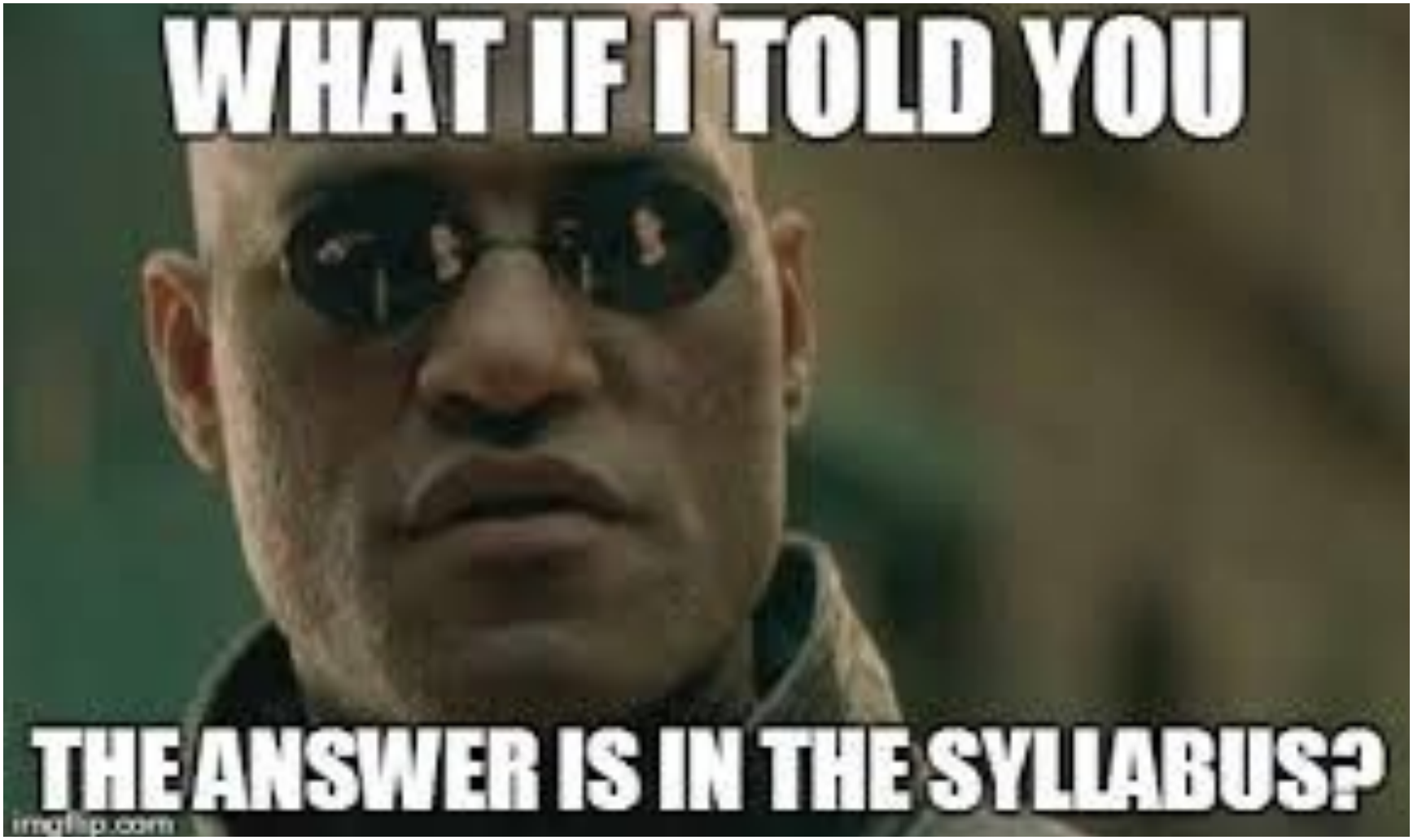
INTRODUCTION

Pattern Recognition

Syllabus



Registration

- Graduate students
  - 12 slots, sec 2
  - If filled, register as V/W only
- For undergrads, sec 22

Tools

- Python
- Python
- Python
- Jupyter
- Numpy
- Pandas
- Keras

Plagiarism Policy

- You shall not show other people your code or solution
- Copying will result in a score of zero for both parties on the assignment
- Many of these algorithms have code available on the internet, do not copy paste the codes

Plagiarism vs. Cheating

What is the difference?

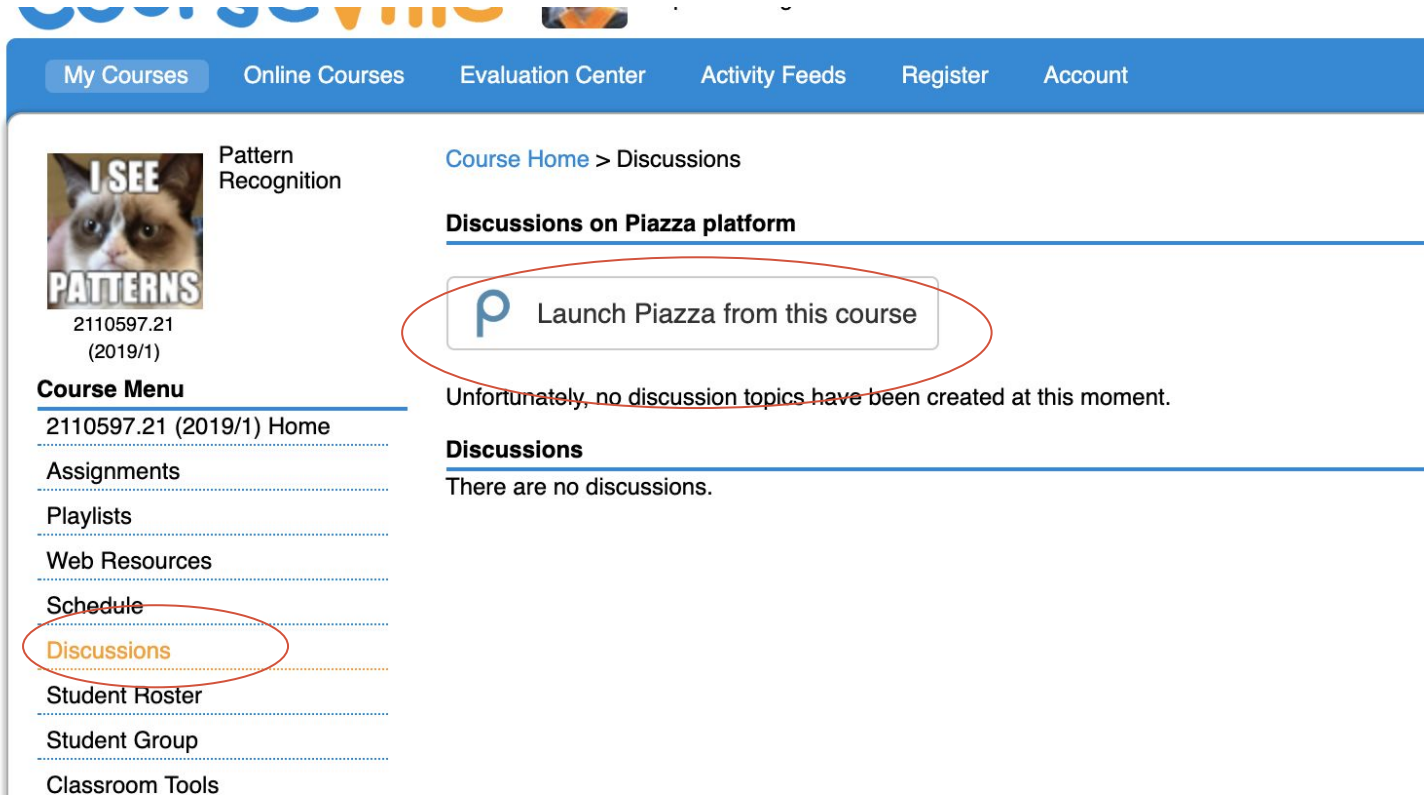
Courseville & Github

- 2110597.21 (2019/1)

Password: cattern

- https://github.com/ekapolc/pattern_2019

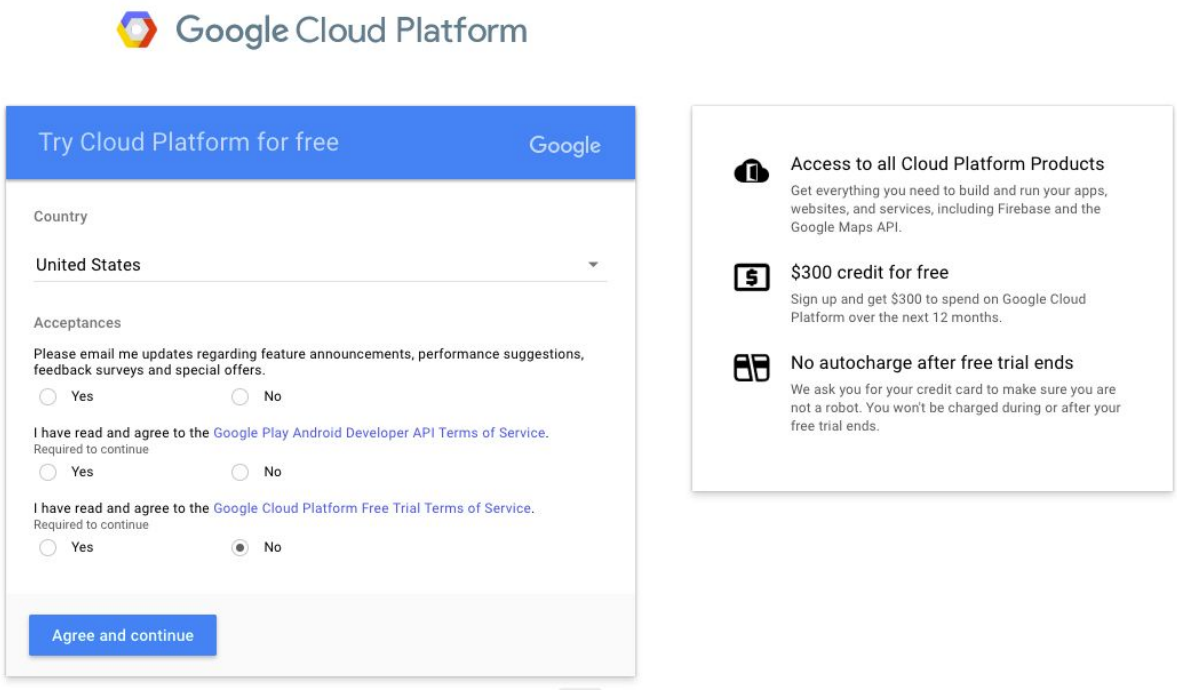
Piazza



- Participation score (5 points) comes from piazza

Cloud

- Gcloud
- Credit card



Course project

- 5 people
- Topic of your choice
  - Can be implementing a paper
  - Extension of a homework
  - Project for other courses with an additional machine learning component
  - Your current research (with additional scope)
  - Or work on a new application
  - Must already have existing data! No data collection!
- Topics need to be pre-approved
  - Details about the procedure TBA

The machine learning trend 2015

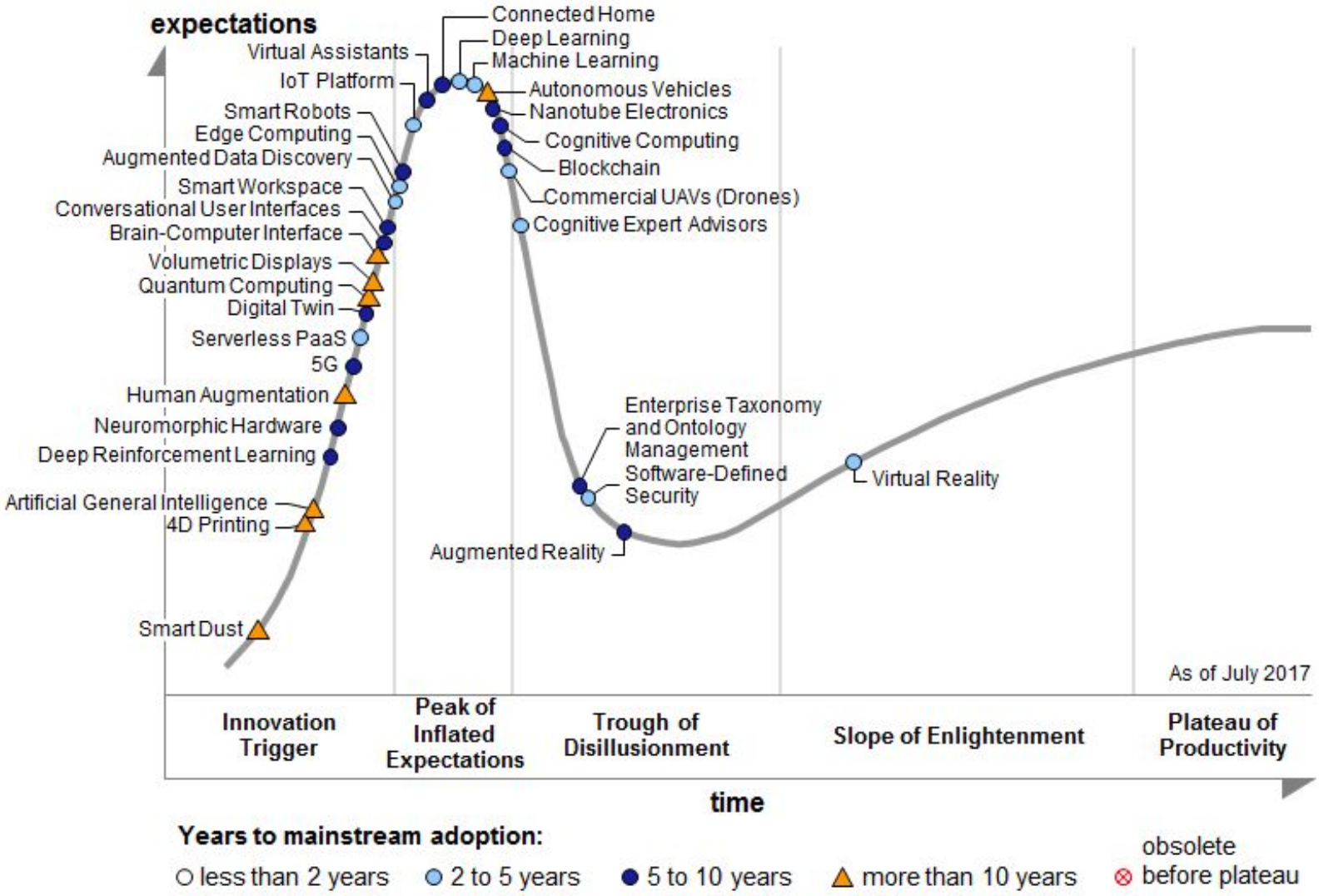
http://www.gartner.com/newsroom/id/3114217

The machine learning trend 2016

Source: Gartner (July 2016)

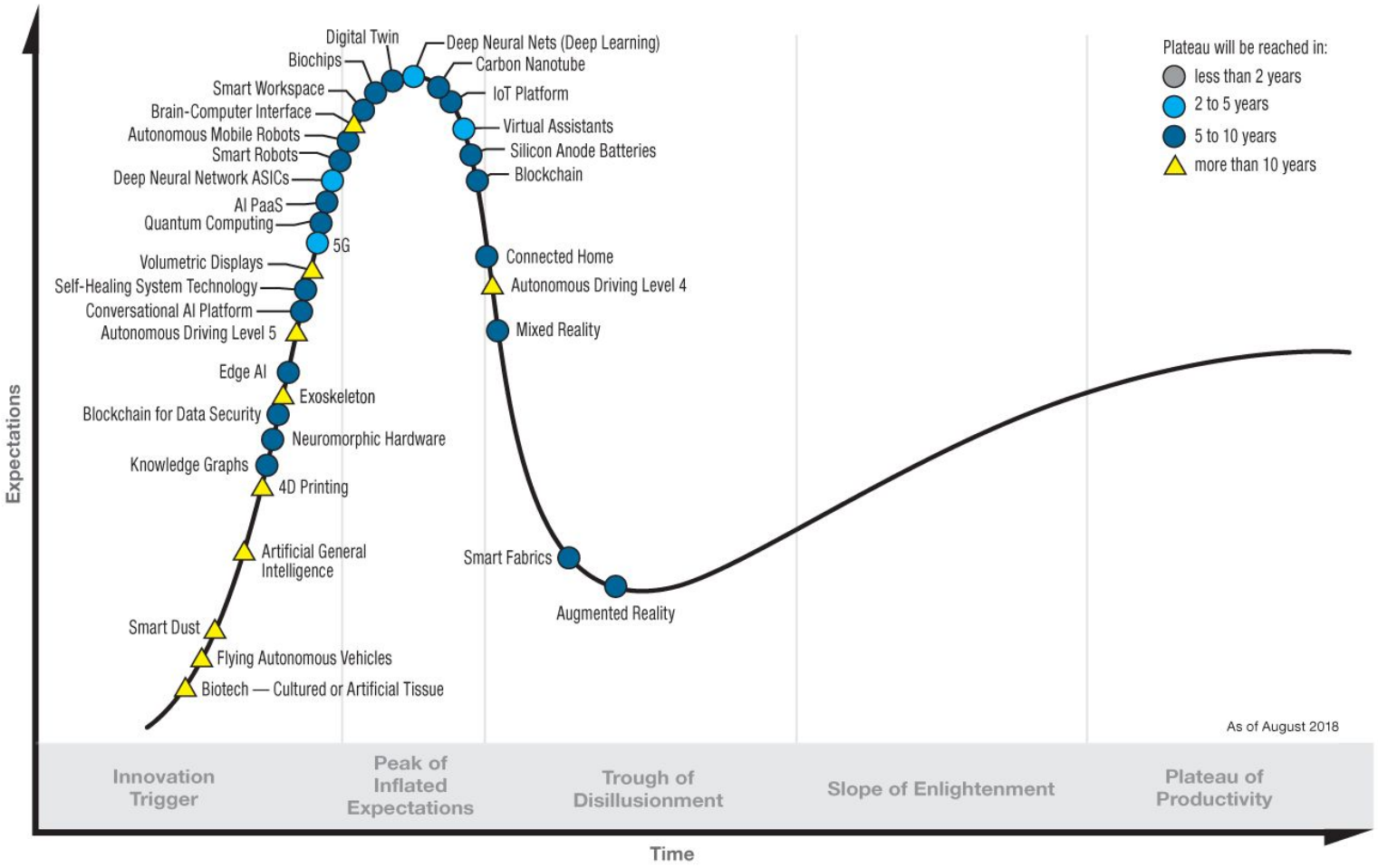
http://www.gartner.com/newsroom/id/3412017

The machine learning trend 2017



http://www.gartner.com/smarterwithgartner/top-trends-in-the-gartner-hype-cycle-for-emerging-technologies-2017/

The machine learning trend 2018



https://www.gartner.com/smarterwithgartner/5-trends-emerge-in-gartner-hype-cycle-for-emerging-technologies-2018/

MACHINE INTELLIGENCE 3.0

ENTERPRISE INTELLIGENCE

VISUAL
Orbital Insight, planet, clarifai, DEEPVISION, cortica, sigocion, SPACE_KNOW, Captricity, netra, deepomatic

AUDIO
Gridspace, TalkIQ, nexidia, twilio, CAPIO, Expect Labs, Clover, Mobvoi, Qurious.AI, popup archive

SENSOR
PREDIX, GE IoT, MAANA, Sentenai, PLANET OS, UPTAKE, IMUBIT, Preferred Networks, thingworx, KONUX, Alluvium

INTERNAL DATA
PRIMER, IBM WATSON, Cycorp, Palantir, ARIMO, Alation, sapho, Outlier, Digital Reasoning

MARKET
mattermark, Quid, DataFox, PREMISE, Bottlenose, MOTIVA, enigma, CB INSIGHTS, Tracxn, predata

ENTERPRISE FUNCTIONS

CUSTOMER SUPPORT
DigitalGenius, Kasisto, ELOQUENT, Wise.io, ACTIONIQ, zendesk, Preact, CLARABRIDGE

SALES
collective[i], 6sense, fuse|machines, AVISO, salesforce, INSIDE SALES.COM, clari, Zensight

MARKETING
MINTIGO, Lattice, RADIUS, Liftigniter, PERSADO, brightfunnel, retention science, COGNICOR, AIRPR, msg.ai

SECURITY
CYLANCE, DARKTRACE, ZIMPERIUM, deepinstinct, Sentinel, DEMISTO, graphistry, drawbridge, SignalSense, AppZen

RECRUITING
textio, entelo, Wade & Wendy, hi, unitive, SpringRole, GIGSTER, HireVue

AUTONOMOUS SYSTEMS

GROUND NAVIGATION
drive.ai, AdasWorks, ZOOX, MOBILEYE, UBER, Google, TESLA, nuTonomy, Auro Robotics

AERIAL
SKYDIO, SHIELD AI, Airware, dji, LILY, DroneDeploy, pilot.ai, SKYCATCH

INDUSTRIAL
JAYBRIDGE, OSARO, CLEARPATH, fetch, KINDRED, HARVEST, rethink robotics

AGENTS

PERSONAL
amazon alexa, Cortana, Allo, facebook, Siri, Replika

PROFESSIONAL
butter.ai, pogo, SKIPFLAG, clara, x.ai, slack, talla, Zoom, sudo

INDUSTRIES

AGRICULTURE
BLUE RIVER, mavrx, tule, TRACE GENOMICS, Pivot Bio, Terravion, AGRI-DATA, Descartes Labs, ag analytics

EDUCATION
KNEWTON, volley, gradescope, CTI, coursera, UDACITY, altschool

INVESTMENT
Bloomberg, sentient, SENTIUM, KENSHO, alphasense, Dataminr, CEREBELLUM CAPITAL, Quandl

LEGAL
blueJ, BEAGLE, Everlaw, RAVEL, seal, ROSS, LEGAL ROBOT

LOGISTICS
NAUTO, Acerta, PRETECKT, Routific, clearmetal, MARBLE, PITSTOP

INDUSTRIES CONT'D

MATERIALS
zymergen, Citrine, Eigen Innovations, SIGHT MACHINE, GINKGO BIOWORKS, nanotronics, CALCULARIO

RETAIL FINANCE
TALA, zestfinance, Lendo, earnest, affirm, MIRADOR, wealthfront, Betterment

HEALTHCARE

PATIENT
PULSE, CareSkore, ZEPHYR HEALTH, IBM Watson Health, Onc24, SENTRIAN, Atomwise, Numerate

IMAGE
BUTTERFLY, 3SCAN, ARTERYS, enlitic, BAYLABS, imagia, Google DeepMind

BIOLOGICAL
iCarbonX, color, GRAIL, deep genomics, RECURSION, LUMINIST, Numerate, Atomwise, verily, WHOLE BIOME

TECHNOLOGY STACK

AGENT ENABLERS
OCTANE.AI, howdy, Maluuba, KITT.AI, OpenAI Gym, Kasisto, AUTOMAT, semanticmachines

DATA SCIENCE
DOMINO, SPARKBEYOND, rapidminer, kaggle, DataRobot, yhat, AYASDI, data iku, seldon, yseop, bigml

MACHINE LEARNING
CognitiveScale, GoogleML, context relevant, Cycorp, HyperScience, narralogics, minds.ai, H2O.ai, SCALED INFERENCE, sparkcognition, Loop, GEOMETRIC INTELLIGENCE, deepsense.io, reactive, skymind, bonsai

NATURAL LANGUAGE
agolo, AYLIEN, LEXALYTICS, Narrative Science, spaCy, LUMINOSO, cortical.io, MonkeyLearn

DEVELOPMENT
SIGOPT, HyperOpt, fuzzyio, pkite, rainforest, lobe, Anodot, Signifai, LAYER 6, bonsai

DATA CAPTURE
CrowdFlower, diffbot, CrowdAI, import.io, Paxata, DATASIFT, amazon mechanical turk, enigma, WorkFusion, DATALOGUE, TRIFACTA, parsehub

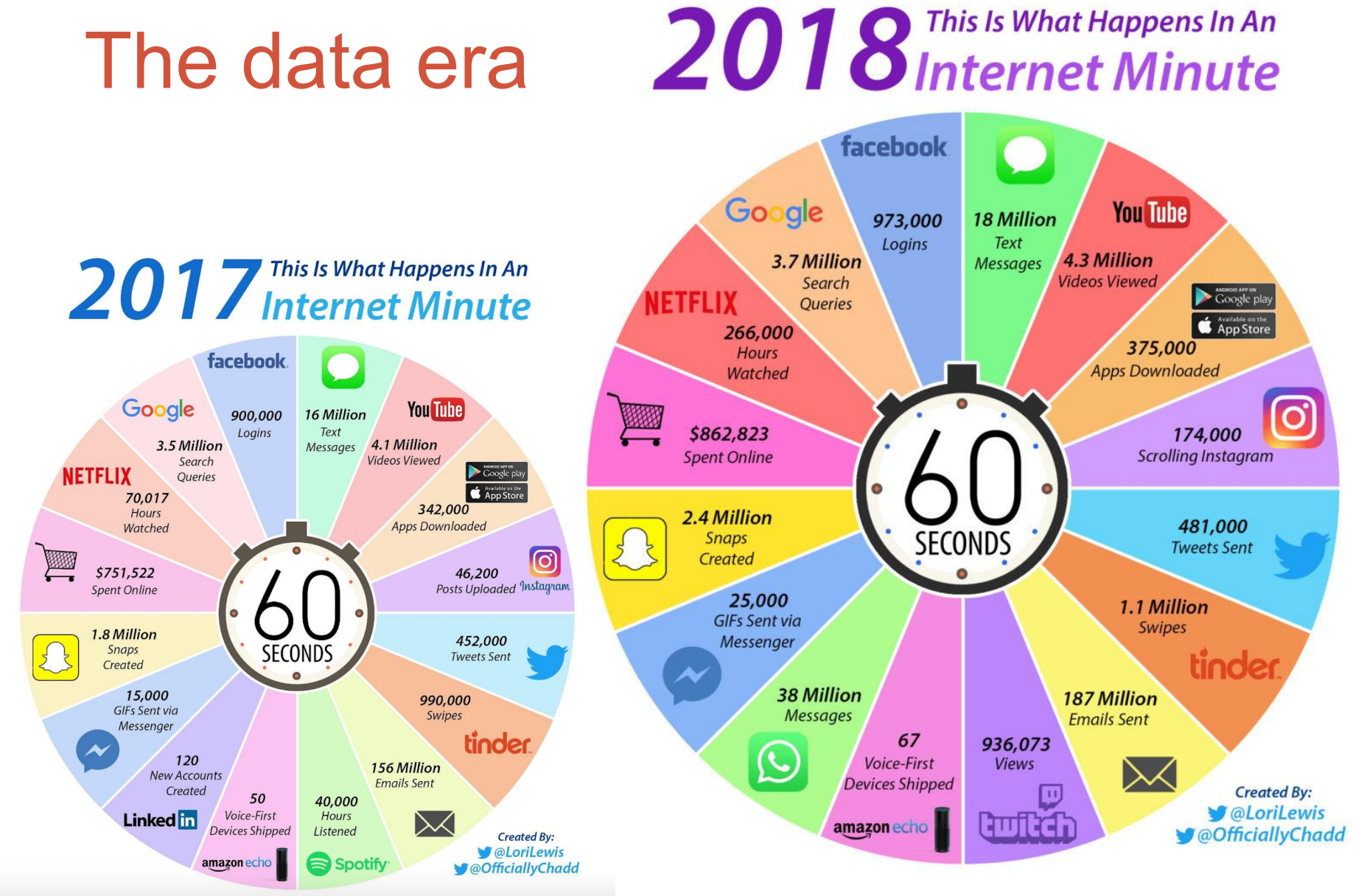
OPEN SOURCE LIBRARIES
Keras, Chainer, CNTK, TensorFlow, Caffe, H20, DEEPLEARNING4J, theano, torch, DSSTNE, Scikit-learn, AzureML, neon, MXNet, DMTK, Spark, PaddlePaddle, WEKA

HARDWARE
KNUPATH, TENSTORRENT, Cirrascale, NVIDIA, intel nervana, Movidius, tensilica, GoogleTPU, 10^26 Labs, Qualcomm, Cerebras, Isosemi

RESEARCH
OpenAI, maluuba, ELEMENT AI, vicarious, KNOGGIN, Numenta, Kimera Systems, Cogitai

shivonzilis.com/MACHINEINTELLIGENCE · Bloomberg BETA

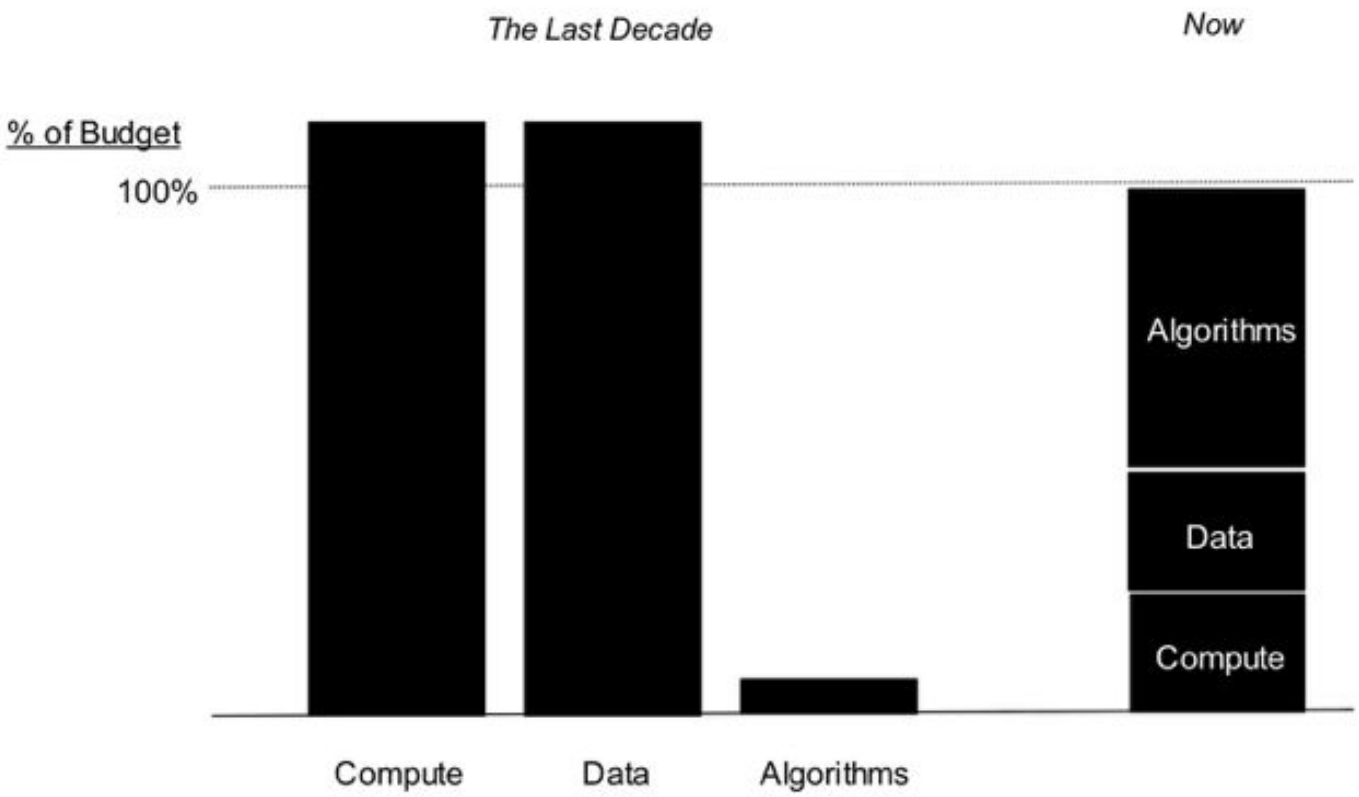
The data era



http://www.visualcapitalist.com/internet-minute-2018/

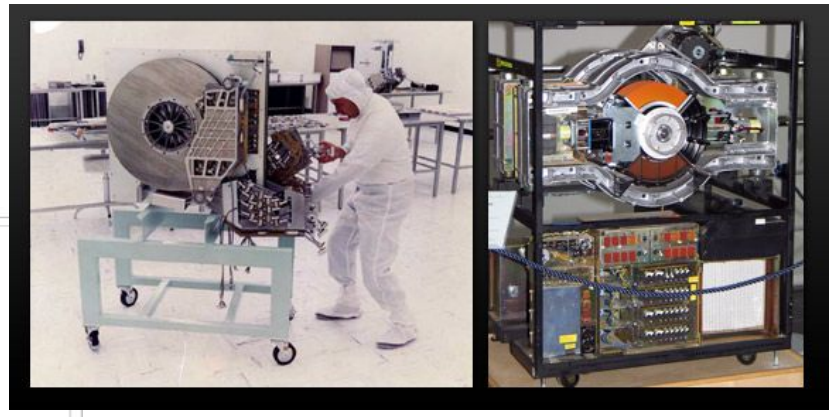
Factors for ML

- Data
- Compute
- Algo



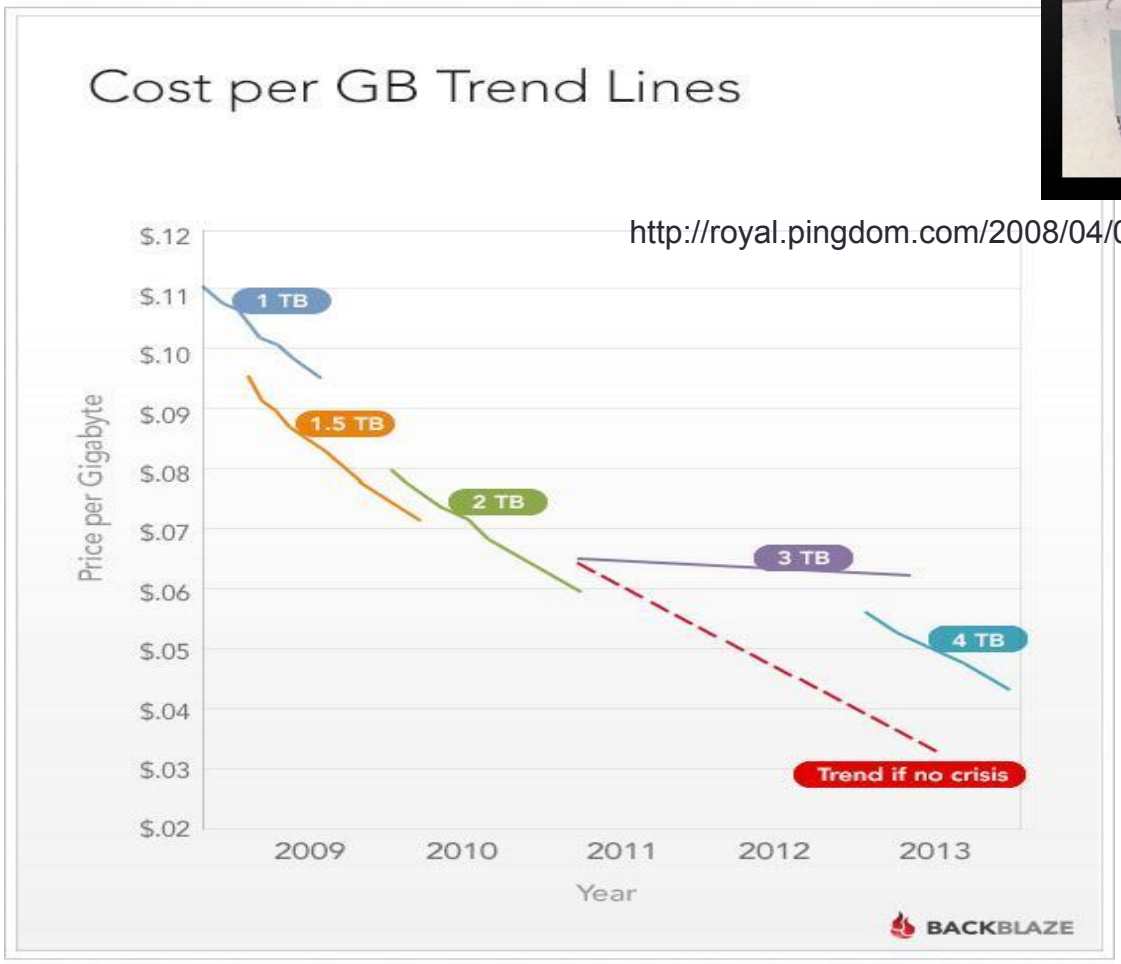
http://www.kdnuggets.com/2017/06/practical-guide-machine-learning-understand-differentiate-apply.html

The cost of storage



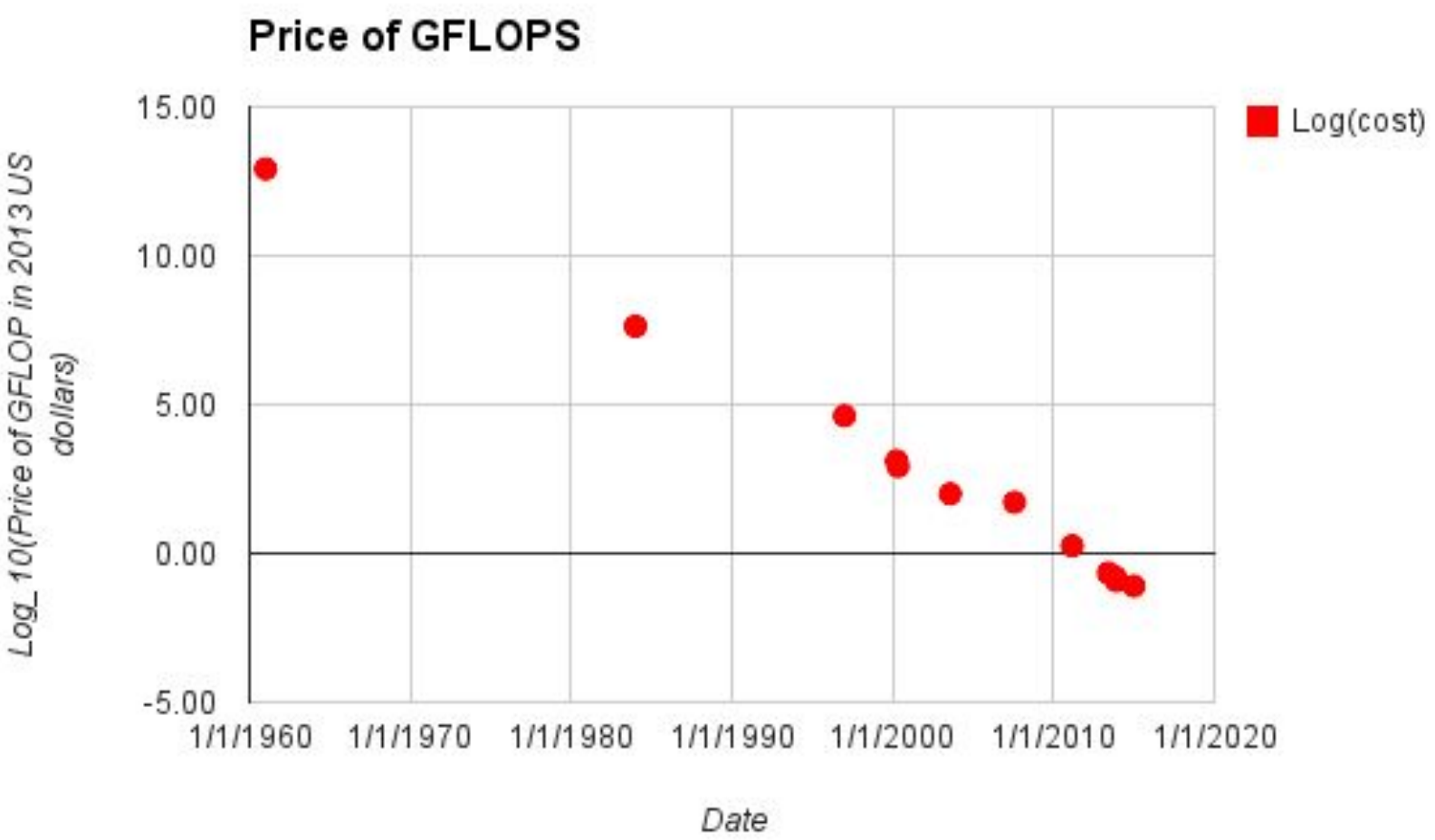
http://royal.pingdom.com/2008/04/08/the-history-of-computer-data-storage-in-pictures/

1980 250MB hard disk drive
250 kg 100k USD (300k USD in today's dollar)



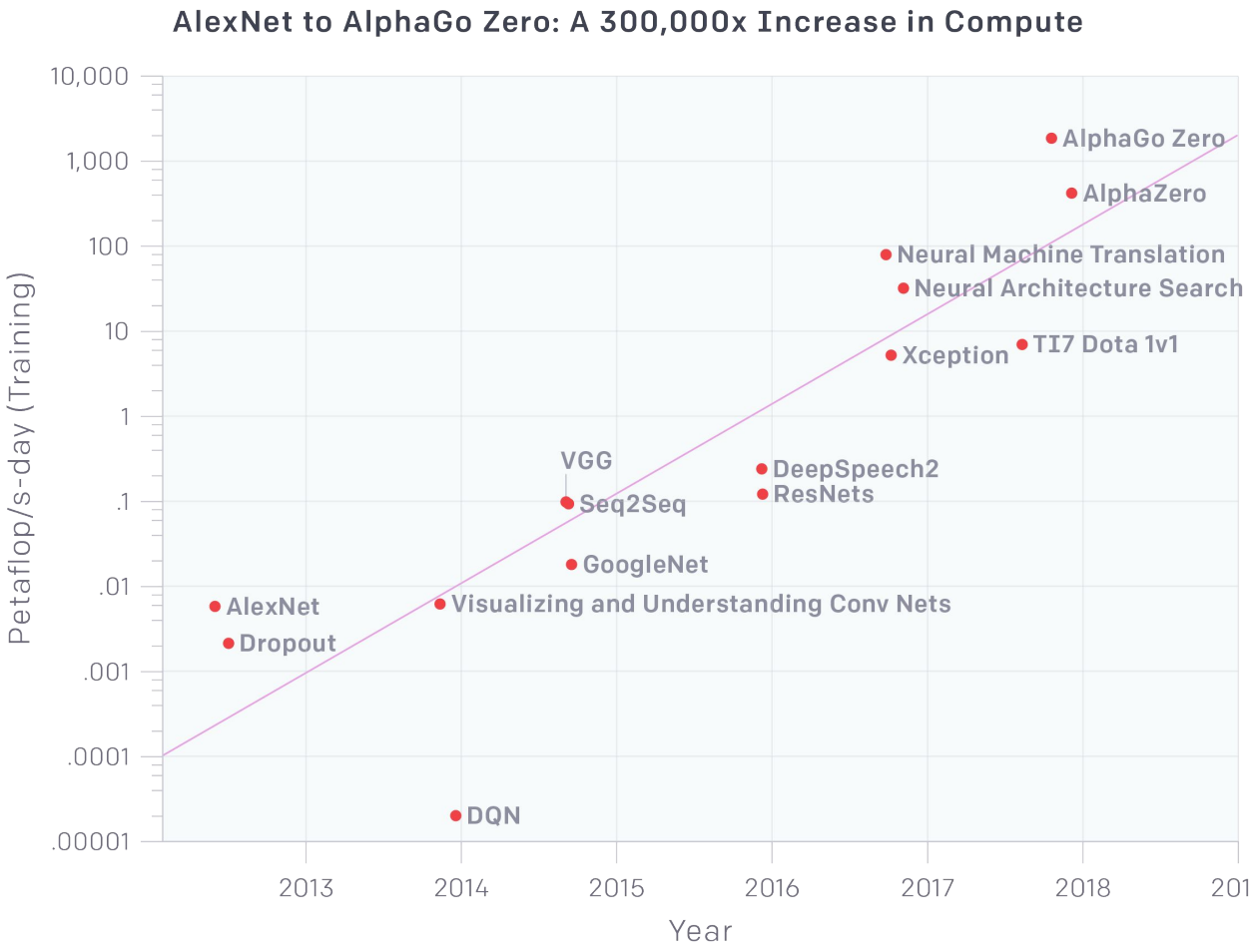
https://www.backblaze.com/blog/farming-hard-drives-2-years-and-1m-later/

The cost of compute



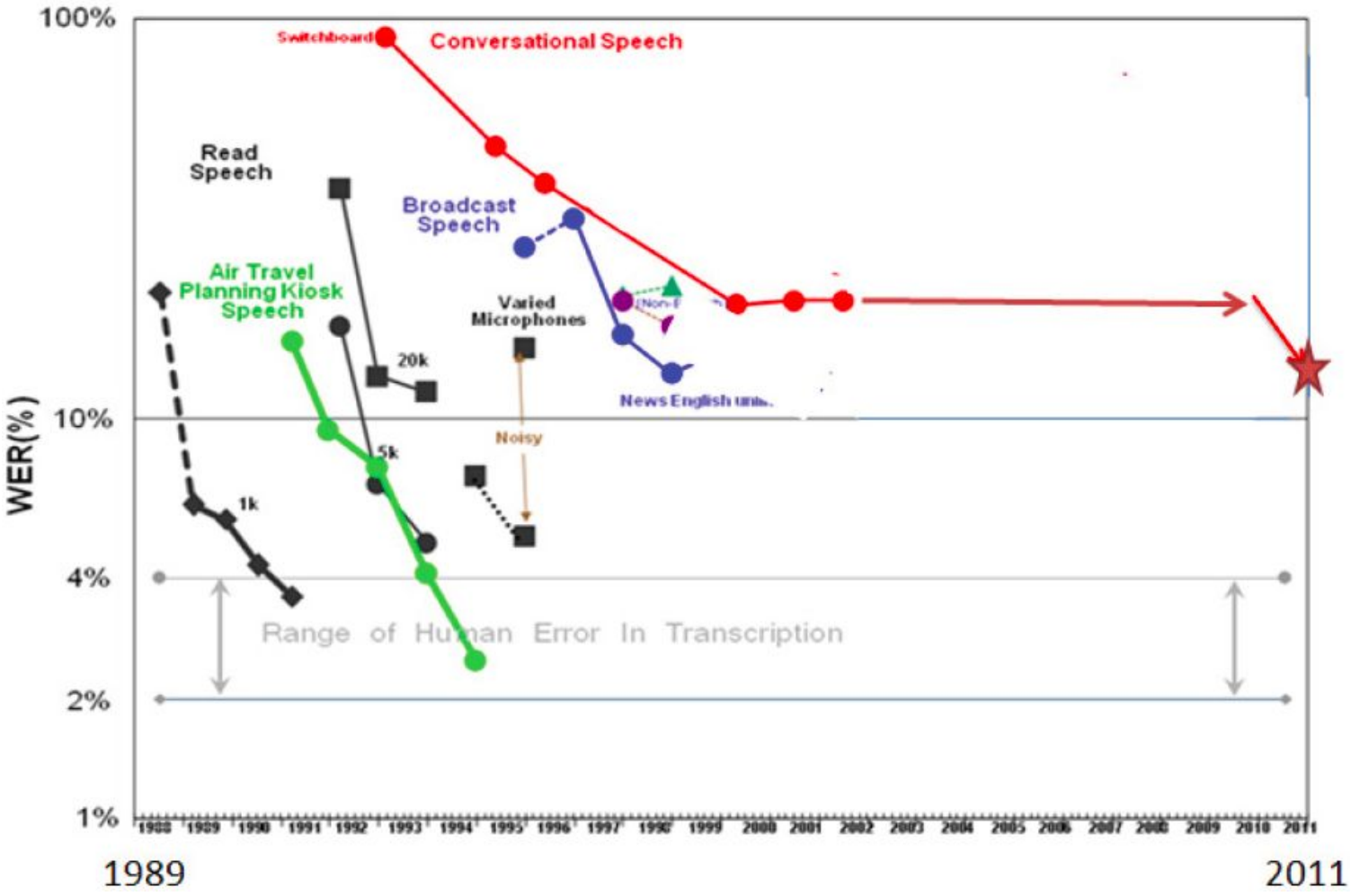
http://aiimpacts.org/trends-in-the-cost-of-computing/

Deep learning and compute



https://blog.openai.com/ai-and-compute/

Hitting the sweet spot on performance



http://recognize-speech.com/acoustic-model/knn/benchmarks-comparison-of-different-architectures

Hitting the sweet spot in performance

PRESS RELEASE
OCTOBER 4, 2011

Apple Launches iPhone 4S, iOS 5 & iCloud

iPhone 4S Features Dual-Core A5 Chip, All New Camera, Full 1080p HD Video Recording & Introduces Siri

CUPERTINO, California—October 4, 2011—Apple® today announced iPhone® 4S, the most amazing iPhone yet, packed with incredible new features including Apple's dual-core A5 chip for blazing fast performance and stunning graphics; an all new camera with advanced optics; full 1080p HD resolution video recording; and Siri™, an intelligent assistant that helps you get things done just by asking. With the launch of iPhone 4S

Now time for videos

https://www.youtube.com/watch?v=wiOopO9jTZw

2017

Now time for videos

https://blog.openai.com/openai-five/
https://youtu.be/eHipy_j29Xw

2018

Now time for videos

https://deepmind.com/blog/article/alphastar-mastering-real-time-strategy-game-starcraft-ii

Now time for videos

https://ai.facebook.com/blog/pluribus-first-ai-to-beat-pros-in-6-player-poker/

https://www.youtube.com/watch?v=u90TbxK7VEA

- “If I were to guess like what our biggest existential threat is, it’s probably that. So we need to be very careful with the artificial intelligence. There should be some regulatory oversight maybe at the national and international level, just to make sure that we don’t do something very foolish.”

- “I think people who are naysayers and try to drum up these doomsday scenarios — I just, I don’t understand it. It’s really negative and in some ways I actually think it is pretty irresponsible”

Darren Cunningham @dcunni · 6h

Zuckerberg blasts @elonmusk warnings against artificial intelligence as 'pretty irresponsible' bizjournals.com/sanjose/news/2… @svbizjournal #ai

Facebook CEO Mark Zuckerberg blasts Tesla CEO Elon Musk's warn...

"People who are naysayers and try to drum up these doomsday scenarios — I just, I don't understand it," the Facebook CEO said. "It's really negative

bizjournals.com

30 296 566

Elon Musk ✓

@elonmusk

Following

Replying to @dcunni @SVbizjournal

I've talked to Mark about this. His understanding of the subject is limited.

8:07 AM - 25 Jul 2017

Poll

What is Pattern Recognition?

- "Pattern recognition is a branch of machine learning that focuses on the recognition of patterns and regularities in data, although it is in some cases considered to be nearly synonymous with machine learning."

wikipedia

- What about
  - Data mining
  - Knowledge Discovery in Databases (KDD)
  - Statistics
  - Data science

ML vs PR vs DM vs KDD

- “The short answer is: None. They are … concerned with the same question: how do we learn from data?”

  Larry Wasserman – CMU Professor

- Nearly identical tools and subject matter

History

- Pattern Recognition started from the engineering community (mainly Electrical Engineering and Computer Vision)
- Machine learning comes out of AI and mostly considered a Computer Science subject
- Data mining starts from the database community

Different community viewpoints

- A screw looking for a screw driver
- A screw driver looking for a screw

Different applications

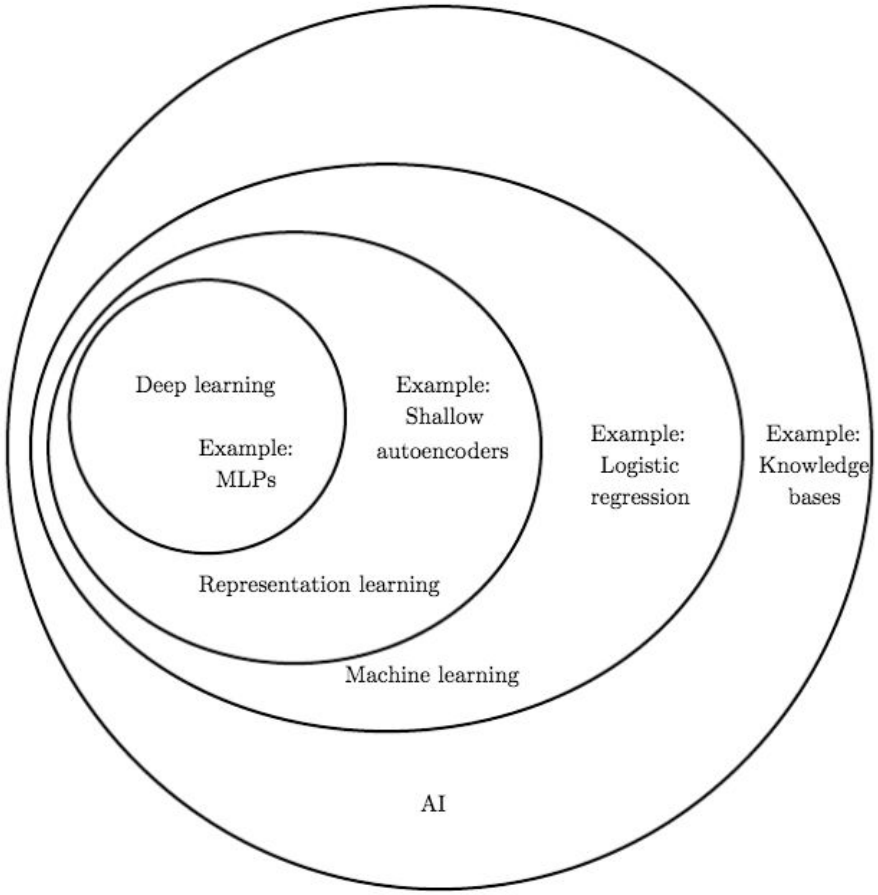
Different tools

The Screwdriver and the Screw

DM PR ML AI

Distinguishing things

- DM – Data warehouse, ETL
- AI – Artificial General Intelligence
- PR – Signal processing (feature engineering)



http://www.deeplearningbook.org/

Different terminologies

http://statweb.stanford.edu/~tibs/stat315a/glossary.pdf

Machine learning	Statistics
network, graphs	model
weights	parameters
learning	fitting
generalization	test set performance
supervised learning	regression/classification
unsupervised learning	density estimation, clustering
large grant = $1,000,000	large grant= $50,000

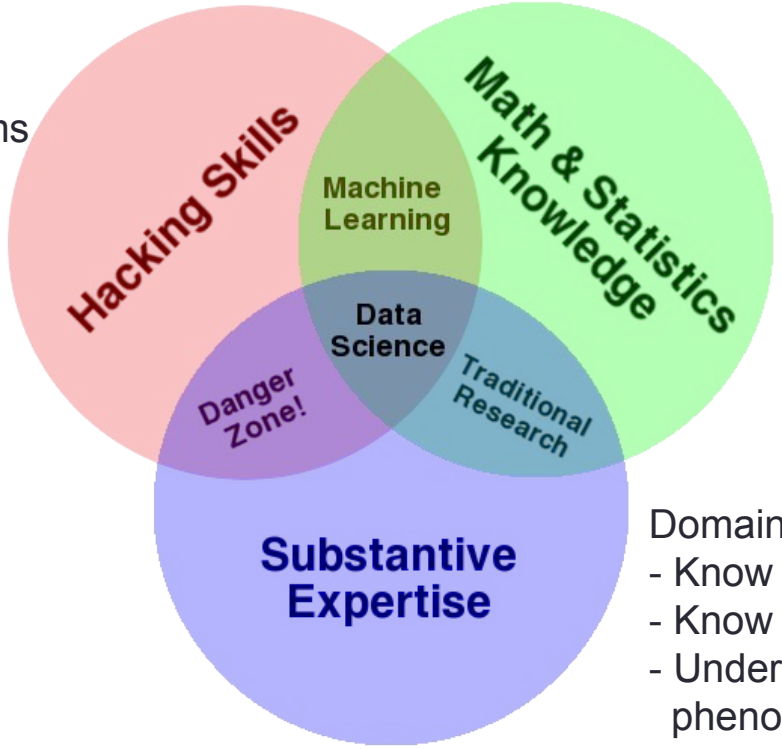
Merging communities and fields

- With the advent of Deep learning the fields are merging and the differences are becoming unclear

Data science

- How to get value from data
  - Data-driven decision making

Programming
- Algorithms
- Distributed systems
- Database
- Hardware

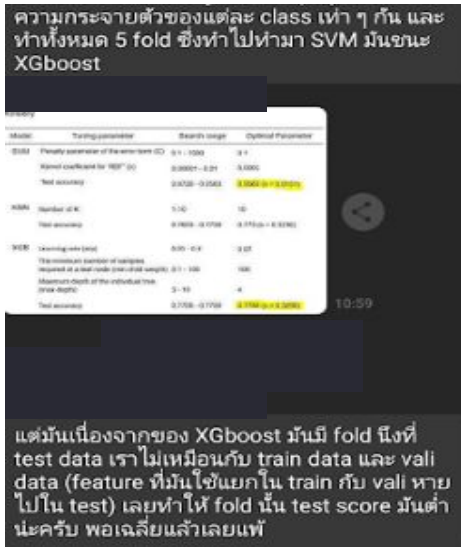


Math knowledge
- Statistics
- Probability
- Optimization
- Linear algebra

Domain knowledge
- Know the right questions to ask
- Know the relevant information
- Understand the underlying phenomena

http://drewconway.com/zia/2013/3/26/the-data-science-venn-diagr

The danger zone



อาจารย์ครับ ผมมีประโยคตัดเป็นคำ ใส่เป็น word bag ที่มีขนาดเท่ากับ vocab
ใส่เป็นความถี่ของคำในประโยค
แล้วผ่าน TF-IDF แล้วเข้า Multinomial NB มันพังกว่าไม่ผ่าน TF-IDF ครับ ผมสงสัยว่าเป็นเพราะอะไร ผมรู้แค่ว่าค่า prob มันเปลี่ยนจริงๆ มันน่าจะดี เพราะผมทำ TF-IDF บน SVM แล้วผลมันดีครับ

ผมใช้ laplace ด้วยครับ แต่ผลไม่น่าพังขนาดนั้น alpha=1 ครับ

อาจารย์ครับ ควรแบ่ง data ไงดีครับตอนทำ neural net
ผมมี data อยู่ของช่วงวันที่ 9-16 ครับ
คือ ตอนที่ทำ linear กับ pca ผมใช้ train เป็นช่วงวันที่ 9 - 13 ครับ ส่วน test ผมใช้ช่วง 14-15
ตอนทำ neural net ถ้าผมใช้เป็นวันที่ 16 วันเดียวพอไหมครับ หรือควรแบ่ง data ใหม่ครับ
ตอนนี้ training set ผมมีประมาณ 360000 ครับ ส่วน test set มีประมาณ 150000 ครับ

แล้วจึงนำภาพ3D 11440 ไป Train เข้า CNN แล้ว Classify ว่า เป็น 1(Depression Group) หรือ 0(Control Group)

ซึ่งผมก็พยายามปรับพารามิเตอร์ต่าง ๆ ก็ได้ Acc สูงสุดที่ 65%

ผมเลยมามองว่าจะใช้ GRU ด้วย

ปล.การ Train ครั้งก่อน ทำการ shuffle ตาต้าเรียบร้อยแล้วนะครับ

คราวนี้ GRU จะต้องรับ input ยังไงเหรอครับ คือ รับเปน 11440 โดยไม่ shuffle แล้วตั้ง batch=143 เพื่อให้มันมองเป็น คน ๆ ไปเหรอครับ ?

หรือผมต้องมี 1 timesteps มาคั่นระหว่าง sample เพื่อให้มันแยกได้

อย่างถ้าเป็นใน HW1 ของ NLP มันจะให้เราทำ model ที่บอกว่าขึ้นประโยคใหม่หรือยัง แต่ของผมจะเป็นแบบนอกจากบอกว่าขึ้นประโยคใหม่แล้ว ยังต้องบอกด้วยว่าเป็น Noun, Verb, หรือ Adj. ประมาณนั้นอะครับ เพราะของผมคือต้องบอกว่าเป็น control หรือ depression โดยใช้ทั้งหมด 143 timesteps

อีกอย่างที่สงสัยคือตอน test น่ะครับ อย่างงี้หมายความว่าถ้าผม test แต่ละ sample ด้วย 143 timestep input ผมก็จะได้ 143 outputs ใช่มั้ยครับ แต่ที่ผมต้องการคือ output ตัวเดียวที่บอกว่าเป็น Depression หรือ Control เพียงค่าเดียว

Driving a car analogy

- Just drive without knowing where you are going
- Getting there vs getting there effectively
- Putting the wrong fuel into the car

Types of machine learning

1. Supervised learning
2. Unsupervised learning
3. Reinforcement learning

0. Pre-machine learning: rule-base

Pre-machine learning: 7-segment display

- **Input:** 7 binary values (0,1) forming a disply
- Given **x** = (A, B, C, D, E, F, G)
- **Output:** y, either 0, 1, …, 9 or not a number
- **Task:** write a program (a function F) that maps **x** to y; F(**x**) = y

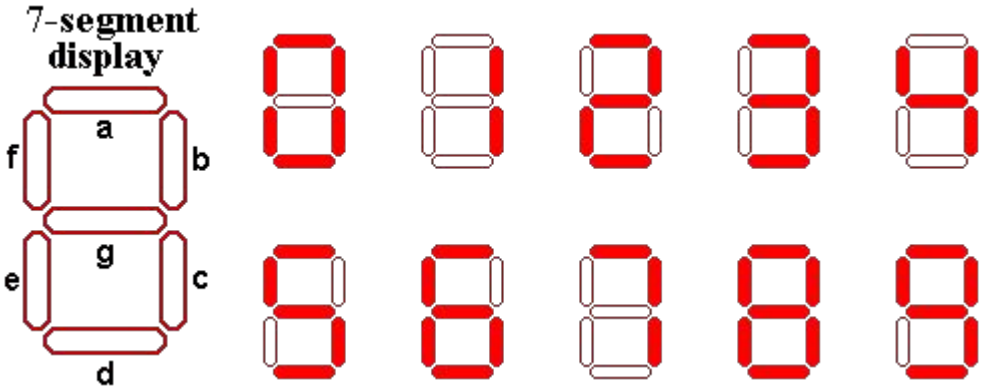


Image from
http://www.physics.udel.edu/~watson/scen103/colloq2000/7-seg.html

Mapping function

Y ⟵ X ⟶

Count	a	b	c	d	e	f	g
0	●	●	●	●	●	●	
1		●	●				
2	●	●		●	●		●
3	●	●	●	●			●
4		●	●			●	●
5	●		●	●		●	●
6	●		●	●	●	●	●
7	●	●	●				
8	●	●	●	●	●	●	●
9	●	●	●	●		●	●

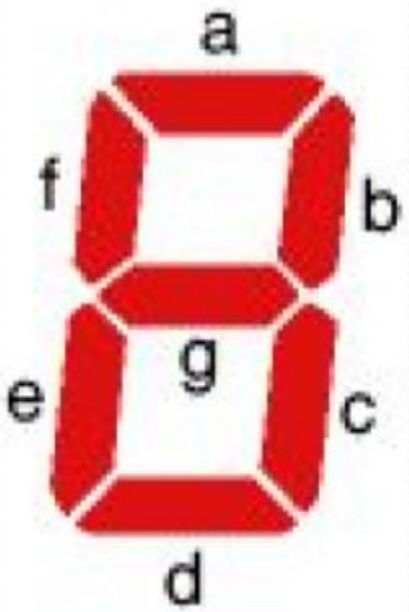
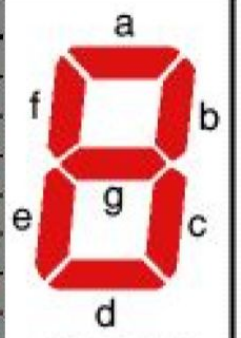


Image from: http://www.instructables.com/id/DIY-7-Segment-Display/

Mapping function

Count	a	b	c	d	e	f	g
0	●	●	●	●	●	●	
1		●	●				
2	●	●		●	●		●
3	●	●	●	●			●
4		●	●			●	●
5	●		●	●		●	●
6	●		●	●	●	●	●
7	●	●	●				
8	●	●	●	●	●	●	●
9	●	●	●	●		●	●



7-segment display

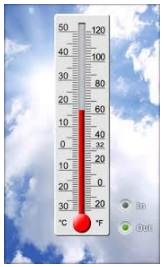
- IF A==1 && B==1 && C==1 && D==1 && E==1 && F==1 && G==0, THEN output(0).
- IF B==1 && C==1, THEN output(1)
- …..
- OTHERWISE, output(“not number”)

F(x)

Learning from data

- Machine learning requires identifying the same ingredients
  - Input, Output, Task

Real world observations



This is the hardest part of data science and the last part to be replaced by machines.

An example

- Handwritten digit recognition
- Input: **x** = 28 x 28 pixel image
- Output: y = digit 0 to 9
- Task: find F(**x**) such that y ≈ F(**x**)

Goal of machine learning is to find the best F(**x**) automatically from data

Supervised learning

- Learn a classifier F from a training set (input-output pairs)
  - $\{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), (\mathbf{x}_3, y_3), \ldots, (\mathbf{x}_n, y_n)\}$

x	y
	0
	1
	2

Need a training set for training.
Training = finding (optimizing) a good function f

Labeling (i.e., assigning y for each **x** in the training set) is typically done manually.

Types of machine learning

1. Supervised learning

   Learn a model F from pairs of (**x**,y)

2. Unsupervised learning

   Discover the hidden structure in unlabeled data **x** (no y)

3. Reinforcement learning

   Train an agent to take appropriate actions in an environment by maximizing rewards

Typical workflow of machine learning

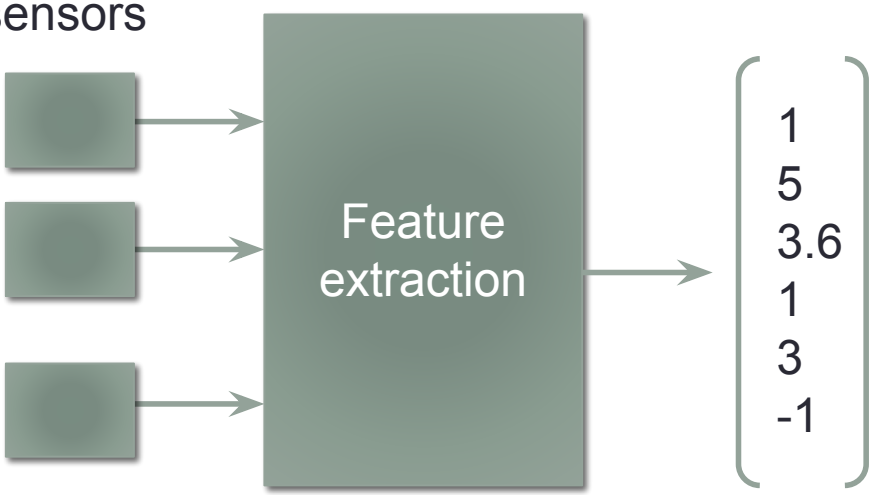
1. Feature extraction (getting the **x**)
2. Modeling
   - Training (getting the function F)
3. Evaluation
   - Metrics (defining what's the best function F)
   - Testing (getting the y for unseen inputs)

Typical workflow of machine learning

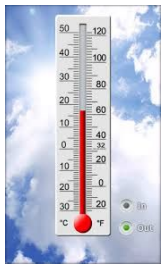
- The typical workflow

Real world observations

sensors



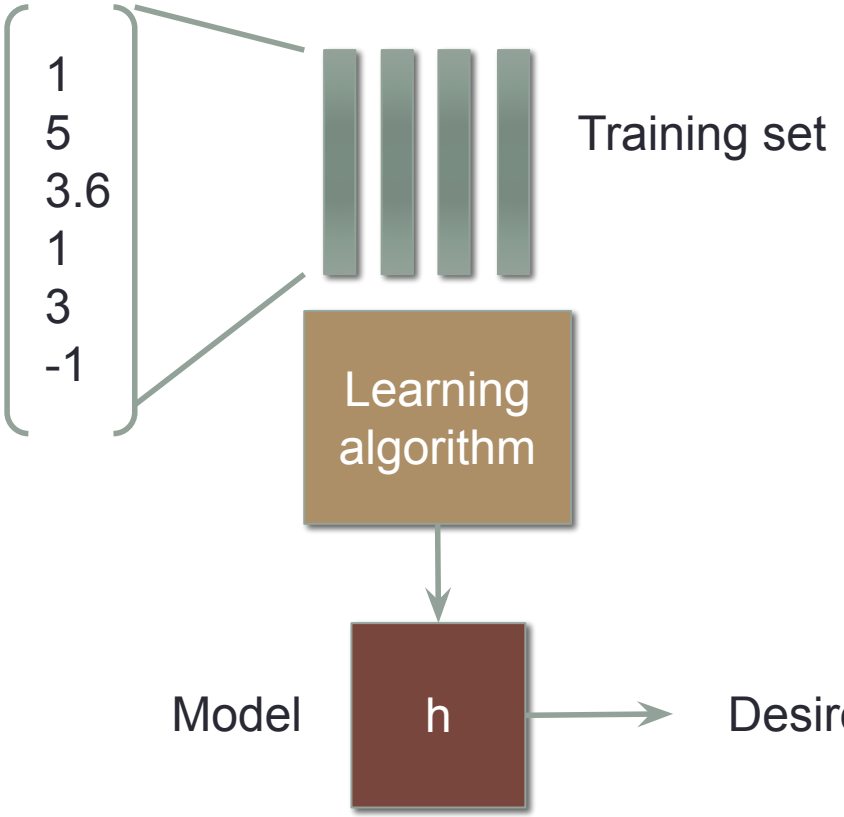
Feature vector
x



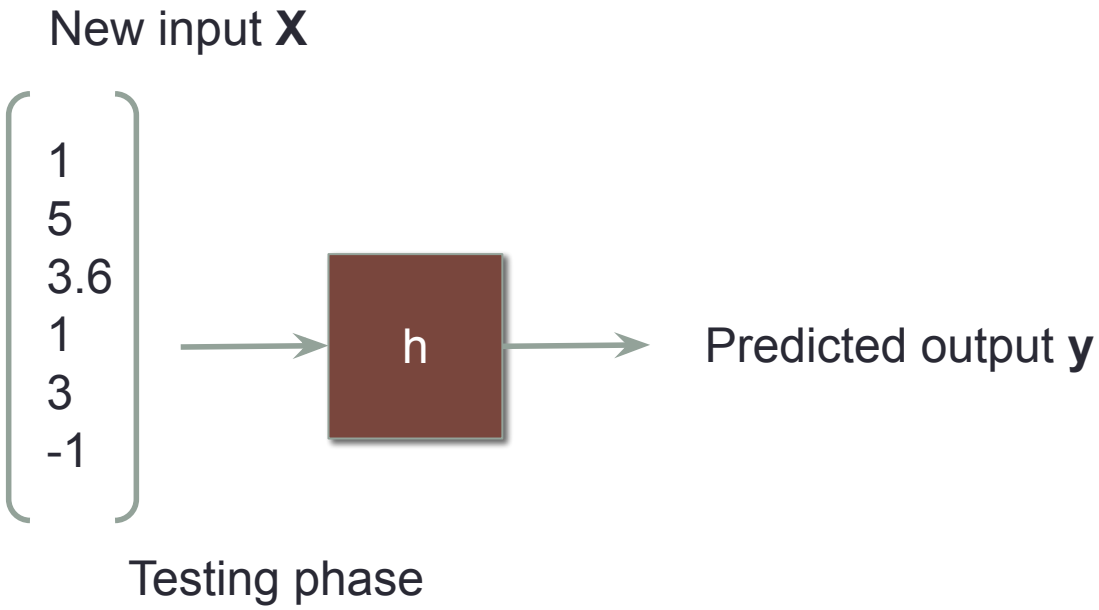
Hotdog classifier

https://www.youtube.com/watch?v=pqTntG1RXSY

How do we learn from data?



How do we learn from data?

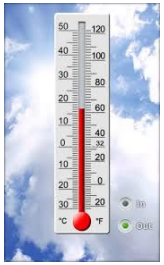


Feature extraction

- The process of extracting meaningful information related to the goal
- A distinctive characteristic or quality
- Example features

data1

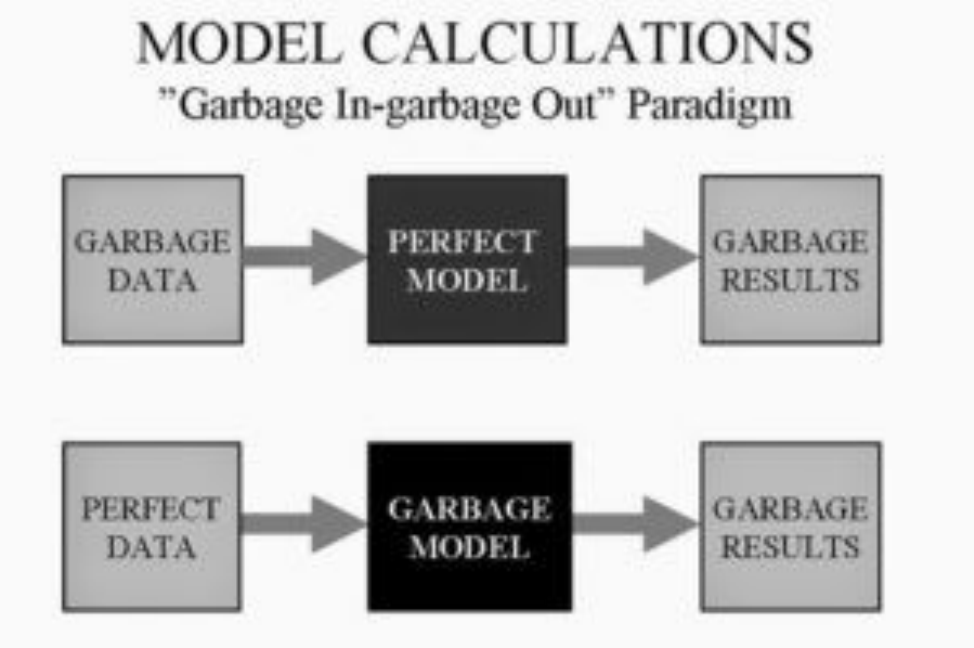
data2



data3

Garbage in Garbage out

- The machine is as intelligent as the data/features we put in
- “Garbage in, Garbage out”
- Data cleaning is often done to reduce unwanted things



https://precisionchiroco.com/garbage-in-garbage-out/

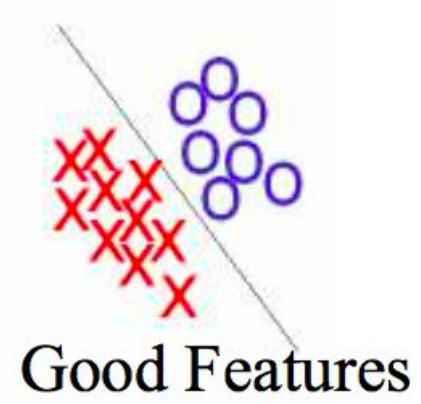
The need for data cleaning

However, good models should be able to handle some dirtiness!

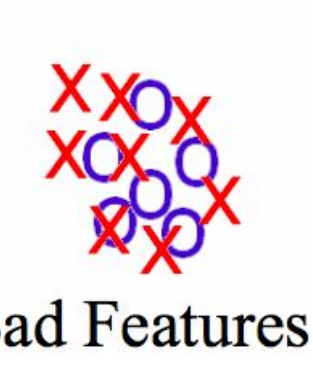
https://www.linkedin.com/pulse/big-data-conundrum-garbage-out-other-challenges-business-platform

Feature properties

- The quality of the feature vector is related to its ability to discriminate samples from different classes



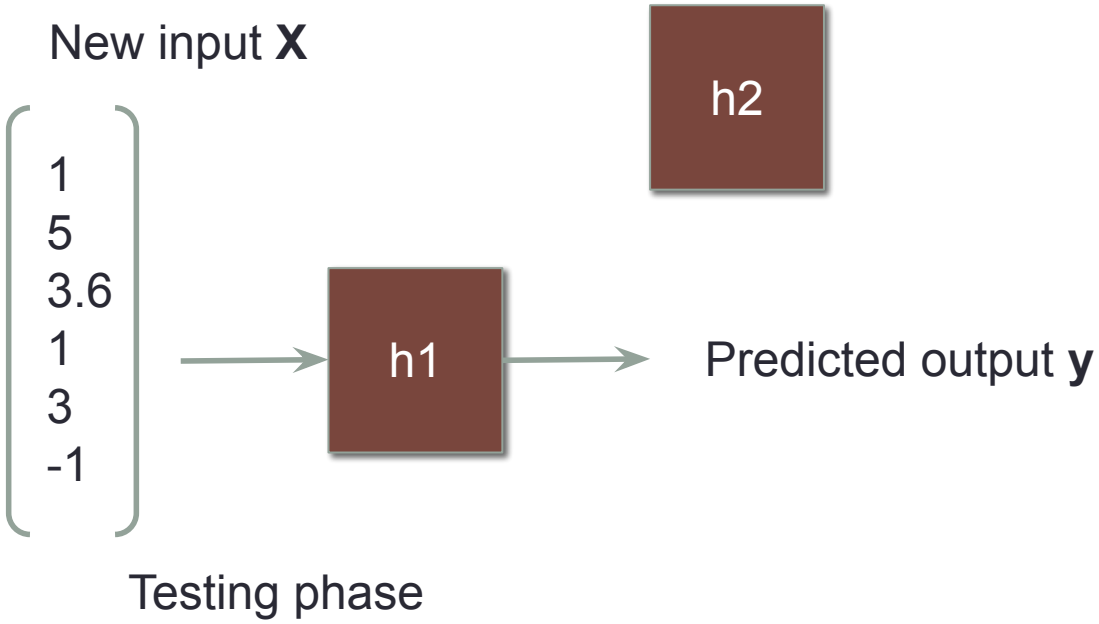
Good Features



Bad Features

Model evaluation

How to compare h1 and h2?



Metrics

- Compare the output of the models
  - Errors/failures, accuracy/success
- We want to quantify the error/accuracy of the models
- How would you measure the error/accuracy of the following

Ground truths

- We usually compare the model predicted answer with the correct answer.
- What if there is no real answer?
  - How would you rate machine translation?

ไป
ไหน
Model A: Where are you going?
Model B: Where to?

Designing a metric can be tricky, especially when it's subjective

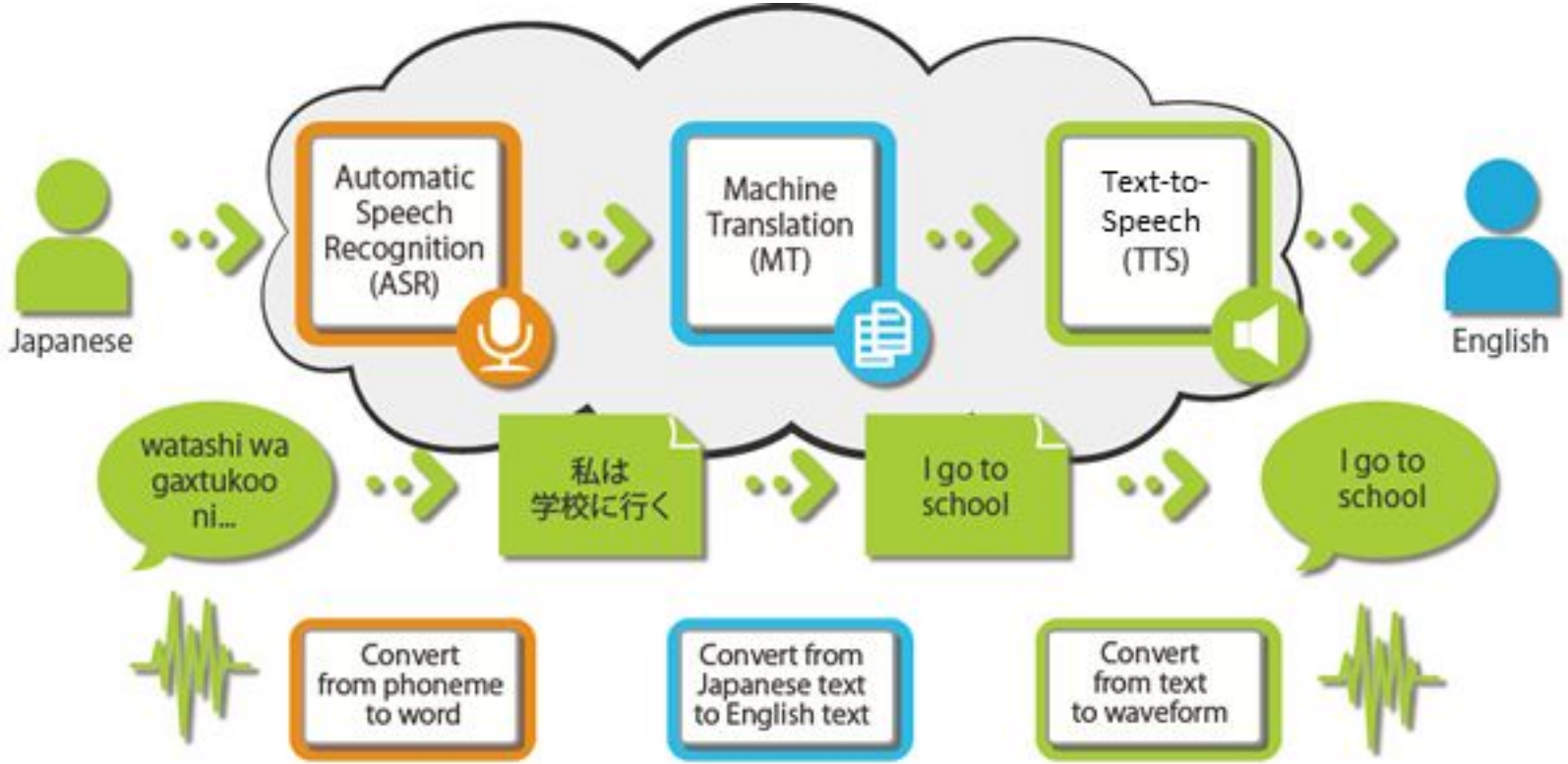
Metrics consideration 1

- Are there several metrics?

- Use the metric closest to your goal but never disregard other metrics.
  - May help identify possible improvements

Metrics consideration 2

- Are there sub-metrics?

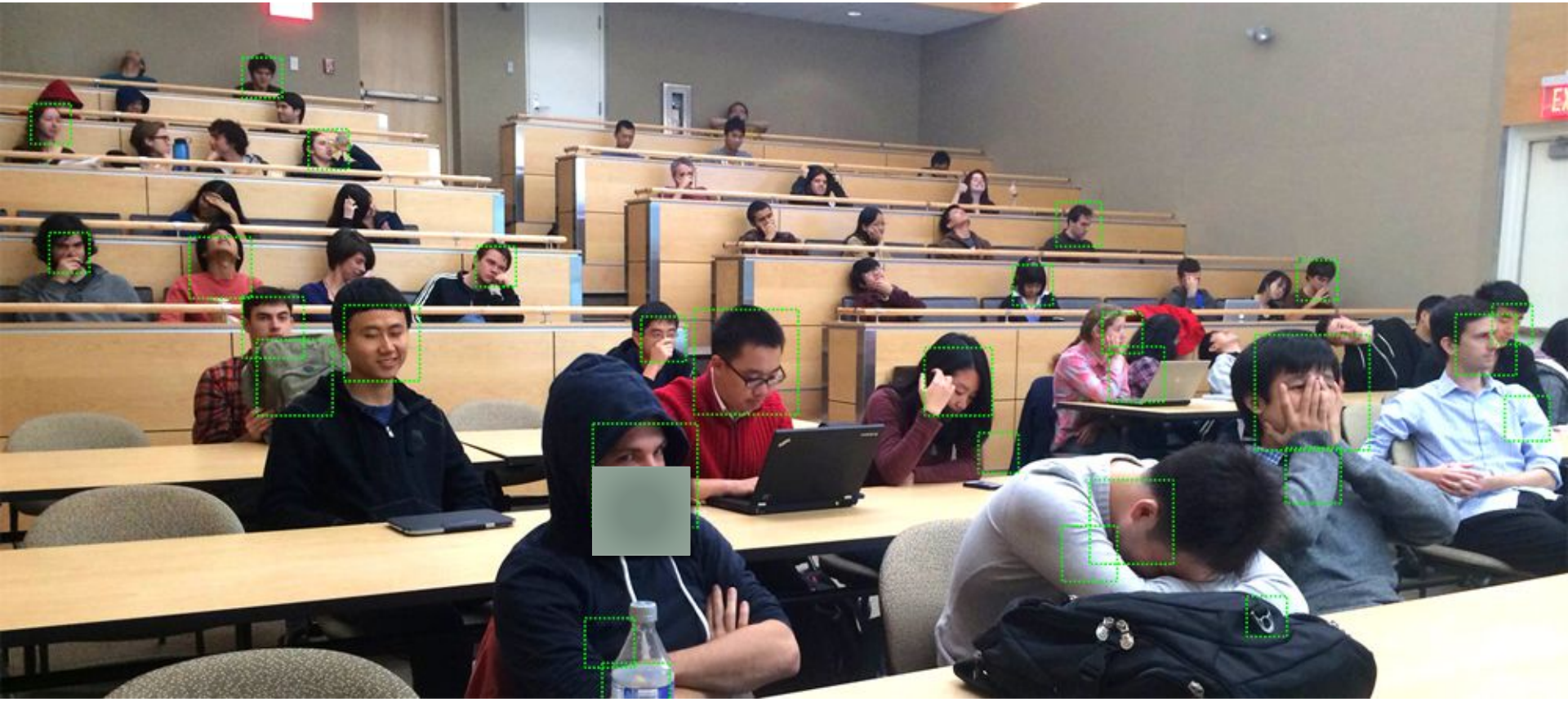


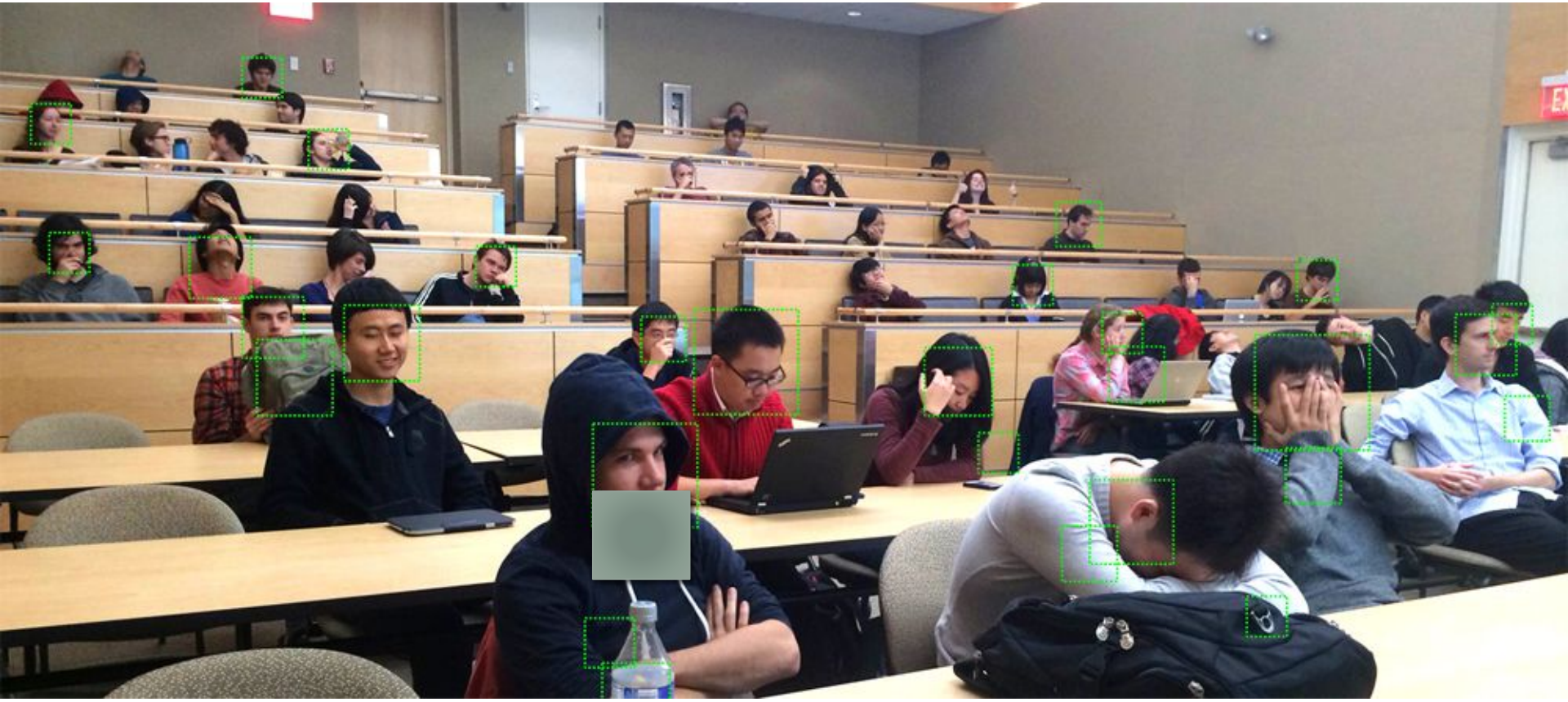
http://www.ustar-consortium.com/qws/slot/u50227/research.html

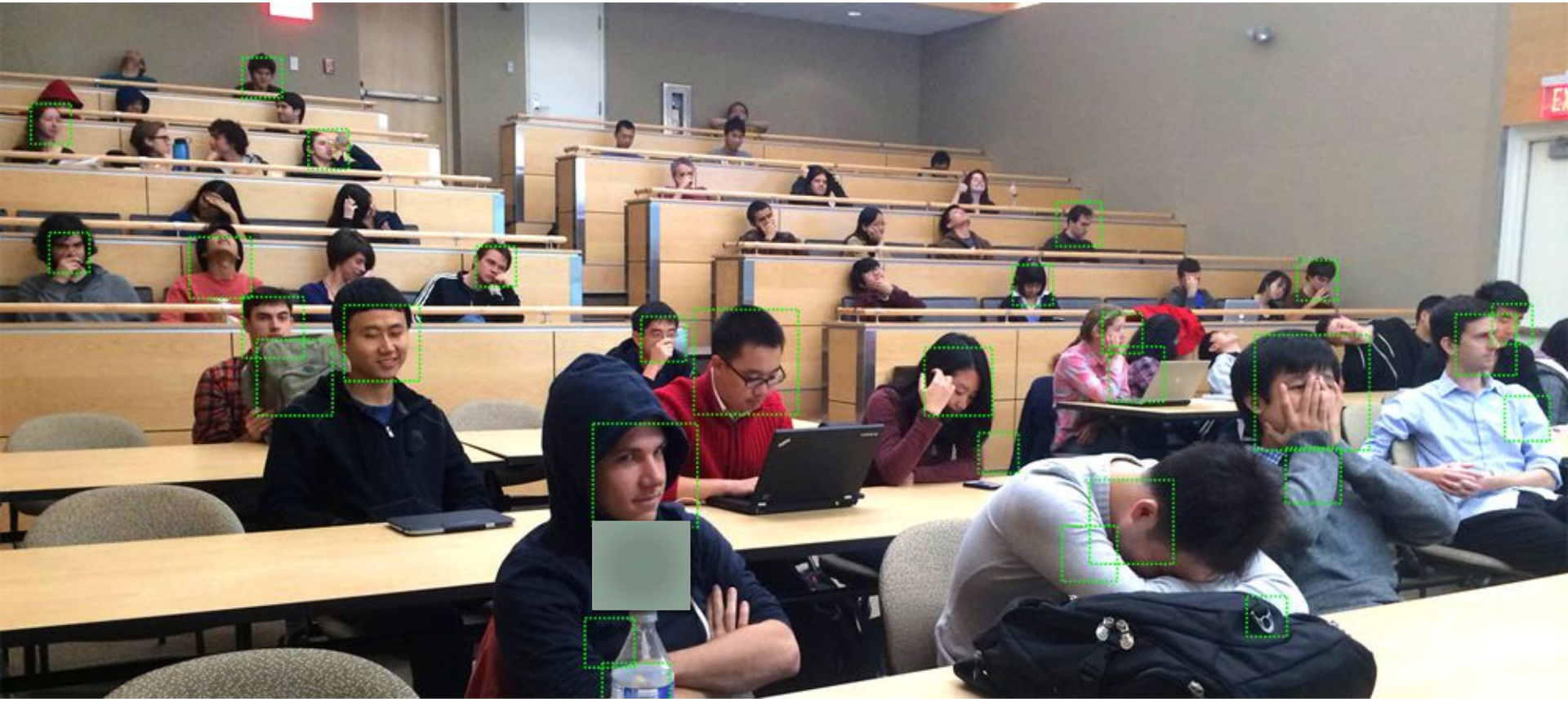
Metrics definition

- Defining a metric can be tricky when the answer is flexible

https://www.cc.gatech.edu/~hays/compvision/proj5/



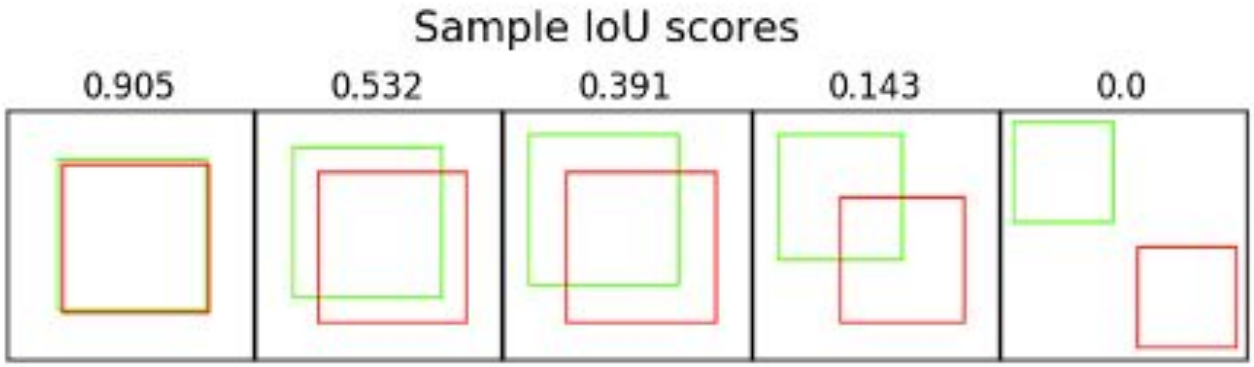




Be clear about your definition of an error before hand!
Make sure that it can be easily calculated!
This will save you a lot of time.

IoU (Intersection over Union)

What IoU score should be considered a detection?



http://ronny.rest/tutorials/module/localization_001/iou/

Commonly used metrics

- Error rate
- Accuracy rate

- Precision
- True positive
- Recall
- False alarm
- F score

A detection problem

- Identify whether an event occur
- A yes/no question
- A binary classifier

Smoke detector

Hotdog detector

Evaluating a detection problem

- 4 possible scenarios

		Detector	
		Yes	No
Actual	Yes	True positive	False negative (Type II error)
	No	False Alarm (Type I error)	True negative

True positive + False negative = # of actual yes
False alarm + True negative = # of actual no

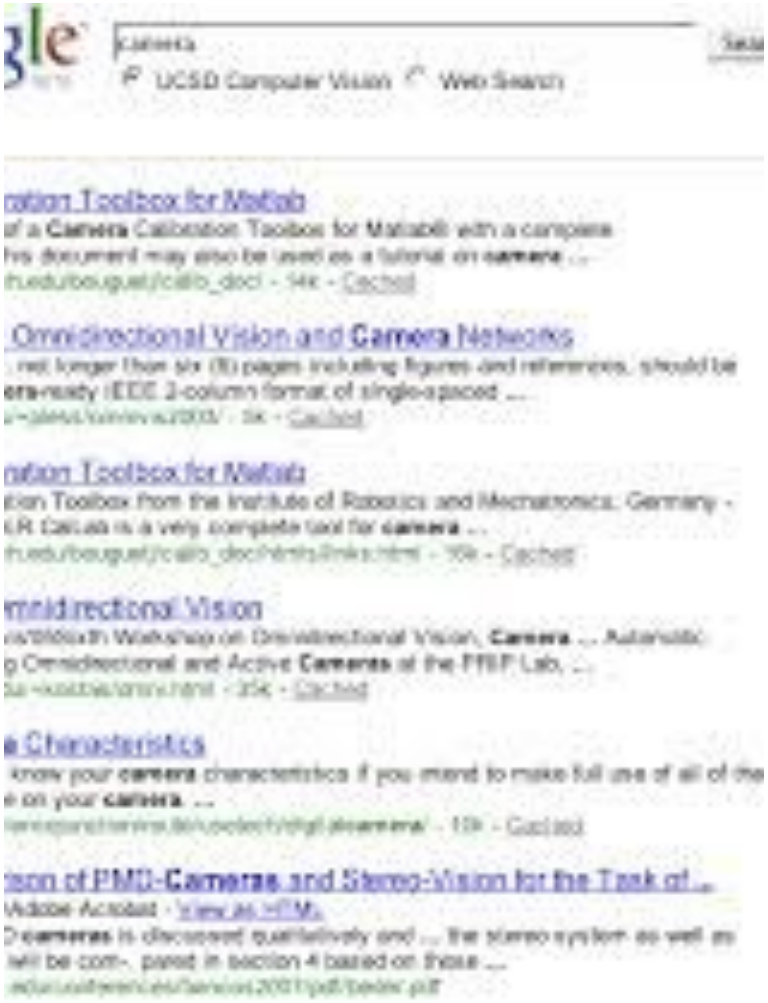
- False alarm and True positive carries all the information of the performance.

Definitions

- True positive rate (Recall, sensitivity)
  = # true positive / # of actual yes
- False positive rate (False alarm rate)
  = # false positive / # of actual no
- False negative rate (Miss rate)
  = # false negative / # of actual yes
- True negative rate (Specificity)
  = # true negative / # of actual no

- Precision = # true positive / # of predicted positive

Search engine example



A recall of 50% means?

A precision of 50% means?

When do you want high recall?
When do you want high precision?

Recall/precision

- When do you want high recall?
- When do you want high precision?

- Initial screening for cancer
- Face recognition system for authentication
- Detecting possible suicidal postings on social media

Usually there's a trade off between precision and recall. We will revisit this later

Definitions 2

- F score (F1 score, f-measure)

$$F_1 = 2 \cdot \frac{1}{\frac{1}{\text{recall}} + \frac{1}{\text{precision}}} = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}.$$

  - A single measure that combines both aspects
  - A harmonic mean between precision and recall (an average of rates)

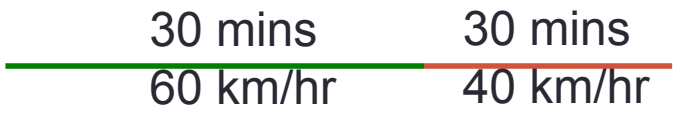
Note that precision and recall says nothing about the true negative

Harmonic mean vs Arithmetic mean

- You travel for half an hour for 60 km/hr, then half an hour for 40 km/hr. What is your average speed?
  - Arithmetic mean = 50 km/hr
  - Harmonic mean

$$\frac{n}{\frac{1}{x_1}+\ldots+\frac{1}{x_n}} = \frac{2}{\frac{1}{40}+\frac{1}{60}} = 48 \text{ km/hr}$$

  - Total distance covered in 1 hour = 30+20 = 50



Harmonic mean vs Arithmetic mean

- You travel for distance X for 60 km/hr, then another X for 40 km/hr. What is your average speed?
  - Arithmetic mean = 50 km/hr
  - Harmonic mean

$$\frac{n}{\frac{1}{x_1}+\ldots+\frac{1}{x_n}} = \frac{2}{\frac{1}{40}+\frac{1}{60}} = 48 \text{ km/hr}$$

  - Total distance covered 2X

Harmonic mean vs Arithmetic mean

- For the arithmetic mean to be valid you need to compared over the same number of hours (denominator)

- For precision and recall, you have different denominators, but the same numerator, which fits the harmonic mean.

True positive rate (Recall, sensitivity)
= # true positive / # of actual yes

Precision = # true positive / # of predicted positive

Evaluating models

- We talked about the training set used to learn the model
- We use a different data set to test the accuracy/error of models – “test set”
- We can still compute the error and accuracy on the training set
- Training error vs Testing error
- We will discuss how we can use these to help guide us later

Other considerations when evaluating models

- Training time
- Testing time
- Memory requirement
- Parallelizability
- Latency

Course walkthrough

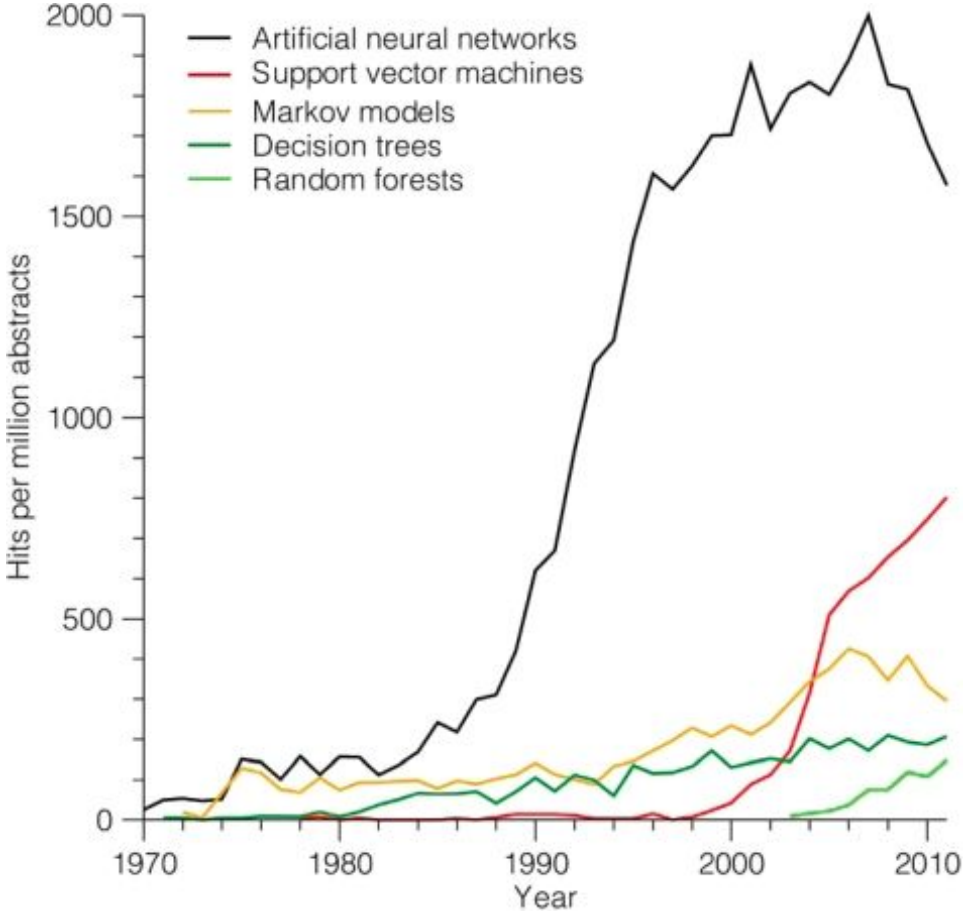
Traditional Machine learning

Deep learning

คาบเรียนที่	เนื้อหา	การบ้านและควิซ
1 - 13/8	Introduction	
2 - 20/8	Regression & Jupyter Notebook demo	เริ่มHW1
3 - 27/8	Probability review	
4 - 3/9	MLE, MAP, and Naive Bayes	ส่งHW1, Quiz 1, เริ่มHW2
5 - 10/9	GMM and EM	
6 - 17/9	Dimensionality reduction and visualization	ส่งHW2, Quiz 2, เริ่มHW3
7 - 24/9	SVM	
8 - 1/10	Neural network basics & Gcloud and Keras demo	ส่งHW3, Quiz 3, เริ่มHW4
7/10-11/10	Midterm week	
9 - 15/10	CNN, Recurrent architectures	
10 - 22/10	Recent Advances in NN	ส่งHW4, Quiz 4, ส่ง course project proposal
11 - 29/10	Reinforcement Learning	
11 - 5/11	Unsupervised methods	Course project progress
12 - 12/11	Probabilistic Graphical models and Causal Inference	
13 - 19/11	Tricks of the trade: machine learning in the real world (with guest lecture)	
14 - 26/11	Project presentation	ส่งcourse project

Why anything else besides deep learning

- The rise and fall of machine learning algorithms



Methods used in bioinformatics papers

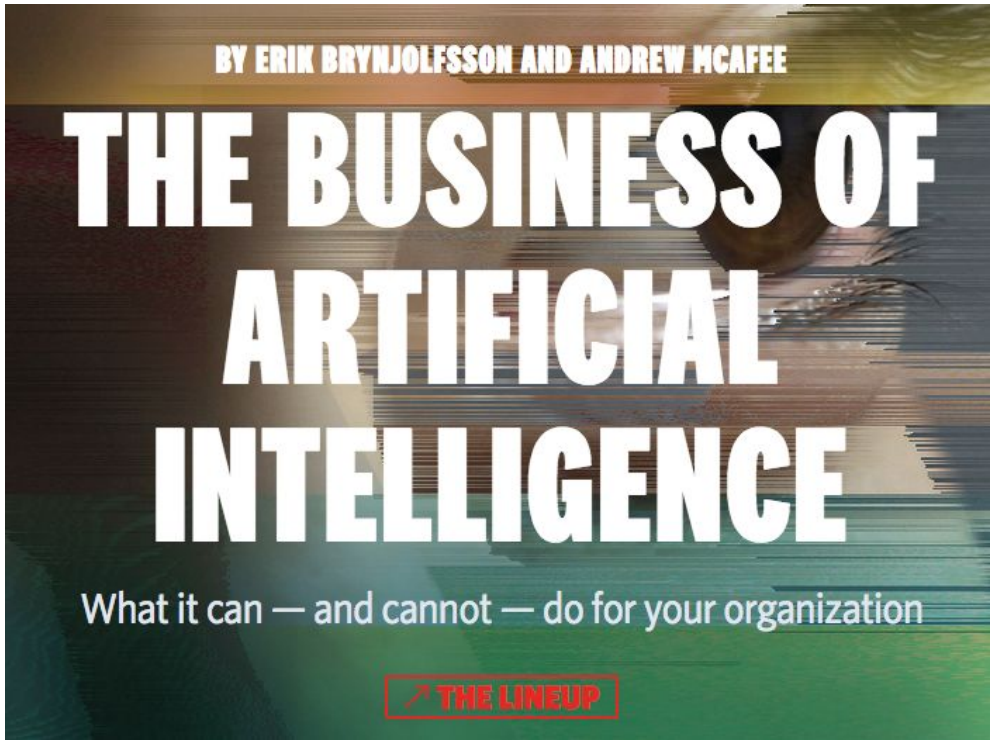
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3232371/figure/F1/

What we will not cover

- Random forest
- Decision trees
- Boosting

Homework

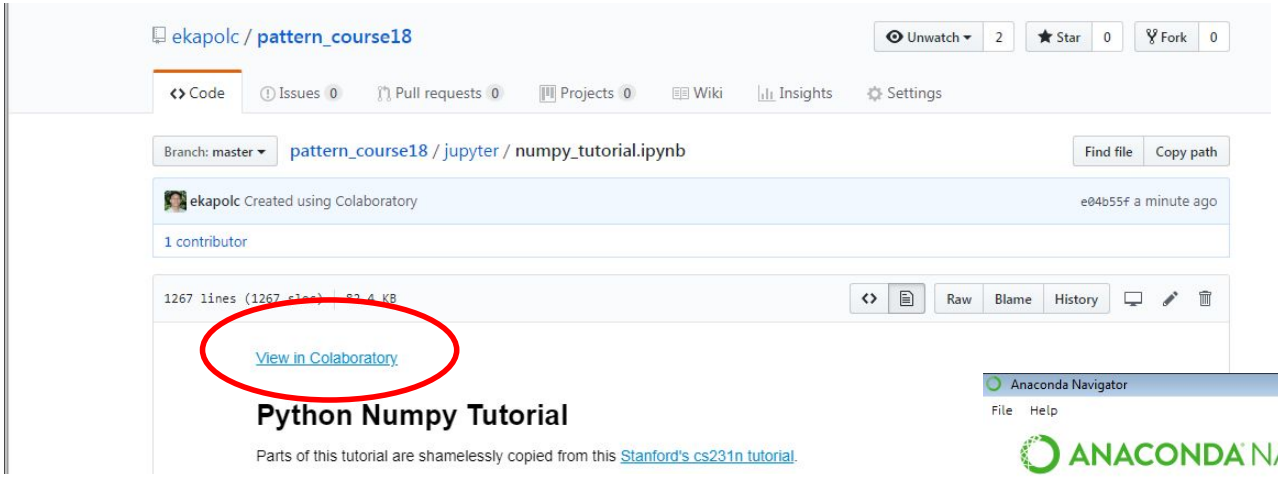
- Reading assignment

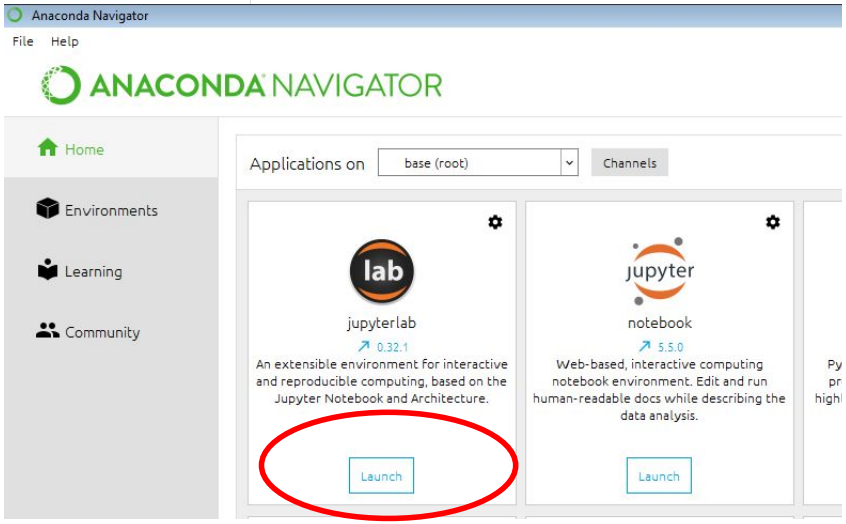


https://web.archive.org/web/20170828143636/https://hbr.org/cover-story/2017/07/the-business-of-artificial-intelligence

Jupyter lab and Colaboratory

- We will use Jupyter lab and Colaboratory for this course





https://www.anaconda.com/download/