

Machine Learning (AIMS) - MT 2017

0. (My) Introduction to ML

Varun Kanade

University of Oxford
November 6, 2017

Machine Learning in Action



(Using <https://www.betafaceapi.com/demo.html>)

Machine Learning in Action



(Using <https://www.betafaceapi.com/demo.html>)

Machine Learning in Action



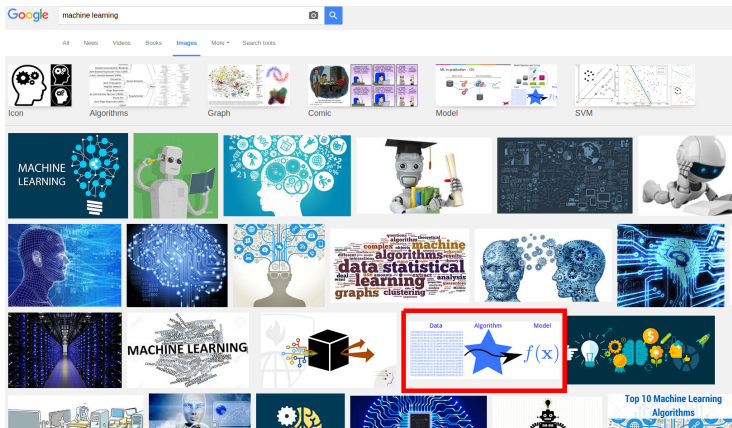
age: 19, beard: no, expression: other,
gender: female, glasses: no, mustache: no,
race: white,



age: 23, beard: no, expression: other,
gender: female, glasses: no, mustache: no,
race: white,

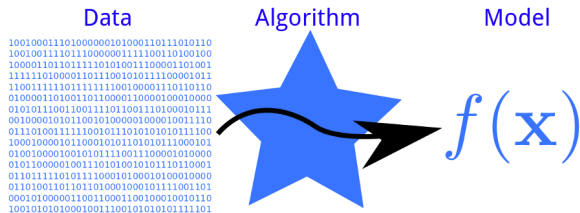
(Using <https://www.betafaceapi.com/demo.html>)

What is machine learning?



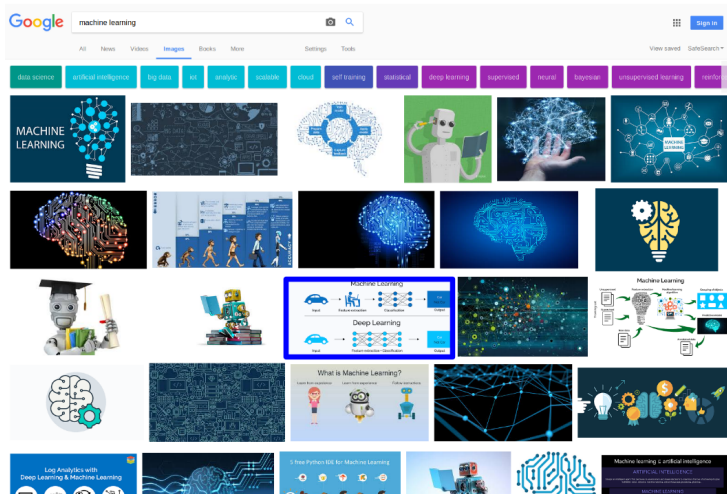
circa October 2016

What is machine learning?



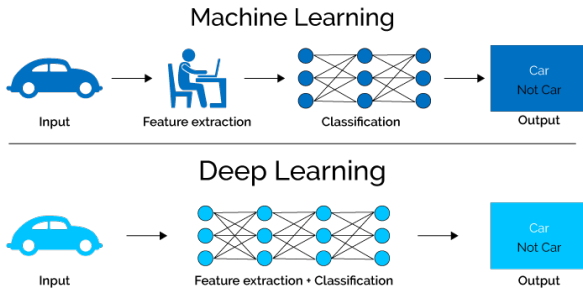
circa October 2016

What is machine learning?



circa October 2017

What is machine learning?



circa October 2017

What is machine learning?

What is artificial intelligence?

"Instead of trying to produce a programme to simulate the adult mind, why not rather try to produce one which simulates the child's? If this were then subjected to an appropriate course of education one would obtain the adult brain."



Turing, A.M. (1950). Computing machinery and intelligence. *Mind*, 59, 433-460.

What is machine learning?

Definition by Tom Mitchell

A computer program 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 .

Face Detection

- ▶ E : images (with bounding boxes) around faces
- ▶ T : given an image without boxes, put boxes around faces
- ▶ P : number of faces correctly identified

Course Information

Website

www.cs.ox.ac.uk/people/varun.kanade/teaching/ML-AIMS-MT2017/

Lectures

Mon-Fri: 10h30 - 12h30 (here)

Practicals

Mon-Fri: 14h-17h

Demonstrators: Philip Lazos, David Martínez

Textbooks

Kevin Murphy - Machine Learning: A Probabilistic Perspective

- ▶ Online access through Bodleian library

A few last remarks about this course



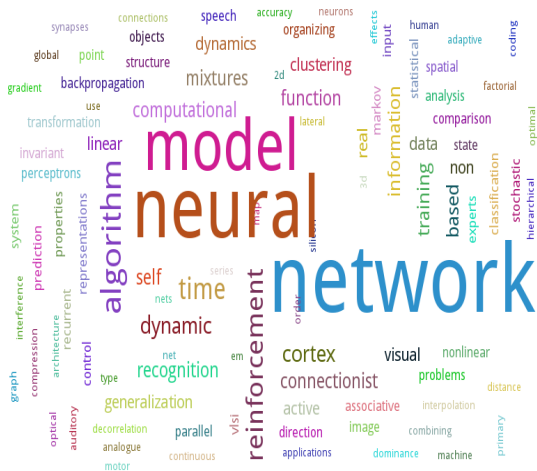
As ML developed through various disciplines - CS, Stats, Neuroscience, Engineering, *etc.*, there is no consistent usage of **notation** or even **names** among the textbooks. At times you may find inconsistencies even within a single textbook.

You will be required to read, both before and after the lectures. I will post suggested reading on the website.

Resources:

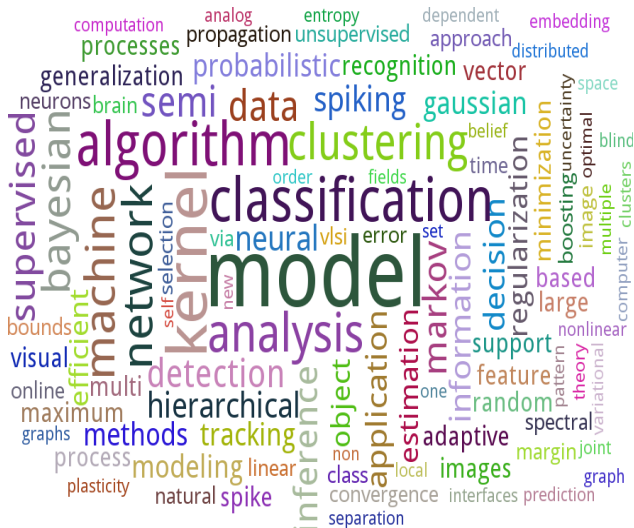
- ▶ Wikipedia has many great articles about ML and background material
- ▶ Online videos: Andrew Ng on coursera, Nando de Freitas on youtube, *etc.*
- ▶ Many interesting blogs, podcasts, *etc.*

NIPS Papers!



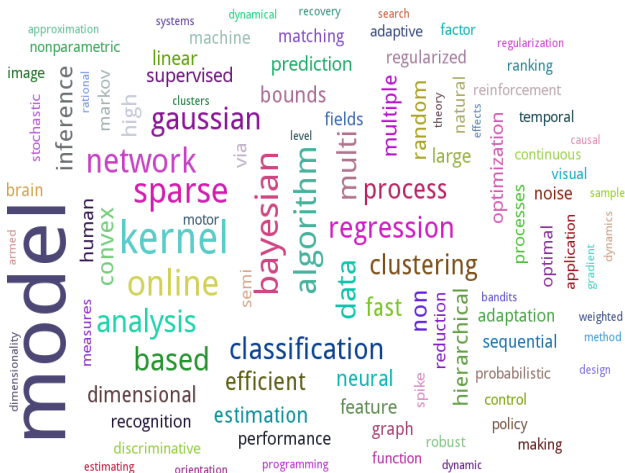
Advances in Neural Information Processing Systems 1995

NIPS Papers!



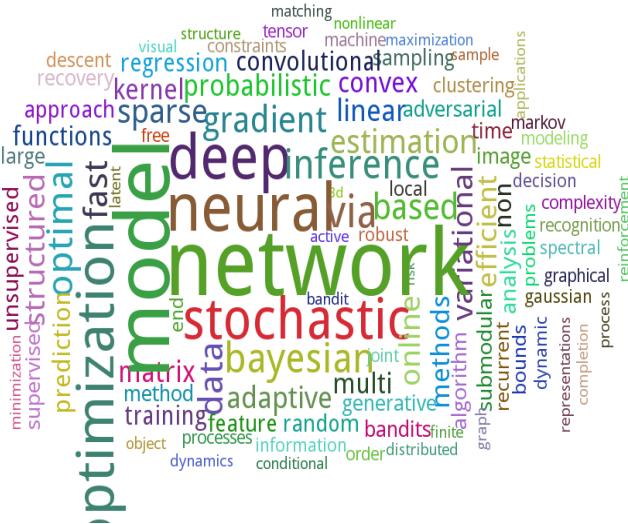
Advances in Neural Information Processing Systems 2005

NIPS Papers!



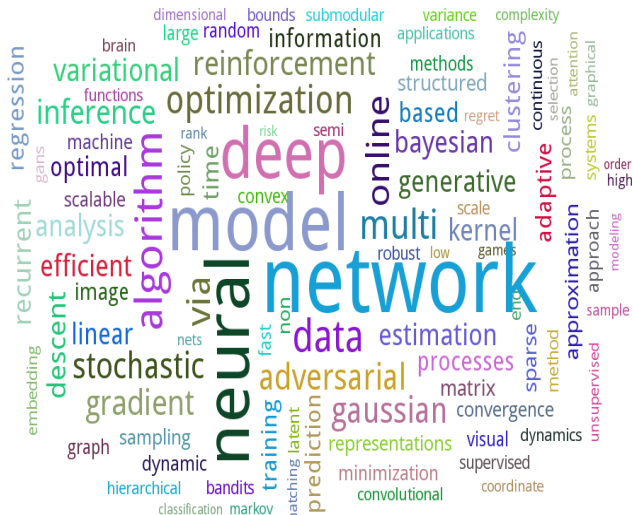
Advances in Neural Information Processing Systems 2009

NIPS Papers!



Advances in Neural Information Processing Systems 2016

NIPS Papers!

Advances in Neural Information Processing Systems 2017 [\[video\]](#)

Application: Boston Housing Dataset

Numerical attributes

- ▶ Crime rate per capita
- ▶ Non-retail business fraction
- ▶ Nitric Oxide concentration
- ▶ Age of house
- ▶ Floor area
- ▶ Distance to city centre
- ▶ Number of rooms

Categorical attributes

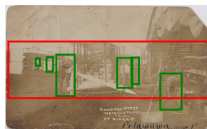
- ▶ On the Charles river?
- ▶ Index of highway access (1-5)

Predict house cost



Source: [UCI repository](#)

Application: Object Detection and Localisation



- ▶ 200-basic level categories
- ▶ Here: Six pictures containing airplanes and people
- ▶ Dataset contains over 400,000 images
- ▶ Imagenet competition (2010-)
- ▶ All recent successes through very deep neural networks!

Supervised Learning

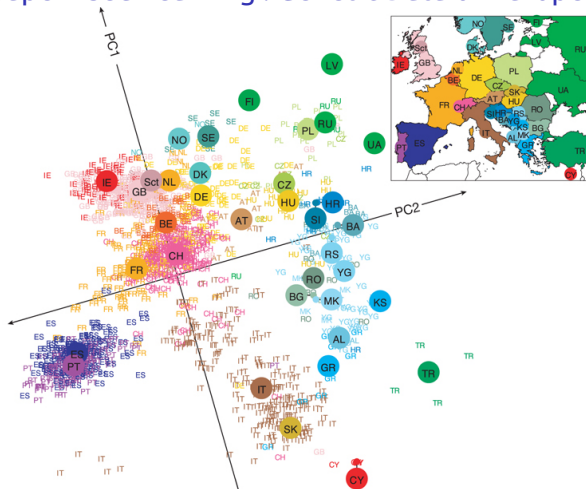
Training data has inputs x (numerical, categorical) as well as outputs y (target)

Regression: When the output is real-valued, *e.g.*, housing price

Classification: Output is a category

- ▶ Binary classification: only two classes *e.g.*, spam
- ▶ Multi-class classification: several classes *e.g.*, object detection

Unsupervised Learning : Genetic Data of European Populations



Experience (E)

Task (T)

Performance (P)

Source: [Novembre et al., Nature \(2008\)](#)

Dimensionality reduction - Map high-dimensional data to low dimensions

Clustering - group together individuals with similar genomes

Unsupervised Learning : Group Similar News Articles

The screenshot shows the Google News homepage. The browser title is "Google News - Mozilla Firefox (Private Browsing)". The address bar shows "https://news.google.com". The page layout includes a left sidebar with "Top Stories" and a list of categories (Donald Trump, Google, Florida, Nobel Prize, Brexit, Formula One, Samsung Electronics Limited, Wayne Rooney, Oculus Rift, PlayStation VR, Oxford, England, World, U.K., Business, Technology, Entertainment, Sports, Science, Health, Spotlight). The main content area features several news articles with headlines and images. The right sidebar includes a "Sign in to get news on topics you care about" prompt, a "Weather for Oxford, England" section with a table showing temperatures for Today, Sun, Mon, and Tue, and an "Editors' Picks" section featuring the "Mirror" newspaper.

Top Stories

- Donald Trump
- Google
- Florida
- Nobel Prize
- Brexit
- Formula One
- Samsung Electronics Limited
- Wayne Rooney
- Oculus Rift
- PlayStation VR
- Oxford, England
- World
- U.K.
- Business
- Technology
- Entertainment
- Sports
- Science
- Health
- Spotlight

US election: Donald Trump says he will not quit over video

US presidential candidate Donald Trump has said he will not withdraw from the race in phone interviews with US media. Mr Trump has been under pressure after a tape of him making lewd sexual comments and bragging about groping and kissing women ...

VIEWERS' GUIDE: Time for candidates to keep calm, debate on Daily Mail

Donald Trump defiant as groping tape drives growing Republican calls to quit The Guardian

Live Updating: Second presidential debate 2016: What time, how to watch and live stream online, CBS News

German city on lock down as police investigate bomb plot threat

Police are questioning three people after explosives were found hidden in an apartment officers raided in connection with a Syrian man suspected of planning a bomb attack on a German airport.

Trump vows to stay in race after calls for him to quit over lewd remarks

NEW YORK/WASHINGTON, Oct 8 (Reuters) - U.S. Presidential nominee Donald Trump vowed on Saturday to remain in the race even as his campaign was thrown into crisis as his own running mate criticized him and some prominent Republicans withdrew ...

A weakening Matthew rakes Atlantic coast; US death toll at 4

CHARLESTON, S.C. (AP) - A fast-weakening Hurricane Matthew continued its march along the Atlantic coast Saturday, lashing two of the South's most historic cities and some of its most popular resort islands, battering trees, swamping streets and ...

Derby County part company with Nigel Pearson by mutual agreement

Derby have announced they have parted company with their manager Nigel Pearson with immediate effect. The former Leicester boss had been suspended last month and the club began investigating a number of disciplinary issues.

Weather for Oxford, England

Today	Sun	Mon	Tue
14° 6°	14° 4°	13° 6°	15° 7°

The Weather Channel - Weather Underground - AccuWeather

Oxford, England »

VIDEO: Controlled explosion came out after mortar found in Oxford river

Oxford Mail - 53 minutes ago

Half marathon effort follows friends' 'miraculous' recovery

The Oxford Times - 6 hours ago

MICHAEL APPLETON COLUMN: A change in routine but we will be ready for tomorrow

The Oxford Times - 4 hours ago

Editors' Picks

Mirror

What to watch on Netflix in October including the best new movies, shows... mirror

Google Pixel and Pixel XL: Specs, features, price and release date of... mirror

Disney World closed for first time in 11 years as hurricane puts Florida...

Group similar articles into categories such as politics, music, sport, etc.

In the dataset, there are no labels for the articles

Active and Semi-Supervised Learning

Active Learning

- ▶ Initially all data is unlabelled
- ▶ Learning algorithm can ask a human to label some data



Semi-supervised Learning

- ▶ Limited labelled data, lots of unlabelled data
- ▶ How to use the two together to improve learning?



Collaborative Filtering : Recommender Systems

Movie / User	Alice	Bob	Charlie	Dean	Eve
The Shawshank Redemption	7	9	9	5	2
The Godfather	3	?	10	4	3
The Dark Knight	5	9	?	6	?
Pulp Fiction	?	5	?	?	10
Schindler's List	?	6	?	9	?

Netflix competition to predict user-ratings (2008-09)

Any individual user will not have used most products

Most products will have been use by some individual



Reinforcement Learning

- ▶ Automatic flying helicopter; self-driving cars
- ▶ Cannot conceivably program by hand
- ▶ Uncertain (stochastic) environment
- ▶ Must take **sequential decisions**
- ▶ Can define **reward functions**
- ▶ Fun: Playing Atari breakout! [\[video\]](#)

