Machine Learning

Nick Malleson, Dan Birks, Natacha Chenevoy

School of Geography & LIDA, University of Leeds

UNIVERSITY OF LEEDS

N8 POLICING RESEARCH PARTNERSHIP

Aims of the day

By the end of today, participants should:

- Understand what is meant by '*machine learning*';
- Be familiar with some common machine learning methods, including *LDA* and *neural networks*;
- Have an appreciation for how *natural language processing* and *image recognition* could be used to derive new insight from police data;
- Be able to run some simple machine learning algorithms in Python.

Aims of the day

But you won't:

- Be an expert in python!
- Know how to apply these techniques in earnest

The aim of today is just to provide an *introduction*

UNIVERSITY OF LEEDS

N8 POLICING RESEARCH PARTNERSHIP

Computer Programming

"The spread of computers and the Internet will put jobs in two categories: people who tell computers what to do, and people who are told by computers what to do."

(Marc Andreessen)

UNIVERSITY OF LEEDS

N8 POLICING RESEARCH PARTNERSHIP

Outline for the day

Time	Activity
11:00	Registration and introduction to the workshop
11:30	**An Introduction to Machine Learning and Text Analysis** Presentation 1a: What is machine learning? Presentation 1b: Natural Language Processing
12:30	Lunch
13:30	Hands-on Session 1: Exploring Different Types of Modus Operandi using Machine Learning
14:45	Break
15:00	Hands-on Session 2: Hate Speech Classification
16:00	Optional Presentation 3: Image Recognition and ANPR
16:30	Close

UNIVERSITY OF LEEDS

N8 POLICING RESEARCH PARTNERSHIP

Resources Used

I can take almost no credit for the content here!

Code examples:

- Natacha Chenevoy (https://github.com/mednche)
- Current Leeds PhD student
- Winner of the 2018 International Association of Law Enforcement Intelligence Analysts (IALEIA) *Award for Excellence*
- https://github.com/nickmalleson/n8-prp-ml-practicals

Machine learning for image processing ('Machine Learning is Fun!)

- https://medium.com/@ageitgey/machine-learning-is-fun-80ea3ec3c471

NLTK and Neural Networks

- Andy Evans' fantastic programming resources
- https://www.geog.leeds.ac.uk/courses/computing/materials/extra/index.html
- https://www.geog.leeds.ac.uk/courses/computing/

UNIVERSITY OF LEEDS

N8 POLICING RESEARCH PARTNERSHIP

Questions?

UNIVERSITY OF LEEDS

N8 POLICING RESEARCH PARTNERSHIP

Session 1
An Introduction to Machine Learning and Text Analysis

Presentation 1a: What is machine learning?

UNIVERSITY OF LEEDS

N8 POLICING RESEARCH PARTNERSHIP

What is machine learning?

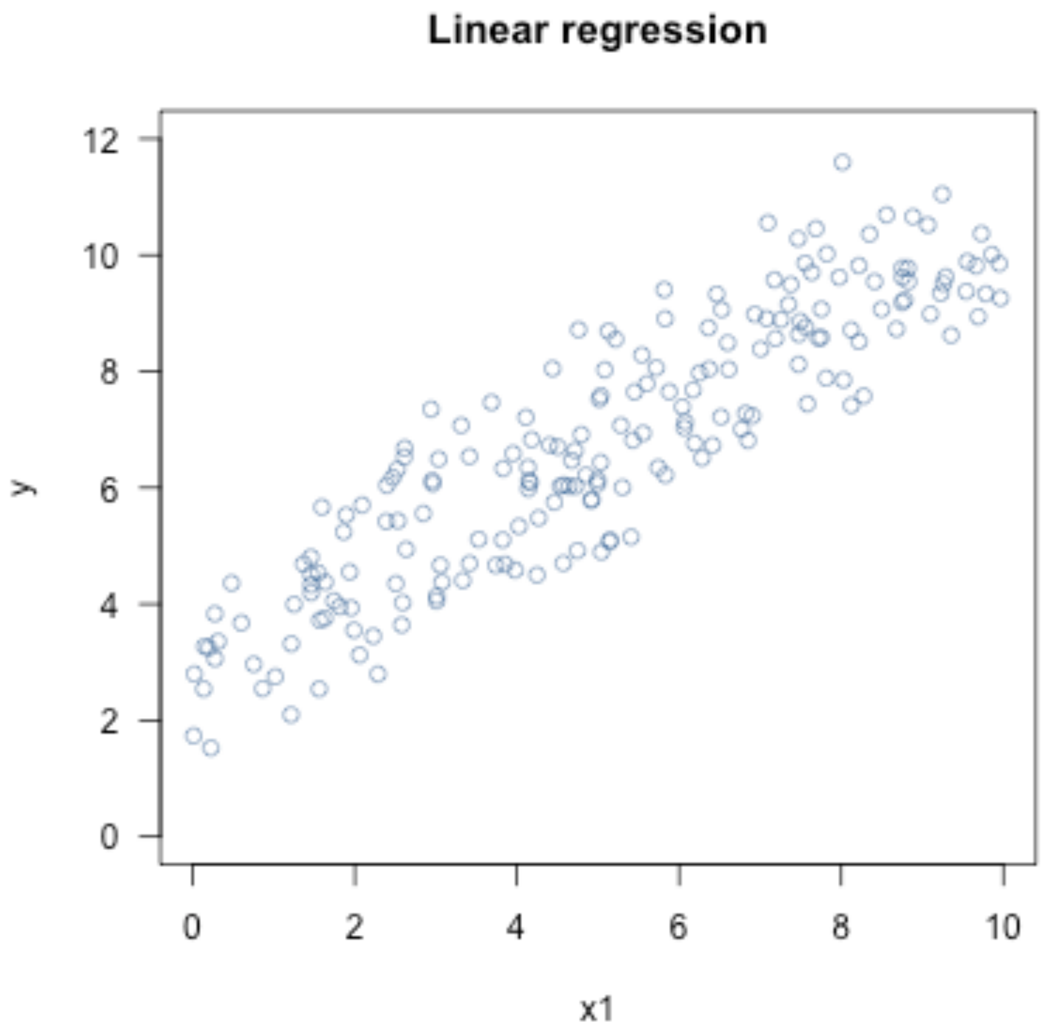
“Machine learning is the idea that there are generic algorithms that can tell you something interesting about a set of data without you having to write any custom code specific to the problem. Instead of writing code, you feed data to the generic algorithm and it builds its own logic based on the data.”

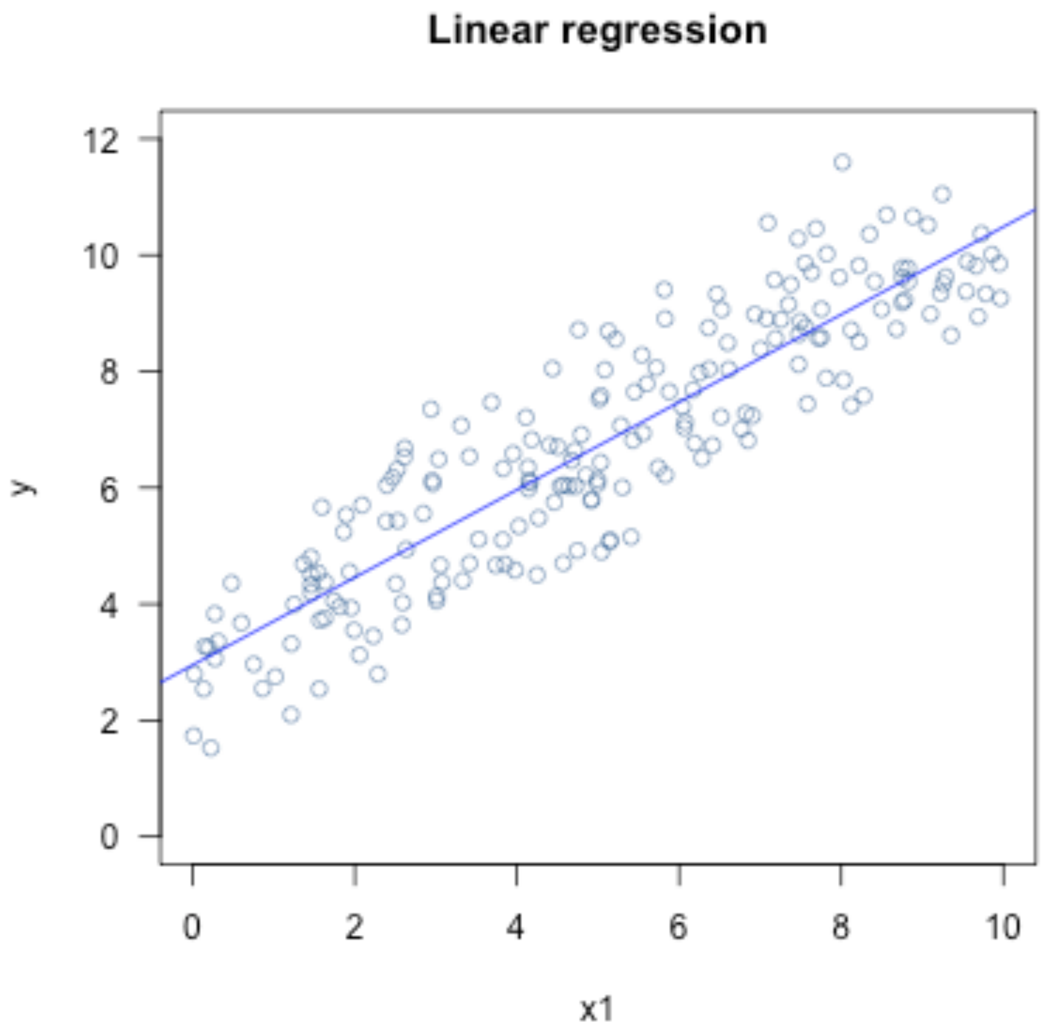
Question: name a machine learning algorithm that you have used

(One) Answer: regression!

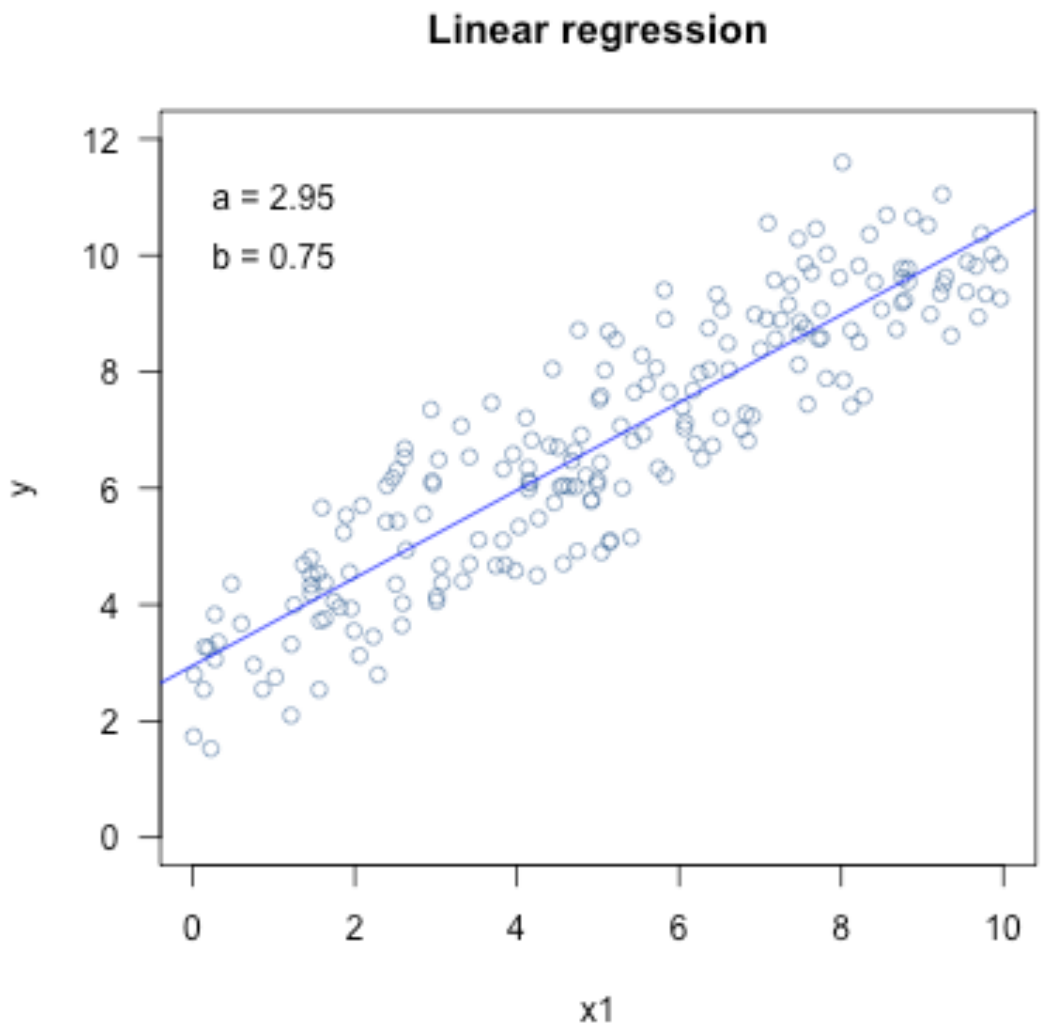
UNIVERSITY OF LEEDS

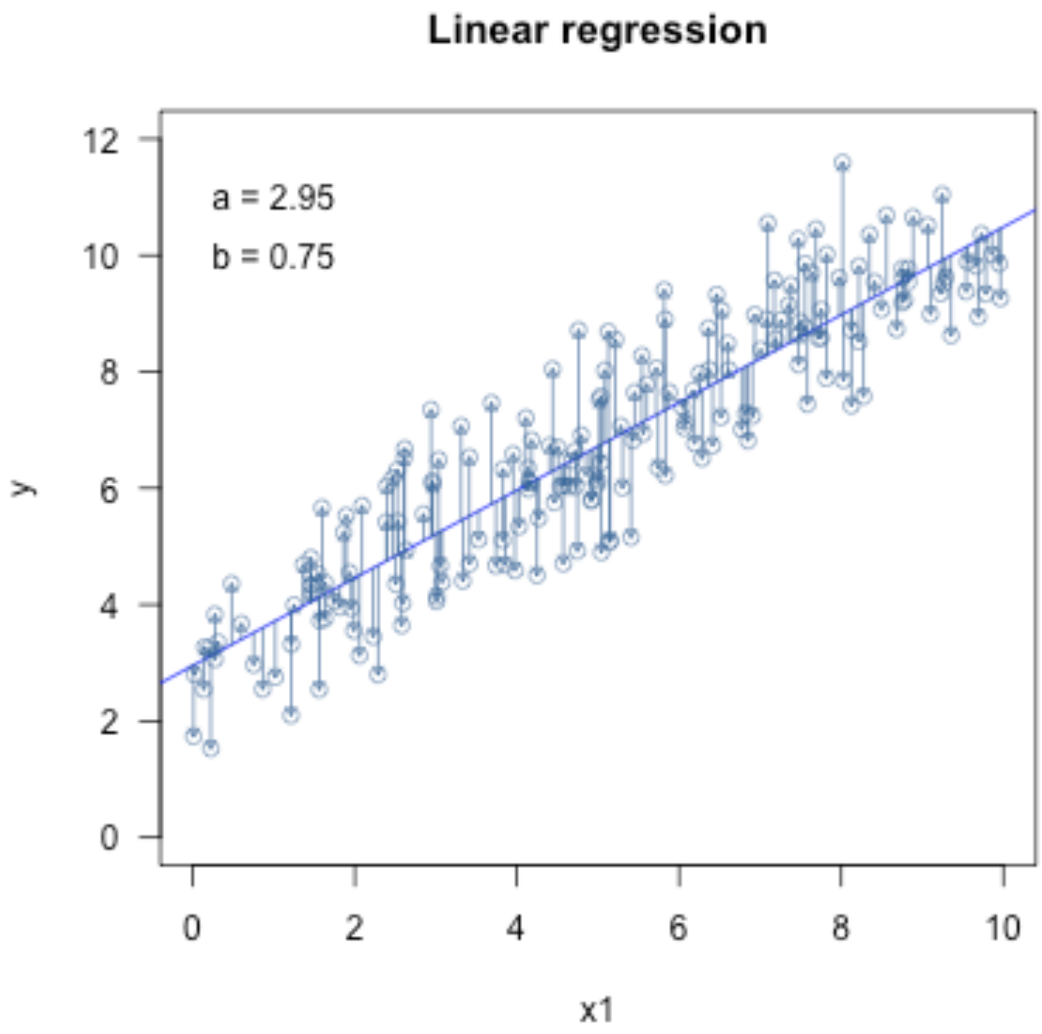
POLICING RESEARCH PARTNERSHIP


Linear regression
y
x1
0
2
4
6
8
10
12
0
2
4
6
8
10



Linear regression
y
x1
0
2
4
6
8
10
12
0
2
4
6
8
10


Linear regression
a = 2.95
b = 0.75
y
x1
0
2
4
6
8
10
12


Linear regression
a = 2.95
b = 0.75
y
x1
0
2
4
6
8
10
12

Linear regression / OLS

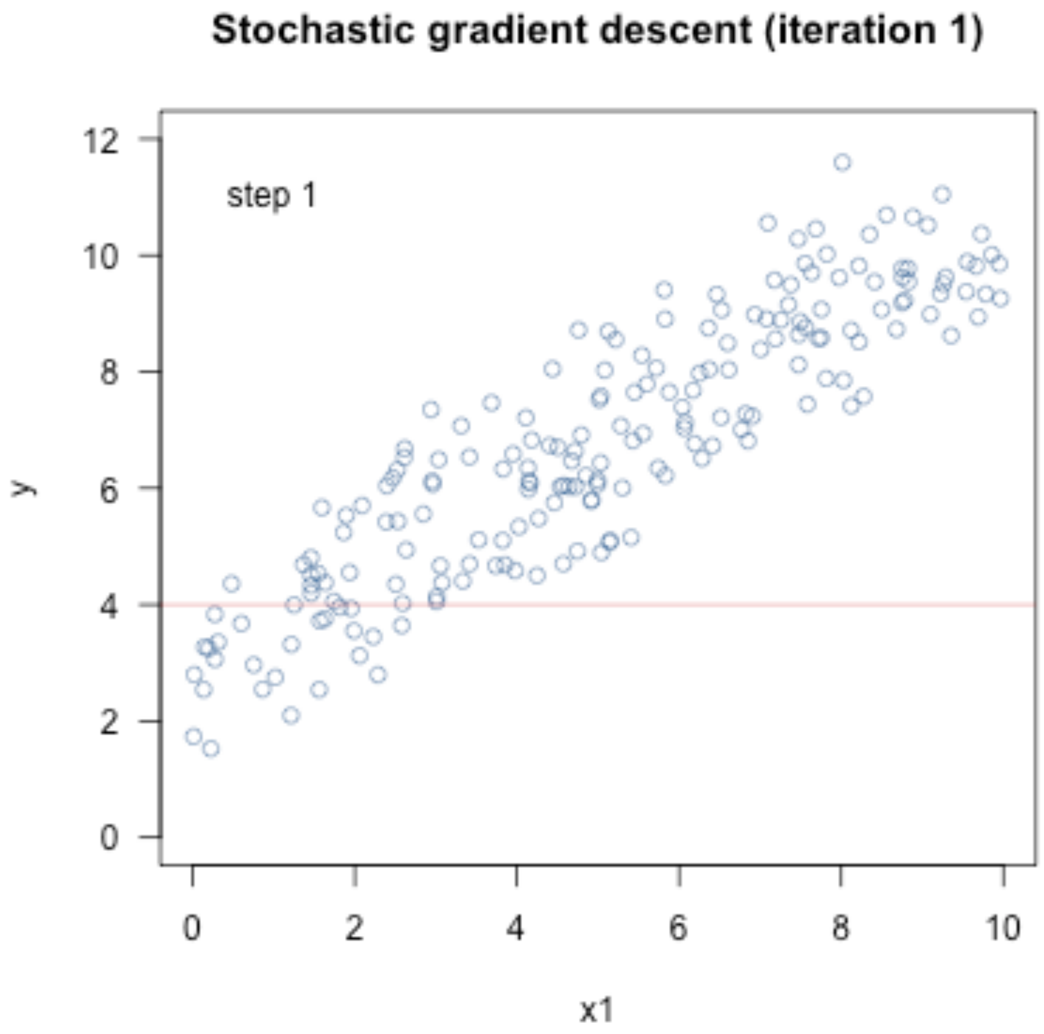
Parameter estimate minimizes the sum of squared residuals

Usually this is solved with a matrix inversion

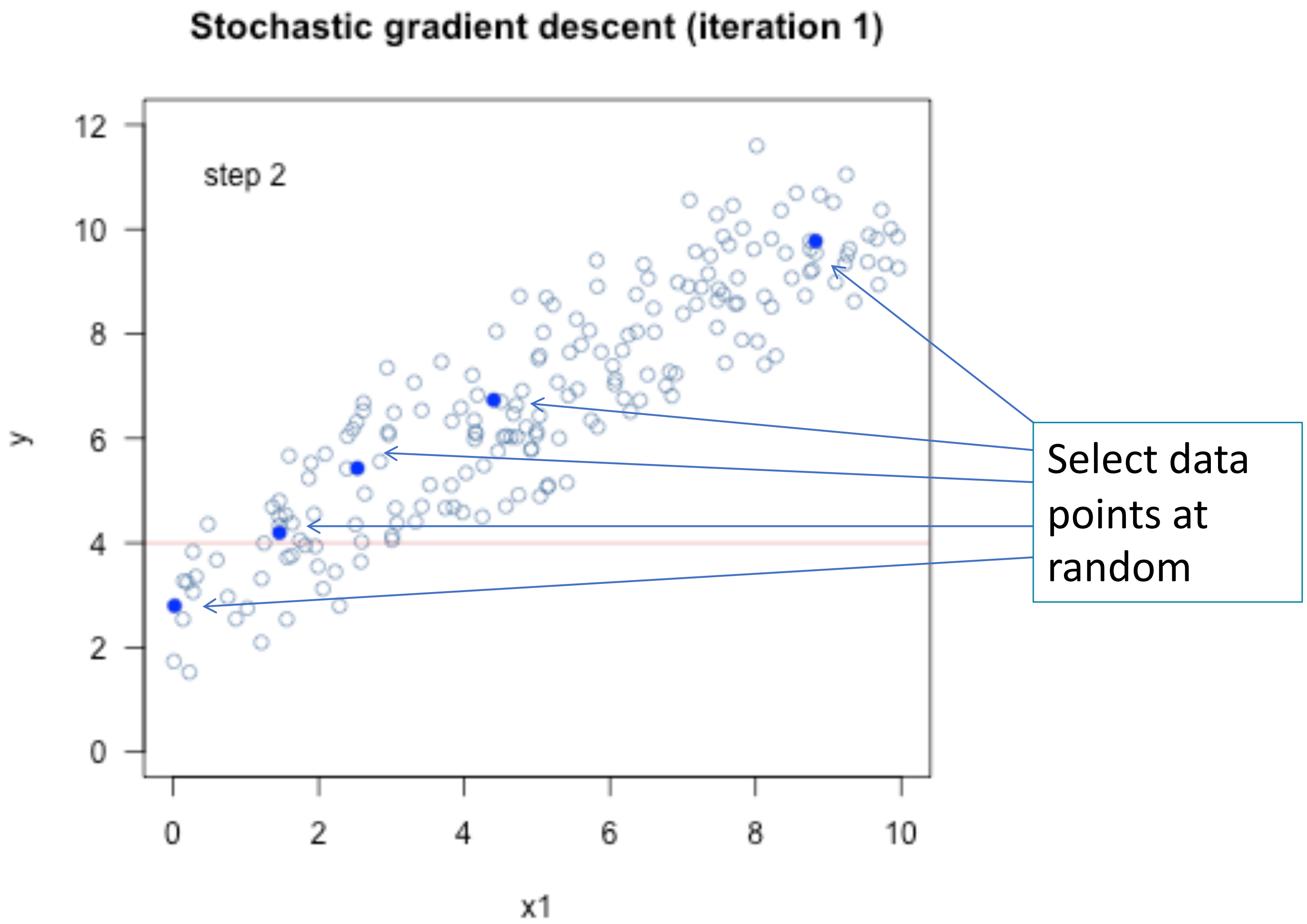
Which could be a BIG problem with BIG data!

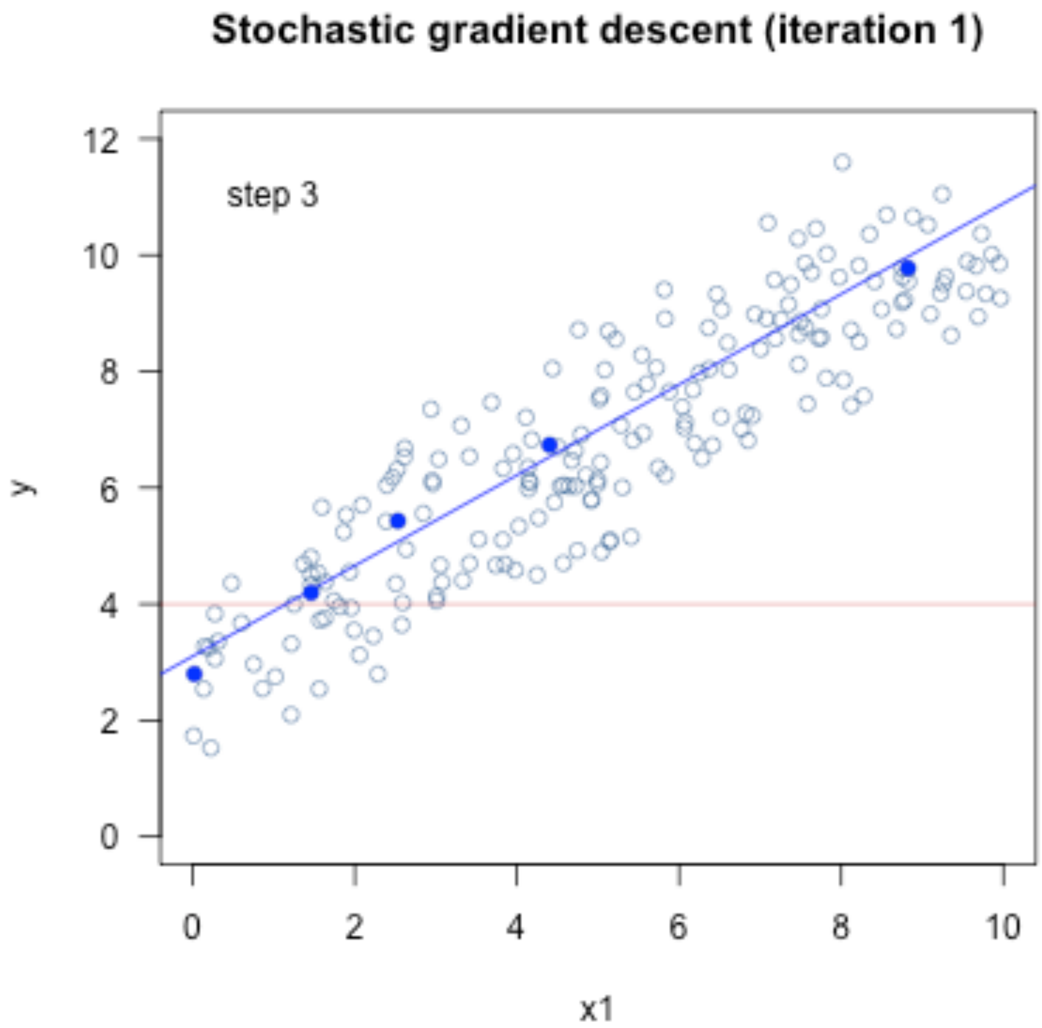
Machine learning solves this by ‘learning’

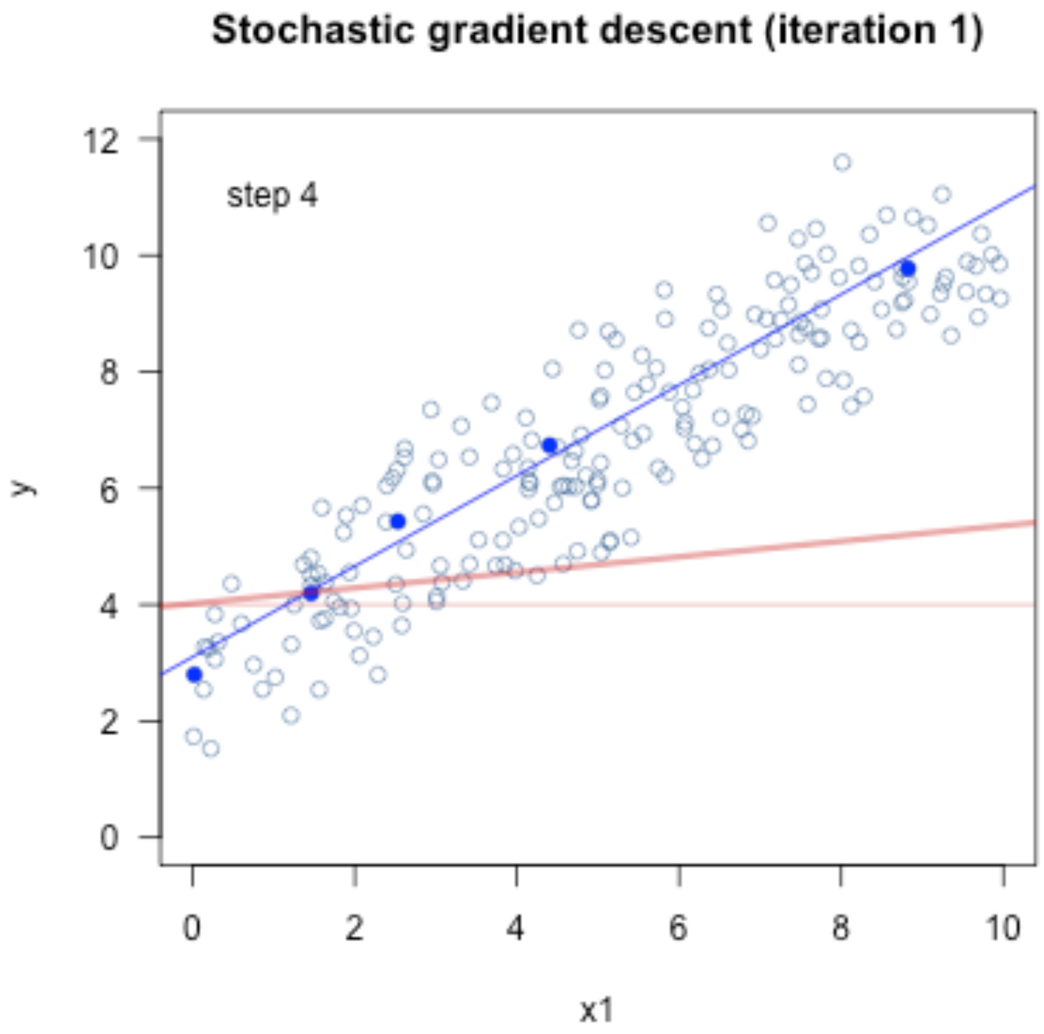
Example: stochastic gradient descent



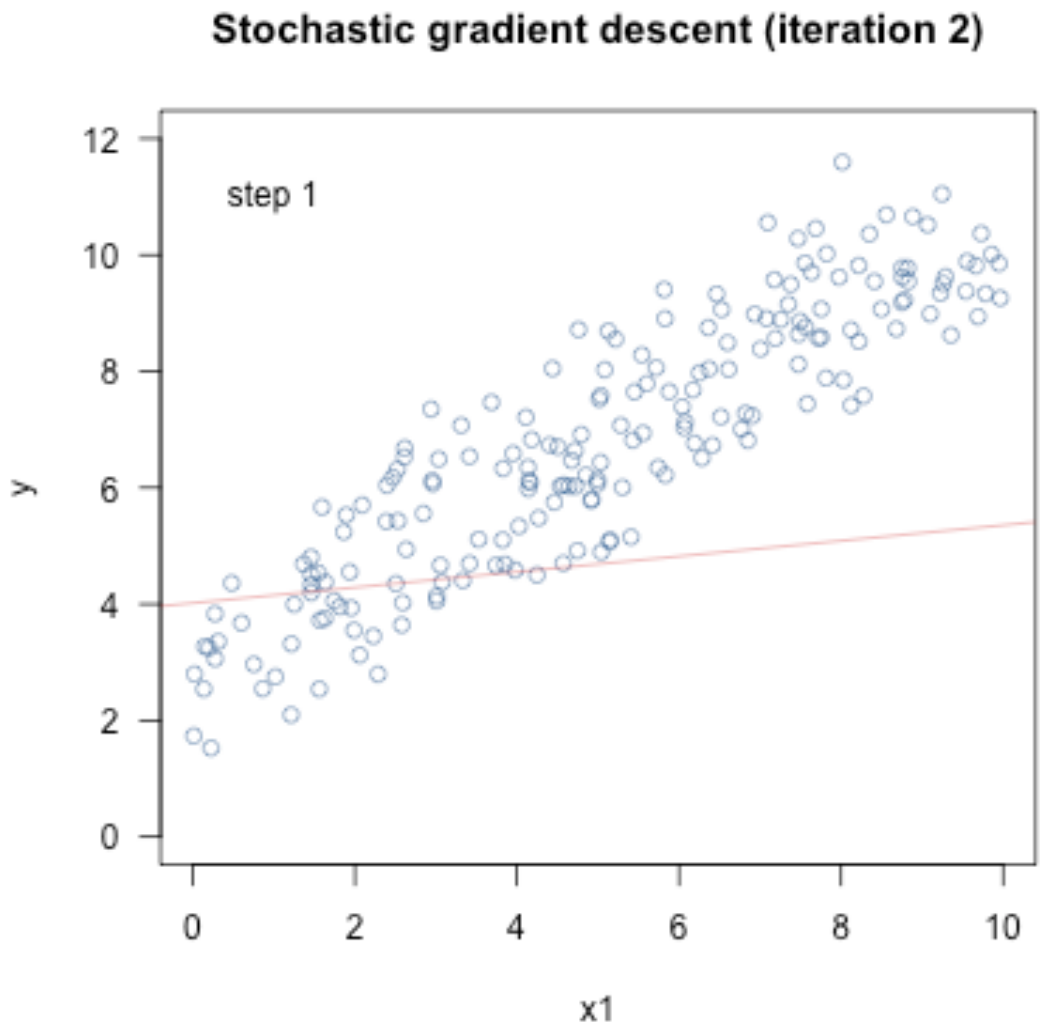
Some random start values of intercept and slope


Stochastic gradient descent (iteration 1)
step 2
y
x1
0
2
4
6
8
10
12
Select data points at random

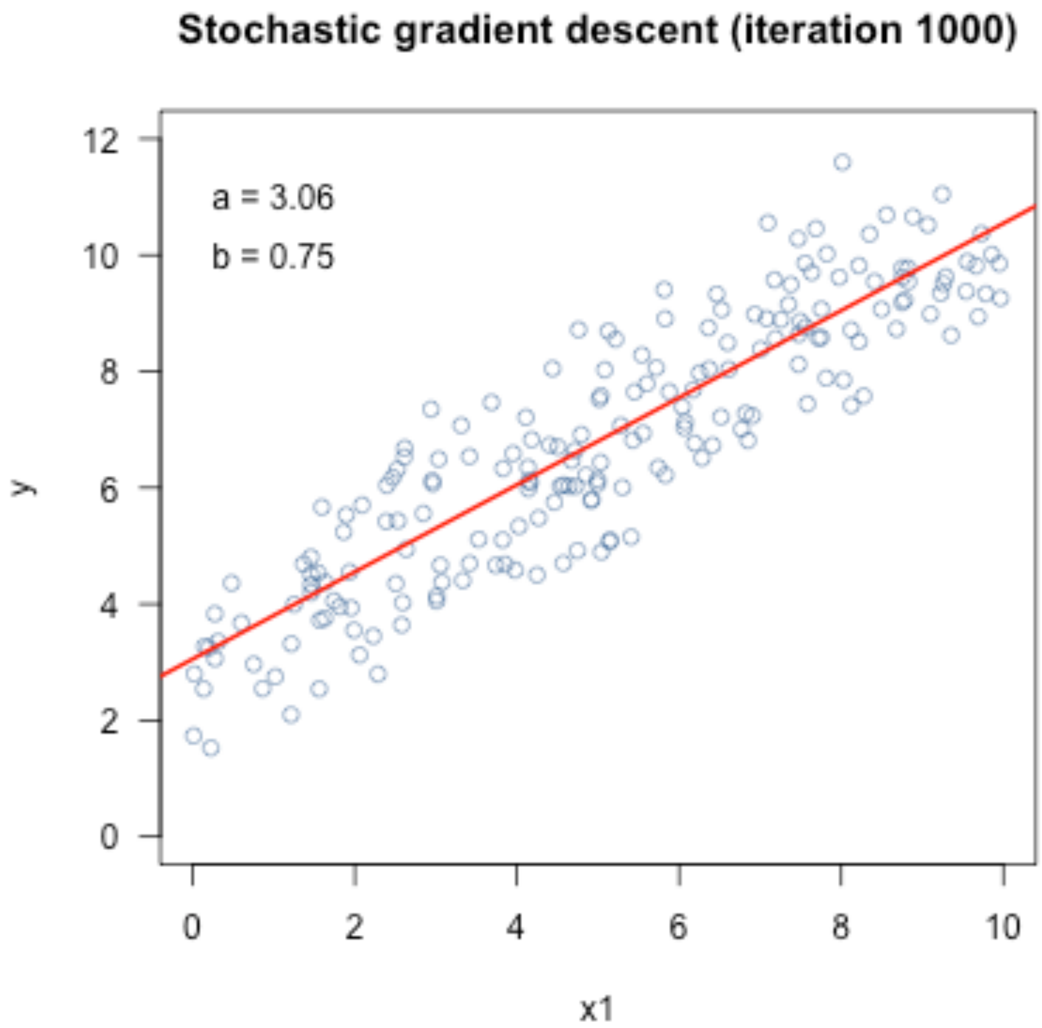

Stochastic gradient descent (iteration 1)
step 3
y
x1
0
2
4
6
8
10
12
0
2
4
6
8
10



New best guess of intercept & slope

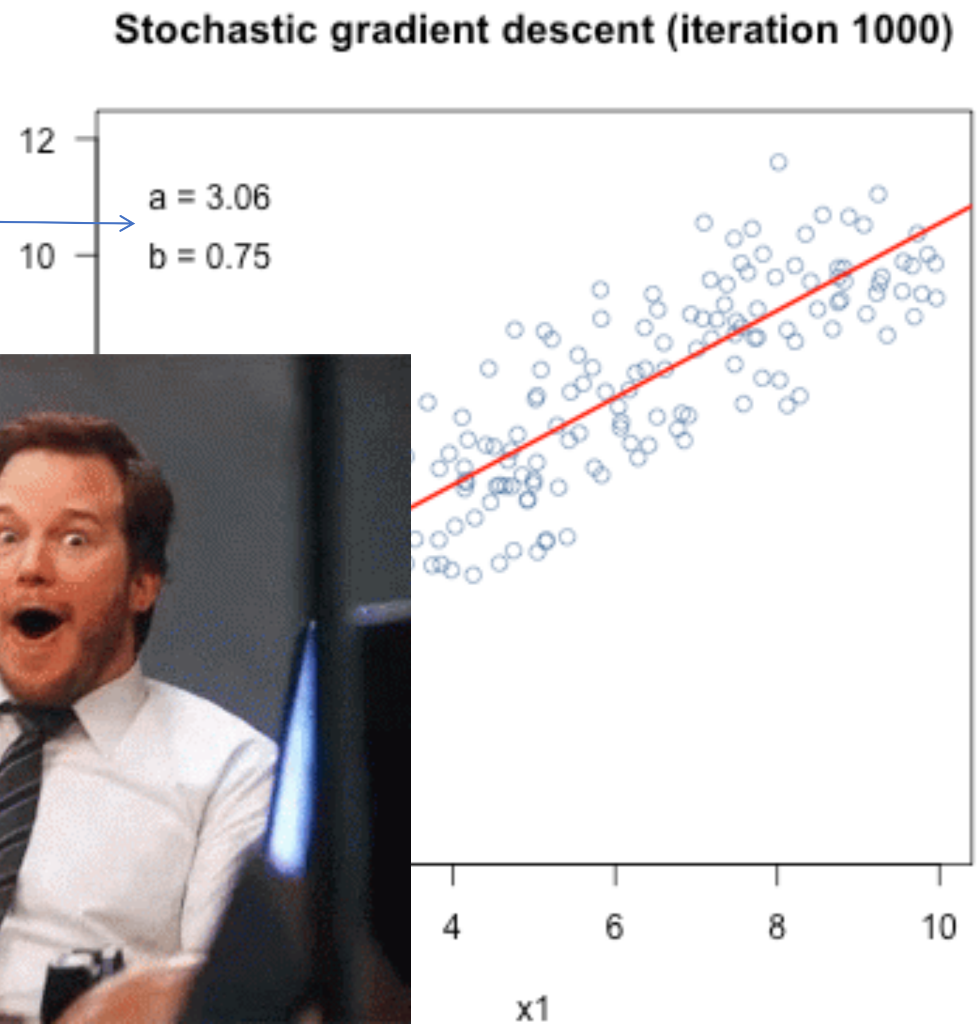


Starting value of intercept & slope in next iteration



Stochastic gradient descent (iteration 1000)
a = 3.06
b = 0.75
y
x1
0
2
4
6
8
10
12

The algorithm has “learned” from the data



Two kinds of machine learning: **Supervised** and Unsupervised

With **supervised** learning, the algorithm uses *training* data to learn how to make *predictions* given a certain input.

E.g. for crime, we might have information on each neighbourhood which we can use to train our algorithm to try to predict the crime rate.

Area	Deprivation	Proportion of young males	Policing budget	Crime Rate
A	15	0.6	500,000	5
B	190	0.5	350,000	80
C	60	0.7	500,000	70
D	20	0.3	550,000	15
E	120	0.5	120,000	30

Two kinds of machine learning: **Supervised** and Unsupervised

And can then use the algorithm to estimate the crime rate in unknown areas, or in areas where something changes (e.g. a decrease in the police budget)

Area	Deprivation	Proportion of young males	Policing budget	Crime Rate
F	20	0.55	400,000	??

Two kinds of machine learning: Supervised and **Unsupervised**

With **unsupervised** learning, instead of trying to make predictions based on our data, we look for groups of similar patterns.

E.g. in retail, we might try to classify stores into particular types given their characteristics

Store	Floor square area	No. parking spaces	Distance to the city centre	Store category
A	15	10	500	?
B	190	500	350	?
C	60	210	5,000	?
D	20	0	10	?
E	120	90	1,500	?

UNIVERSITY OF LEEDS

N8 POLICING RESEARCH PARTNERSHIP

Two kinds of machine learning: Supervised and Unsupervised

The practical sessions today will go through examples of both

Morning: Natural Language Processing (**unsupervised**)

- Read some text (crime notes) and try to group similar crimes
- The groups are not predetermined; the algorithm is not given the answer beforehand

Afternoon: Hate Speech Classification (**supervised**)

- Read some text (tweets) that have been labelled as ‘hate speech’
- The algorithm looks for patterns of words/phrases that are indicative of hateful speech.
- It then tries to determine whether a new tweet (which it hasn’t seen before) is hateful (or not)

UNIVERSITY OF LEEDS

N8 POLICING RESEARCH PARTNERSHIP

Problems

Overfitting

Is the problem solvable with the data?

- If we had data on the number of people with first names beginning with ‘S’ then we could not predict crime!

Is the model appropriate?

- E.g. the previous example (basically linear regression) assumes that the inputs have a linear effect on the outputs
  - If police budgets go up, the crime rate goes down.
- But what if areas with very high deprivation and lots of young men actually have higher community cohesion and lower crime?

Data

Machine learning algorithms typically need **loads** of data.

- 10,000+ images of a single object for image recognition?

Unexpected benefit: big AI/tech companies a releasing their algorithms (free and open source). E.g.

- Tensor Flow (Google, machine learning)
- DeepMask and SharpMask (Facebook, image recognition)

The value is in the *data*, not the algorithms

The police have loads of data!

Some common machine learning algorithms

Neural Networks

Regression Trees & Random Forest

Naive Bayes

K-Nearest Neighbors

Learning Vector Quantization

Support Vector Machines

https://towardsdatascience.com/a-tour-of-the-top-10-algorithms-for-machine-learning-newbies-dde4edffae11

UNIVERSITY OF LEEDS

A Tour of The Top 10 Algorithms for Machine Learning Newbies

Conclusion: What is machine learning?

It is not wildly different to statistical methods that you already know

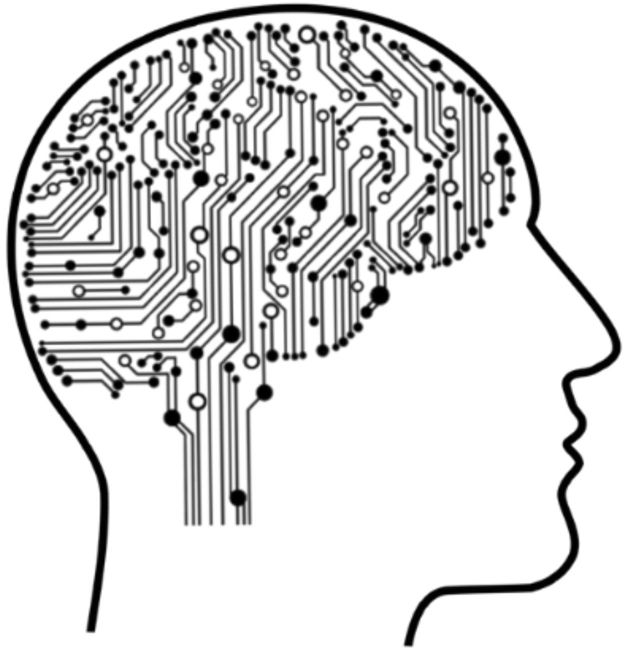
It's about 'learning' patterns from data

It isn't hugely complicated

- Although, as you'll see, it can be fiddly

There is a huge variety of methods that suit different tasks

- Although many are quite similar



UNIVERSITY OF LEEDS

N8 POLICING RESEARCH PARTNERSHIP

If you want more ...

Read the blog on medium

(Search for 'Machine Learning is Fun')

Do Andrew Ng's Machine Learning course on Coursera *(it's really good)*

https://www.coursera.org/learn/machine-learning

I also took loads of material from this blog post:

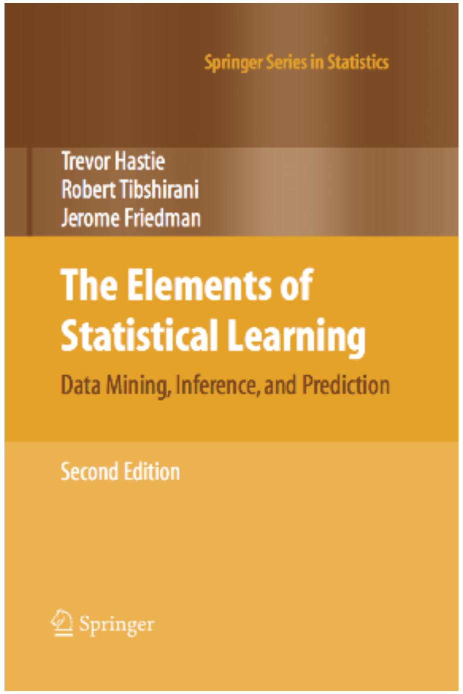
https://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets/

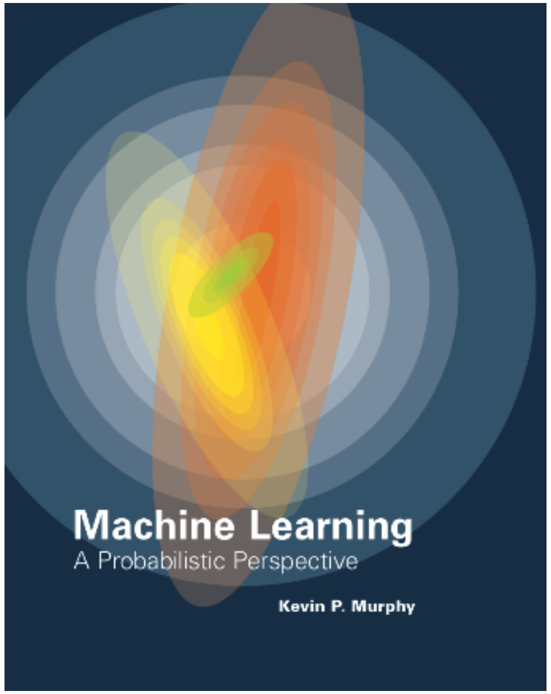
There are lots of practical text books

UNIVERSITY OF LEEDS

N8 POLICING RESEARCH PARTNERSHIP

If you want more …





A First Encounter with Machine Learning

Max Welling
Donald Bren School of Information and Computer Science
University of California Irvine

April 21, 2010

UNIVERSITY OF LEEDS

N8 POLICING RESEARCH PARTNERSHIP

Session 1
An Introduction to Machine Learning and Text Analysis

Presentation 1b: Natural Language Processing

UNIVERSITY OF LEEDS

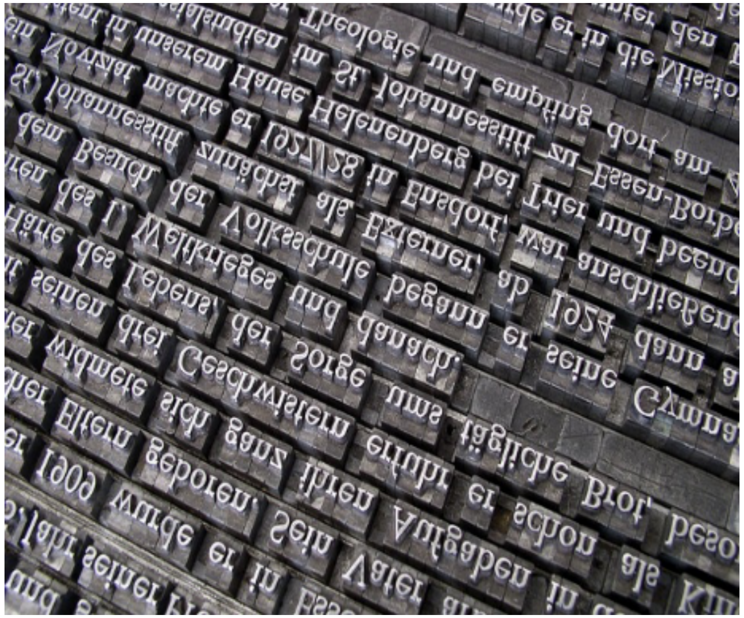
N8 POLICING RESEARCH PARTNERSHIP

Natural Language Processing

The field of study that links **human language** and **computers**.

- Computer science
- Artificial intelligence
- Computational linguistics

Obtain meaning from human language *in a useful way*.



Source: https://blog.algorithmia.com/introduction-natural-language-processing-nlp/

UNIVERSITY OF LEEDS

N8 POLICING RESEARCH PARTNERSHIP

Common Uses of Natural Language Processing

Automatic text summarization
- Summarise the main points in a document

Sentiment analysis
- Opinion mining; determining how someone feels by what they say

Topic extraction
- Clustering; trying to identify distinct topics in text

Named entity recognition
- E.g. people, objects, places, organisations, times, ..

Stemming etc.
- Identifying the roots of words.
- E.g. argue, argued, argues, arguing

UNIVERSITY OF LEEDS

Source: https://blog.algorithmia.com/introduction-natural-language-processing-nlp/

N8 POLICING RESEARCH PARTNERSHIP

NLP Tasks: Whitelisting

Crime notes may contain identifying information such as people’s names

These can be hard to anonymise manually, especially with large volumes

Fortunately, a few words usually make up most of the text

Whitelisting:

Identify these common words

- Check that they’re not disclosive
- Remove *all other words* from the crime notes

There is an example of whitelisting in the appendix

UNIVERSITY OF LEEDS

N8 POLICING RESEARCH PARTNERSHIP

NLP Tasks: Cleaning the Text

Data usually need to be cleaned before analysis

This includes

- Removing strange characters
- Removing case (e.g. make all lower-case)
- Taking out non-words (e.g. URLs, hash tags, etc.)
- Removing *stop words* (the, a, him, her, with, by, would)

UNIVERSITY OF LEEDS

N8 POLICING RESEARCH PARTNERSHIP

NLP Tasks: Tokenizing / Segmenting

Splitting a long sequence of text into distinct elements
- Paragraphs, sentences, words, etc.

Can be surprisingly difficult!
- Full stop used to mark abbreviations (e.g. ‘She travelled from the U.K. to France.’)
- And to stop the sentence (e.g. ‘But, after Brexit, she wasn’t allowed back to the U.K.’).
- Etc.

Fortunately, ‘tokenizers’ exist
- Someone else has done the hard work

Here, we use a word tokenizer to break down the crime notes into all of the distinct words

Drawback: Sentences are ignored

Useful Preliminary Analysis

Common words

The most common words in the text

Colocations

Pairs (or 2+) words that occur commonly together. In our data:

taxi driver; city centre; mobile phone; without paying; causing damage; calls police.

Concordance

The *context* that words appear in

Can be useful to see if words often occur in similar (or different) context. E.g.:

taxi driver given **phone** back
victim leaves **phone** taxi
passenger take **phone** leaving taxi
victim leaves **phone** car

UNIVERSITY OF LEEDS

N8 POLICING RESEARCH PARTNERSHIP

Topic Modelling

Topic modelling is a way to organise separate texts into similar *topics*

"Topics explain co-occurrences of words in documents with sets of semantically related words, called topics" (http://topicmodels.west.uni-koblenz.de/)

Topics can be *interpreted by humans* (?)

Question: is this **supervised** or **unsupervised**?

Example – Consider three sentences:

I eat fish and vegetables.
Fish are pets.
My kitten eats fish.

(from https://algobeans.com/2015/06/21/laymans-explanation-of-topic-modeling-with-lda-2/)

Topic Modelling: Latent Dirichlet Allocation (LDA)

LDA is a commonly used topic modelling algorithm

LDA might identify two topics:

Red: Topic A

Blue: Topic B

What would you name the topics?

The algorithm would classify as follows:

Sentence 1: 100% Topic A

Sentence 2: 100% Topic B

Sentence 3: 67% Topic A and 33% Topic B

1. I eat fish and vegetables.
2. Fish are pets.
3. My kitten eats fish.

UNIVERSITY OF LEEDS

Source: https://algobeans.com/2015/06/21/laymans-explanation-of-topic-modeling-with-lda-2/)

N8 POLICING RESEARCH PARTNERSHIP

LDA: How it works

1. Define the number of topics
2. Assign each word to a temporary topic.
3. Loop over every word and calculate:
   - How prevalent is that word across topics?
   - How prevalent are topics in the document?

   Weight the above conclusions and consider which topic to assign the word
4. Reassign words as appropriate, then repeat (3)

1. I eat fish and vegetables.
2. Fish are pets.
3. My kitten eats fish.

UNIVERSITY OF LEEDS

Source: https://algobeans.com/2015/06/21/laymans-explanation-of-topic-modeling-with-lda-2/)

N8 POLICING RESEARCH PARTNERSHIP

LDA: Advantages and Disadvantages

Advantages:

LDA provides a *distribution* over topics

- We don't lose information about the uncertainty with each prediction

Disadvantages:

Need to define the number of topics

- Trade-off between interpretability and statistical certainty
- How many is correct?
- Changing the number of topics can vastly alter the results

UNIVERSITY OF LEEDS

Source: https://algobeans.com/2015/06/21/laymans-explanation-of-topic-modeling-with-lda-2/)

N8 POLICING RESEARCH PARTNERSHIP

An alternative approach: Classification

Similar to topic modelling, but **supervised**

Give the algorithm a load of documents, *with labels*

Create a classification by 'learning' which words/phrases lead to a particular classification

Predict the classification of new documents

Classification Case Study:

Hate in England and Wales

UNIVERSITY OF LEEDS

Online hate: a proxy for hate crime

UNIVERSITY OF LEEDS

N8 POLICING RESEARCH PARTNERSHIP

Locations of tweets

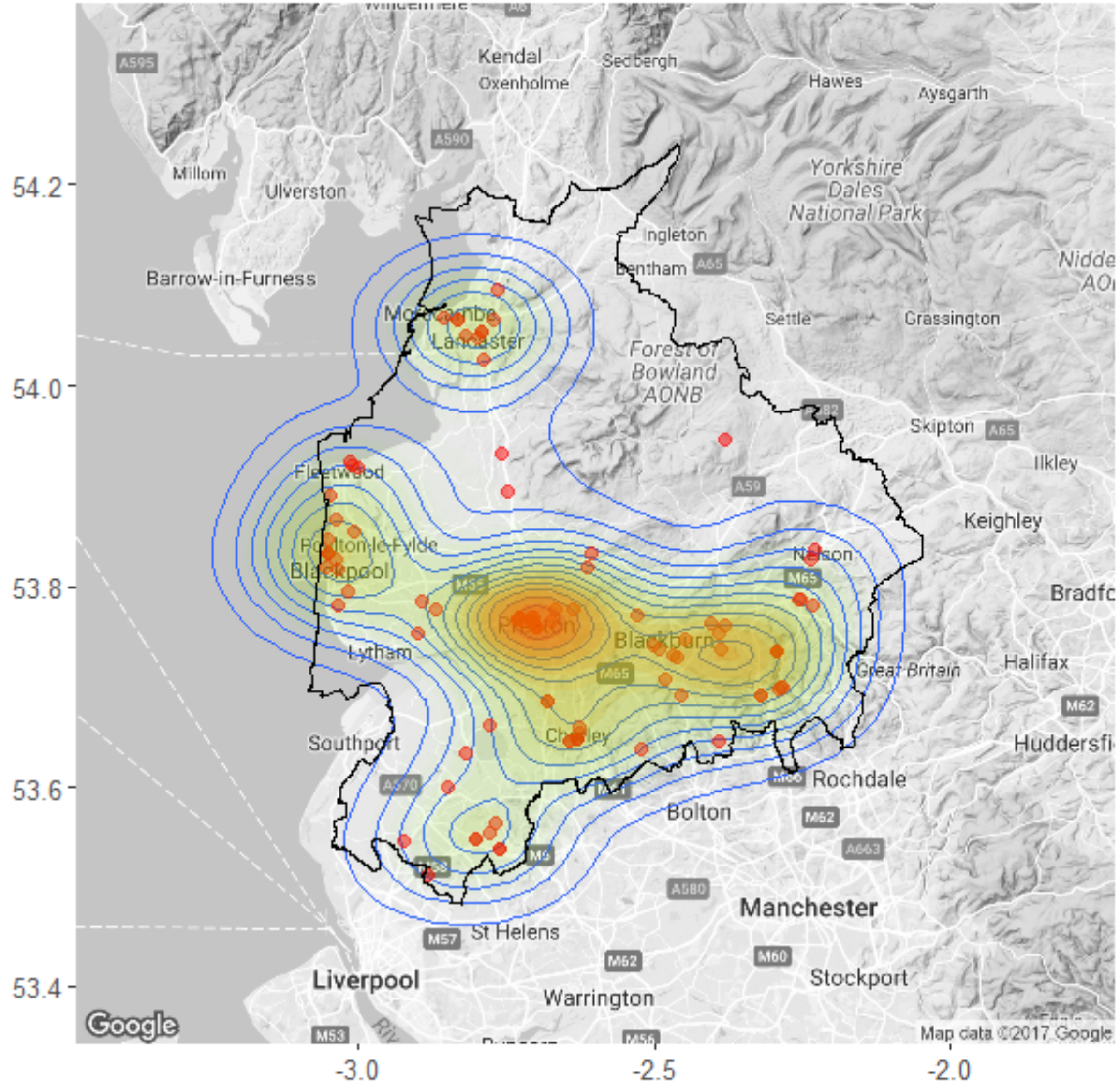
1. Hometown location

2. Geotag

UNIVERSITY OF LEEDS

N8 POLICING RESEARCH PARTNERSHIP

Locations of hateful tweet



UNIVERSITY OF LEEDS

Where to start?

UNIVERSITY OF LEEDS

Where to start?

"I HATE YOU!"

UNIVERSITY OF LEEDS

Some of the many hateful terms in English

UNIVERSITY OF LEEDS

N8 POLICING RESEARCH PARTNERSHIP

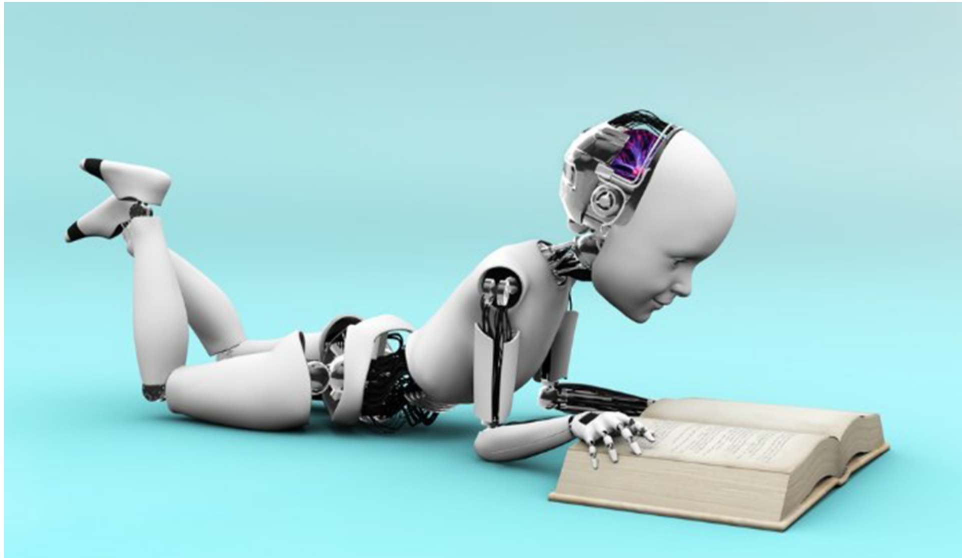
Training a ML algorithm to identify hate speech in tweets

Natural Language Processing

↓

Sentiment Analysis

(supervised learning)



UNIVERSITY OF LEEDS

N8 POLICING RESEARCH PARTNERSHIP

How do we do this?

RT @lildurk: I figured this shit out	0
When will be Kyungie\'s turn_ @exotic_tiny ?	0
#RT #Follow #TopStories Spain\'s Economy Expands 0.7\% Because Or Despite Political Deadlock? - Forbes https://t.co/uuv5KQ5vuY	0
RT @AllyBrooke: ...out and I should be fine for tomorrow :) I LOVE YOU	0
Probably gonna recolor/shade this head_ crep ã€Œã‚¨ãƒˆã‚»ãƒˆãƒ©ã€?	0
@MoriTaheripour shut up nigger whore! Hope u get raped by one of those animals. Might change your tune.	1
Fuck dykes	1
@sizzurp__ @ILIKECATS74 @yoPapi_chulo @brandonernandez @bootyacid at least i dont look like jefree starr faggot	1
@elaynay your a dirty terrorist and your religion is a fucking joke, you go around screaming Allah akbar doing terrorist shit. Dirty faggot.	1
I hate faggots like you	1

1. Train

Look for presence/absence of important features to differentiate hate versus non-hate.

2. Test

Make predictions and validate them against true labels.

UNIVERSITY OF LEEDS

N8 POLICING RESEARCH PARTNERSHIP

A few of the challenges

UNIVERSITY OF LEEDS

N8 POLICING RESEARCH PARTNERSHIP

A few of the challenges

UNIVERSITY OF LEEDS

N8 POLICING RESEARCH PARTNERSHIP

A few of the challenges

UNIVERSITY OF LEEDS

N8 POLICING RESEARCH PARTNERSHIP

A few of the challenges

UNIVERSITY OF LEEDS

POLICING
RESEARCH
PARTNERSHIP

Discussion: Uses for NLP in Policing

UNIVERSITY OF LEEDS

N8 POLICING RESEARCH PARTNERSHIP

Discussion points

How useful is this?

- Does it actually tell us anything useful?
- Might you use it in your analysis?
- What if it was possible in Excel?

Any improvements to the process?

- Have we chosen a suitable number of topics? Are they stable?
- Have we identified the correct stop words, etc.?

UNIVERSITY OF LEEDS

N8 POLICING RESEARCH PARTNERSHIP

Session 1
An Introduction to Machine Learning and Text Analysis

Practical 1: Exploring Different Types of Modus Operandi using Machine Learning

UNIVERSITY OF LEEDS

POLICING RESEARCH PARTNERSHIP

Practical 1: Exploring Different Types of Modus Operandi using Machine Learning

Aims:

- Introduce some of the natural language processing tools that are commonly used.

Will only touch on the most common/basic methods

Not an exhaustive tutorial!

Think about: what else could I do?

- I know what I might use this stuff for, but what is its relevance for policing?

UNIVERSITY OF LEEDS

N8 POLICING RESEARCH PARTNERSHIP

Natural Language Toolkit (NLTK)

The NLTK is a python library for analysing text

The NLTK Book is available online

- http://www.nltk.org/book
- It’s generally well written, but assumes some knowledge of python syntax

The NLTK library will do most of the work

UNIVERSITY OF LEEDS

N8 POLICING RESEARCH PARTNERSHIP

Jupyter Notebooks

Jupyter Notebooks
- Are documents that combine text with python code
- The practicals are written using Jupyter Notebooks

When you ’launch’ a notebook, it reads a `.pynb` file and loads it in a web browser.

The document is broken up into ‘chunks’
- These can be sections of python code or plain text (in a format called *markdown*).

You can then edit and run the code chunks interactively
- There will be some optional activities that will require you to add code to empty cells, and then run the code.

The easiest way to learn is by doing …

Python and it’s friends

Python is a multipurpose programming language.

- It does most of the normal things that programming languages do (making graphical interfaces, allowing computers to communicate, etc.)
- And also now very popular for *data analysis*

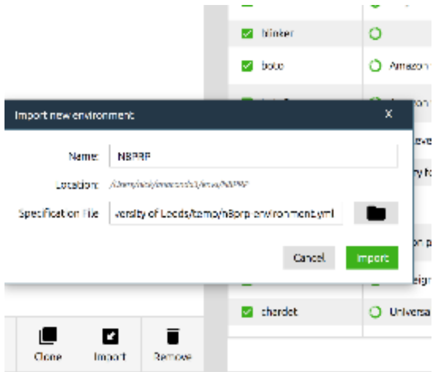
Anaconda

- Is a program that helps to manage python versions and install python libraries
- We’ll use Anaconda to install the libraries that we need and to run python

iPython

- A version of python that is easier to interact with than normal python (we’ll actually use this rather than normal python)

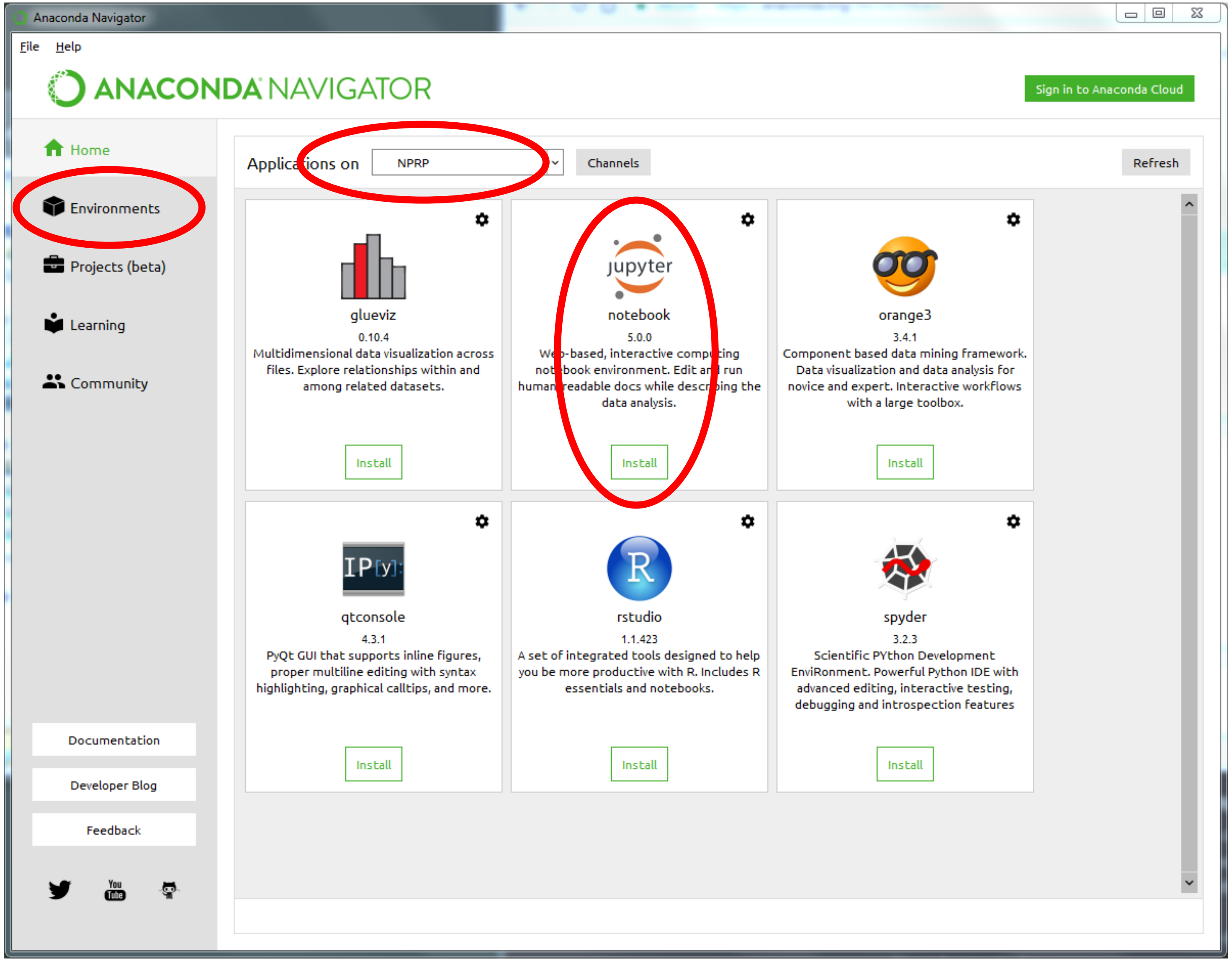
Practical 1: Instructions



1. Download the practical files and save them in your Downloads folder
   - https://tinyurl.com/y5y2t84k
2. Extract the zip file
3. Launch Anaconda Navigator
   - Press start button -> then type `anaconda-navigator` and press return
4. Import the N8 PRP 'environment' (this tells Anaconda which libraries we will need)
   - Click on '**Environments**'
   - Click on **Import**
   - Find the file called **`n8prp-environment.yml`**
   - Give the environment a sensible name **(try 'PRP')**
   - (If you get a question about 'allowing a program to make changes' click 'No')
5. Click on 'Home', make sure the *N8PRP* environment is selected
6. Start Jupyter Notebook (this will open a web browser)
7. Navigate to '`NLP.ipynb`' (not the '`COMPETE`' one)

UNIVERSITY OF LEEDS

N8 POLICING RESEARCH PARTNERSHIP


Anaconda Navigator
File Help
ANACONDA NAVIGATOR
Sign in to Anaconda Cloud
Home
Environments
Projects (beta)
Learning
Community
Applications on
NPRP
Channels
Refresh
glueviz
0.10.4
Multidimensional data visualization across files. Explore relationships within and among related datasets.
Install
notebook
5.0.0
Web-based, interactive computing notebook environment. Edit and run human-readable docs while describing the data analysis.
Install
orange3
3.4.1
Component based data mining Framework. Data visualization and data analysis for novice and expert. Interactive workflows with a large toolbox.
Install
qtconsole
4.3.1
PyQt GUI that supports inline figures, proper multiline editing with syntax highlighting, graphical calltips, and more.
Install
rstudio
1.1.423
A set of integrated tools designed to help you be more productive with R. Includes R essentials and notebooks.
Install
spyder
3.2.3
Scientific PYthon Development EnviRonment. Powerful Python IDE with advanced editing, interactive testing, debugging and introspection features
Install
Documentation
Developer Blog
Feedback

Session 3
Image Recognition with Machine Learning

Presentation 2: Image recognition using Machine Learning

UNIVERSITY OF LEEDS

N8 POLICING RESEARCH PARTNERSHIP


Machine Learning is Fun!
The world’s easiest introduction to Machine Learning

Machine Learning is Fun!

Fantastic series of blog posts on *medium*

https://medium.com/@ageitgey/machine-learning-is-fun-80ea3ec3c471

I'm basically just copying material from those blog posts

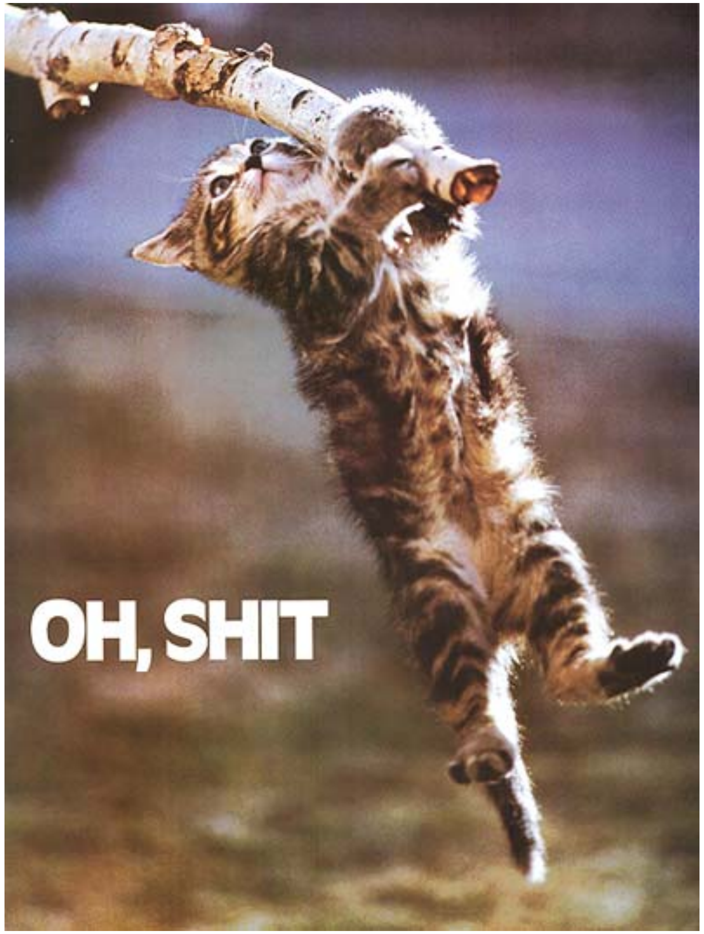
What I'm going to cover:

- **What is machine learning?** *(Everyone in this room probably knows this already, even if you don't realise it)*
- **What are neural networks?**
- **What are convolutional neural networks?**

Why?

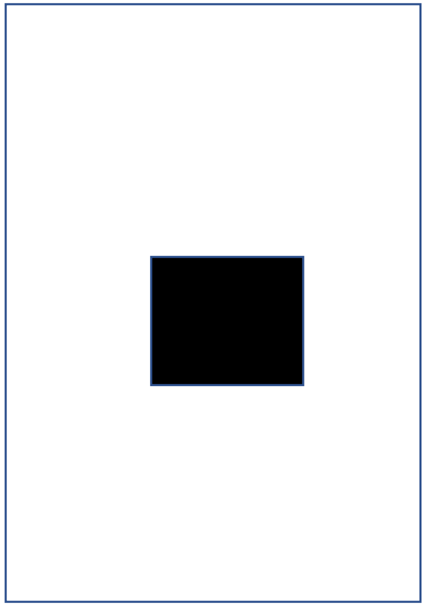
- Convolutional neural networks are absolutely amazing!!! (geek interest)
- And will be useful in policing (more serious interest)

[redacted]


OH, SHIT

(Recap) What is machine learning?

"Machine learning is the idea that there are generic algorithms that can tell you something interesting about a set of data without you having to write any custom code specific to the problem. Instead of writing code, you feed data to the generic algorithm and it builds its own logic based on the data."



UNIVERSITY OF LEEDS

N8 POLICING RESEARCH PARTNERSHIP

Why do we need machine learning?

The real world is:

- Noisy
- Full of outliers
- Multivariate causes
- Patchy (missing) data

Therefore:

- Hard to classify
- Hard to model
- Hard to understand

UNIVERSITY OF LEEDS

N8 POLICING RESEARCH PARTNERSHIP

Solution: learn from nature

Animals (including us!) live in this world, yet cope.

They recognise patterns in the chaos and noise.

So, we need a system that acts like an animal.

Neural networks take their inspiration from nature

UNIVERSITY OF LEEDS

N8 POLICING RESEARCH PARTNERSHIP

What are neural networks?

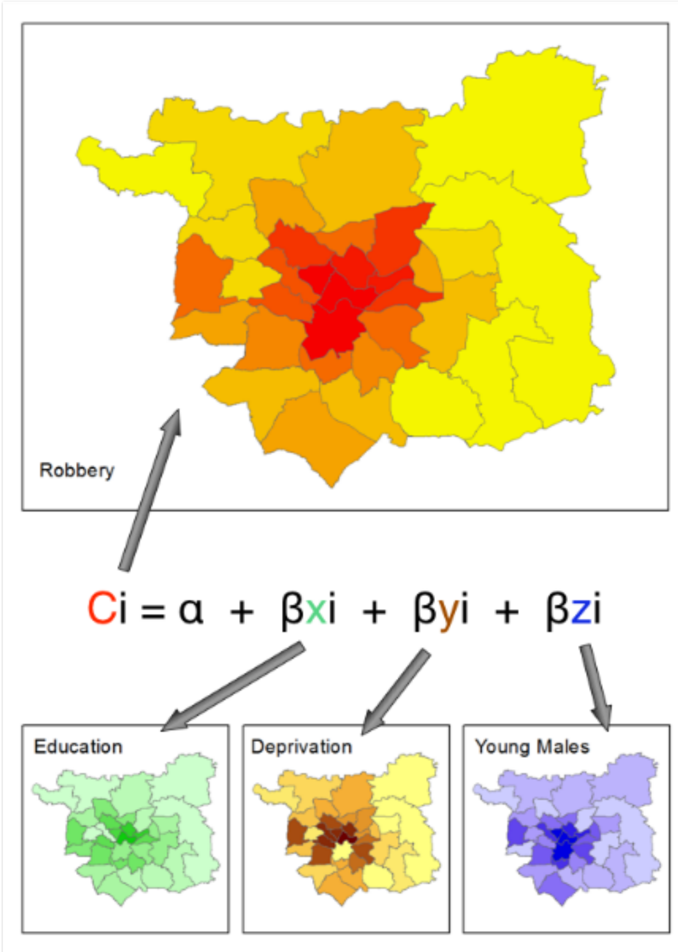
Example: regression for crime analysis

Predict crime rates

Quantify the impact of different factors on crime rates

Very commonly used!

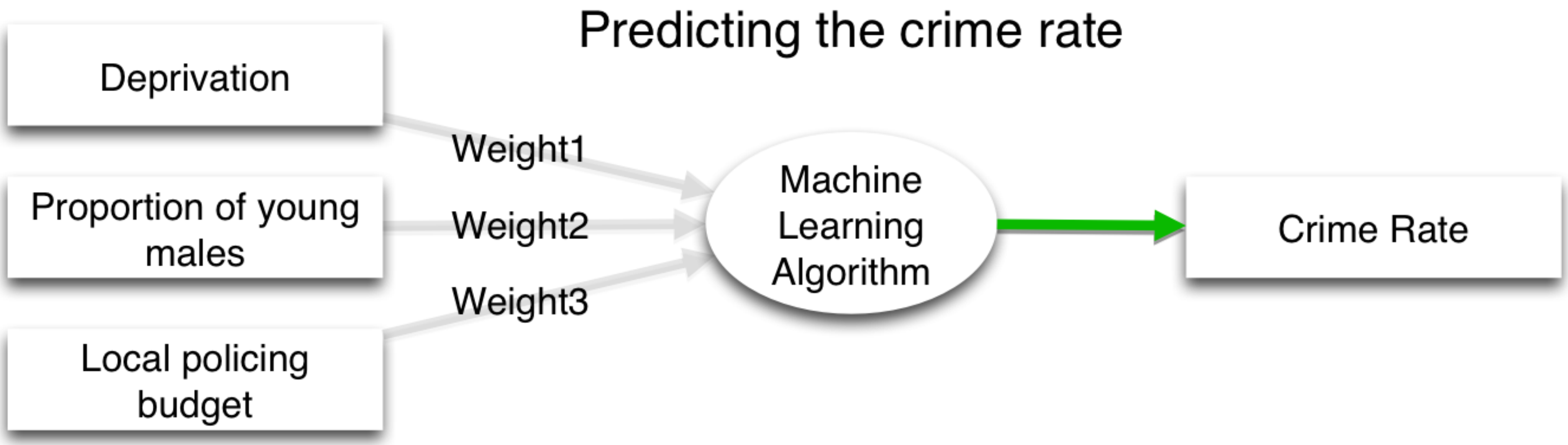
Great explanation: https://medium.com/@ageitgey/machine-learning-is-fun-part-2-a26a10b68df3



UNIVERSITY OF LEEDS

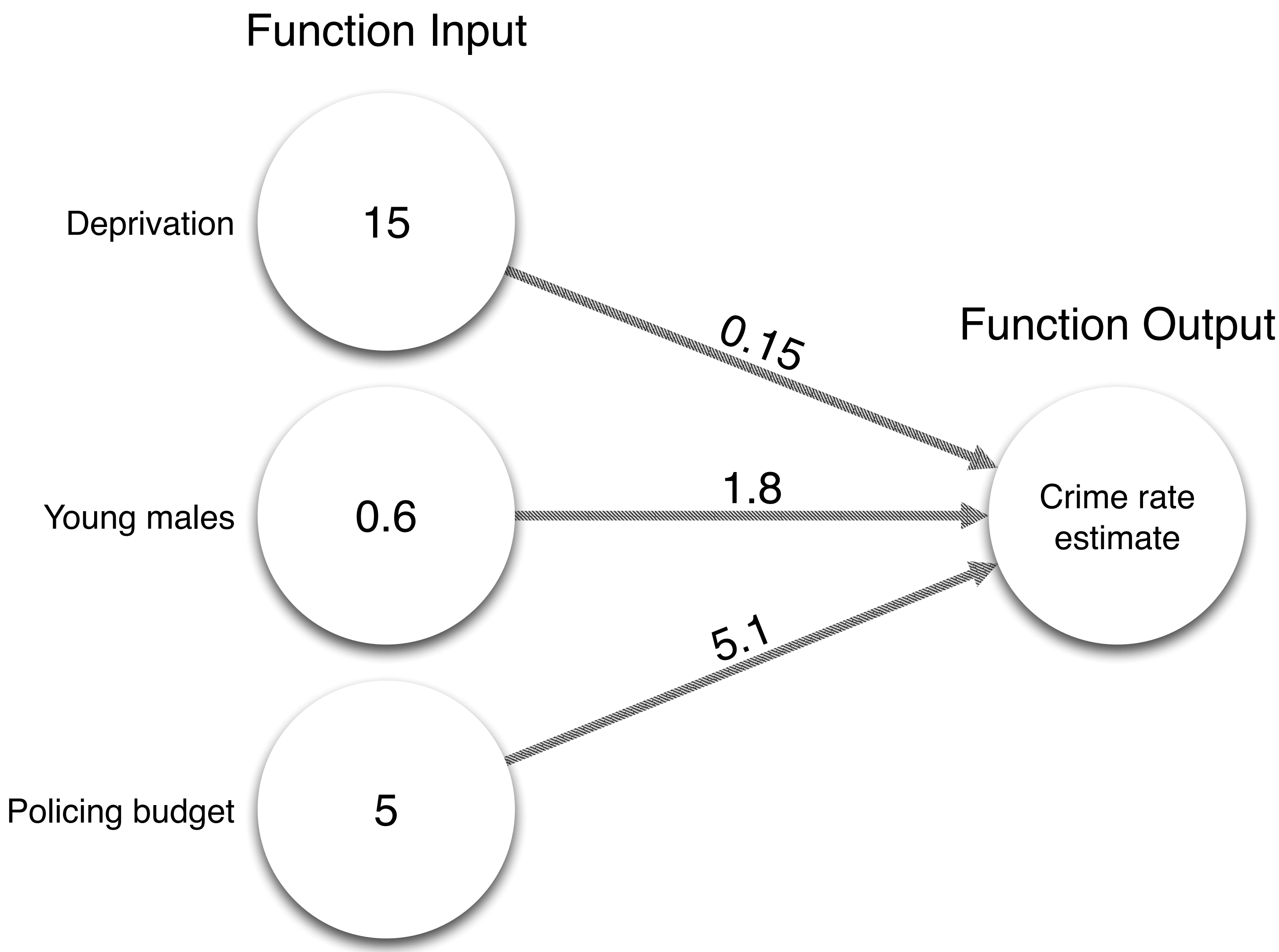
N8 POLICING RESEARCH PARTNERSHIP

What are neural networks?



UNIVERSITY OF LEEDS

N8 POLICING RESEARCH PARTNERSHIP

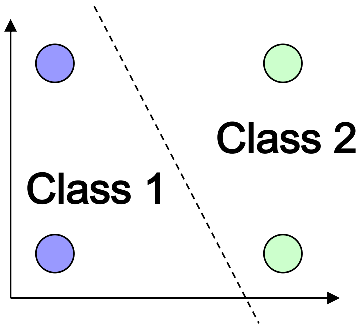

Function Input
Deprivation
15
Young males
0.6
Policing budget
5
0.15
1.8
5.1
Function Output
Crime rate
estimate

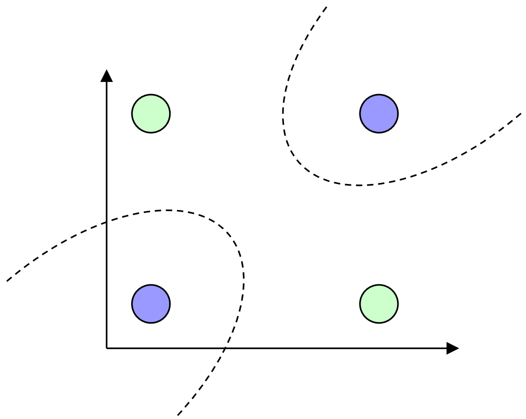
Problems

How to model non-linearity?

E.g. high deprivation -> community cohesion -> lower crime ?

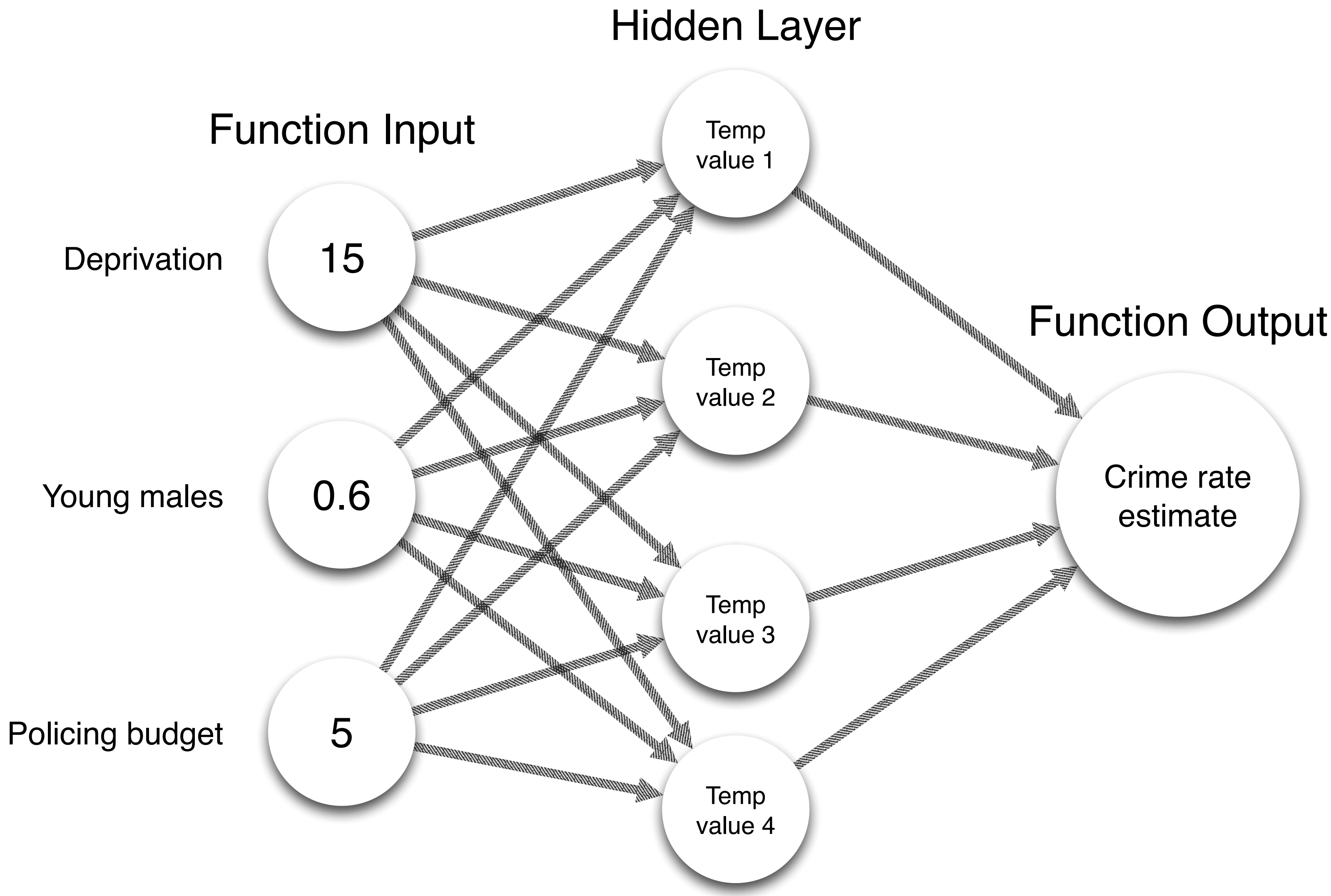
How to classify when straight lines are no good?





UNIVERSITY OF LEEDS

N8 POLICING RESEARCH PARTNERSHIP


Hidden Layer
Function Input
Function Output
Deprivation
15
Young males
0.6
Policing budget
5
Temp value 1
Temp value 2
Temp value 3
Temp value 4
Crime rate estimate

What are neural networks?

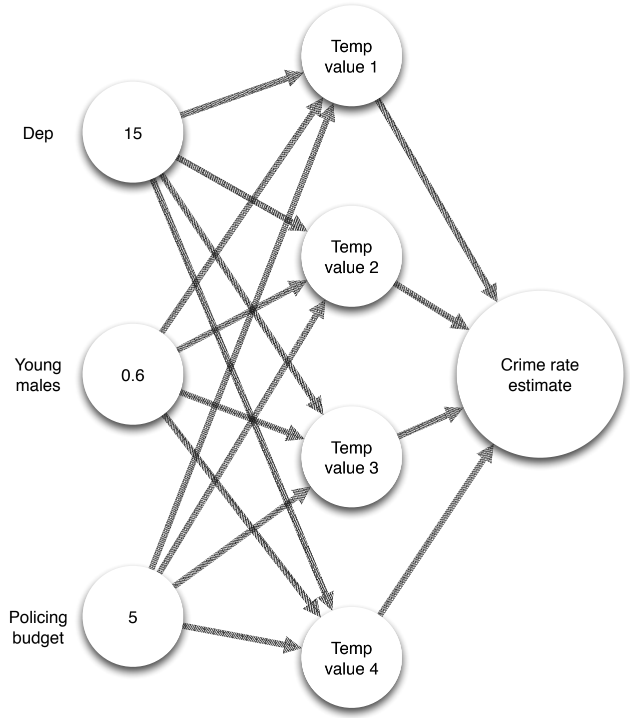
This is a simple neural network

We just add another layer of nodes in between the input and output

Then the task is to find values for all of the new weights, based on our input data.

You can do this with *backcasting* (I wont try to explain it though!)

And what about ‘**deep learning**’? You just add more hidden layers!



UNIVERSITY OF LEEDS

N8 POLICING RESEARCH PARTNERSHIP

What are convolutional neural networks?

In short – they are deep neural networks with clever hidden layers

Feed them loads of images of certain objects, and they should be able to learn which images contain the object and which don't

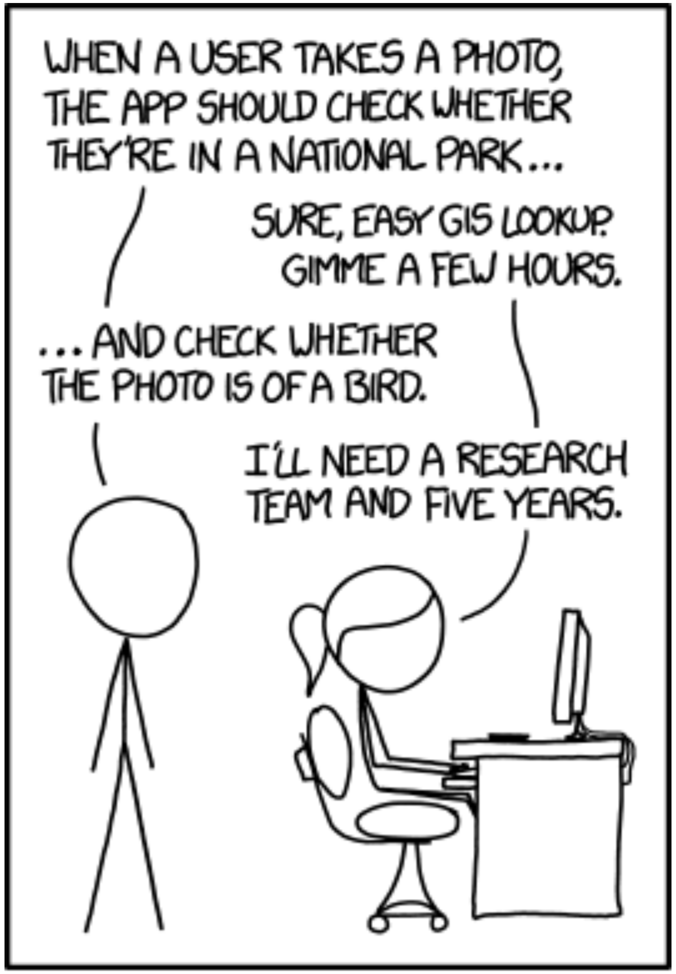
Credits:

- https://medium.com/@ageitgey/machine-learning-is-fun-part-3-deep-learning-and-convolutional-neural-networks-f40359318721
- I also draw heavily on: https://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets/
- Thank you to those blog authors!!

UNIVERSITY OF LEEDS

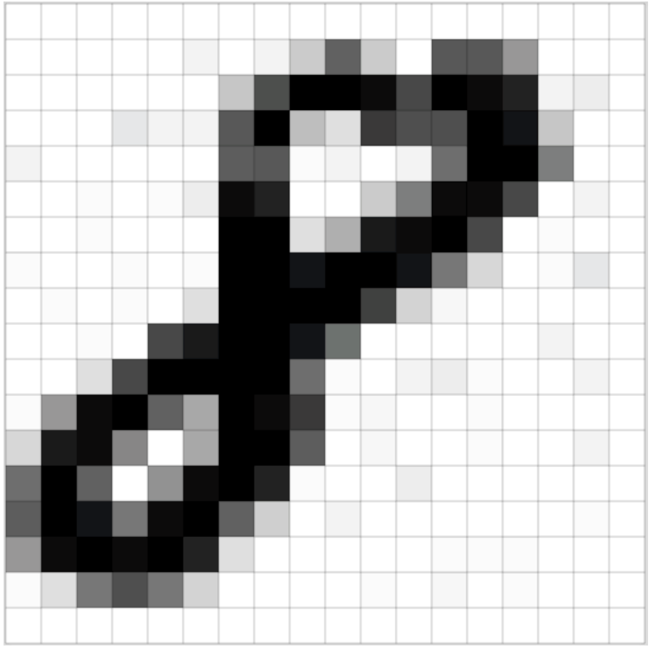
N8 POLICING RESEARCH PARTNERSHIP

https://xkcd.com/1425/



IN CS, IT CAN BE HARD TO EXPLAIN THE DIFFERENCE BETWEEN THE EASY AND THE VIRTUALLY IMPOSSIBLE.

On a computer, images are stored as lists of numbers



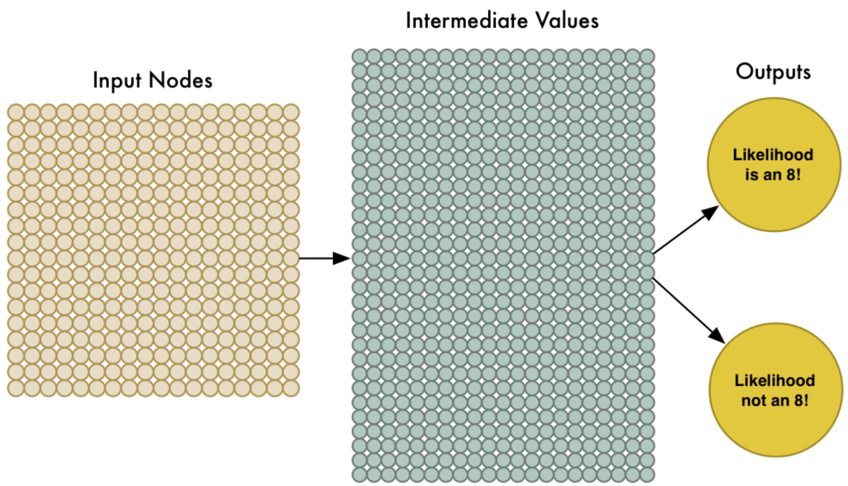
=

```
[0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 12, 0, 11, 39, 137, 37, 0, 152, 147, 84, 0, 0, 0, 0,
0, 1, 0, 0, 0, 41, 160, 250, 255, 235, 162, 255, 238, 206, 11, 13, 0, 0, 0, 0, 16, 9, 9, 150, 251, 45, 21, 184, 159, 154, 2
55, 233, 40, 0, 0, 10, 0, 0, 0, 0, 0, 145, 146, 3, 10, 0, 11, 124, 253, 255, 107, 0, 0, 0, 0, 3, 0, 4, 15, 236, 216, 0, 0,
38, 109, 247, 240, 169, 0, 11, 0, 1, 0, 2, 0, 0, 0, 253, 253, 23, 62, 224, 241, 255, 164, 0, 5, 0, 0, 6, 0, 0, 4, 0, 3, 252
, 250, 228, 255, 255, 234, 112, 28, 0, 2, 17, 0, 0, 2, 1, 4, 0, 21, 255, 253, 251, 255, 172, 31, 8, 0, 1, 0, 0, 0, 0, 0, 4,
 0, 163, 225, 251, 255, 229, 120, 0, 0, 0, 0, 0, 11, 0, 0, 0, 0, 21, 162, 255, 255, 254, 255, 126, 6, 0, 10, 14, 6, 0, 0, 9
, 0, 3, 79, 242, 255, 141, 66, 255, 245, 189, 7, 8, 0, 0, 5, 0, 0, 0, 0, 26, 221, 237, 98, 0, 67, 251, 255, 144, 0, 8, 0, 0
, 7, 0, 0, 11, 0, 125, 255, 141, 0, 87, 244, 255, 208, 3, 0, 0, 13, 0, 1, 0, 1, 0, 0, 145, 248, 228, 116, 235, 255, 141, 34
, 0, 11, 0, 1, 0, 0, 0, 1, 3, 0, 85, 237, 253, 246, 255, 210, 21, 1, 0, 1, 0, 0, 6, 2, 4, 0, 0, 0, 6, 23, 112, 157, 114, 32
, 0, 0, 0, 0, 2, 0, 8, 0, 7, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]
```

UNIVERSITY OF LEEDS

N8 POLICING RESEARCH PARTNERSHIP

What are convolutional neural networks?



On a computer, images are stored as lists of numbers

E.g. black and white picture with 18x18 pixels = 324 numbers

We can enlarge our neural network to have 324 input nodes; one for each pixel

To train the model: input loads of images and let the network try to work out what the image contains.

But, this only works on very simple images

It will break if the images are moved, rotated, etc.

Solution: **convolution**

UNIVERSITY OF LEEDS

N8 POLICING RESEARCH PARTNERSHIP

How convolution works

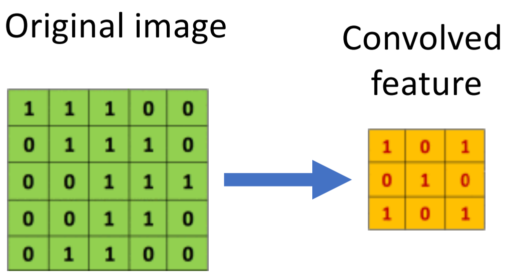
1. The image is broken into many overlapping tiles
2. Each tile is fed into a small neural network.
3. The small network (also called a 'filter') can be configured to pick out 'interesting' features.
4. The output of the filter is *convolved feature*.
5. Continue filtering, using different filters and some other clever image manipulations…

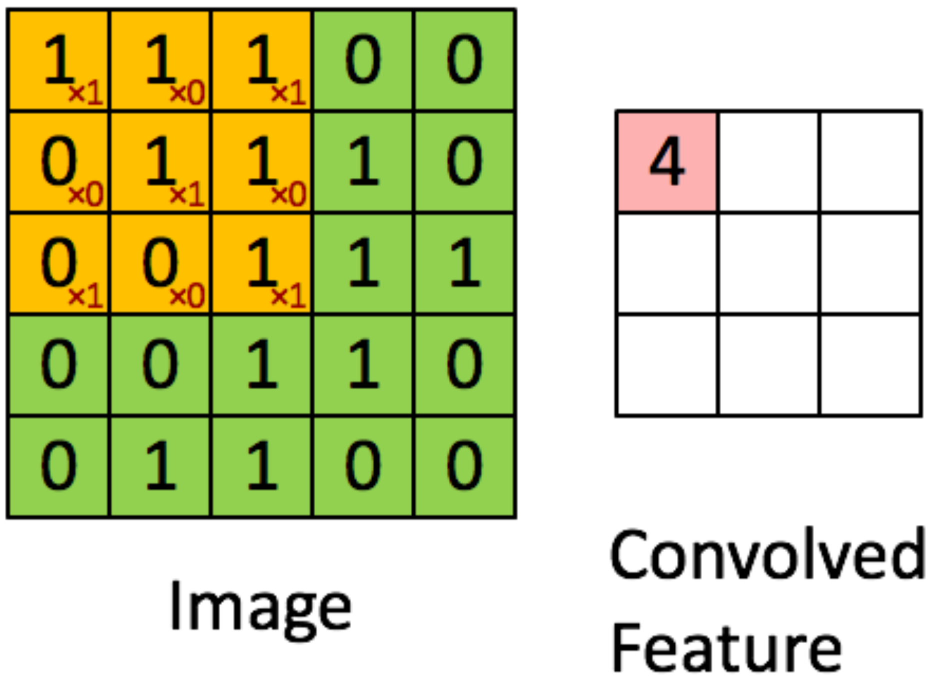
The really amazing thing: **the algorithm identifies the features that it needs to look for!!**

UNIVERSITY OF LEEDS

POLICING RESEARCH PARTNERSHIP

Creating the convolved feature





https://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets/

Input

https://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets/

Create a pipeline

Convolution

Non Linearity (ReLU)

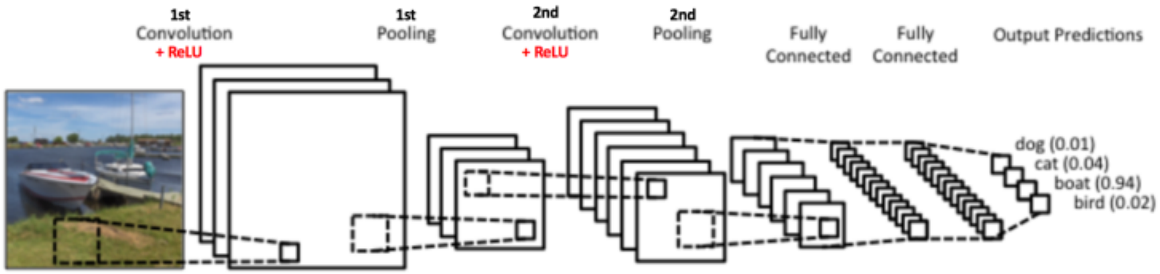
Introduce some non-linearity after convolution

Pooling

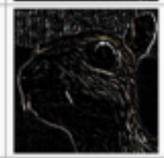
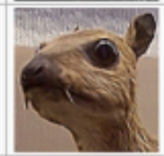
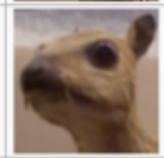
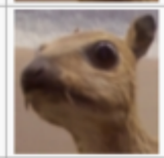
(down-sample, e.g. min, max, avg..)

Classification (fully connected layer)

use previously identified features to classify the image



https://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets/

Operation	Filter	Convolved Image
Identity	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$	
Edge detection	$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}$	
	$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$	
	$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$	
Sharpen	$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$	
Box blur (normalized)	$\frac{1}{9}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$	
Gaussian blur (approximation)	$\frac{1}{16}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$	

Non Linearity (ReLU)

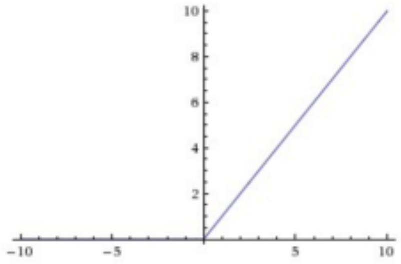
"Rectified Linear Unit"

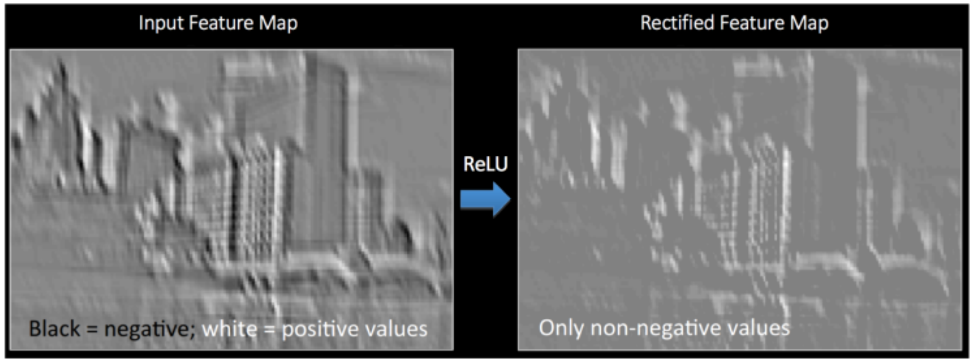
Introduce some non-linearity after convolution

Most data that the algorithm will be trained on are non-linear

Assign all negative values 0

Output = Max(zero, Input)





UNIVERSITY OF LEEDS

https://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets/

N8 POLICING RESEARCH PARTNERSHIP

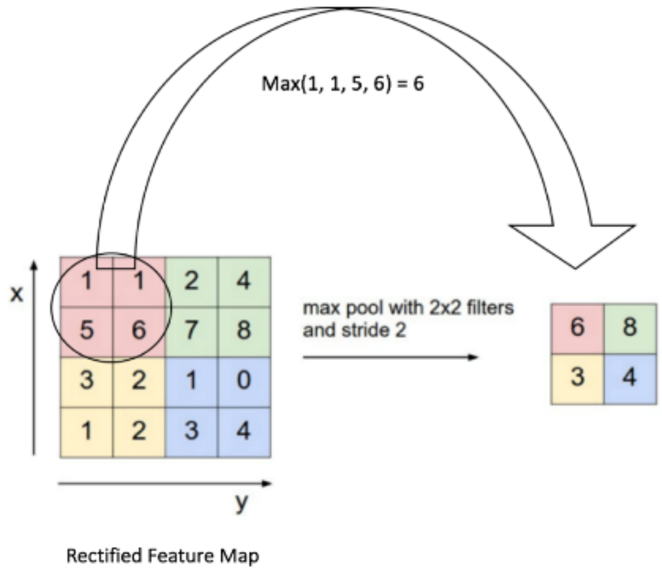
Pooling

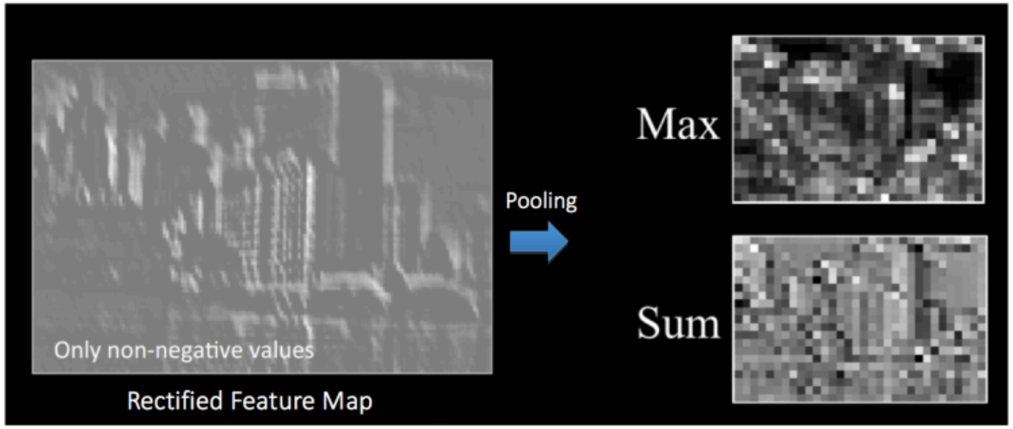
Aka subsampling or downsampling

Reduces the amount of information in a group op pixels by aggregating

e.g. min, max, avg..

Reduces no. parameters and makes the algorithm less susceptible to small variations





https://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets/

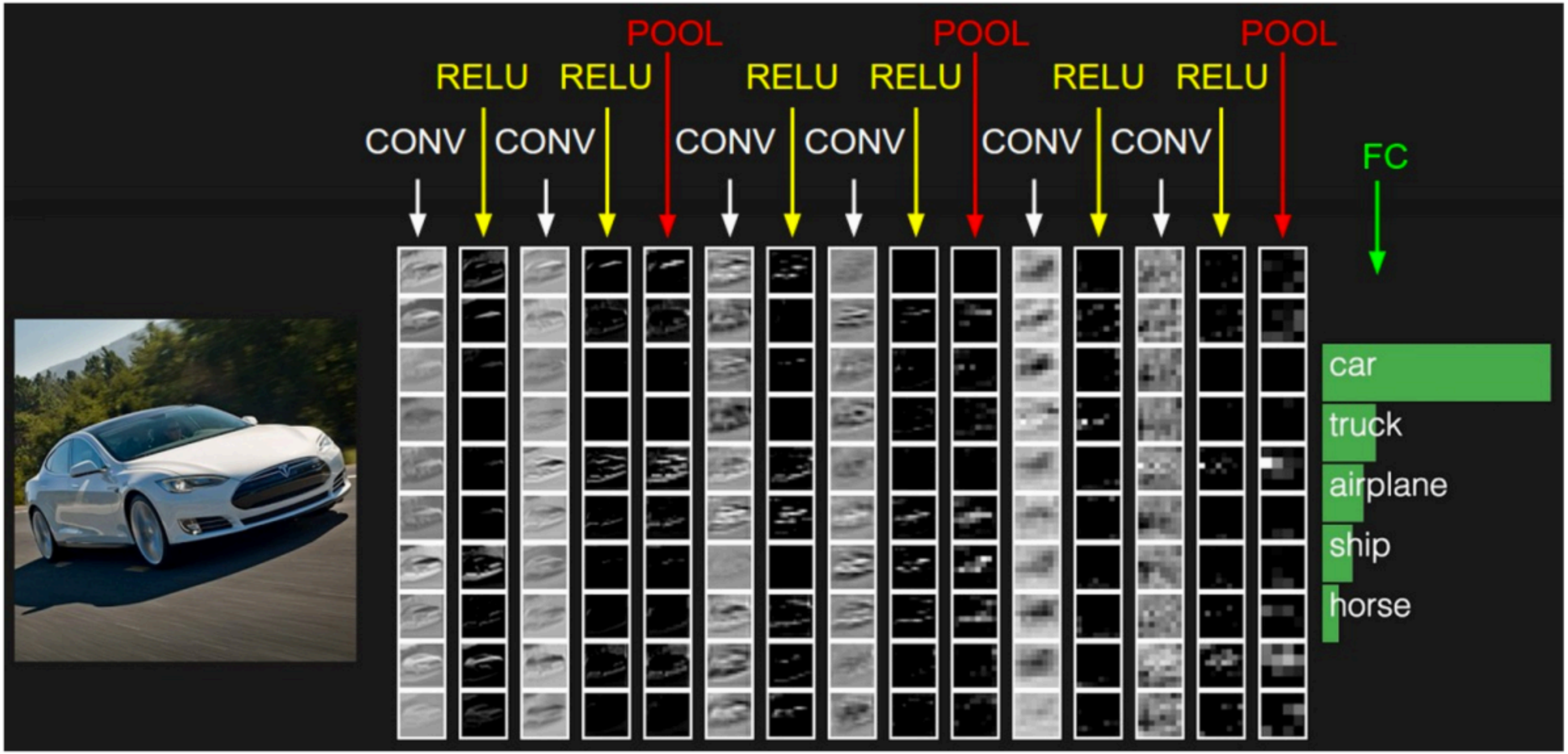
Classification with a fully connected layer

Previous steps identify *high-level features*

The *fully connected layer* uses these features to classify the image

A ‘normal’ neural network

‘Fully connected’ means that every node in one layer is connected to every node in the next one



https://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets/

The really amazing thing: the algorithm identifies the features that it needs to look for!!

With a cleverly configured neural network, computers can pick objects out of images

Data

10,000s images might not be sufficient

Google et al use millions to train their networks.

"In machine learning, having more data is almost always more important that having better algorithms."

This is why Google, Facebook, etc. now release lots of their algorithms publicly. The value is in the data, not the code.

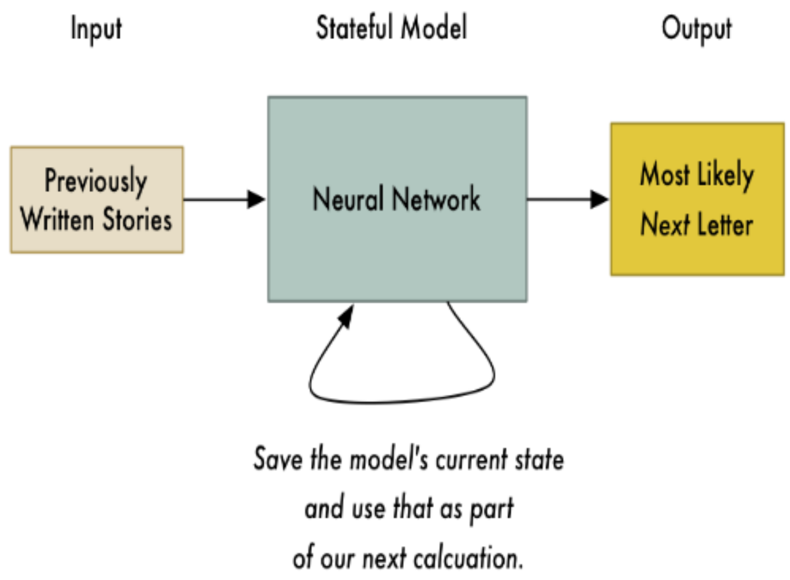
Another neat trick – adding memory

Recurrent neural networks keep track of their state

Output depends on current input, *and previous inputs*

Can be trained on published text and then write new text in the style of the author (kind of)

Possible application: predicting traffic?

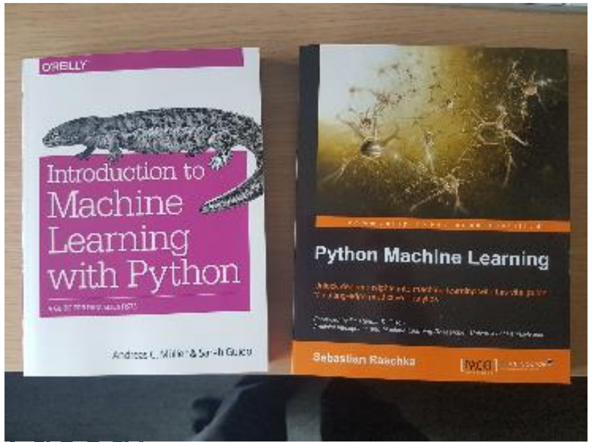


UNIVERSITY OF LEEDS

N8 POLICING RESEARCH PARTNERSHIP

Want more?

Read the blog on medium
(Search for 'Machine Learning is Fun')



Do Andrew Ng's Machine Learning course on Coursera *(it's really good)*
https://www.coursera.org/learn/machine-learning

I also took loads of material from this blog post:

https://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets/

There are lots of practical text books

UNIVERSITY OF LEEDS

N8 POLICING RESEARCH PARTNERSHIP

Image Recognition with Machine Learning - Conclusions

Image recognition is surprisingly easy!

It doesn’t need a complex algorithm to try to identify particular shapes etc.

By cleverly re-organising neural networks, computers are able to recognise complex patterns in data

BUT: as you’ll see, it can be hard to understand exactly what is happening at each step, and therefore hard to optimise an algorithm.

UNIVERSITY OF LEEDS

N8 POLICING RESEARCH PARTNERSHIP

Discussion:
Challenges and next steps to using these methods in practice

UNIVERSITY OF LEEDS

N8 POLICING RESEARCH PARTNERSHIP