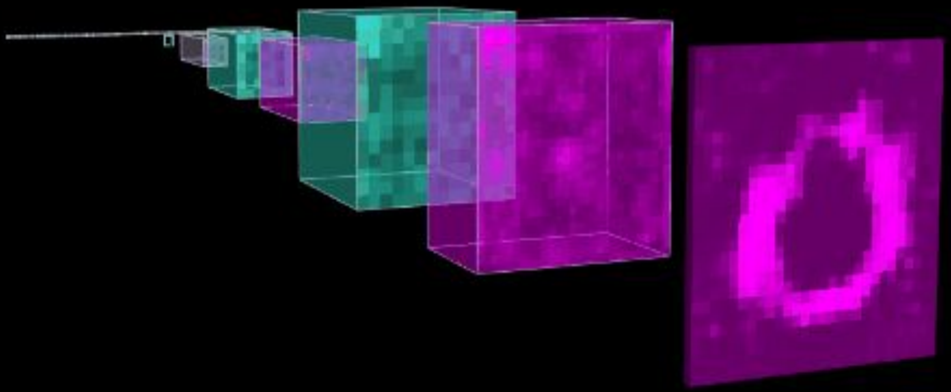
Data, Task and Algorithm Complexity in Deep Learning Projects

MACHINE
LEARNING
TOKYO



Find our slides and code on GitHub:
www.github.com/Machine-Learning-Tokyo/AI-SUM

Suzana Ilic, Dimitris Katsios | Nikkei's AI/SUM

Suzana Ilic

- Founder of MLT, a team of ML Engineers and Researchers and a community of 2500 members
- PhD research Linguistics/NLP: Deep Learning for Affective Computing (Conversational Agents)
- Past: Contract Google Japan, Machine Intelligence: NLU (Google Assistant); National Institute of Informatics, Deep Learning for NLP
- @suzatweet

Dimitris Katsios

- Machine Learning Engineer
- MSc in Intelligent Information Systems
- MSc in Systems Engineering
- MEng in Industrial Engineering
- Deep Learning workshops Instructor
- Currently working as ML Engineer at LPixel (AI on medical data)

MACHINE
LEARNING
TOKYO

2,500+
MEMBERS

50+
HANDS-ON
WORKSHOPS

250+
ML HOURS

MACHINE
LEARNING
TOKYO

Data, Task, Algorithm Complexity in Deep Learning Projects

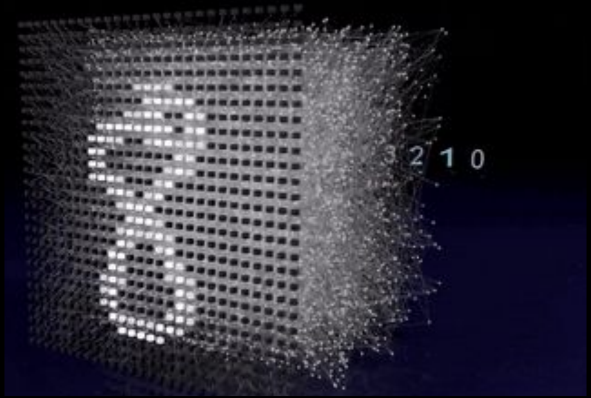
2-hour workshop (talk + code examples)

1. AI successes
2. Learning in Deep Networks
3. Evaluating Data, Task and Algorithm Complexity
4. Hands-on: Training models in different scenarios
5. Applications
6. Recap

Takeaways

- Clear picture of the pipeline when structuring DL projects
- Realistic expectations towards the capabilities and limitations of Deep Learning models
- Identify opportunities and dead ends in AI

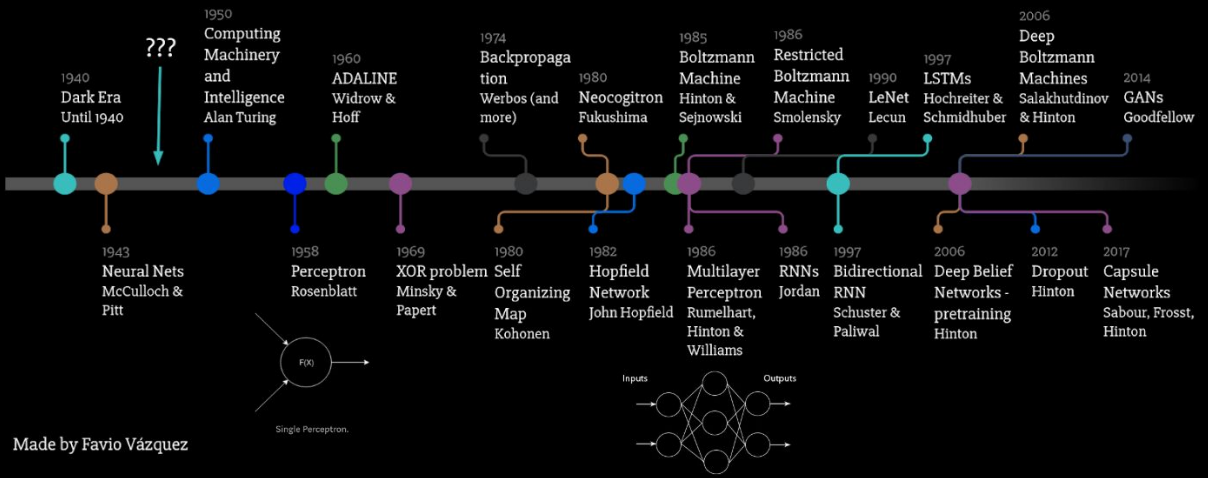
0. Deep Learning

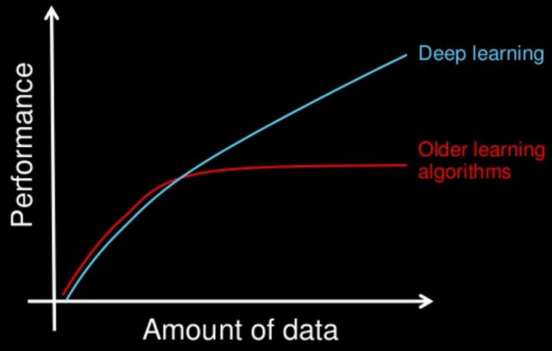


End-to-end learning from complex high-dimensional data: images, text, speech.

MACHINE LEARNING TOKYO

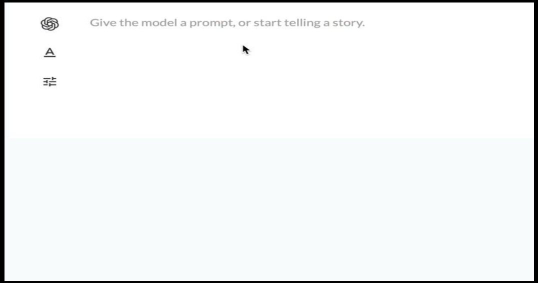
Algorithms have been around for decades.
Deep Learning successes based on access to Massive Datasets + Computing Power.

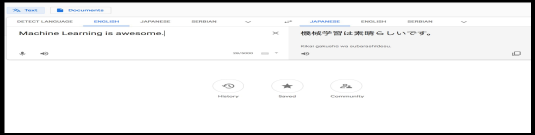


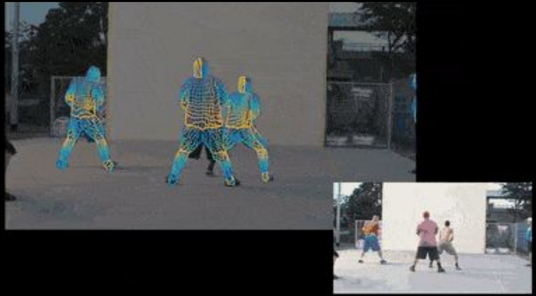


MACHINE
LEARNING
TOKYO

1. Some recent AI successes: Games, LM (GPT-2, BERT, ELMo), GANs, NMT, self-driving cars, chatbots/voice assistants, DensePose, Generative Art, ...

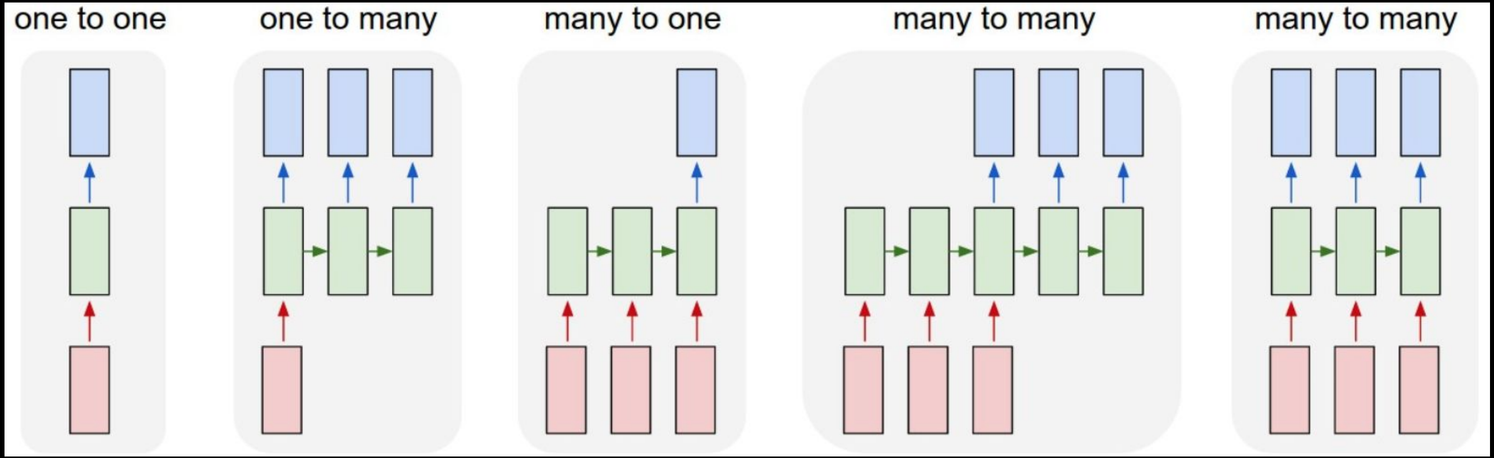






Supervised Learning: Mapping inputs to outputs
Images, video, text, speech:

Classification (binary, multi-class, multi-label), Detection (Recognition, Verification), Segmentation, Regression, Generation, ...

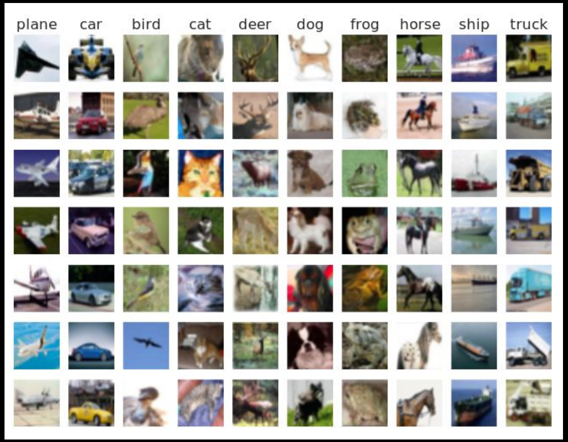


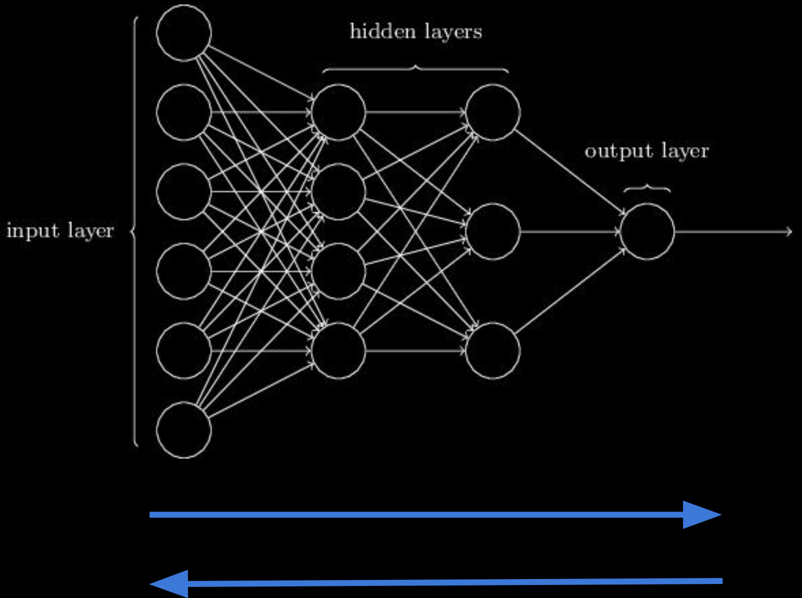
MACHINE
LEARNING
TOKYO

2. Learning in Deep Networks

Supervised Learning: Mapping input to output

Labels y
Training data x

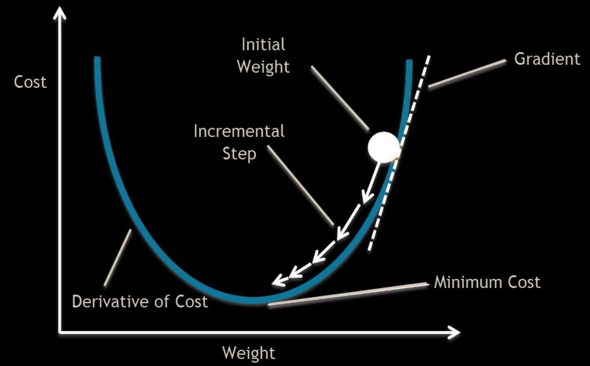


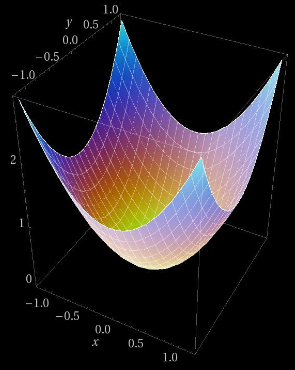


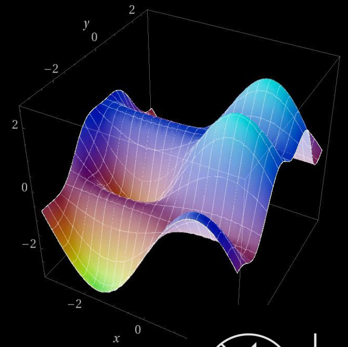
1. Learning in Deep Networks
Supervised Learning: Mapping input to output

Training loop (mini-batch SGD)

(1) **Draw a batch** of training samples *x* and the corresponding targets *y*
(2) **Train the network on *x*** (*forward pass*) to obtain the predictions *y_pred*
(3) **Compute the loss**, a measure of the mismatch between *y_pred* and y
(4) Compute the gradient of the loss with regard to the network's parameters (*backward pass:* ***backpropagation***)
The **optimizer** specifies how the gradient of the loss will be used to update the parameters
(5) **Move the parameters** (magnitude is defined by the **learning rate**): **update the weights** of the network to **reduce the loss**







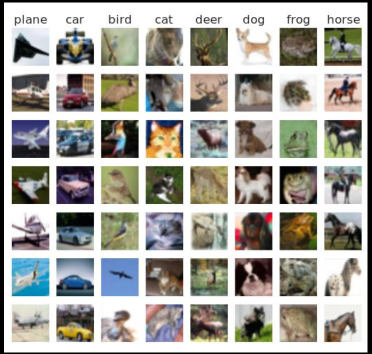
CATS & DOGS

2 classes

MNIST

10 classes
28 x 28
60,000 training images

CIFAR-10



10/100 classes
32 x 32
60,000 training images

IMAGENET

20,000 classes
Average: 469 × 387
14M training images

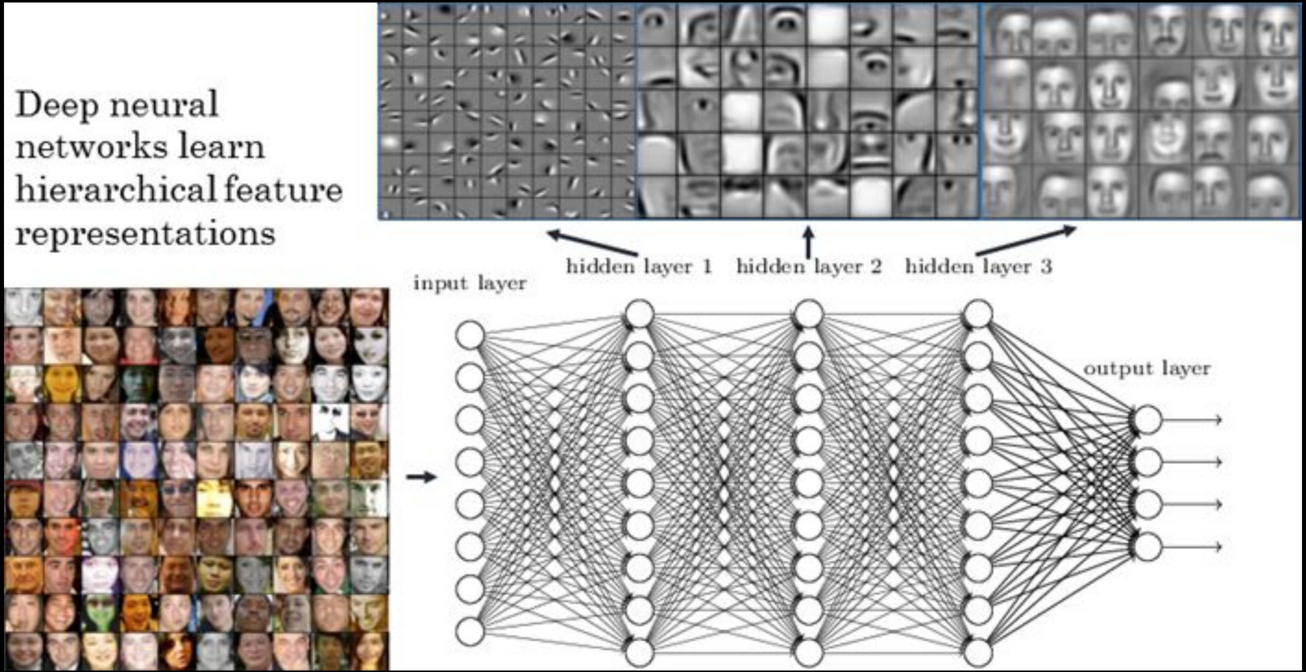
CELEBA

10,177 identities
178×218
200,000 training images

How difficult is the task? e.g. How many classes?
How distinct are the features of the classes compared?
How complex are the patterns within one image?
How nuanced is the color palette?
What's the resolution of the images?
How much irrelevant/distractive information is in the images?
...

This will help us understand:
How much data do we need?
What's our baseline model?

Hierarchical feature representations



MACHINE LEARNING TOKYO

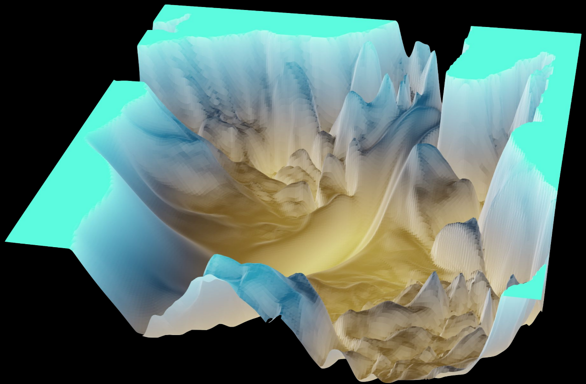
Data and task complexity determine

- Dataset size
- Network type, capacity and depth

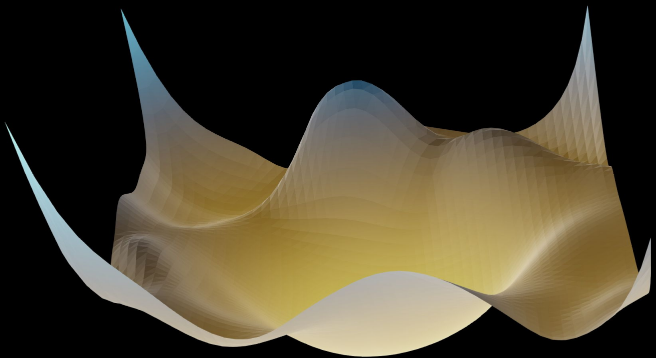
And subsequently

- Resources (compute, training time, ...)

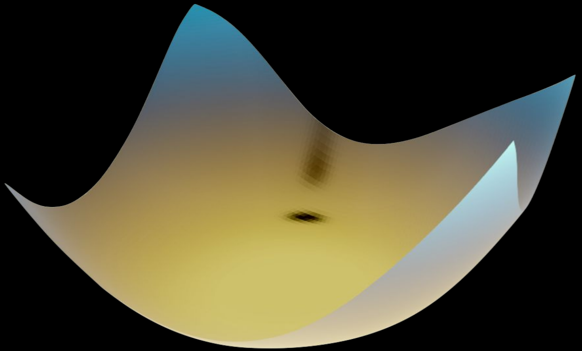
Network Complexity: 3D visualized loss landscapes (CIFAR-10)



VGG-56

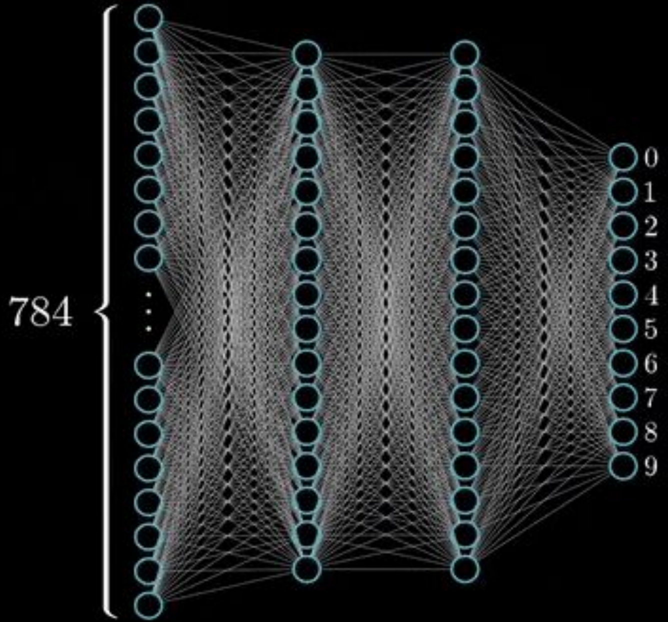


Renset-56

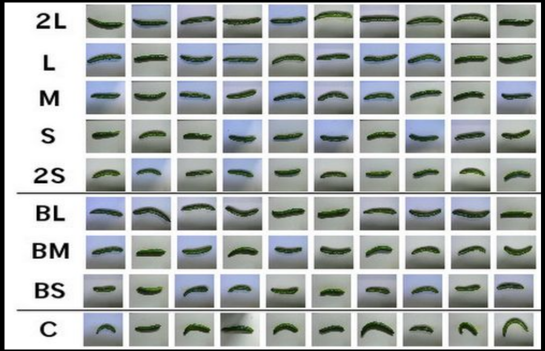


Densenet-121

4. Hands-on: Training a basic neural network



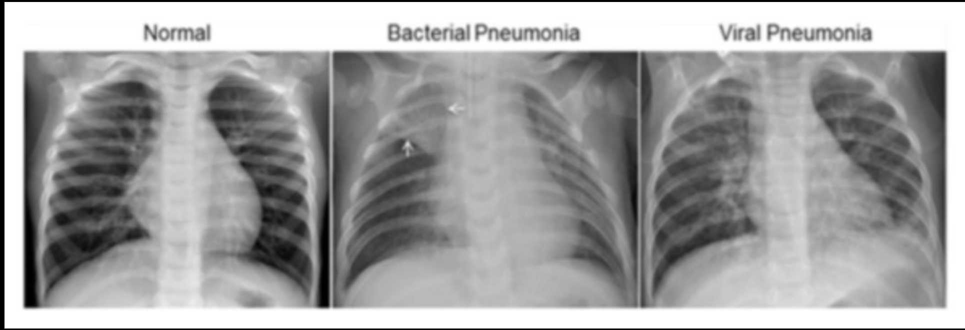
The Japanese Cucumber Farmer: A cucumber sorting system



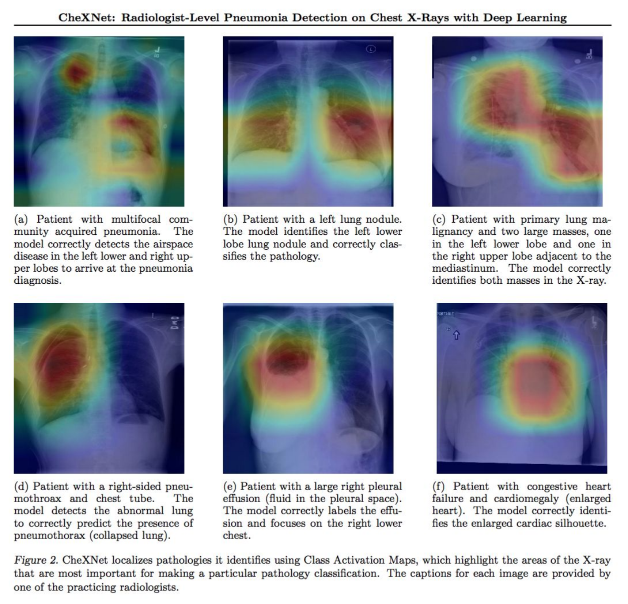
Automating manual effort of sorting cucumbers.

7,000 images. 9 classes based on shape, length and level of distortion.

Chest X-Ray Images



Kaggle: Chest X-Ray Images (Pneumonia)



CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning

Labeled data by Radiologists.

- Binary Classification
- Detection
- Segmentation

What about the severity of the task?

The cucumber farmer used accuracy as his performance metric. That’s good enough for the task. But when looking at chest X-Rays different metrics and standards should be set, such as sensitivity and specificity, especially to avoid false negatives.

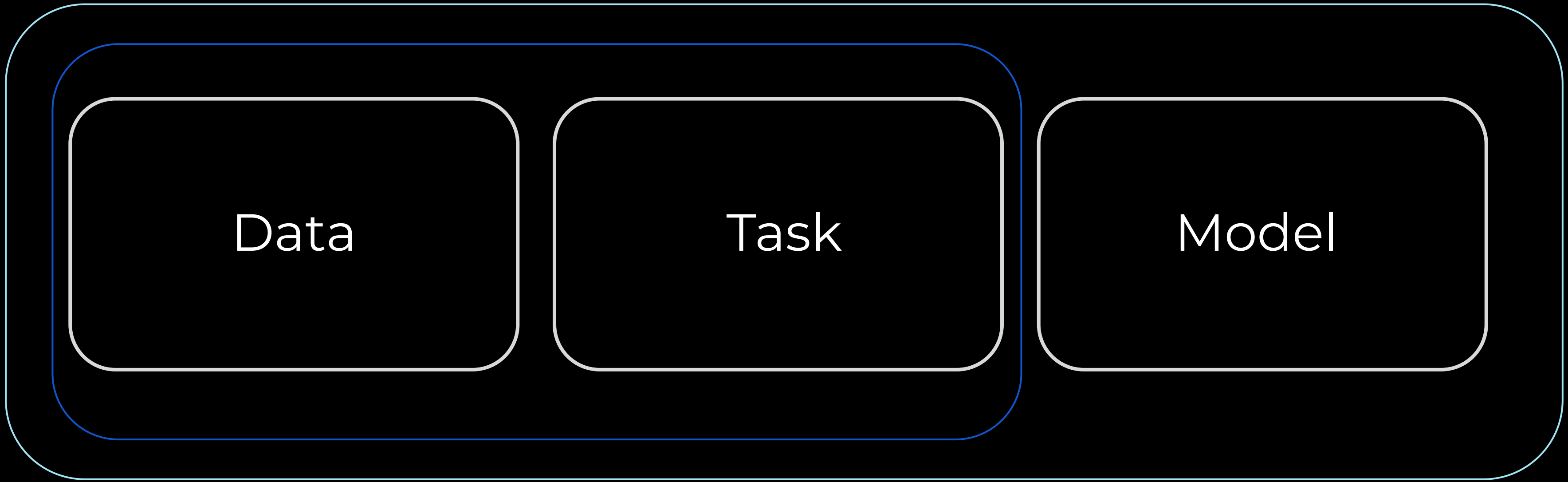

Data
Task
Model

Image sources

Title: https://tensorspace.org/
Loss landscapes: https://www.cs.umd.edu/~tomg/projects/landscapes/
https://arxiv.org/abs/1712.09913
MNIST 1: https://gfycat.com/thatdisloyalcaribou-perceptron-ann
MNIST 2: https://data.world/nrippner/mnist-handwritten-digits and https://www.3blue1brown.com/
Deep Learning Timeline: https://towardsdatascience.com/a-weird-introduction-to-deep-learning-7828803693b0
Gradient descent:
https://www.oreilly.com/library/view/learn-arcore-/9781788830409/e24a657a-a5c6-4ff2-b9ea-9418a7a5d24c.xhtml
GANs:
https://scoutmagazine.ca/2018/12/18/watch-this-artificial-intelligence-program-generate-people-that-dont-exist/
OpenAI FIve, GPT-2:
https://medium.com/syncedreview/humans-call-gg-openai-five-bots-beat-top-pros-og-in-dota-2-8508e59b8fd5
https://openai.com/blog/better-language-models/
Tesla: https://www.tesla.com/autopilot
Healthcare: https://www.ge.com/reports/software-please-doctors-looking-ai-speed-diagnosis/
Chatbots/Virtual Assistants: https://homeautotechs.com/siri-vs-google-vs-alexa-vs-cortana-which-one-is-best-voice-assistant/
Hierarchical Feature Extraction:
https://towardsdatascience.com/a-comprehensive-hands-on-guide-to-transfer-learning-with-real-world-applications-in-deep-learning-212bf3b2f27a
Brain: https://neurofantastic.com/brain/2017/3/27/is-brain-ageing-inevitable
Graph DL-Data: https://www.slideshare.net/ExtractConf
CelebA: https://www.theverge.com/2017/10/30/16569402/ai-generate-fake-faces-celebs-nvidia-gan
Object Detection: https://paperswithcode.com/task/real-time-object-detection
Chest X-Ray: https://stanfordmlgroup.github.io/projects/chexnet/
Note: Colors are inverted compared to the original to match the presentation background.

References and good reads

BLOG POSTS
GPT-2 https://openai.com/blog/better-language-models/
AlphaZero: https://deepmind.com/blog/alphazero-shedding-new-light-grand-games-chess-shogi-and-go/
Japanese Cucumber Farmer:
https://cloud.google.com/blog/products/gcp/how-a-japanese-cucumber-farmer-is-using-deep-learning-and-tensorflow
The unreasonable Effectiveness of RNNs: http://karpathy.github.io/2015/05/21/rnn-effectiveness/

PAPERS
AlphaZero: https://science.sciencemag.org/content/362/6419/1140
DensePose: https://arxiv.org/abs/1802.00434
GPT-2: https://d4mucfpksywv.cloudfront.net/better-language-models/language-models.pdf
BERT: https://arxiv.org/abs/1810.04805
Tesla Autopilot: https://ieeexplore.ieee.org/document/8122757
https://hcai.mit.edu/tesla-autopilot-human-side.pdf
Chexnet: https://arxiv.org/abs/1711.05225

DATASETS
Kaggle Chest X-Ray Images: https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia
MNIST: http://yann.lecun.com/exdb/mnist/index.html
CIFAR-10: https://www.cs.toronto.edu/~kriz/cifar.html
ImageNet: http://www.image-net.org/
CelebA: http://mmlab.ie.cuhk.edu.hk/projects/CelebA.html