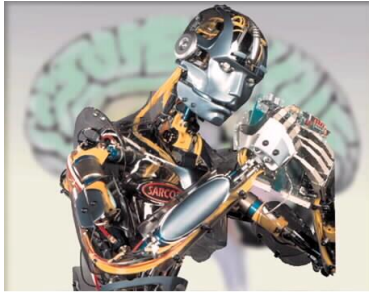
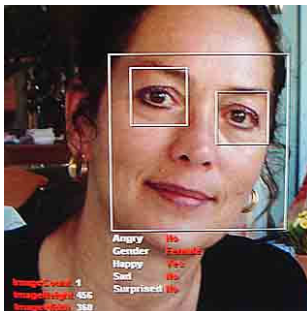


# What is Machine Learning?

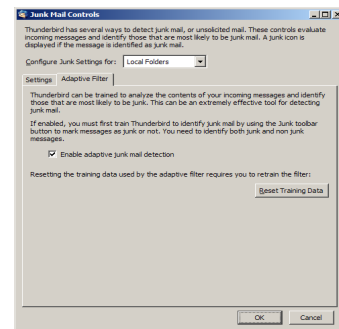


amazon.com.

facebook



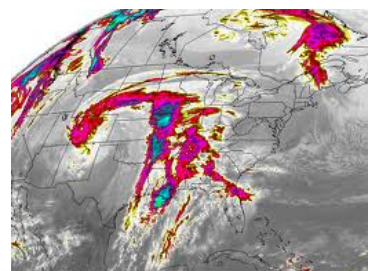
Microsoft®



## Learning from Data

*The world is driven by data.*

- Germany's climate research centre generates 10 petabytes per year
- Google processes 24 petabytes per day
- The Large Hadron Collider produces 60 gigabytes per minute (~12 DVDs)
- There are over 50m credit card transactions a day in the US alone.



# Learning from Data

**Data** is recorded from some real-world phenomenon.  
What might we want to do with that data?

## Prediction

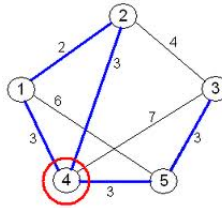
- what can we **predict** about this phenomenon?

## Description

- how can we **describe/understand** this phenomenon in a new way?



AHRQ Prevention Quality Indicators						
Calculation: Admin/Qual/Pop (P-Q-T)						
County Numbers in GREED are significantly lower than the National Average.						
County Numbers in HSA are significantly higher than the National Average.						
County Name	Cases	Population	Crude Rate	UCL	Rate	UCL
Alaska	79	11,718	1.16	4.82	0.99	5.35
Alaska	26	14,291	1.86	1.41	2.46	2.89
Alaska	12	19,493	0.76	0.26	0.94	1.42
Alaska	8	8,688	1.32	0.24	1.00	1.62
Alaska	102	37,112	1.28	2.86	2.03	3.35
Alaska	15	8,688	1.82	0.84	1.00	2.62
Alaska	122	23,005	0.29	4.52	4.93	5.43
Alaska	68	16,348	1.91	0.95	1.34	1.64
Alaska	26	10,241	0.91	0.95	1.36	1.64
Alaska	52	39,283	0.81	0.38	0.77	1.06
Alaska	52	12,887	1.43	0.86	1.34	1.79
Alaska	18	6,790	2.89	1.79	2.43	3.47
Alaska	23	12,887	1.23	0.84	1.34	1.79
Alaska	23	62,112	0.44	0.28	0.50	0.69
Alaska	23	10,241	1.23	0.84	1.34	1.79
Alaska	13	10,241	1.26	0.39	1.00	1.61
Alaska	28	29,388	0.38	0.50	0.90	1.30
Alaska	14	46,477	0.81	0.50	0.90	1.30
Alaska	5	7,261	1.89	0.99	0.99	1.36
Alaska	26	7,960	1.52	1.97	0.96	2.58
Alaska	18	21,968	0.85	0.97	0.90	1.34
Alaska	47	11,646	1.72	2.72	0.96	3.68



# Learning from Data

*How can we extract knowledge from data to help humans take decisions?*

*How can we automate decisions from data?*

*How can we adapt systems dynamically to enable better user experiences?*

Write code to explicitly  
do the above tasks

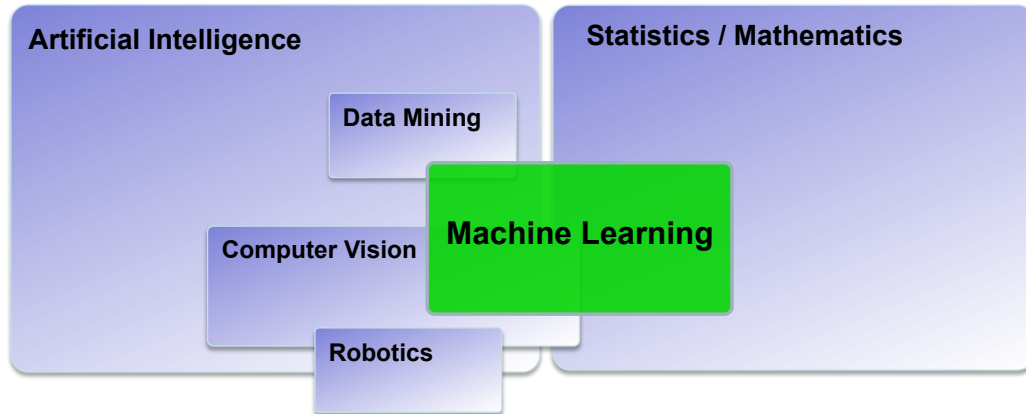


Write code to make the computer  
**learn** how to do the tasks

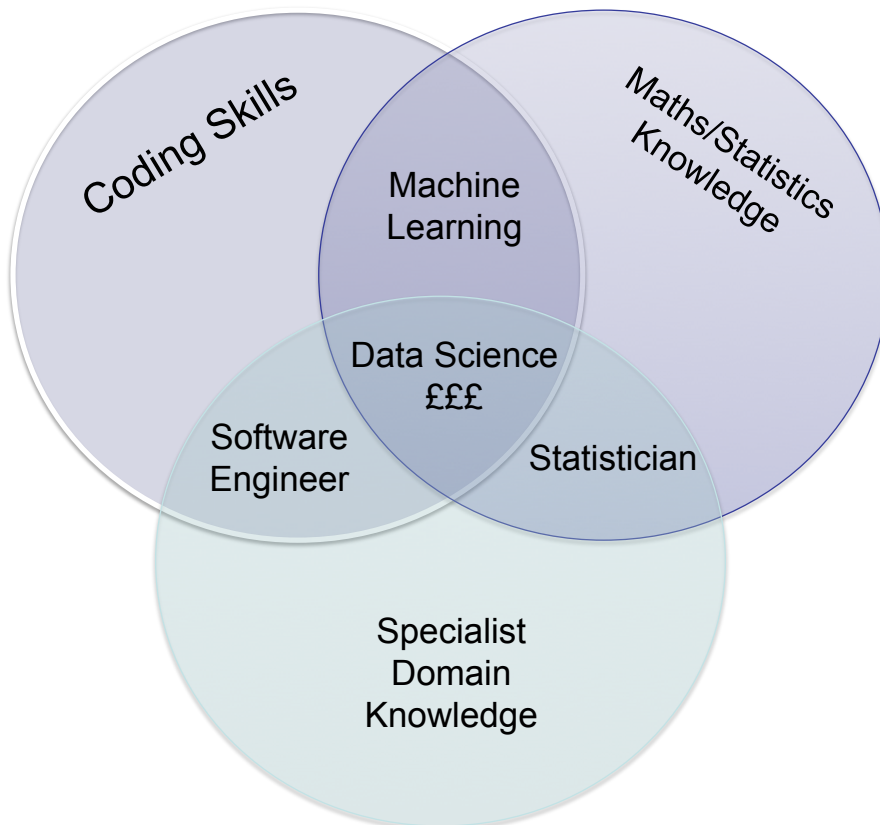


# Machine Learning

*Where does it fit? What is it **not**?*



(No definition of a field is perfect – the diagram above is just one interpretation, mine ;-)



Humans can:

- think, **learn**, see, understand language, reason, etc.



Artificial Intelligence aims to reproduce these capabilities.  
Machine Learning is **one** part of Artificial Intelligence.

**COMP14112** Fundamentals of Artificial Intelligence

**COMP24111** Introduction to Machine Learning  
**COMP24412** Symbolic AI



**COMP37212** Computer Vision  
**COMP34512** Knowledge Representation/Reasoning  
**COMP34411** Natural Language Systems  
**COMP34120** Artificial Intelligence and Games

## ***Introduction to Machine Learning***

<http://studentnet.cs.manchester.ac.uk/ugt/COMP24111>

50% lab / coursework

- Ex1 (due this week) .... 10%
- Ex2 (due end of Oct) ..... 20%
- Ex3 (due end of Nov) .....20%

50% January exam

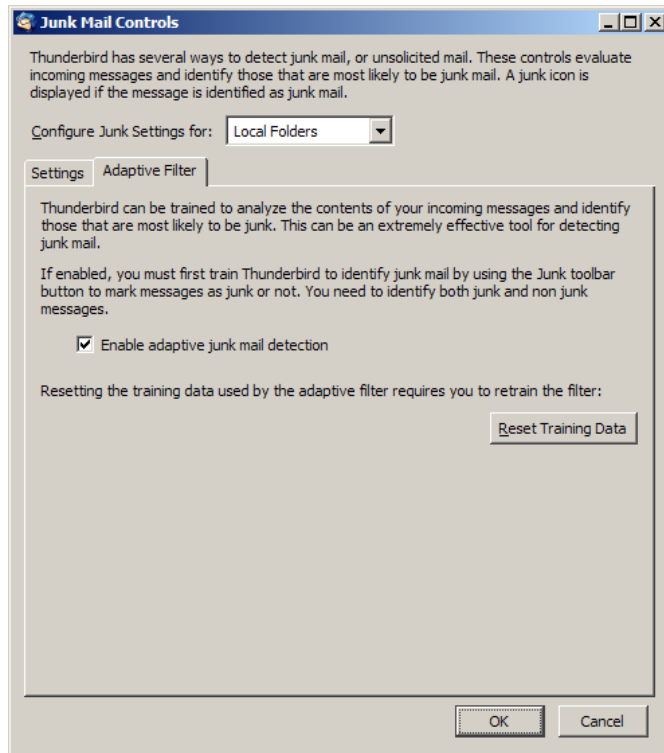
---

**Programming** : Matlab (no experience required)  
**Maths** : A little bit – would help you to revise A-level.

See notes/slides on course website.

---

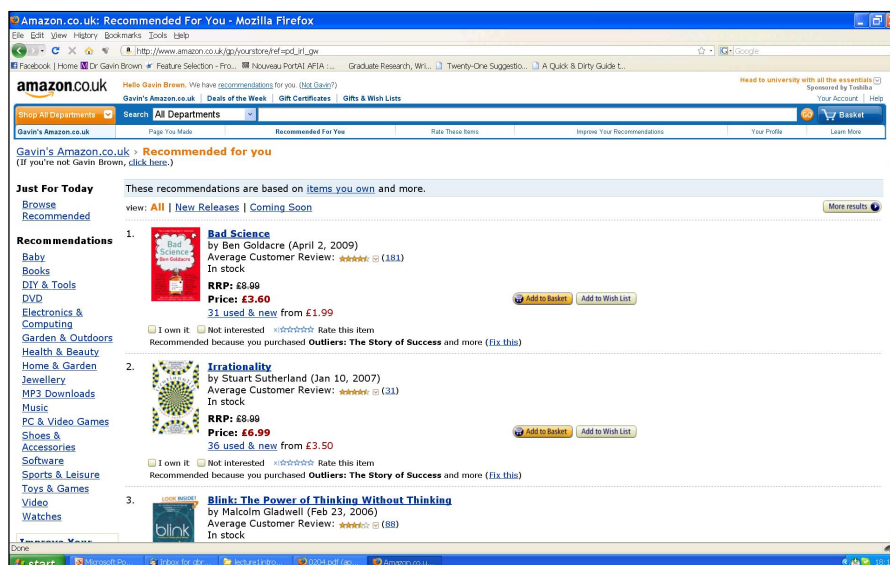
- Using machine learning to detect spam emails.



To: you@gmail.com  
 GET YOUR DIPLOMA TODAY!  
 If you are looking for a fast and cheap way to get a diploma, this is the best way out for you. Choose the desired field and degree and call us right now: For US: 1.845.709.8044 Outside US: +1.845.709.8044 "Just leave your NAME & PHONE NO. (with CountryCode)" in the voicemail. Our staff will get back to you in next few days!

**ALGORITHM**  
 Naïve Bayes  
 Rule mining

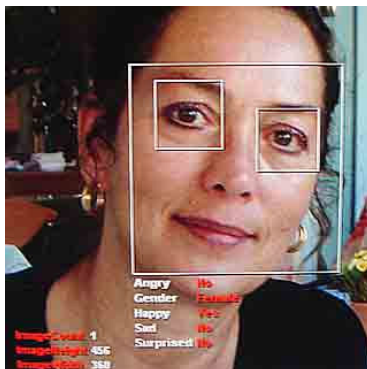
- Using machine learning to recommend books.



**ALGORITHMS**  
 Collaborative Filtering  
 Nearest Neighbour  
 Clustering

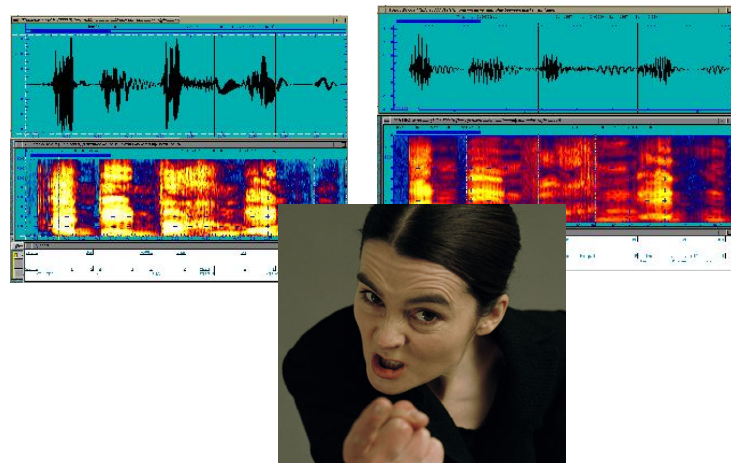
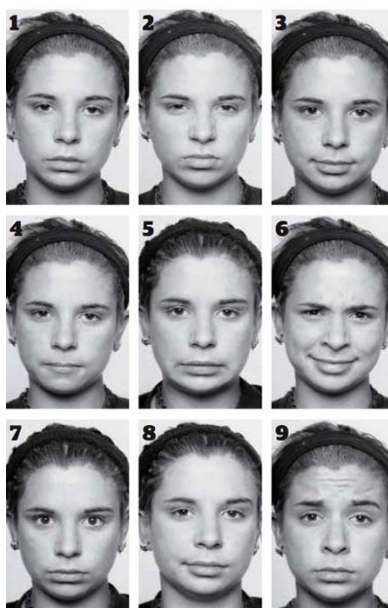


- Using machine learning to identify faces and expressions.



**ALGORITHMS**  
Decision Trees  
Adaboost

- Using machine learning to identify vocal patterns



**ALGORITHMS**  
Feature Extraction  
Probabilistic Classifiers  
Support Vector Machines  
+ many more....

- ML for working with social network data:  
detecting fraud, predicting click-thru patterns,  
targeted advertising, etc etc etc .

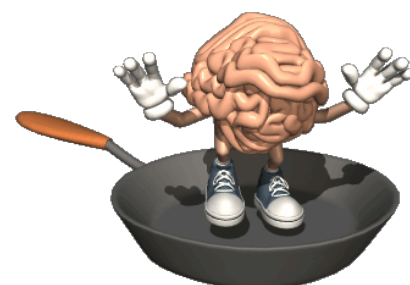


#### ALGORITHMS


Support Vector Machines  
Collaborative filtering  
Rule mining algorithms  
Many many more....

**Driving a car**  
**Recognising spam emails**  
**Recommending books**  
**Reading handwriting**  
**Recognising speech, faces, etc**

How would you write these programs?  
Would you want to?!?!?!?



Many applications are immensely hard to program directly.  
These almost always turn out to be “pattern recognition” tasks.

1. Program the computer to do the pattern recognition task directly. 

1. Program the computer to be able to learn from examples.

2. Provide “training” data.



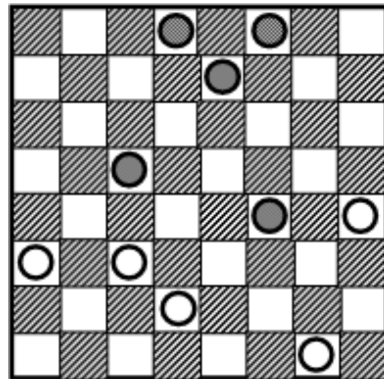
## Definition of Machine Learning

- self-configuring data structures that allow a computer to do things that would be called “intelligent” if a human did it
- “making computers behave like they do in the movies”



# A Bit of History

- Arthur Samuel (1959) wrote a program that **learnt** to play draughts (“checkers” if you’re American).



1940s

Human reasoning / logic first studied as a formal subject within mathematics (Claude Shannon, Kurt Godel et al).

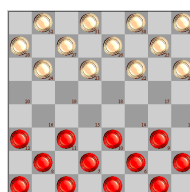
1950s

The “Turing Test” is proposed: a test for true machine intelligence, expected to be passed by year 2000. Various game-playing programs built. 1956 “Dartmouth conference” coins the phrase “artificial intelligence”.

1960s

A.I. funding increased (mainly military). Famous quote: “Within a generation ... the problem of creating 'artificial intelligence' will substantially be solved.”

Ax. 1.  $P(\varphi) \wedge \Box \forall x[\varphi(x) \rightarrow \psi(x)] \rightarrow P(\psi)$   
 Ax. 2.  $P(\neg\varphi) \leftrightarrow \neg P(\varphi)$   
 Th. 1.  $P(\varphi) \rightarrow \Diamond \exists x [\varphi(x)]$   
 Df. 1.  $G(x) \iff \forall \varphi[P(\varphi) \rightarrow \varphi(x)]$   
 Ax. 3.  $P(G)$   
 Th. 2.  $\Diamond \exists x G(x)$   
 Df. 2.  $\varphi \text{ ess } x \iff \varphi(x) \wedge \forall \psi\{\psi(x) \rightarrow \Box \forall x[\varphi(x) \rightarrow \psi(x)]\}$   
 Ax. 4.  $P(\varphi) \rightarrow \Box P(\varphi)$   
 Th. 3.  $G(x) \rightarrow G \text{ ess } x$   
 Df. 3.  $E(x) \iff \forall \varphi[\varphi \text{ ess } x \rightarrow \Box \exists x \varphi(x)]$   
 Ax. 5.  $P(E)$   
 Th. 4.  $\Box \exists x G(x)$



1970s

A.I. “winter”. Funding dries up as people realise it’s hard.  
Limited computing power and dead-end frameworks.

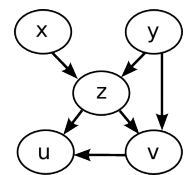
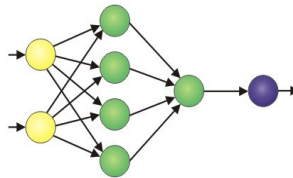
1980s

Revival through bio-inspired algorithms: Neural networks, Genetic Algorithms.  
A.I. promises the world – lots of commercial investment – mostly fails.  
Rule based “expert systems” used in medical / legal professions.

1990s

AI diverges into separate fields: Computer Vision, Automated Reasoning,  
Planning systems, Natural Language processing, **Machine Learning**...

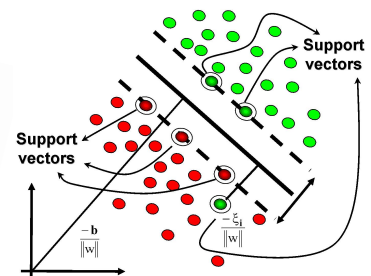
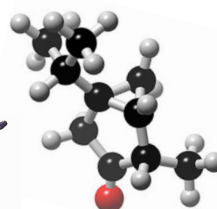
...Machine Learning begins to overlap with statistics / probability theory.



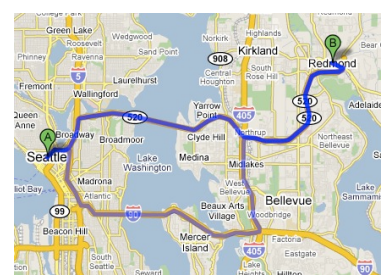
$$P(A|B) = \frac{P(B|A) P(A)}{P(B)}.$$

2000s

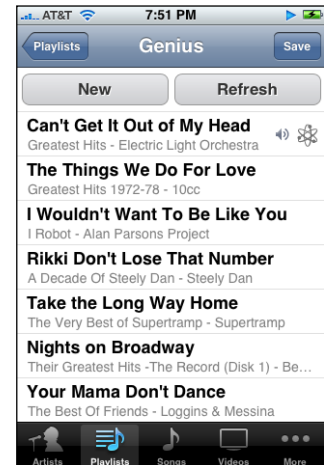
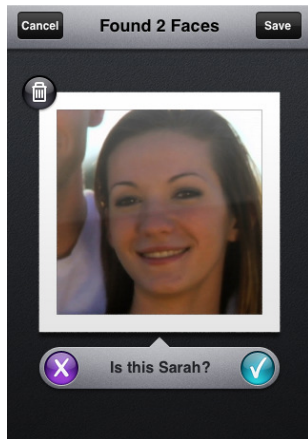
ML merging with statistics continues. Other subfields continue in parallel.  
First commercial-strength applications: Google, Amazon, computer games, route-finding, credit card fraud detection, etc...  
Tools adopted as standard by other fields e.g. biology



**2010s.... ??????**



# The future?



[http://www.youtube.com/watch?v=NS\\_L3Yyv2RI](http://www.youtube.com/watch?v=NS_L3Yyv2RI)

**Microsoft®**

Microsoft Research - Mozilla Firefox

History Bookmarks Tools Help

<http://research.microsoft.com/en-us/press/cmbishop.aspx>

Search Microsoft Research

Projects Publications People Downloads

WorldWide Telescope | Aphelion  
Download the latest release

Home Our Research Collaboration Careers  
About Us News Media Resources Events Community

Media Resources > Chris Bishop

## Chris Bishop

**Chief Research Scientist  
Microsoft Research Cambridge  
Microsoft Corp.**

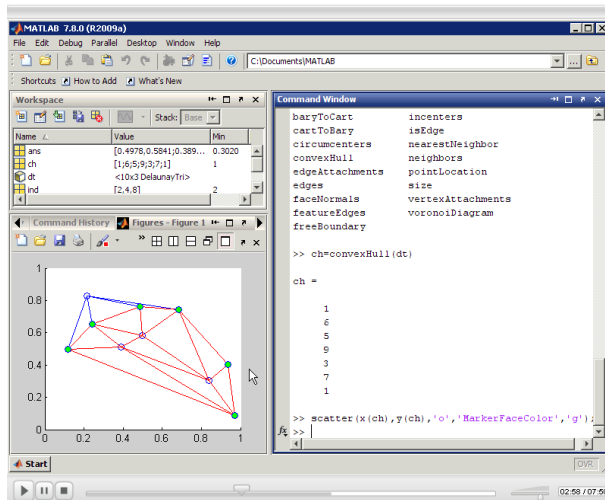
Chris Bishop is a Chief Research Scientist at Microsoft Research Cambridge, where he helps with the strategic direction and planning for the lab, and jointly leads the Machine Learning and Perception Group. His research interests include probabilistic approaches to machine learning, as well as their application to fields such as biomedical sciences and healthcare.

Chris is also Professor of Computer Science at the University of Edinburgh, where he is a member of the Institute for Adaptive and Neural Computation in the School of Informatics. Chris is a Fellow of Darwin College, Cambridge and a Fellow of the British Computer Society. In 2004 he was elected Fellow of the Royal Academy of Engineering, and in 2007 he was elected Fellow of the Royal Society of Edinburgh. In 2008 he was selected as the Royal Institution Christmas Lecturer.

Chris obtained a BA in Physics from Oxford, and a PhD in Theoretical Physics from the University of Edinburgh, with a thesis on quantum field theory. He then joined

**Microsoft has a MAJOR worldwide investment in Machine Learning**

# Programming language : “Matlab”



## MATrix LABoratory

- Interactive scripting language
- Interpreted (i.e. no compiling)
- Objects possible, not compulsory
- Dynamically typed
- Flexible GUI / plotting framework
- Large libraries of tools
- Highly optimized for maths

## ***Introduction to Machine Learning***

<http://studentnet.cs.manchester.ac.uk/ugt/COMP24111>

Now – short break – [prompt!](#)

**After the break:**

**Your first machine learning algorithm.**