EECS 545: Machine Learning

Introduction

Mert Pilanci

September 6, 2017

Outline

- Introduction
- Administrative
- What is Machine Learning?

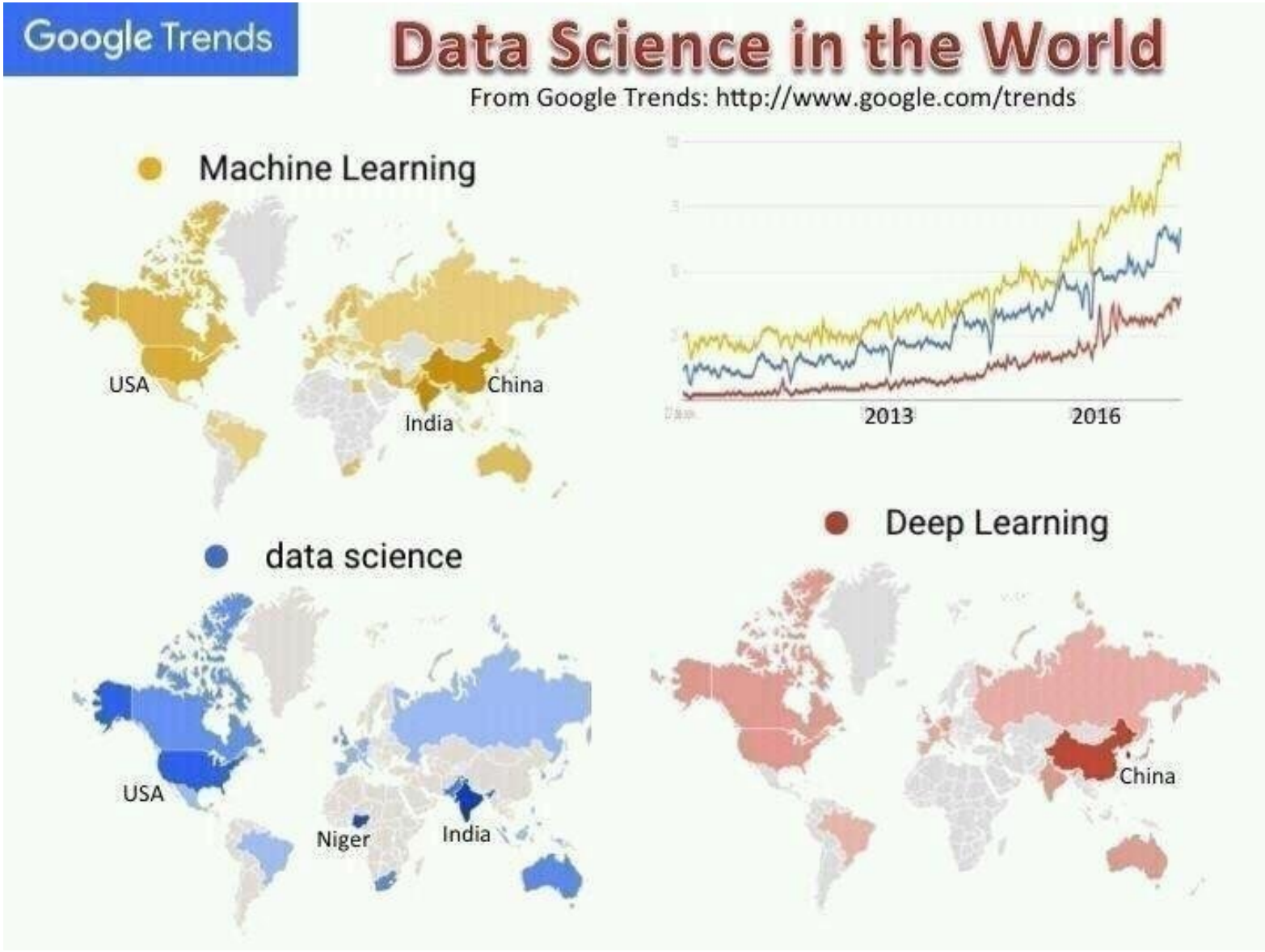
Administrative

Teaching staff

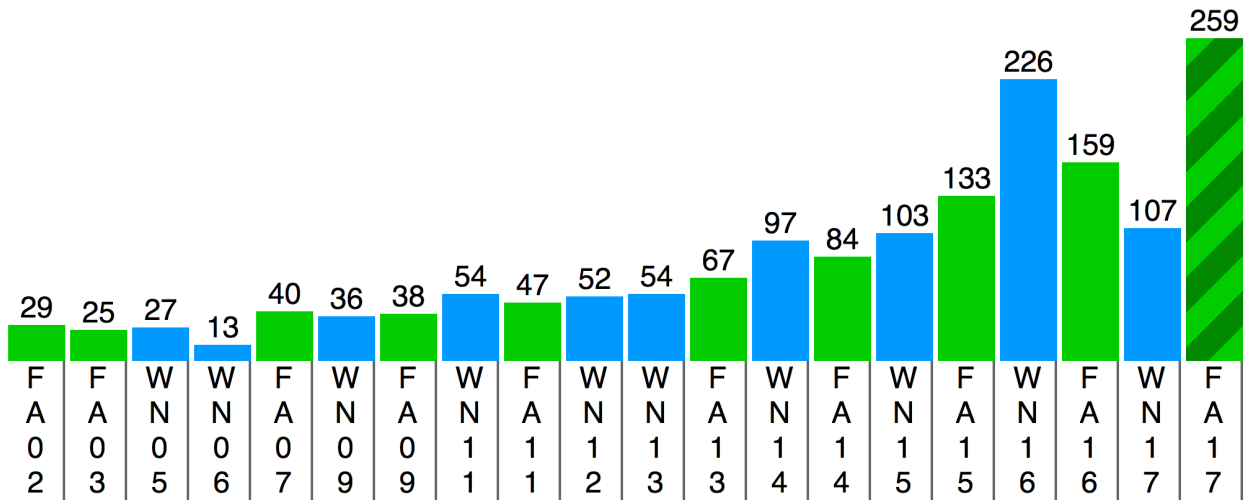
- Instructor: Mert Pilanci
  - Email: pilanci@umich.edu
  - Office hour: Tuesday 2:30pm-4:00pm in EECS 4419
  - Extra office hour today: 1:00pm-2:00pm in EECS 2311
- GSIs: Haozhu Wang
  - Email: hzwang@umich.edu
  - Office hour: GGBL 1025 Wed 18:00-21:00
  - Other GSIs TBA...

For all questions please use Piazza
piazza.com/umich/fall2017/eecs545
no email policy

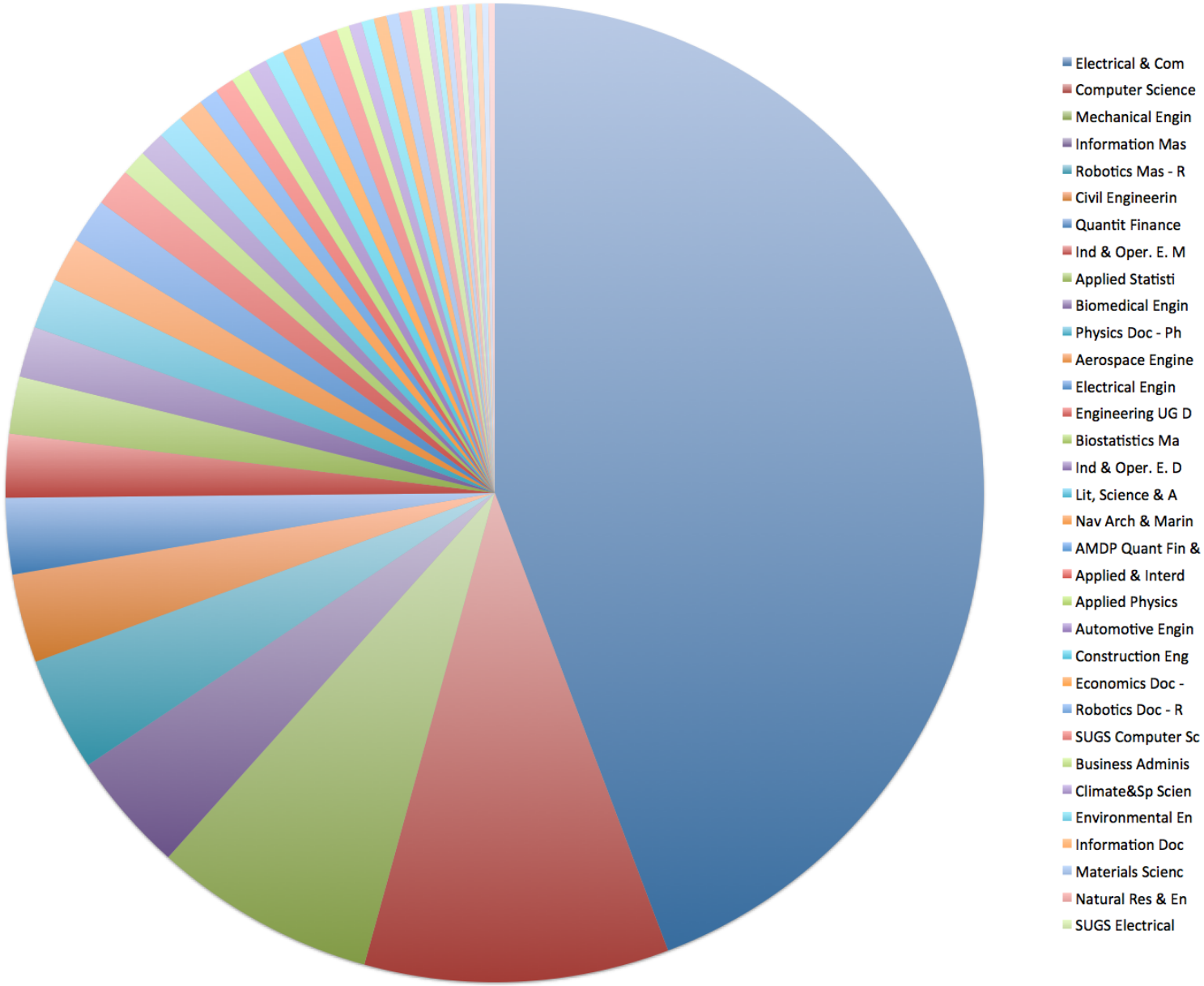
Machine Learning Trends



- EECS 545 enrollment by years



- Where do you people come from?



Some Quotes:

- If you were a current computer science student what area would you start studying heavily?
  Answer: Machine Learning.

Some Quotes:

- If you were a current computer science student what area would you start studying heavily?

  Answer: Machine Learning.

  Bill Gates, Reddit AMA

  "A breakthrough in machine learning would be worth ten Microsofts."

Some Quotes:

- If you were a current computer science student what area would you start studying heavily?
  Answer: Machine Learning.
  Bill Gates, Reddit AMA
  "A breakthrough in machine learning would be worth ten Microsofts."
- Machine learning is the next Internet
  Tony Tether, Director, DARPA
- Machine learning is todays discontinuity
  Jerry Yang, CEO, Yahoo
- Machine learning and AI is undergoing an amazing renaissance
  Jeff Bezos, CEO, Amazon

Machine Learning

- "We are drowning in information and starving for knowledge." **Rutherford D. Roger**

Comparison

- Data scientist and machine learning jobs?

Comparison

- Data scientist and machine learning jobs?

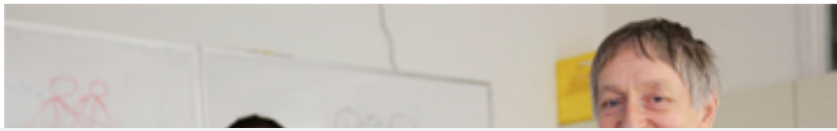
Google snaps up object recognition startup DNNr

Google has ac
Toronto, who

by Josh Lowensohn

2 / 0

Google has acqui
research compan



« Search needs a shake-up

Songbirds use grammar rules »

Machine Learning Startup Acquired by ai-one

Press Release

For Immediate Release: August 4, 2011

San Diego artificial intelligence startup acquired by leading
ning SDKs as market for advanced

oday that it acquired Auto-Semantics, a local start-up
es to corporate IT departments. The acquisition is the
nd acquisitions by ai-one that consolidates its
ging market for machine learning technologies.

Topic: Cloud

Follow via:

Microsoft acquires legal-focused machine-learning vendor Equivio

Summary: *Microsoft has purchased Equivio, maker of a machine-learning platform for the legal industry, for an undisclosed amount.*

By Mary Jo Foley for All About Microsoft | January 20, 2015 -- 16:24 GMT (08:24 PST)

Follow @maryjofoley 95.6K followers Get the ZDNet Cloud newsletter now

Microsoft has purchased Equivio, an eDiscovery/compliance vendor with a specialization in text analysis, for an undisclosed amount.

Microsoft officials announced the acquisition of the Israeli company -- its first acquisition of 2015 using more of its offshore cash -- on January 20.

Update: The Wall Street Journal reported back in October last year that Microsoft planned to buy Equivio for $200 million.

Update No. 2: A Microsoft spokesperson said the $200 million estimate was inflated and incorrect, but declined to provide a different figure.

FOUNDED
2011

OVERVIEW
DeepMind is a cutting edge artificial intelligence company. We combine the best techniques from machine learning and systems neuroscience to build powerful general-purpose learning algorithms. Founded by Demis Hassabis, Shane Legg and Mustafa Suleyman, the company is based in London and supported by some of the most iconic technology entrepreneurs and investors of the past decade. Our first commercial ...

Slide Credit: Dhruv Batra

About EECS 545

- Our goal in this course is to help you to:
  - Understand fundamentals of machine learning.
  - Learn technical details of ML algorithms.
  - Learn how to implement some important algorithms.
  - Use machine learning algorithms for your research and applications.
- We focus on foundations of machine learning:
  - Supervised learning
  - Unsupervised learning
  - Kernel methods
  - Various advanced topics: Graphical models, learning theory, sparsity, feature selection, Bayesian techniques, on-line learning, neural networks, deep learning etc.

Recommended text books

- Murphy, Machine Learning: A Probabilistic Perspective, MIT Press, 2012.
- Hastie, Tibshirani, and Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Springer, Second Edition, 2010.
- Bishop, Pattern Recognition and Machine Learning, Springer, 2006.
- Mohri, Rostamizadeh, and Talwalkar, Foundations of Machine Learning, MIT Press, 2012.

• No text book is required.

Prerequisites

- The current official requirement is listed as EECS 492, Artificial Intelligence, but this is inaccurate:
- Probability: jointly distributed random variables, multivariate densities and mass functions, expectation, independence, conditional distributions, Bayes rule, the multivariate normal distribution (e.g. EECS 401 Probabilistic Methods).
- Linear algebra: rank, nullity, linear independence, inner products, orthogonality, positive (semi-) definite matrices, eigenvalue decompositions. (e.g. MATH 217 Linear Algebra , Math 417 Matrix Algebra, EECS 551 Matrix Methods).
- Multivariable calculus: partial derivatives, gradients, chain rule.
- Programming skills in Matlab and Python (optional).

- It is expected that students will have a good working knowledge of these topics.

- If you have most but not all of this background, you should be able to catch up during the semester with some additional effort, otherwise it is strongly recommended that you finish them first before taking this course.

Grading policy

- Homework: 30%, submission via gradescope, lowest dropped.
- Midterm exam: 30%, Tuesday Nov. 28, 18:30-21:30, Stamps Auditorium.
- Quizzes 10% (will be announced)
- Project: 30% :
  - Proposal submission, due Friday Oct. 6, at 17:00.
  - Midterm project report, due Friday Nov. 3, at 17:00.
  - Final project, due Sunday Dec. 17, at 17:00, reviews due Wednesday Dec. 20, at 17:00.
- Extra credit: 5-10%, who answer questions in Piazza and significantly enhance the course experience through their contributions.

Homework

- Assigned homeworks will be bi-weekly.
- The problem sets will also include programming assignments to implement algorithms covered in the class.
- MATLAB is the official programming language of the course but we also support those prefer Python.
- Homework #1 is out already (see Canvas or `goo.gl/xZSQdA`) – due 9/22, at 17:00.
- Please start on homework early.
  (Warning: cramming does not work).

Group Study

- Homework:
  - Working in groups is allowed, but each member must submit their own writeup.
  - Write the members of your group on your solutions (Up to four people are allowed).
- Project:
  - You will be asked to form groups of about 4 people.
  - You will be asked to select a paper on a methodology not covered in class, and implement the method.
  - You will grade each others projects.

For details see Canvas!

Any questions?

Definition of ML

- Samuel (1959): Field of study that gives computers the ability to learn without being explicitly programmed.

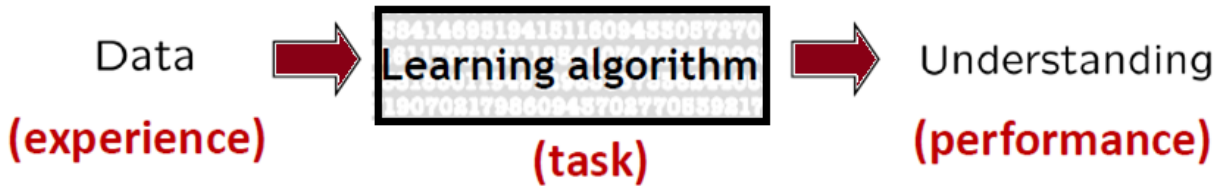
Definition of ML

- Samuel (1959): Field of study that gives computers the ability to learn without being explicitly programmed.
- Kevin Murphy (2012): Algorithms that automatically detect patterns in data use the uncovered patterns to predict future data or other outcomes of interest

Definition of ML

- Samuel (1959): Field of study that gives computers the ability to learn without being explicitly programmed.
- Kevin Murphy (2012): Algorithms that automatically detect patterns in data use the uncovered patterns to predict future data or other outcomes of interest
- Mitchell (1998)- Formal Definition: **A** is said to learn from experience **E** with respect to some class of tasks **T** and performance measure **P**, if its performance at tasks in **T**, as measured by **P**, improves with experience **E.**

- Informal Definition: Algorithms that improve their prediction performance at some task with experience (or data).



E * T = P

Experience	*	Task	=	Performance
Input Data: • Housing prices • Customer transactions • Clickstream data • Images		**Task:** • Predict prices • Segment customers • Optimize user flows • Categorize images		**Performance:** • Accurate prices • Coherent groupings • KPI lifts • Correctly sorted images

Typical Examples

• Spam filtering: Suppose your email program watches which emails you do or do not mark as spam, and based on that learns how to better filter spam:

- Task: Classifying emails as spam or not spam.
- Experience: Watching you label emails as spam or not spam.
- Performance measure: The number (or fraction) of emails (disjoint from the training emails) correctly classified as spam/not spam.

We wish to congratulate you once again, for being among the Twelve (12) selected winners in the ongoing E-mail Electronic Online Sweepstakes. Hence we do believe with your prize, you will continue to be active in your patronage to Google and its Products. A Bank Cheque has been issued in your favour, hence you have won for yourself the sum of £950.000.00 (Nine Hundred and Fifty Thousand Great British Pounds Sterling), One Google Nexus 10 Tablet and also you have been enlisted as One of the Google Ambassadors for 2015.

To claim your reward, please contact our Foreign Payment Bureau officer below by neatly filling the verification and funds release form below, as your payment will be released and arranged by our United Kingdom Office.

MANDATORY FOREIGN PAYMENT RELEASE FORM.

(1) Your Contact Address:
(2) Your Contact Telephone/Mobile Number:
(3) Your Nationality/Country:
(4) Your Full Names:
(5) Occupation:
(6) Age/Gender:
(7) Marital Status:
(8) Private Email Address:
(9) Ever Won An Online Lottery?
(10) How Do You Feel As A Winner?
(11) Your Preferred mode of prize remittance from the two options below:

(a) Cash Pick-Up (You as the Beneficiary coming Down to UK to receive your Award Personally, available to only British citizens and residents).
(b) Courier Delivery of your certified winning cheque in your name and other Winning documents safely to you.

Contact our Foreign Payment Bureau officer below:
Eric E. Schmidt
Executive Chairman Google Board with these E-mail accounts as follows,
Email: mrschmidt.gpa12@googlemail.com , schmidt.apa11@googlemail.com

Note: You can either fill your claims verification form by printing and manually filling out the requested details or you can fill directly on e-mail, or provide the details on Microsoft Word.

NOTE!!! For security reasons, you are advised to keep your winning information confidential till your claims are processed and your money remitted to you. This is part of our precautionary measure to avoid double claiming and unwarranted abuse of this Program by some unscrupulous elements. Please be WARNED!!!!

Congratulations from the Staff & Members of Google Board Commission.

11/06/14

Office of the Director Google United Kingdom

Google Incorporation World wide

06/11/15

MD Matt Britt
Chairman of the Board and Managing Director,

Larry Page
Co-Founder & CEO

money	free	conference	algorithm	*data*
100	10	0	0	1

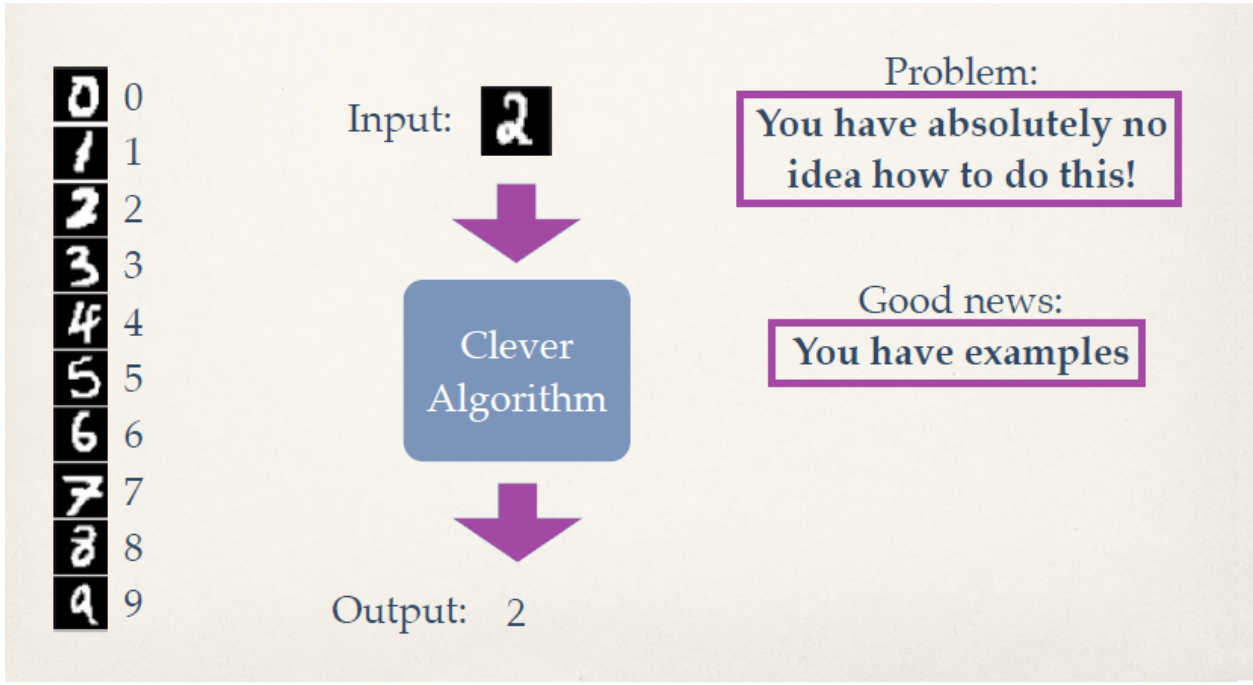
money	free	conference	algorithm	data
0	0	2	1	3

Dear Mert,

We hope all is well with you!

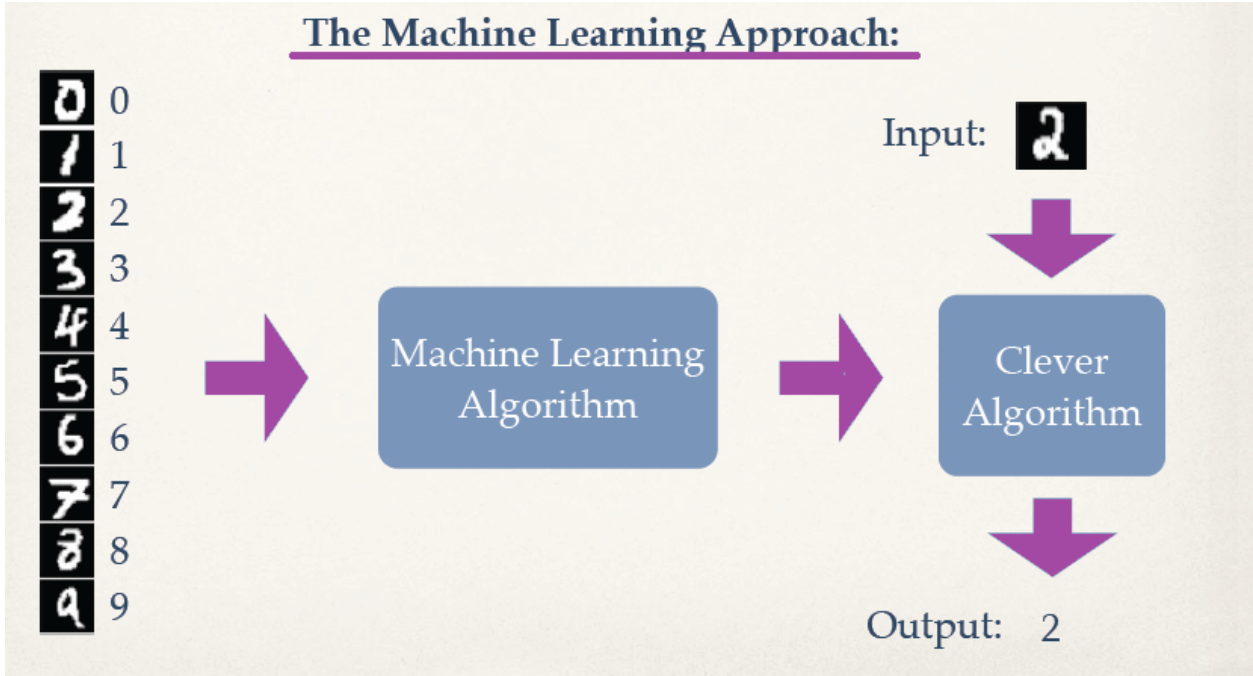
We're organizing a minisymposium (MS) in the SIAM Conference on Applied Linear Algebra 2018 (May 04 -- 08, 2018, Hong Kong Baptist University; http://www.math.hkbu.edu.hk/siam-ala18/). The proposed MS shall focus on novel theory, algorithms, and applications in exploiting low-complexity structures for modern data analysis.

• Hand-written Digit: Given an image of a handwritten digit, how the learning algorithms can be used to recognize images of hand-written digits?



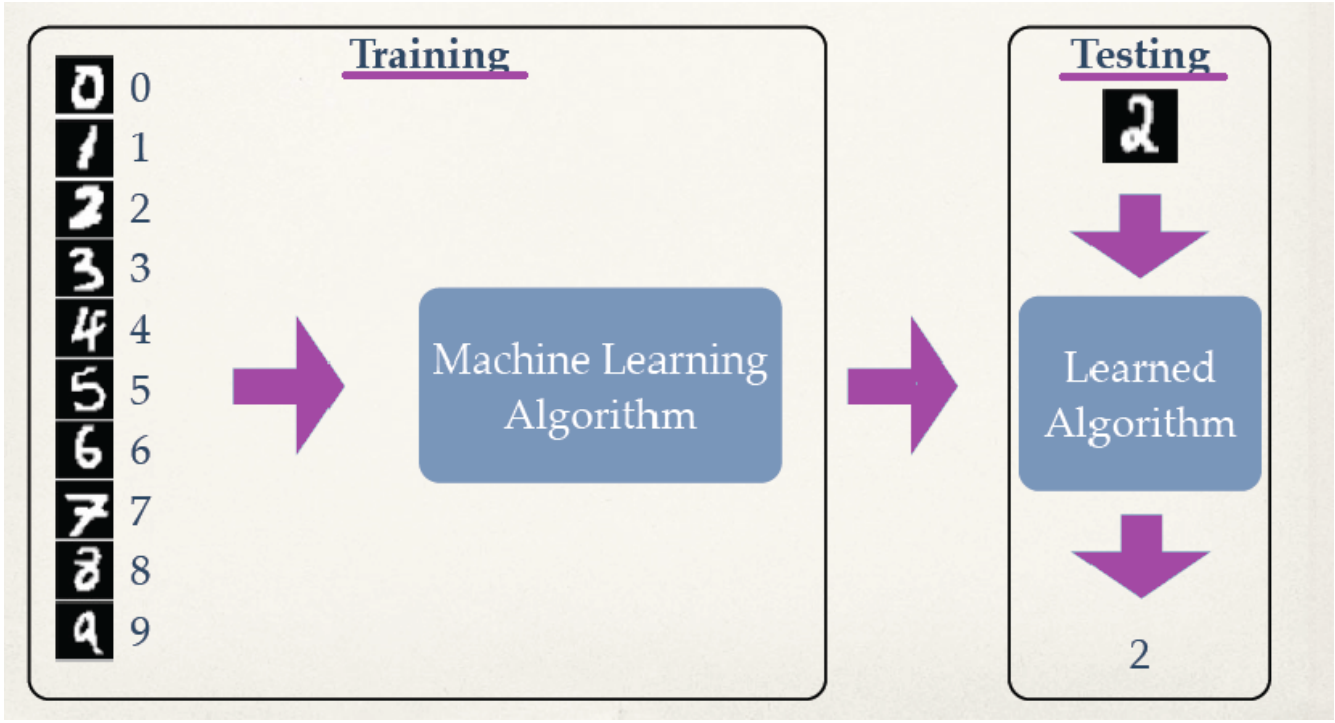
Slide Credit: Jake Abernethy

• Hand-written Digit: Given an image of a handwritten digit, how the learning algorithms can be used to recognize images of hand-written digits?



Slide Credit: Jake Abernethy

• Hand-written Digit: Given an image of a handwritten digit, how the learning algorithms can be used to recognize images of hand-written digits?

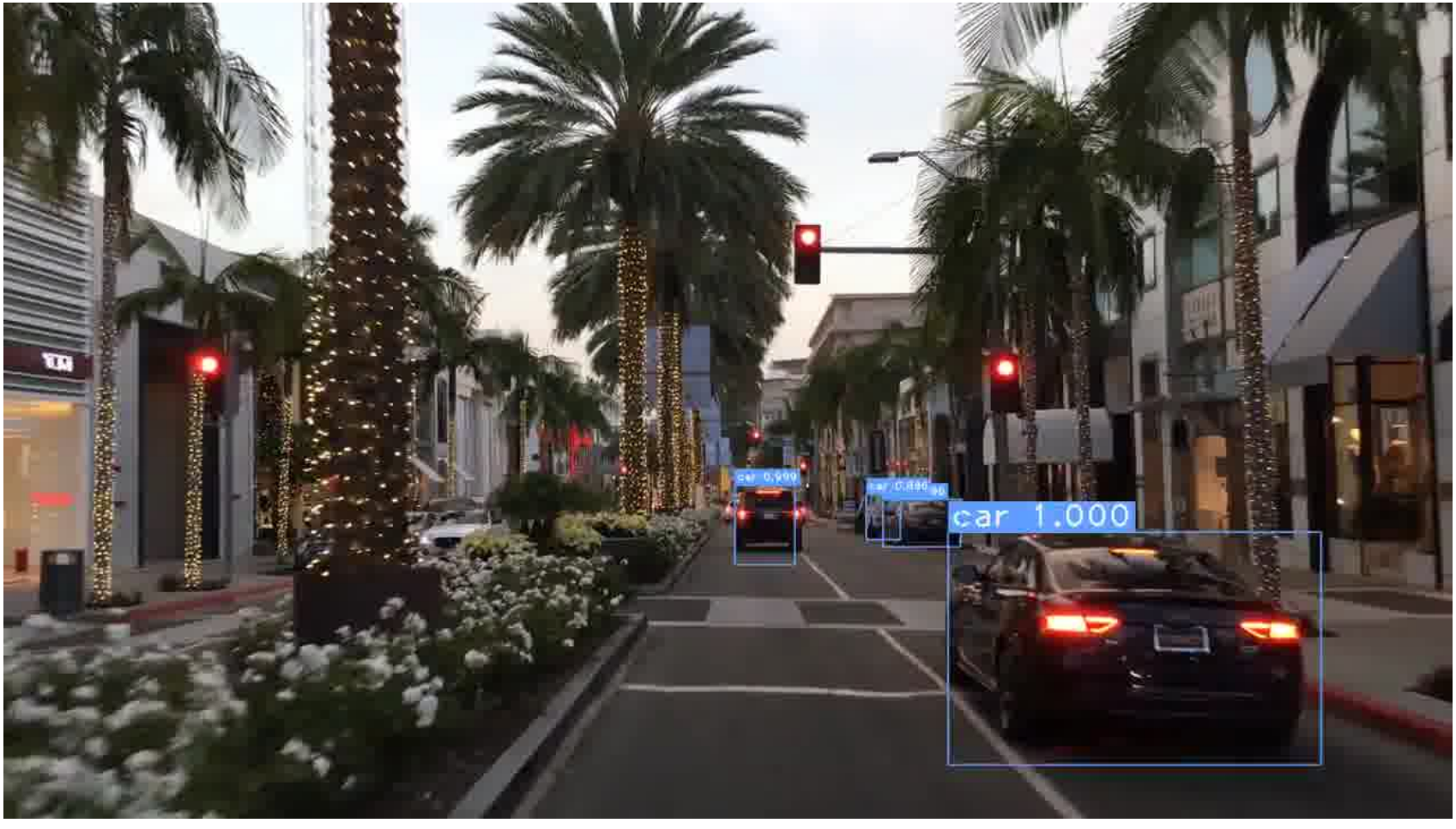


Slide: Jake Abernethy

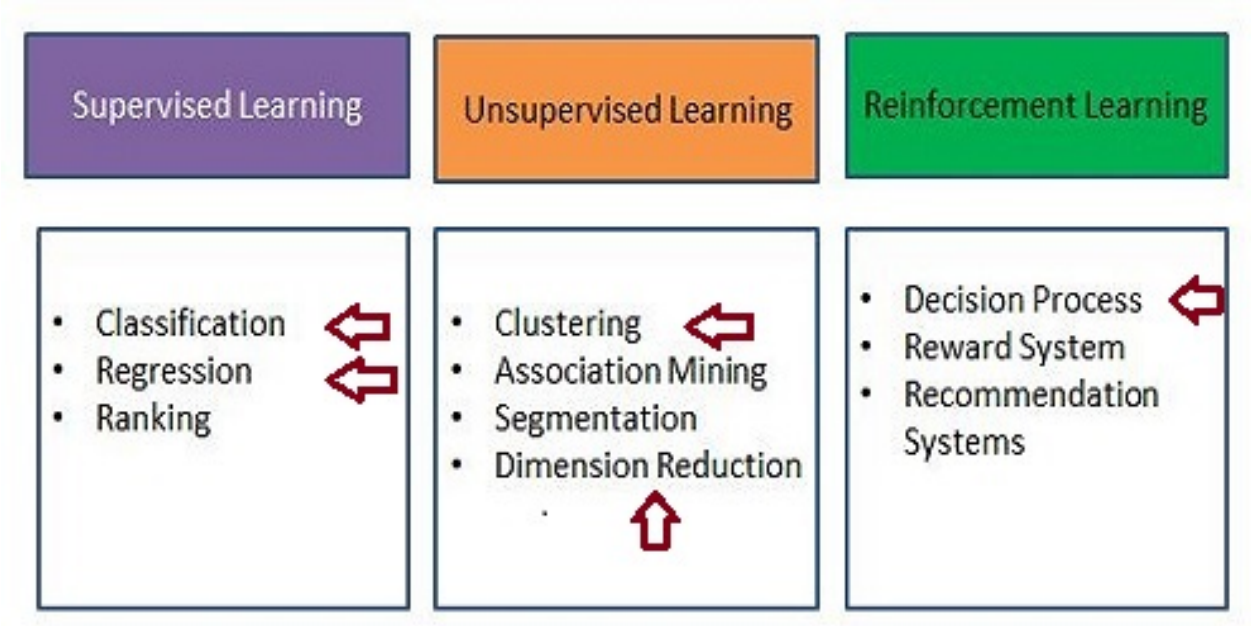
Example: Netflix Challenge

- Predict how a viewer will rate a movie
- 10% improvement = \$1M

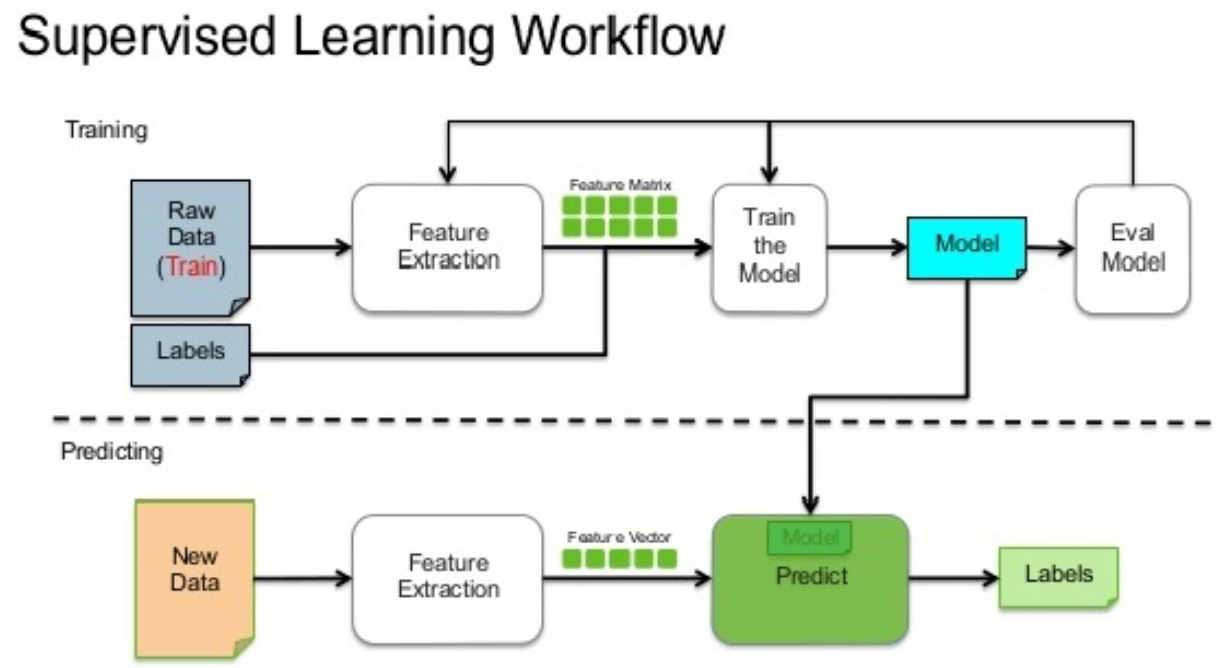
Example: Self-driving cars



Machine Learning Tasks



Supervised Learning



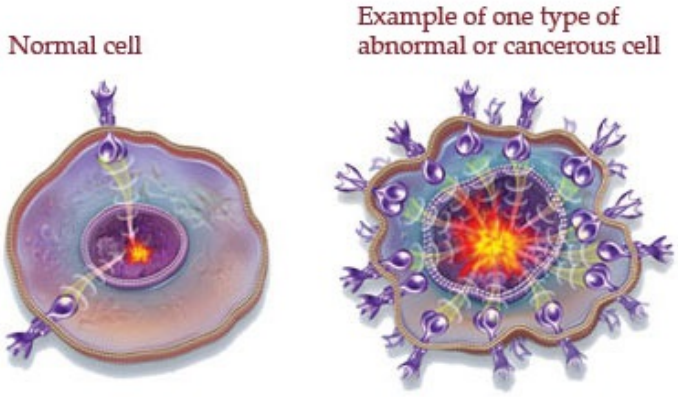
- Discrete labels: **classification**
- Continuous labels: **regression**

  Task: Given $\mathbf{X} \in \mathcal{X}$ (Feature Space), predict label $Y \in \mathcal{Y}$ (Label Space).

Classification (discrete label)

- Bio-Classification

Malignant cell or Normal cell



- Multi-Classification

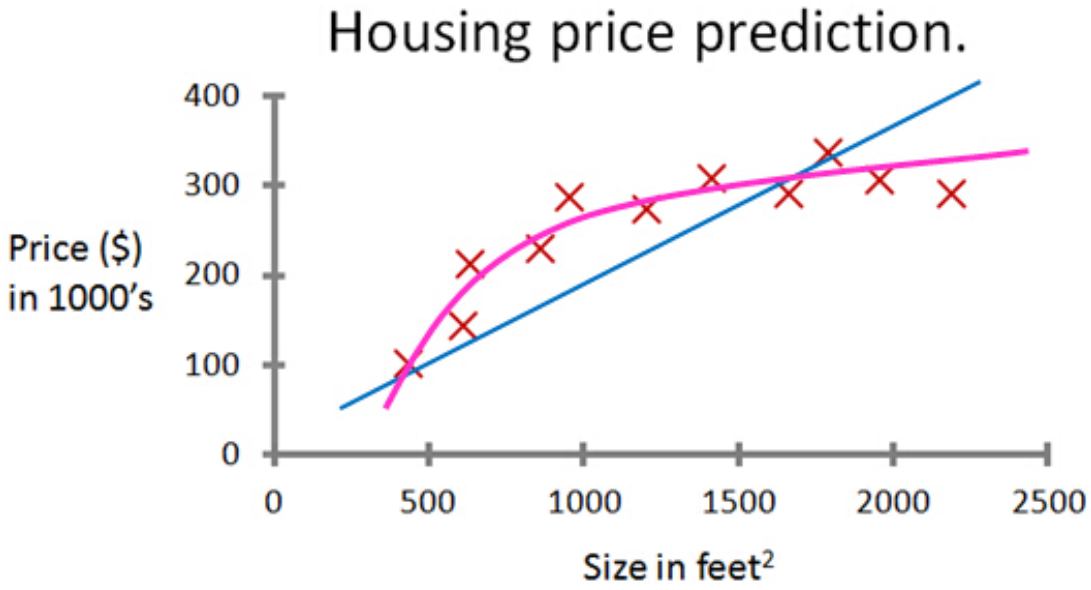
Predicting results in a **discrete** output. In other words, we are trying to map input variables into **discrete** categories.

$$h_\Theta(x) \in \mathbb{R}^4$$

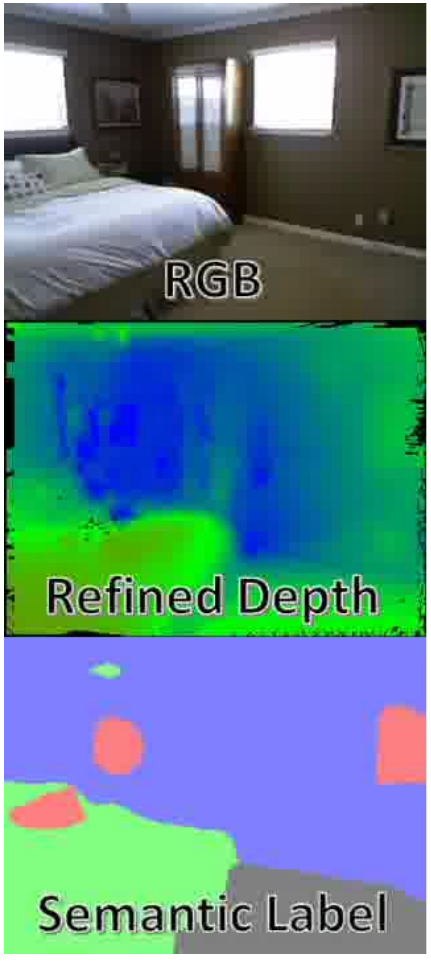
Want $h_\Theta(x) \approx \begin{bmatrix}1\\0\\0\\0\end{bmatrix}$, $h_\Theta(x) \approx \begin{bmatrix}0\\1\\0\\0\end{bmatrix}$, $h_\Theta(x) \approx \begin{bmatrix}0\\0\\1\\0\end{bmatrix}$, etc.

when pedestrian when car when motorcycle

Regression (continuous label)

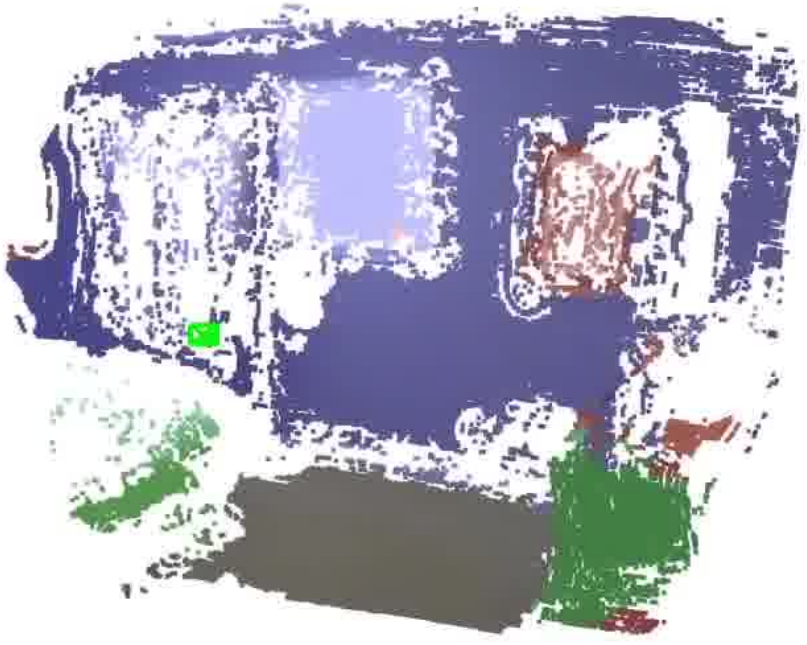


Application: Self-navigating drones (regression and classification)



FPS: 32.679134

:Floor :Vertical structure/Wall
:Large structure/furniture :Small structure



Result of dense 3D reconstruction and semantic label fusion

Other applications

- Forecasting (e.g. energy demand prediction, sales)
- Detecting anomalies (e.g. intruders, virus mutations)
- Classifying (e.g. credit risk assessment, cancer diagnosis)
- Ranking (e.g. Google search, personalization)
- Summarizing (e.g. News, social media sentiment)
- Decision making (e.g. AI, robotics, compiler tuning, trading)

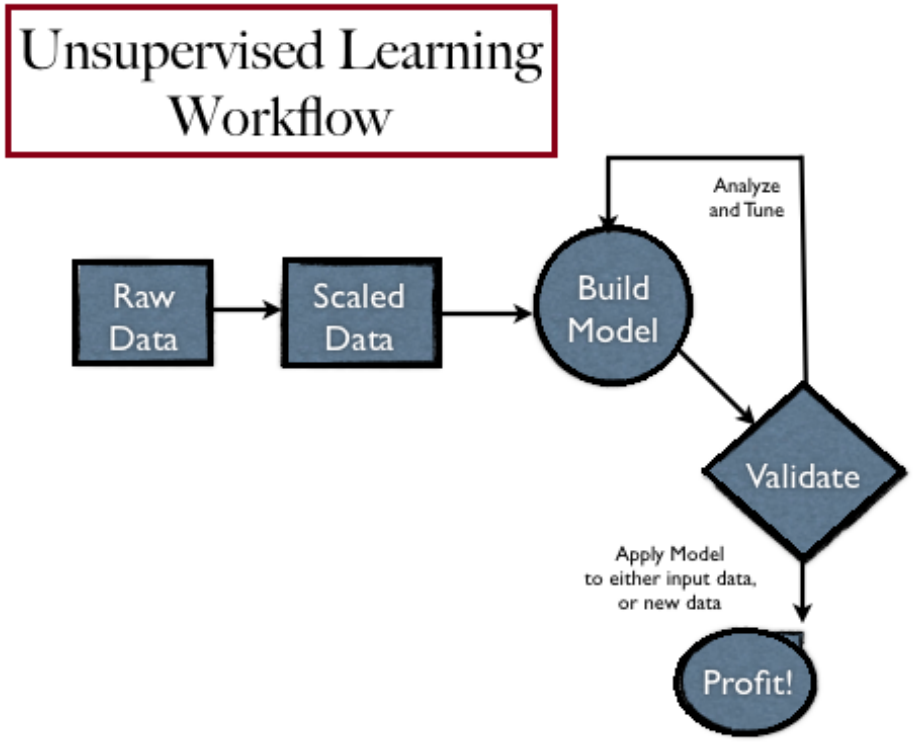
Slide Credit: Nando de Freitas

When to apply machine learning

- Human expertise is absent (e.g. navigating on Mars)
- Humans are unable to explain their expertise
  (e.g. Speech recognition, vision, language)
- Solution changes with time
  (e.g. Tracking, temperature control, preferences)
- Solution needs to be adapted to particular cases
  (e.g. biometrics, personalization)
- The problem size is to vast for our limited reasoning capabilities
  (e.g. Calculating webpage ranks, matching ads to Facebook pages)

Slide Credit: Nando de Freitas

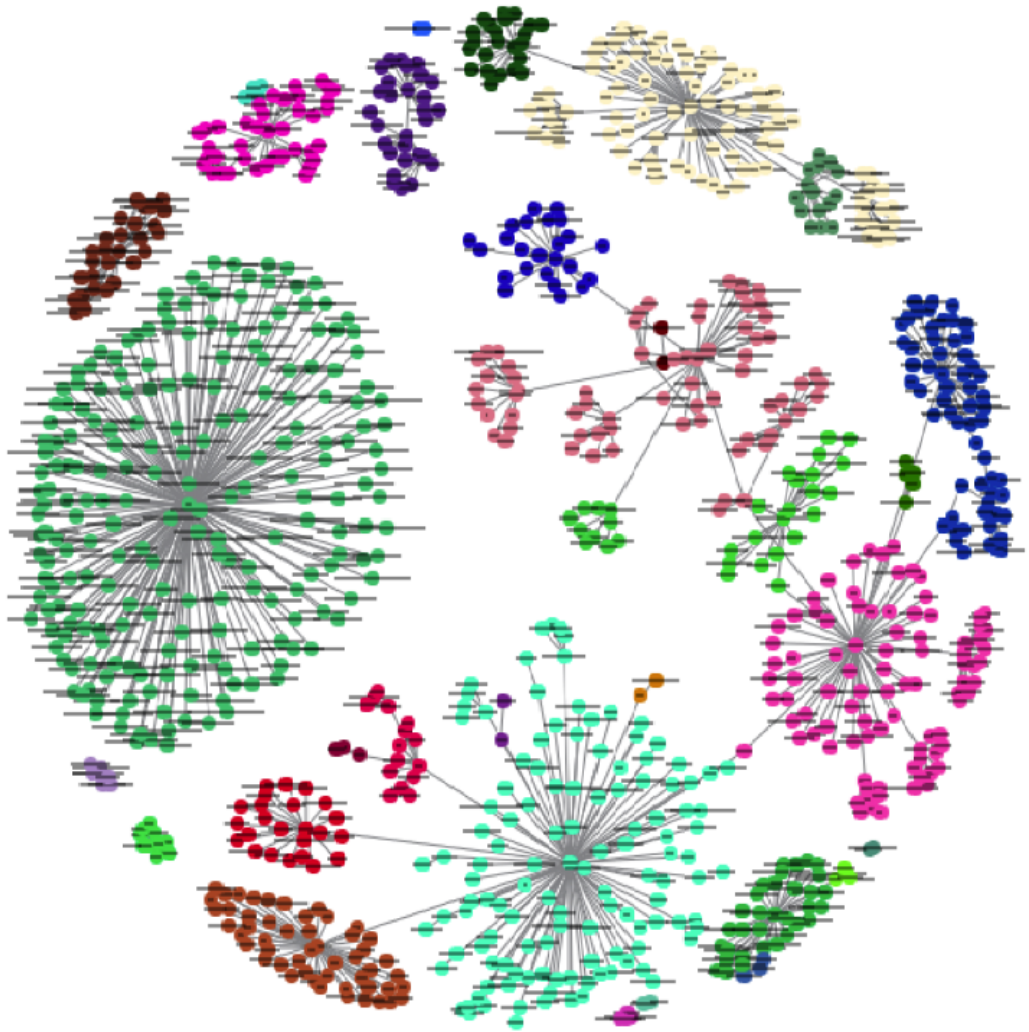
Unsupervised Learning



- Goal: Given data **X** without any labels.
- Learn the structures of the data:
  - Clustering
  - Probability distribution (density) estimation
  - Embedding and neighborhood relations

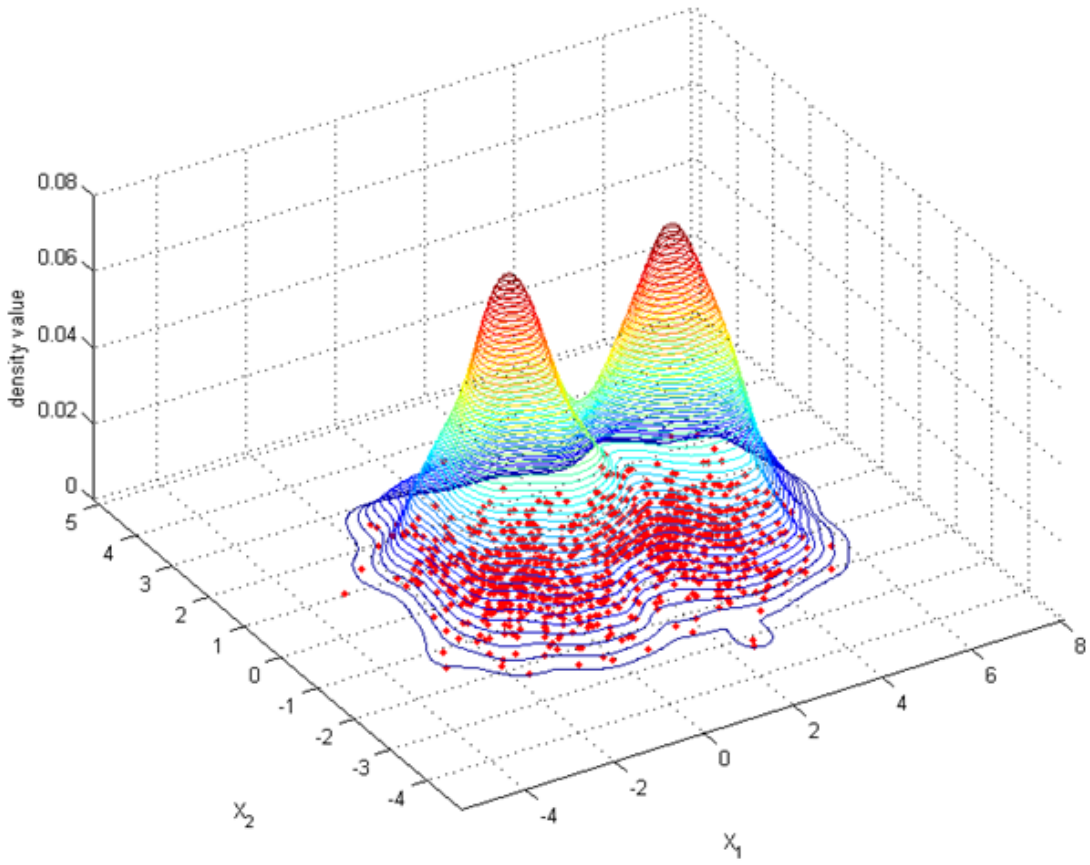
Clustering

- A cluster is a collection of objects which are similar between them and are dissimilar to the objects belonging to other clusters.



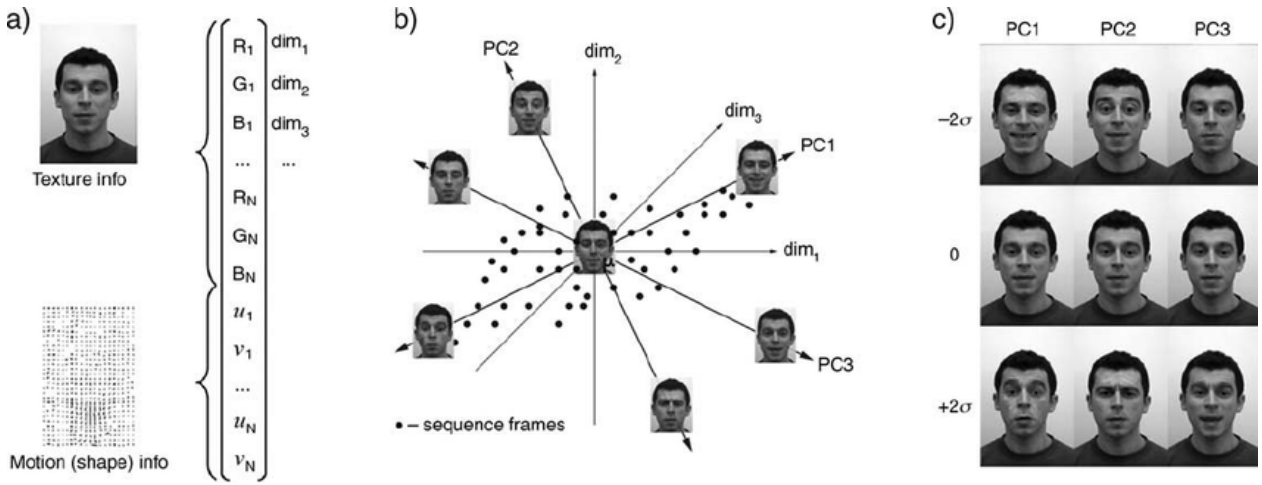
Density Estimation

- Multivariate Kernel Density Estimation



Learning-Embedding and Dimensionality Reduction

- Vectorization and application of Principal component analysis (PCA) on facial motion sequence:



F. Berisha et al (2010)

Example: Twitter

Elon Musk @elonmusk · Sep 4

Replying to @elonmusk

China, Russia, soon all countries w strong computer science. Competition for AI superiority at national level most likely cause of WW3 imo.

3.1K 18K 43K

Darren Cunningham @dcunni · 4h

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

Elon Musk @elonmusk · 4h

Replying to @dcunni @SVbizjournal

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

320 2.1K 7.6K

Example: Twitter

Elon Musk @elonmusk · Sep 4

Replying to @elonmusk

China, Russia, soon all countries w strong computer science. Competition for AI superiority at national level most likely cause of WW3 imo.

3.1K 18K 43K

Darren Cunningham @dcunni · 4h

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

Elon Musk @elonmusk · 4h

Replying to @dcunni @SVbizjournal

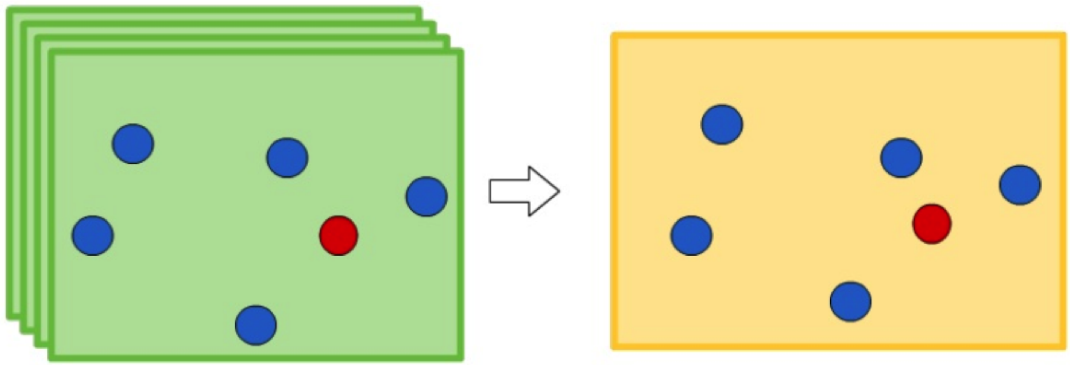
I've talked to Mark about this. His understanding of the subject is limited.

320 2.1K 7.6K

- document $\times$ term matrix

AI	...	China	...	computer	...	Mark	science	talked	Russia	WW3
1	0	1	0	1	0	0	1	0	1	1
0	0	0	0	0	0	1	0	1	0	0

Dimension Reduction



demo

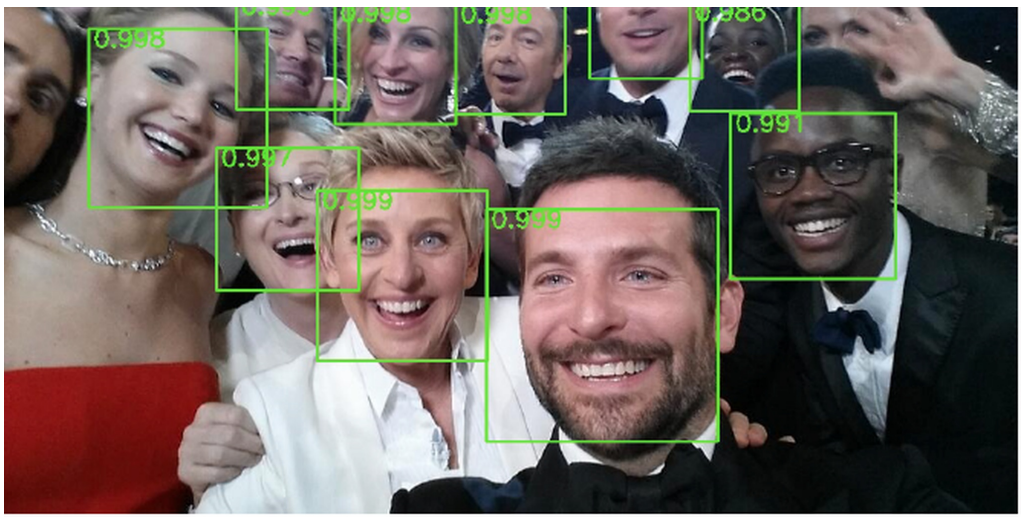
Feature Extraction(FE)

- Feature Extraction (FE) is any algorithm that transformation raw data into features that can be used as an input for a learning algorithm.
- Features are statistics or attributes that describe the data:

- Examples:
  - Construct bag-of-words vector from an email
  - Remove stopwords in a sentence
  - Apply PCA projection to high-dimensional data

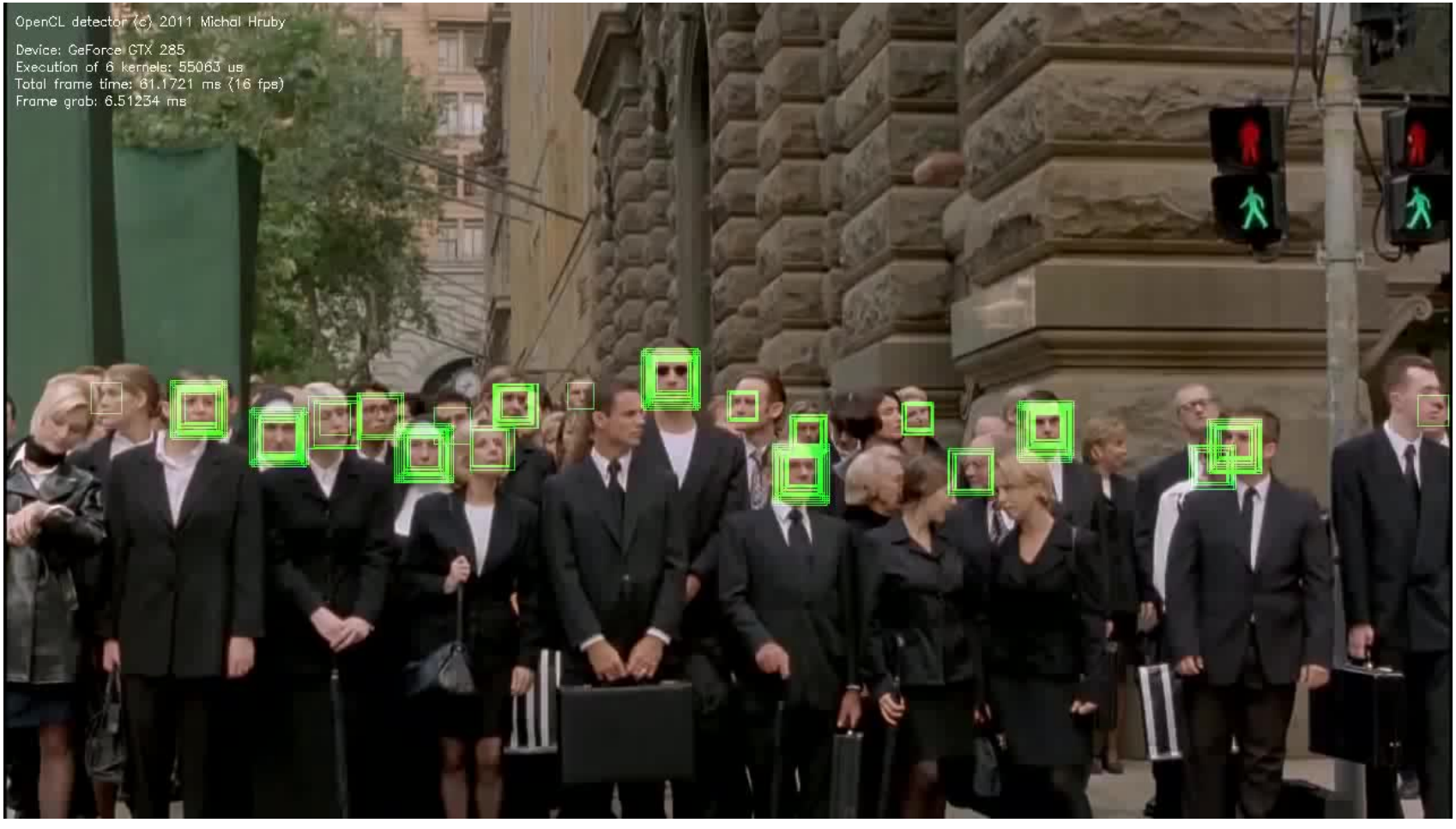
ML Application: Computer Vision

- Face Detection
  - Face-detection algorithms focus on the detection of frontal human faces.
  - A reliable face-detection approach based on the genetic algorithm and the eigen-face technique.



Viola and Jones face detector (2001)

Real-time face detection



ML Application: Speech Recognition

- Voice Search (e.g., Google), Speech Transcription

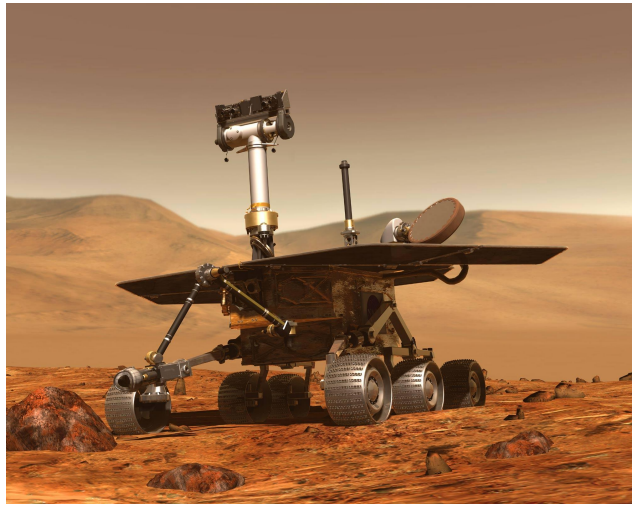
- Text to Speech

Baidu Deep Voice Demo

ML Application: Robotics

- Helicopter control

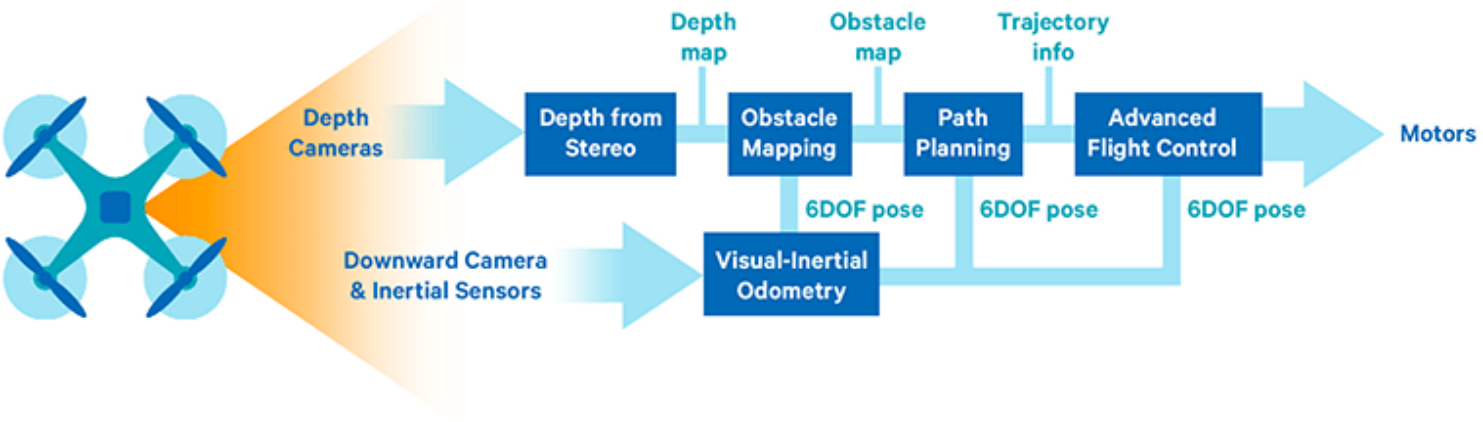
- Robot perception and navigation



ML Application: Robotics

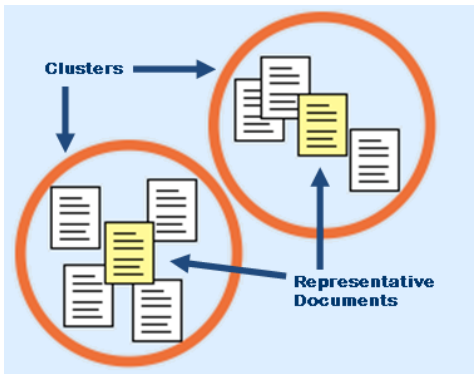
- Self-navigating drones

- Learning depth map and semantic segmentation



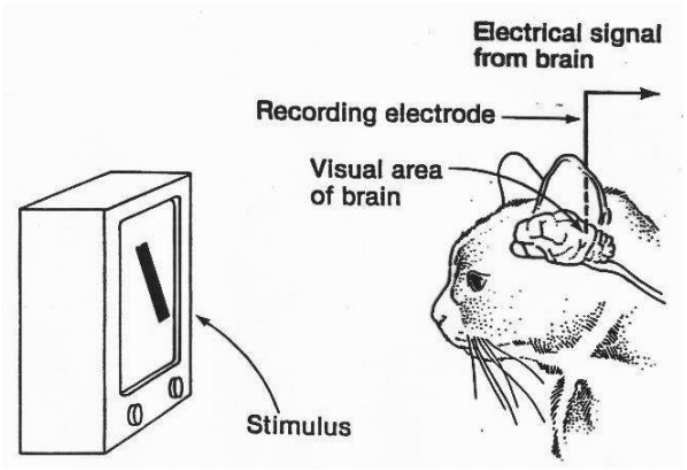
Text Processing/Data Mining

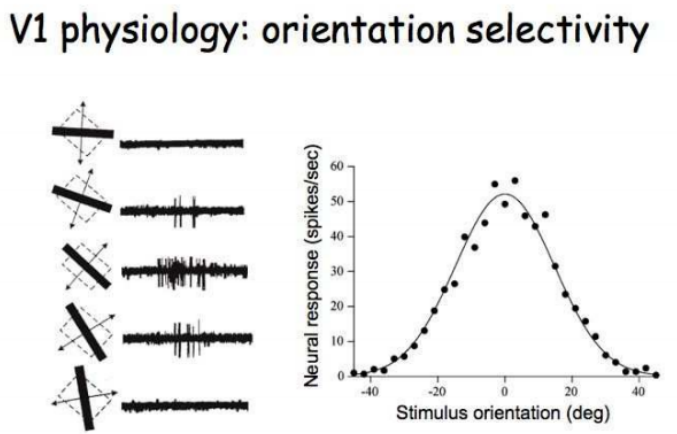
- Spam filtering
- Document clustering
  - Given news articles, group them into different categories (google news).
- Web search
  - Given query, predict which document will be clicked on.
- Product matching/Advertisement
  - Given user profile, predict which item user will like to purchase (e.g., Netflix, Amazon, etc.)

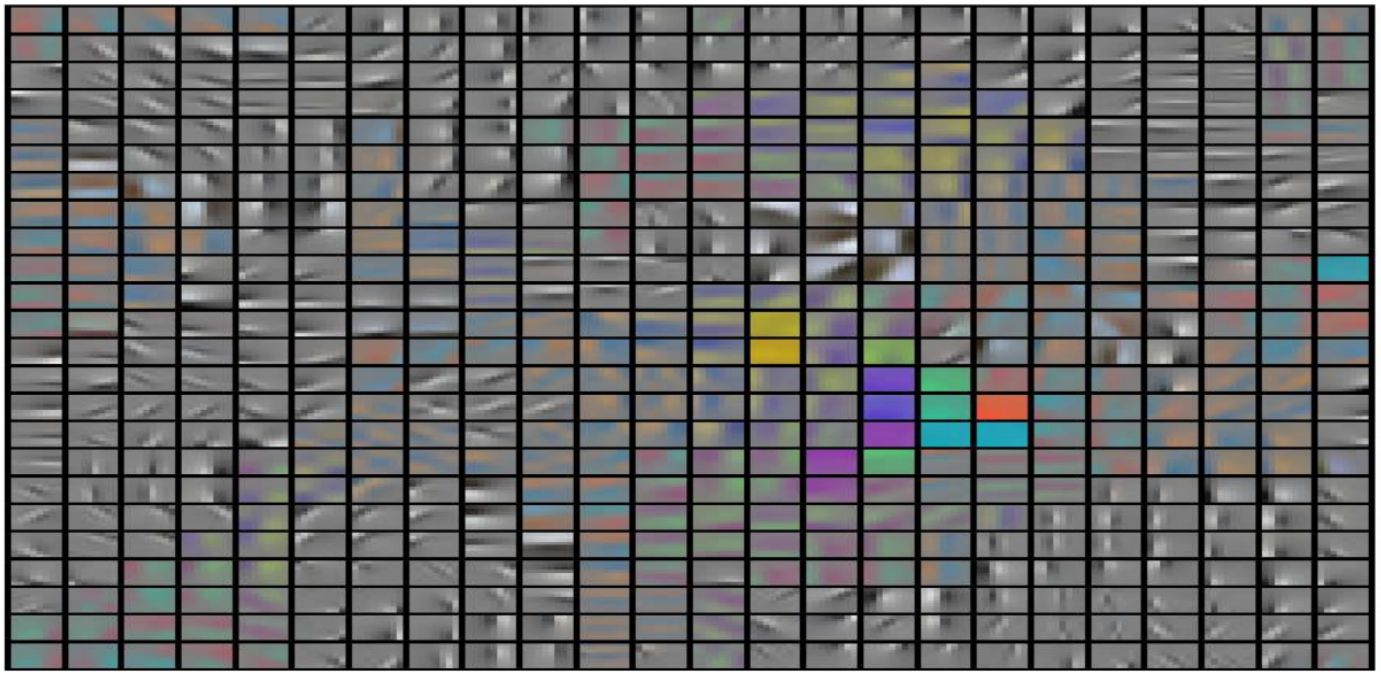


Neural Nets

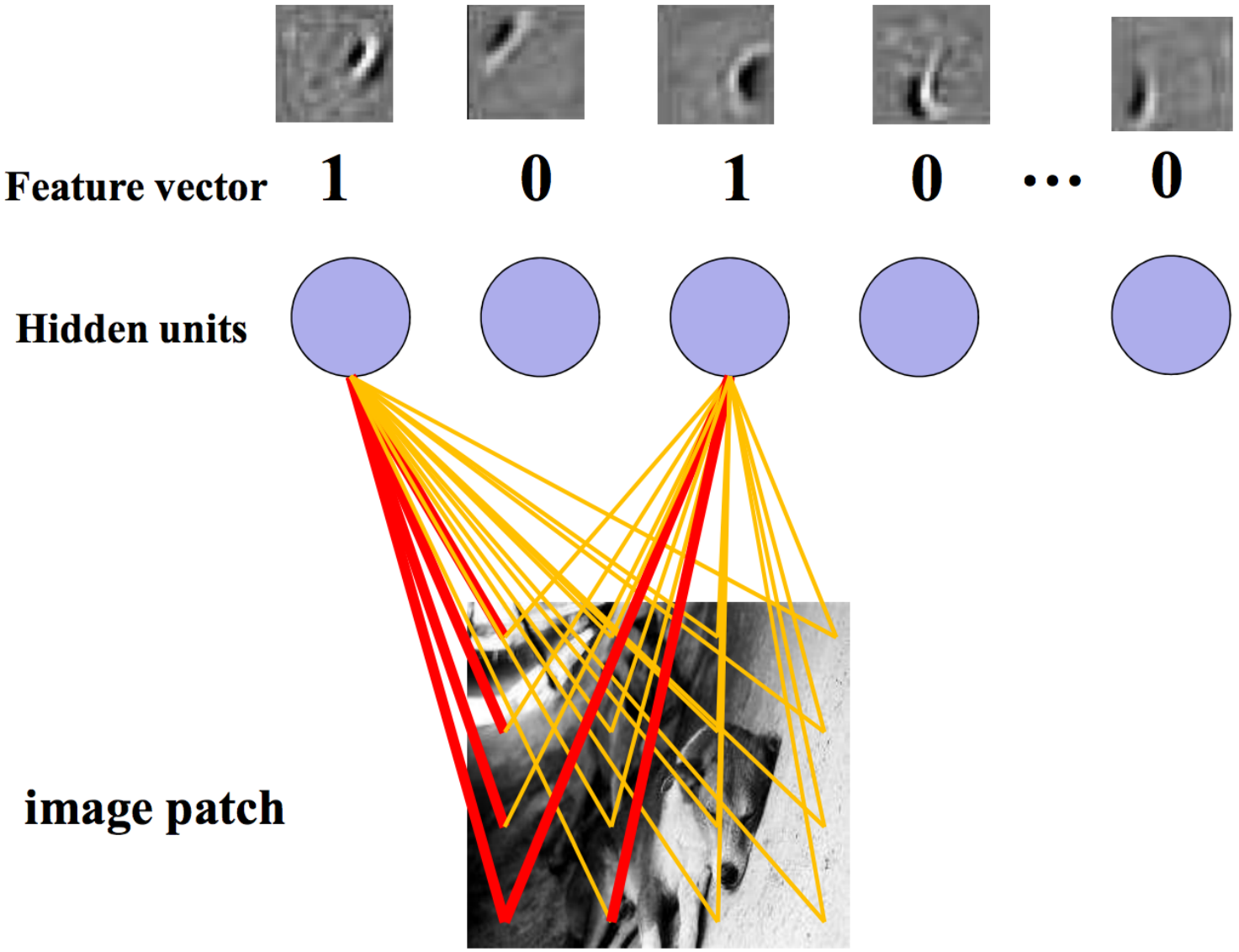
Selectivity and Topographic maps in V1



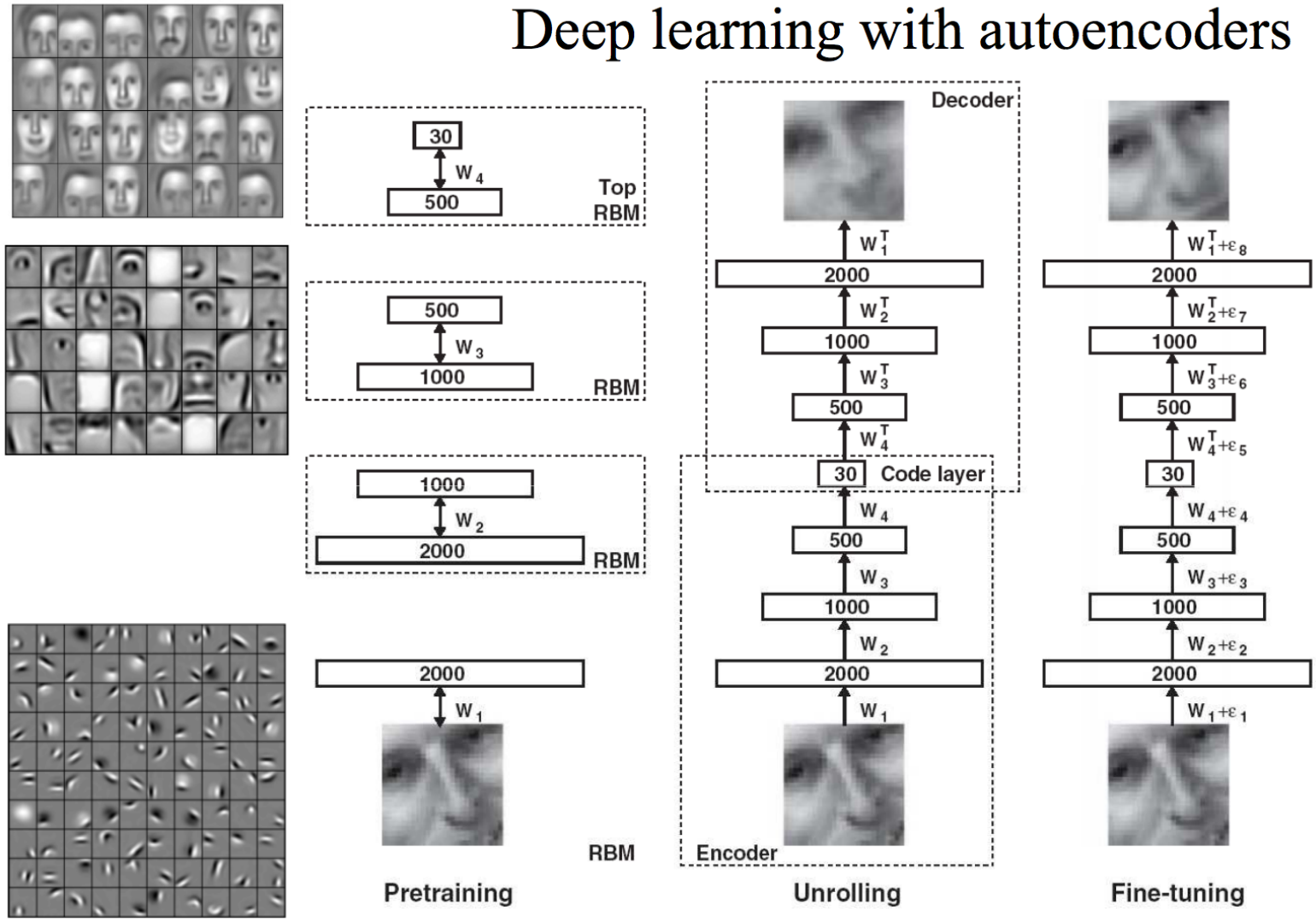




Neural Nets

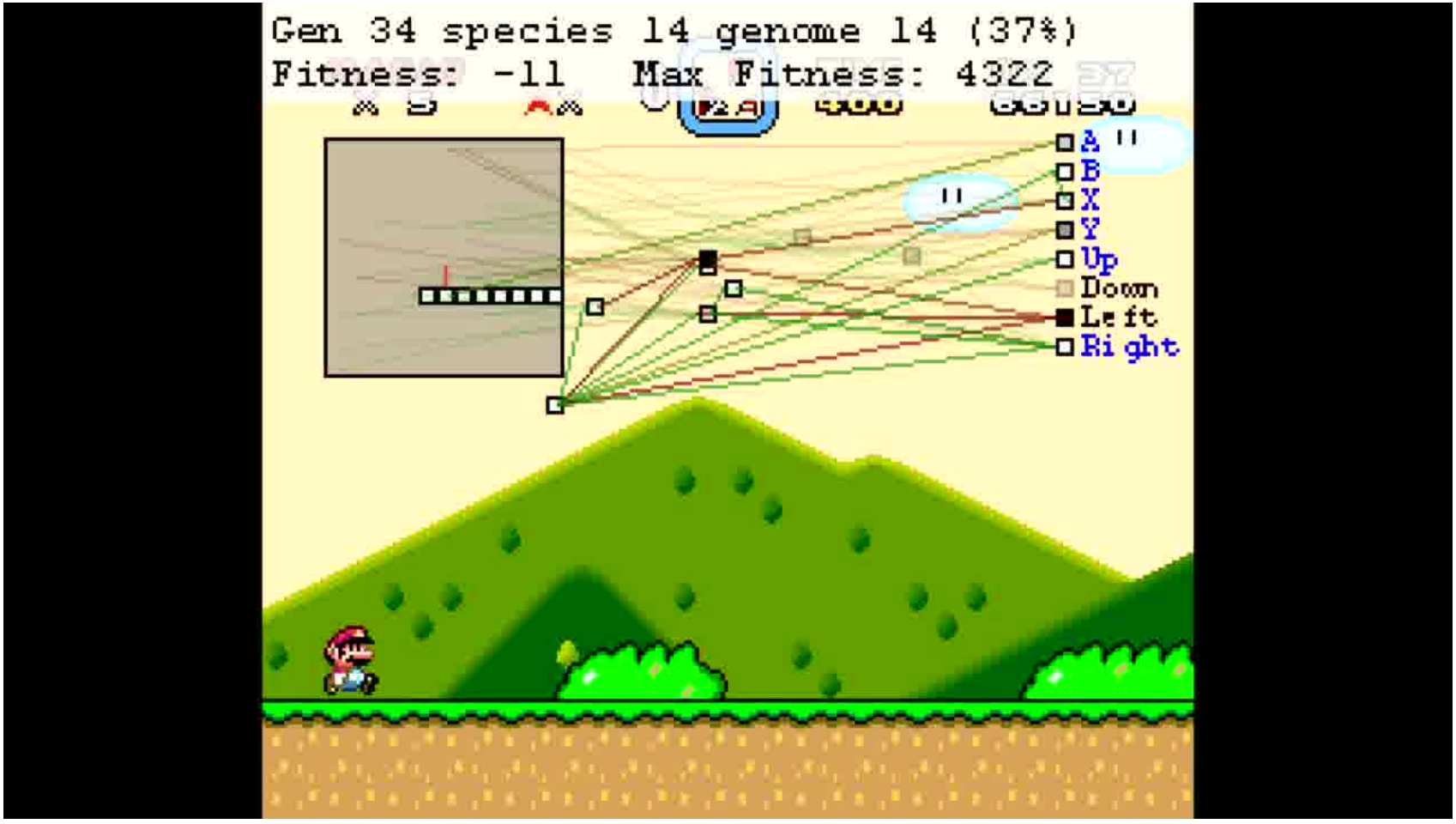


Neural Nets



[Russ Salakhutdinov, Geoff Hinton, Yann Lecun, Yoshua Bengio, Andrew Ng, ...]

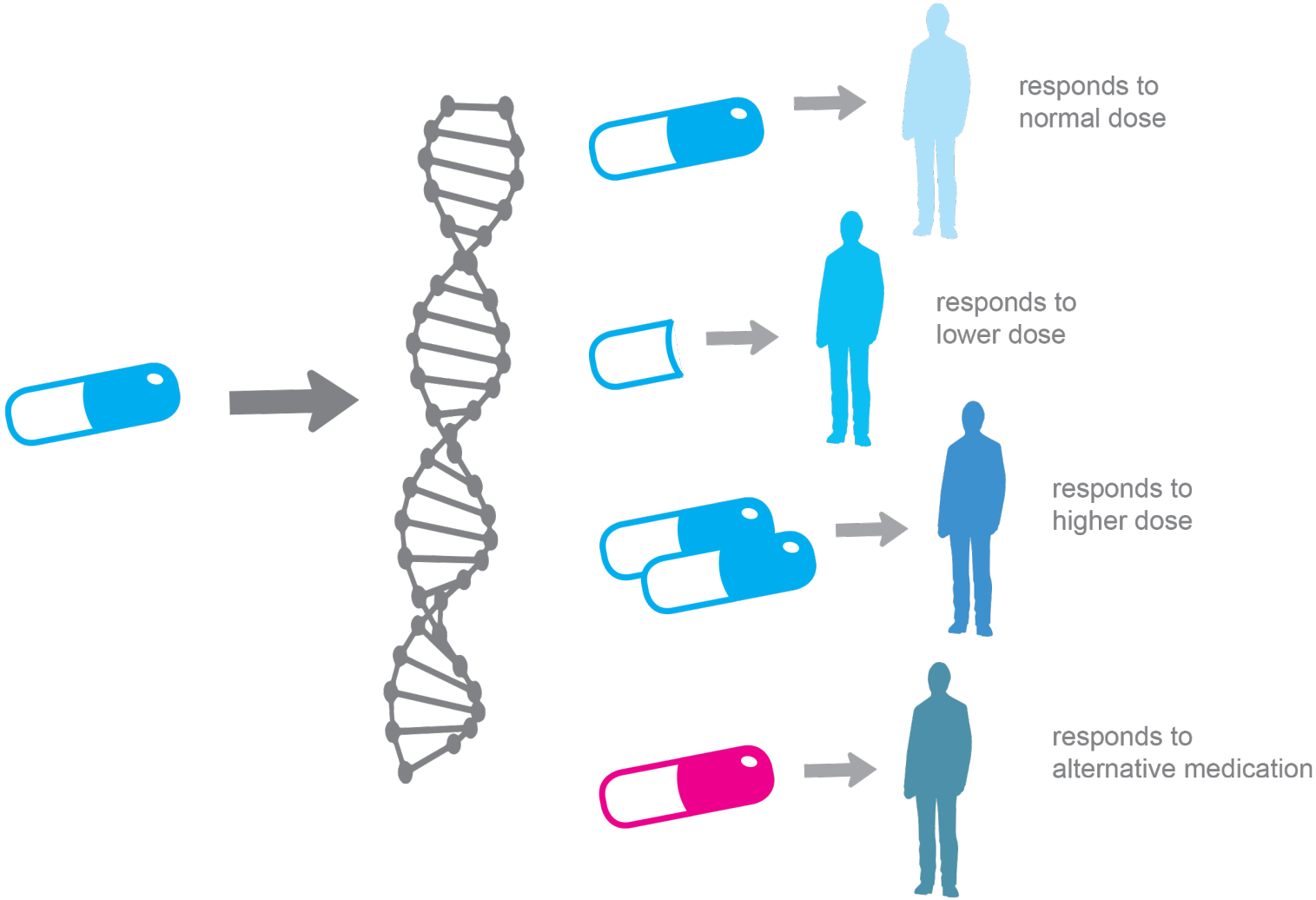
Neural nets playing video games



video by Seth Bling youtube.com/user/sethbling

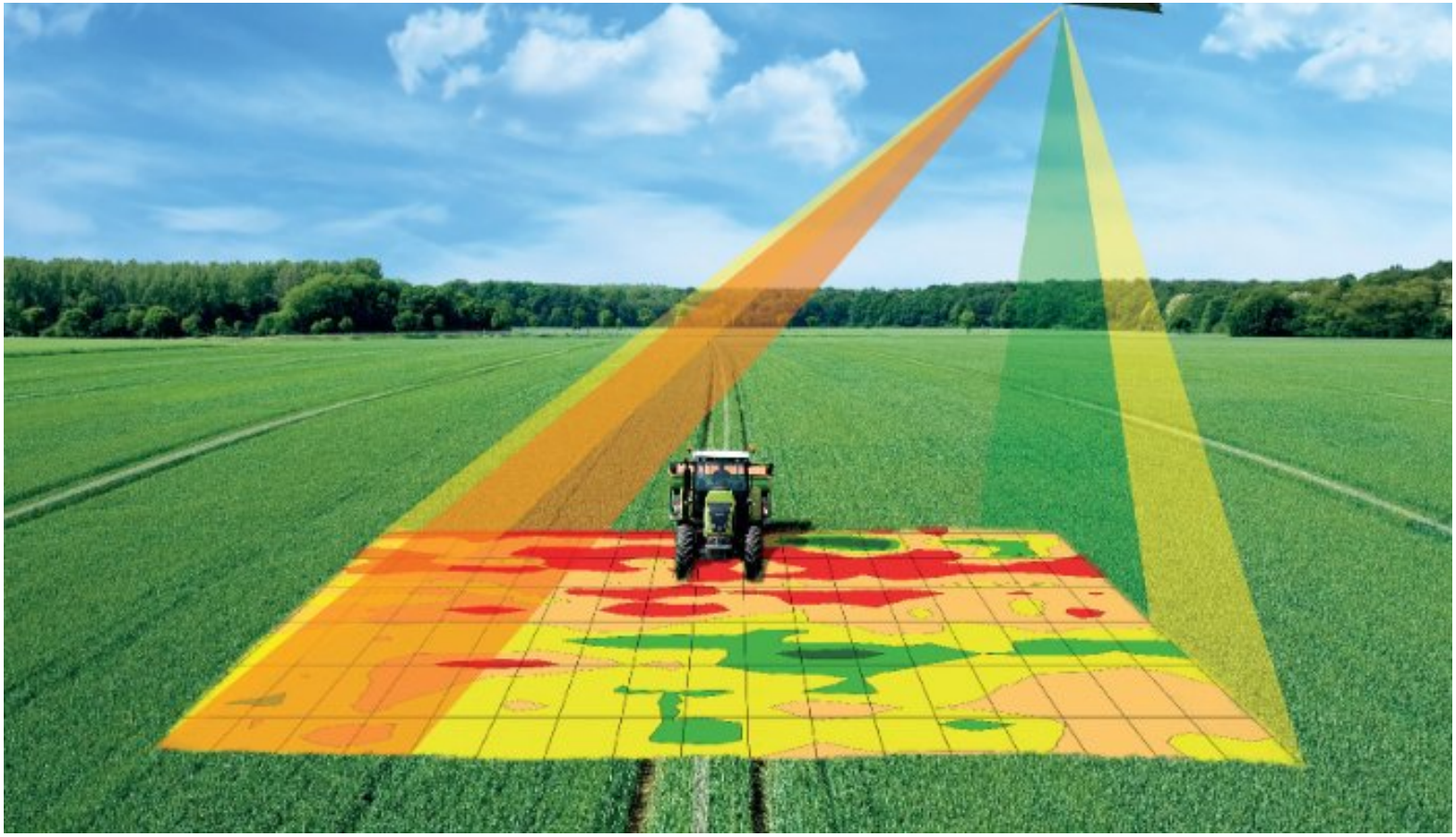
Healthcare

Personalized/precision medicine



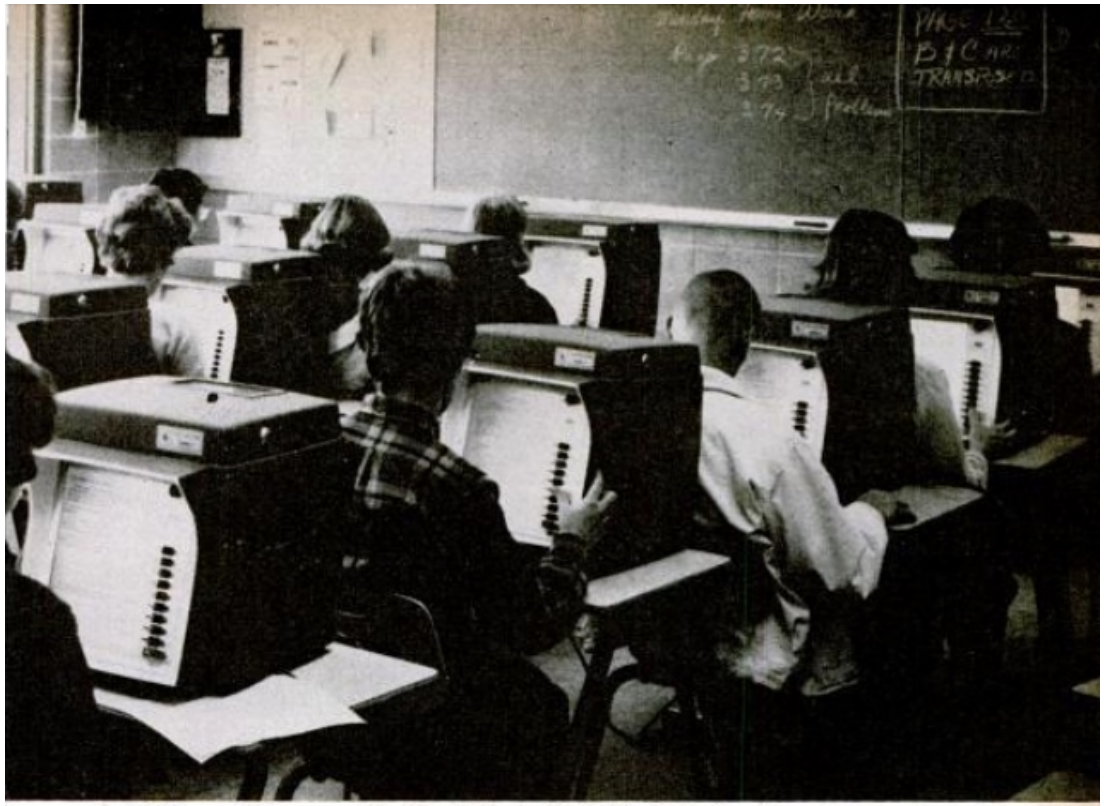
Agriculture & Ecosystem

- Precision farming
- Natural resources



Education

Personalized learning



By 1965, predicts one authority, half of all U. S. students will make use of machines.

Teaching Machines —Do They or Don't They?

Robot teachers are stirring up more to-do in education circles than anything since the invention of chalk

No free lunch!

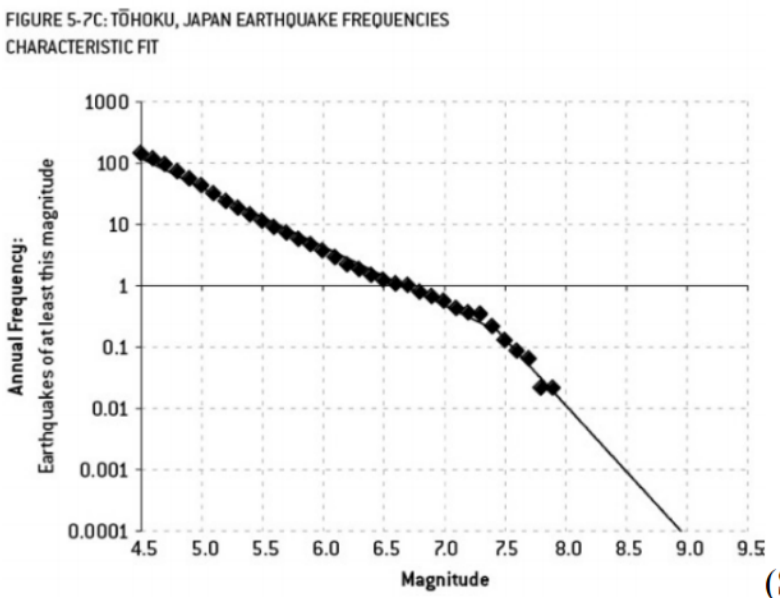
- There is no standard machine learning machine algorithm which can be applicable for every problem and get the best result.
- "Essentially, all models are wrong, but some are useful." George E. P. Box

Failure of Predictive Models

- The Fukushima power plant disaster

Failure of Predictive Models

- Overfitting

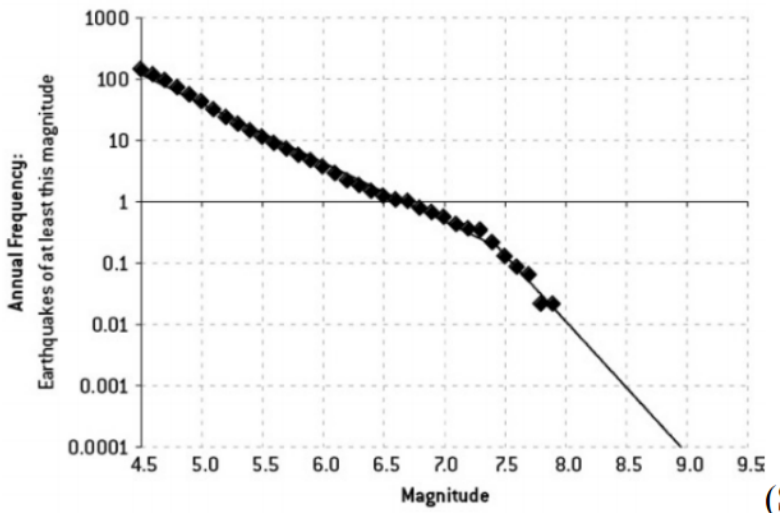


(Silver, N, 2012)

Failure of Predictive Models

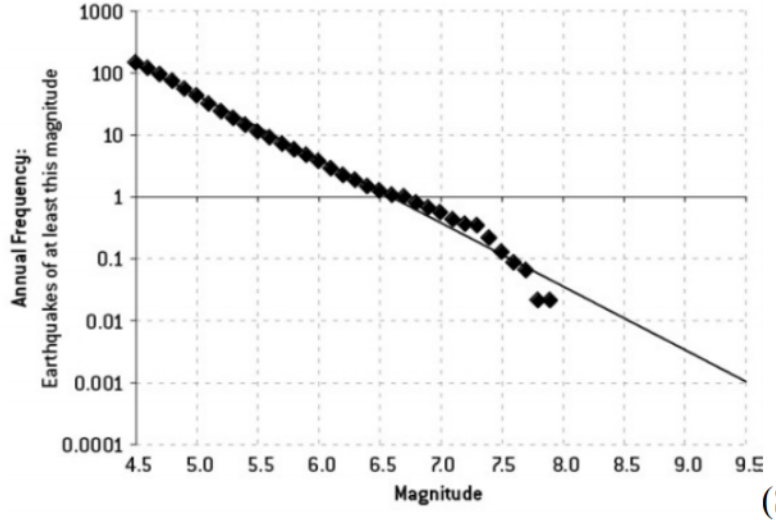
- Overfitting

FIGURE 5-7C: TŌHOKU, JAPAN EARTHQUAKE FREQUENCIES
CHARACTERISTIC FIT

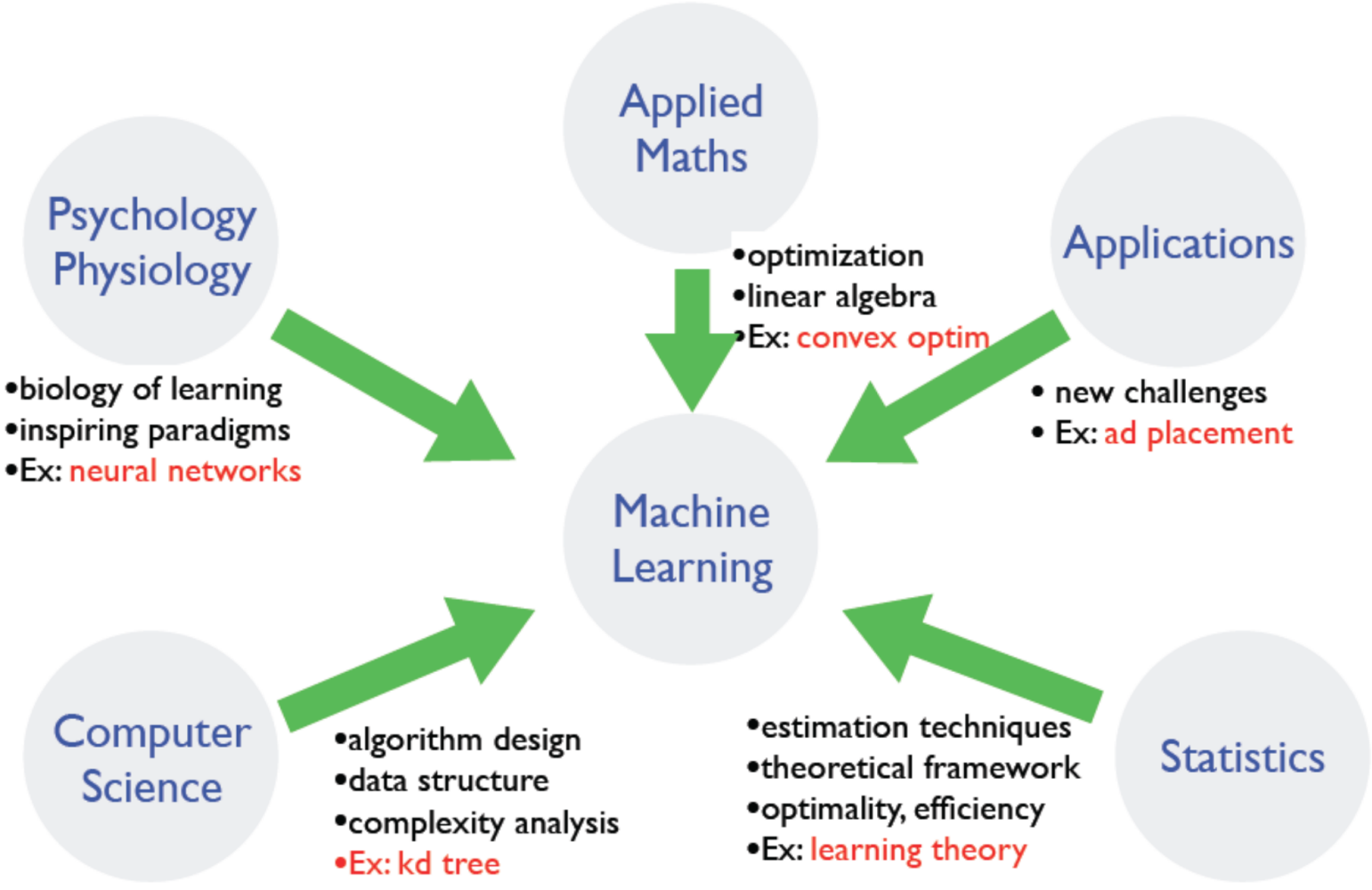


(Silver, N, 2012)

FIGURE 5-7B: TŌHOKU, JAPAN EARTHQUAKE FREQUENCIES
GUTENBERG-RICHTER FIT



(Silv



Slide Credit: Fei Sha

Questions?

Questions?

TODO:

- Homework 1 due 9/22
- Readings: linear algebra and probability background, see Canvas