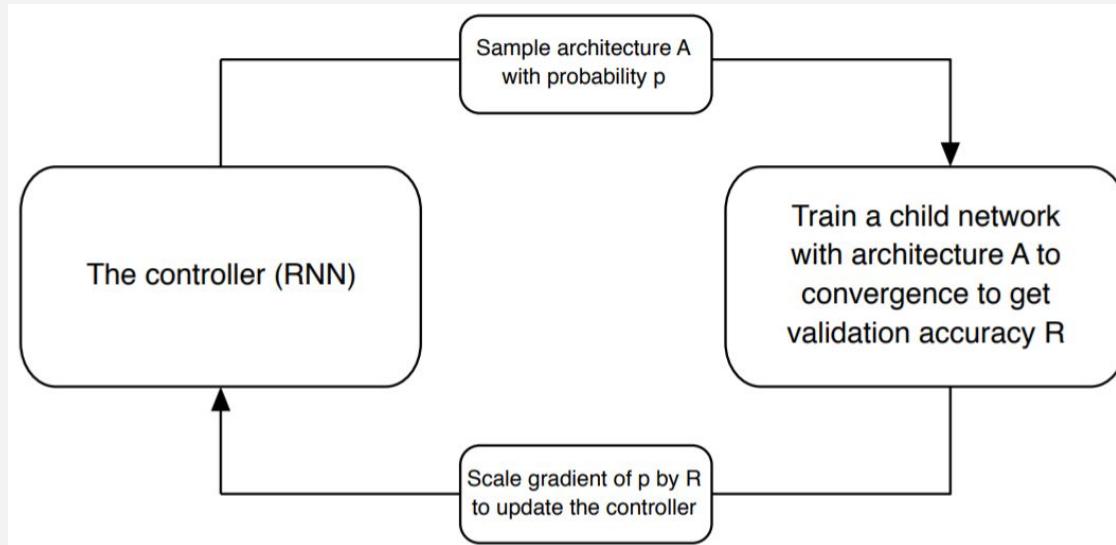
Lec 5: Neural Network Architectures

Model search: How to find the best structure for your problem



Lec 5: Neural Network Architectures

Model search: Google AutoML



<https://arxiv.org/abs/1611.01578>

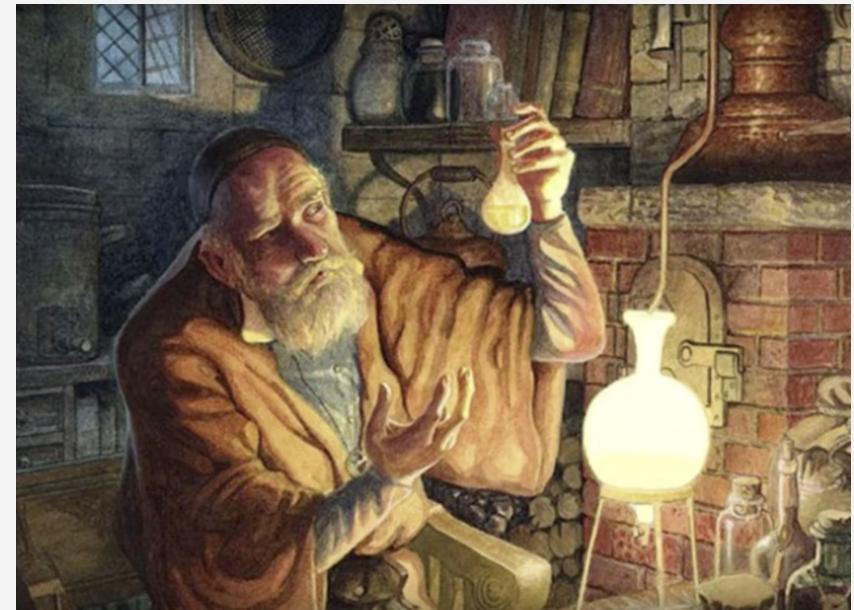
<https://arxiv.org/pdf/1707.07012.pdf>

Lec 6: Tricks on Training Neural Networks

Training a (good) neural network is an art.

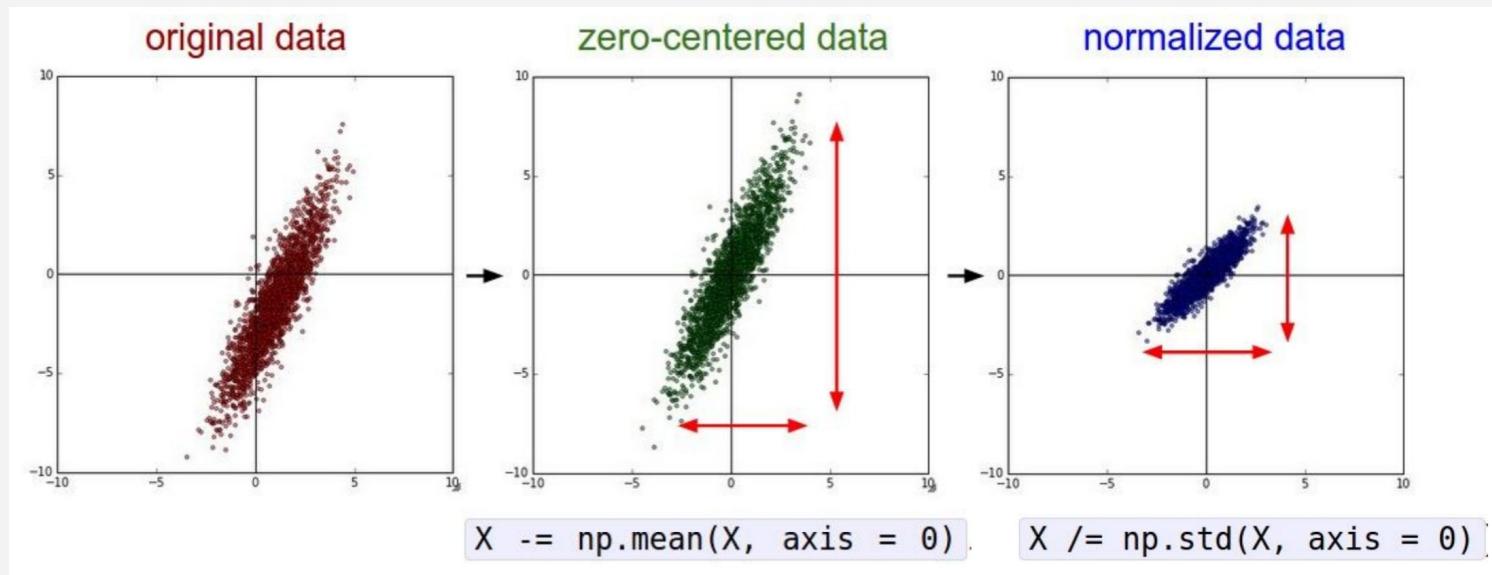
There is no principal approach that ‘guarantee to work’.

But there are tricks people found useful.



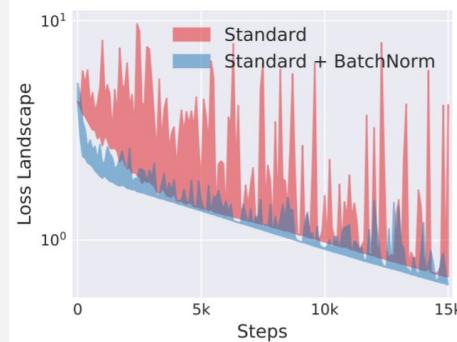
Lec 6: Tricks on Training Neural Networks

Data Pre-processing



Lec 6: Tricks on Training Neural Networks

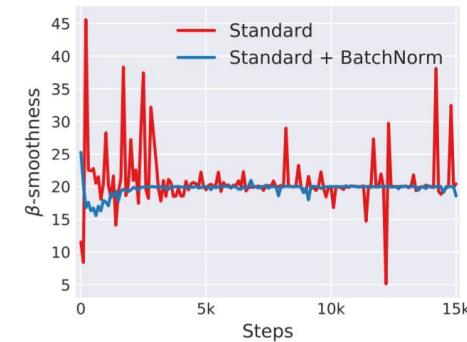
Adding special layers such as BatchNorm



(a) loss landscape



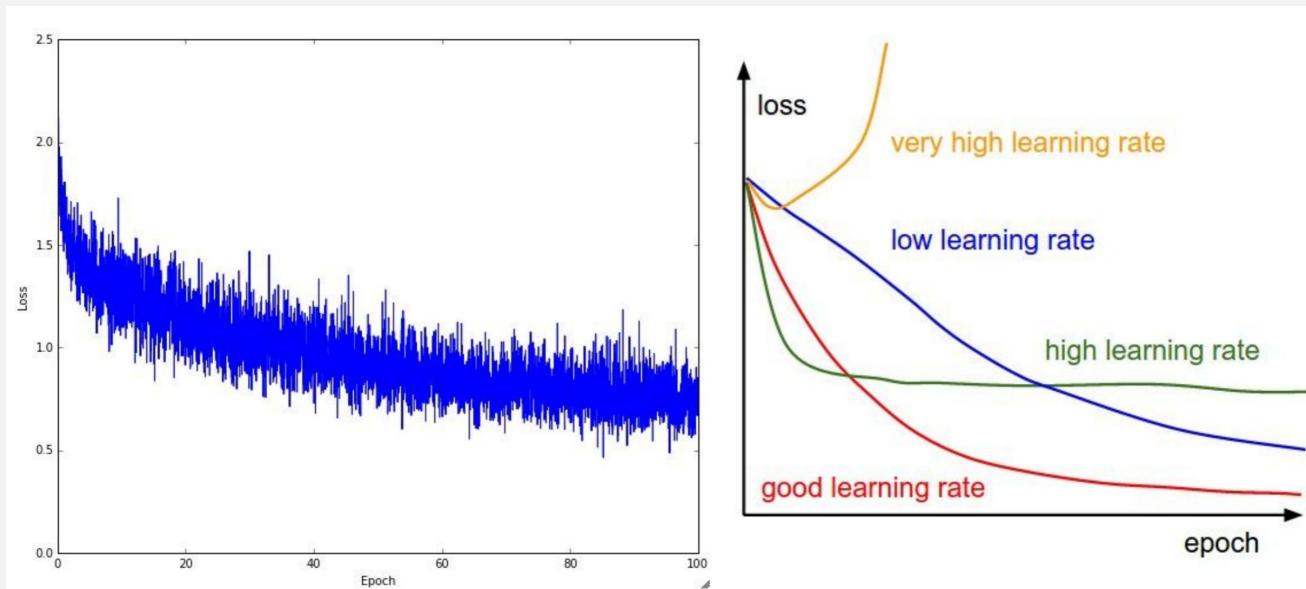
(b) gradient predictiveness



(c) “effective” β -smoothness

Lec 6: Tricks on Training Neural Networks

How to diagnose training curves



Lec 7: Machine Learning Model Evaluation

How do I know which model is the best?

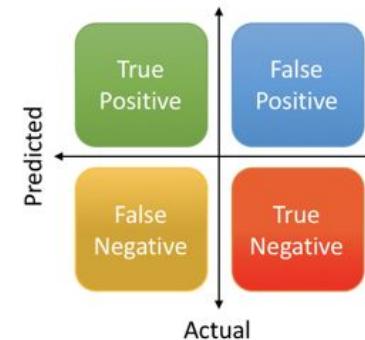
The screenshot shows the COCO Detection Leaderboard page. At the top, there is a navigation bar with the COCO logo, the text "Common Objects in Context", and links for "Home", "People", "Dataset", "Tasks", and "Evaluate". The "Evaluate" link is highlighted with a green background. Below the navigation bar, the title "Detection Leaderboard" is displayed. There are two sets of tabs for filtering results: "BBOX" and "SEGMENTATION". Under "BBOX", the "Dev" tab is selected, while "Standard15", "Chal15", "Chal16", and "Chal17" are shown in blue. Under "SEGMENTATION", all tabs ("Dev", "Standard15", "Chal15", "Chal16", "Chal17", "Chal18") are shown in blue. Below the tabs, there are buttons for "Copy to Clipboard" and "Export to CSV", and a search bar. The main content is a table with the following columns: Rank, Model Name, AP, AP⁵⁰, AP⁷⁵, AP^S, AP^M, AP^L, AR^I, AR^T, AR¹⁰⁰, AR^S, AR^M, AR^L, and date. The table lists 13 models, each represented by a green circular icon with a white plus sign. The data for the first few rows is as follows:

Rank	Model Name	AP	AP ⁵⁰	AP ⁷⁵	AP ^S	AP ^M	AP ^L	AR ^I	AR ^T	AR ¹⁰⁰	AR ^S	AR ^M	AR ^L	Date
1	Megvii (Face++)	0.526	0.730	0.585	0.343	0.556	0.660	0.391	0.645	0.689	0.513	0.727	0.827	2017-10-05
2	CM-CV AR	0.525	0.717	0.578	0.352	0.550	0.642	0.392	0.647	0.689	0.528	0.720	0.821	2019-07-26
3	Night owl	0.519	0.704	0.570	0.342	0.548	0.647	0.391	0.640	0.680	0.498	0.713	0.824	2019-08-05
4	Alibaba Turing Lab	0.514	0.694	0.563	0.336	0.540	0.639	0.388	0.638	0.679	0.499	0.711	0.818	2019-08-05
5	UCenter	0.510	0.705	0.558	0.326	0.539	0.648	0.392	0.640	0.678	0.497	0.720	0.829	2017-10-05
6	MSRA	0.507	0.717	0.566	0.343	0.529	0.627	0.379	0.638	0.690	0.524	0.720	0.824	2017-10-05
7	DL-61	0.507	0.708	0.569	0.337	0.534	0.626	0.383	0.639	0.686	0.516	0.718	0.821	2018-08-15
8	FAIR Mask R-CNN	0.503	0.720	0.558	0.328	0.537	0.627	0.380	0.622	0.659	0.485	0.704	0.800	2017-10-05
9	fanglinpu	0.491	0.693	0.542	0.304	0.527	0.617	0.374	0.612	0.654	0.456	0.693	0.804	2019-08-02
10	KPLab	0.488	0.691	0.543	0.316	0.520	0.613	0.374	0.622	0.670	0.478	0.708	0.819	2018-08-14
11	kth0307	0.486	0.669	0.530	0.319	0.513	0.595	0.366	0.614	0.670	0.484	0.703	0.811	2019-07-31

Lec 7: Machine Learning Model Evaluation

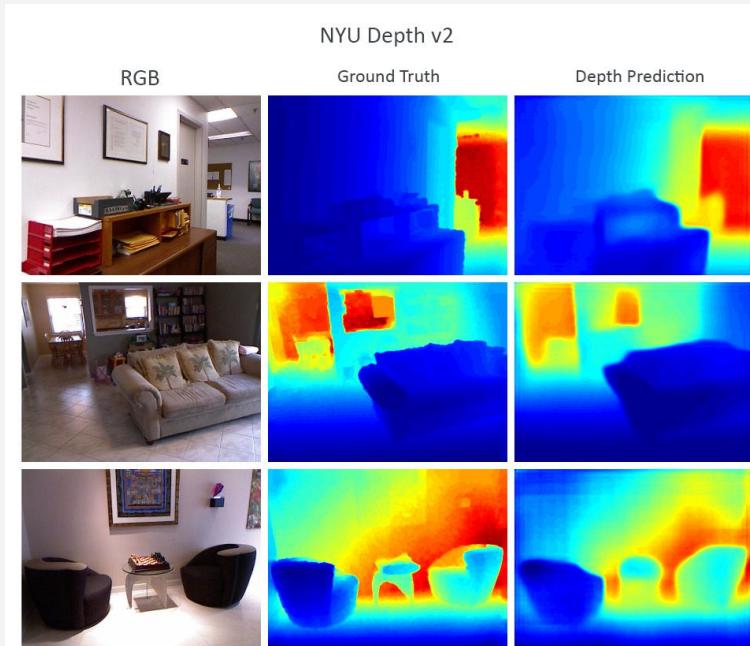
Classification metric

$$\text{Precision} = \frac{\text{True Positive}}{\text{Actual Results}} \quad \text{or} \quad \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$$
$$\text{Recall} = \frac{\text{True Positive}}{\text{Predicted Results}} \quad \text{or} \quad \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$$
$$\text{Accuracy} = \frac{\text{True Positive} + \text{True Negative}}{\text{Total}}$$



Lec 7: Machine Learning Model Evaluation

Regression metric



Threshold: % of y_i s.t. $\max\left(\frac{y_i}{y_i^*}, \frac{y_i^*}{y_i}\right) = \delta < thr$

Abs Relative difference: $\frac{1}{|T|} \sum_{y \in T} |y - y^*| / y^*$

Squared Relative difference: $\frac{1}{|T|} \sum_{y \in T} \|y - y^*\|^2 / y^*$

RMSE (linear): $\sqrt{\frac{1}{|T|} \sum_{y \in T} \|y_i - y_i^*\|^2}$

RMSE (log): $\sqrt{\frac{1}{|T|} \sum_{y \in T} \|\log y_i - \log y_i^*\|^2}$

RMSE (log, scale-invariant): The error Eqn. 1

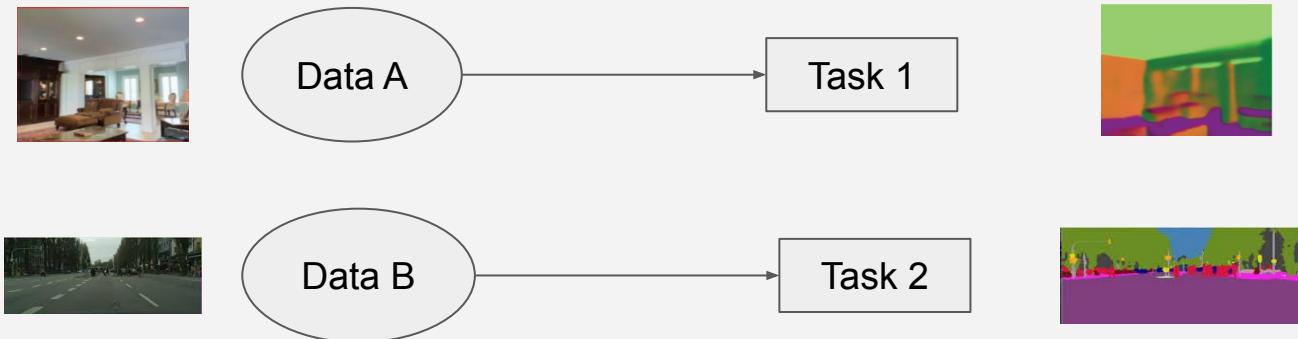
Lec 7: Machine Learning Model Evaluation

Task-dependent metric

SYSTEM A:	Israeli officials	responsibility of	airport	safety
	2-GRAM MATCH			
REFERENCE: Israeli officials are responsible for airport security				
SYSTEM B:	airport security	Israeli officials are responsible		
	2-GRAM MATCH	4-GRAM MATCH		
Metric	System A	System B		
precision (1gram)	3/6	6/6		
precision (2gram)	1/5	4/5		
precision (3gram)	0/4	2/4		
precision (4gram)	0/3	1/3		
brevity penalty	6/7	6/7		
BLEU	0%	52%		

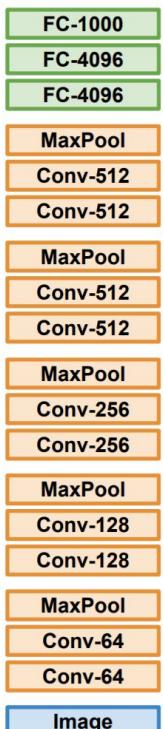
Lec 8: Case Study: Transfer Learning

- ❑ Fine-tuning
- ❑ Domain Transfer / Adaptation
- ❑ Multi-Task Learning
- ❑ Lifelong Learning

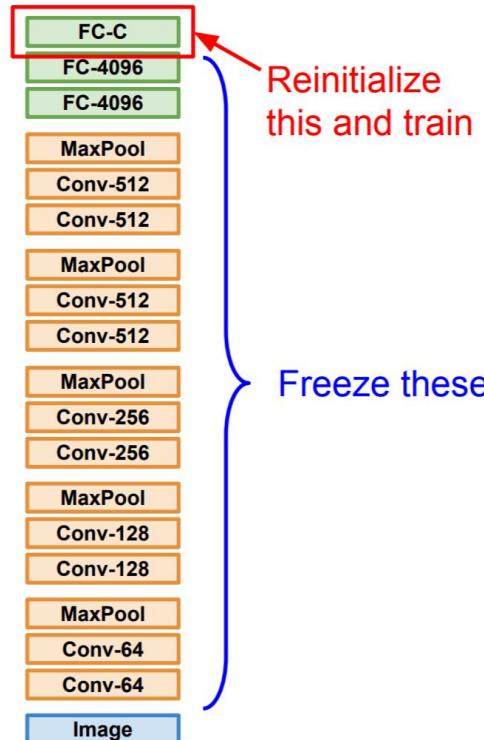


Lec 8: Case Study: Transfer Learning

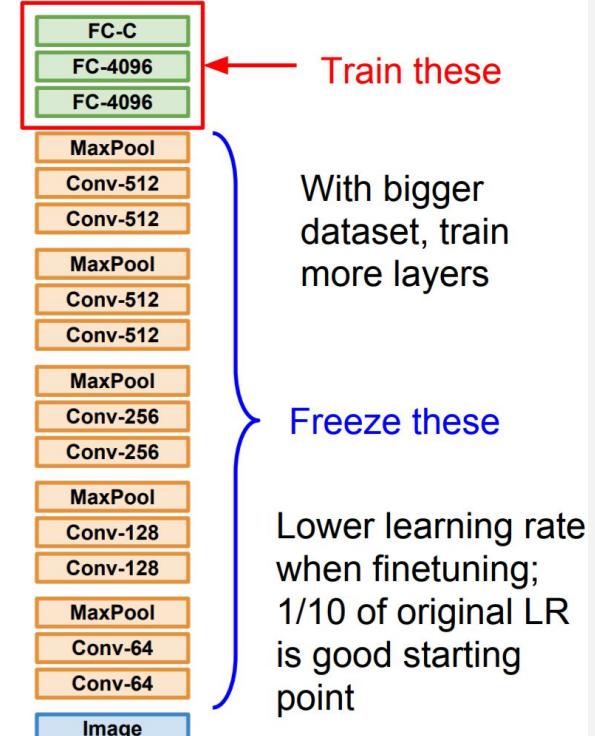
1. Train on Imagenet



2. Small Dataset (C classes)



3. Bigger dataset



Lec 8: Case Study: Transfer Learning

“What you saw is not what you get”

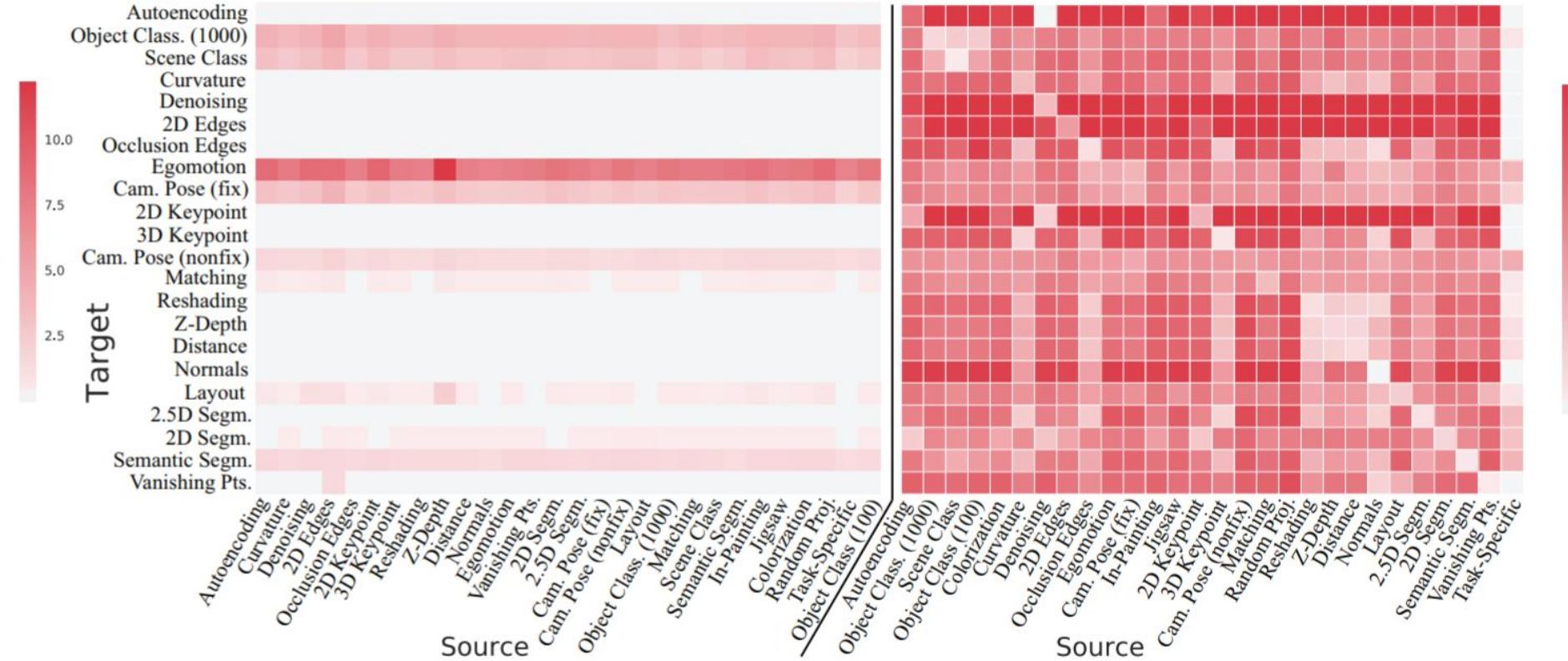


What your net is trained on

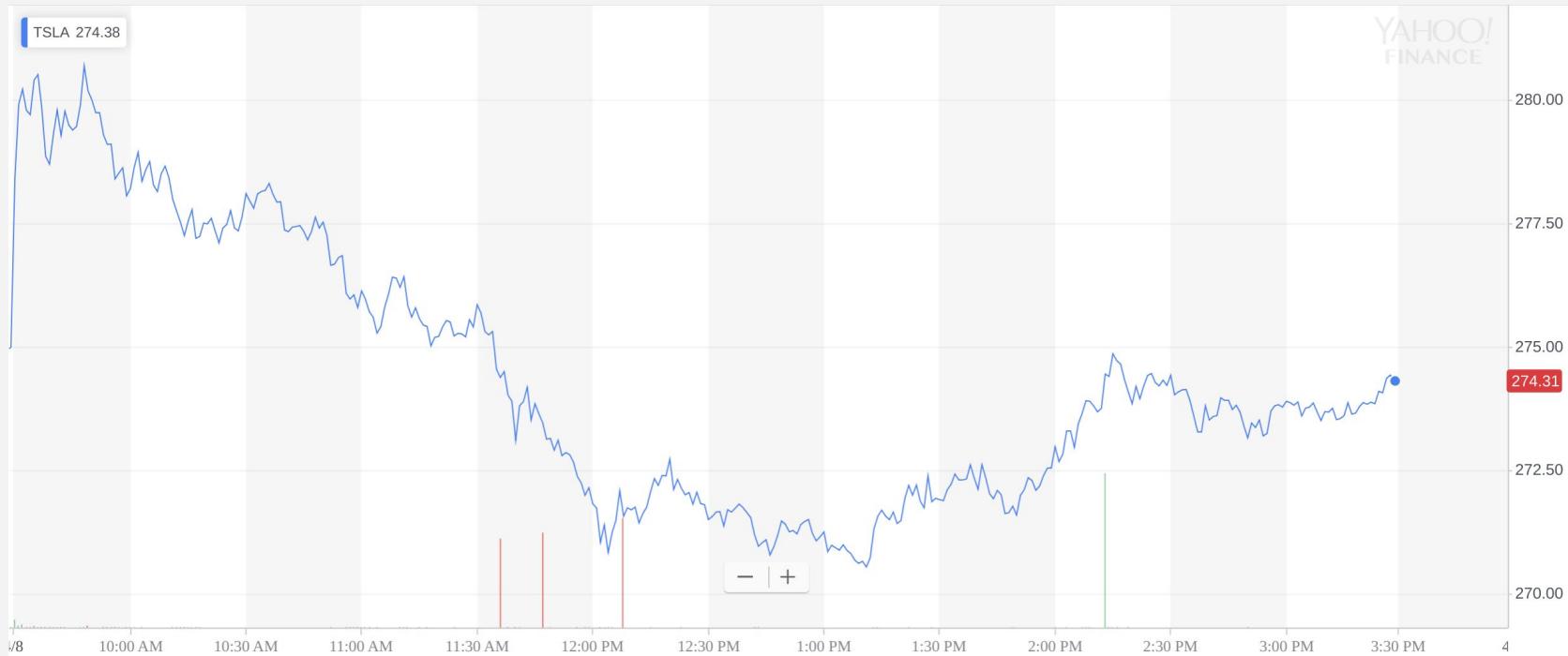


What it's asked to label

Lec 8: Case Study: Transfer Learning

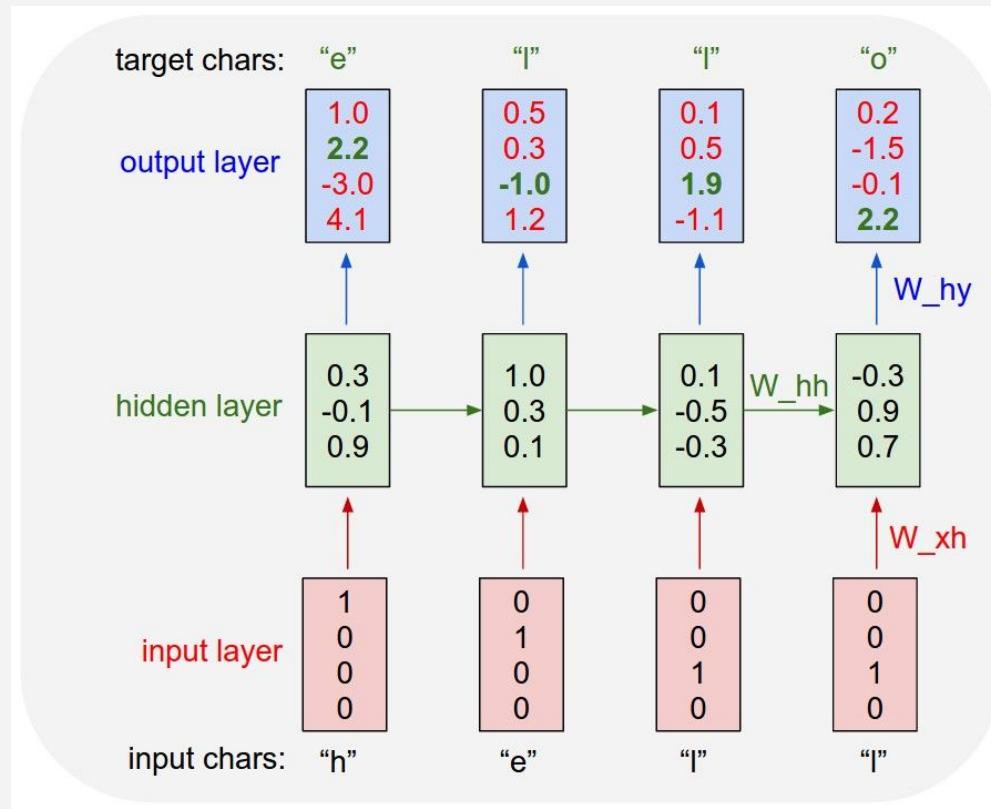


Lec 9: Case Study: Sequential Data Modeling



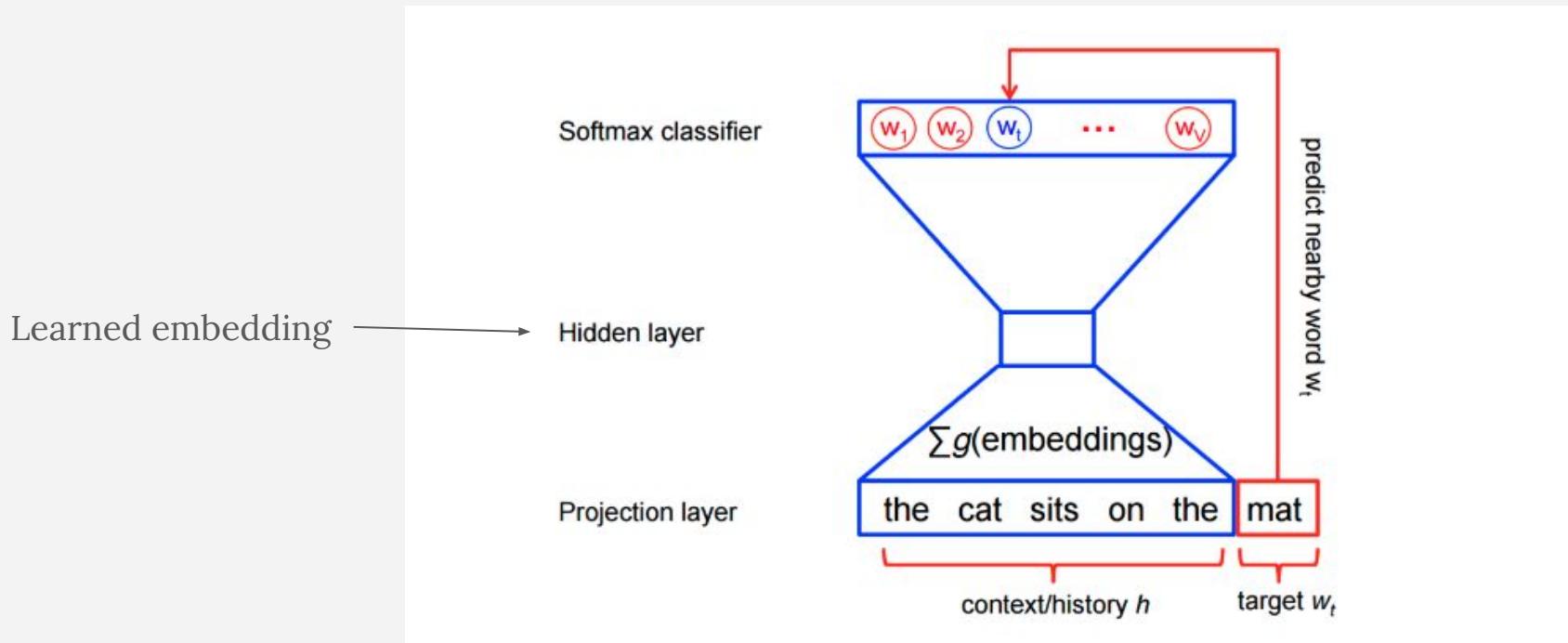
Lec 9: Case Study: Sequential Data Modeling

Simple RNNs
LSTM, GRU



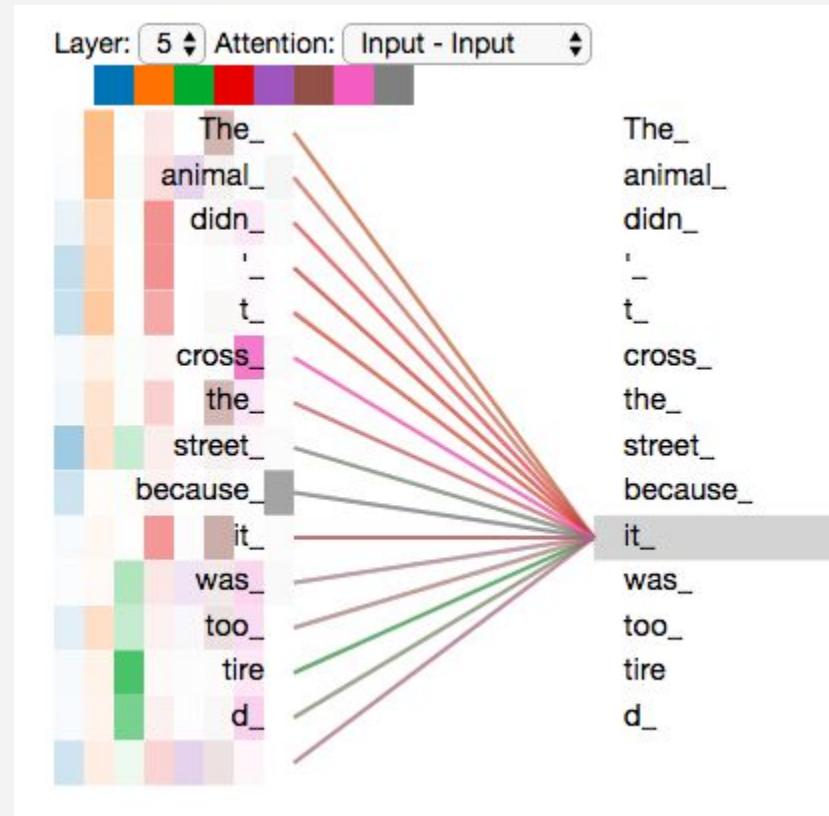
Lec 9: Case Study: Sequential Data Modeling

NLP Word2Vec Model



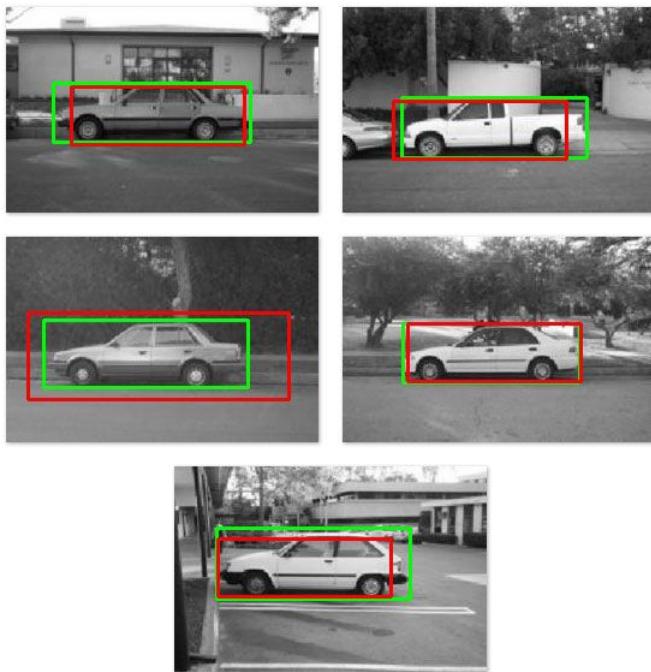
Lec 9: Case Study: Sequential Data Modeling

Attention Models

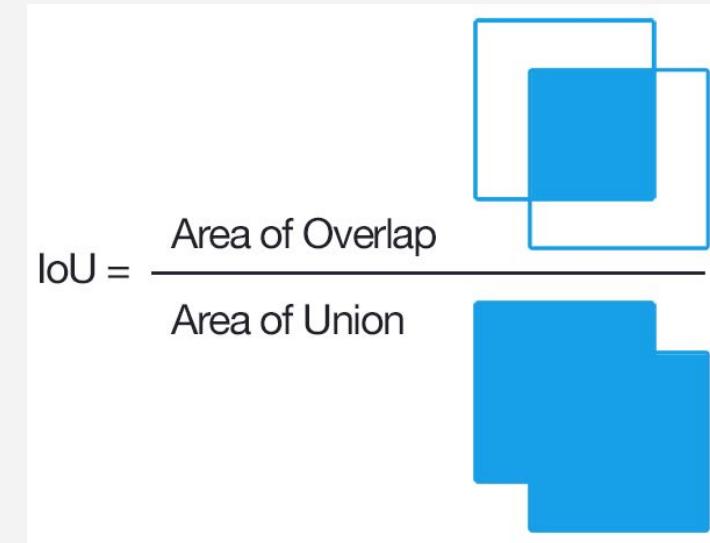


Lec 10: Case Study: Object Detection and Segmentation

Object Detection

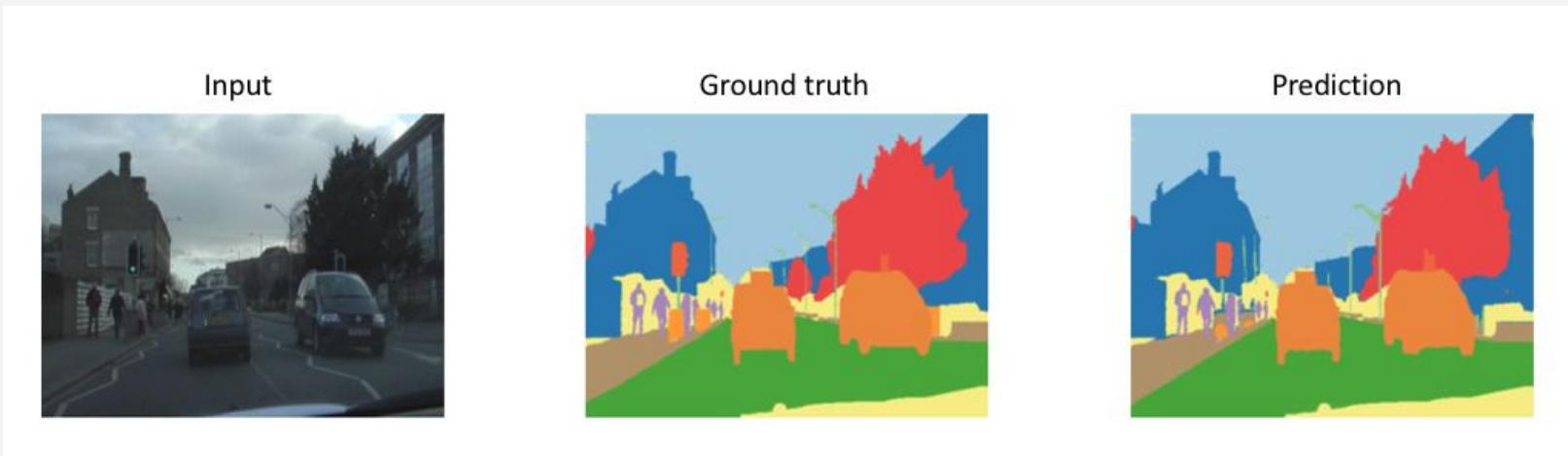


$$\text{IoU} = \frac{\text{Area of Overlap}}{\text{Area of Union}}$$



Lec 10: Case Study: Object Detection and Segmentation

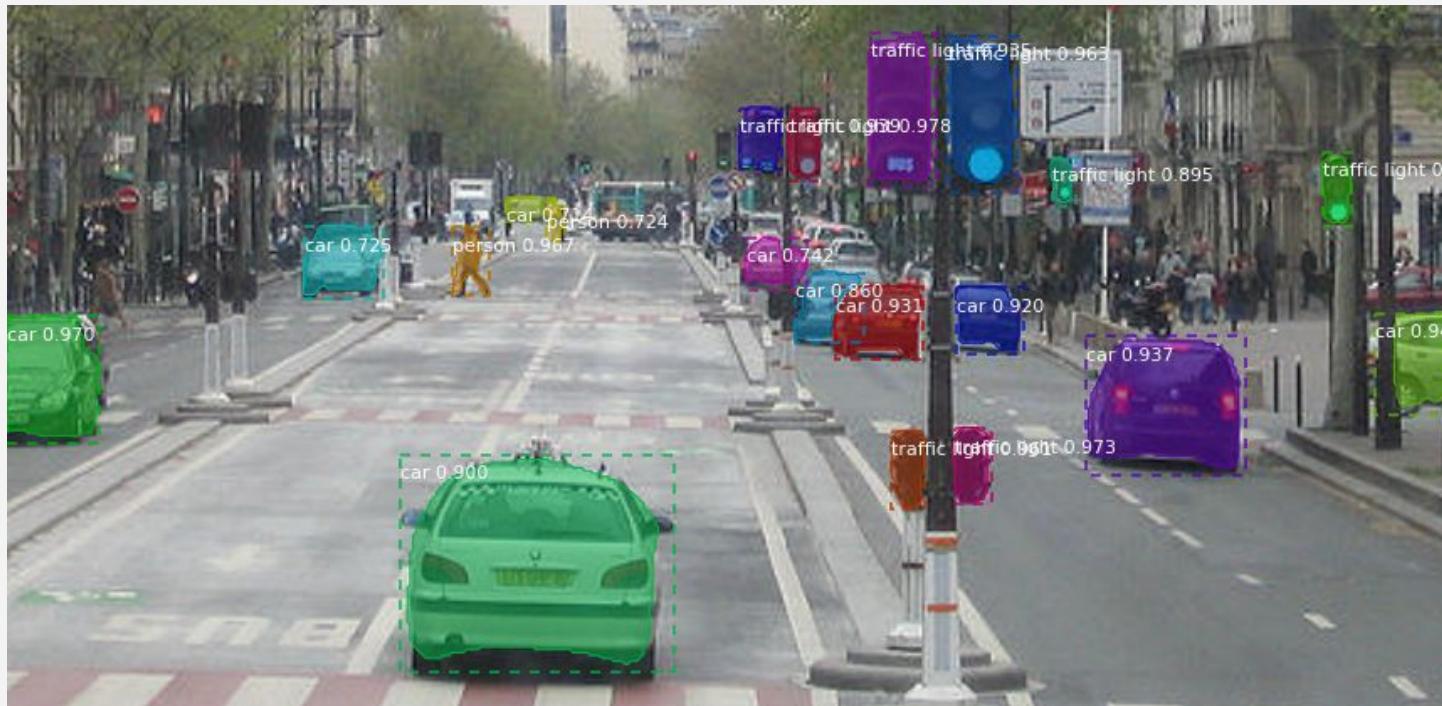
Semantic Segmentation



Green: road
Orange: car
Purple: person

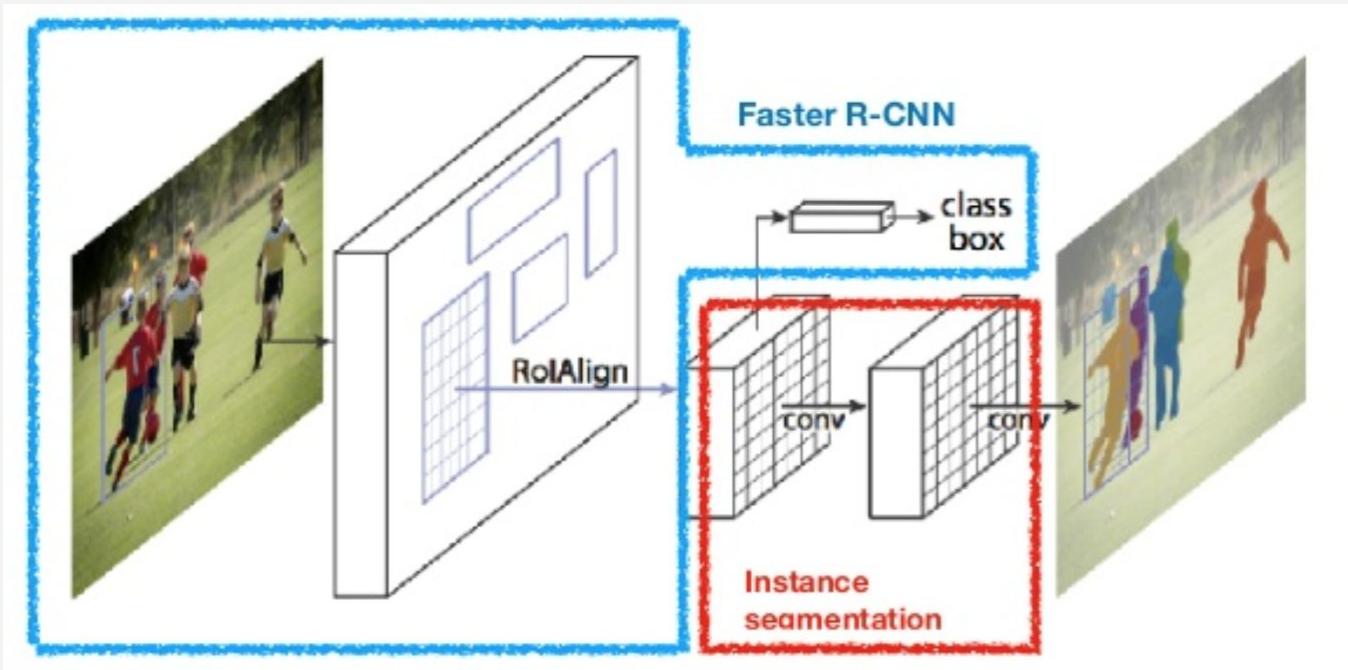
Lec 10: Case Study: Object Detection and Segmentation

Instance Segmentation

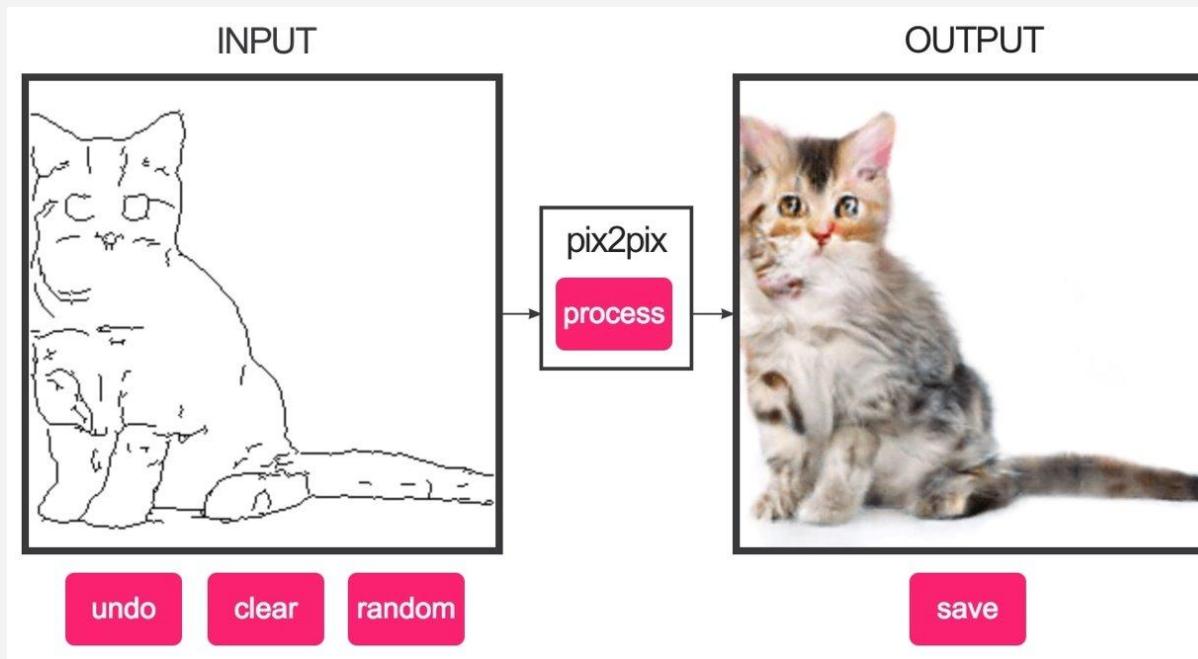


Lec 10: Case Study: Object Detection and Segmentation

Mask RCNN



Lec 11: Case Study: Generative Adversarial Networks



<https://affinelayer.com/pixsrv/>

Lec 11: Case Study: Generative Adversarial Networks

Fill/brush color: Brush shape: Brush size: 24

NVIDIA GauGAN Beta

Building
Ground
Landscape
Plant

Clouds
Fog
Hill
Mountain
River
Rock
Sea
Sky
Snow
Stone
Water







?

















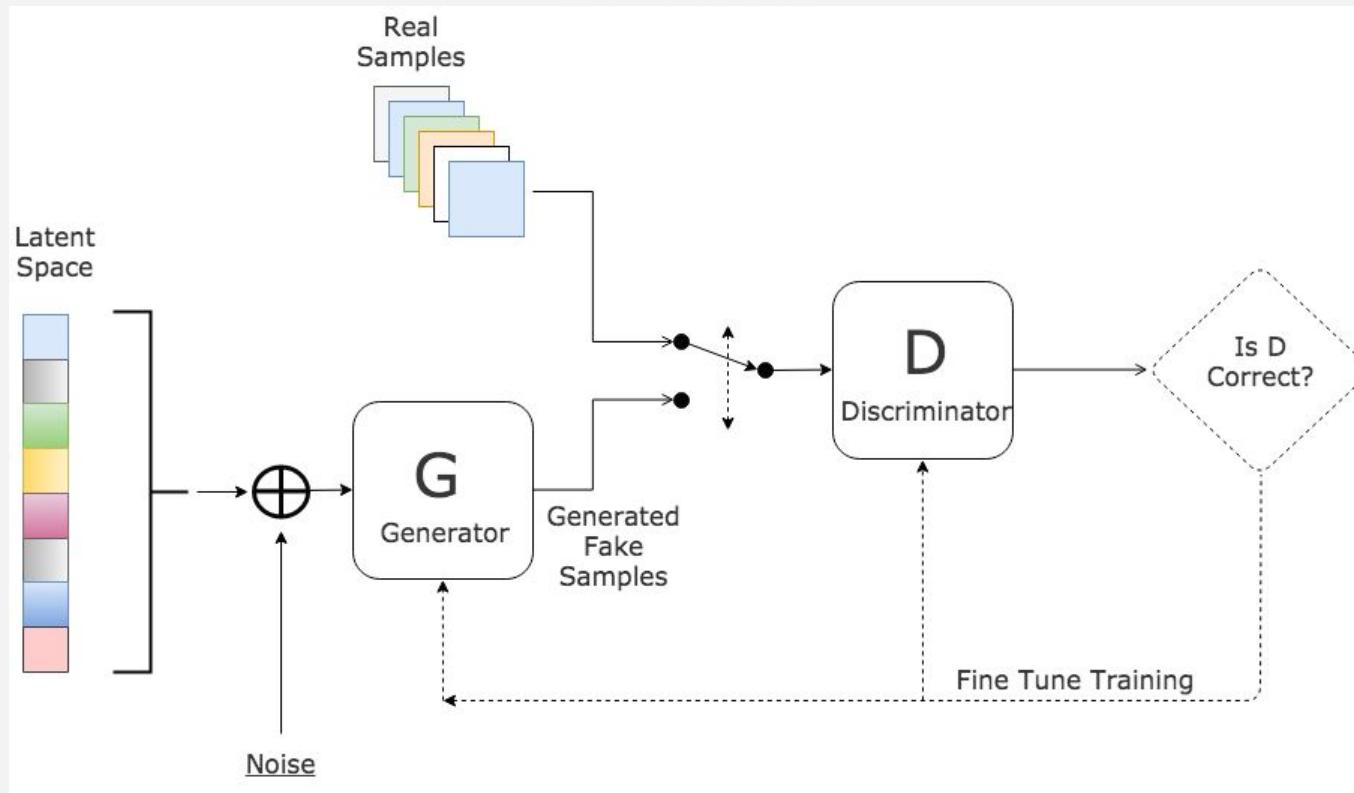






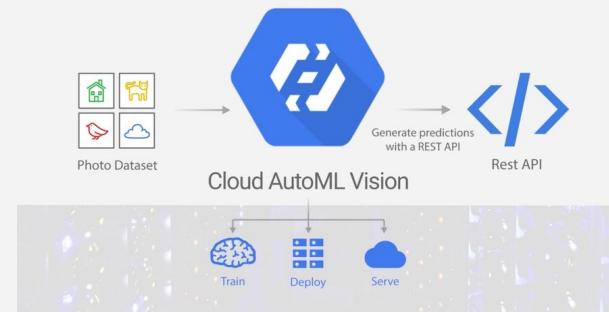
<http://nvidia-research-mingyuliu.com/gaugan>

Lec 11: Case Study: Generative Adversarial Networks



Lec 12: Real-World Ready Machine Learning Tools

Tools that are widely used and proven to be effective in real-world problems



Today

- ❑ Overview of the class
 - ❑ Background
 - ❑ Logistics
- ❑ Lectures preview
- ❑ **Office hours preview**
- ❑ FAQ

Office Hours Preview



Me

Your Deep Learning Concerns:

- Feasibility Evaluation
- Modeling and Task Formulation
- Network Search and Design
- Training Guidance

Case I

You have some idea about starting a research project or product studio project.

You'd like to know:

- 1) Is my problem too hard?
- 2) How to formulate my problem into a learning problem?
- 3) How to collect my data? How much data is needed?

Case II

You are working on your project involving in training some neural networks.

You were wondering:

- 1) How can I improve my network's performance?
- 2) What other models I can try?
- 3) My network is not training well, what's wrong?
- 4) Should I collect more data?

Today

- ❑ Overview of the class
 - ❑ Background
 - ❑ Logistics
- ❑ Lectures preview
- ❑ Office hours preview
- ❑ FAQ

FAQ

Can I attend only selected sessions?

Sure! But I recommend to you to attend the first few introduction lectures to grasp some basic understanding about DL/ML.

I don't have a strong technical background, can I still attend?

Of course! Come for the lectures that provide high level idea and how to use existing tools for quick prototyping.

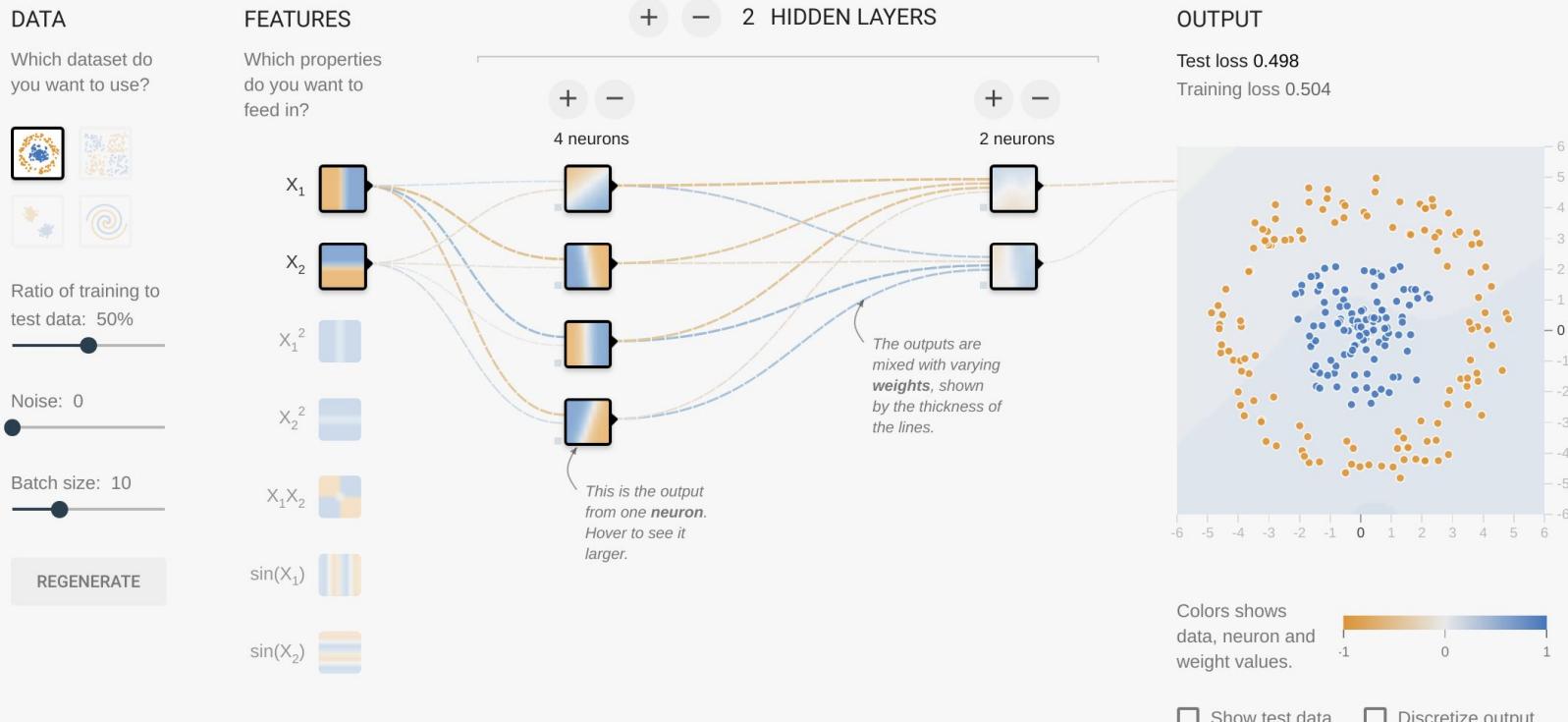
Next Week:

Deep Learning Frameworks

Extra

- ❑ Overview of the class
 - ❑ Background
 - ❑ Logistics
- ❑ Lectures preview
- ❑ Office hours preview
- ❑ A simple interactive machine learning example

A Simple Interactive Machine Learning Example



A Neural Network Playground [Link](#)

DATA

Which dataset do you want to use?



Ratio of training to test data: 50%



Noise: 0



Batch size: 10



REGENERATE

Data:

(x,y) 2D Points

Binary Label

Train/Test Split

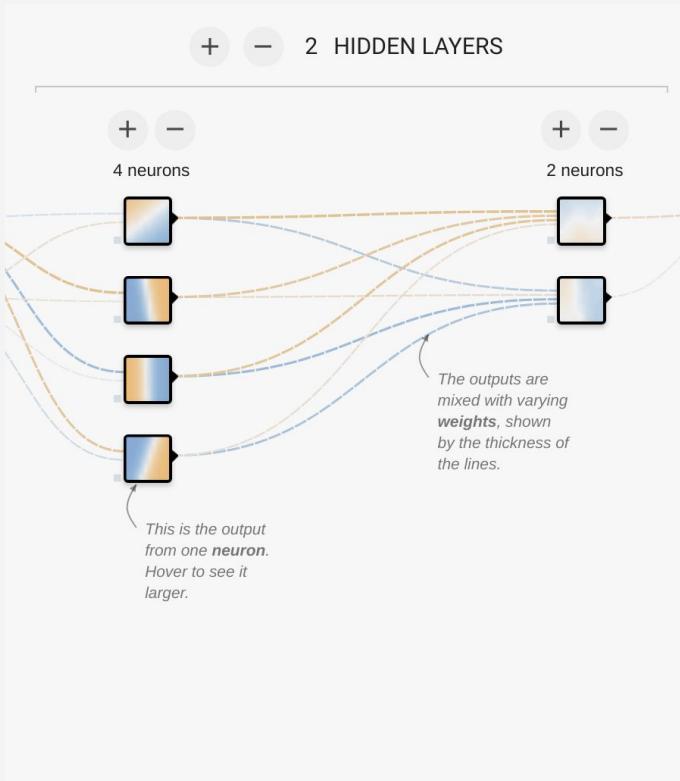
Noise Level

Batch Size



Feature Representation:

Learning problem becomes easier/harder with different feature representations, even with the same data!



A Learning Model:

Network Structure

Layers

Connectivity

No network works for all the problems!



Epoch	000,000	Learning rate	0.03	Activation	Tanh	Regularization	None	Regularization rate	0	Problem type	Classification
-------	---------	---------------	------	------------	------	----------------	------	---------------------	---	--------------	----------------

Training:

Train/Test Loss

Epochs

Optimization Algorithm

Evaluation:

Metric (accuracy, distance, ...)

Cross-validation

