CS 737
Deep Learning (3-1-0)

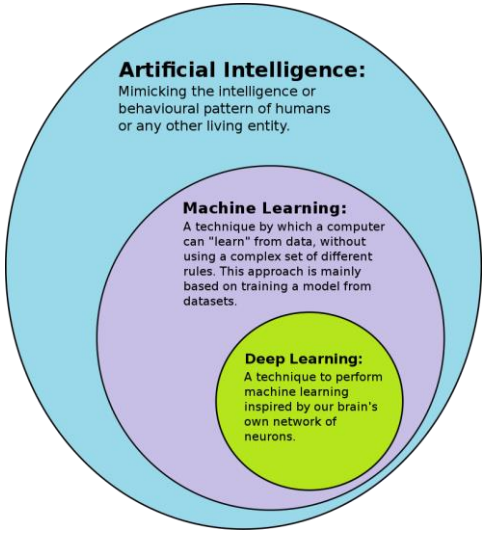
Dr. Jeny Rajan
Dept. of Computer Science and Engg.
National Institute of Technology Karnataka
Surathkal, Mangalore

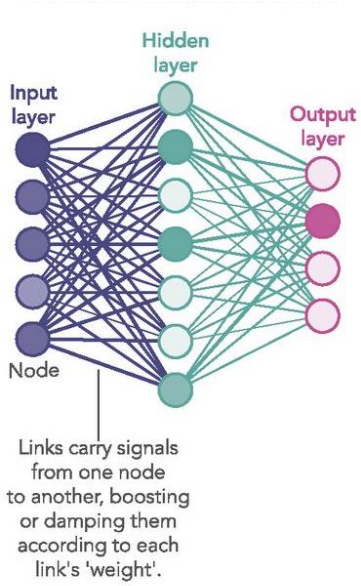
What is Deep Learning ?

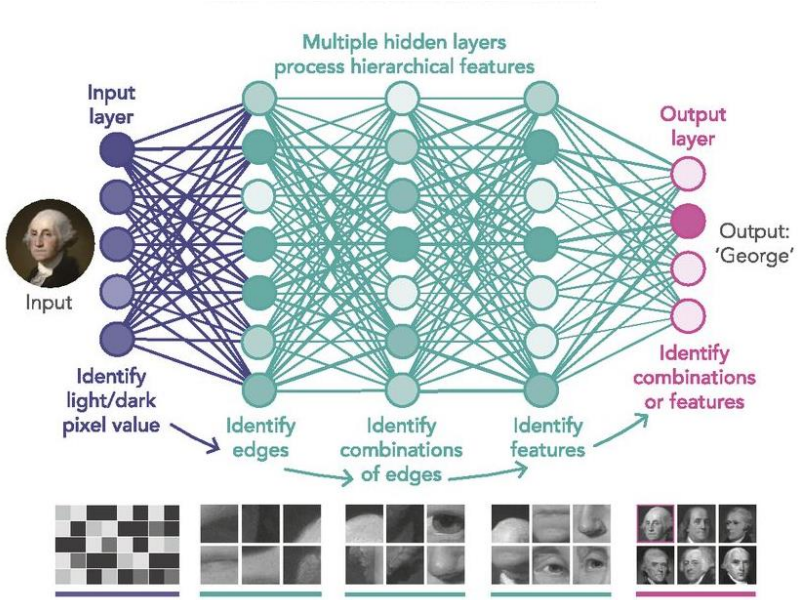
Deep learning is a subset of machine learning, which is essentially a neural network with three or more layers.

The term “deep” refers to the number of layers in the network—the more layers, the deeper the network.

Artificial neural networks (ANNs) were inspired by information processing and distributed communication nodes in biological systems.





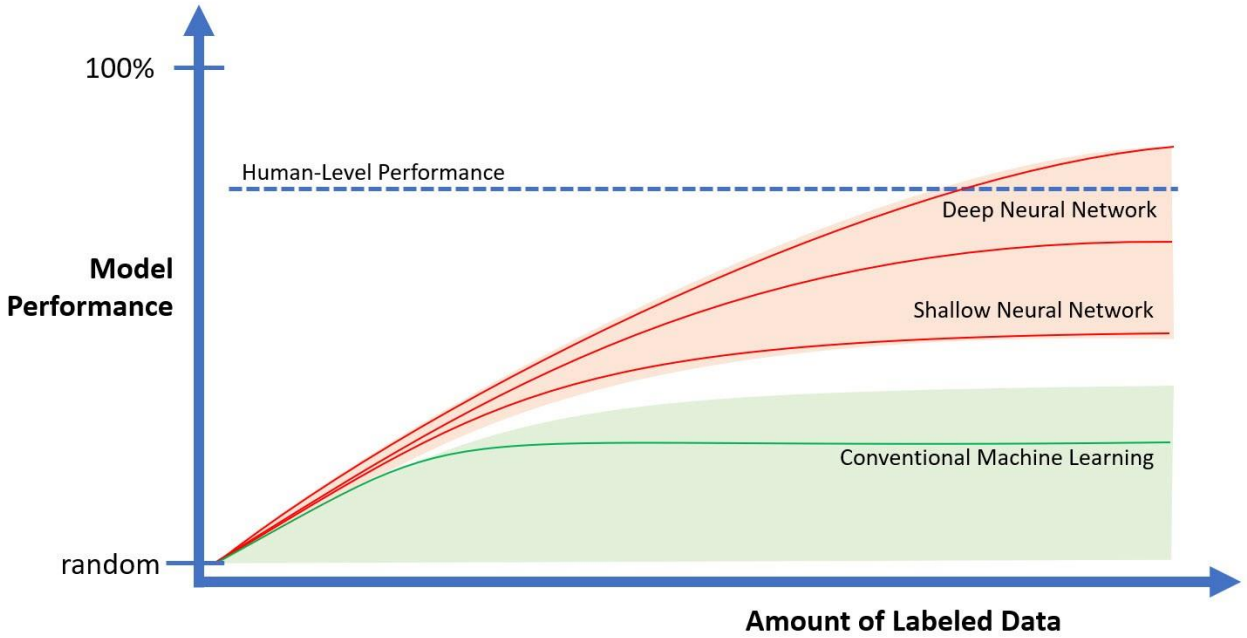


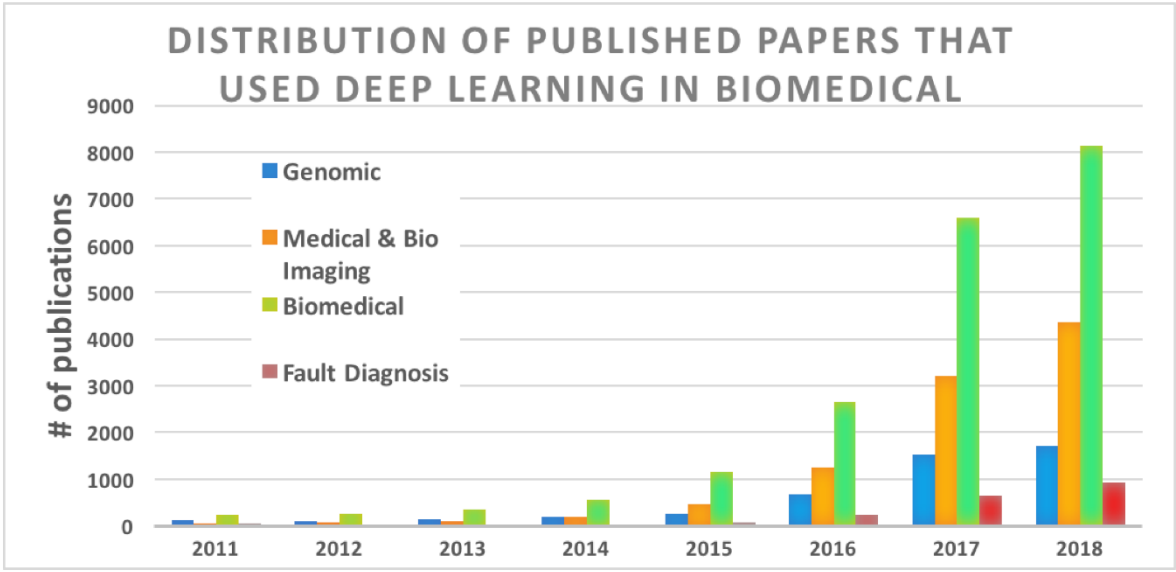
https://www.pnas.org/content/116/4/1074
https://in.mathworks.com/campaigns/offers/next/deep-learning-ebook.html
https://en.wikipedia.org/wiki/Deep_learning

Why Deep Learning ?

- Artificial Neural Networks, especially Convolutional Neural Networks currently provide the best solutions to many problems in image processing, computer vision and natural language processing.

- The ability to process large numbers of features makes deep learning very powerful when dealing with unstructured data.





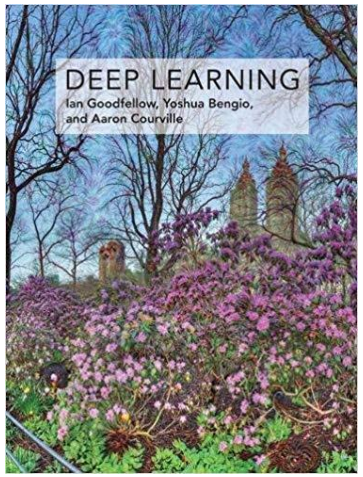
https://henrikhain.io/slides/Recent-Advancements-in-Deep-Reinforcement-Learning/Recent-Advancements-in-Deep-Reinforcement-Learning.html#32

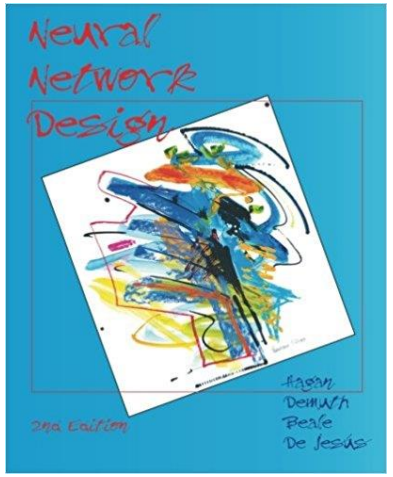
Learning is all about finding useful abstractions or concepts that describe the world.

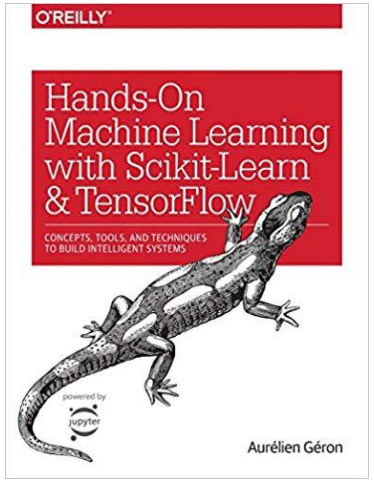
Ref. materials

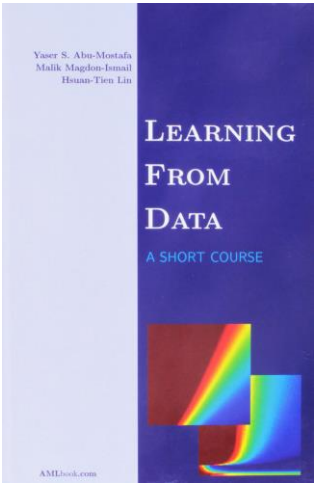
Books:

1. Deep Learning by Ian Goodfellow, Yoshua Bengio and Aaron Courville, MIT Press, 2016.
2. Martin T Hagan et.al, Neural Network Design (2nd Edition).
3. Aurelien Geron, Hands-on Machine learning with Scikit-Learn & TensorFlow, O'Reilly Media, 2017.
4. Yaser S. Abu-Mostafa, Learning From Data, AMLBook,









Course Syllabus

Machine learning basics, Basic neural network models [McCulloch-Pitts Model of Neuron, Perceptron], Adaline, linear and non linear activation functions, loss functions, gradient descent method, back propagation algorithm, Deep feed forward networks, Regularization for deep learning, Convolutional neural networks, Optimization for training deep models, RNN, Autoencoders, Popular deep learning architectures published in the last 10 years, Limitations of CNN, Semi-supervised deep learning, Applications (image classification and segmentation).

Teaching Assistants

Prabuddha Paul

Panhalkar Shweta Prashant

Course Evaluation

Mid Sem – 20 %

End Sem – 40 %

Viva-voce – 20%

Assignments / Course Project/ New ideas – 20%

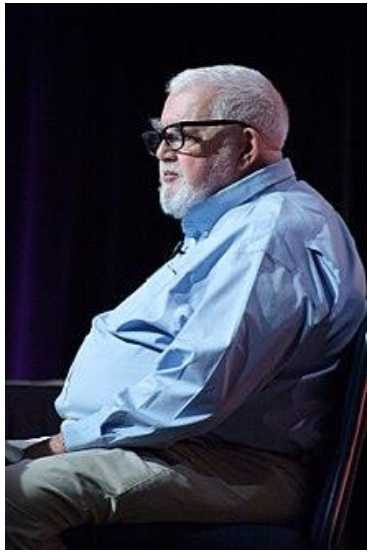
AI and machine learning

AI was introduced as an academic discipline in 1950s to create machines that can think.

John McCarthy, widely recognized as one of the godfathers of AI, defined it as the science and engineering of making intelligent machines.

Old-Fashioned AI were mainly based on rules. (eg. Chess-playing system).

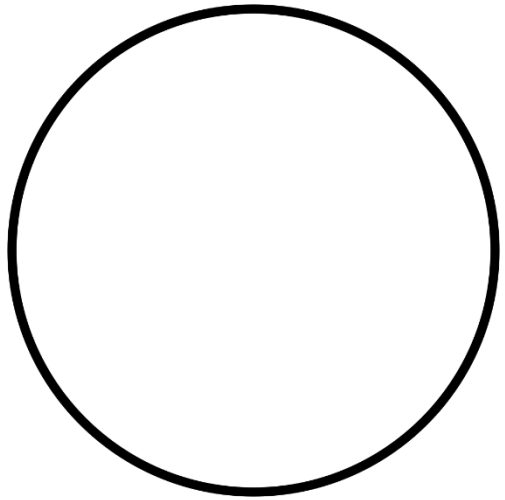
Machine learning is the field of study that gives computers the ability to learn without being explicitly programmed (Arthur Samuel, 1959).

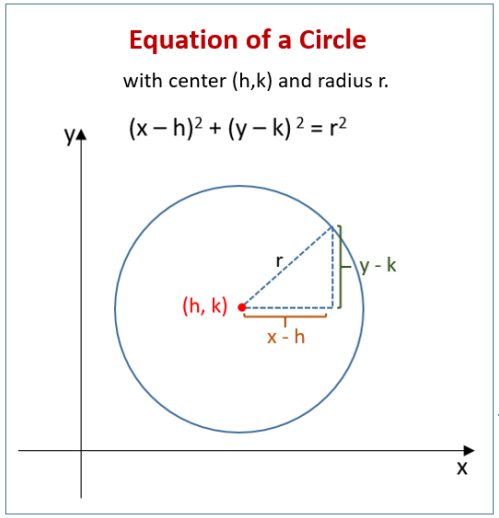


John McCarthy

Arthur Lee Samuel

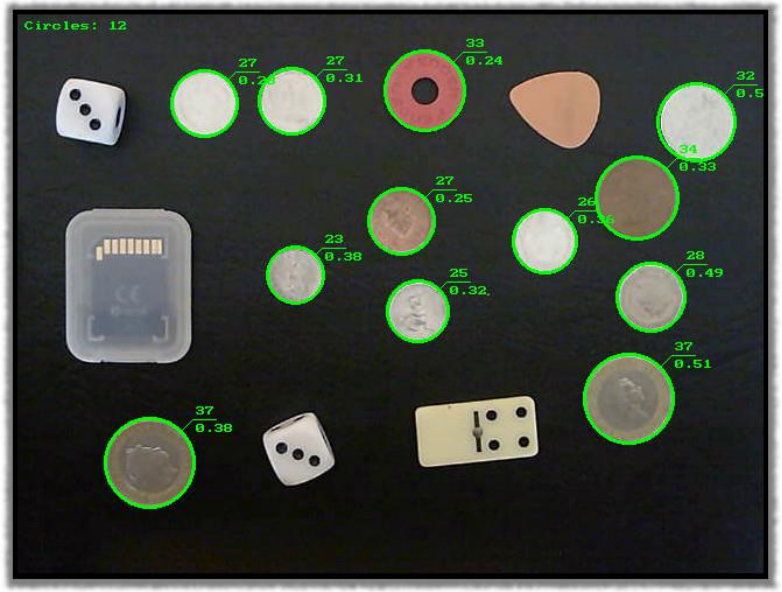
Rules to define face ?

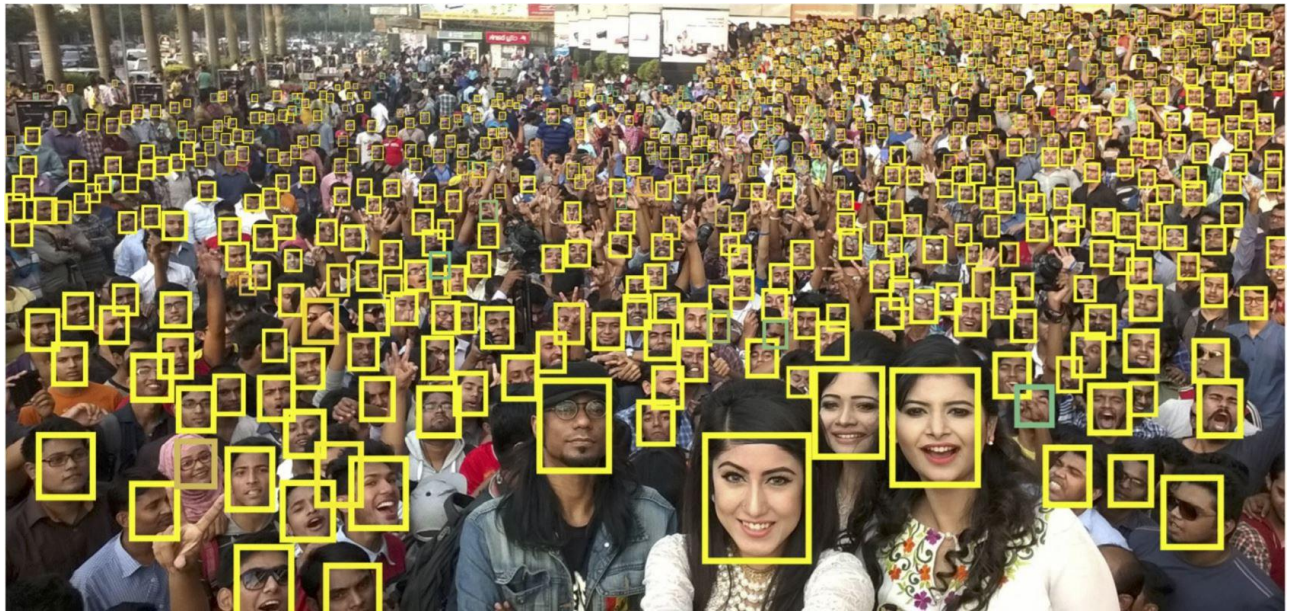


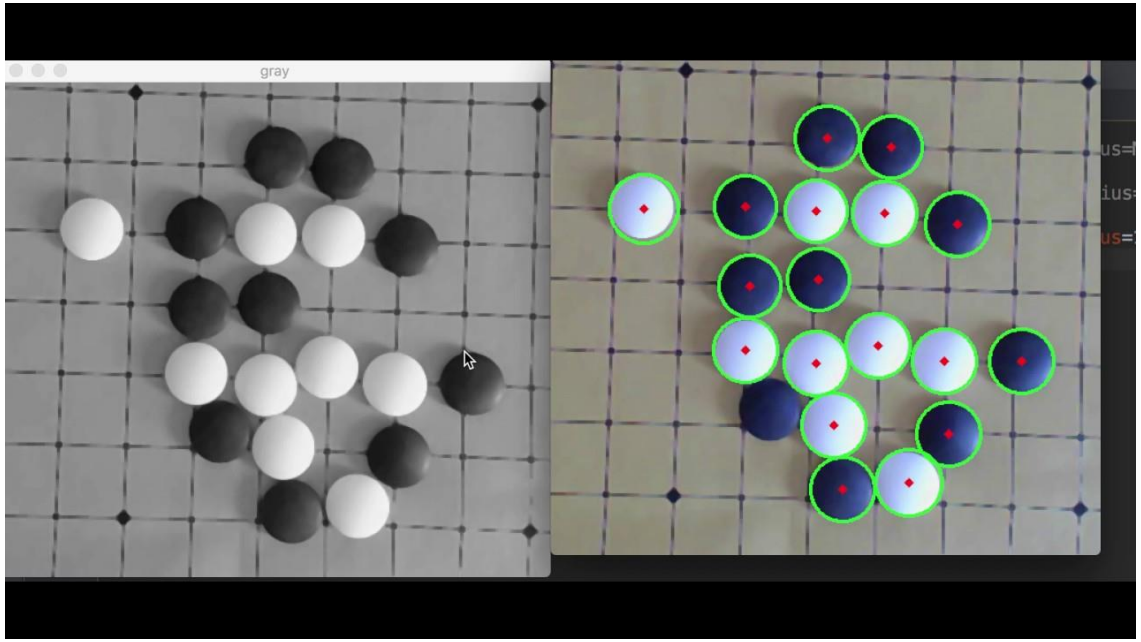


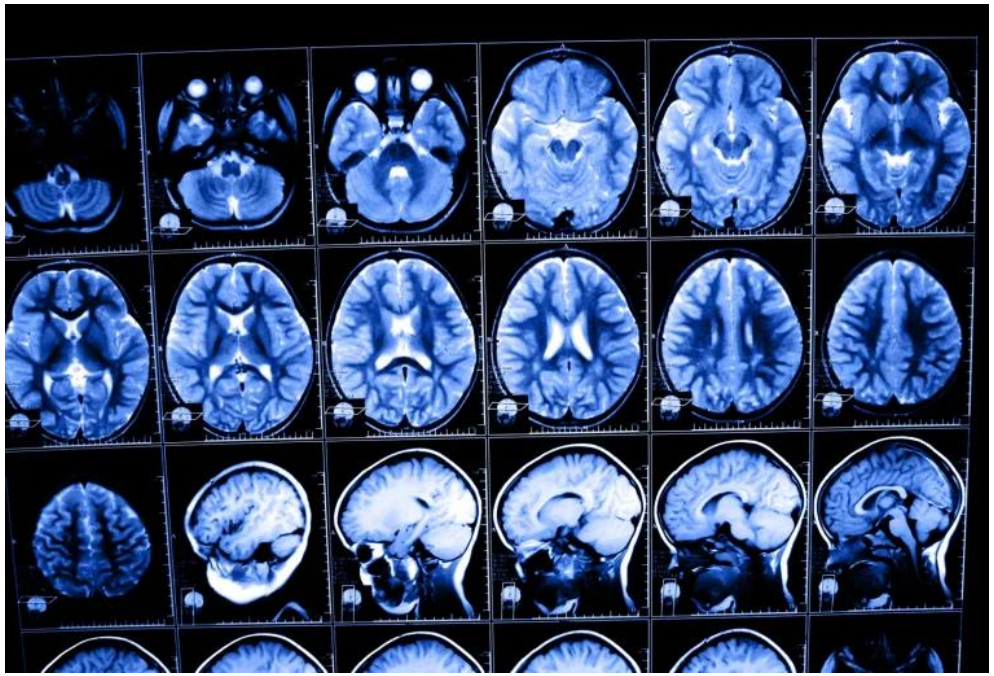
Equation of a Tree ?

We learned trees by looking at trees, not by studying its mathematical definition. In other words, we learned from data.


Circles: 12
27
0.29
27
0.31
33
0.24
32
0.5
34
0.33
27
0.25
26
23
0.38
25
0.32
28
0.49
37
0.51
37
0.38

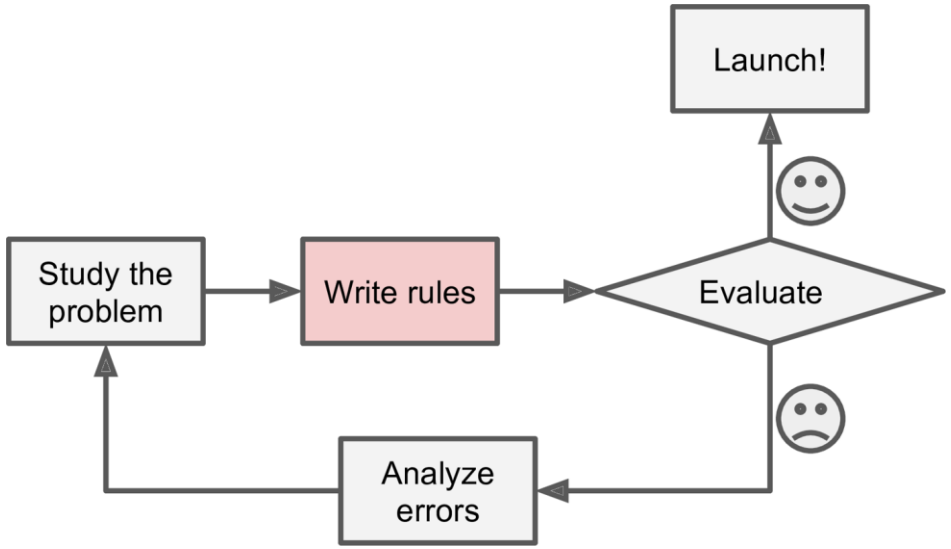



gray

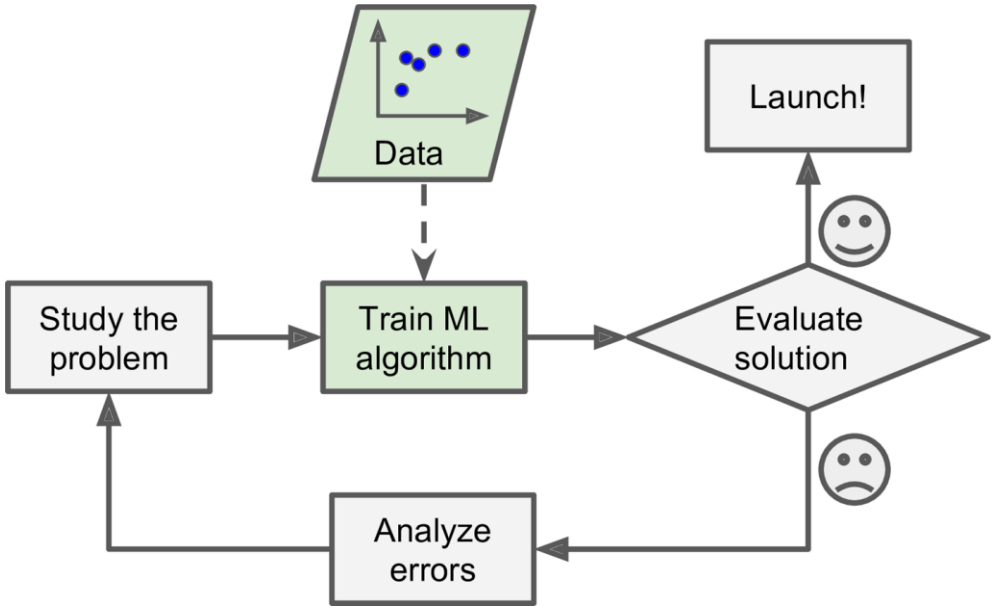


Machine learning methods learns from data.

- Learning from data is used in situations where we don't have an analytic solution, but we do have data that we can use to construct an empirical solution.
- Learning from data is one of the most widely used techniques in science, engineering, and economics, among other fields.
- Machine learning is the science and art of programming computers so they can learn from data.



Traditional Approach



Machine Learning Approach

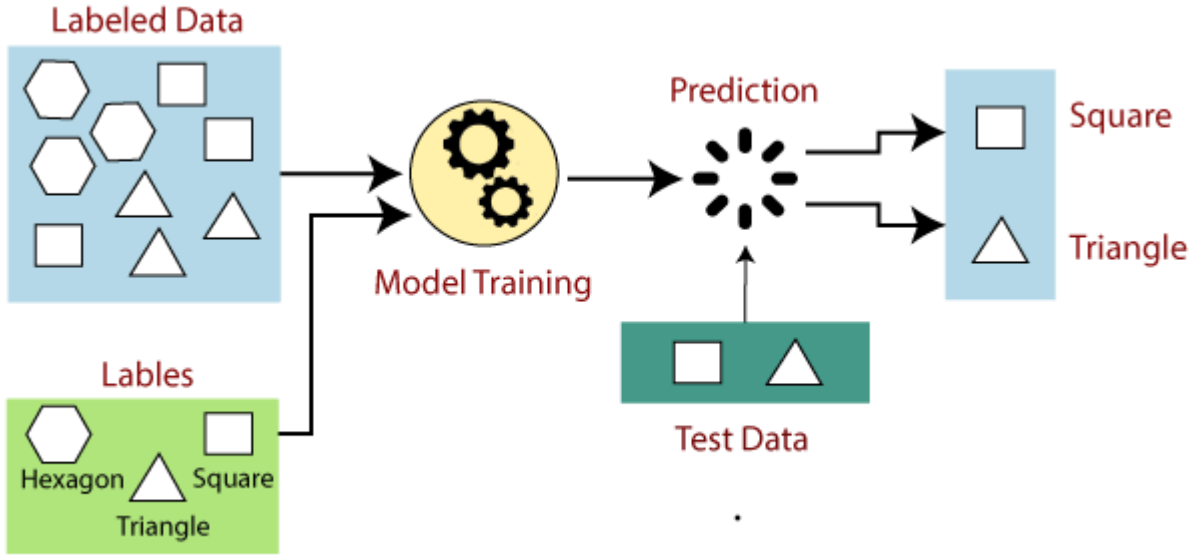
https://www.oreilly.com/library/view/hands-on-machine-learning/9781491962282/ch01.html

Types of Learning

- **Supervised learning** is the machine learning task of learning a function that maps an input to an output based on example input-output pairs.

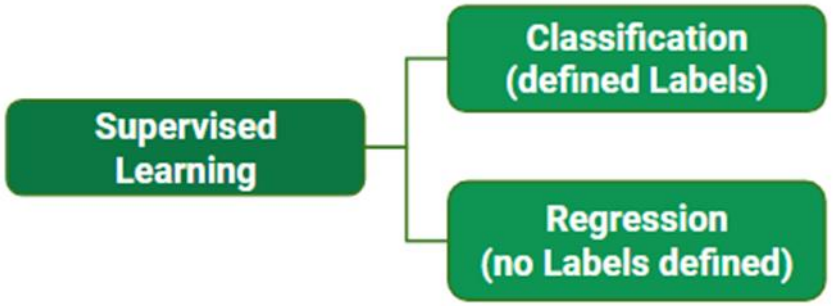
- It infers a function from *labeled* *training data* consisting of a set of *training examples*.

- In supervised learning, each example is a *pair* consisting of an input object (typically a vector) and a desired output value (also called the *supervisory signal*).



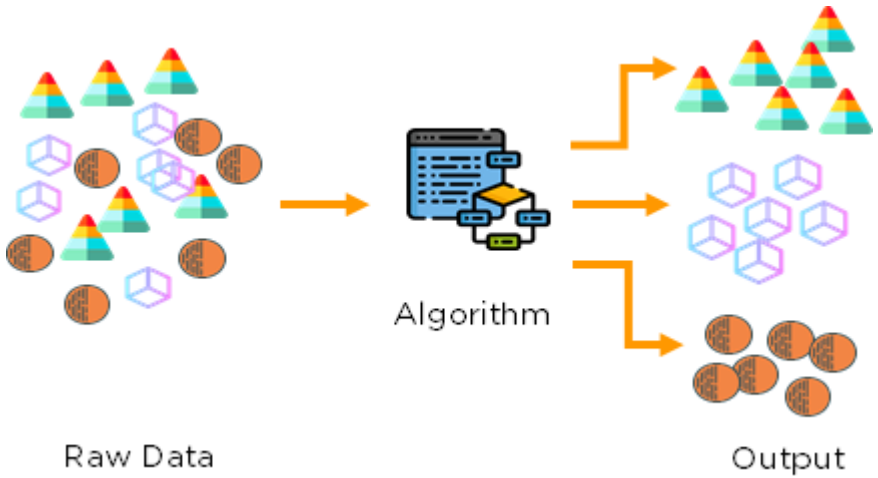
Some important supervised learning methods

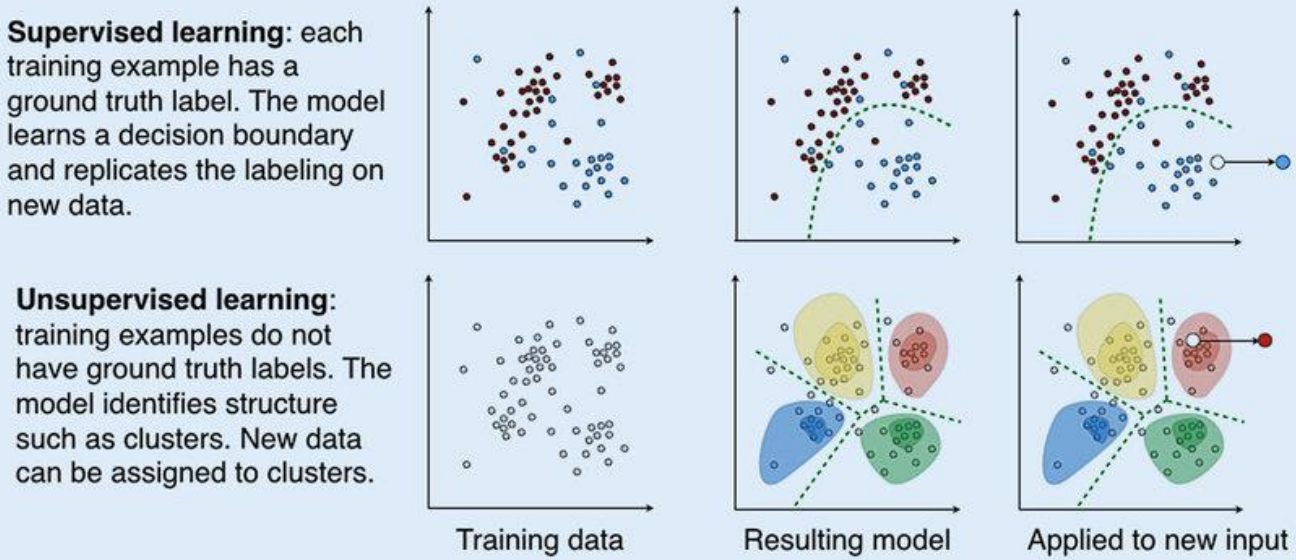
- Linear regression
- Logistic regression
- Support Vector Machines (SVM)
- Decision Trees and Random Forests
- Neural Networks



Unsupervised Learning

- In the unsupervised setting, the training data does not contain any output information at all.
- In contrast to supervised learning, it's not always easy to come up with metrics for how well an unsupervised learning algorithm is doing.





Some important unsupervised learning methods

- Clustering
- Dimensionality reduction

Main challenges of machine learning

Insufficient quantity of training data

- Many machine learning algorithms require large amounts of data before they begin to give useful results.
- Even for very simple problems you typically need thousands of examples, and for complex problems such as image or speech recognition you may need millions of examples (unless you can reuse parts of an existing model).

The Unreasonable Effectiveness of Data

In a famous paper published in 2001, Microsoft researchers Michele Banko and Eric Brill showed that very different Machine Learning algorithms, including fairly simple ones, performed almost identically well on a complex problem of natural language disambiguation[8] once they were given enough data (as you can see in Figure 1-20).

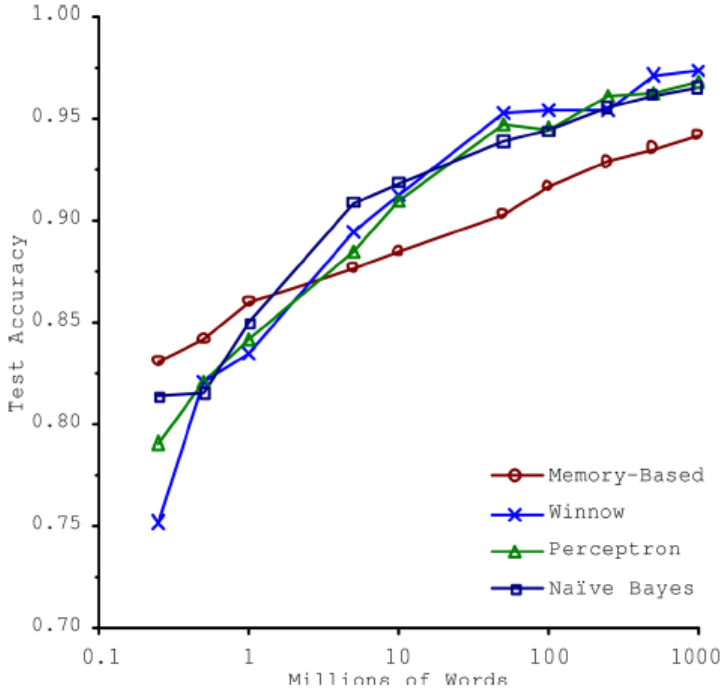
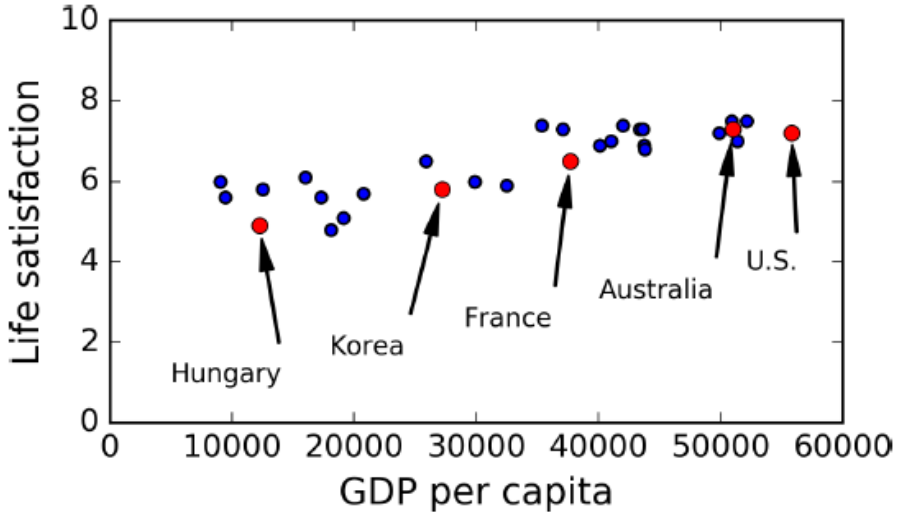
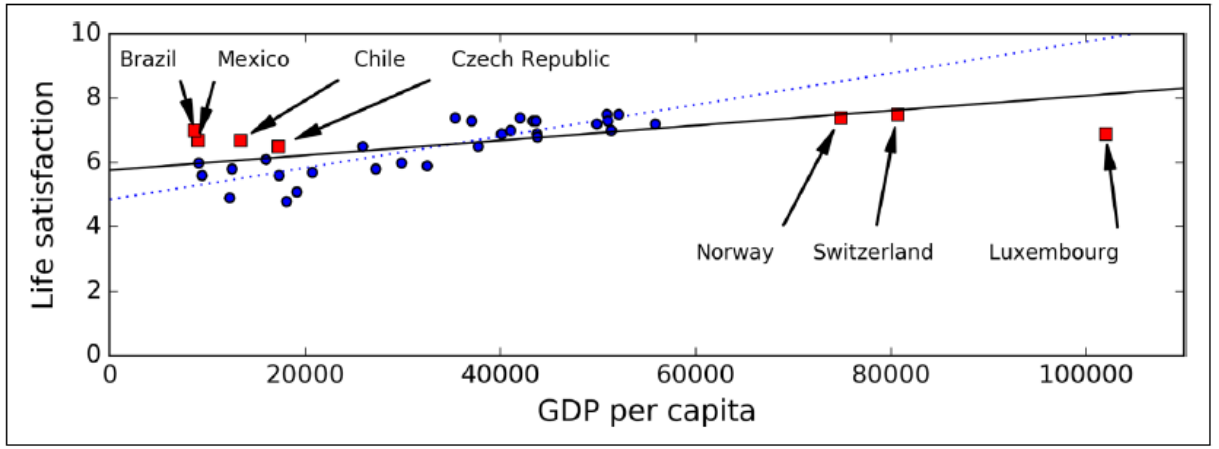


Figure 1-20. The importance of data versus algorithms[9]

Nonrepresentative Training data

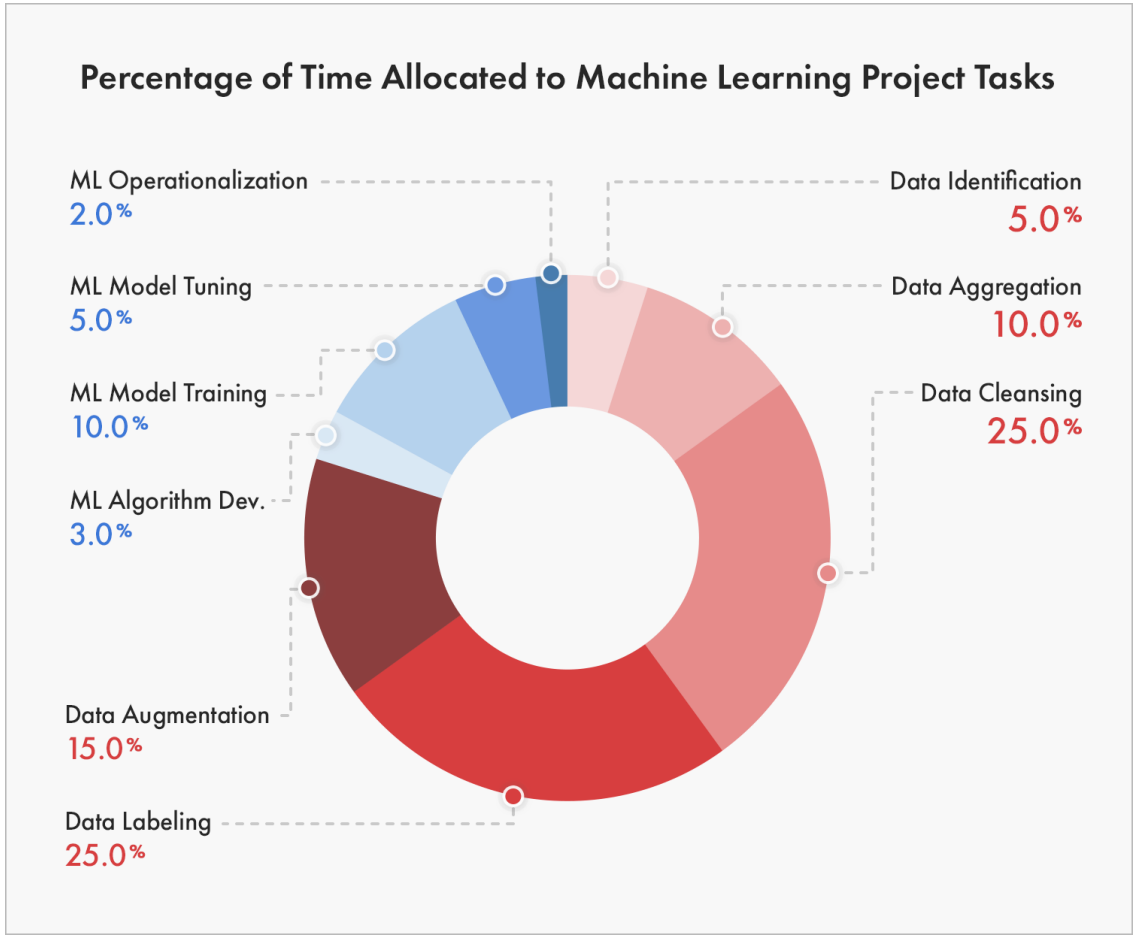
- In order to generalize well, it is crucial that your training data be representative of the new cases you want to generalize to.
- It is crucial to use a training set that is representative of the cases you want to generalize to.
- If the sample is too small, you will have *sampling noise* (i.e., nonrepresentative data as a result of chance), but even very large samples can be nonrepresentative if the sampling method is flawed. This is called *sampling bias*.





Poor-Quality data

- If your training data is full of errors, outliers, and noise (e.g., due to poor quality measurements), it will make it harder for the system to detect the underlying patterns, so your system is less likely to perform well.
- Clean the data before using it for training/testing
- If some instances are clearly outliers, it may help to simply discard them or try to fix the errors manually.

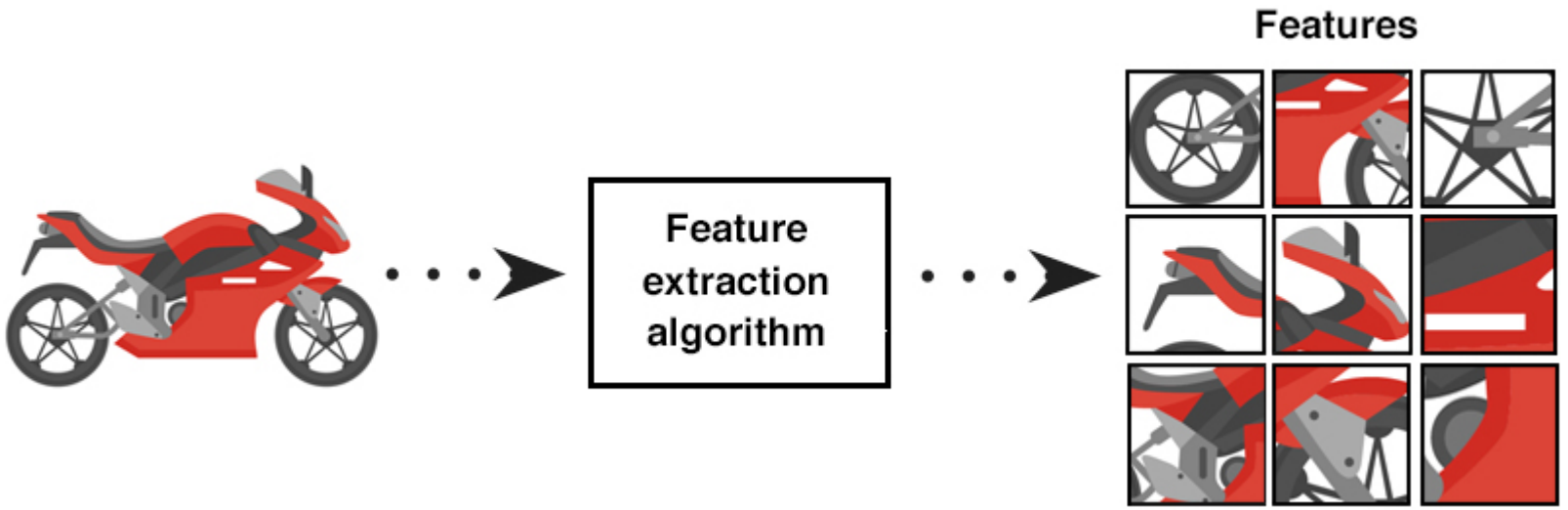


Irrelevant Features

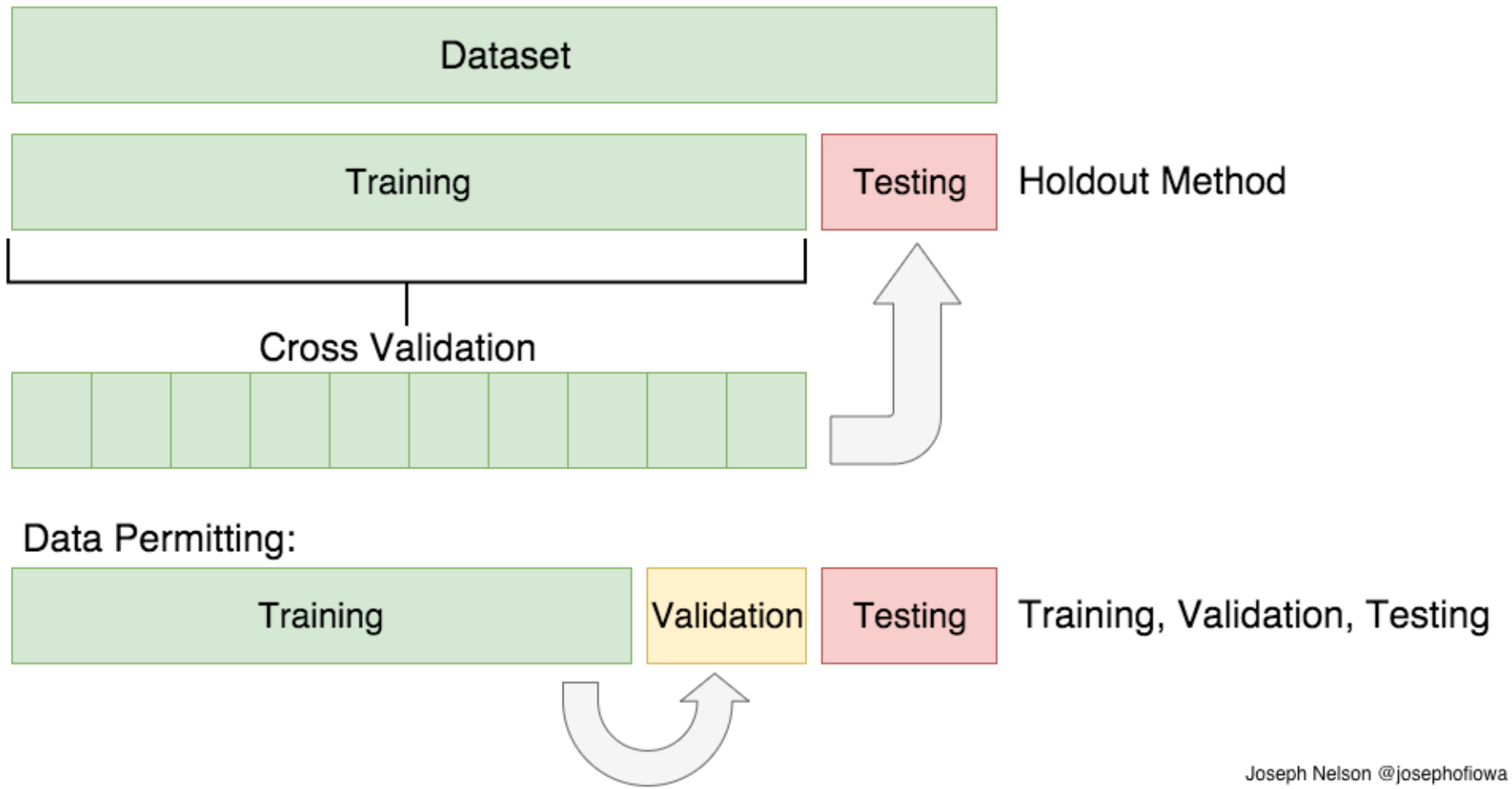
A critical part of the success of a Machine Learning project is coming up with a good set of features to train on.

This process, called *feature engineering*, involves:

- *Feature selection*: selecting the most useful features to train on among existing features.
- *Feature extraction*: combining existing features to produce a more useful one (as we saw earlier, dimensionality reduction algorithms can help).
- Creating new features by gathering new data



Training, Validation and Test data



https://towardsdatascience.com/train-test-split-and-cross-validation-in-python-80b61beca4b6

Overfitting and underfitting

Overfitting : The model performs well on the training data, but does not generalize well.

Underfitting : The model is too simple to learn the underlying structure of the data.