

# AI circa 2018

- While we wait to start, think about problems in the right...



- How to solve quadratic equation?
  - $ax^2 + bx + c = 0$
- Who is learning better in new airplane mission simulator?

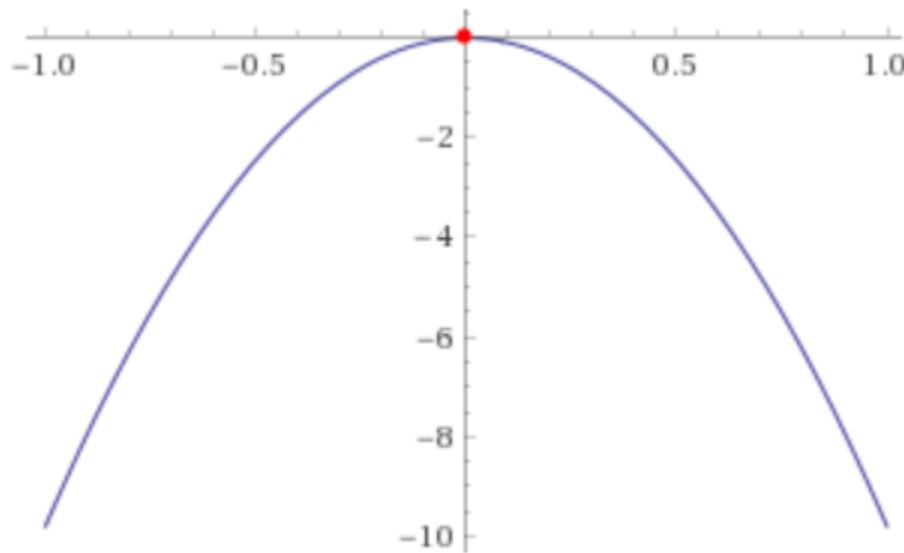
Pilot	Week 1	Week 2
Olivia	0%	75%
Oliver	25%	100%

- What is the next number in the sequence?  
11, 21, 1211, 111221, ?

# Quadratic equation solution

- From Algebra

- $x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$



- Why do we care?



# Going to higher orders...

- Cubic equation... Anyone?
  - $ax^3 + bx^2 + cx + d = 0$
- Specific example...
  - $x^3 - 2x^2 - 5x + 6 = 0$
- What if...
  - We plot it
  - Or could solve it!

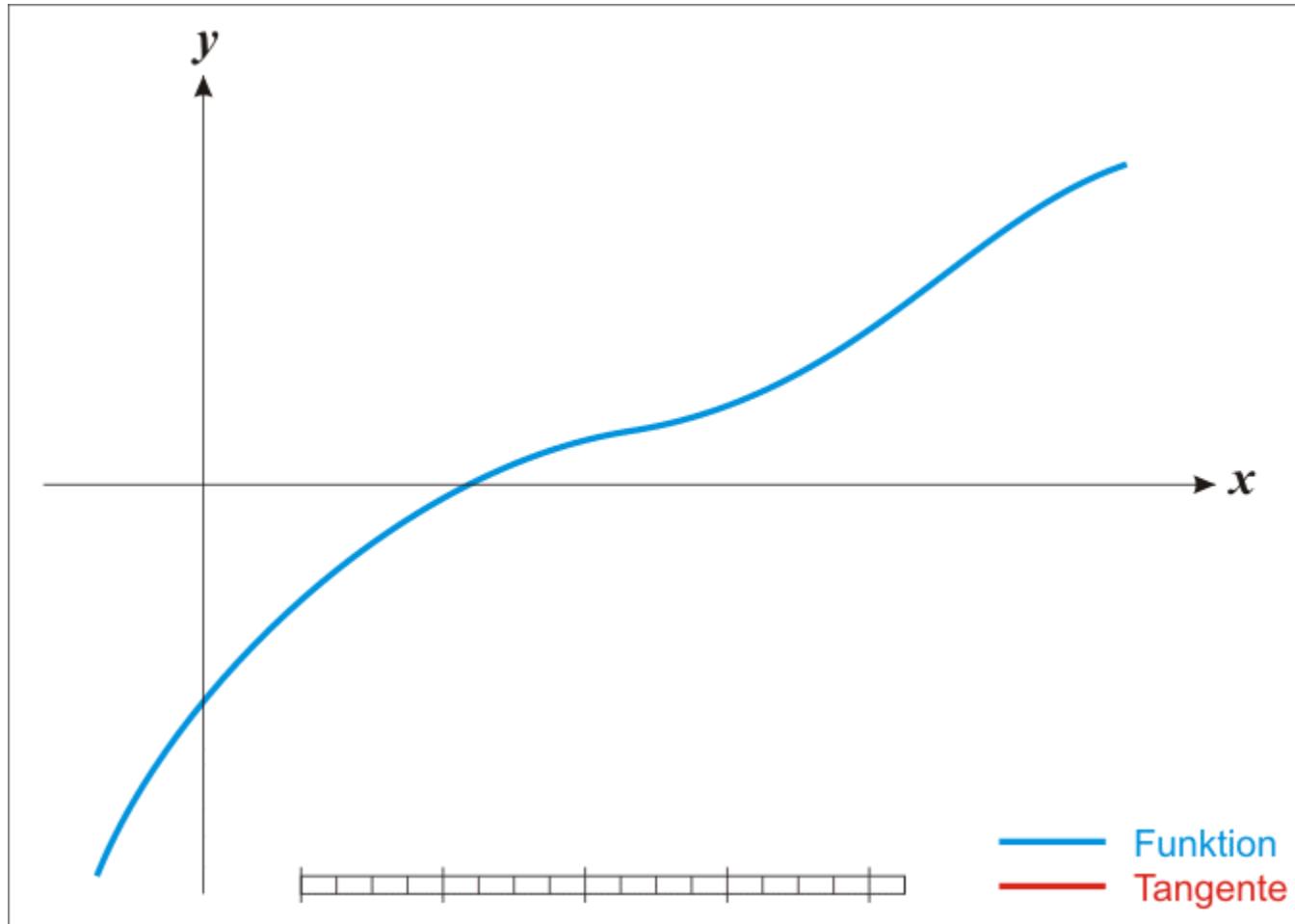
# Evolution of AI

- Math
- Memory
- Models

# Which AI?

- Today
  - Applied AI (a.k.a. “weak AI” or “narrow AI”)
- Out of scope
  - Artificial General Intelligence, a.k.a., “strong AI” or “full AI”
  - Consciousness
    - Like in movies and TV series: [2001: A Space Odyssey](#), [AI](#), [Her](#), [Humans](#), [Westworld](#), ...

# Numerical Calculus: Newton (~1685-1740)

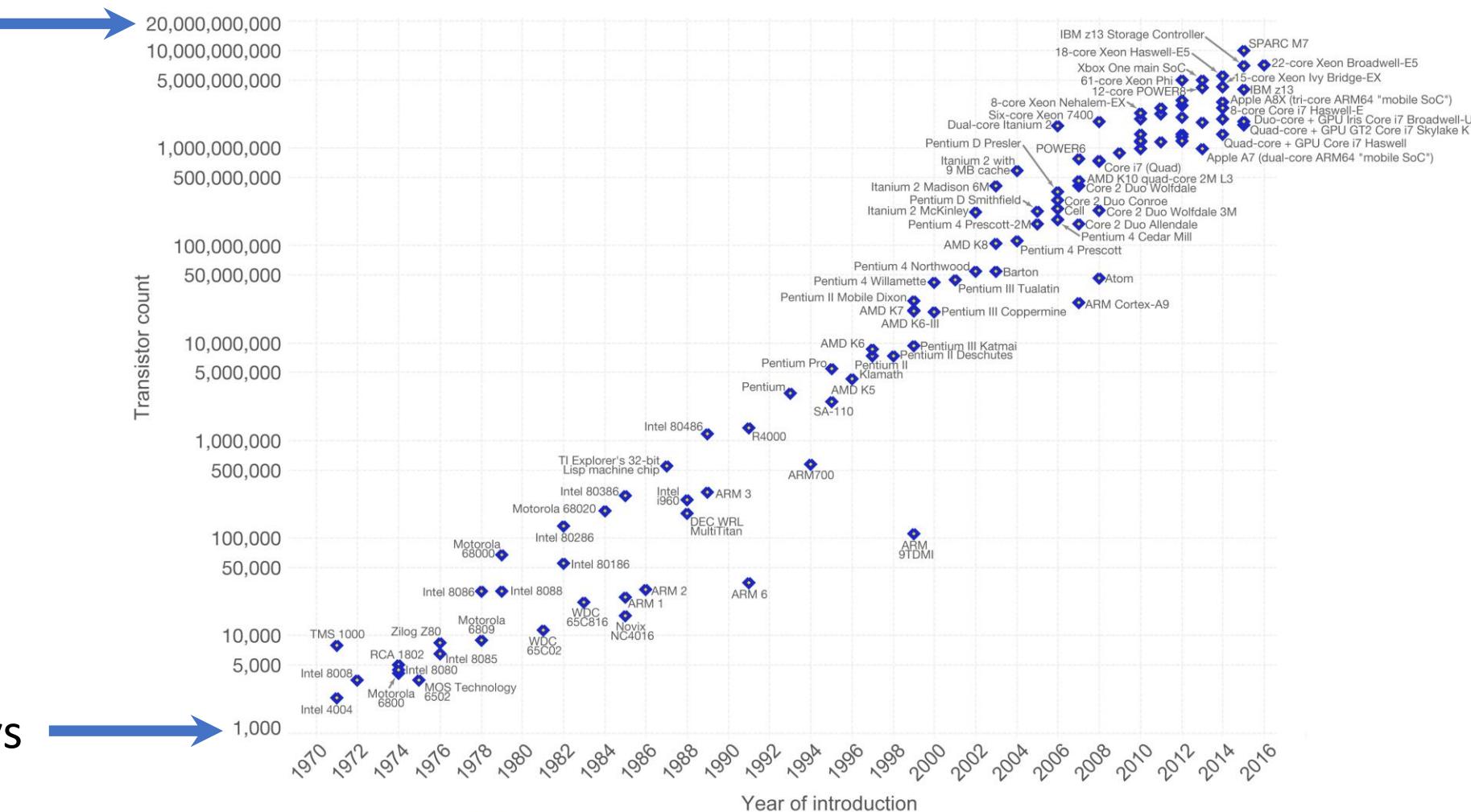


# Moore's Law – The number of transistors on integrated circuit chips (1971-2016)

Our World  
in Data

Moore's law describes the empirical regularity that the number of transistors on integrated circuits doubles approximately every two years. This advancement is important as other aspects of technological progress – such as processing speed or the price of electronic products – are strongly linked to Moore's law.

1 second



Data source: Wikipedia ([https://en.wikipedia.org/wiki/Transistor\\_count](https://en.wikipedia.org/wiki/Transistor_count))

The data visualization is available at OurWorldinData.org. There you find more visualizations and research on this topic.

Licensed under CC-BY-SA by the author Max Roser.

# Usual development

# ML development



# Computational power enabled brute force...

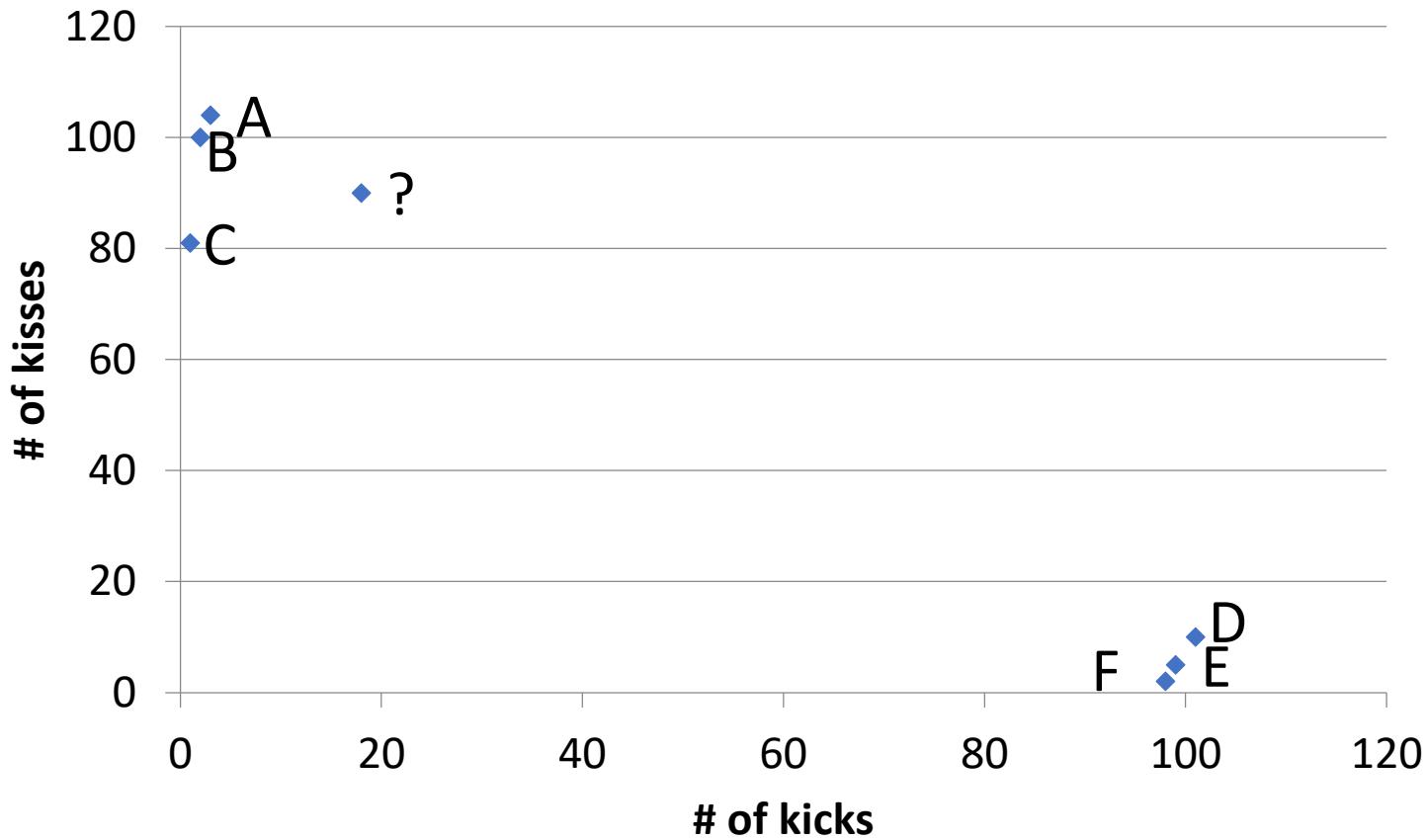
- k-Nearest Neighbors
  - Pros: accuracy, insensitive to outliers, little “data preparation”
  - Cons: computationally expensive
  - Basic idea: “Tell me who you walk with...”

# Movies: data

Movie	# of kicks	# of kisses	Type
A	3	104	Romance
B	2	100	Romance
C	1	81	Romance
D	101	10	Action
E	99	5	Action
F	98	2	Action
?	18	90	Unknown

# Movies: Scatter chart

$$d = \sqrt{(x_a - x_b)^2 + (y_a - y_b)^2}$$



# Movies: Distance

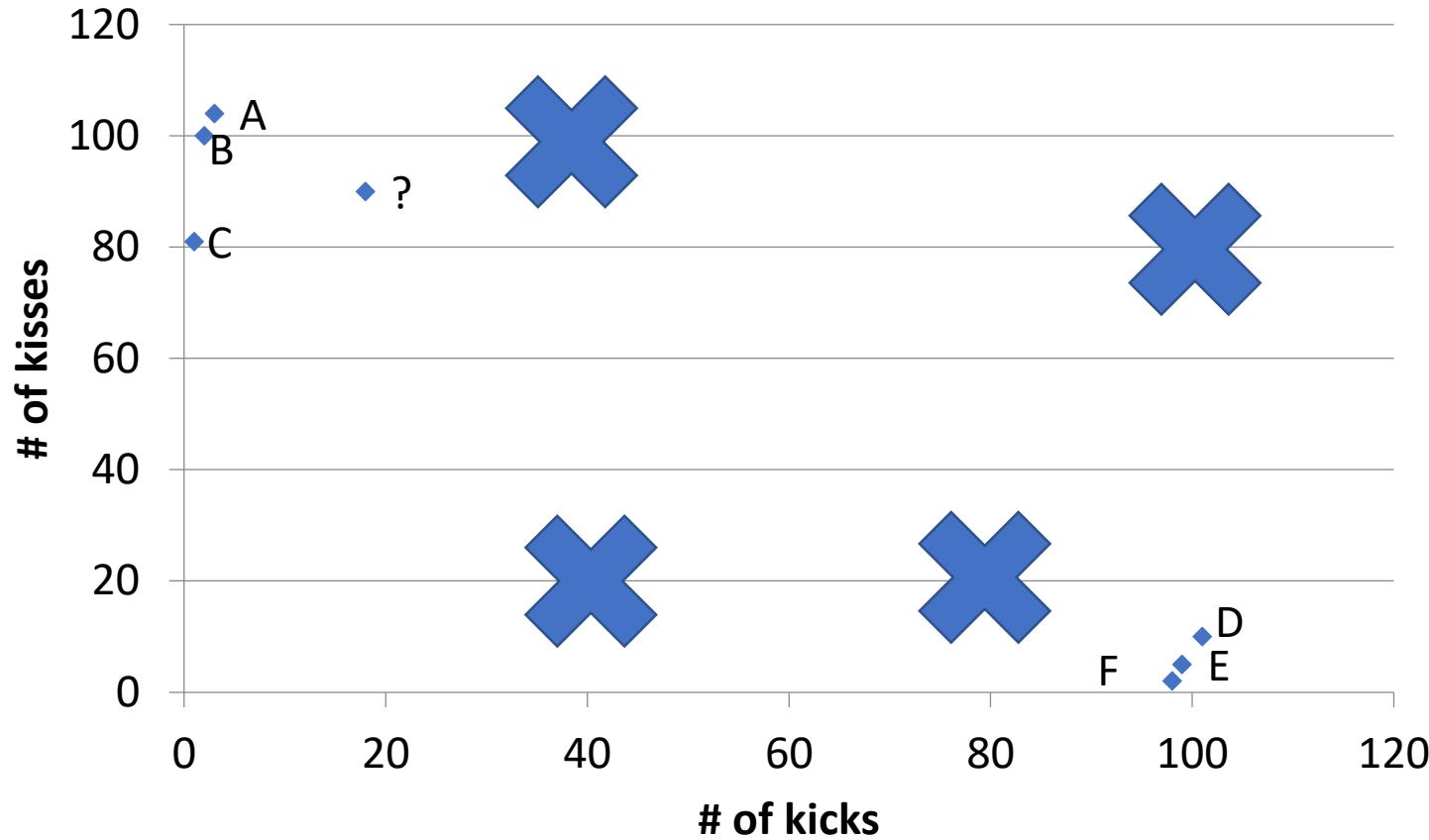
Movie	# of kicks	# of kisses	Type	d?
?	18	90	Unknown	0.0
A	3	104	Romance	20.5
B	2	100	Romance	18.9
C	1	81	Romance	19.2
D	101	10	Action	115.3
E	99	5	Action	117.4
F	98	2	Action	118.9

# Movies: Sorted Distance

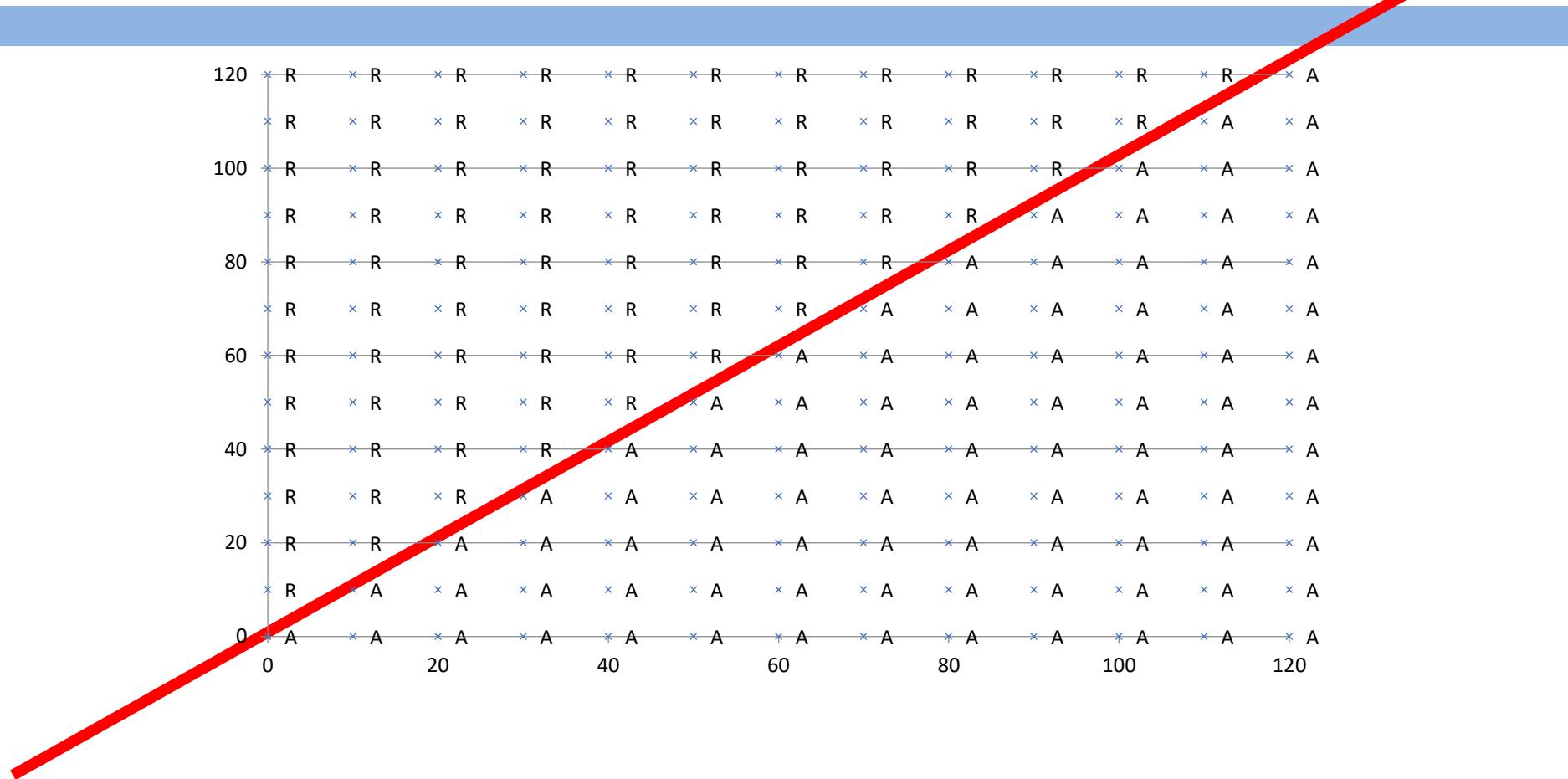
Movie	# of kicks	# of kisses	Type	d?
?	18	90	Unknown	0.0
B	2	100	Romance	18.9
C	1	81	Romance	19.2
A	3	104	Romance	20.5
D	101	10	Action	115.3
E	99	5	Action	117.4
F	98	2	Action	118.9

if kNN = 3NN then...

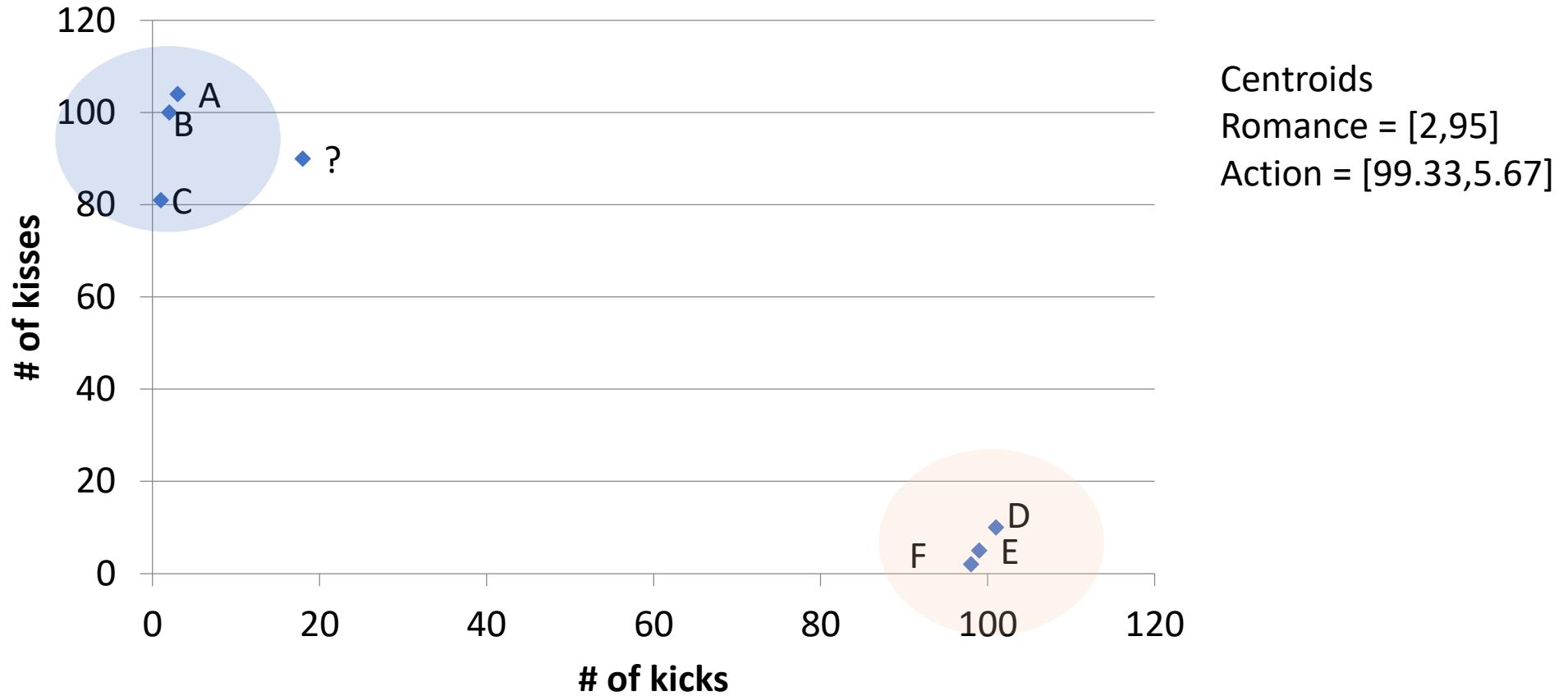
# What If? (BI Scenarios)



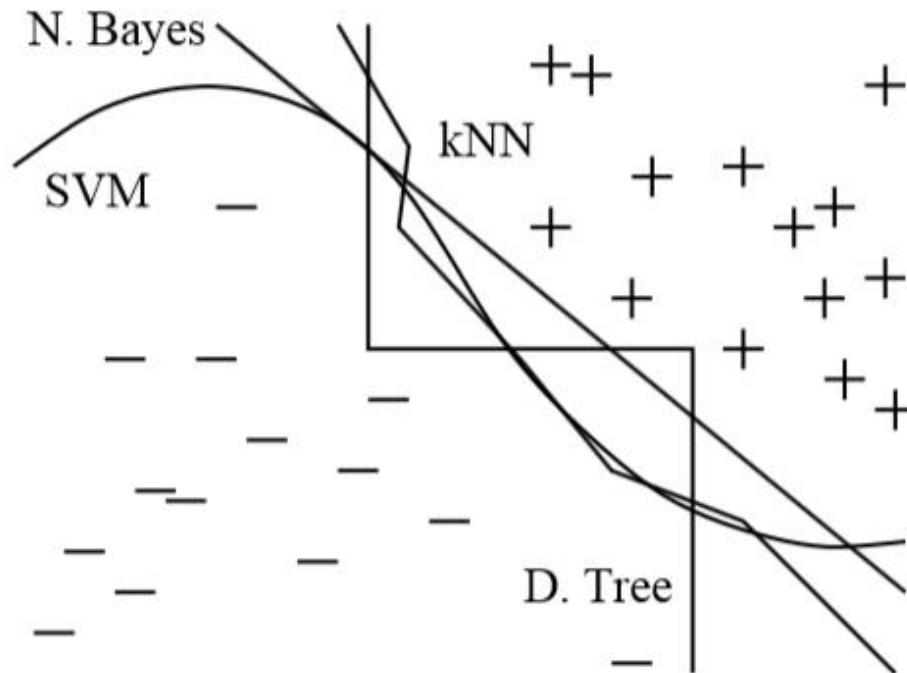
# Pre-calculated “border”



# Cluster representation



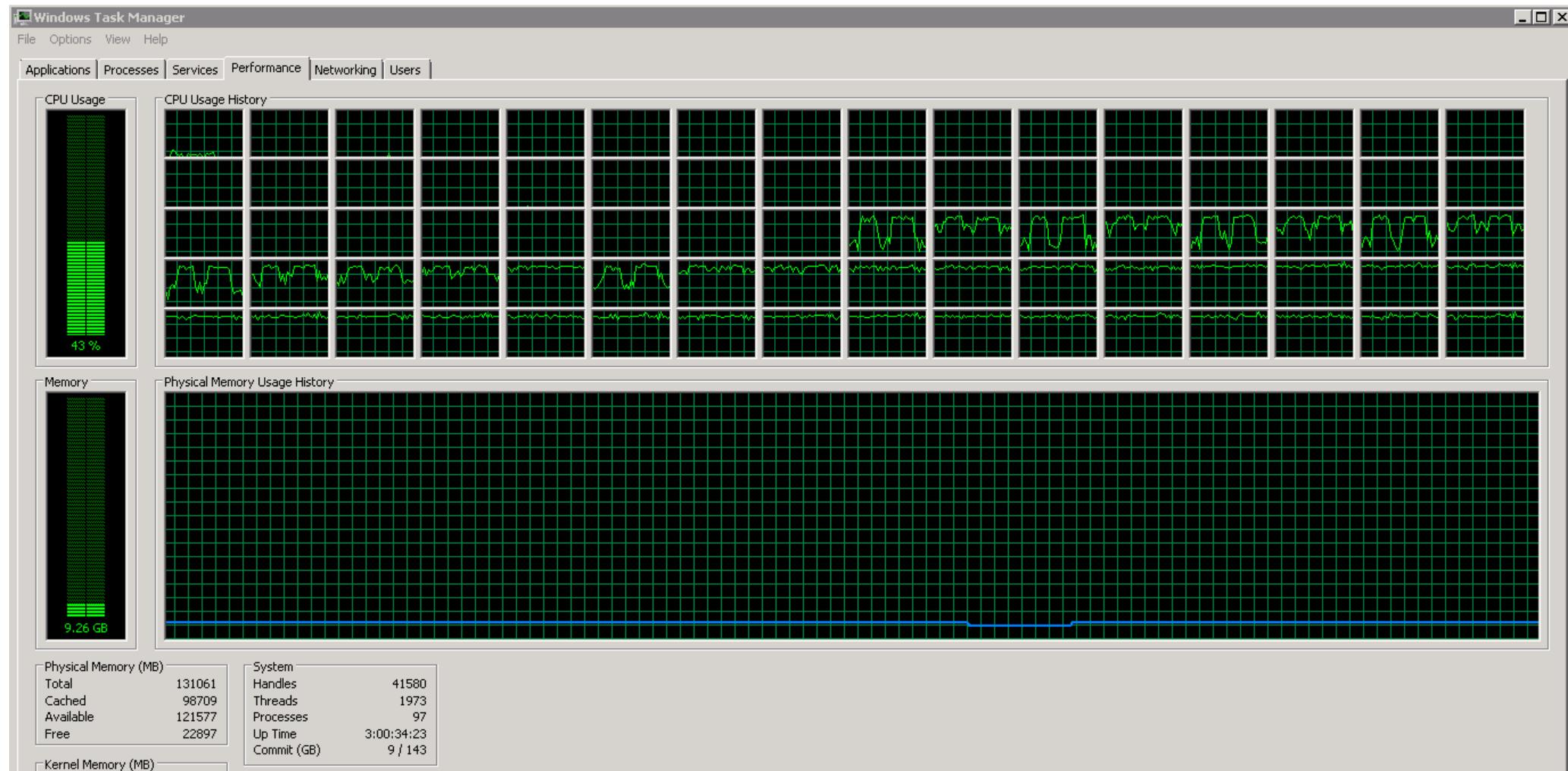
# Right method and correctly implemented



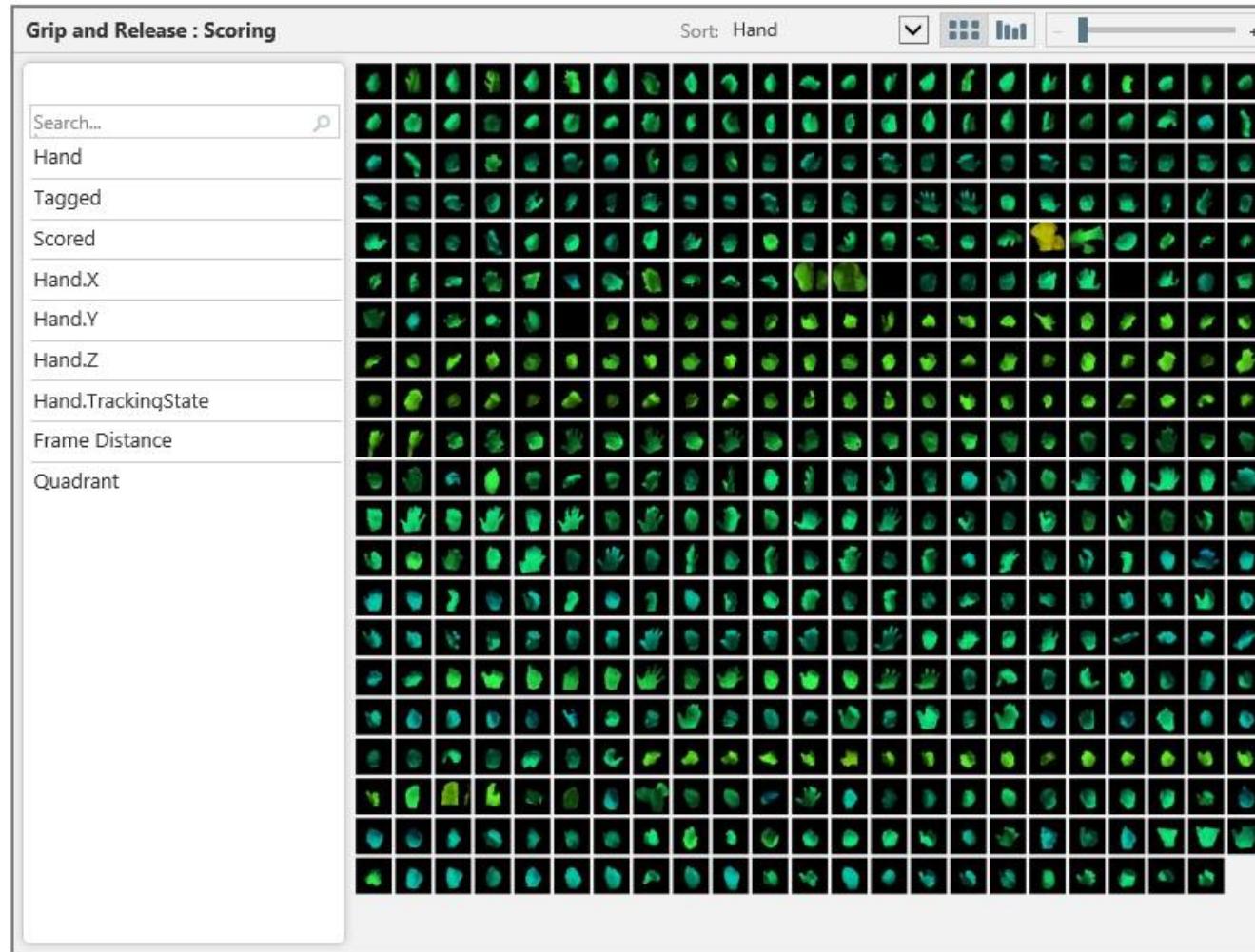
[A Few Useful Things to Know About Machine Learning,](#)  
by Pedro Domingos

**Figure 3:** Very different frontiers can yield similar class predictions. (+ and – are training examples of two classes.)

# Training is costly...

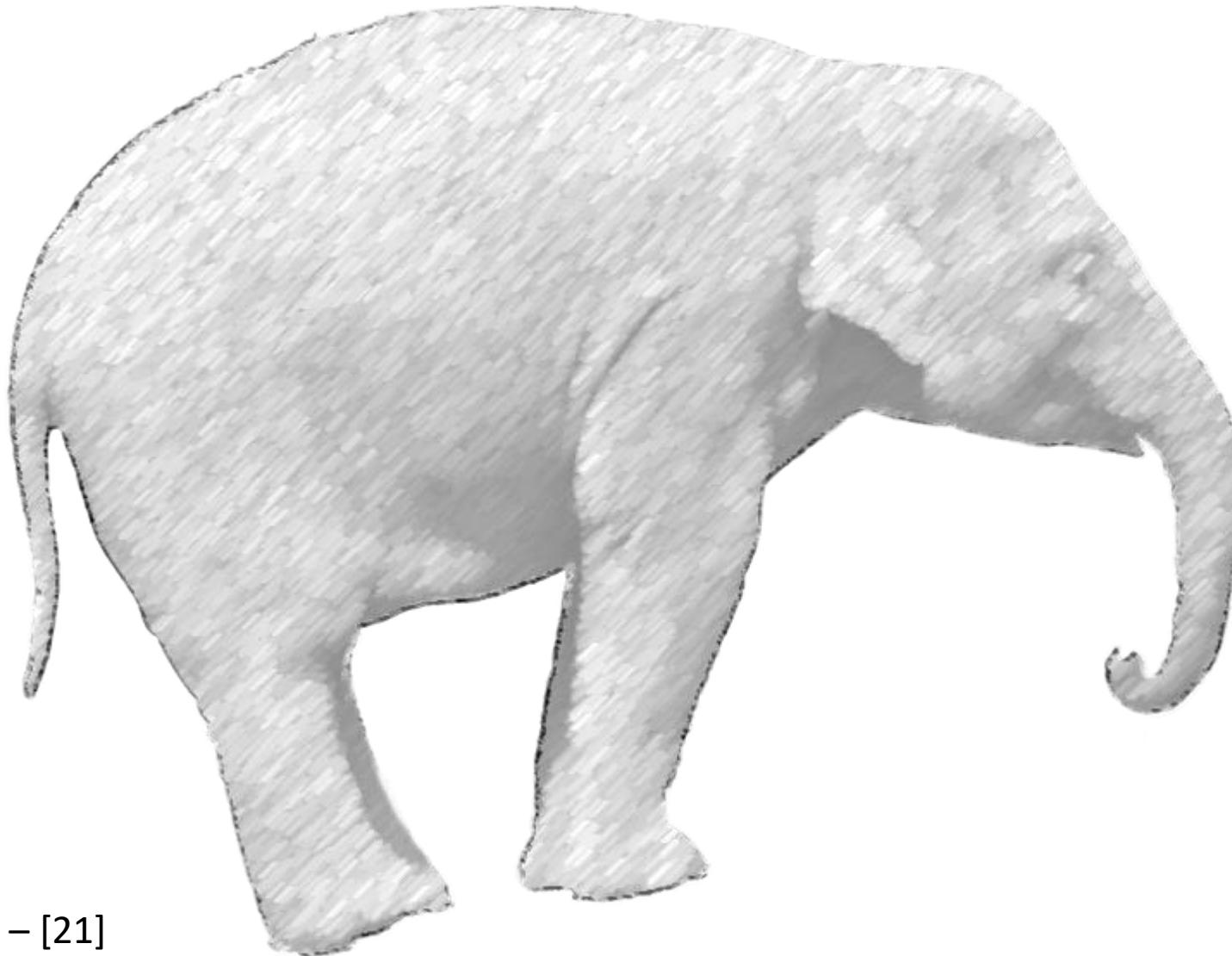


# Data analysis for an experiment



From Kinect for Windows  
presentation during  
Microsoft Build 2013 event

# What is this?



# What is Deep Learning?



# What is deep learning?

- Deep learning (also known as deep structured learning, hierarchical learning or deep machine learning) is a class of machine learning algorithms that: use a cascade of many layers of nonlinear processing units for feature extraction and transformation.
- But what \*is\* a Neural Network?
- Visualize: <http://playground.tensorflow.org/>

# Who is learning better?

- Two pilots have been learning how to complete a difficult mission in an airplane simulator. Rates of success per week are in the table.

Pilot	Week 1	Week 2	Aggregated
Olivia	0%	75%	
Oliver	25%	100%	

# Simpson's paradox

- Trend in different groups of data disappears or reverses when these groups are combined (a.k.a. **reversal** or **amalgamation** paradox)

Pilot	Week 1	Week 2	Aggregated
Olivia	(0/1) 0%	(3/4) 75%	(3/5) 60%
Oliver	(1/4) 25%	(1/1) 100%	(2/5) 40%

# Accuracy paradox

Confusion matrix

	Predictive Positive	Predictive Negative
Positive samples	True Positive	False Negative
Negative samples	False Positive	True Negative

Accuracy

$$\bullet A = \frac{TP+TN}{TP+FP+TN+FN}$$

# Model comparison by accuracy

Model A

	Predictive Positive	Predictive Negative
Positive samples	100	50
Negative samples	150	9,700

$$A_{(A)} = \frac{100 + 9,700}{100 + 150 + 9,700 + 50} = 0.98$$

Model B

	Predictive Positive	Predictive Negative
Positive samples	1	149
Negative samples	1	9,849

$$A_{(B)} = \frac{1 + 9,849}{1 + 1 + 9,849 + 149} = 0.985$$

# F1 Score

## Confusion matrix

	Predictive Positive	Predictive Negative
Positive samples	True Positive	False Negative
Negative samples	False Positive	True Negative

## Measurements (Precision, Recall, F1)

- $P = \frac{TP}{TP+FP}$
- $R = \frac{TP}{TP+FN}$
- $F_1 = 2 \frac{P*R}{P+R}$

# Model comparison by F1 Score

Model A

	Predictive Positive	Predictive Negative
Positive samples	100	50
Negative samples	150	9,700

$$P = 0.4, R = 0.67, F_1 = 0.5$$

Model B

	Predictive Positive	Predictive Negative
Positive samples	1	149
Negative samples	1	9,849

$$P = 0.5, R = 0.0067, F_1 = 0.013$$

# Takeaways: Math and AI

- Computers evolved enabling expanded use of Numerical Calculus
- Someone in the team needs to understand the Math
- Math may be right, yet its interpretation may be incorrect

# Any questions before we go to “Memory”?

## ~~• Math~~

- Memory
- Models

# 这是什么？



# 这是什么？



这是什么？



# What just happened?



# Your mind...

- Analyzed images
- Split it into components
- Associated such components with a verbal description

# 你有这段记忆吗？



# 你有这段记忆吗？



# 你有这段记忆吗？



# Playing with Image Search

- <https://images.google.com/>
- <https://www.bing.com/images/>

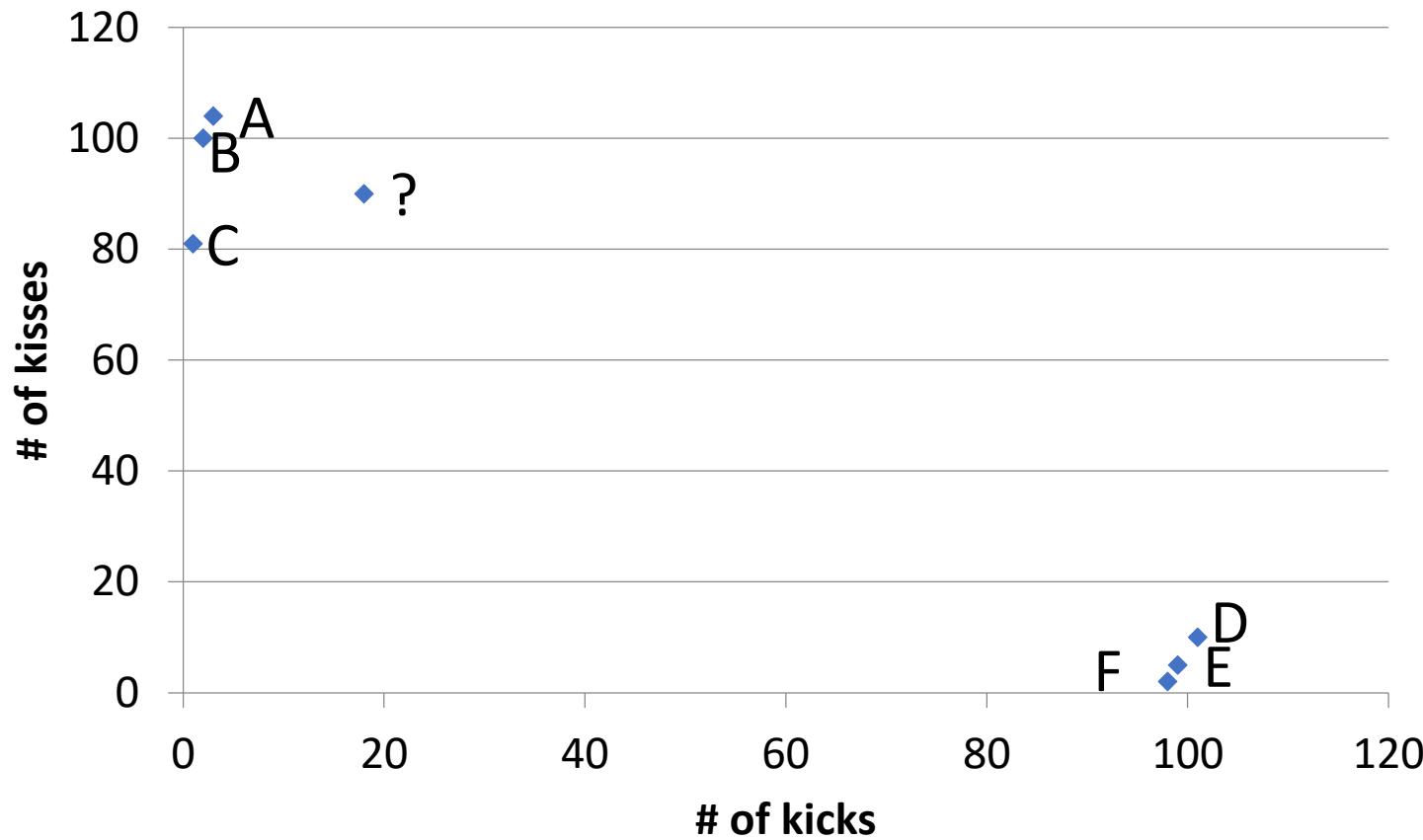
# Your thoughts

- How important are the features for the search?
- “*An image is worth a thousand words*”
  - Would any 1,000 words be equivalent?

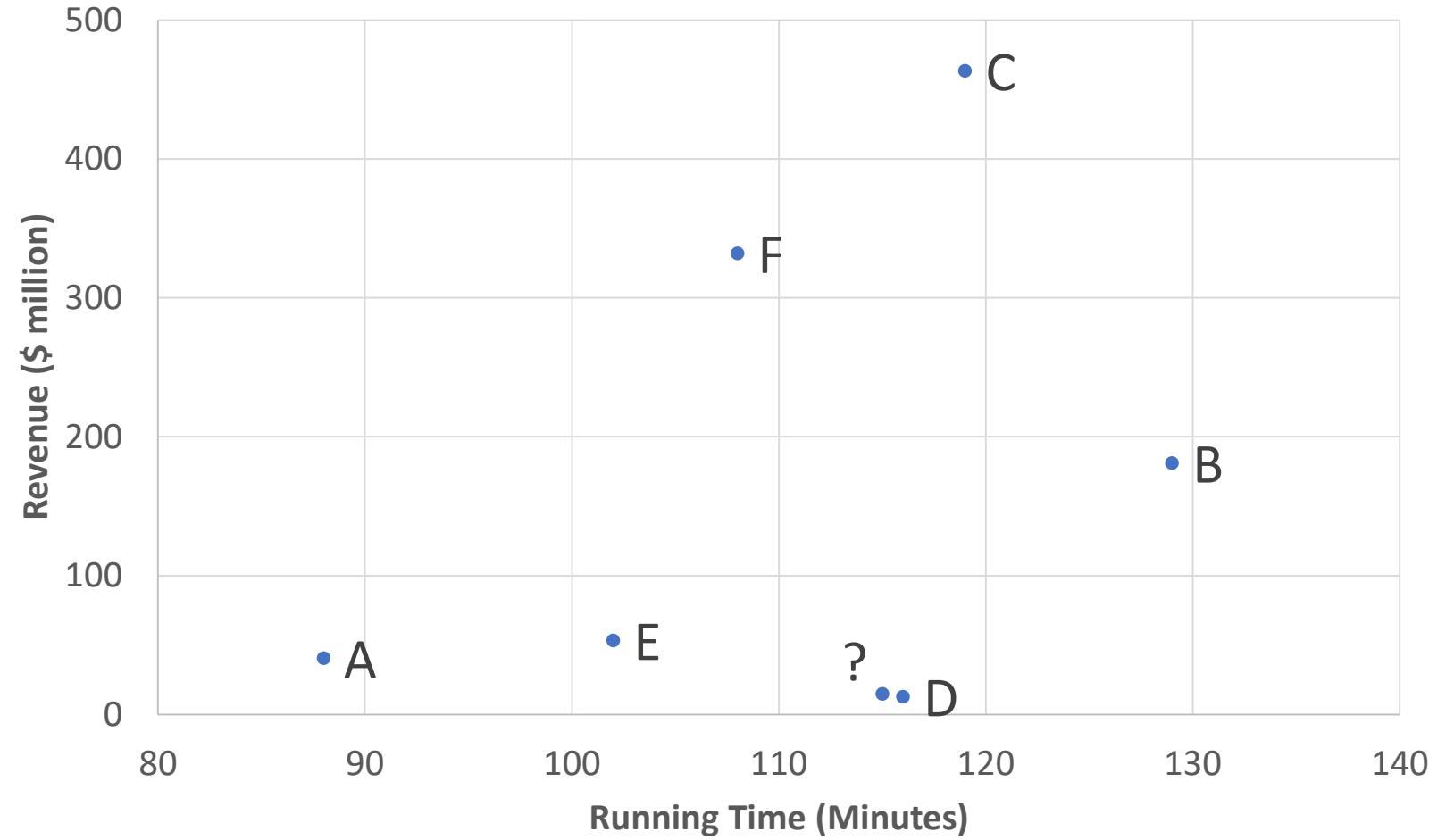
# Movies: another dataset

Movie	Running Time (Minutes)	Revenue (\$ million)	Type
A	88	40.7	Romance
B	129	181.1	Romance
C	119	463.4	Romance
D	116	12.9	Action
E	102	53.4	Action
F	108	332	Action
?	115	15	Unknown

# Movies: Scatter chart: kicks and kisses



# Movies: Scatter chart: revenue and duration



# Orphaned ML/AI projects

- Pattern for failed ML project
  - Data was accumulated: Volume, Variety, Velocity, Veracity
  - Model was built with “potential”: success for initial anecdotes
  - Then: different questions or need for more precise answers
  - Result: project collapses
- Corollary: data is thrown away or lost
  - <https://toolbox.google.com/datasetsearch>

## The Human Brain Project Reboots: A Search Engine for the Brain Is in Sight

The massive €1 billion project has shifted focus from simulation to informatics

By Megan Scudellari



CAN WE  
COPY THE  
BRAIN?

Section 2:  
The Mechanics  
of the Mind

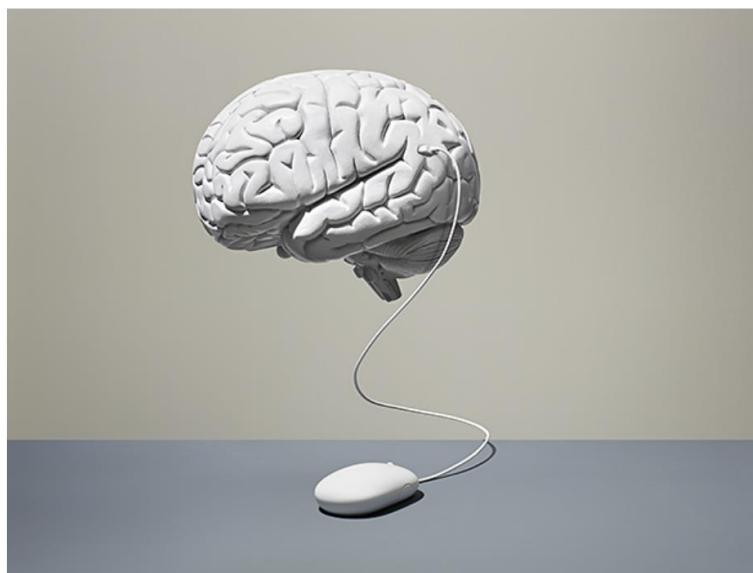


Photo: Dan Saelinger

**Human Brain Project**

Science ▾ Platforms ▾ Collaborate ▾ Follow HBP ▾ About ▾ Education ▾

Welcome to the Human Brain Project

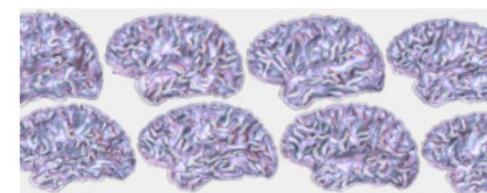
The Human Brain Project aims to put in place a cutting-edge research infrastructure that will allow scientific and industrial researchers to advance our knowledge in the fields of neuroscience, computing, and brain-related medicine

[Learn more about the project](#)

3D-Polarized Light Imaging of the human hippocampus.  
Image: Aixer, Amunts and team, Jülich.

Explore the Brain
Brain Simulation
Silicon Brains
Understanding Cognition
Medicine
Robots
Massive Computing
Social, Ethical, Reflective

### News



TUESDAY, 28 AUGUST 2018

Individual Brain Charting: A high-resolution brain map of cognitive functions

MONDAY, 27 AUGUST 2018



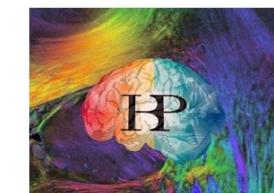
The pioneering partnership of Supercomputing and Neuroscience in Europe

THURSDAY, 23 AUGUST 2018



The Human Brain Project launches voucher programme

THURSDAY, 16 AUGUST 2018



Information concerning the Human Brain Project's Coordination Office

### Events

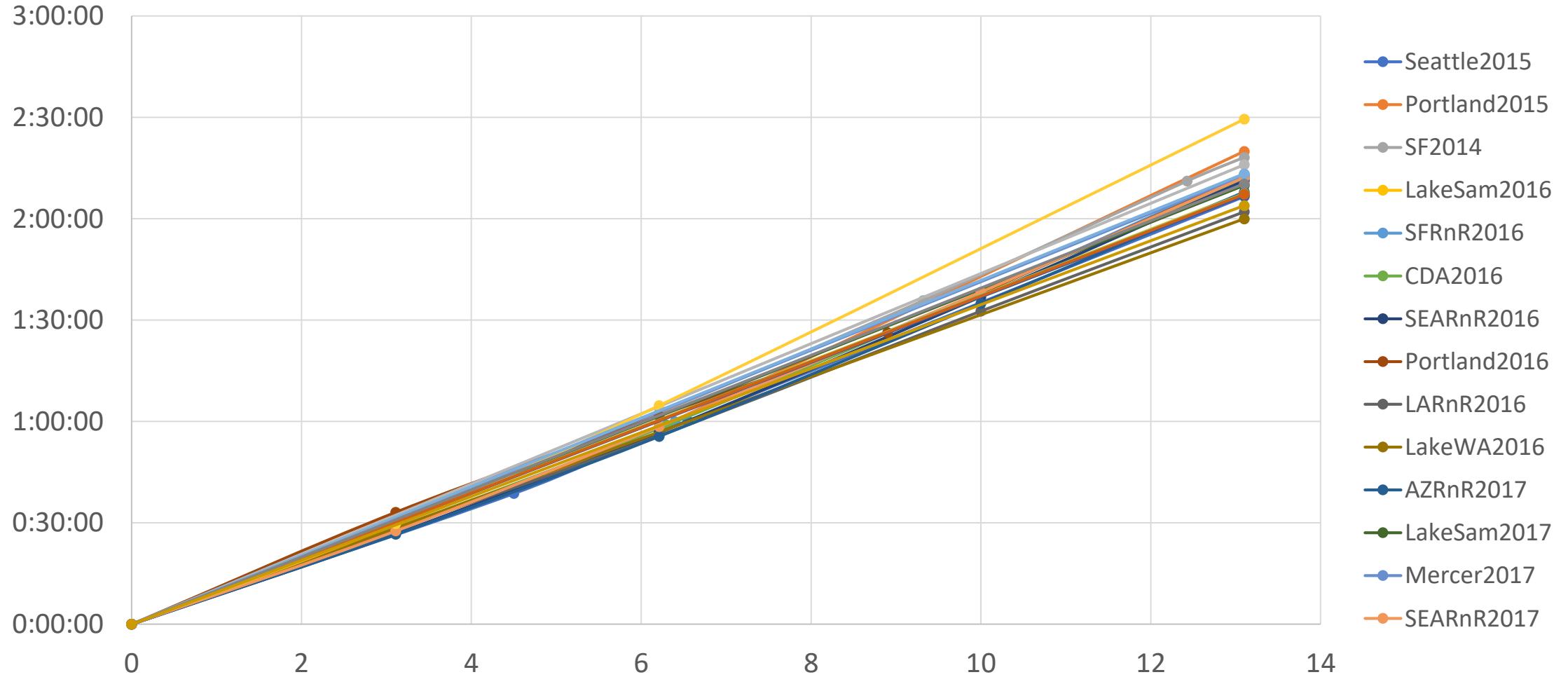
MONDAY, 15 OCTOBER 2018  
HBP Open Day 2018 - Maastricht  
📍 MECC Maastricht

THURSDAY, 4 OCTOBER 2018  
HBP Colloquium at  
Forschungszentrum Jülich  
📍 Central Library, Forschungszentrum Jülich

MONDAY, 17 SEPTEMBER 2018  
The Brain Simulation Platform -  
HBP School  
📍 Mondello (Palermo), Italy

[VIEW ALL EVENTS](#)

# Personal data collection: half-marathons



# Takeaways: Memory and AI

- Human memory is still poorly understood
- Try the reverse of current trend
  - Start with the questions
  - Build a model (proof of concept)
  - Then seek for the data
  - Iterate...
    - Everything: questions, model, data

# Any questions before we go to “Models”?

- ~~Math~~
- ~~Memory~~
- Models

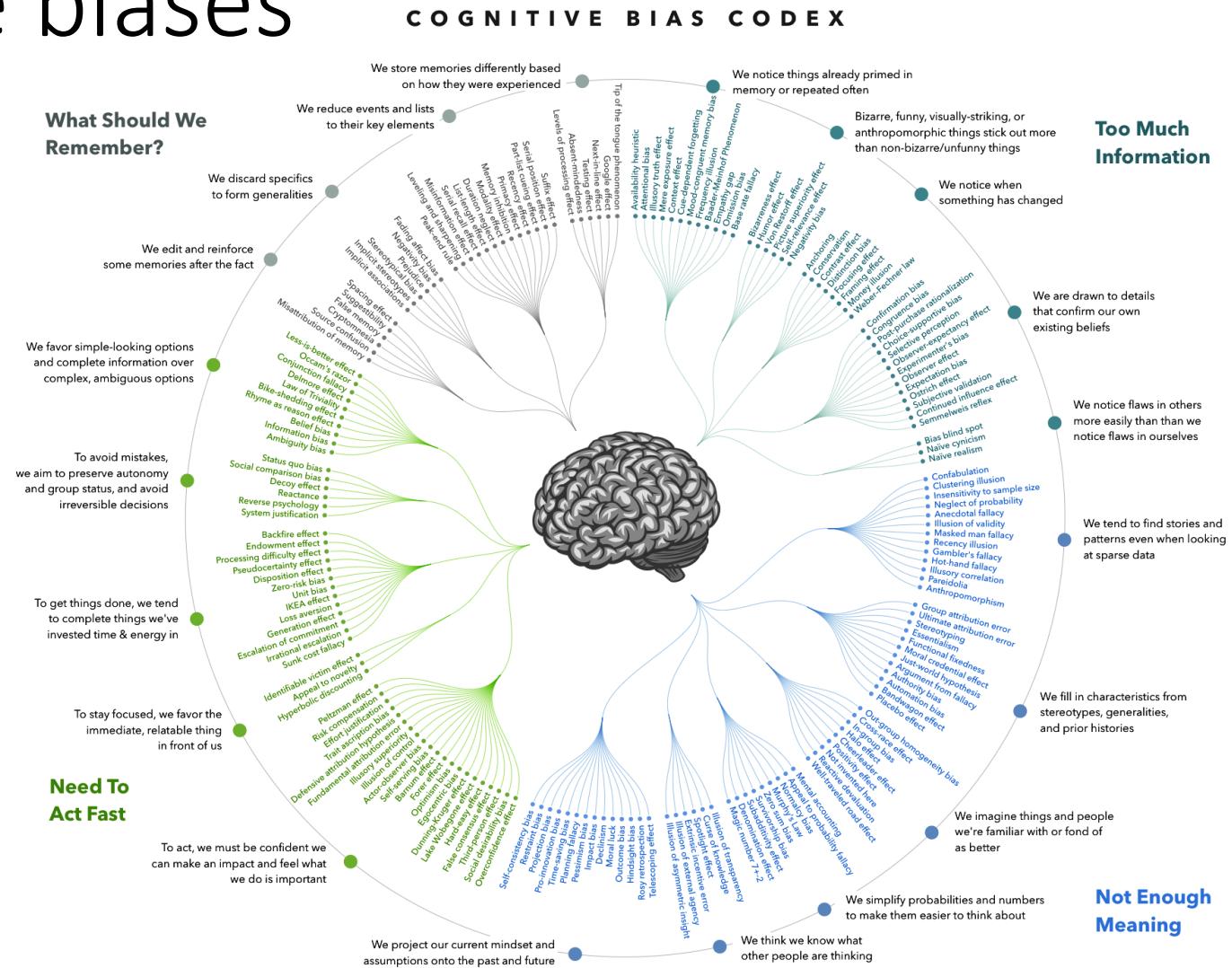
# Humans and modeling: our intuition is bad!

... *perceive*  
or *infer*...

- Intelligence: “... can be more generally described as the ability to **perceive or infer information**, and to retain it as knowledge to be applied towards adaptive behaviors within an environment or context.”
- Earth = center of universe, flat
- Birds fly, have feathers, ... from Icarus to Leonardo da Vinci
- Humans don't understand the “features”
  - Sports, stock market, relationships, ...
  - Models include abstractions: “gravity”

# Cognitive biases

*... perceive  
or infer...*



# Biases and modeling

*... perceive  
or infer...*

- Decision-making, belief, behavior
  - Dunning–Kruger effect
  - Irrational escalation (sunk cost)
  - Parkinson's law of triviality (bikeshedding)
- Social biases
  - Illusory superiority (Lake Wobegon effect)
  - Fundamental attribution error
- Memory errors and biases
  - Bizarreness effect
  - Hindsight bias

# How is Physics applied?

... perceive  
or infer...

- Acceptance of multiple models
  - No wrong or right model: *the best predictor wins the day!*
- Acceptance of imprecise answers
  - Simplifies models, eliminating a lot of noise
  - $F = ma$  and  $E = mc^2$  only for ideal conditions
- Detaching models from data
  - [Nobel Prize in Physics 2017](#): Rainer Weiss, Barry C. Barish, Kip S. Thorne  
LIGO: Laser Interferometer Gravitational-Wave Observatory  
confirmed predictions by Albert Einstein 100 years ago

# Applying AI

*... perceive  
or infer...*

- Do you need to solve the general problem?
- Going from applied ML to applied AI

# Airplane autopilot

- Ship's gyroscopic-compass set.  
# [1,242,065](#), Elmer A Sperry
- Automatic pilot for airplanes  
#[1,707,690A](#), Lawrence B Sperry
  - Control of position, course or altitude of land, water, air, or space vehicles, e.g. automatic pilot with safety arrangements for transition from automatic pilot to manual pilot and vice versa

Stanford University



MS&E 238 Blog

Leading Trends In Information Technology

[Home](#) [About MS&E238](#) [MS&E238 Class Website](#)

#### RECENT POSTS

[Artificial Confucius Intelligence.](#)

[Something controversial](#)

[IoT and AgTech](#)

[Where Eagles Dare](#)

[What's cooking? IoT in the Kitchen](#)

#### RECENT COMMENTS

[Joshua Ching on Blockchain + AI](#)

[Joshua Ching on IoT and AgTech](#)

[ccjw on Why Investors Really Care about Impact Investing](#)

[ccjw on Why Investors Really Care about Impact Investing](#)

[ccjw on Blockchain + AI](#)

#### CATEGORIES

[summer2017 \(277\)](#)

[Uncategorized \(8\)](#)

[Week1 \(49\)](#)

[Week2 \(50\)](#)

[Week3 \(50\)](#)

[Week4 \(46\)](#)

[Week5 \(49\)](#)

#### TAGS

## Flying Smarter: AI & Machine Learning in Aviation Autopilot Systems

by Ben Garlick on August 4, 2017 10:49 am

Categories: [summer2017](#)

Tags: [AI](#), [artificial intelligence](#), [Aviation](#), [machine learning](#)

Many experienced airline travelers have experienced a common scenario- As the boarding time approaches, the gate agent announces that there will be a slight delay since the pilots are running behind schedule on their current flight. The passengers continue to wait as the plane and flight attendants sit idle on the tarmac. Several companies are working to transform the cockpit of the future to reduce or potentially eliminate the requirement for manned pilots, which would make this experience a relic of the past.

The average passenger is often unfamiliar with the level of autonomy currently provided by the autopilot systems in use today in commercial aviation. These systems have been around in various forms since the 1980's. For commercial flights on a modern aircraft with greater than 18 passenger seats, the airplanes flight management system (FMS) and its associated autopilot functions are generally in control of the aircraft from shortly after takeoff until landing and rollout on the runway under normal operations. Instead of flying the aircraft manually through the flight controls, the crew manages the aircraft's systems through the FMS interface. Nearly all major airports utilize the CAT IIIb "autoland" instrument landing system approaches where the pilot is only a backup to the automated landing system in the event of a failure and takes over once the plane is on the ground to taxi to the gate. CAT IIIc approaches that would extend this autonomous control through the taxi process are in development in many places as well. These automated systems do have their limitations. In the event of a mechanical issue or extreme environmental issues such as severe turbulence, the autopilot may disengage and alert the pilot to take manual control of the aircraft. The system is also only as "smart" as the pilot that is inputting the information, and they require careful attention and oversight to ensure that the system is functioning as required. Many high profile incidents, including the tragic [Air France Flight 447](#) crash in 2009, have resulted from miscommunications between the aircrew and the autopilot system. The pilot is still ultimately responsible for flying the aircraft, and these systems function as a kind of advanced aerial cruise control that simplifies and in many cases eliminates the hands-on flight control manipulation to allow the pilots to focus on communicating, navigating, and managing the flight as a whole.

# Applied Machine Learning

... *perceive  
or infer...*

- Get a goal
- ***Make a goal-related model***
- Collect data
- Test hypothesis
- Apply prediction
- Maximize flight occupation
- ***Model for no-shows***
- Gather ticket info
- Check predictions for no-shows
- Actions based on predictions
  - Reminders/penalty for no-show
  - Allow more overbooking

# Passenger-Based Predictive Modeling of Airline No-show Rates

... perceive  
or infer...

Richard D. Lawrence  
IBM T. J. Watson Research Ctr  
P. O. Box 218  
Yorktown Heights, NY 10598  
[ricklawr@us.ibm.com](mailto:ricklawr@us.ibm.com)

Se June Hong  
IBM T. J. Watson Research Ctr  
P. O. Box 218  
Yorktown Heights, NY 10598  
[sjhong@us.ibm.com](mailto:sjhong@us.ibm.com)

Jacques Cherrier  
Air Canada  
P.O. Box 9000  
Dorval, Quebec H4Y 1C2  
[jacques.cherrier@aircanada.ca](mailto:jacques.cherrier@aircanada.ca)

## ABSTRACT

Airlines routinely overbook flights based on the expectation that some fraction of booked passengers will not show for each flight. Accurate forecasts of the expected number of no-shows for each flight can increase airline revenue by reducing the number of spoiled seats (empty seats that might otherwise have been sold) and the number of involuntary denied boardings at the departure gate. Conventional no-show forecasting methods typically average the no-show rates of historically similar flights, without the use of passenger-specific information.

We develop two classes of models to predict cabin-level no-show rates using specific information on the individual passengers booked on each flight. The first of these models computes the no-show probability for each passenger, using both the cabin-level historical forecast and the extracted passenger features as explanatory variables. This *passenger-level* model is implemented using three different predictive methods: a C4.5 decision-tree, a segmented Naïve Bayes algorithm, and a new aggregation method for an ensemble of probabilistic models. The second *cabin-level* model is formulated using the desired cabin-level no-show rate as the response variable. Inputs to this model include the predicted cabin-level no-show rates derived from the various passenger-level models, as well as simple statistics of the features of the cabin passenger population. The cabin-level model is implemented using either linear regression, or as a direct probability model with explicit incorporation of the cabin-level no-show rates derived from the passenger-level model outputs.

The new passenger-based models are compared to a conventional historical model, using train and evaluation data sets taken from over 1 million passenger name records. Standard metrics such as lift curves and mean-square cabin-level errors establish the improved accuracy of the passenger-based models over the historical model. All models are also evaluated using a simple revenue model, and it is shown that

the cabin-level passenger-based model can produce between 0.4% and 3.2% revenue gain over the conventional model, depending on the revenue-model parameters.

## Categories and Subject Descriptors

H.2.8 [Information Systems]: Database Applications—  
*Data mining*

## General Terms

Data mining

## Keywords

Airline overbooking, no-show forecasting, predictive modeling, classification, probabilistic estimation, model aggregation

## 1. INTRODUCTION

The practice of optimizing revenue by controlling the availability and pricing of airline seats is commonly referred to as revenue management[7]. Sophisticated revenue management systems are in use at all major airlines today, and are widely viewed as a critical component of an airline's overall logistics framework. Rather than offering identical seats at a common fare, revenue management systems introduce multiple booking classes differentiated by the offered fare as well as other possible restrictions such as cancellation options or overnight-stay requirements. The number of seats allocated to each booking class is determined by the estimated demand for each class. Sales of tickets in each class are controlled in an attempt to maximize revenue. For example, it is desirable to reserve seats in high-fare classes for last-minute travelers willing to pay higher fares, while limiting the number of seats sold in lower-fare classes earlier in the booking process. Revenue management establishes booking policies to determine whether to accept or reject a booking in a specific booking class, given the current number of bookings and expected additional demand prior to departure.

# End-to-end trip workflow

... perceive  
or infer...

- |                                    |                                       |
|------------------------------------|---------------------------------------|
| 1. Luggage + documentation         | 1. <u>Ouibring</u> + documentation    |
| 2. Passenger gets to airport       | 2. <u>Uber</u>                        |
| 3. Security checkpoints            | 3. <u>Clear</u> , TSA <u>PreCheck</u> |
| 4. Boarding                        | 4. Waiting room + priority check-in   |
| 5. Flight + service                | 5. Classes of service                 |
| 6. Leaving plane + connections     | 6. Class of service + connections     |
| 7. Luggage + customs + immigration | 7. Priority luggage + entry systems   |
| 8. Airport to destination          | 8. <u>Uber</u>                        |
| 9. Unpack...                       | 9. Unpack...                          |

# Workflows and “intelligence”

... perceive  
or infer...

- Intelligence will come from optimizing the end-to-end workflow
- You wanted to play music
  - Before: buy tape/vinyl/CD, put media on device, locate song, press play
  - Then the media disappeared: song went directly to your iPod
  - Then, full process was streamlined: “*Alexa: play [song]*”
- Conversational user interface
  - Trying on assistant: Google: [https://assistant.google.com/explore?hl=en\\_us](https://assistant.google.com/explore?hl=en_us)

# What if airlines could...

*... perceive  
or infer...*

- Predict flights by a class of passengers (holidays, Super Bowl)
- Offer full package: flight, accommodation, etc.
- Alert that a passport is about to expire ahead of trip
- Send transportation to customer houses
- ...
- Summary: anticipate and optimize complete job end-to-end

# Models and data

- Which models/data can be shared?
  - Models about passenger behavior?
  - Models about airline operations? (Crew, equipment, supply chain)
  - Models about services?
- Should models be “regulated”?
  - EU PNR (Passenger Name Record)
  - What about auditing?

# AI and Ethics

*... perceive  
or infer...*

- Why is this a topic at all?
- Security and privacy
- Will the intelligence change the world/model?
- How could intelligence maximize revenue for a hospital?
- Overbooking scenario: which passenger would have to wait?

# Takeaways: Models and AI

*... perceive  
or infer...*

- Requirements
  - Machine scalability
  - Organizational expertise
  - Data
  - Experiments
  - Libraries (software)
- “Think big and long term”

# Q&A

- ~~Math~~
- ~~Memory~~
- ~~Models~~

- Alisson Sol  
[email@AlissonSol.com](mailto:email@AlissonSol.com)
- What was the next number in the sequence?  
11, 21, 1211, 111221, ?