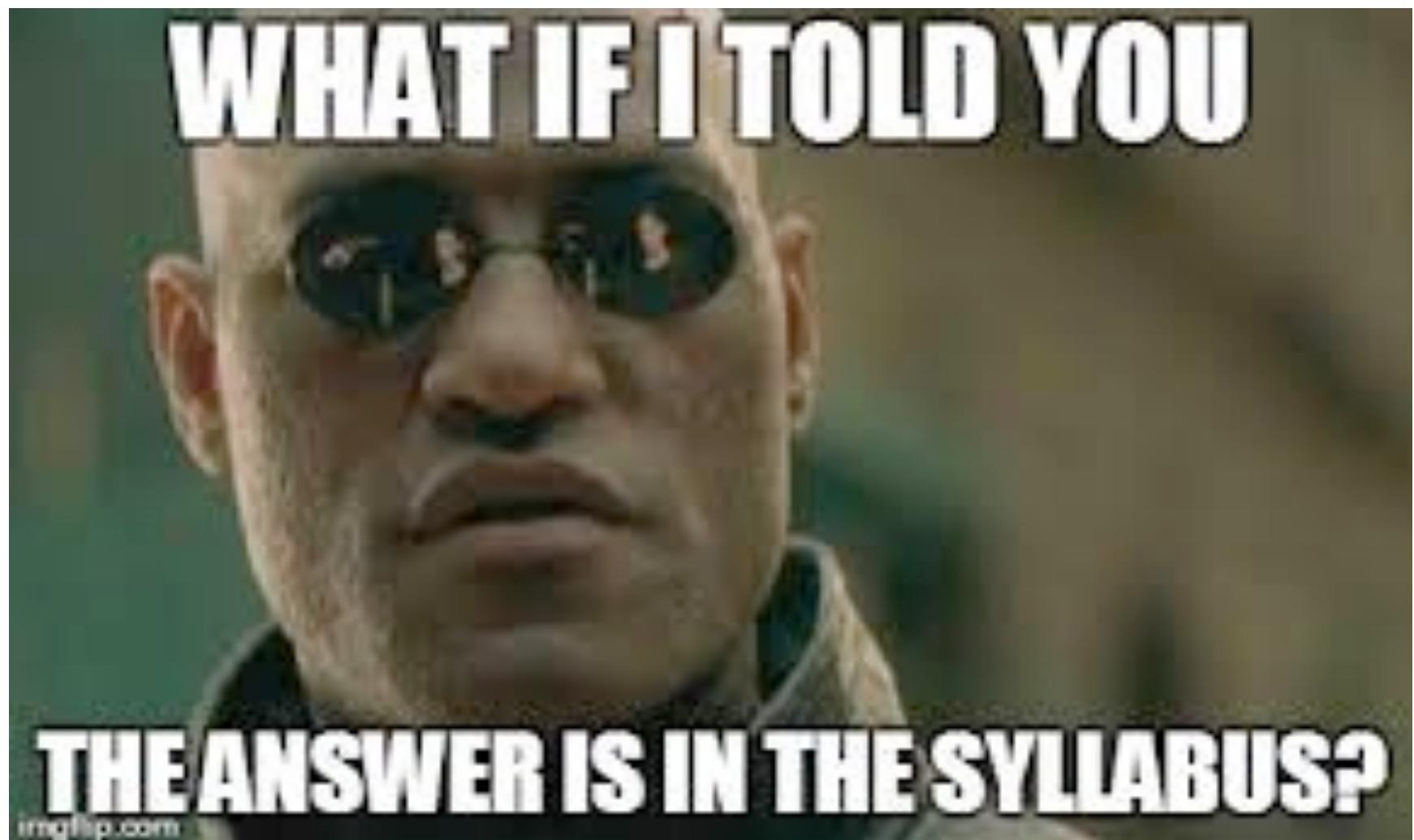


INTRODUCTION

Pattern Recognition

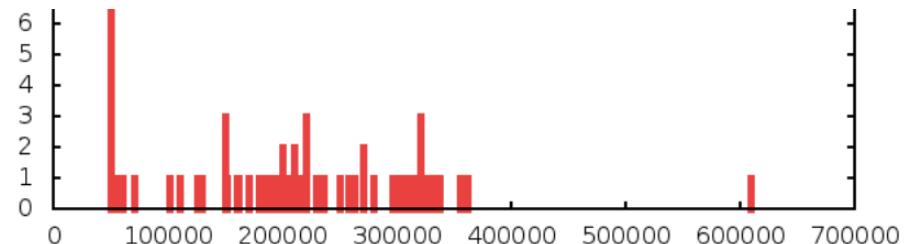
Slides at <https://ekapolc.github.io/slides/L1-intro.pdf>

Syllabus



Registration

- Graduate students
 - 12 slots, sec 2
 - If filled, register as V/W only
- For undergrads, sec 21
- Signup sheet for sit-ins, s/u, v/w going around the room



Tools

- Python
- Python
- Python
- Jupyter
- Numpy
- Scipy
- Pandas
- Tensorflow, Keras



ExxonMobil

ການວິຊາວິទະກຽມກອບພົວເປົວ
ການວິດວິກຽມທາງສອງ
ອັພົາອັກຮູນທ່າວິດຈາລັຍ

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

- 2110597.21 (2017/1)
- https://www.courseville.com/?q=courseville/course/register/2110597.21_2017_1&spin=on



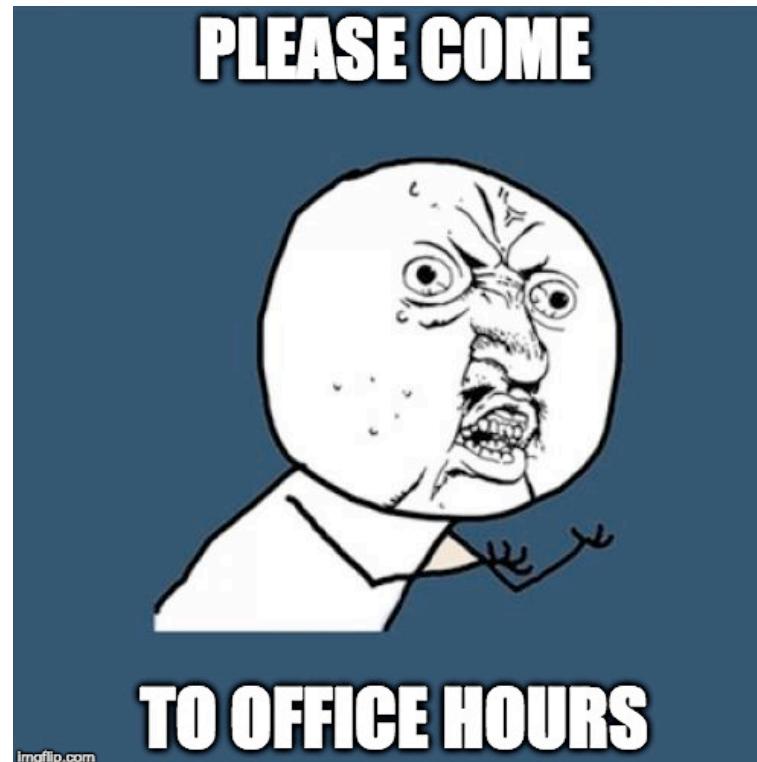
Password: cattern

Piazza

- <http://piazza.com/chula.ac.th/fall2017/2110597>
- Requires chula.ac.th email
- 5 points of participation score comes from piazza

Office hours

- Thursdays 16.30-18.30 starting from Aug 31st
- Location TBA



Cloud

- Gcloud
- Credit card

 Google Cloud Platform

Try Cloud Platform for free

Country
United States

Acceptances

Please email me updates regarding feature announcements, performance suggestions, feedback surveys and special offers.

Yes No

I have read and agree to the [Google Play Android Developer API Terms of Service](#).
Required to continue

Yes No

I have read and agree to the [Google Cloud Platform Free Trial Terms of Service](#).
Required to continue

Yes No

Agree and continue

 Access to all Cloud Platform Products
Get everything you need to build and run your apps, websites, and services, including Firebase and the Google Maps API.

 \$300 credit for free
Sign up and get \$300 to spend on Google Cloud Platform over the next 12 months.

 No autocharge after free trial ends
We ask you for your credit card to make sure you are not a robot. You won't be charged during or after your free trial ends.

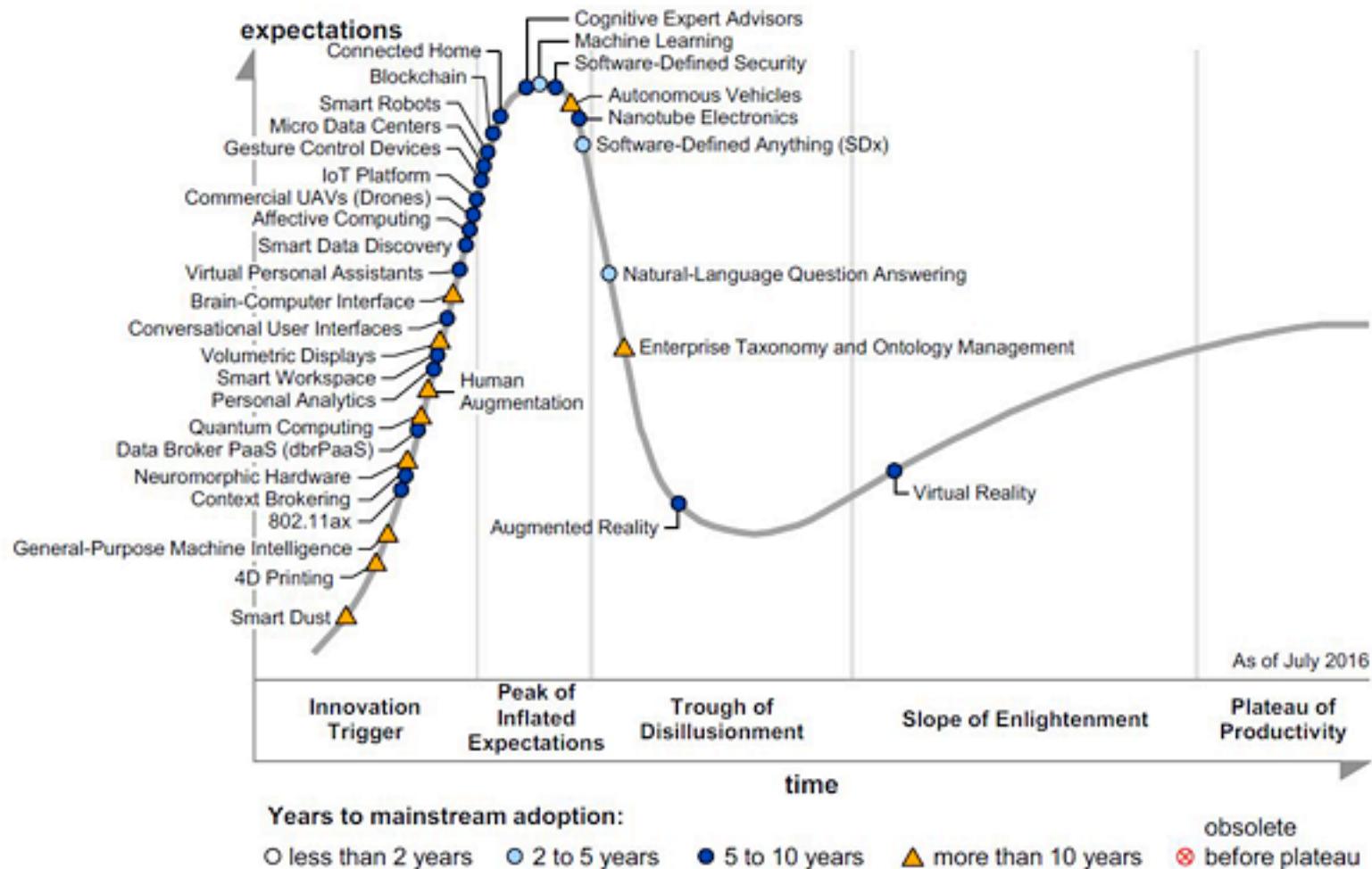
Course project

- 3-4 people (exact number TBA)
- 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



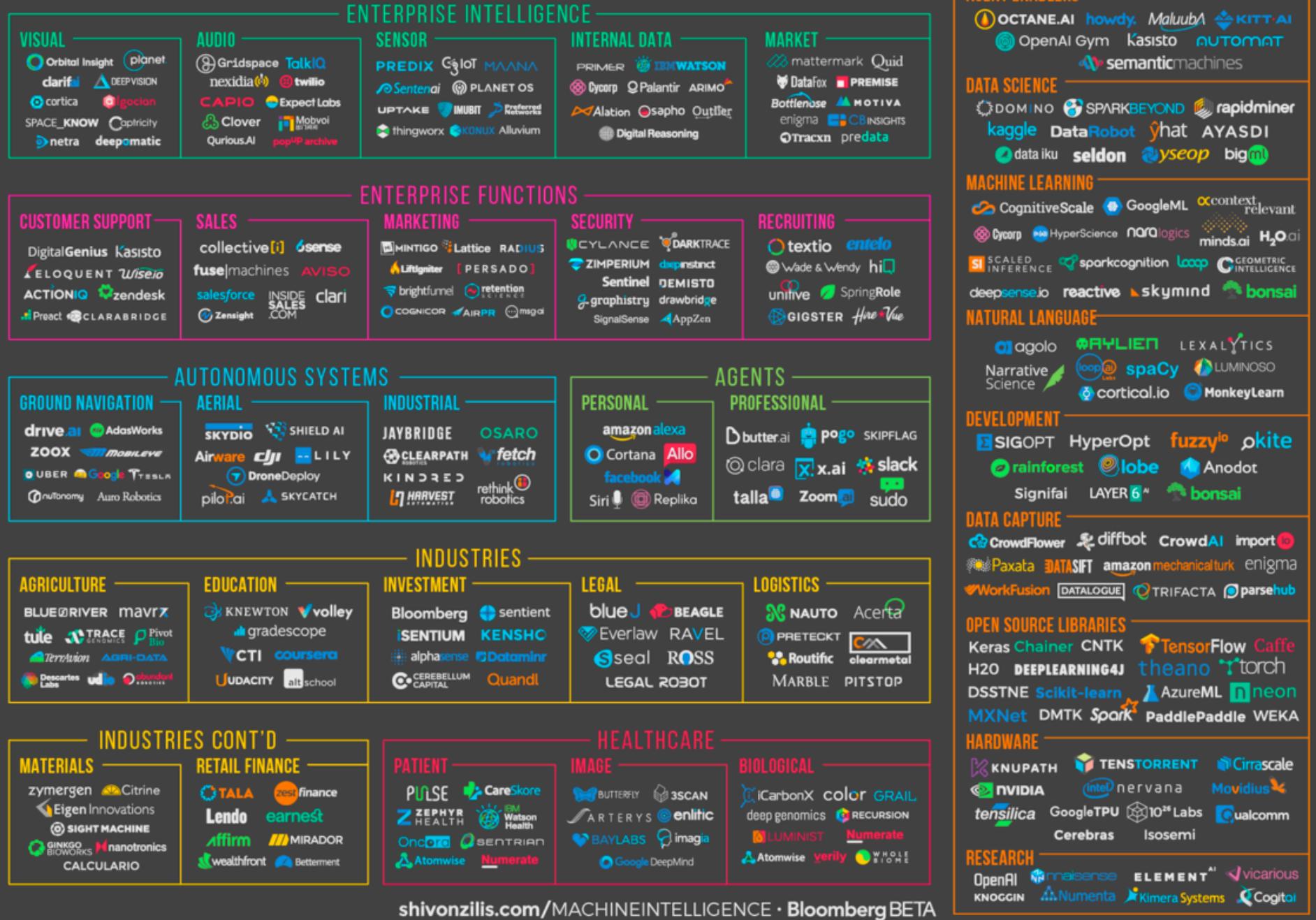
The machine learning trend



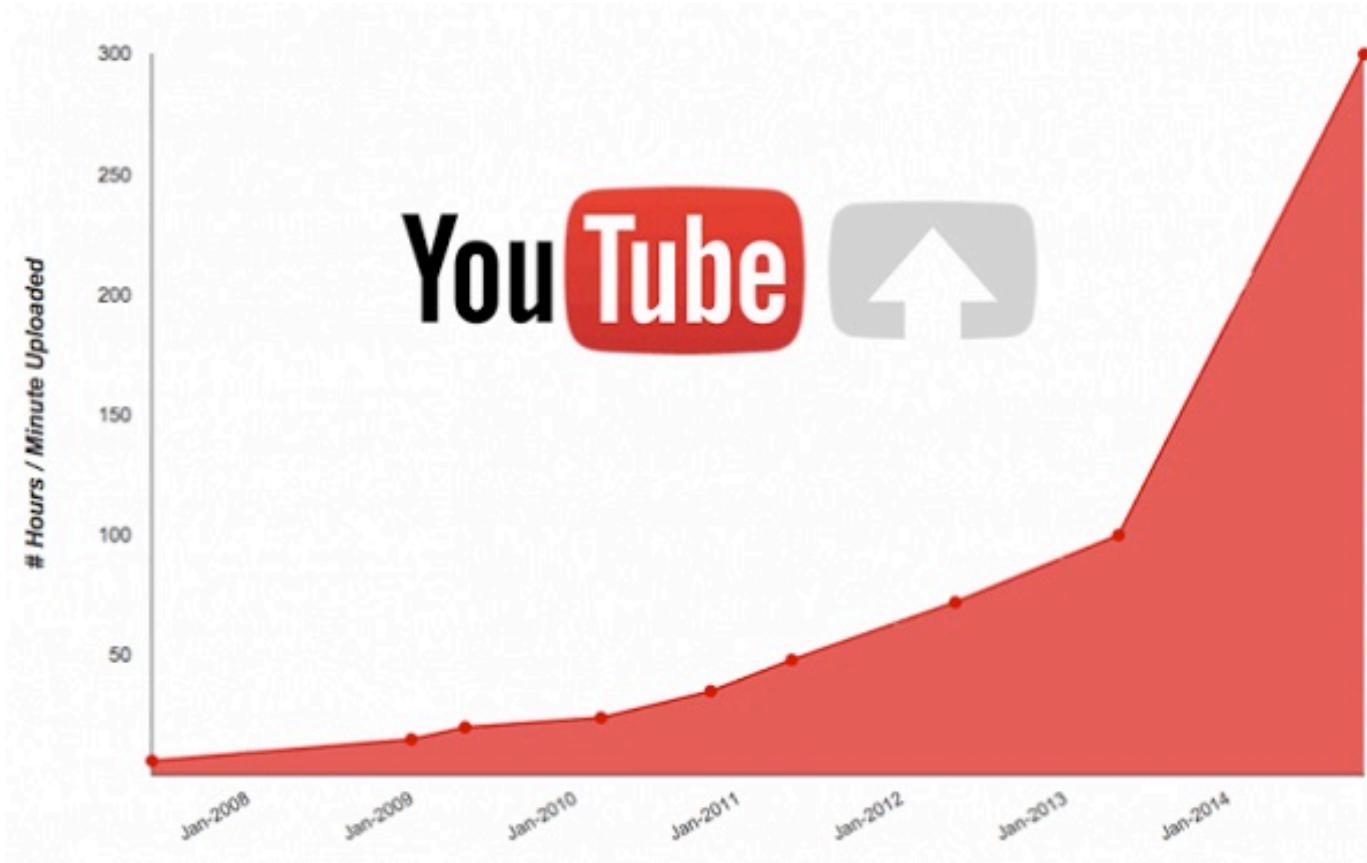
Source: Gartner (July 2016)

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

MACHINE INTELLIGENCE 3.0



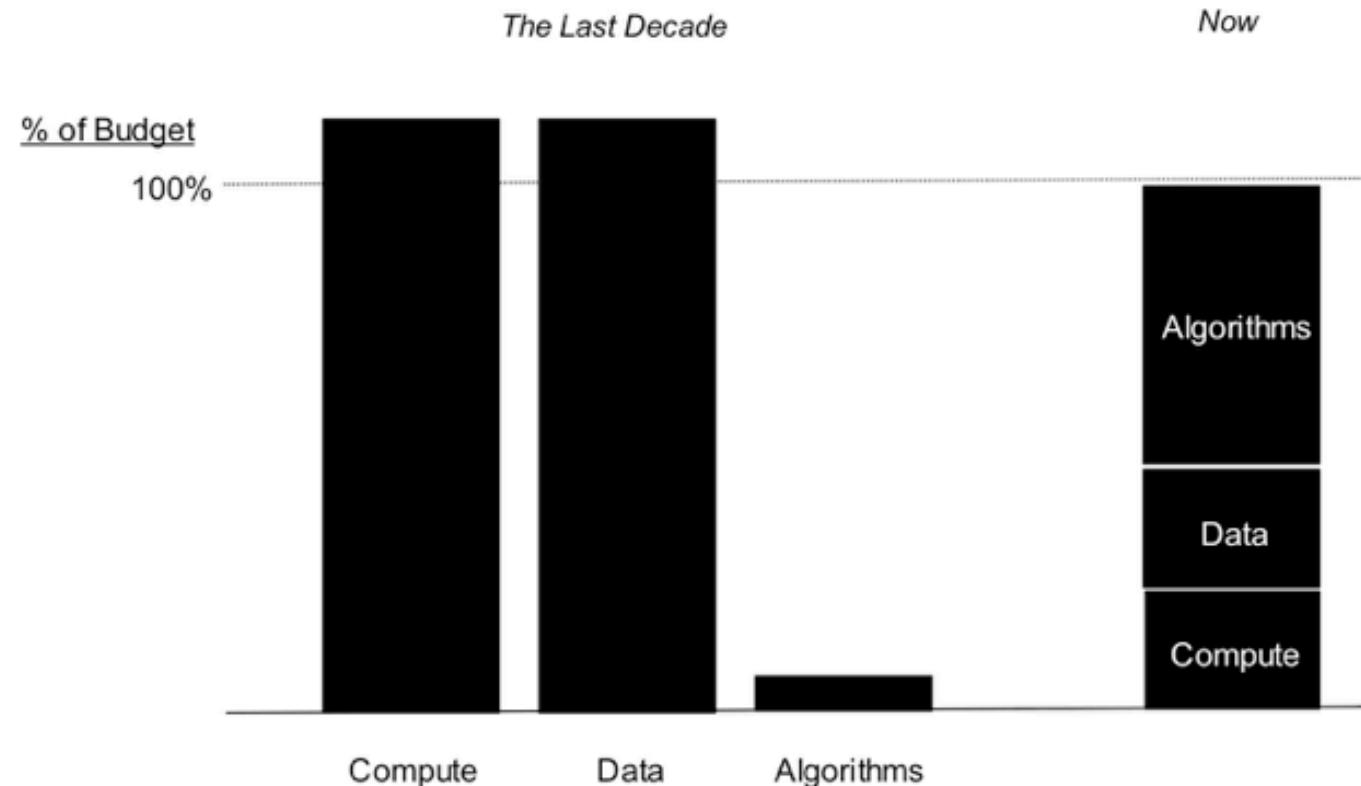
The data era



2017 numbers = 400 hours/min

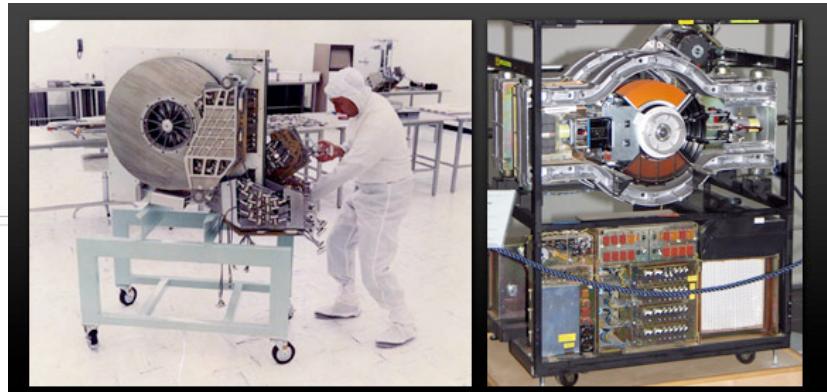
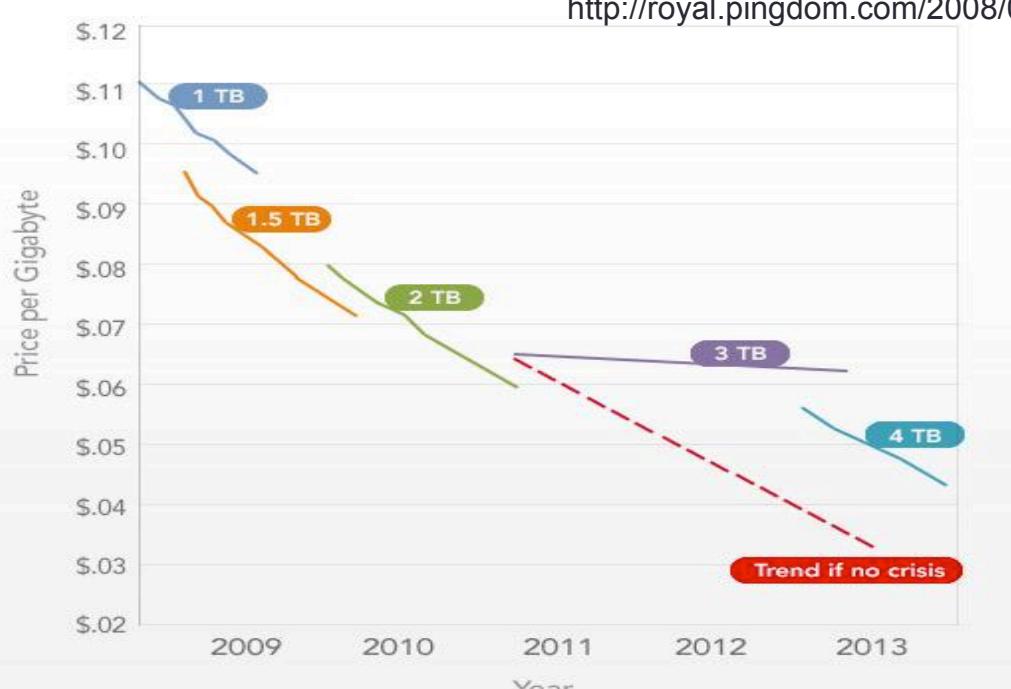
Factors for ML

- Data
- Compute
- Algo



The cost of storage

Cost per GB Trend Lines

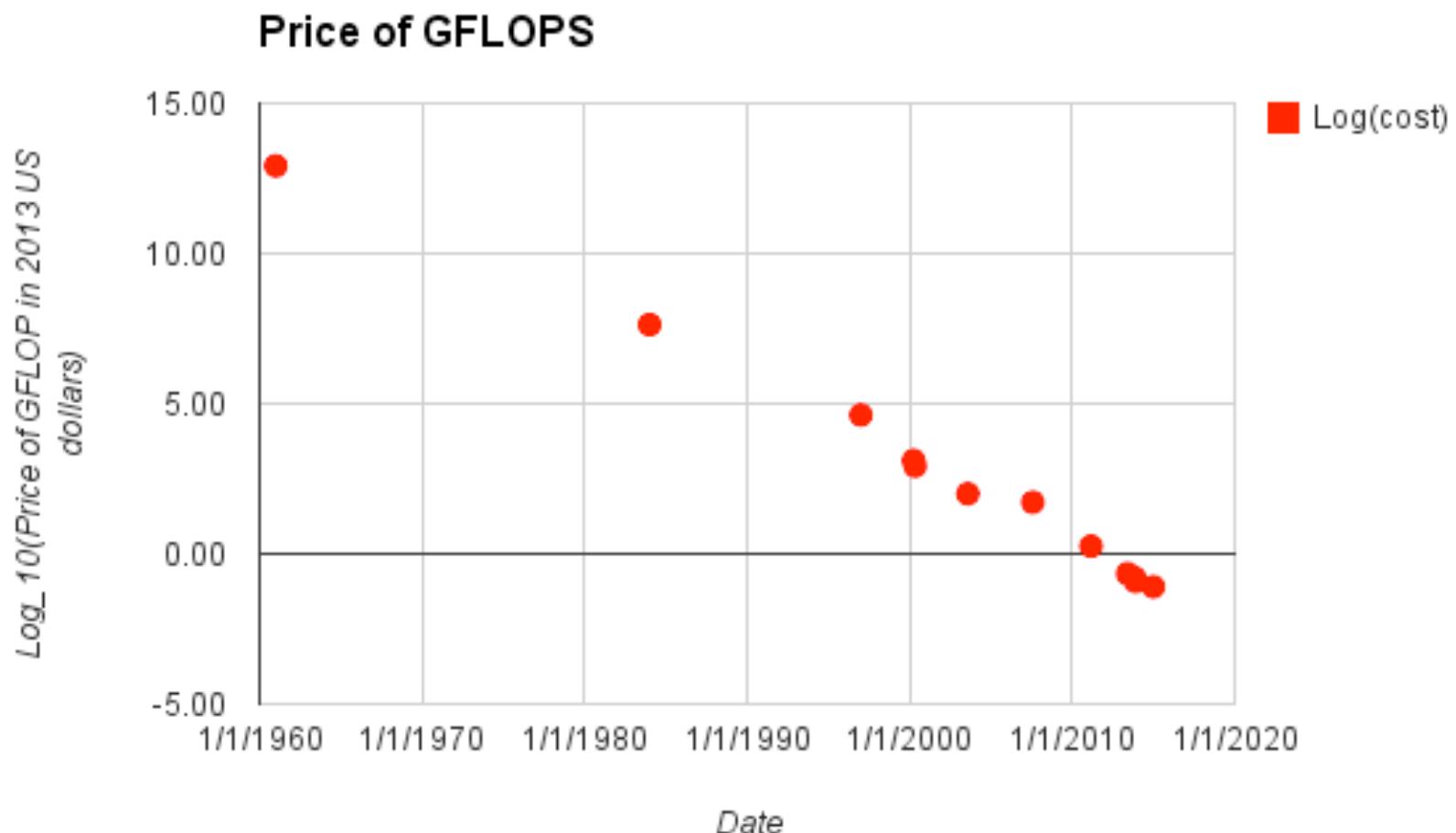


<http://royal.pingdom.com/2008/04/08/the-history-of-computer-data-storage-in-pictures/>

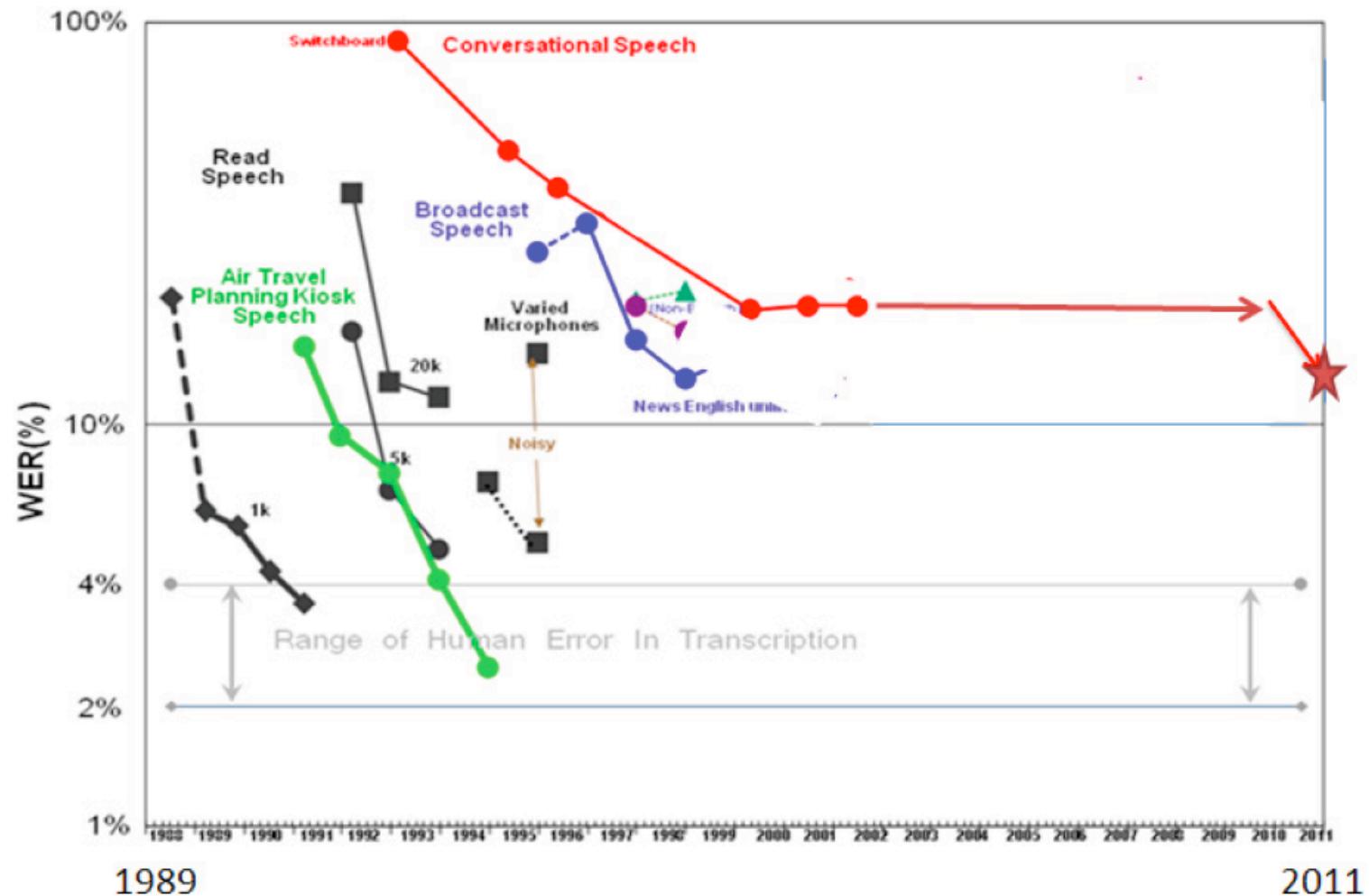
1980 250MB hard disk drive
250 kg 100k USD (300k USD in today's dollar)

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

The cost of compute



Hitting the sweet spot on performance



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 a video



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



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.

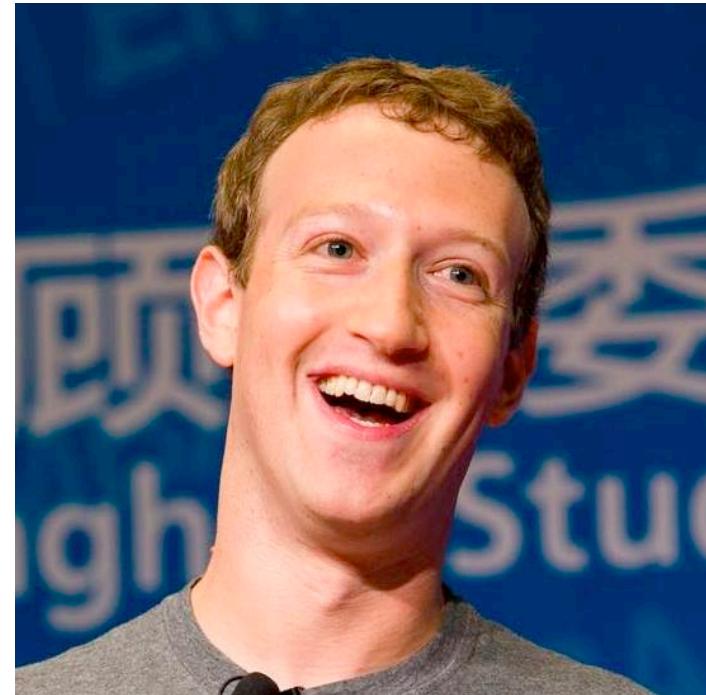
8:07 AM - 25 Jul 2017

© Twitter

- “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”



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

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

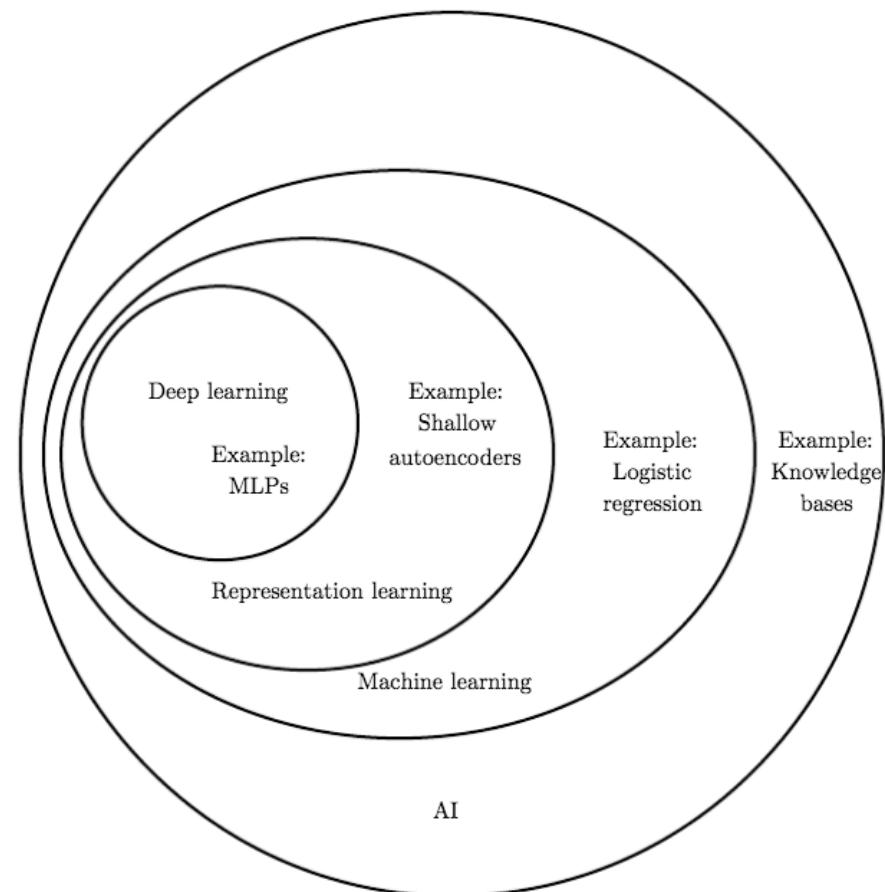
The Screwdriver and the Screw

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



How do we learn from data?

- The typical workflow



Real world observations

Squire Trelawny, Dr. Livesey,
and the rest of these gentlemen
having asked me to write down
the whole particulars about Trellis
and his crew, I have taken
to the end, keeping nothing back
but the bearings of the island;
and that only because there is still
more to tell than I can take up
in my pen in the year of grace 17—
and go back to the time when my
father kept the Admiral Pinnace
in the harbor of Port Royal.
With the salve cut fine took up his
lodging under our roof.

I remember him as it is
yesterday; a boy, a bold pell-mell
to the door, his wavy chest
following behind him in a hand-
barrow, a tall, strong, heavy,
and brown man; his tarry pigtail
falling over his shoulder; his
soiled blue coat, his hands ragged
and scoured, with black, broken

nails, and the salve-cut across
one cheek, a dirty, livid white. I
remember him looking round the
cove, and whistling to himself an
old sailor's tune, and singing out
in that old sea-song that he sang
so often afterwards:

'Tillien was on the deck wearin'
about, when he heard a voice
'east' in the high, old sombre
voice that seemed to have been
tuned and broken at the captain
box, and which sounded like
a bit of stick like a bit of twigs
that he carried, and when my fa-
ther appeared, called roughly for
a glass of rum. This, when I was
bent to him, he did not shun
like a commoner, looking on
the tank and still looking about
him at the cliffs and up at me
again.'

'This is a handy cove,' says he
at length; 'and a pleasant sitty and

grog-shop. Much company, mate?'
My father told him so, very
little company, the more was the
pleasure.

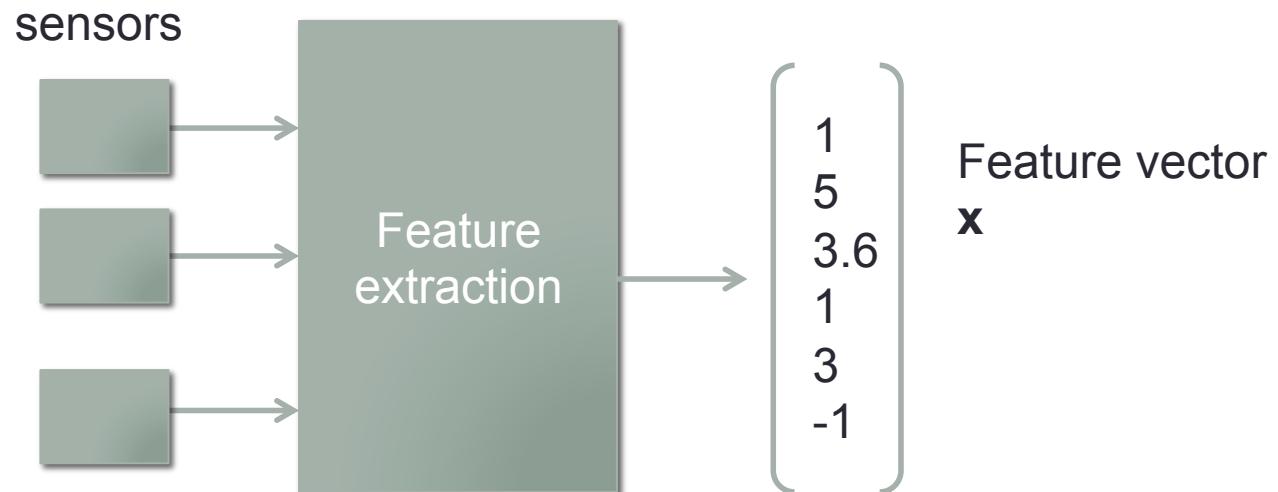
'Well, then,' said he, 'this is the
birth for me. Here you, money.'

he cried to the man who trundled

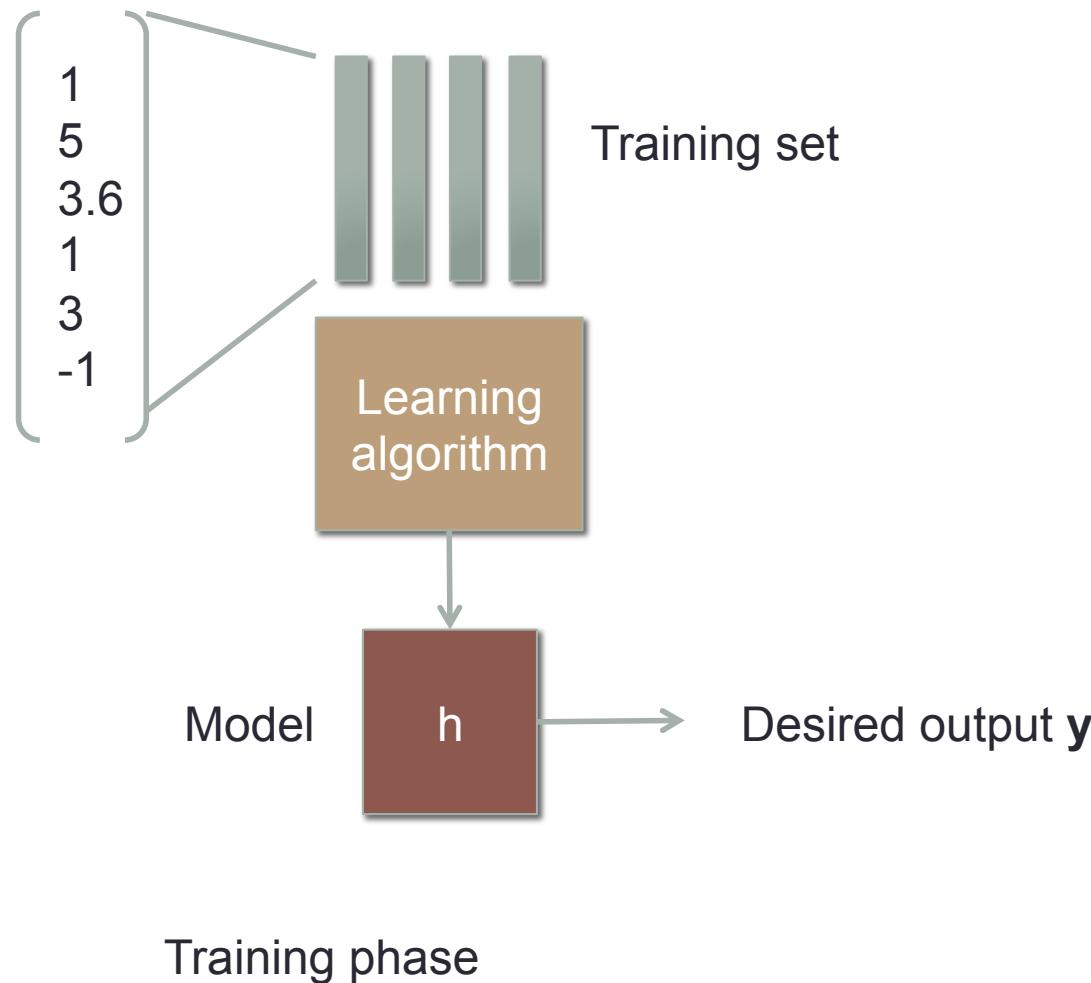
the barrel, 'bring up alongside
and lay me out a bottle of grog
here a hot,' he continued. 'I'm a
plain man; rare ham and bacon and
eggs is what I want, and that head
on a board, and a bottle of grog.'

'What you might call me? You
might call me captain. Oh, I
see what you're at — there! and
he pointed to the three or four gold
pieces on the thwart. 'You can
tell me when I've worked through
that,' says he, looking as fierce as
a commoner.

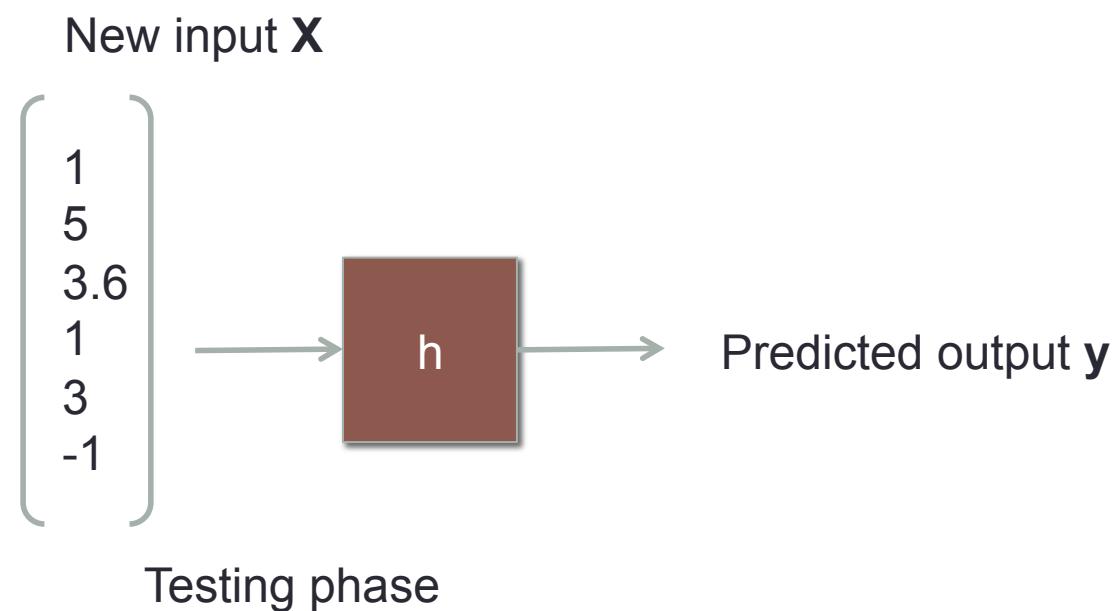
And indeed that as his clothes
were and coarse as he spoke, he
had none of the appearance of a



How do we learn from data?



How do we learn from data?



A task

The raw inputs and the desired output defines a machine learning task



Squire Trelawny, Dr. Livesey,
and the rest of these gentlemen
having asked me to write down
the whole particulars about Trese-
vant and his crew, I will do so
to the end, keeping nothing back
but the bearings of the island;
and that only because there is still
more than one year to take up
my pen in the year of grace 17—
and go back to the time when my
father kept the Admiral Benbow
inn, and how the crew, who
with the salvo cut fine took up his
lodging under our roof.

I remember him as if it were
yesterday; a tall, gaunt pibbledy
to the door, his sword-hilt
following behind him in a hand-
barrow, a tall, strong, heavy,
and brown man; his tarry plaid
faded and shabby; his
soiled blue coat, his hands ragged
and scoured, with black, broken

nails, and the salvo-cut across
one cheek, a dirty, red white. I
remember him looking round the
cockpit and whistling to himself as
he did when he was a boy, singing
in that old sea-song that he sang
so often afterwards:

'Fifteen men on the dead man's
chest—Yo-ho-ho, and a bottle of
grog—Ship, Mah company, mate!²⁷
My father told him so, very
little company, the more was the
pleasure.'

'Well, then,' said he, 'this is the
birth for me. Here you, money,'

he cried to the man who trundled

the barrel, 'bring up alongside

and lay it here, and lay me
here a hot,' he continued. 'I'm a
plain man; rye and bacon and
eggs is what I want, and that head
of mine is burning like a fire.'

'What you might call me? You
might call me captain. Oh, I
see what you're at—there,' and

he pointed to the three or four gold

pieces on the thwart. 'You can
tell me when I've worked through
that,' says he, looking as fierce as

a carbine.'

And indeed that as his clothes
were and coarse as he spoke, he
had none of the appearance of a

'This is a handy cove,' says he
at length; 'and a pleasant sitty and



Magic

→ Predicted output y



Predicting After You stock
price with CCTV image,
facebook posts, and daily
temperature



Key concepts

- Feature extraction
- Evaluation

Feature extraction

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



Squire Trelawny, Dr. Livesey, and the rest of the gentlemen had been there to see me off, and the whole particulars about Treas are bland, from the beginning to the end, keeping nothing back but the name of the ship, and that only because there is still treasure yet to find. I take up my pen in the year of grace 17— and I will tell you all about it. My father kept the Admiral Venturer arm and the brown old seaman with the salter cut first took up his lodgings at the Admiral.

I remember him as if it were yesterday, as he came plodding to the inn door, his sea chest following him, his hands bound, a tall, strong, heavy-set brown man, his tarry pigtail falling over the shoulder of his sodden blue coat, his hands rugged and scarred, with black, broken

nails, and the salter cut across one cheek, a dirty, lard-worn cover-sore between the two, the cover and wheeling to himself as he did so, and then breaking out that old sea-song that he sang all the afternoon.

Yiffer men on the dead man's when No-As-ho, and a hand of them the high old sea-veteran who had been twice at the Cape, and had been turned and fended at the captain boxes. Then he rapped on the door with a bit of stick like a handspike, and carried on when he was ushered in, coffee-soup for a glass of rum. This, when it was brought to him, he drank slowly, his fingers interlaced, his eyes looking through the time and still holding about him at the cliffs and up at our signboard.

"This is a handy cove," says he at length; "and a pleasant abode.

shop. Much company, mate?"

My father told him so, very fine company, the more was the D.R.

data1



data2

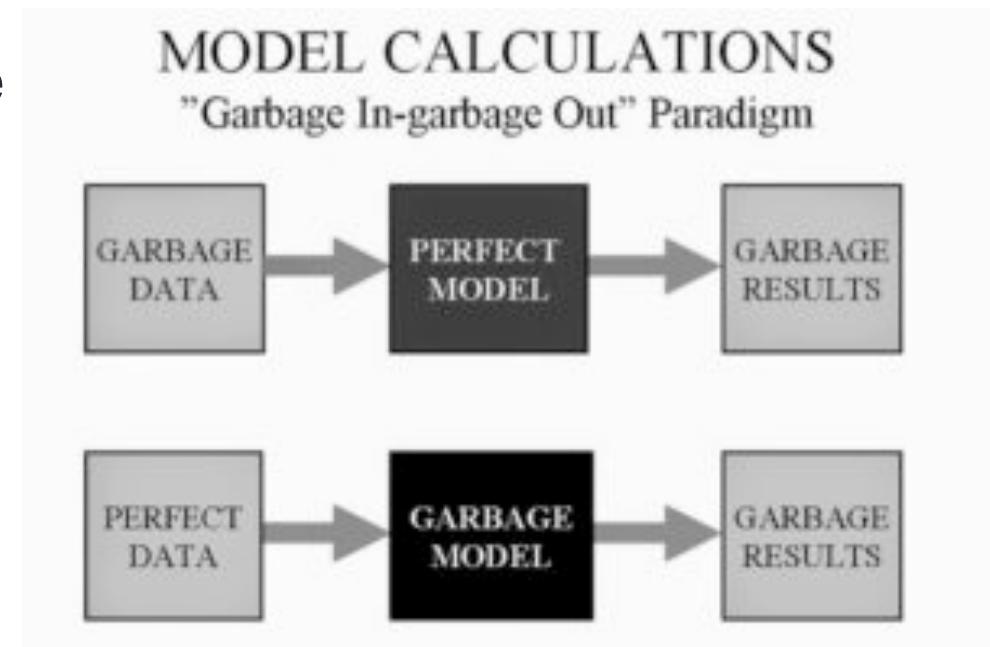


data3



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



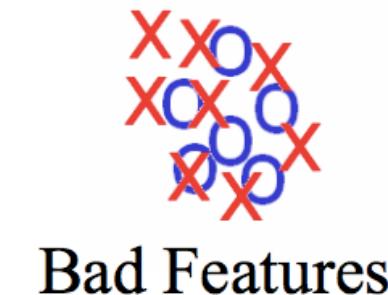
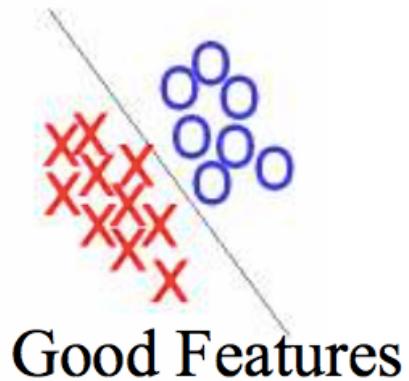
The need for data cleaning



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

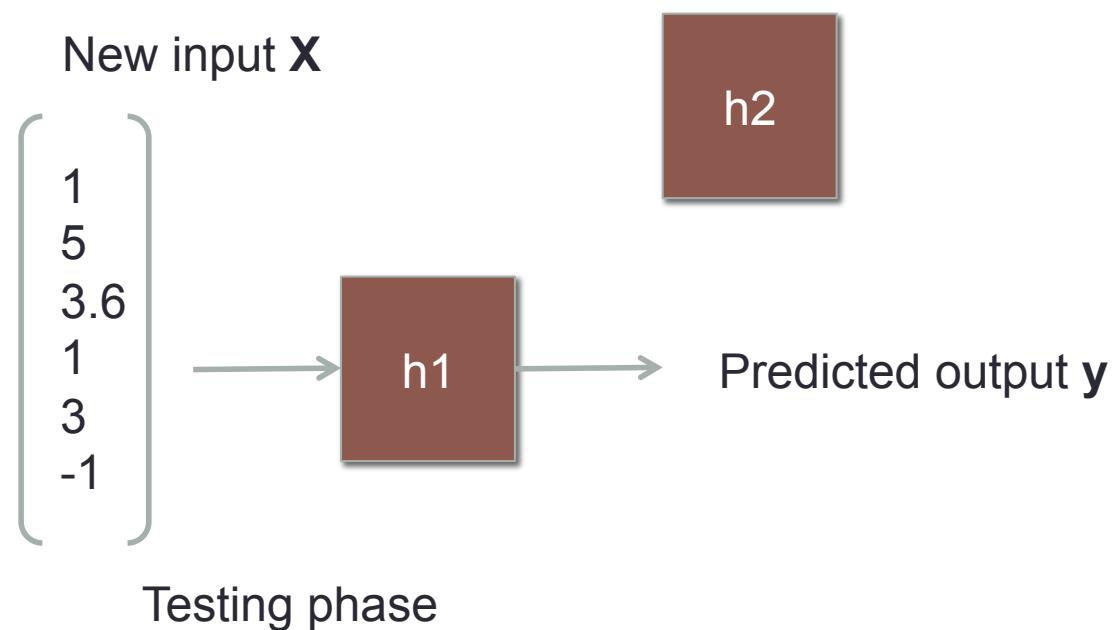
Feature properties

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



Model evaluation

How to compare h_1 and h_2 ?



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?

ໃບໃໝ່

Model A: Where are you going?

Model B: Where to?

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

Metrics consideration 1

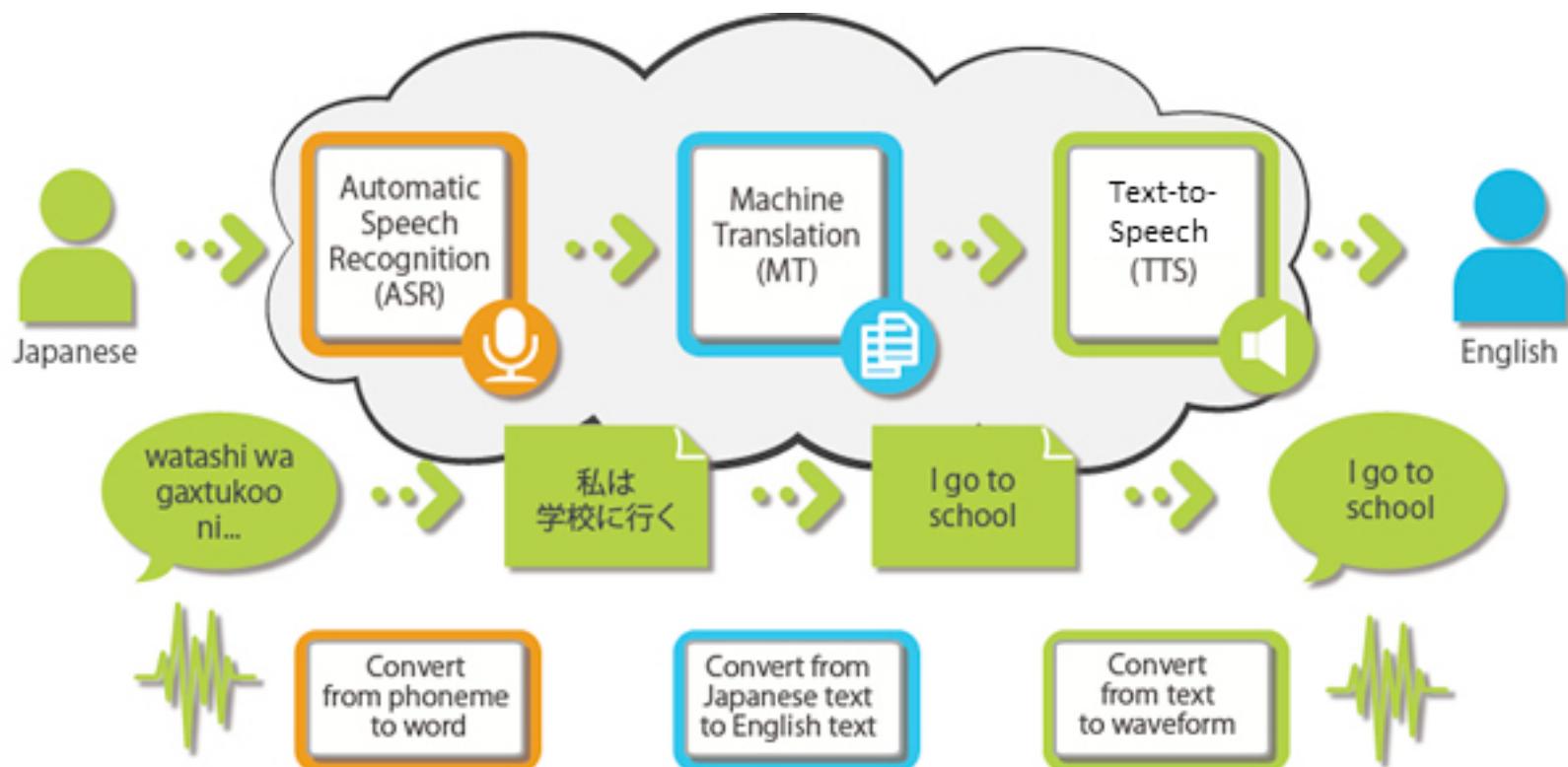
- Are there several metrics?



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

Metrics consideration 2

- Are there sub-metrics?



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.

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

The screenshot shows a search results page with the following results:

- Camera Calibration Toolbox for Matlab**
of a Camera Calibration Toolbox for Matlab with a complete ...
This document may also be used as a tutorial on camera ...
<http://www.vision.ee.ntu.edu.tw/~tak/calib.html> - 14K - Cached
- Omnidirectional Vision and Camera Networks**
not longer than six (6) pages including figures and references, should be ...
era-ready (IEEE 2-column format of single-spaced ...
<http://www.vision.ee.ntu.edu.tw/~tak/> - 5K - Cached
- Camera Calibration Toolbox for Matlab**
Toolbox from the Institute of Robotics and Mechatronics, Germany -
CR Calib is a very complete tool for camera ...
<http://www.vision.ee.ntu.edu.tw/~tak/calib.html> - 10K - Cached
- Omnidirectional Vision**
2000th Workshop on Omnidirectional Vision, Camera ... Automatic ...
g Omnidirectional and Active Cameras of the FRTF Lab, ...
<http://www.vision.ee.ntu.edu.tw/~tak/> - 25K - Cached
- Characteristics**
Know your camera characteristics if you intend to make full use of all of the ...
on your camera ...
<http://www.vision.ee.ntu.edu.tw/~tak/calib.html> - 11K - Cached
- Introduction of PMD-Cameras and Stereo-Vision for the Task of ...**
Videocon Abstract - www.vision.ee.ntu.edu.tw/~tak/
D cameras is discussed quantitatively and ... The stereo system as well as ...
will be compared in section 4 based on those ...
<http://www.vision.ee.ntu.edu.tw/~tak/intro.pdf>

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 re-visit 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

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

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

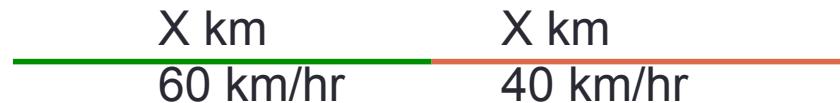


Harmonic mean vs Arithmetic mean

- 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

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

- Total distance covered 2X



Harmonic mean vs Arithmetic mean

- For the arithmetic mean to be valid you need to compare 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

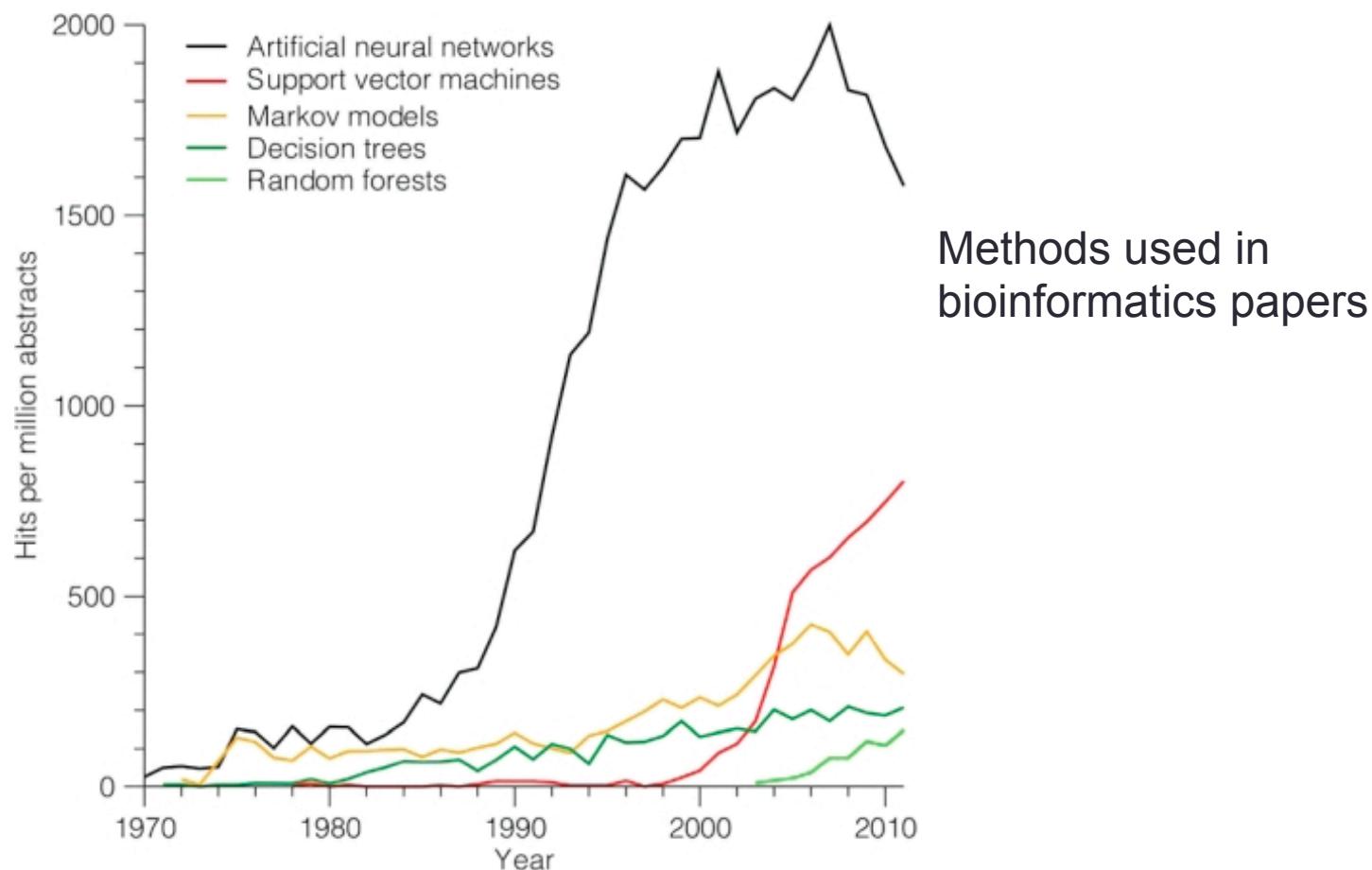
- Training time
- Testing time
- Memory requirement
- Parallelizability
- Latency

Course walkthrough

1 - 15/8	Introduction
2 - 22/8	Applications in Bioinformatics
3 - 29/8	Regression & Jupyter Notebook demo
4 - 5/9	Probability review
5 - 12/9	MLE, MAP, and Naive Bayes
6 - 19/9	GMM and EM
7 - 26/9	Dimensionality reduction and visualization
	Midterm week - No classes
8 - 9/10	SVM
9 - 16/10	Neural network basics & Gcloud, TensorFlow, and Keras demo
10 - 23/10	CNN, Recurrent architectures
11 - 30/10	Recent Advances in NN
12 - 7/11	Reinforcement Learning
13 - 14/11	Unsupervised methods
14 - 21/11	Tricks of the trade: machine learning in the real world
15 - 28/11	Project presentation

Why anything else besides deep learning

- The rise and fall of machine learning algorithms



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

What we will not cover

- Random forest
- Decision trees
- Boosting
- Graphical models

Homework

- Reading assignment



<https://hbr.org/cover-story/2017/07/the-business-of-artificial-intelligence>