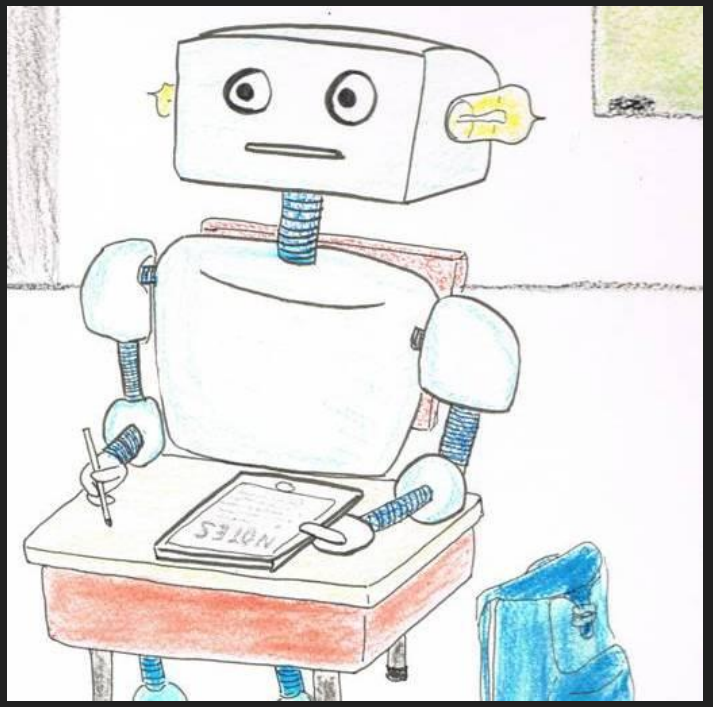
Introduction to Machine Learning

Oisin Mac Aodha, Michael Firman

What is Machine Learning?

Algorithms that can learn from data



http://gureckislab.org/blog/?p=2648

What is Machine Learning?

Algorithms that can learn from data so they can make sensible decisions when faced with novel scenarios.

https://swordrock.files.wordpress.com/2010/08/robot_soccer_robotgames-720918.jpg

What is Machine Learning?

Teaching computers how to perform a task
without having to explicitly program them to do it.

Using data to parameterize models.

Why Learning?

Many tasks are too complex to describe by a set of rules, so it’s often easier to teach by example.

Relationship to other fields

Statistics
Pattern Recognition
Information Retrieval
Data Mining
Computer Vision
Data Science
Artificial Intelligence
...

Origins

Beginnings in more general field of Artificial Intelligence in 1950/60s.

Started to flourish in the 1990s.

Bigger impact on industry in 2010s.

http://www.thechessdrum.net/newsbriefs/2002/NB_photos/Kasparov-DeepBlue.jpg
http://static.guim.co.uk/sys-images/Guardian/Pix/pictures/2013/1/11/1357910581869/IBM-supercomputer-Watson-011.jpg
https://commons.wikimedia.org/wiki/File:Google_self_driving_car_at_the_Googleplex.jpg

Rise of Machine Learning

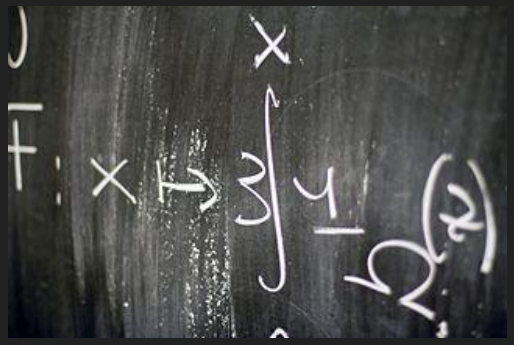
Computation

+

Data

+

Algorithms

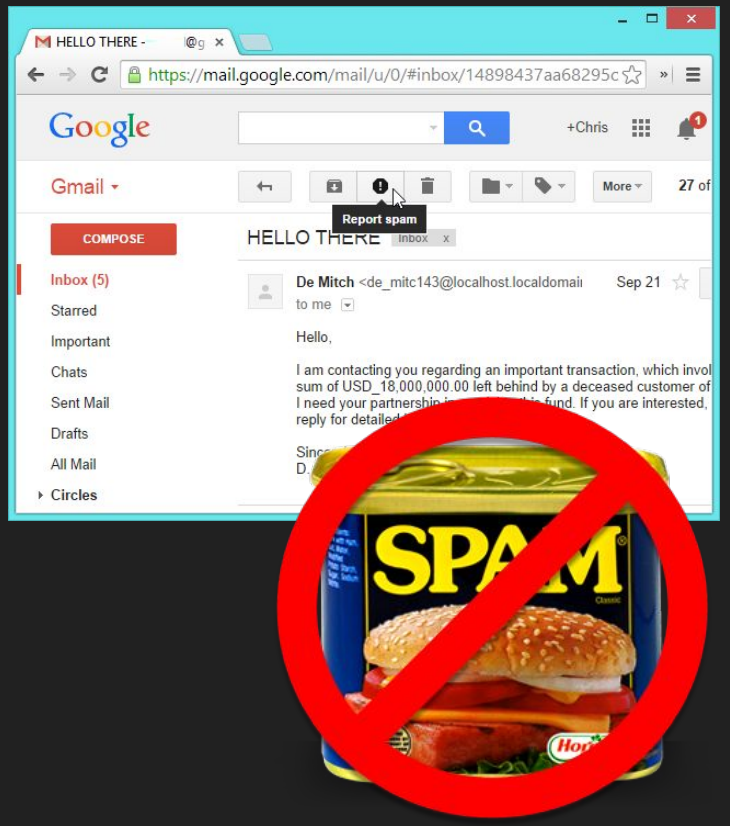


http://www.extremetech.com/wp-content/uploads/2013/12/Google-Datacenter.jpg
https://commons.wikimedia.org/wiki/File:0009-cmls-2012_preview_ecran.jpg

Spam Filtering

In Q1 2015 60% of all email traffic was spam [1].

Gmail claims to filter 99.9% of this [2].

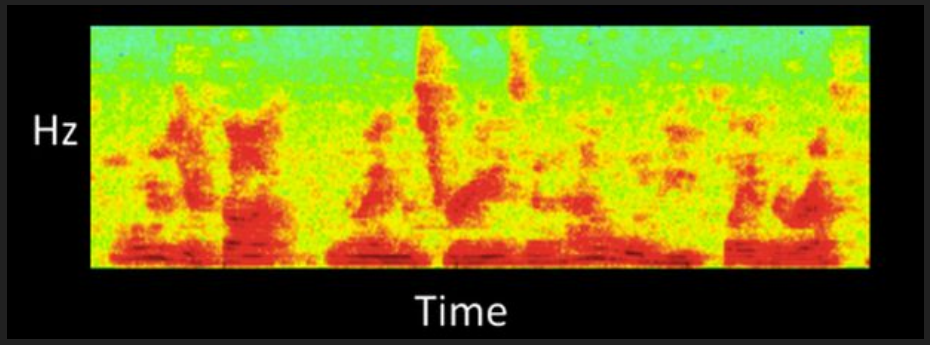


[1] https://securelist.com/analysis/quarterly-spam-reports/69932/spam-and-phishing-in-the-first-quarter-of-2015/
[2] http://gmailblog.blogspot.co.uk/2015/07/the-mail-you-want-not-spam-you-dont.html

http://blogs.longwood.edu/files/2011/09/antispam.png
http://cdn5.howtogeek.com/wp-content/uploads/2014/10/report-spam-email-in-gmail.png

Speech Recognition e.g. Siri

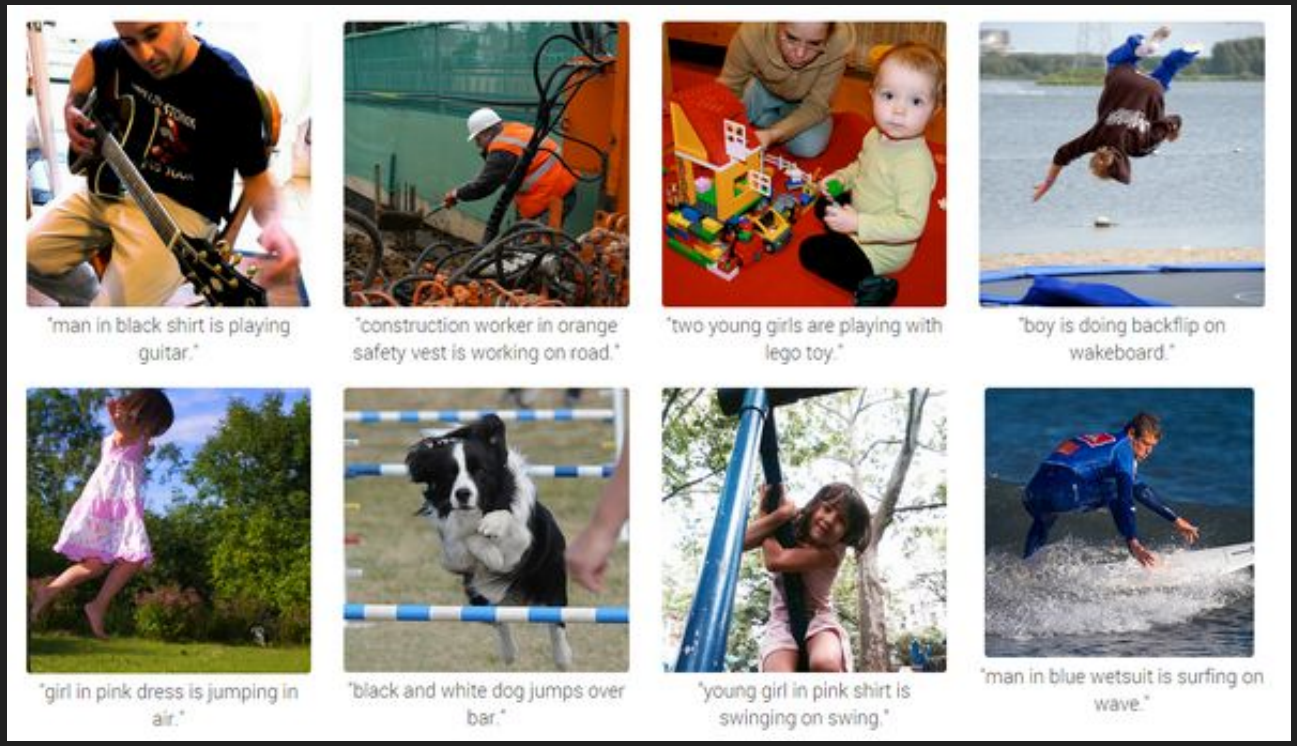
Error rates for Mandarin speech recognition are down to 6% [1].



[1] http://usa.baidu.com/multi-lingual-speech-recognition/.uk/2015/07/the-mail-you-want-not-spam-you-dont.html

http://usa.baidu.com/multi-lingual-speech-recognition/

Automatic Image Captioning



[1] http://cs.stanford.edu/people/karpathy/deepimagesent/

Prediction

Want to make inference about new data.

Not necessarily trying to explain our data or understand the process that generated it.

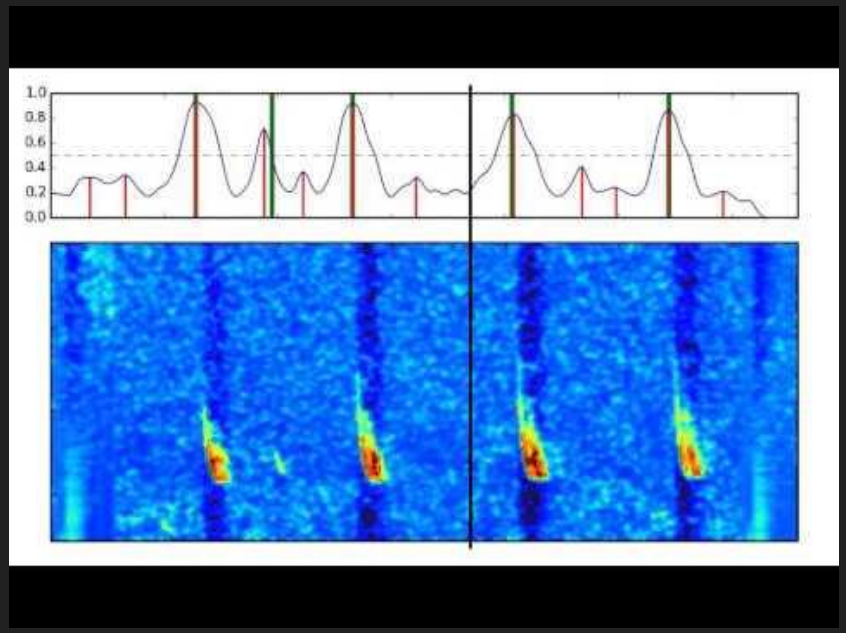
When is Machine Learning Useful?

When you can easily collect lots of data but the manual processing of the data is slow and laborious.

http://www.wwfindia.org/?7322/Camera-traps-capture-exciting-animal-behaviour-across-India

Automated Species Monitoring from Audio

Automatically detecting bats in audio recordings.



Bat
DETECTIVE

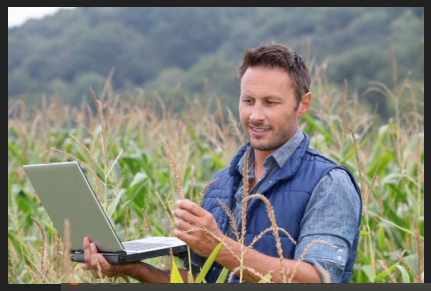
Butterfly Recognition from Images

Classify British butterflies into different species using museum data as training.

http://www.nhm.ac.uk/natureplus/blogs/behind-the-scenes/tags/butterfly

Machine Learning - Data Collection

Data Collection



Liberation Sans 10

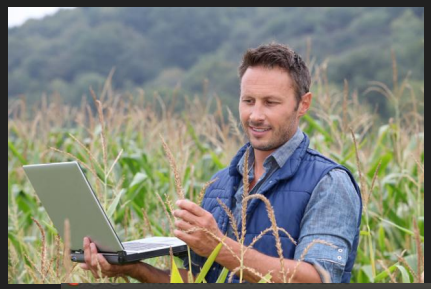
C7

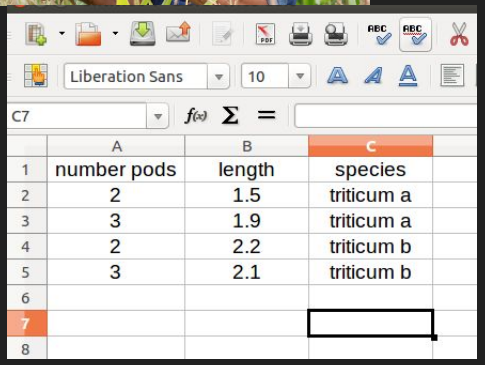
	A	B	C
1	number pods	length	species
2	2	1.5	triticum a
3	3	1.9	triticum a
4	2	2.2	triticum b
5	3	2.1	triticum b
6			
7			
8			

http://hdwyn.com/wallpaper_1024x768/ecologist_field_notebook_1024x768_hd-wallpaper-77094.jpg

Machine Learning - Model Training

Data Collection





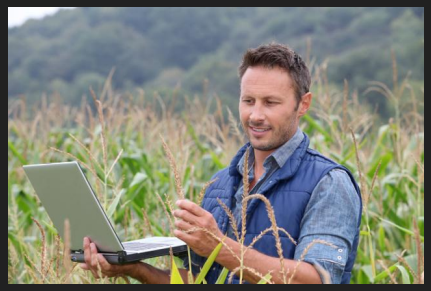
	A	B	C
1	number pods	length	species
2	2	1.5	triticum a
3	3	1.9	triticum a
4	2	2.2	triticum b
5	3	2.1	triticum b
6			
7			
8			

Train Model

http://hdwyn.com/wallpaper_1024x768/ecologist_field_notebook_1024x768_hd-wallpaper-77094.jpg
https://commons.wikimedia.org/wiki/File:C64c_system.jpg

Machine Learning - Prediction

Data Collection



Train Model

Make Predictions

Liberation Sans 10

C15

	A	B	C
1	number pods	length	species
2	2	1.5	?
3	3	1.9	?
4	2	2.2	?
5	3	2.1	?
6	2	1.1	?
7	3	1.3	?
8	3	2.3	?
9	4	2.1	?
10	2	2.5	?
11	3	1.9	?

http://hdwyn.com/wallpaper_1024x768/ecologist_field_notebook_1024x768_hd-wallpaper-77094.jpg
https://commons.wikimedia.org/wiki/File:C64c_system.jpg

Goals

Introduction to machine learning
- intuition not technical depth

Some practical examples of typical problems
- classification and regression

What we won’t cover

Algorithmic details

Probabilistic interpretation

Deep learning

Code and Slides

Download from here:

www.cs.ucl.ac.uk/staff/O.MacAodha/ml_intro