

Machine Learning

CS-433

Martin Jaggi & Rüdiger Urbanke
EPFL

Course Logistics

Assessment

- ✿ **Project 1 (10%), due Oct 30th**
- ✿ **Project 2 (30%), due Dec 21st**
- ✿ **Final exam (60%)**

Course Logistics

Lectures

tentative
schedule

Date	Topics Covered	Lectures	Exercises	Projects
19/9	Introduction		01a,01b	
21/9	Linear Regression, Cost functions		Lab 1	
26/9	Optimization			
28/9	Optimization		Lab 2	
03/10	Least Squares, ill-conditioning, Max Likelihood			Project 1 details
05/10	Overfitting, Ridge Regression, Lasso		Lab 3	
10/10	Cross-Validation			
12/10	Bias-Variance decomposition		Lab 4	
17/10	Classification			
19/10	Logistic Regression		Lab 5	
24/10	Generalized Linear Models			
26/10	K-Nearest Neighbor			Q&A for proj.
31/10	Support Vector Machines			Proj. 1 due 30.10.
02/11	Kernel Regression		Lab 7	
07/11	Unsupervised Learning, K-Means			Project 2 details
09/11	K-Means, Gaussian Mixture Models		Lab 8	
14/11	Mock Exam			
16/11	Gaussian Mixture Models, EM algorithm		Mock exam	
21/11	Matrix Factorizations			
23/11	Text Representation Learning		Lab 10	
28/11	SVD and PCA			
30/11	SVD and PCA and Neural Networks - Basics		Lab 11	
05/12	Neural Networks - Representation Power			
07/12	Neural Networks - Backpropagation, Activation Functions			Q&A for proj.
12/12	Neural Networks - CNN, Regularization, Data Augmentation, Dropout			
14/12	Graphical Models -- Bayes Nets		Lab 13	
19/12	Graphical Models -- Factor Graphs			
21/12	Graphical Models -- Inference and Sum-Product Algorithm		Lab 14	Project 2 due 21.12.

Course Logistics

Lectures

Tuesday 17:15 - 19:00, Room: [CO1](#)

Thursday 16:15 - 18:00, Room: [SG1](#)

Clicker:

For some active participation in the lectures, please point your browser to this
[SpeakUp room](#)

We provide PDF lecture notes on our webpage and GitHub

Teaching material mostly by Emtyiaz Khan, Rüdiger Urbanke, Martin Jaggi

(SpeakUp room)

speakup.info

Room Nr. 87054

Q1: Are you new to EPFL this semester

- A. yes
- B. no

(SpeakUp room)

speakup.info

Room Nr. 87054

Q2: You are from

- A. IC (Comp Sci.)
- B. SB (Basic Sciences)
- C. STI (Engineering)
- D. SV (Life Sciences)
- E. ENAC (Arch., Env. Sci.)
- F. CDM (Management)
- G. CDH (Humanities)
- H. Exchange Student
- I. others

Course Logistics

Exercises

Thursday 14:15 - 16:00

Rooms: [INF119](#) (A-Ch), [INJ218](#) (Ci-Kj), [INM11](#) (Ko-Pe), [INM202](#) (Pi-Z),
according to last-name.

All labs and projects will be in **Python**.

See the first lab to get started.

Code Repository for Labs: github.com/epfml/ML_course

Course Logistics

TA team

contact us: epfmlcourse@gmail.com

Barbier Jean

Kamireddy Sai

Liu Wei

Lu Jun

Maksai Andrii

Zhou Ruofan

Ajalloeian Ahmad

Benyahia Yassine

Borgeaud dit Avocat William

Castellon Arevalo Joel

Champenois Bertrand

Gucevska Natalija

Kunstner Frederik

Moreau Hugo

Course Logistics

Projects

- **Project 1 (10%), due Oct 30th**
- **Project 2 (30%), due Dec 21st**

Real-world problems, Python, Groups of 3 Students

(SpeakUp room)

speakup.info

Room Nr. 87054

Q3: Programming Experience

sect.	avg
MA	5.36
SV	4.98
SC	4.73
IN	4.67

- A. **some in Python**
- B. **some in another language**
- C. **none**

What to expect?

- ❖ **overview over ML**
- ❖ **basic understanding of most important ML methods and concepts**
- ❖ **experience how ML is done on a practical problem**

What not to expect?

- ✿ becoming an **expert** data scientist / ML researcher / statistician / large-scale ML hacker / ML software expert / domain expert
- ✿ learn about this week's newest **hack in deep learning** you saw on blog *xyz*.
- ✿ build your own **self-flying** helicopter

Introduction

What is Machine Learning?

What is Machine Learning?

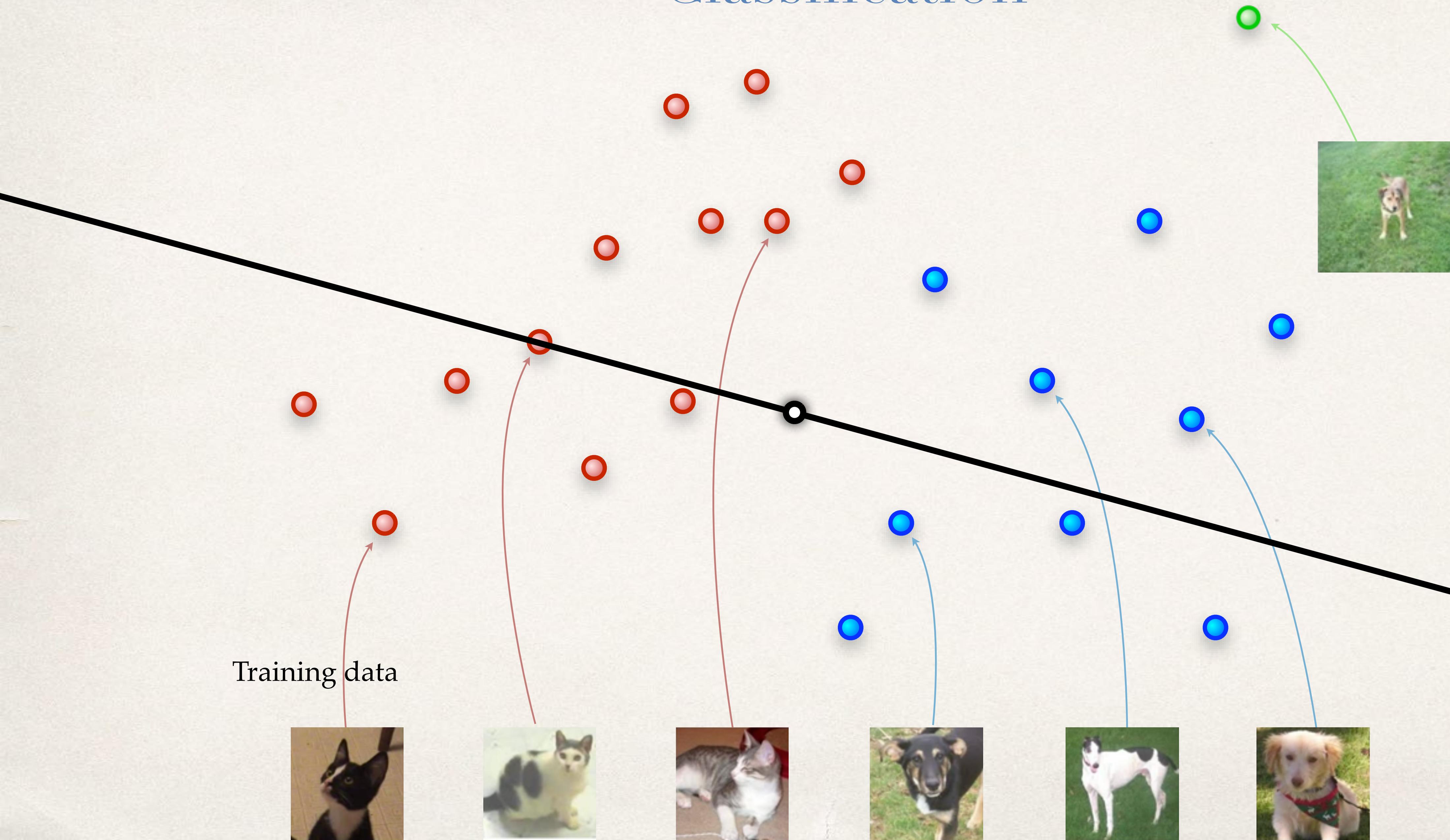
algorithms that can

learn from data

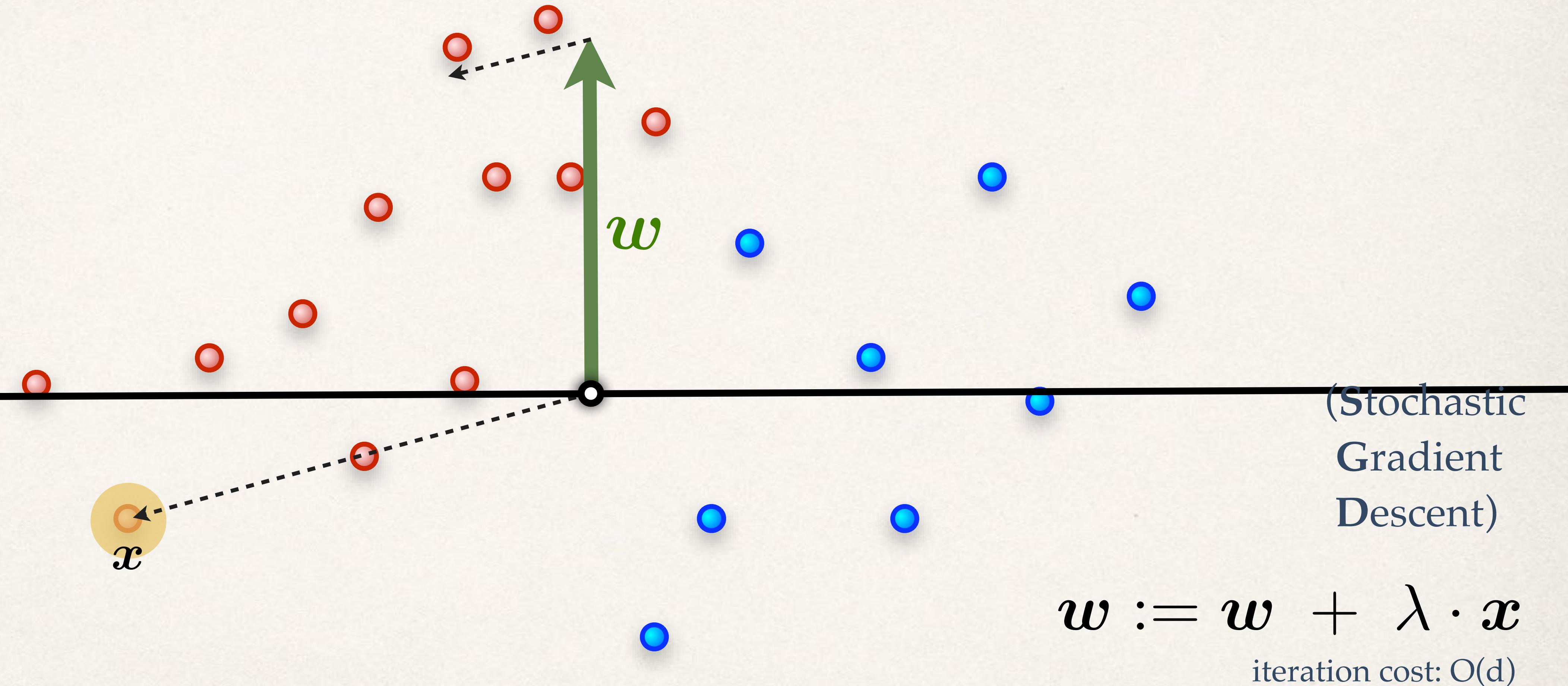


[image source](#)

Classification



The Learning Algorithm



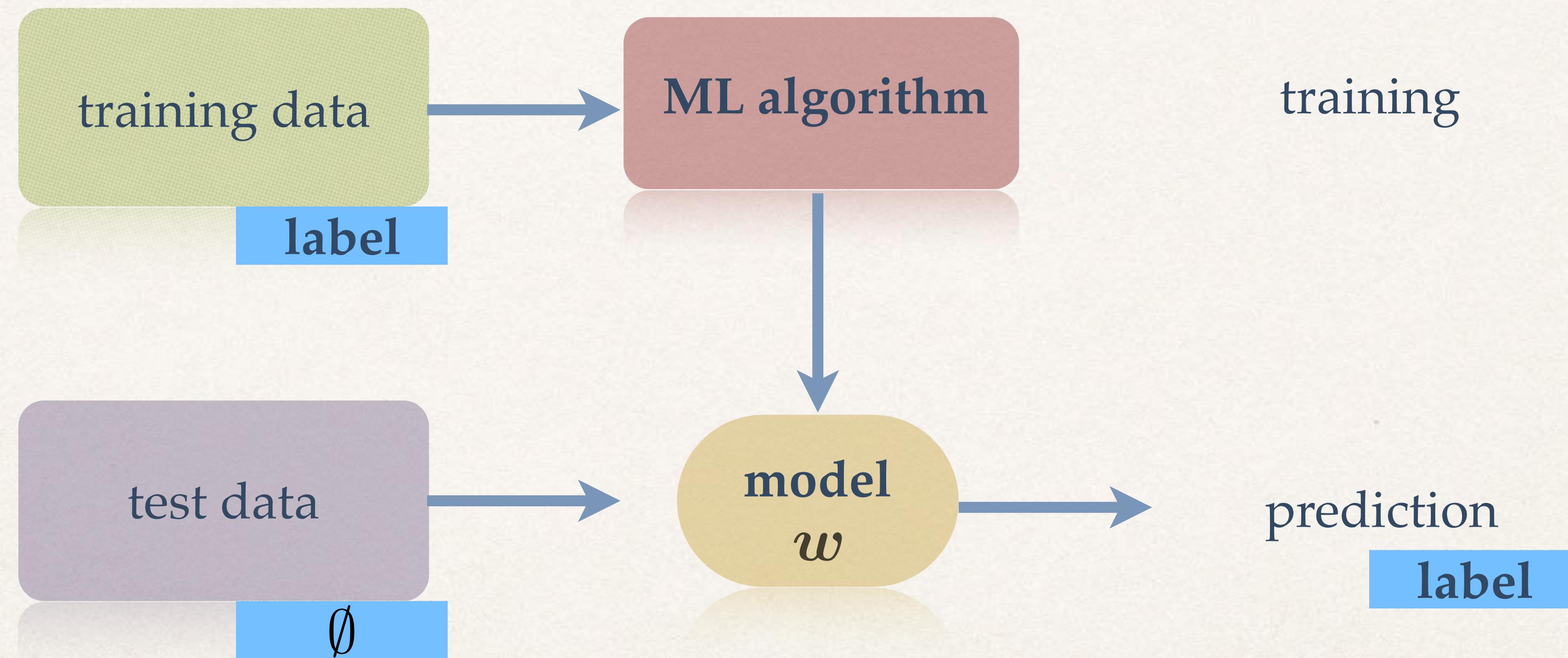
Perceptron

(Rosenblatt 1957)

Support-Vector-Machine

(Cortes & Vapnik 1995)

Machine Learning Fundamentals



what Evaluation Metric?



towards...
understanding intelligence
?

if-then-else

\neq

intelligence



towards...
understanding intelligence
?

Machine Learning

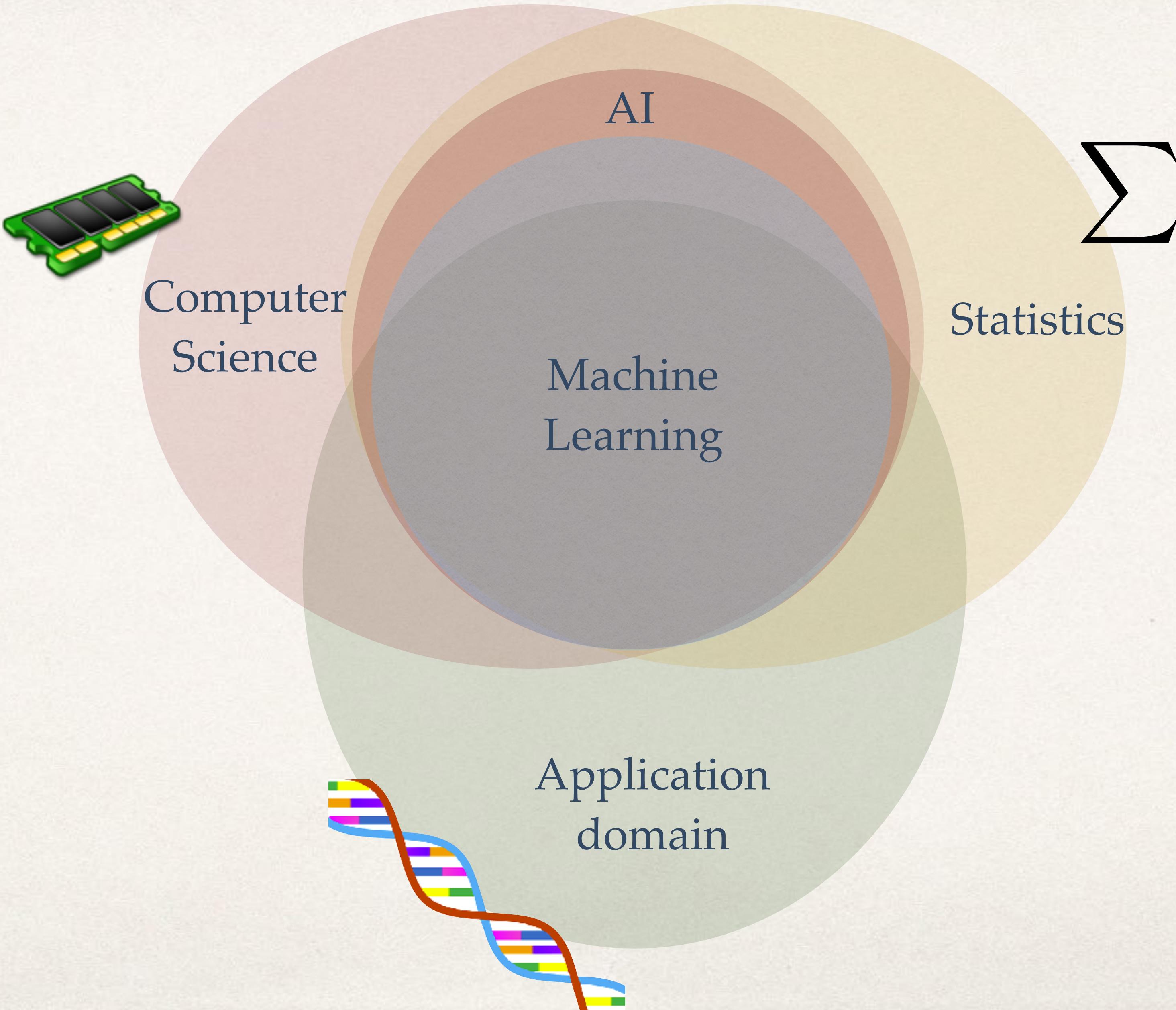


vs

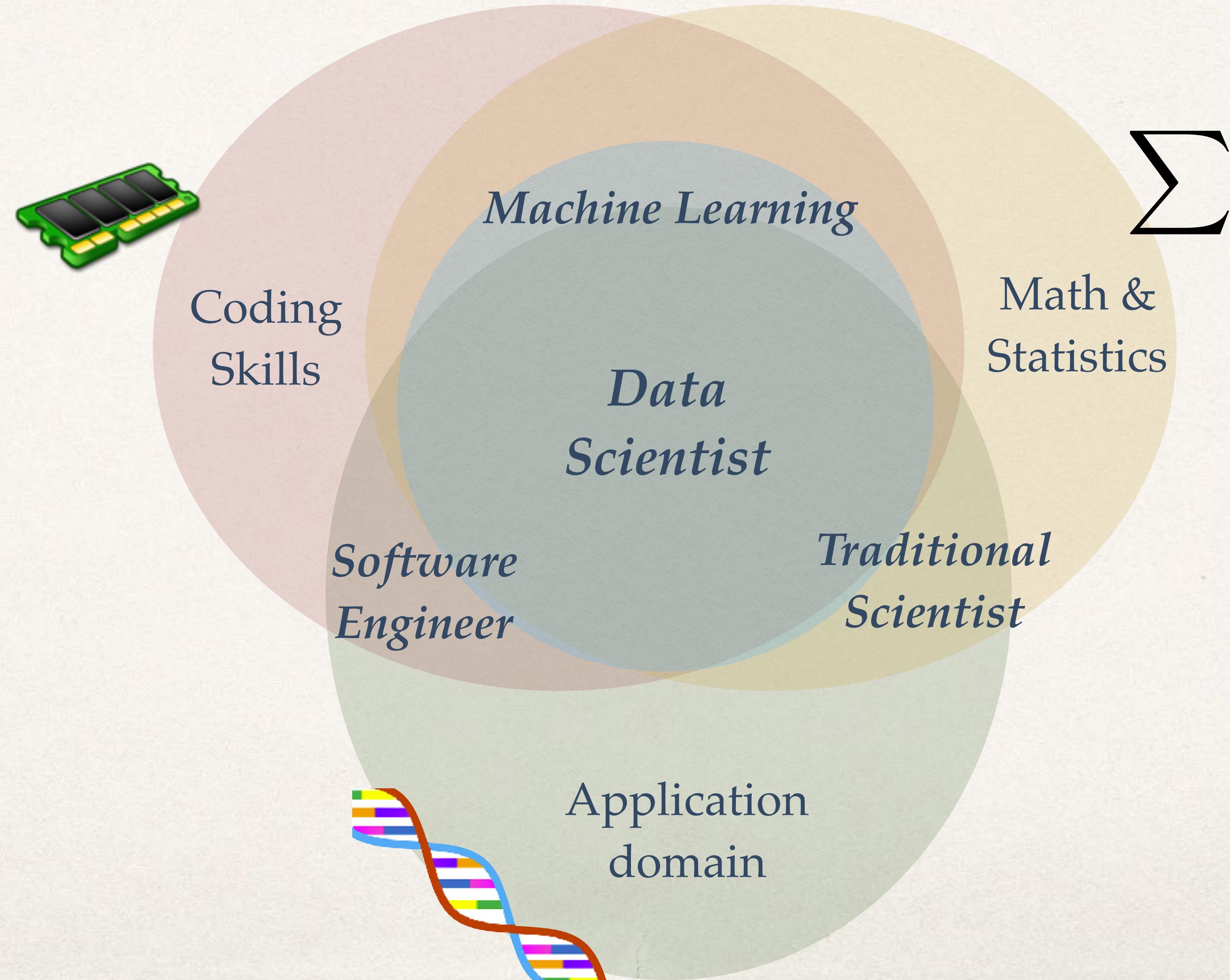
Neuroscience / HBP



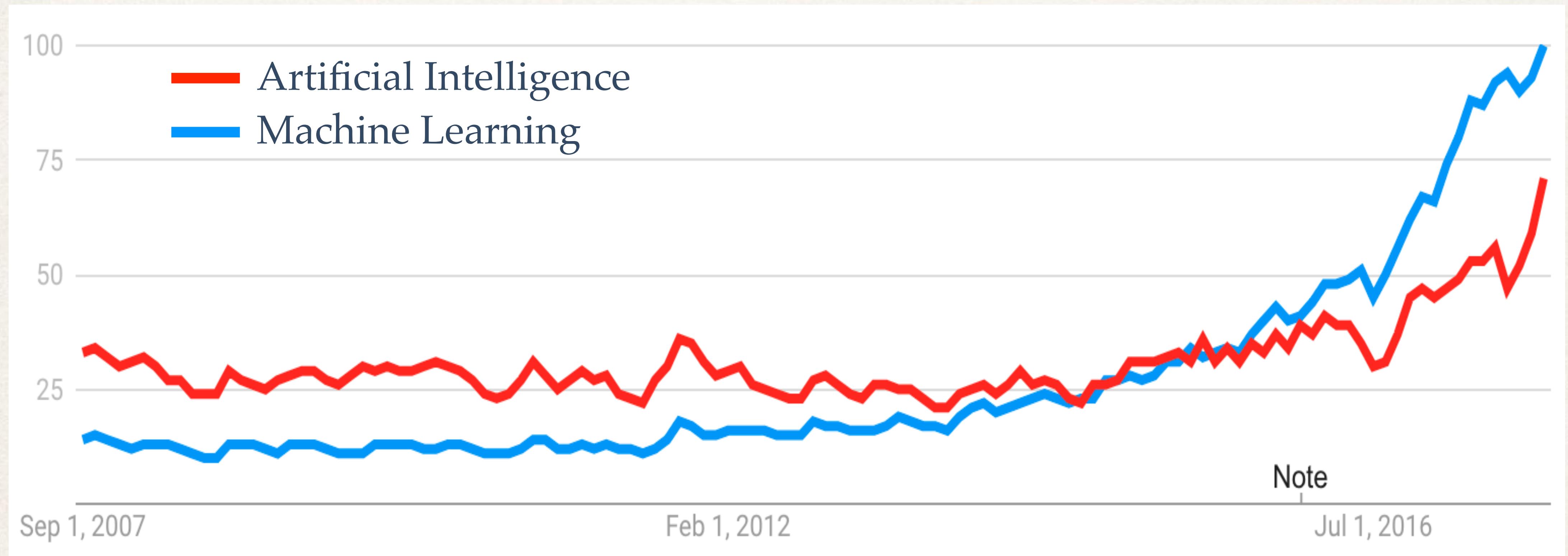
What is the difference between
Artificial Intelligence, Data Mining,
Statistics, Machine Learning?



Job Skills



Evolution of Interest



source

why ML?

Applications

Industry Applications

- ✿ not only the
'usual suspects'



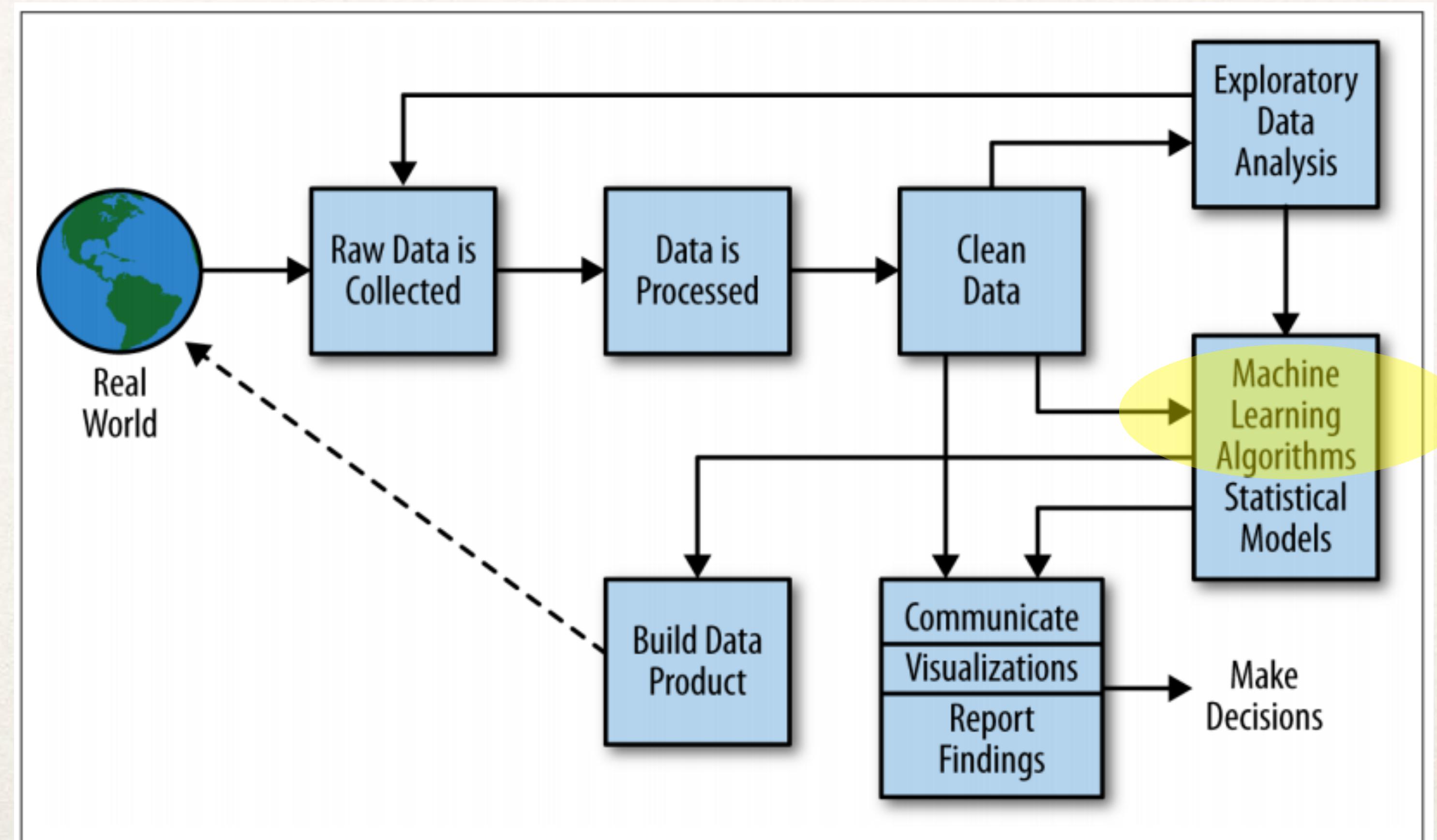
- ✿ majority of industries, originally not 'digital':
 - ✿ agriculture, car & bike sharing, booking.com, manufacturing, sports, personalized health, call centers, ...

Applications in Other Sciences

- ✿ increasingly data driven
 - ✿ ... Psychology, Economy, Social sciences
 - ✿ science of X → *digital* science of X

en.wikipedia.org/wiki/Index_of_branches_of_science

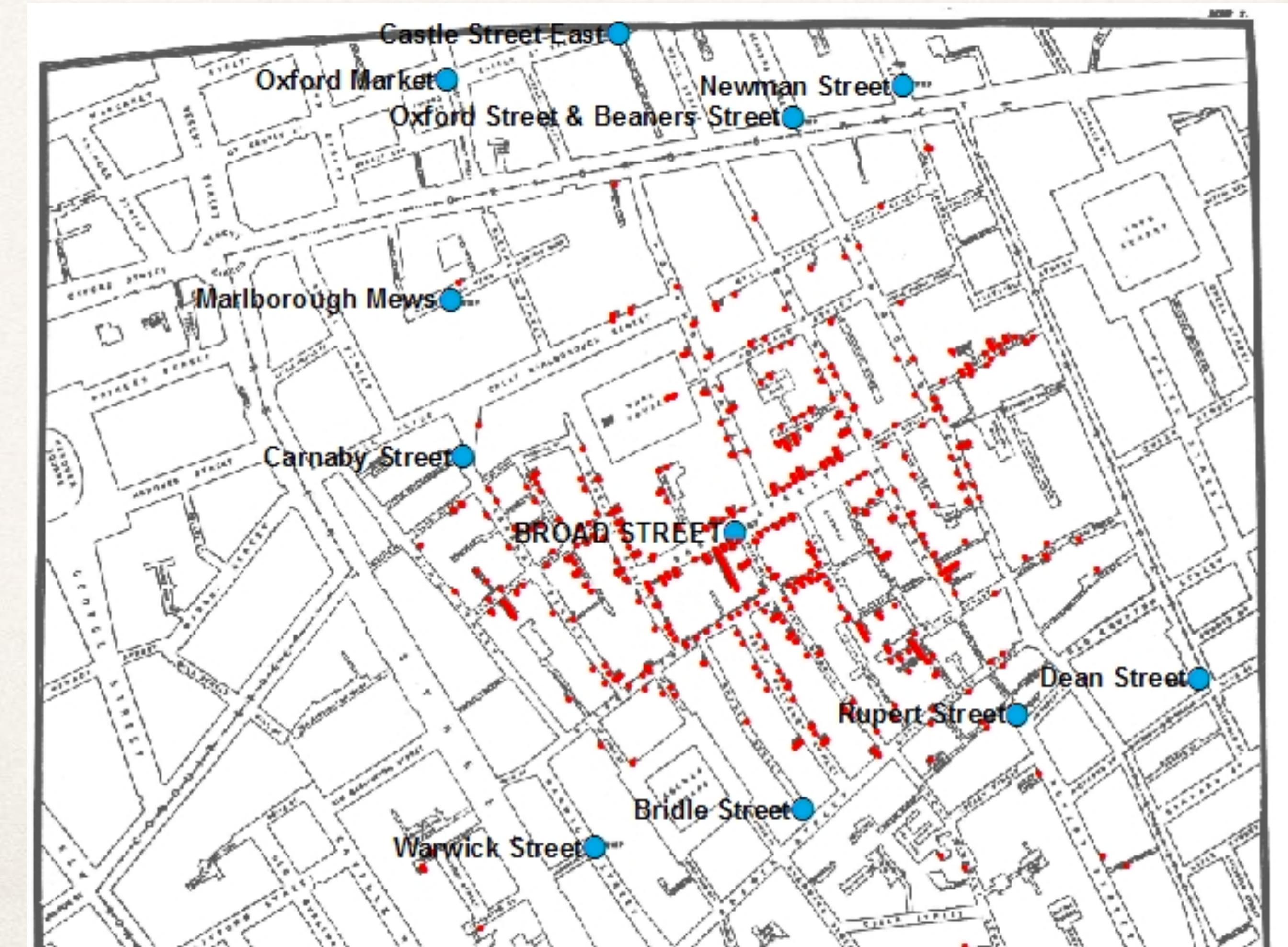
ML is only a small part!



History

- ✿ ML is not new!

London 1854
cholera outbreak



History

- ❖ ML is **not new!**
- ❖ the early days - 1950^{ies} and 1960^{ies}
 - ❖ Neural networks
 - ❖ Turing



What has changed?

1950s: 10^3 FLOPS

2017: 10^{17} FLOPS

“the embryo of an electronic computer that ... will be able to walk, talk, see, write, reproduce itself and be conscious of its existence.”

1958



Challenges

- ❖ Hype
 - ❖ cycles of AI popularity
- ❖ Data Ethics, Privacy, Fairness

[Alle Kategorien ansehen](#)Suche ▾[Sport & Freizeit](#)[Erweiterte Suche](#)[Camping & Outdoor](#)[Fitness](#)[Fußball](#)[Golf](#)[Pulsuhren & GPS](#)

R



Baseballschläger Aluminium "32" von Tysonz

von [Unbekannt](#)Noch keine Kundenrezensionen vorhanden: [Schreiben Sie die erste!](#)Preis: **EUR 21,90**

Alle Preisangaben inkl. MwSt.

Auf Lager.Verkauf und Versand durch [knockout24](#). Für weitere Informationen klicken Sie auf den Verkäufernamen.

Noch 4 Stück auf Lager.

[2 neu ab EUR 21,90](#)[Größeres Bild](#)[Für Kunden: Stellen Sie Ihre eigenen Bilder ein.](#)

Kunden, die diesen Artikel gekauft haben, kauften auch



[Pfeffer 2442 KO
Verteidigungsspray JET
40 ml](#) von Pfeffer KO

EUR 5,22



[Sturmhaube 3-loch
Acryl schwarz](#) von
Unbekannt

EUR 1,38



[Pfefferspray KO-FOG
40ML](#) von Pfeffer KO

EUR 5,19



[KO-CS
Verteidigungsspray 40
ml](#) von KO-CS

EUR 4,22

Challenges

- ✿ Hype
 - ✿ cycles of AI popularity
- ✿ Data Ethics, Privacy, Fairness
- ✿ Lack of Interpretability
 - ✿ example: medical applications of deep learning
- ✿ Social Implications,
Threats from Super-human AI?
 - ✿ see Nick Bostrom

need: Scientific Method, Reproducible
Research, Open Source and Open Data

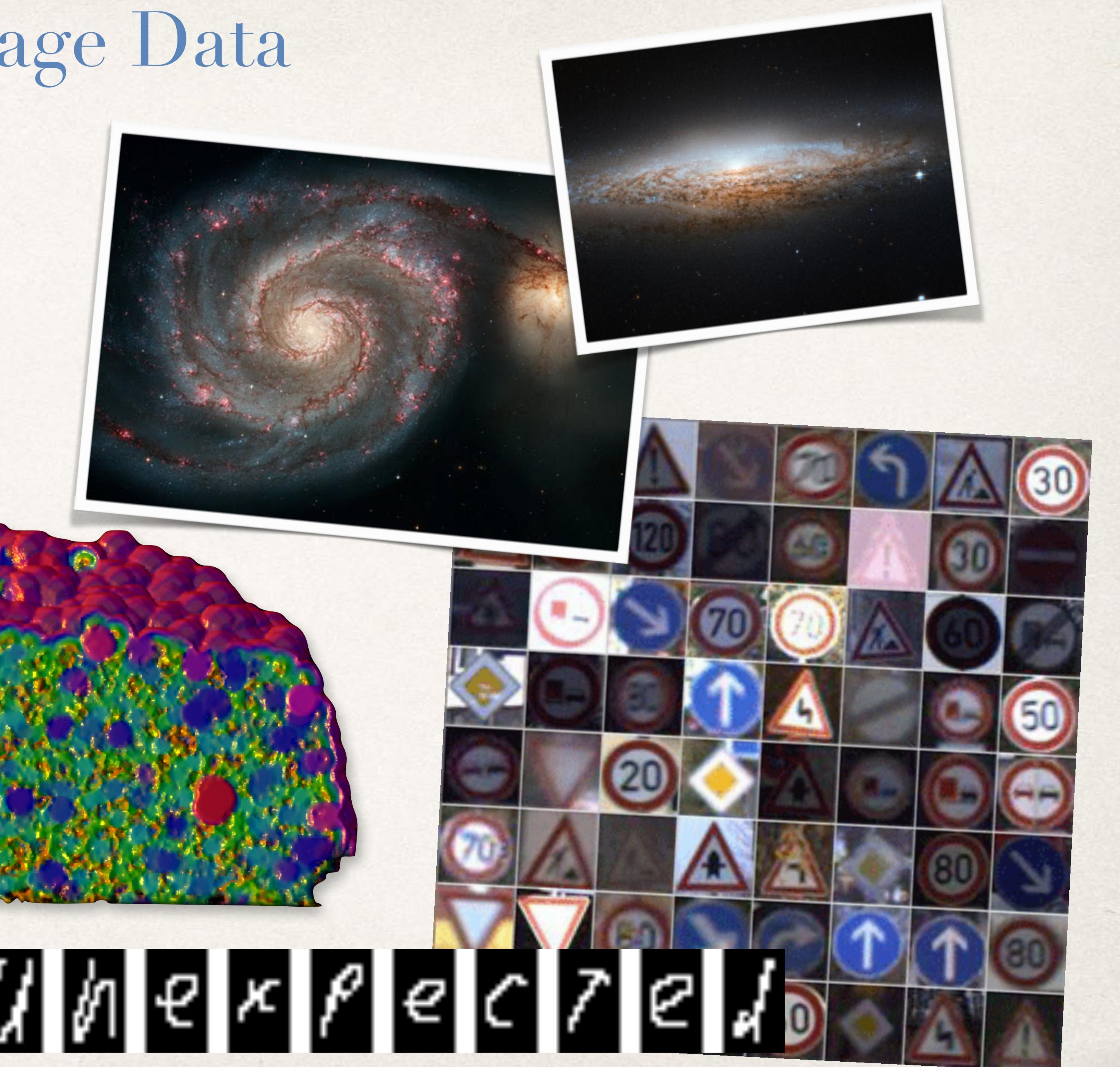
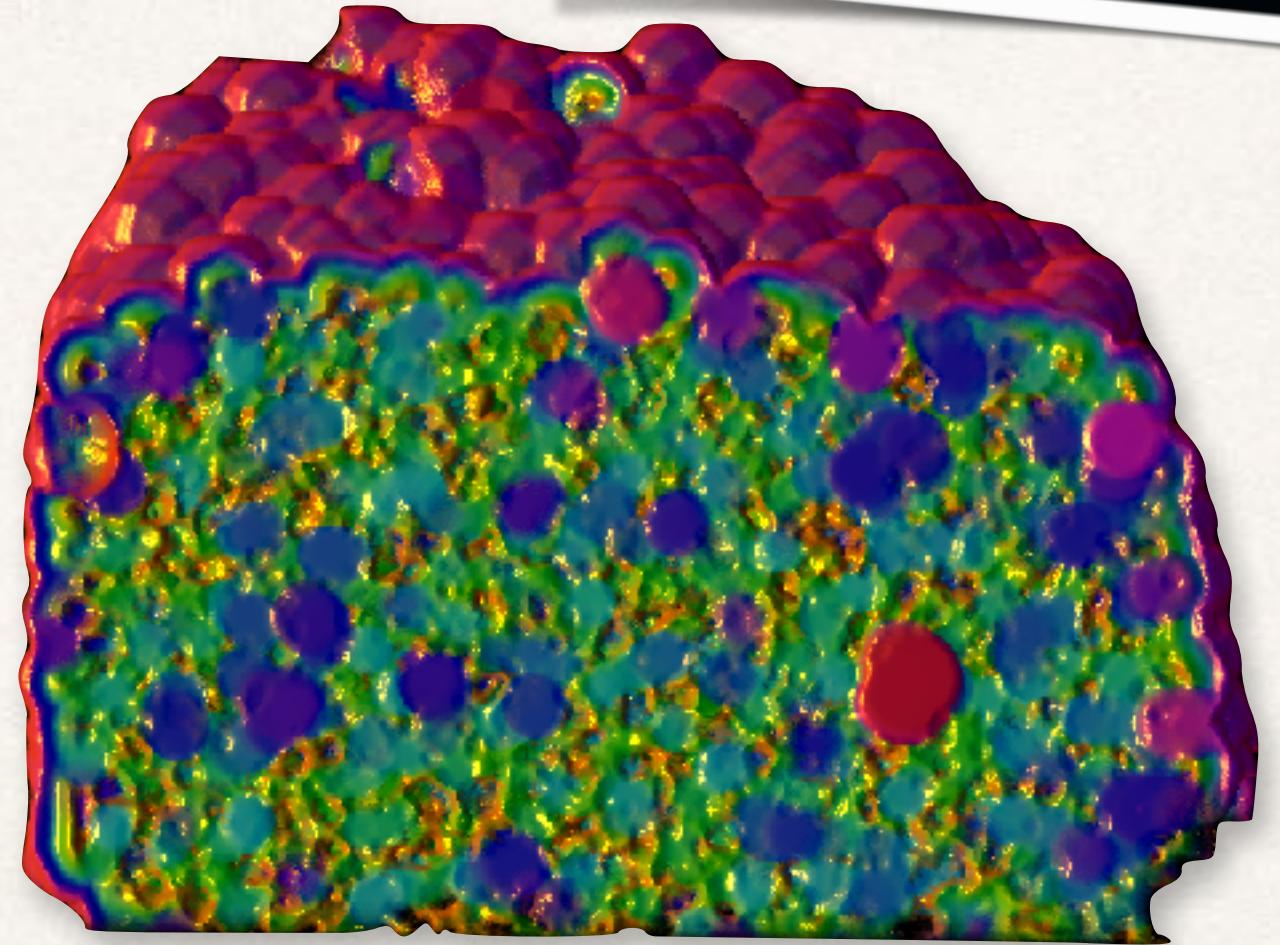
why ML?

Applications, cont.

Image Data

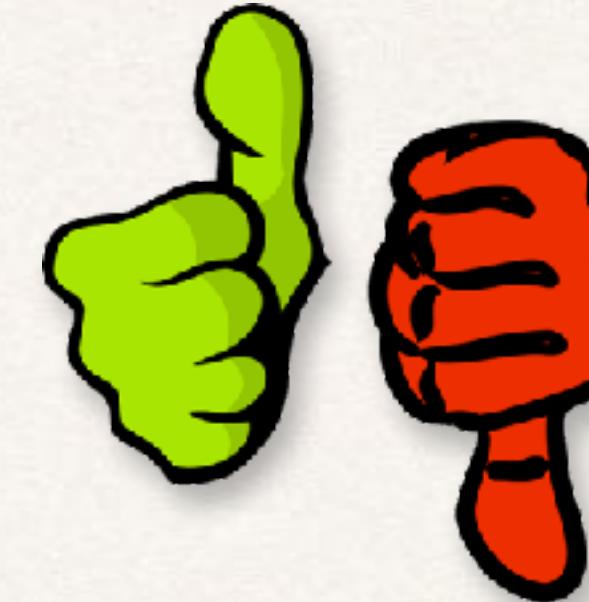
- ❖ Astronomy
- ❖ Face recognition
- ❖ 2D + 3D medical imaging
- ❖ OCR
- ❖ self-driving cars

how-old.net

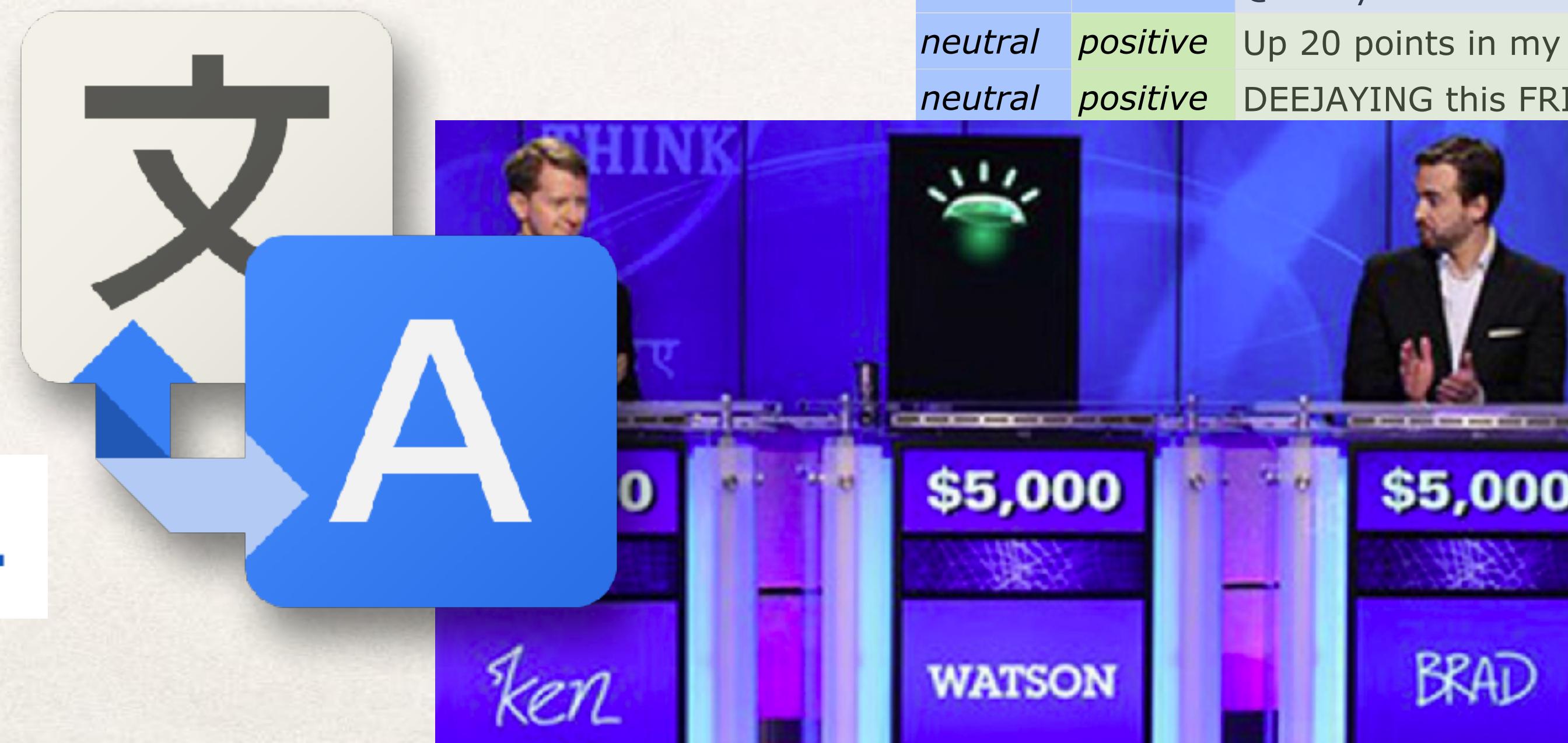


Text Data

- ✿ Spam Detection
- ✿ User Content
- ✿ Medical Data
- ✿ Machine Translation

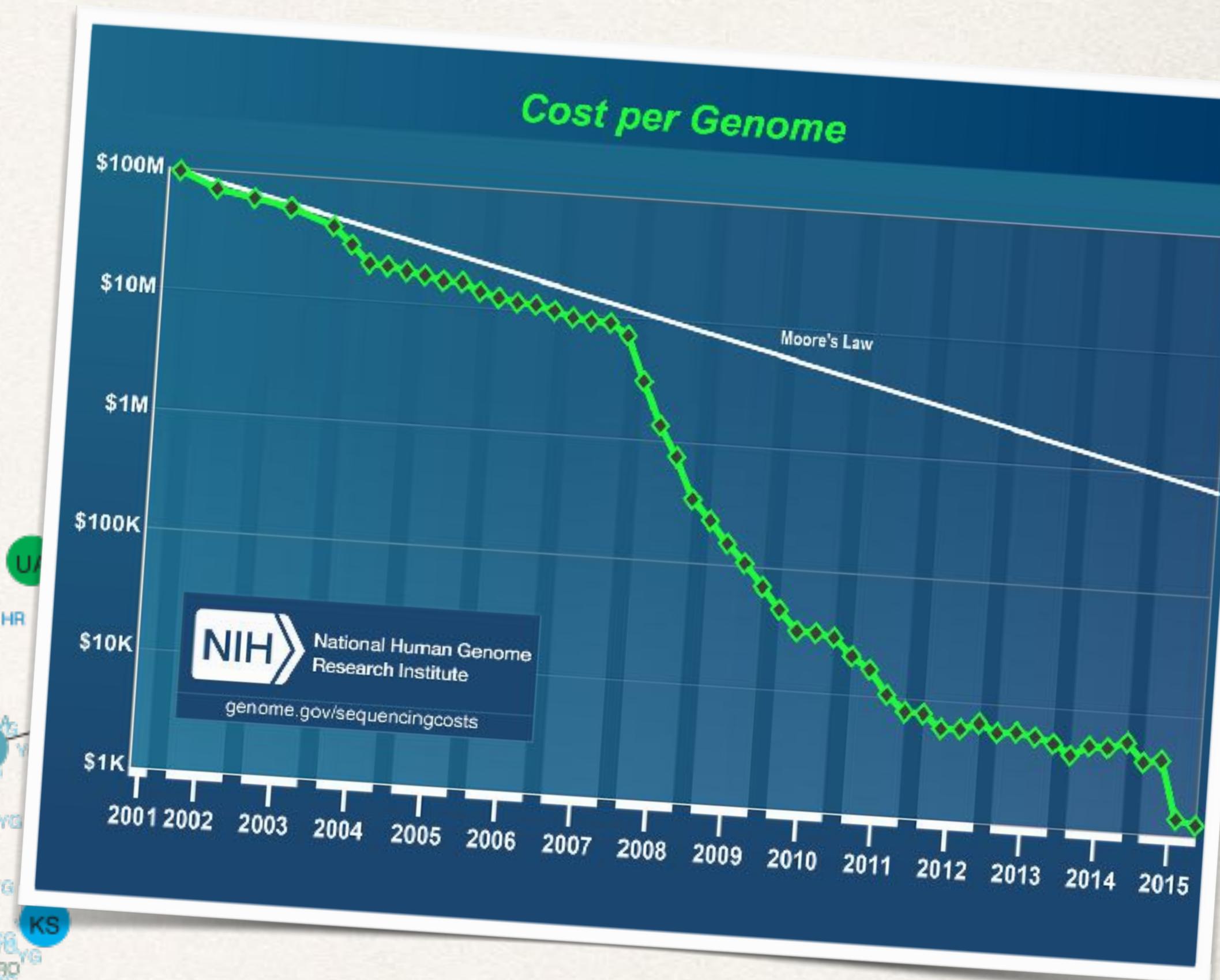
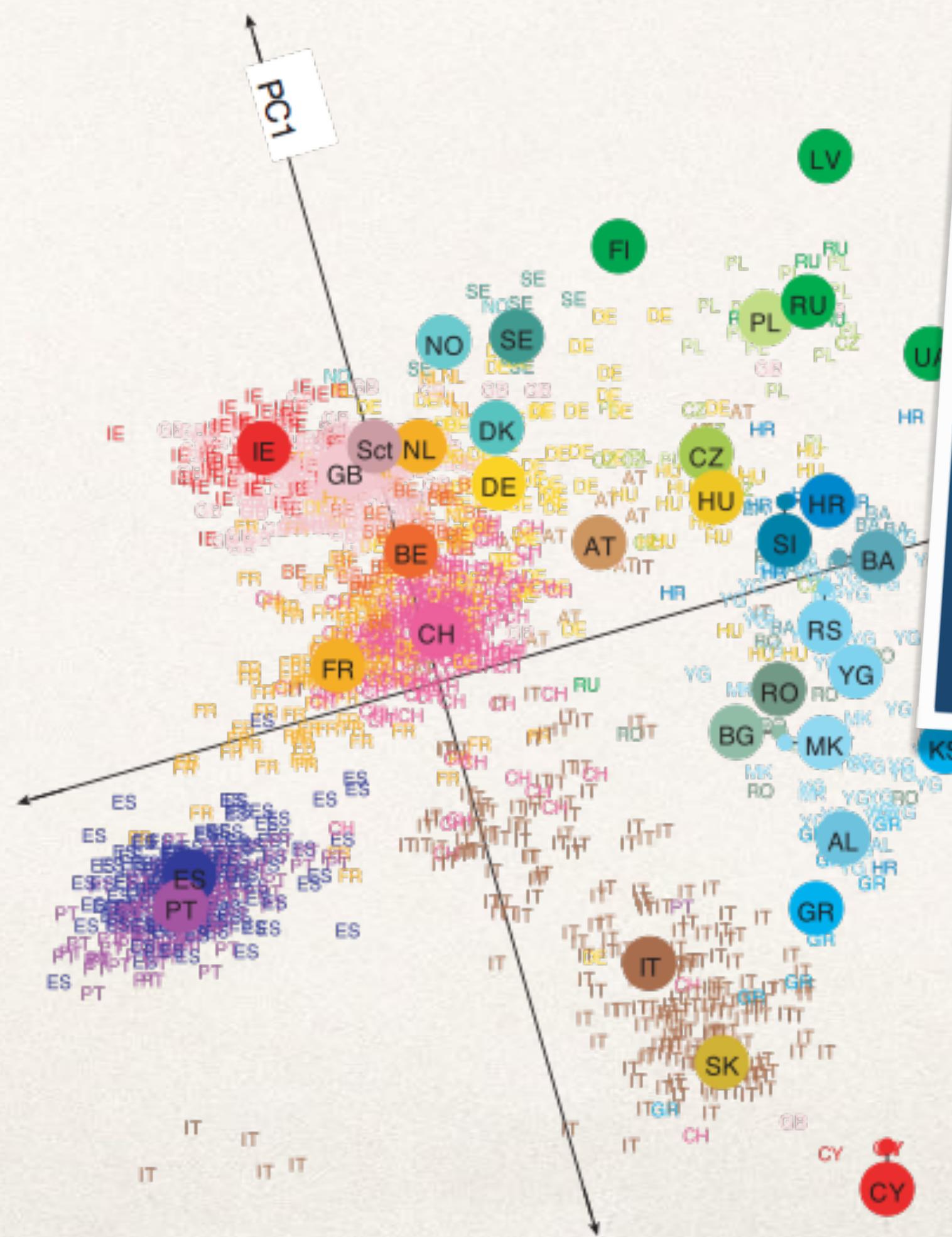
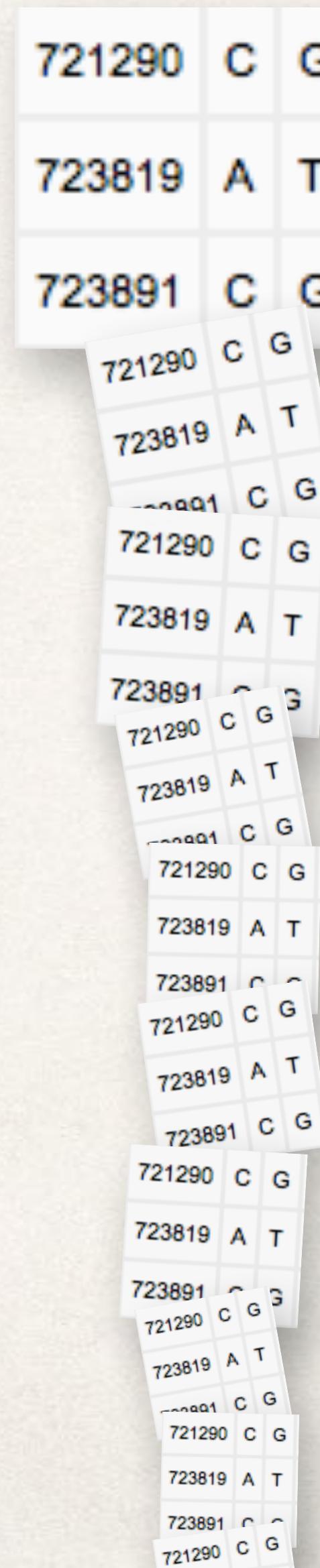


negative	neutral	But i wanna wear my Concords tomorrow though but i don't feel like it
positive	neutral	Gonna watch Grey's Anatomy all day today and tomorrow(:
negative	neutral	@CoachVac heey do you know anything about UVA's fallll fest loll they invited me
neutral	neutral	@DustyEf when that sun is high in that Texas sky, I'll be buckin it to county fair. A
neutral	positive	Up 20 points in my money league with Vernon Davis and L. Fitz still to go tomorrow
neutral	positive	DEEJAYING this FRIDAY in THE FIRST CHOP it's CHRIS actual SMITH with a smash



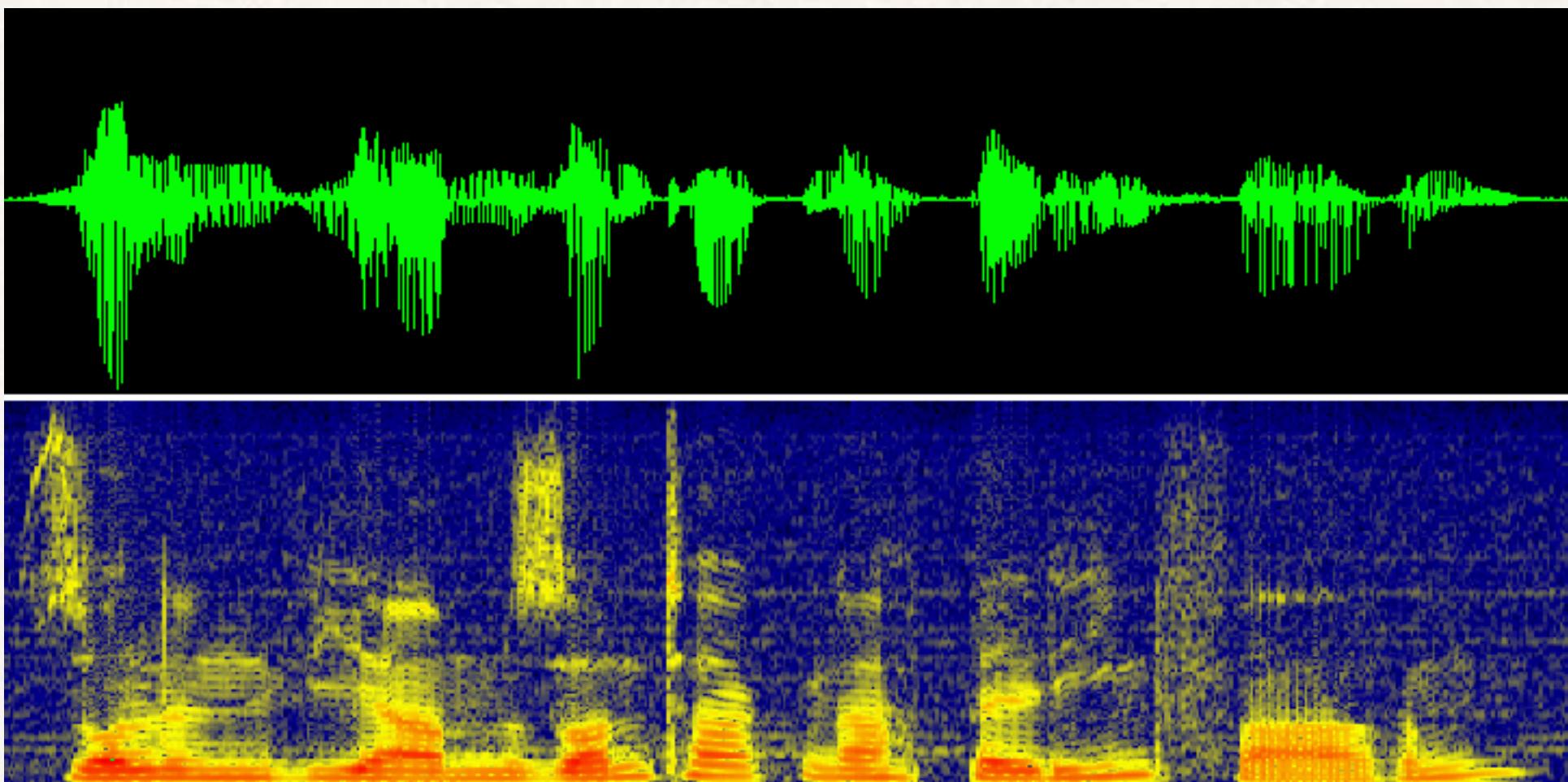
DeepL

Medical: Genetic Data



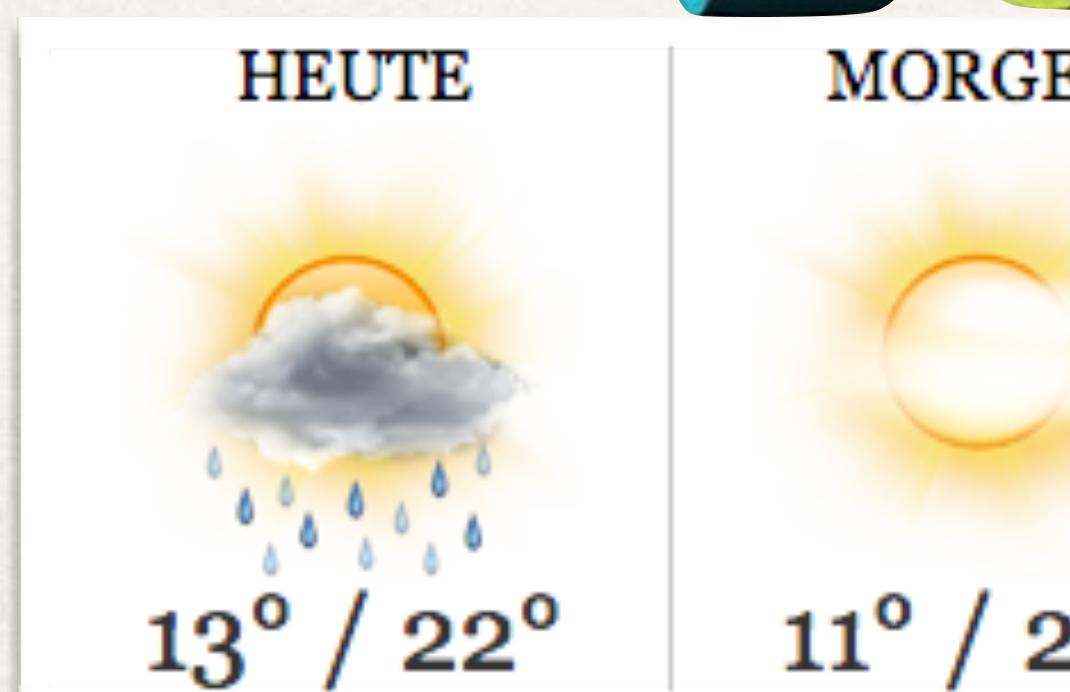
Audio Data

- ❖ Hearing aids
- ❖ Voice Recognition
- ❖ Automatic Translation



Numerical / Sensor Data

- ❖ Cern
- ❖ SKA, Telescopes
- ❖ Fitness Trackers
- ❖ Weather Forecast
- ❖ Robotics
- ❖ Kinect



Games & Simulations

- ❖ Immediate Feedback
- ❖ Chess, Go
- ❖ Physical World



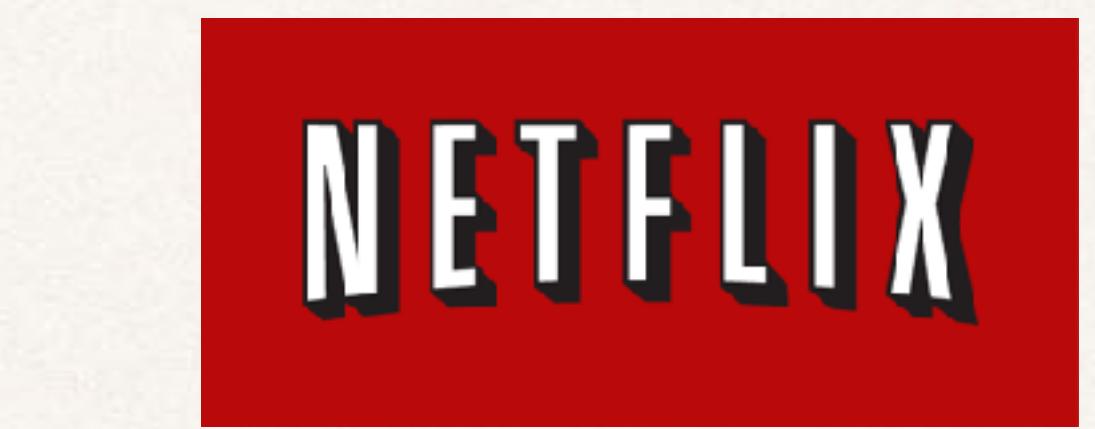
Internet Data



WIKIPEDIA



Google



amazon.com

Prediction Markets

		Products / Words			
		Customers / Words			
Customers / Words	Products / Words	★	★★		
			★		
Customers / Words	Products / Words		★★	★★	
		★★		★	★★
Customers / Words	Products / Words	★★			★★
			★★	★	★★

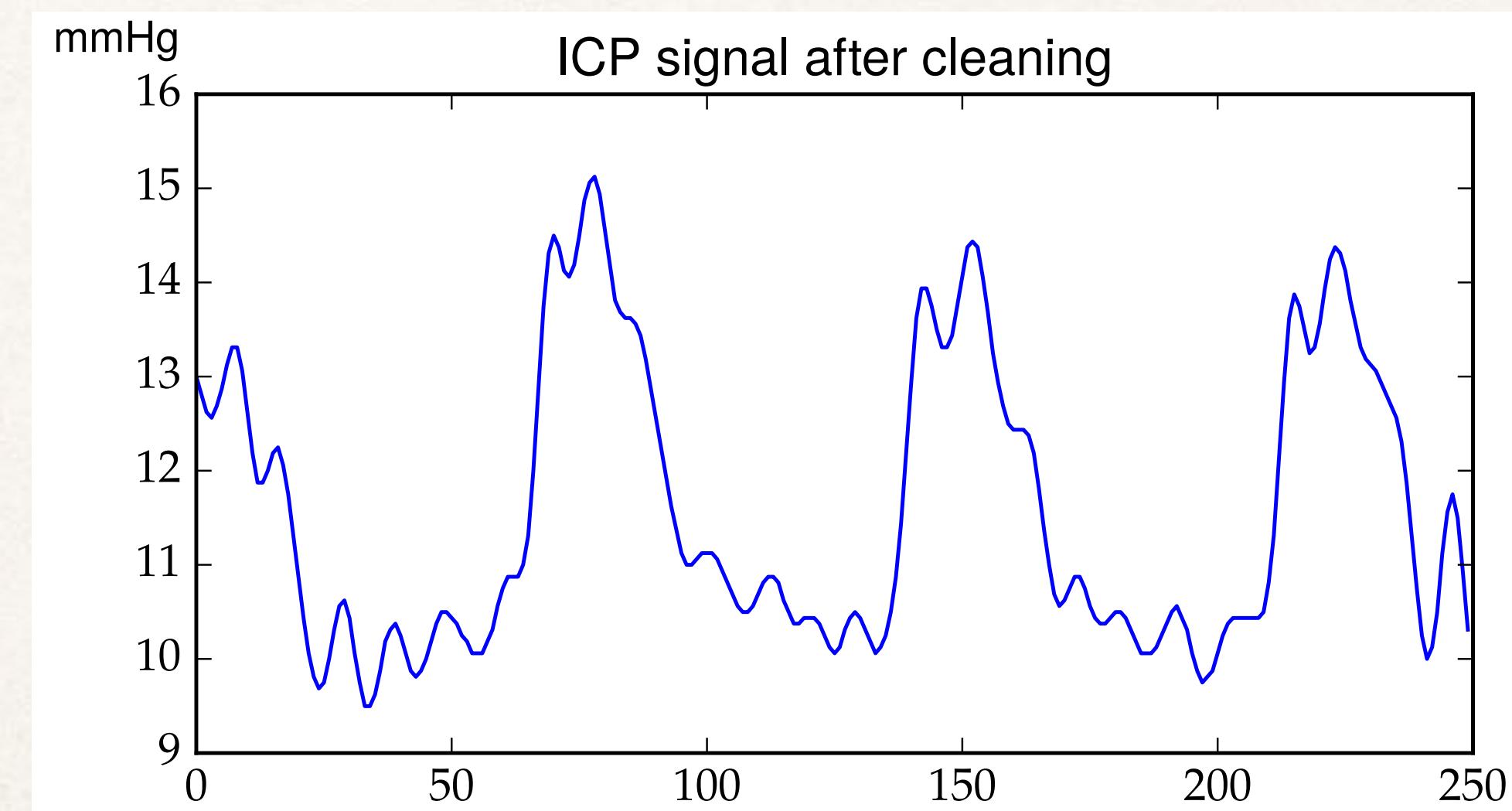
ML Applications by Master Students

Generating Maps from Satellite Images



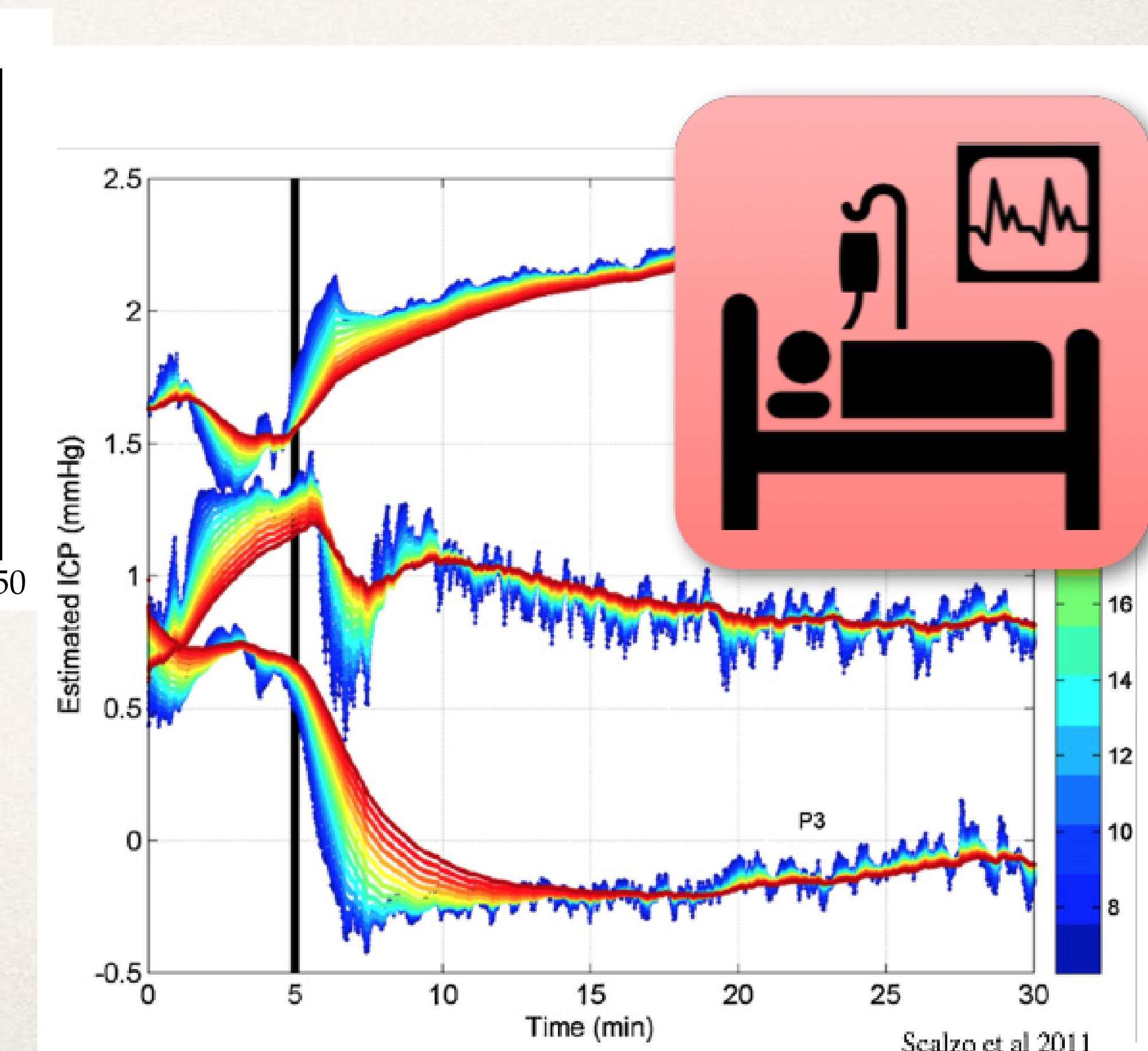
Medical Applications in Intensive Care

Forecasting of intracranial hypertension



43 GB of raw CSV text data
67 days of ICP signal
sampling rate 125 Hz

MSc Thesis Projects
Matthias Hüser & Adrian Kündig



Text Understanding

sentiment of tweets



SemEval 2016:
winner out of 34
teams, >20 countries

MSc Thesis Projects
Jan Deriu & Maurice Gonzenbach

negative	neutral	But i wanna wear my Concord's tomorrow though but i don't feel like it
positive	neutral	Gonna watch Grey's Anatomy all day today and tomorrow(:
negative	neutral	@CoachVac heey do you know anything about UVA's fallll fest loll they invited me
neutral	neutral	@DustyEf when that sun is high in that Texas sky, I'll be buckin it to county fair. A
neutral	positive	Up 20 points in my money league with Vernon Davis and L. Fitz still to go tomorrow
neutral	positive	DEEJAYING this FRIDAY in THE FIRST CHOP it's CHRIS actual SMITH with a smash
negative	negative	The Rick Santorum signing that was scheduled for tomorrow at the Books A Million
positive	neutral	@dreami9 lol yep looks like it! Was after El Clasico on Sunday. I didn't like her lol
neutral	neutral	Back in Stoke on Trent for the 2nd time today!
neutral	neutral	First Girls Varsity Basketball Game tomorrow at 6:00 pm Then Football Senior night
neutral	neutral	#UFC lightweights @Young__Assassin VS @jamievarner set for TUF 16 Finale on t
neutral	neutral	@OOOOO_WEEEE slide thru sometime this weekend ill have somethin yu can sip o
negative	negative	@DannyB618 Sure absolutely-- I meant out of the Bachmann, Perry, Santorum, H
negative	negative	@RichardGordon48 re Levein discussion on Wed. Can't keep changing boss, but he
neutral	neutral	Today In History November 02, 1958 Elvis gave a party at his hotel before going o
neutral	positive	Hustle cause you got to then kick back n party everyday like its Fri
positive	positive	I can't sleep. Way too excited about Vancouver tomorrow! I'm like a kid at Christma

Audio: Bird Songs & Ecology

another kind of tweets

winner of
BirdCLEF 2016



MSc Thesis Project
Elias Sprengel

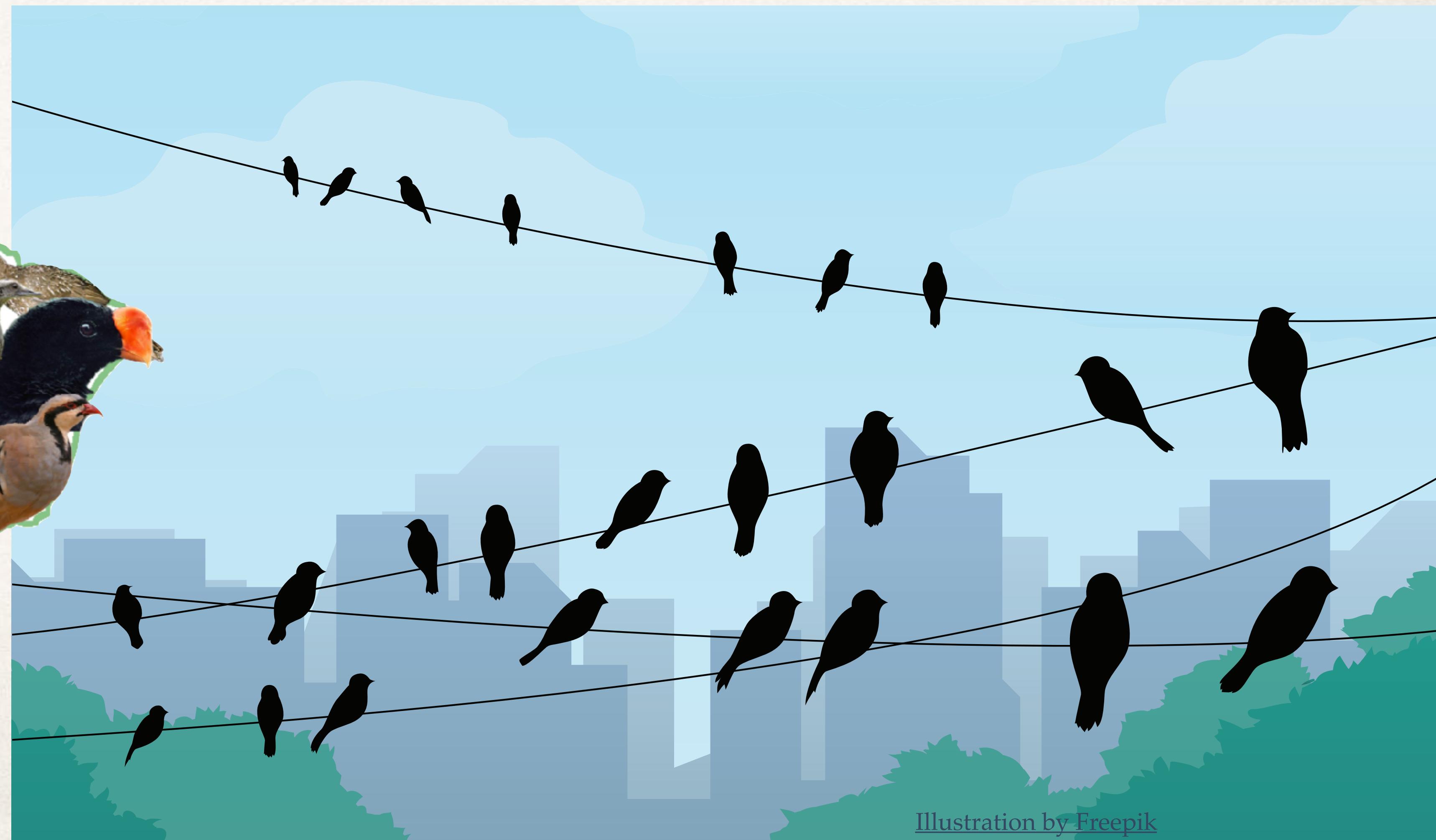
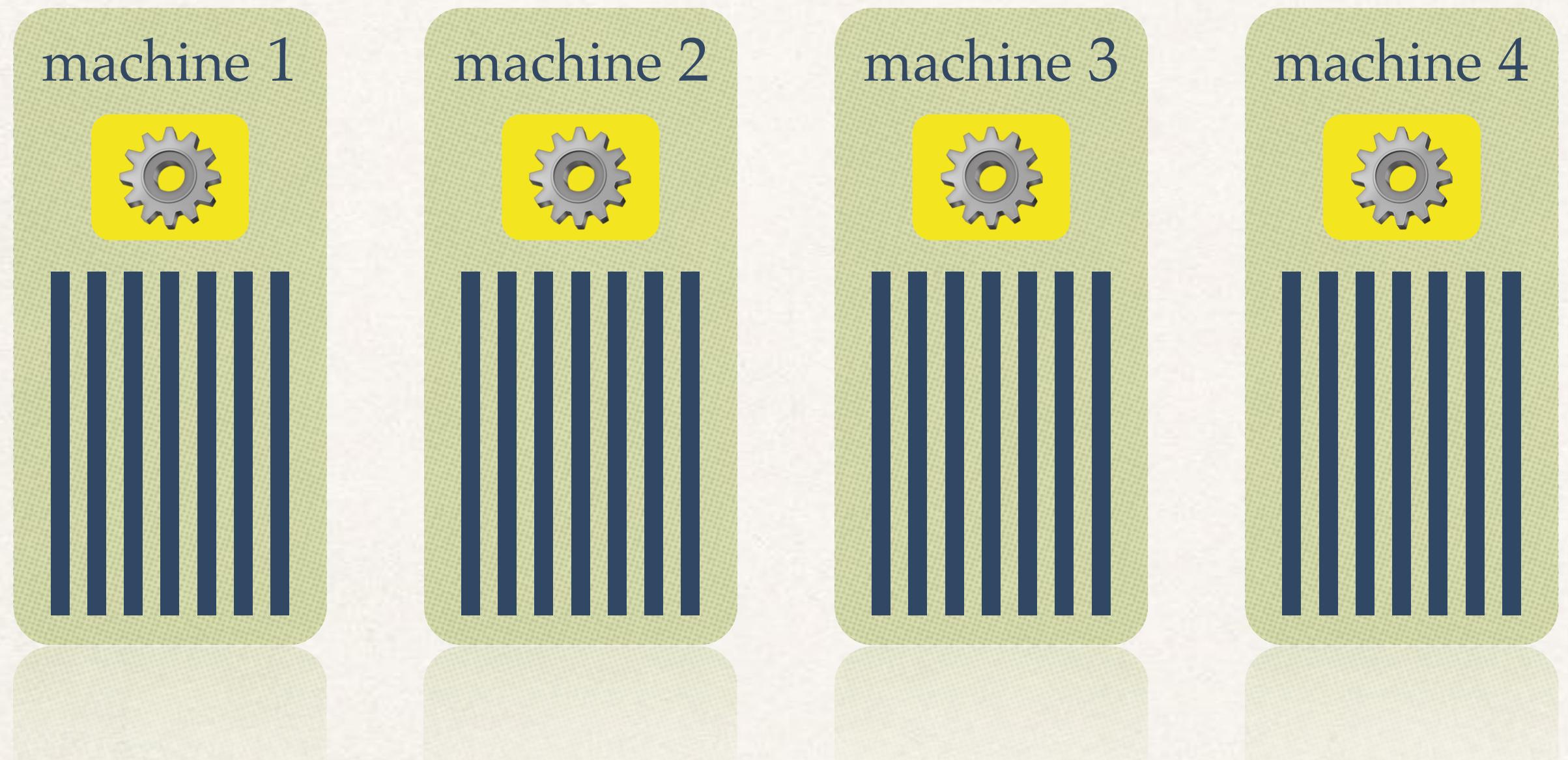
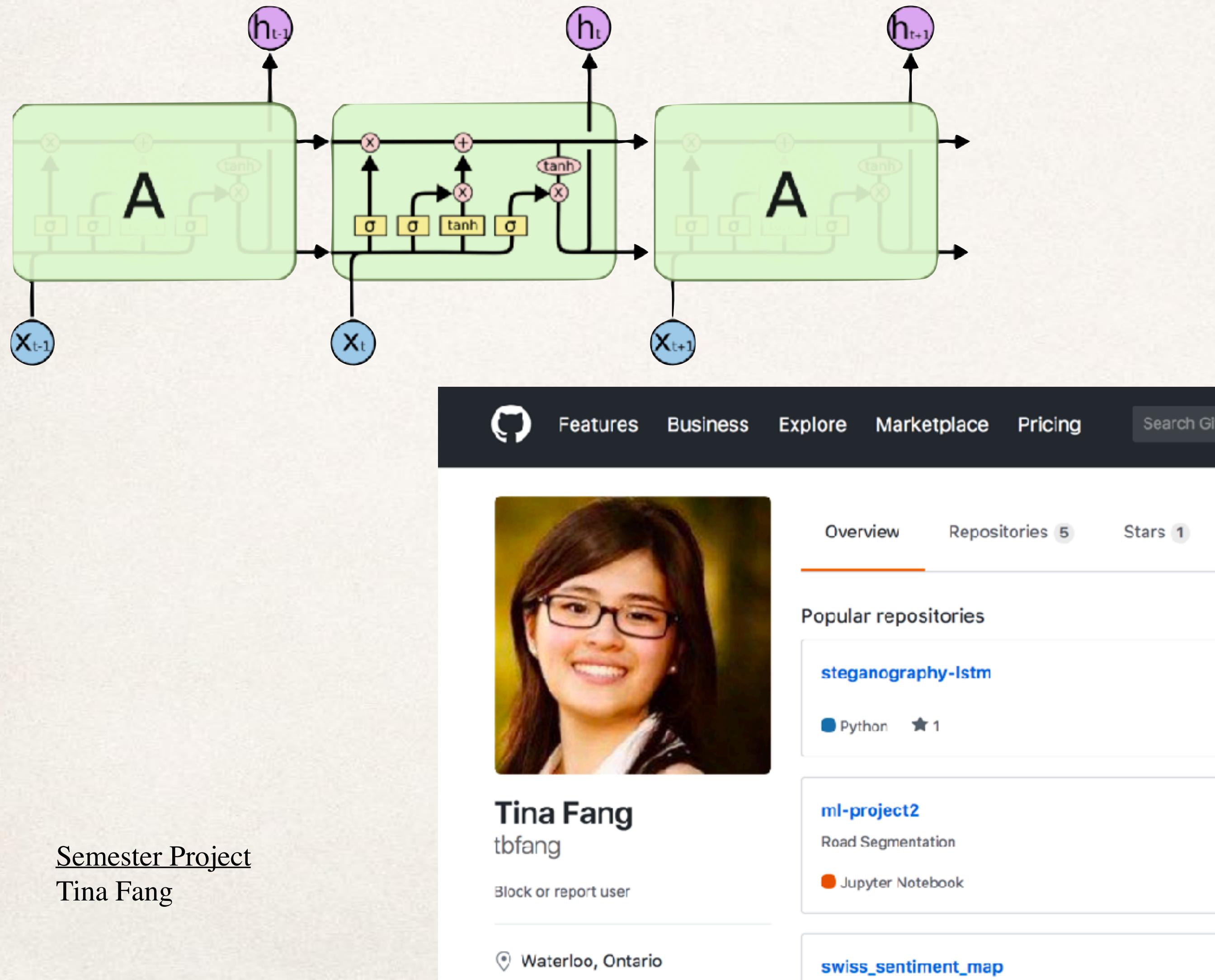


Illustration by Freepik

Distributed Training



Text Steganography with Neural Networks



steganographic tweets:

- *I can't take the whole day to the car, the people I suddenly didnt understand.*
- *"where else were u making?... i feel fine? - e? lol" * does a voice for me & take it to walmart?*
- *Don't think I can trust anymore... but I know is my fault.*
- *I can't believe he is a freshman!*

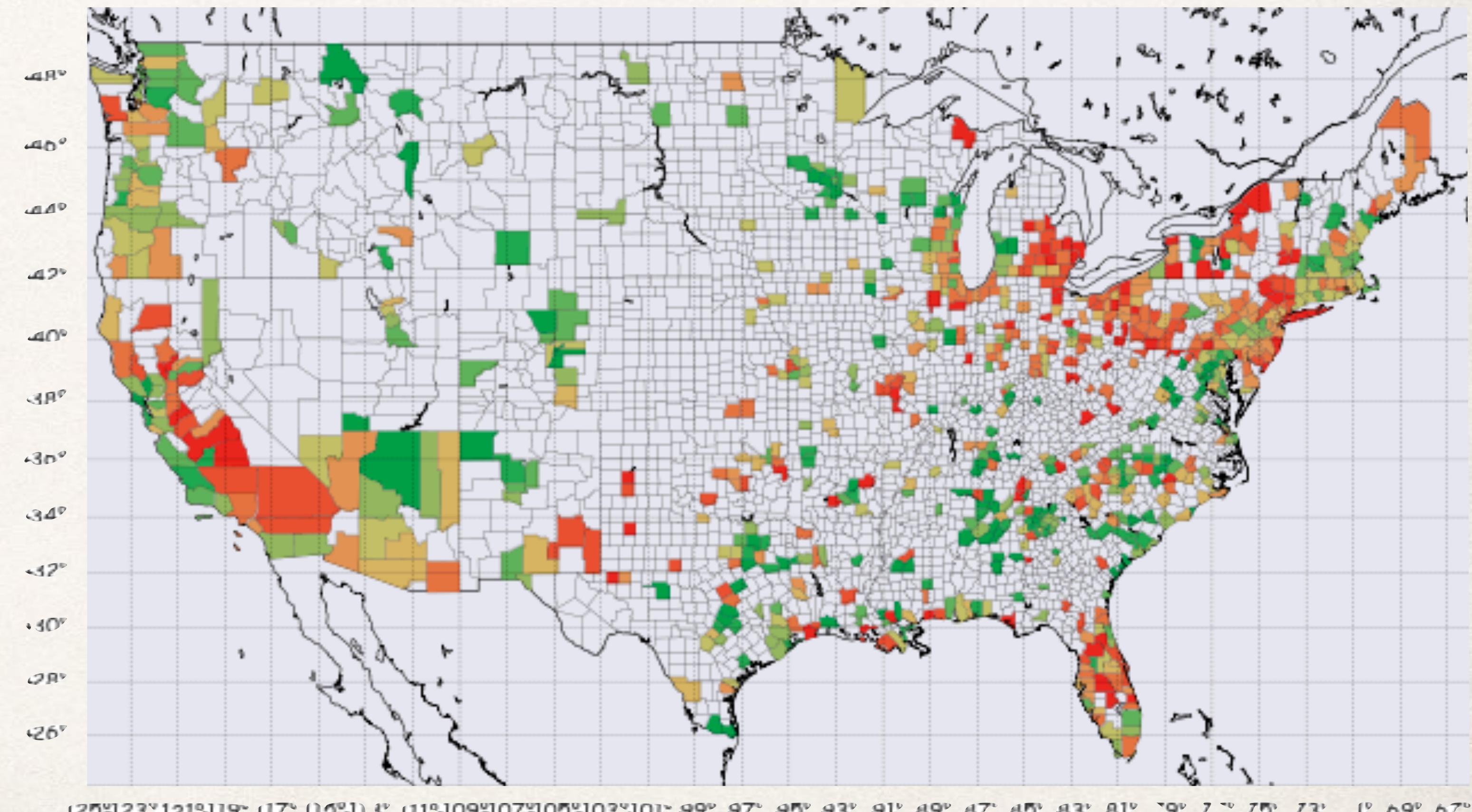
positively correlated



Semester Project
Arno Schnewly

Heart Attacks from Tweets

negatively correlated



up next:

- * Regression
- * Linear Regression
- * Classification
- * ... fundamental concepts of ML