IMPERIAL

Introduction to Machine Learning

Payam Barnaghi
Department of Brain Sciences &
School of Convergence Science (Human and Artificial Intelligence)
Imperial College London
January 2026

ML4NS
Machine Learning for Neuroscience

1

IMPERIAL

Leia: *Yaté. Yaté. Yotó.*
(SUBTITLE: "I have come for the bounty on this Wookiee.")

C-3PO relays this message and Jabba says he'll offer 25,000 for Chewie.

Leia: *Yotó. Yotó.* (SUBTITLE: "50,000, no less.")
C-3PO relays this message and Jabba asks why he should pay so much.

Leia: *Eí yóto.*
The above isn't subtitled, but Leia pulls out a bomb and activates it.

C-3PO **:** Because he's holding a thermal detonator!
Jabba is impressed by this and offers 35,000.

Leia: *Yató cha.*
The above isn't subtitled, but Leia deactivates the bomb and puts it away.

C-3PO **:** He agrees.
Order is restored.

Adapted from: The Art of Language Invention: From Horse-Lords to Dark Elves, The Words Behind World-Building, by David J. Peterson, Penguin, 2015.

2

2

IMPERIAL

Leia: *Yaté. Yaté. Yotó.*
(SUBTITLE: "I have come for the bounty on this Wookiee.")

C-3PO relays this message and Jabba says he'll offer 25,000 for Chewie.

Leia: *Yotó. Yotó.* (SUBTITLE: "50,000, no less.")
C-3PO relays this message and Jabba asks why he should pay so much.

Leia: *Eí yóto.*
The above isn't subtitled, but Leia pulls out a bomb and activates it.

C-3PO **:** Because he's holding a thermal detonator!
Jabba is impressed by this and offers 35,000.

Leia: *Yató cha.*
The above isn't subtitled, but Leia deactivates the bomb and puts it away.

C-3PO **:** He agrees.
Order is restored.

Adapted from: The Art of Language Invention: From Horse-Lords to Dark Elves, The Words Behind World-Building, by David J. Peterson, Penguin, 2015.

3

3

IMPERIAL

$$\forall \rightarrow \textbf{Yotó}:$$

$$if : f([\textbf{Yotó}]) == 1 \rightarrow \textbf{Yotó} \propto "thisWookiee"$$
$$if : f([\textbf{Yotó}]) == 2 \rightarrow \textbf{Yotó} \propto "50,000"$$
$$if : f([\textbf{Yotó}]) == 3 \rightarrow \textbf{Yotó} \propto "noless"$$

$$if : (f([\textbf{Yotó}]) == 3 \,\&\, SeasonOnEarth() == "Spring") \rightarrow \textbf{Yotó} \propto "Welcome!"$$

...

...

4

4

IMPERIAL

However, don't underestimate the human brain

- For example, in German:
  - you must distinguish between "the" and "a" articles
  - and then each one has four case forms
  - three genders
  - and singular and plural forms
  - and then the adjectives have to agree
  - and then there are verbs!
  - or take this as an example:

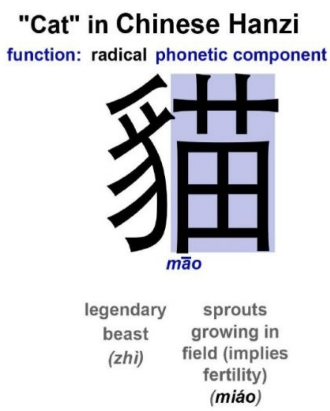
(biáng)

Adapted from: The Art of Language Invention: *From Horse-Lords to Dark Elves, The Words Behind World-Building, by David J. Peterson.*

5

5

IMPERIAL

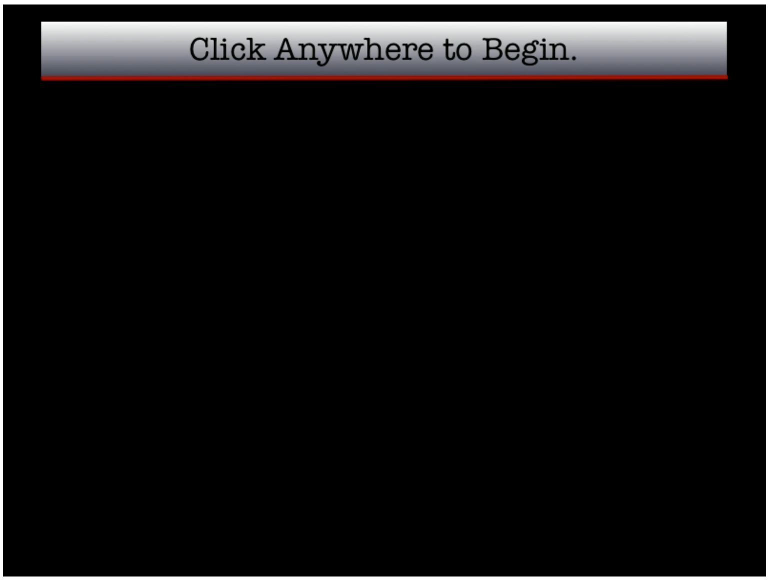


Source: https://www.originofalphabet.com/meow-is-just-another-name-for-cat-2/

6

6

IMPERIAL



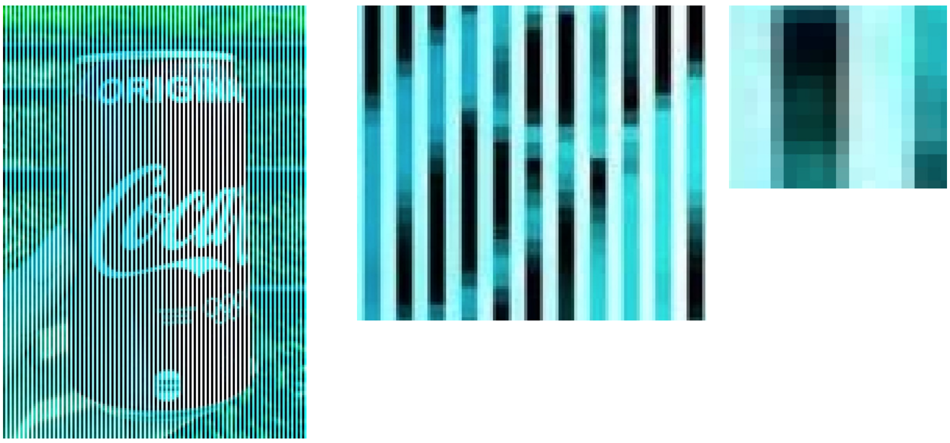
Source: https://cdn.sinauer.com/wolfe4e/wa07.05.html#Movie1

7

7

IMPERIAL

Patterns and generalisation



8

8

Attention

IMPERIAL

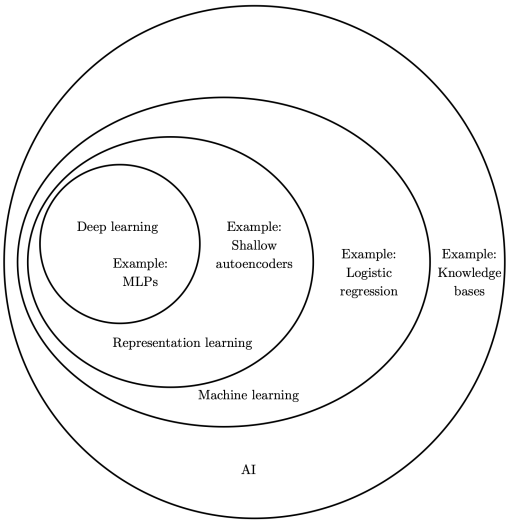
Source: https://www.youtube.com/watch?v=xNSgmm9FX2s
Via: https://www.youtube.com/watch?v=48gBPL7aHJY&t=0s

9

9

AI and Machine Learning

IMPERIAL



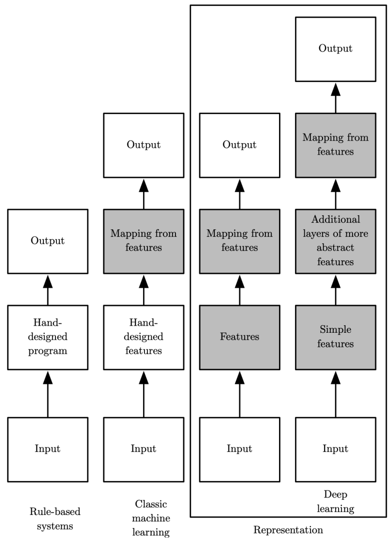
Source: Deep Learning, Ian Goodfellow *et al.*, MIT Press.

10

10

Categories of learning and intelligent systems

IMPERIAL



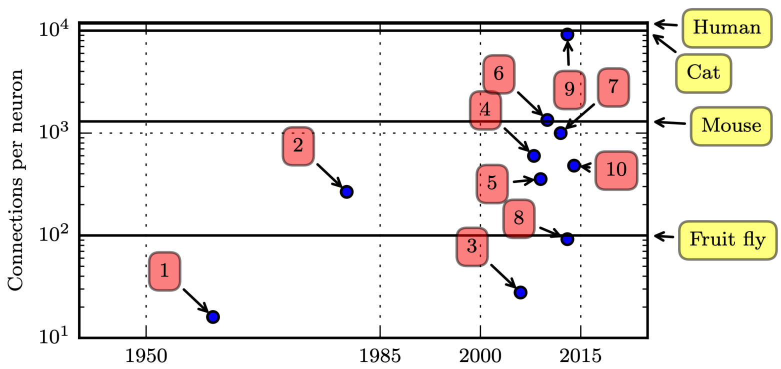
Source: Deep Learning, Ian Goodfellow *et al.*, MIT Press.

11

11

Connections per neuron

IMPERIAL



1. Adaptive linear element (Widrow and Hoff, 1960)

…

8. Multi-GPU convolutional network (Krizhevsky *et al.*, 2012)
9. COTS HPC unsupervised convolutional network (Coates *et al.*, 2013)

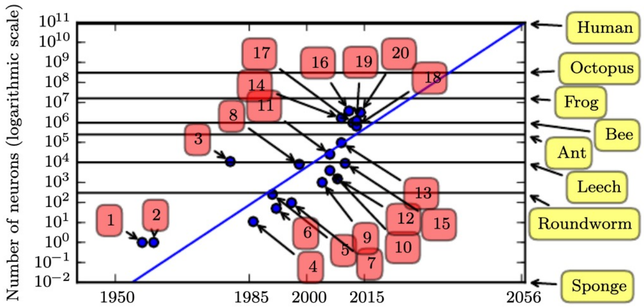
Source: Deep Learning, Ian Goodfellow *et al.*, MIT Press.

12

12

Number of neurons

IMPERIAL



17. Distributed autoencoder (Le *et al.*, 2012)
18. Multi-GPU convolutional network (Krizhevsky *et al.*, 2012)
19. COTS HPC unsupervised convolutional network (Coates *et al.*, 2013)
20. GoogLeNet (Szegedy et al., 2014a)

- The human brain has approximately 86 billion neurons
- The GPT-4 model has approximately 1.8 trillion parameters.
- However, the artificial neurons are not equal to biological neurons

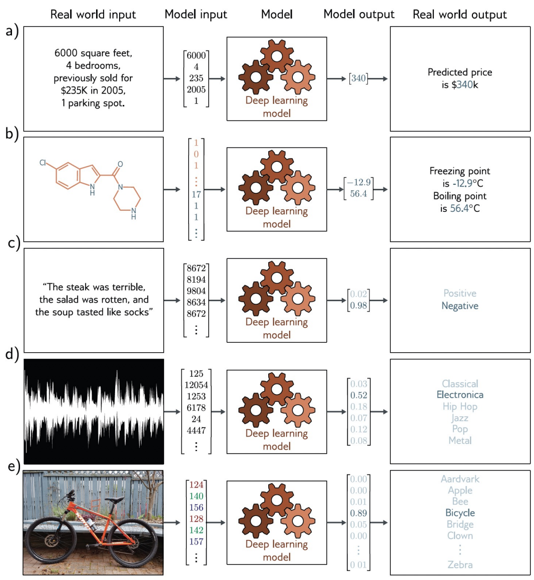
Source: Deep Learning, Ian Goodfellow *et al.*, MIT Press.

13

13

Machine learning models

IMPERIAL



Source: Understanding Deep Learning, Simon J.D. Prince, MIT Press.

14

14

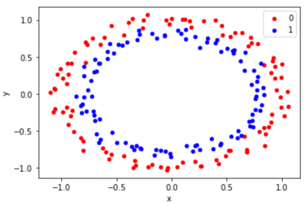
Machine Learning (ML)

IMPERIAL

– Most of machine learning methods are designed as a learning process to learn a function `f*` for which:

$$f^*(x) \approx f^*(x + \varepsilon)$$

- In other words, this means if we know the answer for a training sample x, then that answer is *probably* good in the neighbourhood of x.

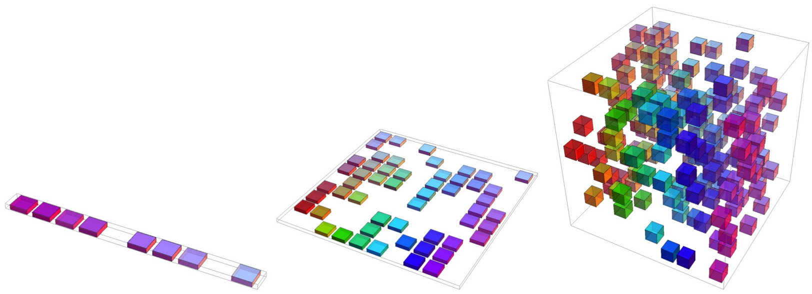


15

15

Curse of dimensionality

IMPERIAL



Overall, for d dimensions with v values we need $O(v^d)$ regions and examples

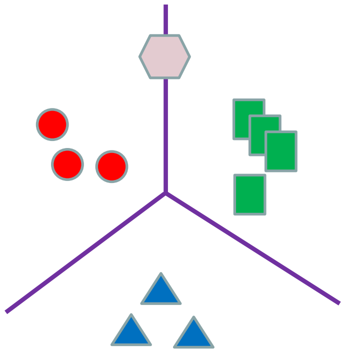
Source: Deep Learning, Ian Goodfellow *et al.*, MIT Press.

16

16

K-nearest neighborhood

IMPERIAL

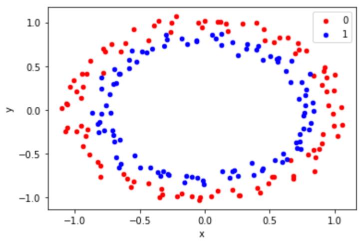


17

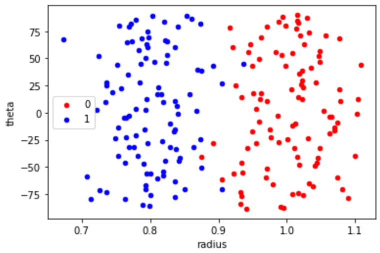
17

Data/sample representation

IMPERIAL



(x,y)



$$r = \sqrt{x^2 + y^2}$$
$$\theta = \tan^{-1}\left(\frac{y}{x}\right).$$

<Code>

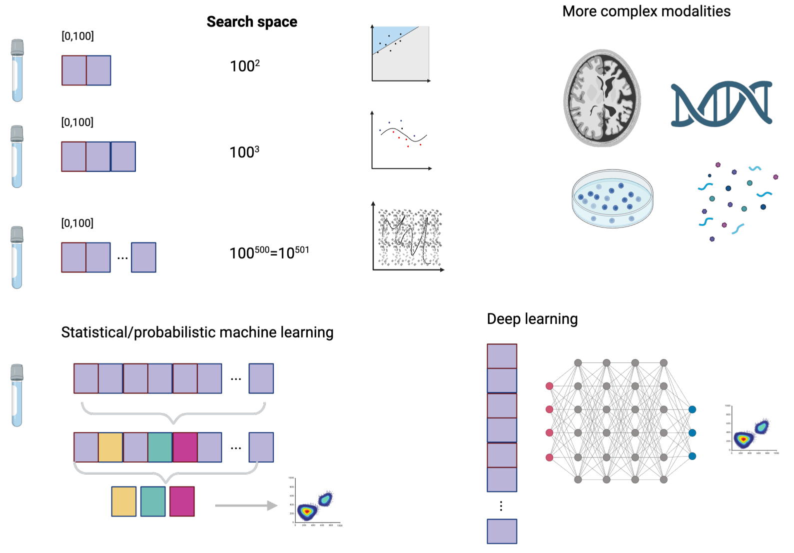
https://github.com/PBarnaghi/Basic-ML-Code

18

18

Machine Learning: looking back, moving forward

IMPERIAL



19

19

What you are going to learn in this module

IMPERIAL

- Core Concepts
  - Principles of machine learning models
  - Designing ML models for neuroscience and medical applications
- Hands-On Skills
  - Select suitable models for specific tasks
  - Design machine learning experiments
  - Apply appropriate metrics to evaluate models
  - Detect and address common issues (e.g., overfitting, underfitting)
  - Evaluate potential bias and fairness in models and their outcomes

20

20

By the end of this module

IMPERIAL

- You should be able to:
  - Construct machine learning experiments
  - Create and run different probabilistic/statistical learning and deep learning models
  - Read and interpret existing works and assess their applicability/validity [to different use cases]
  - Be familiar with common concepts such as training/test, cross-validation, and performance evaluation metrics
  - Learn about common issues and pitfalls in developing ML models
  - Obtain practical experience in developing ML models and running (simple) experiments.

21

21

Key topics covered

IMPERIAL

- Introduction to machine learning
- Regression models and linear prediction
- Probability and information theory
- Bayesian models and maximum likelihood estimation
- Linear models, ensemble models and kernel functions
- Artificial Neural Networks (ANNs) and Multilayer Perceptron (MLP) models
- Deep Neural Networks (DNN) and Deep Learning (DL)
- Convolutional Neural Networks (CNNs)
- Responsible machine learning and ethical considerations
- Applications in neuroscience and neuroscience-inspired machine learning
- Ethical AI/ML: principles and practices

22

22

Module delivery

IMPERIAL

- 10 teaching sessions
  - Primary tutorials and workshops (primary programming and data handling/visualisation skills)
  - 2 hours lectures (Mon-Fri; 10:30-12:30) for two weeks in January 2026.
  - 3 hours lab experiments (supported by TAs)
  - One final project (modelling and machine learning experiment using a real-world dataset)

23

23

Datasets

IMPERIAL

- Syntenic data to demonstrate the concepts
- In-home monitoring data in a dementia cohort study from the UK DRI Minder/TIHM Study (movement, sleep, physiology, and associated healthcare events), in the form of numerical time-series data (Final Project)
- Common benchmark imaging data (non-medical) such as CIFAR-10
- MRI data to build a CNN model
- Publicly available clinical/experimental datasets
  - e.g. Physionet.org

24

24

Tools and languages

IMPERIAL

– *Python* and common machine learning libraries in *Python* including *Scikit – learn* and *PyTorch*

Development/learning framework

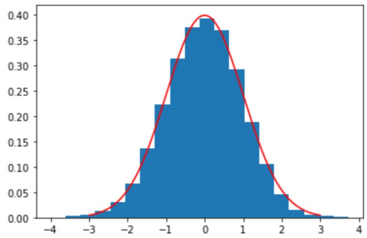
IMPERIAL

```
In [1]:
import numpy as np
import matplotlib.pyplot as plt
import matplotlib
matplotlib.use('TkAgg')

%matplotlib inline
```

Number of sample

```
In [2]:
n= 50

In [4]:
n +5

Out[4]: 55
```

```
In [9]:
mu, sigma = 0,1
dist = gaussian_mcmc(100_000,mu=mu,sigma=sigma)
lines = np.linspace(-3,3,10_000)
normal_curve = [normal(x=l,mu=mu,sigma=sigma) for l in lines]

In [11]:
plt.hist(dist,bins=20,density=True)
plt.plot(lines,normal_curve, color='red')

Out[11]: [<matplotlib.lines.Line2D at 0x7fb7e1816160>]
```



Teaching team

IMPERIAL


Nan Fletcher-Lloyd


Anastasia Gailly De Taurines


Antigone Fogel


Iona Biggart


Payam Barnaghi

27

27

Our Aim

IMPERIAL

- We aim to encourage neuroscientists and clinical scientists to enjoy developing robust and reliable AI/ML models for practical applications. We hope this will help to prosper the practice of medicine and discovery science.

- We hope this module will remove some of the fears naturally felt when starting a career in computational neuroscience and computational modelling in medicine and bioscience, thereby allowing the practitioners to focus on solving unmet health and care needs and advancing knowledge and practice of neuroscience and medicine.

The text is inspired by the preface of the Oxford Handbook of Clinical Medicine by I. Wilkinson *et al.*

28

28

”Patients” and “Patient data”

IMPERIAL

- The terms "patient" and "patient data" will appear throughout this module in the slides and presentations. Please note that we do not take this lightly. Every single piece of data displayed or used from anonymised patient information throughout this module is donated to science by individuals.
- We are grateful for their contributions, which support scientific discovery to enhance existing procedures and to design and develop new treatments and practices.
- Privacy and the ethical use of data and modelling are essential elements of any AI/ML design.

Inspired by the Oxford Handbook of Clinical Medicine by I. Wilkinson *et al.* and a talk given by Junaid Bajwa, Flagship Pioneering.

29

29

Frequently Asked Questions (I)

IMPERIAL

- I don’t know Python. Can I take this module?
  - Yes, we have provided an introductory lecture/workshop on Python (at the level required for this module).
- I haven’t done any programming before. Can I still take this module?
  - Yes, but you should be prepared to learn scripting/programming and get familiar with the concepts quickly. We will provide examples/templates to follow, and there is plenty of online material, but there will be a steep learning curve.

30

30

Frequently Asked Questions (II)

IMPERIAL

- Machine learning seems to involve a lot of mathematics, especially statistics and probability. Do I need to know all these before taking this module?
  - You need to be familiar with the basic concepts of statistics and probability, but we have provided a brief refresher online tutorial.
  - The lecture content/concepts will not be mathematically focused, but when needed, the underlying mathematical concepts will be described/shown to elaborate on the concepts.
  - Learning mathematical concepts in a group is often easier.
  - We will focus on concepts and skills that won't be easy to find or learn from a series of online videos.

31

31

Concepts that will not be covered in this module

IMPERIAL

- Image processing and machine learning for medical imaging
- Natural language processing and text analysis
- Advanced machine learning and deep learning concepts such as transfer learning, generative models (some of these will be covered in an optional series after the module)
- Signal processing and feature engineering for specific modalities/data such as EEG, ECG, and Brain imaging

32

32

What will I learn in this module?

IMPERIAL

- Primary concepts and foundations of machine learning (statistical/probabilistic learning and deep learning models).
- Some of the basic/key models to design and develop supervised and unsupervised machine learning.
- Cross-validation, hyperparameter tuning, and evaluation metrics
- Hands-on skills and working with real-world data, and designing/evaluating models with different real-world datasets.

- Ultimately, this module (hopefully) will help you learn fundamental concepts on how ML models are constructed and applied to real-world problems.

33

33

Module assessment

IMPERIAL

- 2 Marked Lab assignments (50%)
  - Lab 5 (Friday)
    - This assessment will consist of pen-and-paper multiple-choice questions. You may prepare and bring one A4 double-sided cheat sheet.
  - Lab 8 (Wednesday)
    - This assessment will consist of a series of questions on a use case to be reviewed/evaluated.
- Final project (50%)
  - In the 2nd week, we will give you a dataset and ask you to:
    - Conduct a set of exploratory data analyses
    - Develop a simple ML model and evaluate it
    - Code with markdown in Jupyter Notebook will be sufficient.

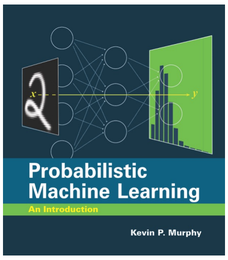
- **Assessments will be on Blackboard**.

34

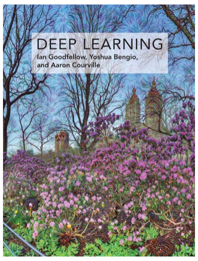
34

Reference and sources

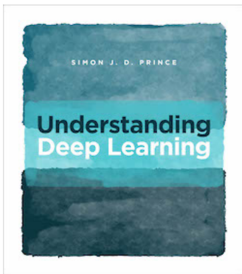
IMPERIAL



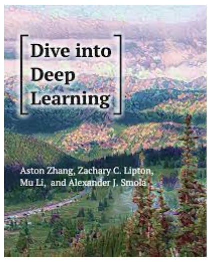
Probabilistic Machine Learning
Kevin P. Murphy, MIT Press



Deep Learning
Goodfellow et al., MIT Press



Understanding Deep Learning,
Simon J.D. Prince, MIT Press



Dive into Deep Learning
Aston Zhang et al.

Links to Preprints (PDFs/online versions):
- Probabilistic Machine Learning: An Introduction, https://probml.github.io/pml-book/book1.html
- Deep learning: https://www.deeplearningbook.org
- Understanding Deep Learning, https://udlbook.github.io/udlbook/
- Dive into deep learning, https://d2l.ai/index.html

35

35

Some other online sources

IMPERIAL

- Andrew Ng lecture notes:
  https://cs229.stanford.edu/lectures-spring2022/main_notes.pdf
- https://towardsdatascience.com/machine-learning/
- Good online sources for questions/troubleshooting
  - https://stackoverflow.com
  - https://stackexchange.com

36

36

MS Teams Link

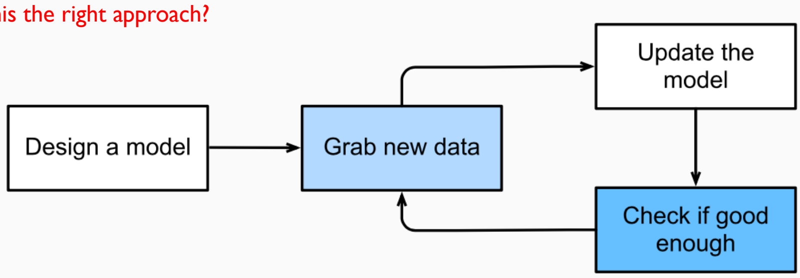
Typical machine learning approaches

- Supervised learning
- Unsupervised learning
- Reinforcement learning

Machine learning development

IMPERIAL

Is this the right approach?



Source: Dive into Deep Learning, Aston Zhang *et al.*

Key components of ML development

IMPERIAL

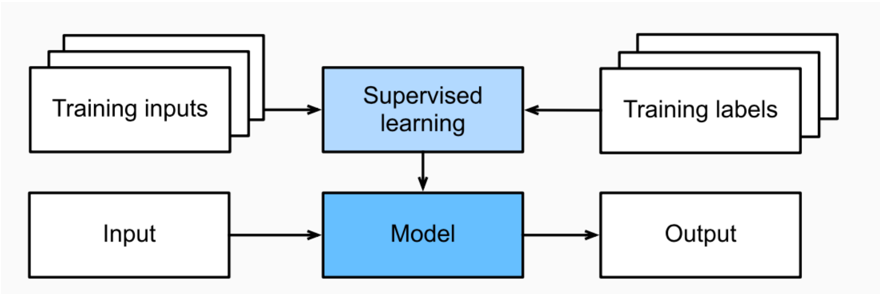
- The *data* that we can learn from.
- A *model* of how to transform the data.
- An *objective function* that quantifies how well (or badly) the model is doing.
- An *algorithm* to adjust the model's parameters to optimise the objective function.

Source: Dive into Deep Learning, Aston Zhang *et al.*

Supervised learning

IMPERIAL

– In supervised learning the data comes with additional attributes (i.e. labels) that we want to predict.



Source: Dive into Deep Learning, Aston Zhang *et al.*

41

41

Classification

IMPERIAL

– Samples belong to two or more classes, and we want to learn from already labelled data how to predict the class of unlabelled data.
– An example of a classification problem is activity recognition, in which the aim is to assign each input vector to one of a finite number of discrete categories.
– Another way to think of classification is as a discrete (as opposed to continuous) form of supervised learning where one has a limited number of categories, and for each of the n samples provided, one is to try to label them with the correct category or class.

Source: Scikit learn documentation

42

42

Regression

IMPERIAL

- The task is called regression if the desired output consists of one or more continuous variables.
- An example of regression problems:
  - How long will it take for an Alzheimer's Disease AD patient to decline to an MMSE* score of x? (judging by previously seen data)
  - How long will it take for a patient to decline from one stage to another stage (in a neurodegenerative condition)?
  - Predicting the length of hospital stay for a patient

* MMSE: The Mini-Mental State Examination (MMSE)

43

43

Unsupervised learning

IMPERIAL

- In unsupervised learning, the training data consists of a set of input vectors x without any corresponding target values.
- The goal in such problems may be to discover groups of similar examples within the data, where it is called clustering, or to determine the distribution of data within the input space, known as density estimation, or to project the data from a high-dimensional space down to two or three dimensions for the purpose of visualisation.

Source: Dive into Deep Learning, Aston Zhang *et al.*

44

44

Training set and testing set

IMPERIAL

- Machine learning is about learning some properties of a data set and applying them to new data.

- This is why a common practice in machine learning to evaluate an algorithm is to split the data at hand into two sets, one that we call the training set, on which we learn data properties and one that we call the testing set, on which we test these properties.

Source: Dive into Deep Learning, Aston Zhang *et al.*

45

45

Dataset

IMPERIAL

7.1.3. Diabetes dataset

Ten baseline variables, age, sex, body mass index, average blood pressure, and six blood serum measurements were obtained for each of n = 442 diabetes patients, as well as the response of interest, a quantitative measure of disease progression one year after baseline.

Data Set Characteristics:

Number of Instances::	442
Number of Attributes::	First 10 columns are numeric predictive values
Target::	Column 11 is a quantitative measure of disease progression one year after baseline
Attribute Information::	• age age in years • sex • bmi body mass index • bp average blood pressure • s1 tc, total serum cholesterol • s2 ldl, low-density lipoproteins • s3 hdl, high-density lipoproteins • s4 tch, total cholesterol / HDL • s5 ltg, possibly log of serum triglycerides level • s6 glu, blood sugar level

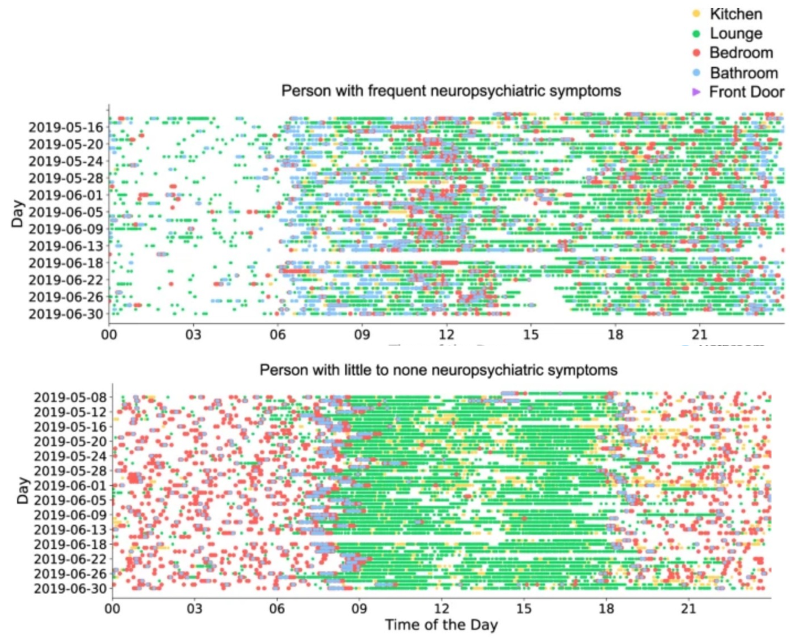
https://www4.stat.ncsu.edu/~boos/var.select/diabetes.html

46

46

Datasets

IMPERIAL



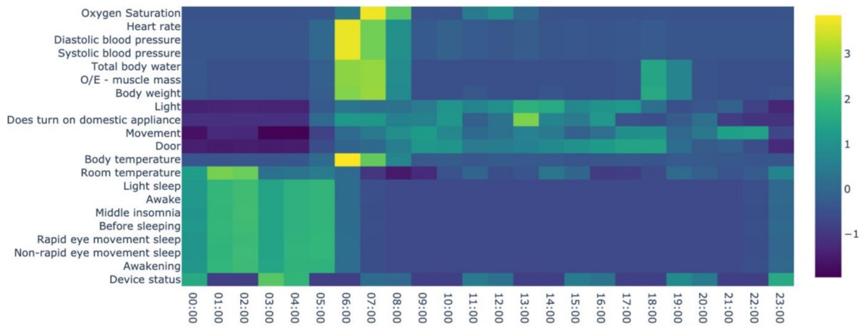
Palermo, F., Chen, Y., Capstick, A. *et al.* TIHM: An open dataset for remote healthcare monitoring in dementia. Sci Data, 10, 606 (2023).

47

47

A sample data for in-home activity

IMPERIAL



48

48

Dataset in the code/programming environment

IMPERIAL

```
from sklearn import datasets

diabetes_X, diabetes_y = datasets.load_diabetes(return_X_y=True)
✓ 0.4s
```

```
diabetes_X
✓ 0.6s
```

```
array([[ 0.03807591,  0.05068012,  0.06169621, ..., -0.00259226,
         0.01990842, -0.01764613],
       [-0.00188202, -0.04464164, -0.05147406, ..., -0.03949338,
        -0.06832974, -0.09220405],
       [ 0.08529891,  0.05068012,  0.04445121, ..., -0.00259226,
         0.00286377, -0.02593034],
       ...,
       [ 0.04170844,  0.05068012, -0.01590626, ..., -0.01107952,
        -0.04687948,  0.01549073],
       [-0.04547248, -0.04464164,  0.03906215, ...,  0.02655962,
         0.04452837, -0.02593034],
       [-0.04547248, -0.04464164, -0.0730303 , ..., -0.03949338,
        -0.00421986,  0.00306441]])
```

49

49

Multivariate data

IMPERIAL

- In general, a learning problem involves a set of n data samples and aims to predict the properties of previously unseen data.
- If each sample is more than a single number and, for instance, a $multi-dimensional$ entry (aka multivariate data), it is said that the data has several attributes or features.

50

50

Feature vector

IMPERIAL

51

51

Training, Test and Validation

IMPERIAL

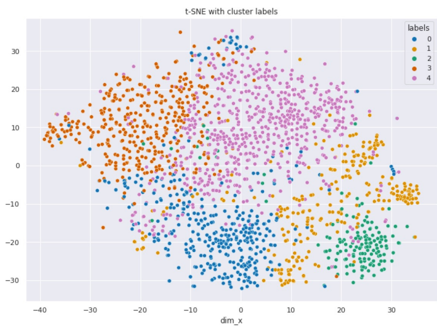
- Training: is used to train the models and determine their parameters (e.g. weights in a neural network).
- Validating: Once you have a collection of algorithms, you need to pick one model. Sometimes, the model that performs best on the validation set is picked up.
- Testing: We use a test set after a model is chosen, but we don't yet know its performance on unseen data. We can apply the chosen model to a test set to see its performance on unseen data.
- How to divide the data? e.g. 60%, 20%, 20%; or 70%, 20%, 10% (need to make sure/check about potential bias and balance/imbalance- more on this later during this module).

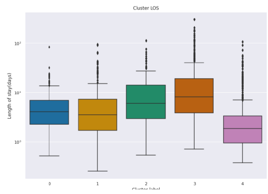
52

52

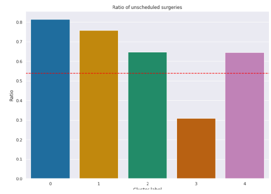
IMPERIAL

Unsupervised learning





Length of stay (log scale)



Rate of emergency surgeries

Image source: J. Li *et al.*, BMC Medical Informatics and Decision Making, 2024.

53

53

IMPERIAL

Clustering

- The problem that needs to be solved in clustering:
  - For example, given a dataset, if we knew there were 3 groups but did not have access to a taxonomy (i.e., labels) to directly identify them, we could try a clustering task.
  - This means splitting the data into well-separated groups called clusters.

54

54

K-means clustering

IMPERIAL

- The $K - means$ algorithm clusters data by trying to separate samples into a set of equal variance, minimising a criterion known as the inertia or within-cluster sum-of-squares.
- This algorithm requires the number of clusters to be specified.
- It scales well to a large number of samples and has been used across a wide range of application areas.
- The algorithm divides a set of N samples into K disjoint clusters C_i, each described by a mean μ_i.

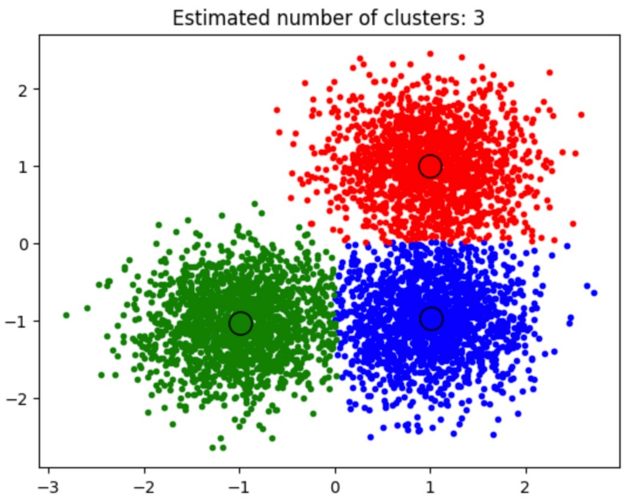
Source: Scikit learn documentation

55

55

K-means

IMPERIAL



56

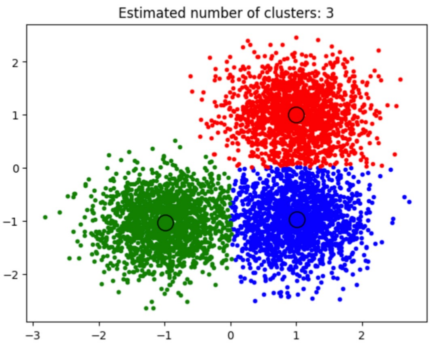
56

K-Means

IMPERIAL

- The means are commonly called the cluster "centroids"; note that they are not, in general, points from X, although they live in the same space.
- The K-means algorithm aims to choose centroids that minimise the inertia, or within-cluster sum of squared criterion:

$$\sum_{i=0}^{n} \min_{\mu_j \in C}(\|\mu_j - \mu_i\|^2)$$



57

57

How K-means works

IMPERIAL

- K-means has three steps.
- The first step chooses the initial centroids, with the most basic method being to choose k samples from the dataset X.
- After $initialisation$, K-means consists of looping between the two other steps.
- The first step assigns each sample to its nearest centroid. The second step creates new centroids by taking the mean value of all the samples assigned to each previous centroid.
- The difference between the old and the new centroids are computed, and the algorithm repeats these last two steps until this value is less than a threshold. In other words, it repeats until the centroids do not move significantly.

58

58

K-Means Algorithm

IMPERIAL

- First, the initial cluster for each of the points is calculated using the current centroids.
- Secondly, the centroids are updated to the mean of each segment.

- The algorithm then repeats this until a stopping criterion is fulfilled (e.g. error rate below a certain threshold).
- Or, for example, iteration stops when centroids move less than a threshold.

59

59

Convergence in K-means

IMPERIAL

- Performing enough iterations, K-means will converge; however, it may converge to a local minimum. This is highly dependent on the initialisation of the centroids.
- As a result, the computation is often done several times, with different initialisations of the centroids.
- Generally, repats are very important in developing robust machine learning models (more on this in the coming days).

60

60

Convergence in K-means

IMPERIAL

- One method to address this issue is the "k-means++" initialisation scheme, which is implemented in scikit-learn (use the init='kmeans++' parameter).
- This initialises the centroids to be (generally) distant from each other, leading to provably better results than random initialisation, as shown in the reference.

Source: Scikit learn documentation

61

61

K-means Example

IMPERIAL

- Before you need to install sci-kit learn
- In Mac OS
  - $sudo\ easy\_install\ pip$
  - $pip\ install\ -U\ numpy\ scipy\ scikit-learn$
- In Linux
  - $sudo\ apt-get\ install\ build-essential\ python-dev\ python-setuptools\ \backslash \text{python}-numpy\ python-scipy\ \backslash \text{libatlas}-dev\ libatlas3gf-base$

62

62

Open discussion

IMPERIAL

- Discussion Scenario:
  - You've built a machine learning model that predicts Alzheimer's risk over the next 5 years based on age, sex, comorbidities, and current MMSE score.
  - For one individual, the model outputs a risk score of 75 (out of 100).
- Question:
  - Your clinician colleague asks: "How should I interpret this result?" What would your response be?

For more information: Biggart I., Fogel A., Barnaghi, P., Predicting Dementia Risk Using Longitudinal Electronic Health Records Data, 39th Conference on Neural Information Processing Systems (NeurIPS 2025) Workshop: Learning from Time Series for Health, https://openreview.net/pdf?id=2FfRHgVSx3

63

63

IMPERIAL

Practice questions

64

64

Practice questions

IMPERIAL

- In a machine learning experiment, we have 1000 samples. We divide the data into training, validation and test sets as below:

- Training samples: 300
- Validation samples: 200
- Test samples: 500

- Is this a good split?
- Please go to menti.com and add your response (code will be given)

65

65

Hospital stay

IMPERIAL

- In a machine learning experiment, you have been asked to design a model that can predict hospital stays for people who are admitted to a neurology ward.
- What type of model will you use to design your model?

66

66

If you have any questions

IMPERIAL

- Please feel free to arrange a meeting or email (p.barnaghi@imperial.ac.uk).
- To arrange a meeting, please email my colleague, Ms Rhiannon Kirby.
- My office: 928, Sir Michael Uren Research Hub, White City Campus.

67

67