University of Exeter

COM1011

Fundamentals of Machine Learning

Introduction

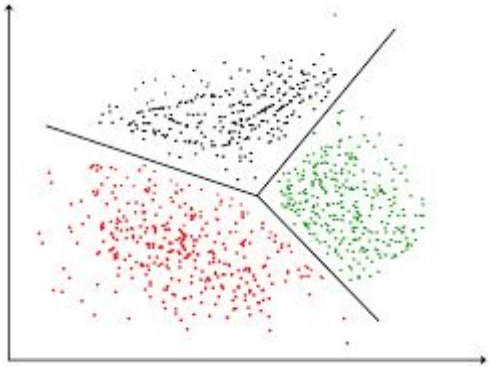
- What this course is about
- What Machine Learning is
- Types of data

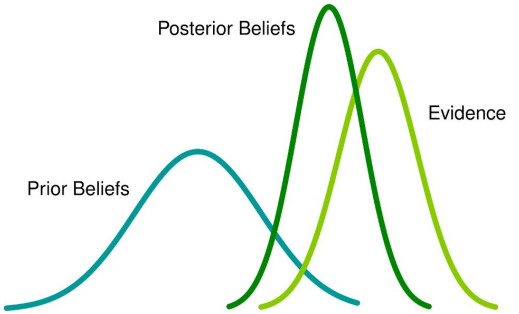
What this course is about

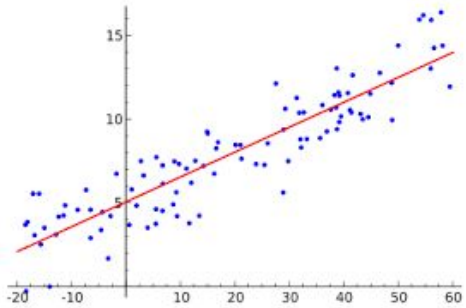
- Core concepts of Machine Learning (ML), Artificial Intelligence (AI) and Data Science (DS)
- How to make machines actually learn from data
- Strengths and weaknesses of different methods
- All at a conceptual level… mostly

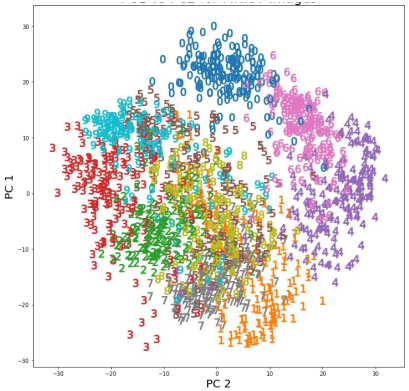
What this course is about

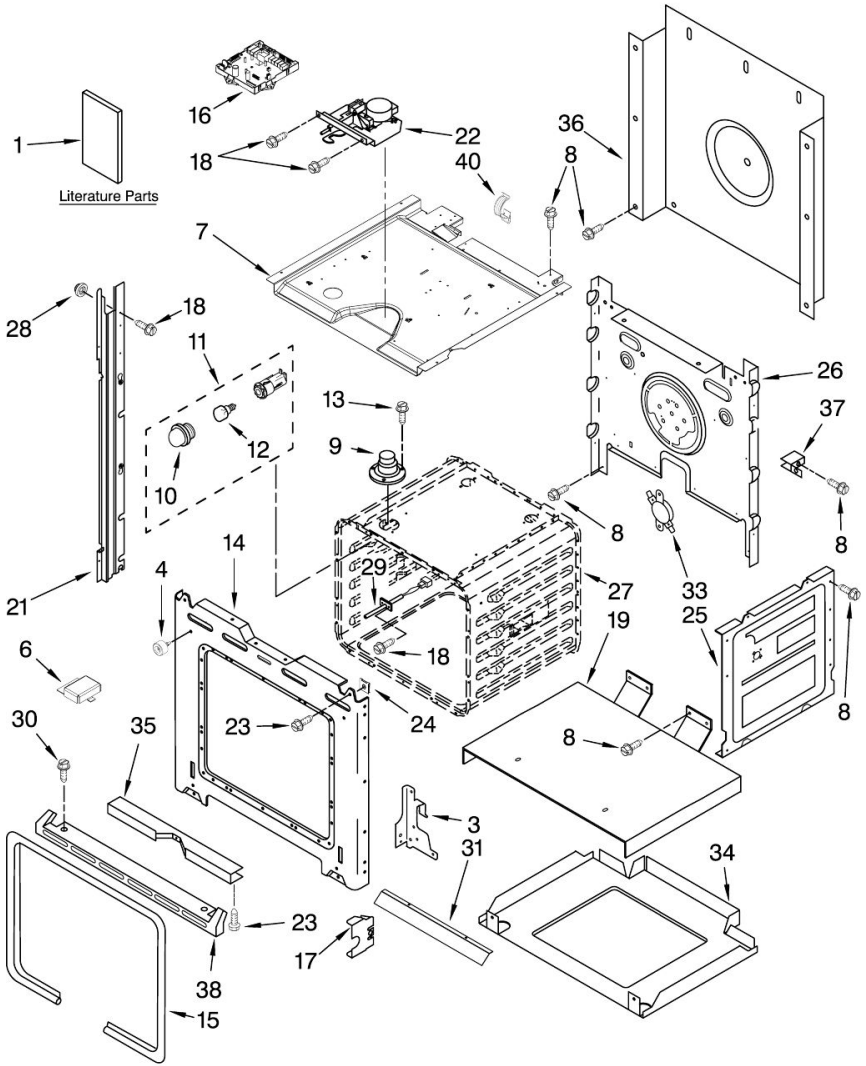
- Regression
- Classification
- Clustering
- Model selection
- Bayesian statistics
- Dimensionality reduction











Algorithm

A step-by-step procedure, specified to a level of detail such that any person (or machine) can follow.

This is basically an algorithm

1. Bring a large pan of water to the boil. Heat a grill pan. Toss the courgette strips with the oil and grill them for 2-3 minutes, turning often. Meanwhile, add the pine nuts to a small saucepan over a medium heat and cook for 2-3 minutes until toasted. Transfer to a plate immediately.

2. Cook the lasagne sheets for 3-4 minutes in the boiling water, then drain well. Carefully cut each lasagne sheet in half width ways so you have 12 pieces. Meanwhile, cook the spinach in 1 tsp water for 2-3 minutes until the leaves have wilted. Drain well, squeezing out the excess moisture with the back of a spoon.

3. Toss the cooked courgette strips in the pesto. Layer them with the lasagne sheets, spinach and dollops of the ricotta. Garnish with the basil and pine nuts.

This is basically an algorithm

1. Bring a large pan of wate
grill them for 2-3 minutes
over a medium heat and

2. Cook the lasagne sheets
each lasagne sheet in ha
1 tsp water for 2-3 minut
moisture with the back of

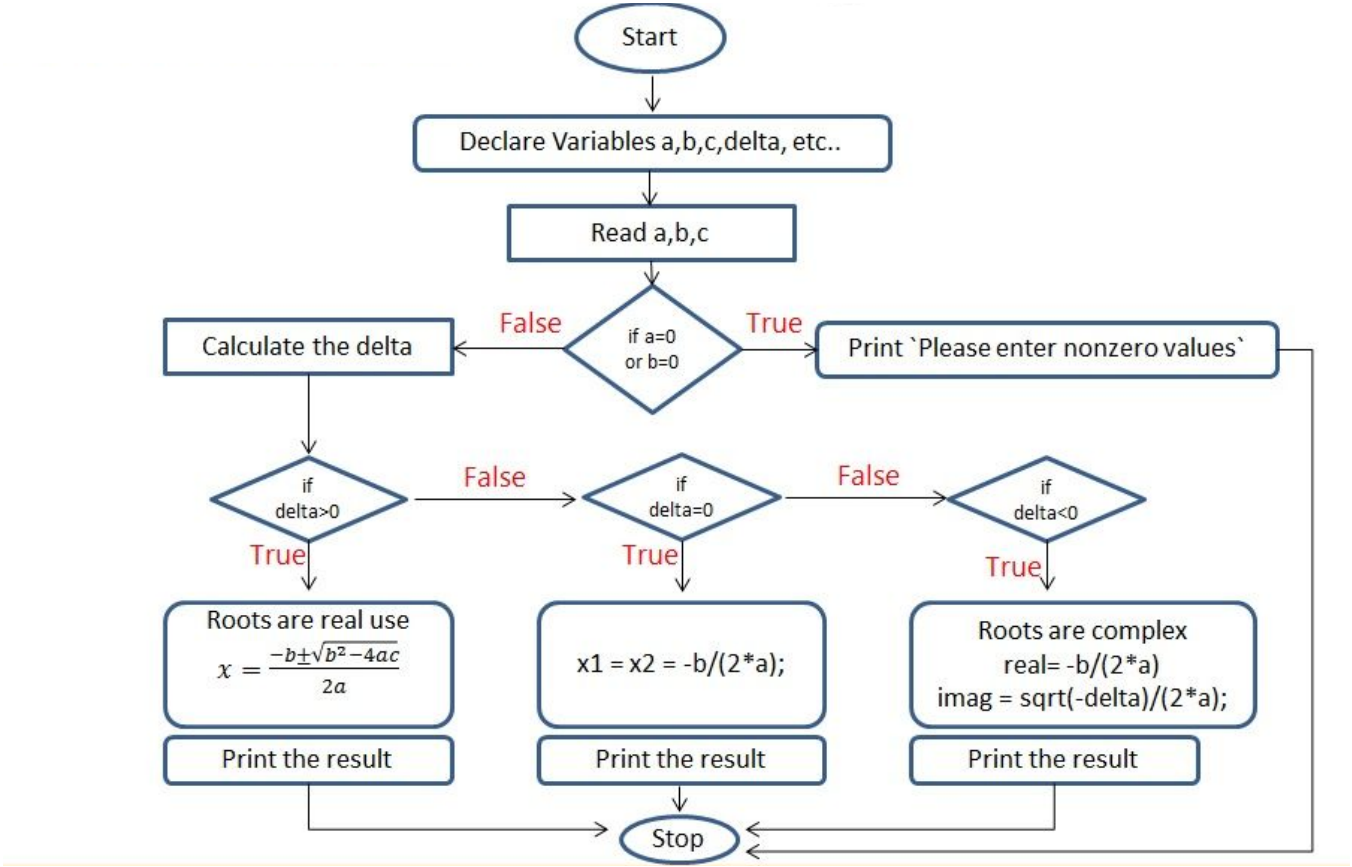
3. Toss the cooked courgette strips in the pesto. Layer them with the lasagne sheets, spinach and dollops of the ricotta. Garnish with the basil and pine nuts.

Algorithms

- Muhammad ibn Musa **al-Khwarizmi**
- Persian mathematician
- Introduced the concept of **algorithm**

  also considered the father of algebra

An algorithm for solving $ax^2+bx+c=0$



An algorithm for solving $ax^2+bx+c=0$

```
#include <iostream>
#include <math.h>
using namespace std;

int main()
{
    double a, b, c;
    double result;

    cout << "\t\tQuadratic Equation Program\n\n";

    cout << "Enter a number for 'a' that is not 0: ";
    cin >> a;

    cout << endl << endl << "Now enter a number for 'b': ";
    cin >> b;

    cout << endl << endl << "Finally, enter a number for 'c': ";
    cin >> c;

    result = (-b + sqrt(pow(b, 2) - 4*a*c))/(2*a);

    return 0;
}
```

Glossary

You can speak of machine learning algorithms that use a mathematics + statistics to find patterns about data.

When those patterns look pretty complicated, people call it artificial intelligence.

(if they’re simple we just don’t bother calling it AI)

Glossary

You can speak of machine learning algorithms that use a mathematics + statistics to find patterns about data.

When those patterns look pretty complicated, people call it artificial intelligence.

(

Glossary

- **Data science / Data-driven science**
  - a mix of mathematics, statistics, and computer science
    ...also often misused as a fancy name for statistics.

- **Machine learning**
  - mathematical foundations and practical applications of systems that learn with data.
    - Zoubin Ghahramani, University of Cambridge, Uber AI

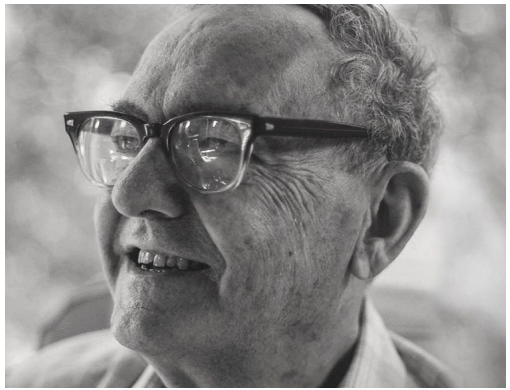
Machine Learning

- What is learning?
- If you had to explain what learning is to an alien (or to a robot), how would you explain it?

Machine Learning

“Learning denotes changes in the system that are adaptive in the sense that they enable the system to do the same task (or tasks drawn from a population of similar tasks) more effectively the next time.”

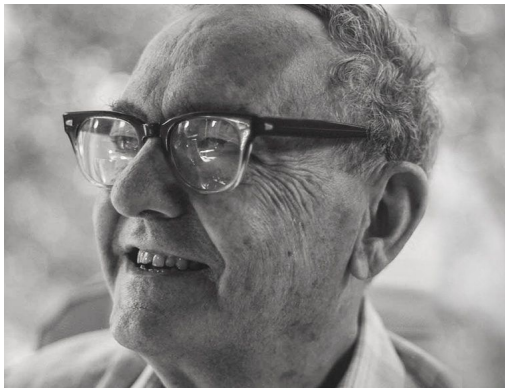
– Herbert Simon, Nobel Prize, Turing Award



Machine Learning

“Learning denotes changes in the system that are adaptive in the sense that they enable the system to do the same task (or tasks drawn from a population of similar tasks) more effectively the next time.”

– Herbert Simon, Nobel Prize, Turing Award

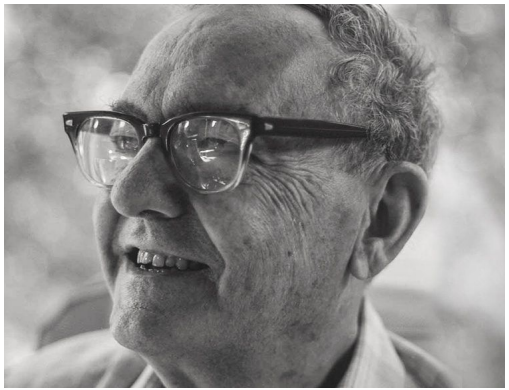


Machine Learning

Some kind of memory

“Learning denotes changes in the system
in the sense that they enable the system to do the same task
(or tasks drawn from a population of similar tasks)
more effectively the next time.”

– Herbert Simon, Nobel Prize, Turing Award



Machine Learning

"Learning denotes changes in the system

in the sense that they enable the system to do the same task

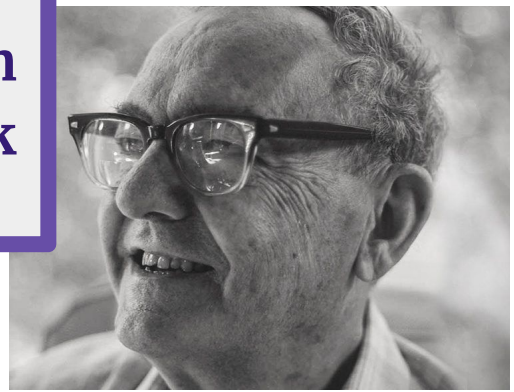
(or tasks drawn from a population of similar tasks)

more effectively the next tim

– Herbert Simon, Nobel Prize,

Some kind of memory

A generalisation of the same task



Machine Learning

"Learning denotes changes in the system

in the sense that they enable the system to do the same task

(or tasks drawn from a population of similar tasks)

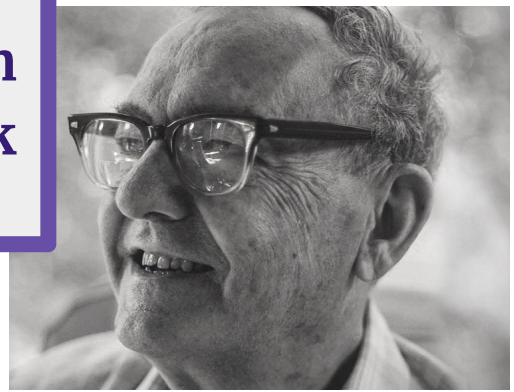
more effectively the next tim

Nobel Prize,

Some kind of memory

Improvement

A generalisation of the same task

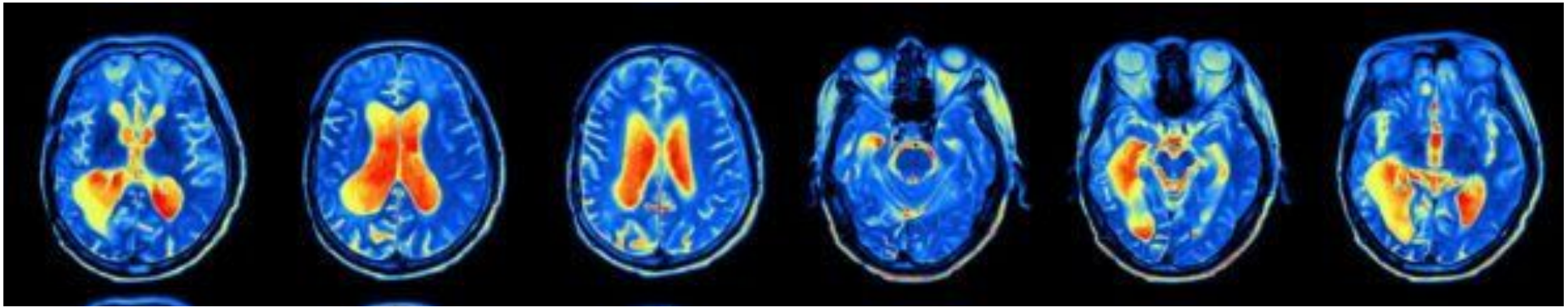


Examples:

- Classification
  - Medical diagnosis: given a brain scan, can the machine identify regions with abnormal behaviour?
- Planning and acting
  - Playing games, driving a car
- Prediction:
  - Predicting sales, income, customers, ad revenue...

Data science requires data.

- Classification
  - Medical diagnosis: given a brain scan, can the machine identify regions with abnormal behaviour?



Quantitative data

It is 180cm tall

Weighs 50 kg

Has 2 dictionaries

Has 3 shelves

Sells for £100

Qualitative data

It is made of wood

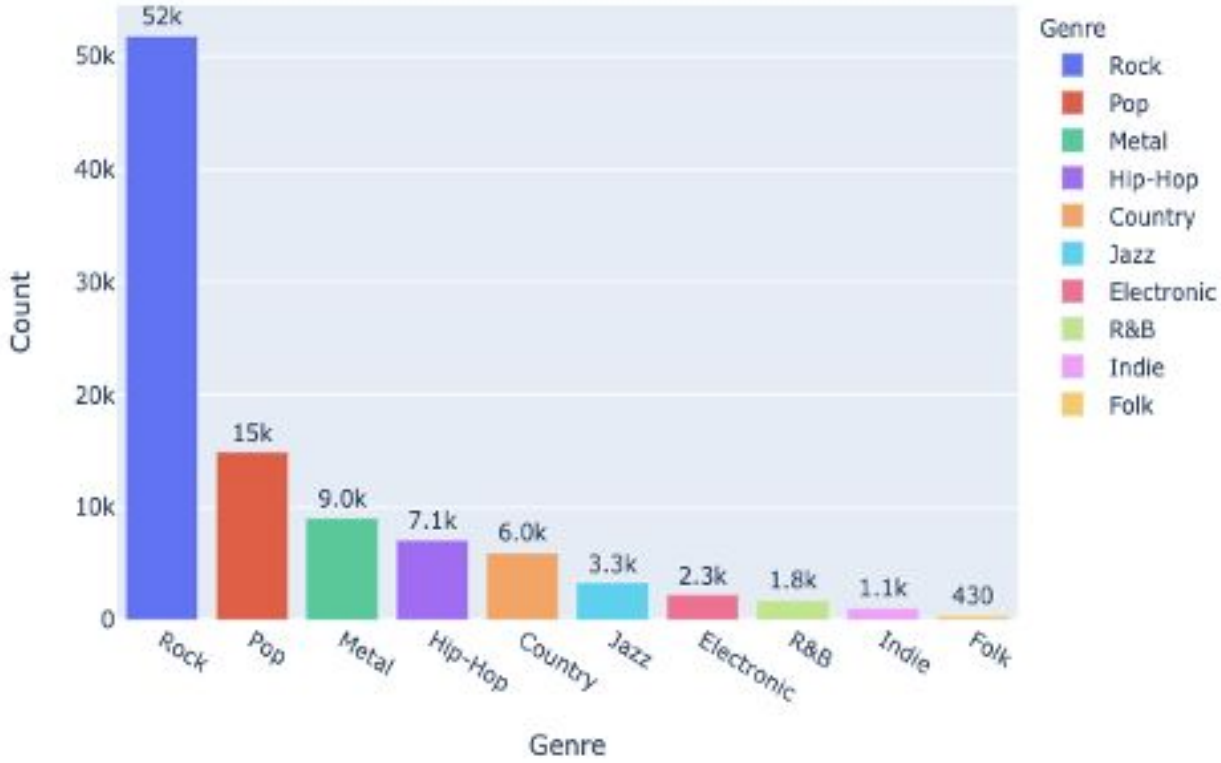
Made in italy

Brown

Smells like oak

Smooth finish

Categorical data



Ordinal data

	Very Good	Good	OK	Poor	Very Poor
Gym	☑	☐	☐	☐	☐
Pool	☐	☐	☑	☐	☐
Cafe	☐	☑	☐	☐	☐

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
I'm planning on buying this product again	○	○	○	○	✔
I'd recommend this product to someone else	○	○	○	○	✔

Numerical, continuous

Numerical, discrete

Mixed

			ype 1	Type 2	Total	HP	Attack	Defense	Sp. Atk	Sp. Def	Speed	Generation	Legendary
				Poison	318	45	49	49	65	65	45	1	False
				Poison	405	60	62	63	80	80	60	1	False
				Poison	525	80	82	83	100	100	80	1	False
				Poison	625	80	100	123	122	120	80	1	False
5	4	Charmander	Fire		309	39	52	43	60	50	65	1	False
6	5	Charmeleon	Fire		405	58	64	58	80	65	80	1	False
7	6	Charizard	Fire	Flying	534	78	84	78	109	85	100	1	False
8	6	CharizardMega Charizard X	Fire	Dragon	634	78	130	111	130	85	100	1	False
9	6	CharizardMega Charizard Y	Fire	Flying	634	78	104	78	159	115	100	1	False
10	7	Squirtle	Water		314	44	48	65	50	64	43	1	False
11	8	Wartortle	Water		405	59	63	80	65	80	58	1	False
12	9	Blastoise	Water		530	79	83	100	85	105	78	1	False
13	9	BlastoiseMega Blastoise	Water		630	79	103	120	135	115	78	1	False
14	10	Caterpie	Bug		195	45	30	35	20	20	45	1	False
15	11	Metapod	Bug		205	50	20	55	25	25	30	1	False
16	12	Butterfree	Bug	Flying	395	60	45	50	90	80	70	1	False
17	13	Weedle	Bug	Poison	195	40	35	30	20	20	50	1	False
18	14	Kakuna	Bug	Poison	205	45	25	50	25	25	35	1	False
19	15	Beedrill	Bug	Poison	395	65	90	40	45	80	75	1	False
20	15	BeedrillMega Beedrill	Bug	Poison	495	65	150	40	15	80	145	1	False
21	16	Pidgey	Normal	Flying	251	40	45				56	1	False
22	17	Pidgeotto	Normal	Flying	349	63	60				71	1	False
23	18	Pidgeot	Normal	Flying	479	83	80	75	70		101	1	False
24	18	PidgeotMega Pidgeot	Normal	Flying	579	83	80	80	135	80		1	False
25	19	Rattata	Normal		253	30	56	35					
26	20	Raticate	Normal		413	55	81	60					
27	21	Spearow	Normal	Flying	262	40	60						
28	22	Fearow	Normal	Flying	442	65	90						
29	23	Ekans	Poison		288	35	60						
30	24	Arbok	Poison		438	60	85						
31	25	Pikachu	Electric		320	35	55						
32	26	Raichu	Electric		485	60	90	55	90	80	110		

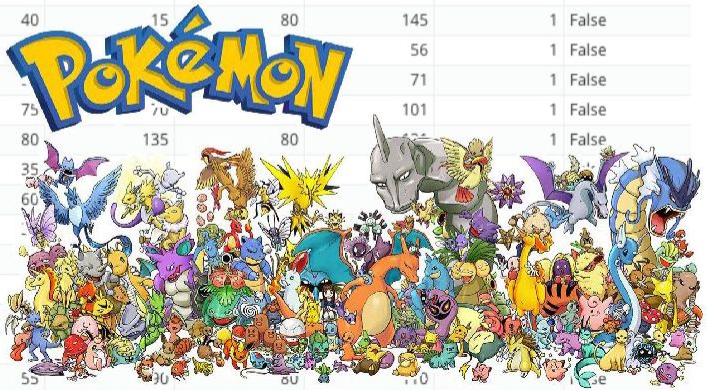
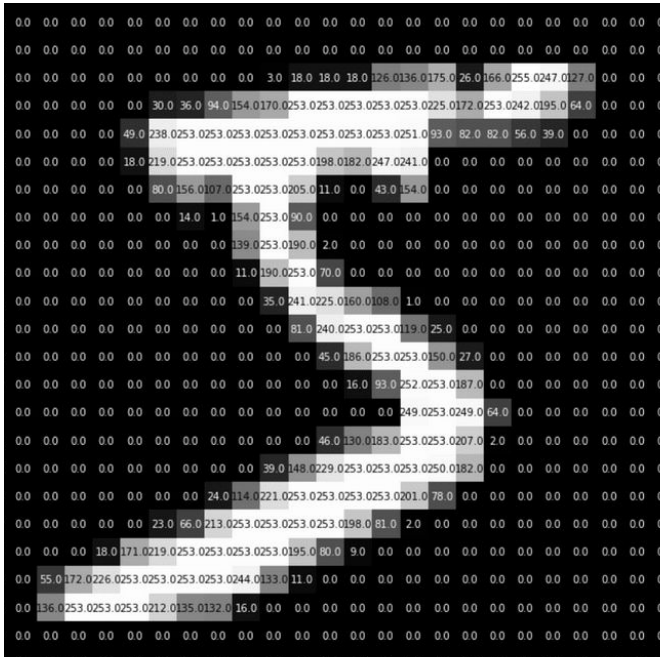


Image data

Images are usually represented as matrices.

A 400 x 400 pixels image can be represented by a 400 x 400 matrix.



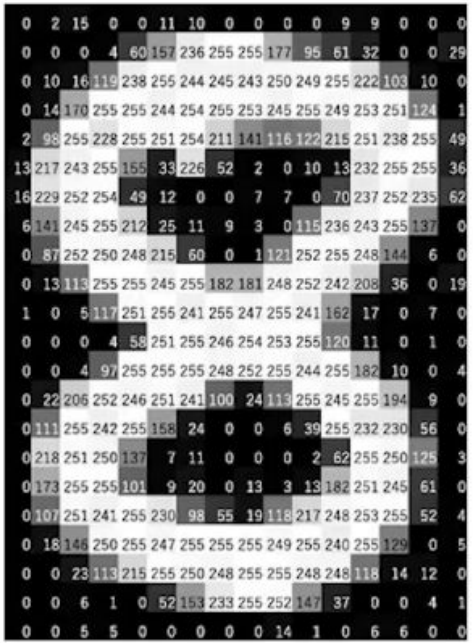
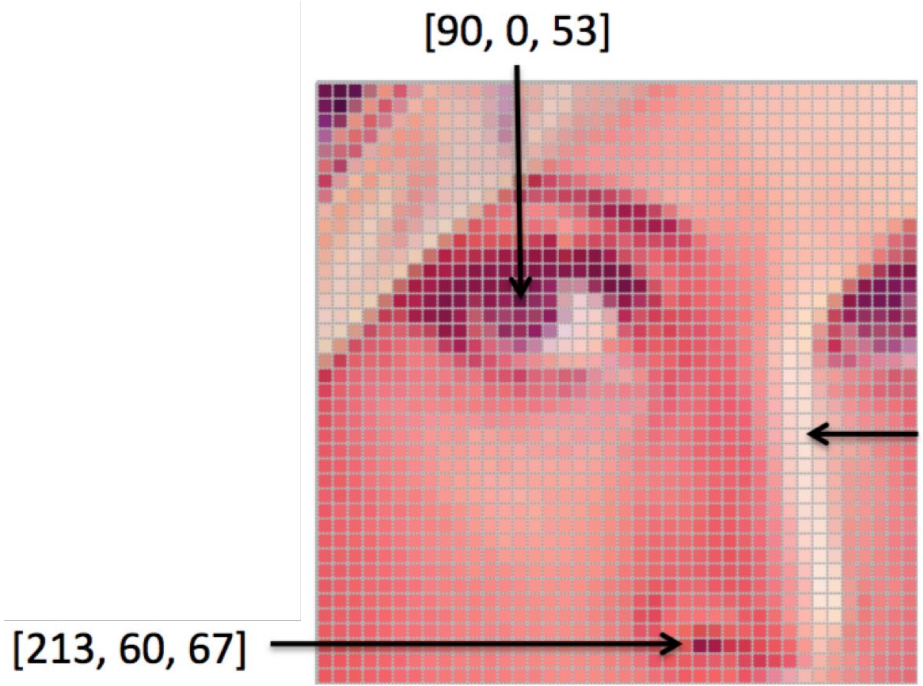


Image data

A 400 x 400 pixels image with colours can be represented by three 400 x 400 matrices, for the **red**, **green** and **blue** channels.



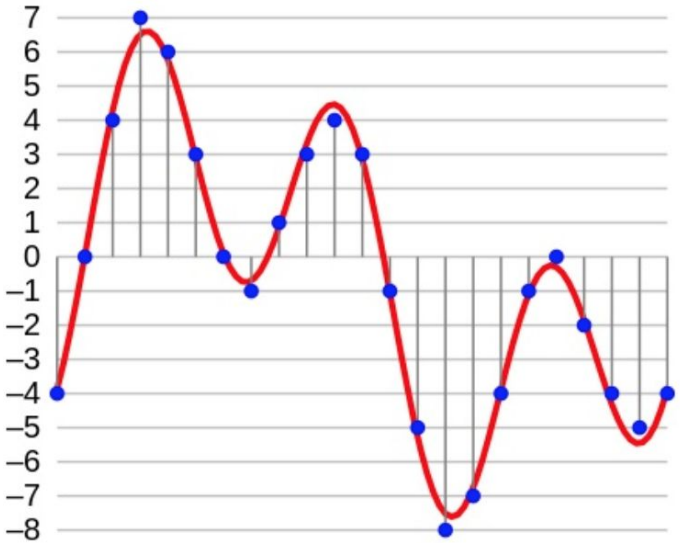
Text data

sales - Notepad

File Edit Format View Help

```
"Country","Salesperson","Order Amount","Quarter"
"UK","Smith",16753,"Qtr 3"
"USA","Johnson",14808,"Qtr 4"
"UK","Williams",10644,"Qtr 2"
"USA","Jones",1390,"Qtr 3"
"USA","Brown",4865,"Qtr 4"
"UK","Williams",12438,"Qtr 1"
"UK","Johnson",9339,"Qtr 2"
"USA","Smith",18919,"Qtr 3"
"USA","Jones",9213,"Qtr 4"
"UK","Jones",7433,"Qtr 1"
"USA","Brown",3255,"Qtr 2"
"USA","Williams",14867,"Qtr 3"
"UK","Williams",19302,"Qtr 4"
"USA","Smith",9698,"Qtr 1"
"USA","Jones",18978,"Qtr 2"
"UK","Brown",9080,"Qtr 4"
```

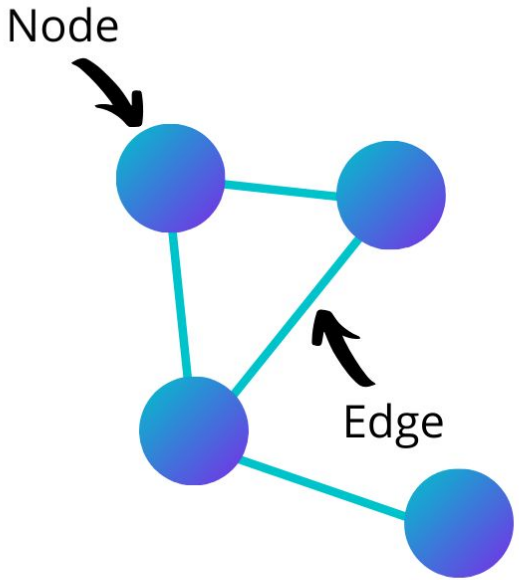
Sound data



Sound wave

Network data, or relational data

```
[
    node1 : node2
    node2 : node3
    node1 : node3
    node4 : node1
]
```



Most of the time: mixed data

```
      age                    job          marital       education   default
 Min.   :23.00   management    :454   married :431   primary  : 39   no :642
 1st Qu.:33.00   technician    : 45   single  :139   secondary: 26   yes:  9
 Median :37.00   entrepreneur : 44   divorced: 81   tertiary :572
 Mean   :39.41   self-employed: 34                  unknown  : 14
 3rd Qu.:44.50   admin.       : 24
 Max.   :75.00   unemployed   : 14
                 (Other)      : 36
```

Next steps

- Next lecture: Linear regressions, Loss functions and R^2

Additional material:

- Tom M. Mitchell, The Discipline of Machine Learning
- Chp. 1, Kevin P. Murphy, M Cambridge Press 2012.
- Chp. 1, Andrew R. Webb, Statistical Pattern Recognition, 2nd edition, Wiley 2002.
- Chp. 1, Dutta et al, Pattern Classification 2nd edition, Wiley 2000
- Chp. 1, Nilsson et al, Introduction to Machine Learning, 1998
- More on ELE