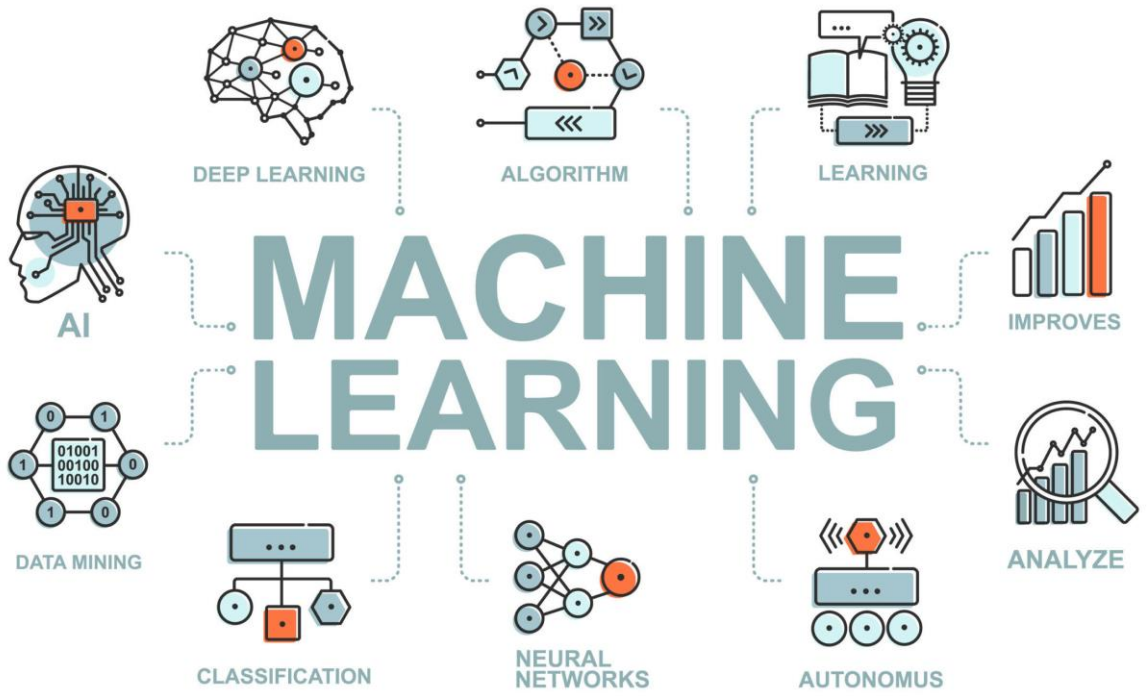
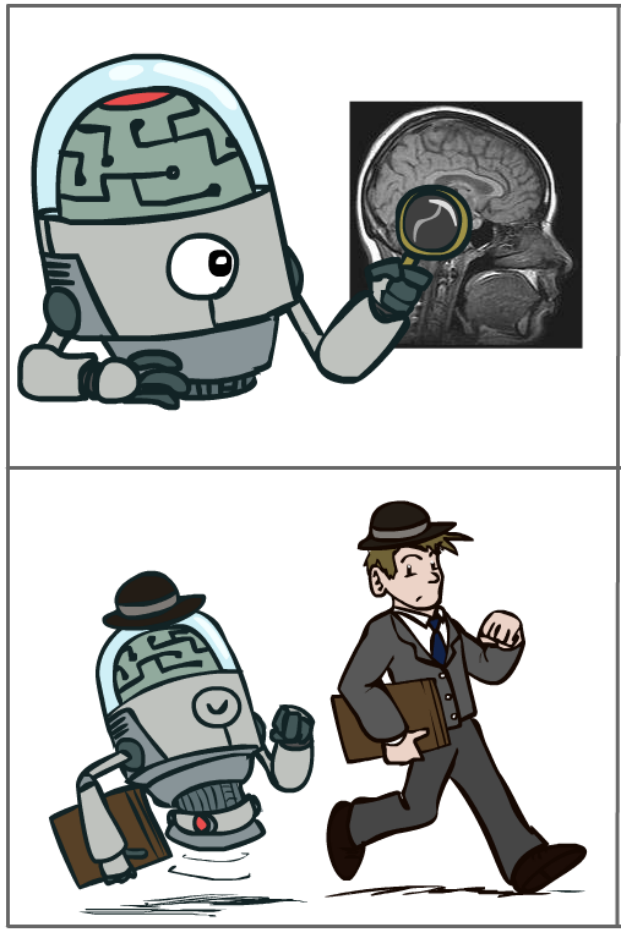
Introduction to Machine Learning



What is AI?

The science of making machines that:



emulate human behavior, enabling them to **learn**, **make decisions**, **recognize patterns**, and **solve complex problems** in a manner akin to human intelligence

AI vs. ML

- **Artificial Intelligence (AI):**
  - AI is a broad field of computer science that aims to create systems that can perform tasks that typically require human intelligence
  - It encompasses a wide range of techniques, approaches, and applications, including problem-solving, natural language understanding, speech recognition, computer vision, and decision-making
- **Machine Learning (ML):**
  - ML is a subset of AI that focuses on developing algorithms and models that enable machines to learn patterns and make predictions or decisions based on data.
  - It involves the use of algorithms to allow machines to improve their performance on a specific task over time without being explicitly programmed.

AI vs. ML vs. DL vs. Generative AI

Artificial Intelligence

AI involves techniques that equip computers to emulate human behavior, enabling them to learn, make decisions, recognize patterns, and solve complex problems in a manner akin to human intelligence.

Machine Learning

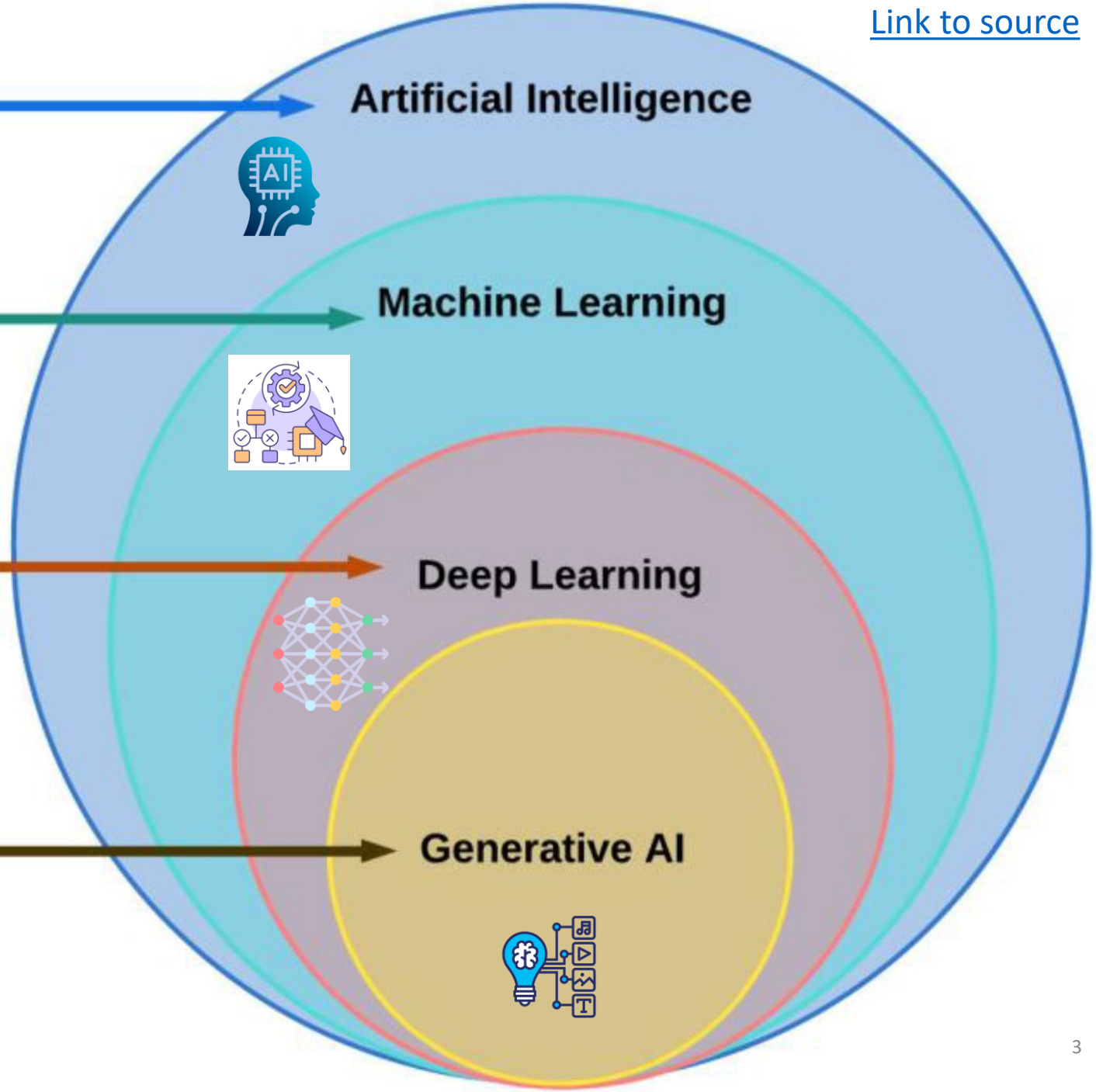
ML is a subset of AI, uses advanced algorithms to detect patterns in large data sets, allowing machines to learn and adapt. ML algorithms use supervised or unsupervised learning methods.

Deep Learning

DL is a subset of ML which uses neural networks for in-depth data processing and analytical tasks. DL leverages multiple layers of artificial neural networks to extract high-level features from raw input data, simulating the way human brains perceive and understand the world.

Generative AI

Generative AI is a subset of DL models that generates content like text, images, or code based on provided input. Trained on vast data sets, these models detect patterns and create outputs without explicit instruction, using a mix of supervised and unsupervised learning.



Link to source

Machine Learning

- **Machine learning** is the subfield of computer science that gives "computers the ability to learn without being explicitly programmed."

  - term coined by Arthur Samuel 1959 while at IBM

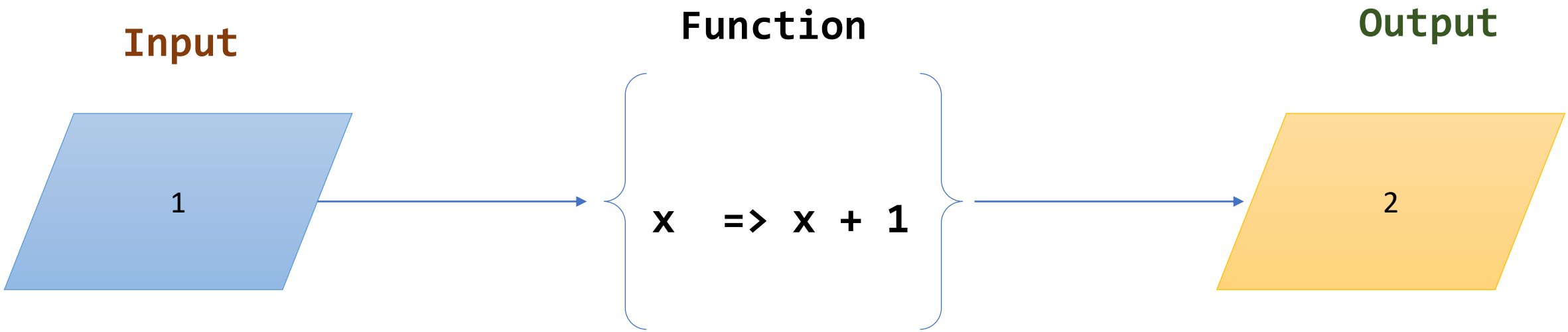
- The study of algorithms that can learn from data.

- Machine learning **focuses on the development of computer programs that can access data and use it learn for themselves**

  - expert.ai

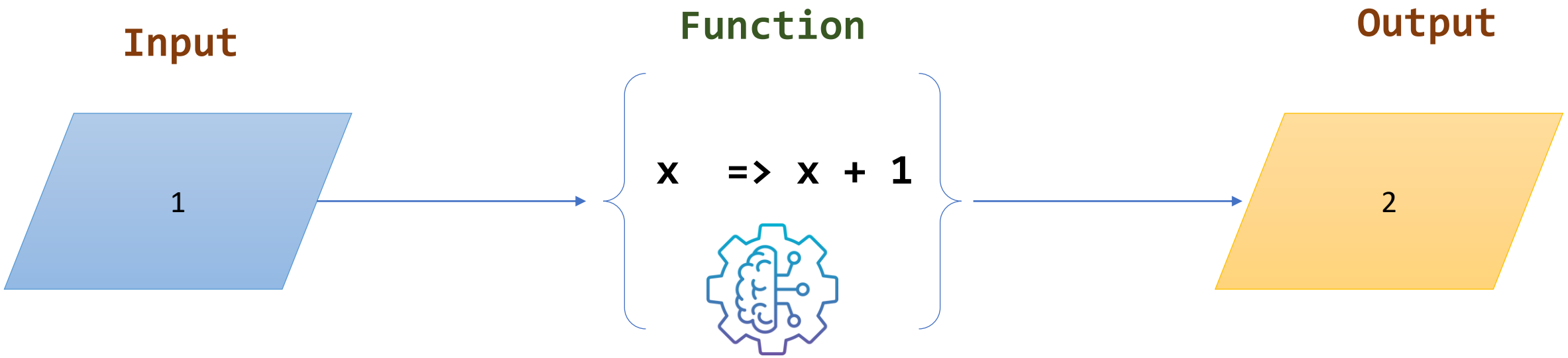
Coding : Input + Function → Output

- In conventional coding, the programming collects the **Input** and write the **Function(s)** to produce the desired **Output**



ML : Input + Output → Function

- ML is all about intelligent guessing/approximation of a function to produce the known **Output** from a given **Input**



Supervised Learning Examples:

Estimate the function f from input/output examples

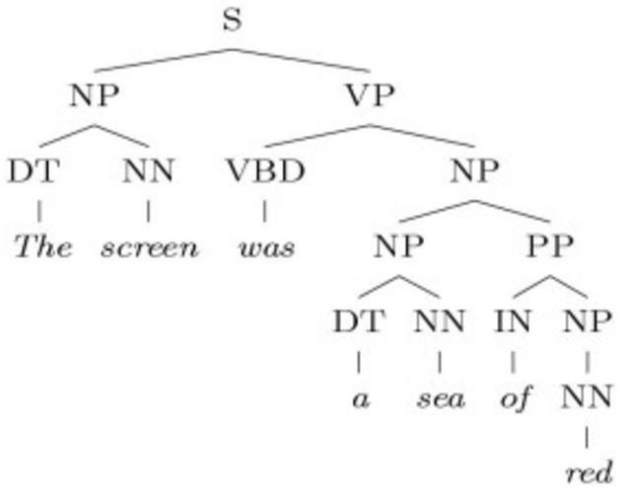
$f($  $) =$ cat

Classification

$f($  $) =$ 

Facial Detection

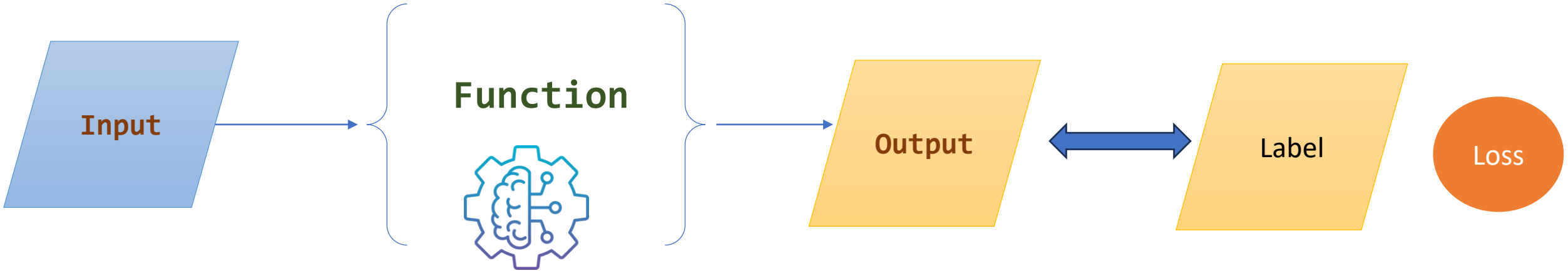
$f($ The screen was a sea of red $) =$



Language Parsing

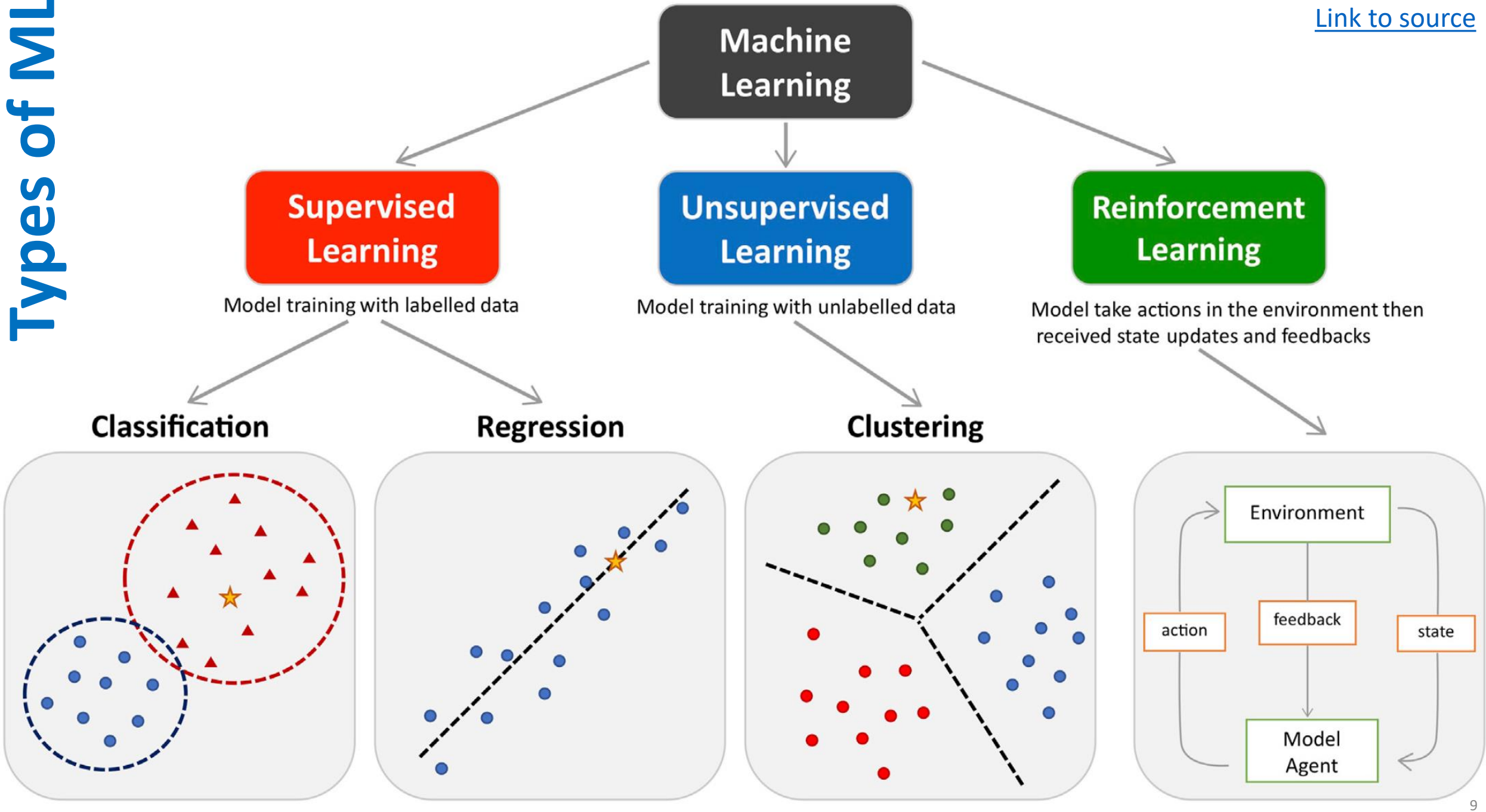
ML: learn a Function that minimizes the loss

- Start with random function parameters
- Repeat intelligent guessing/approximation of the Function parameters such that the difference between the Output the expected Output (i.e., Loss) is reduced



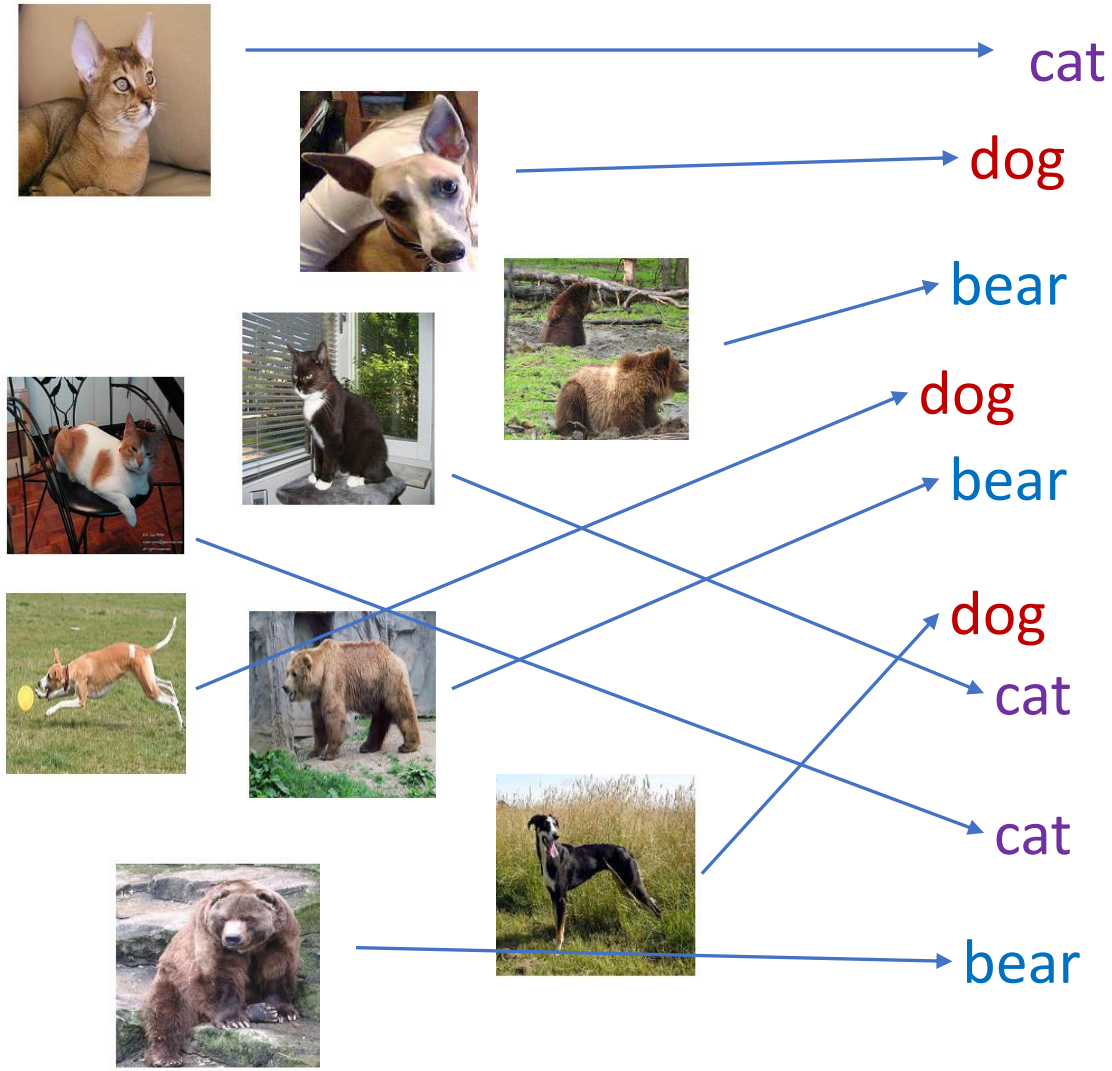
Types of ML

[Link to source]



Supervised Learning vs Unsupervised Learning

$x \rightarrow y$



x

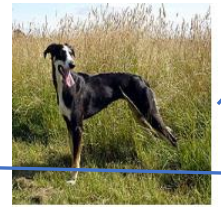
Supervised Learning vs Unsupervised Learning

$x \rightarrow y$








cat
dog
bear
dog
bear
dog
cat
cat
bear

Classification

x





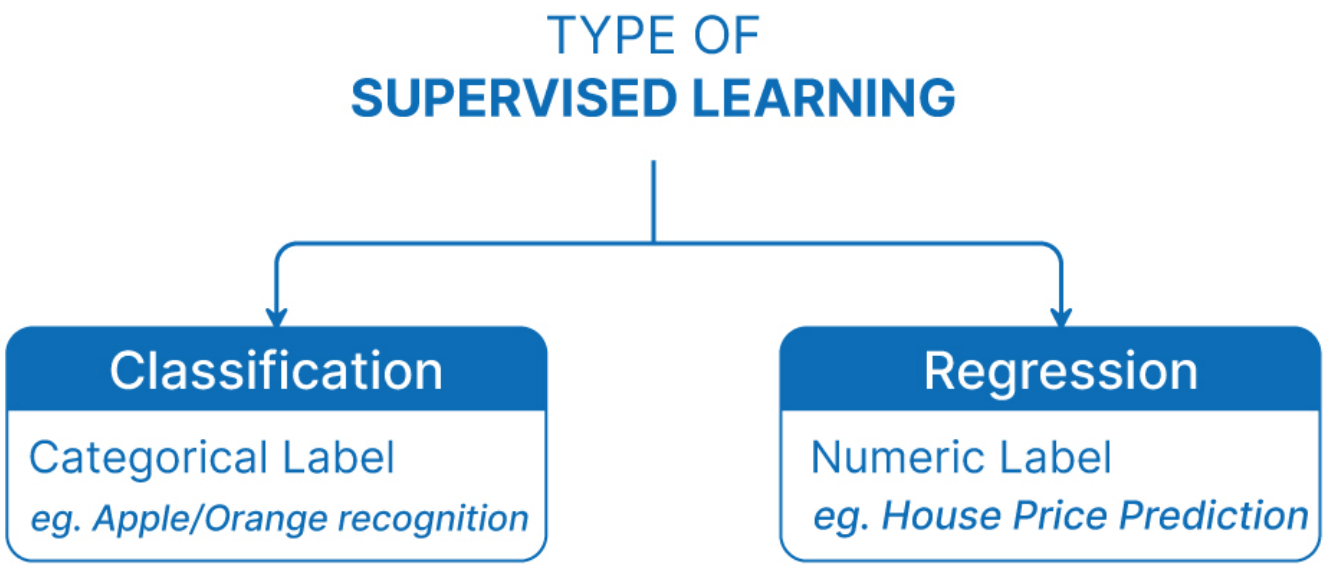




Clustering

Supervised Learning

- Uses a set of **input values** with their **labels** to train the model
  - E.g., classify an email into spam or not spam
  - E.g., using information about a house such as the number of rooms, size, location, to predict the output variable -> the house price



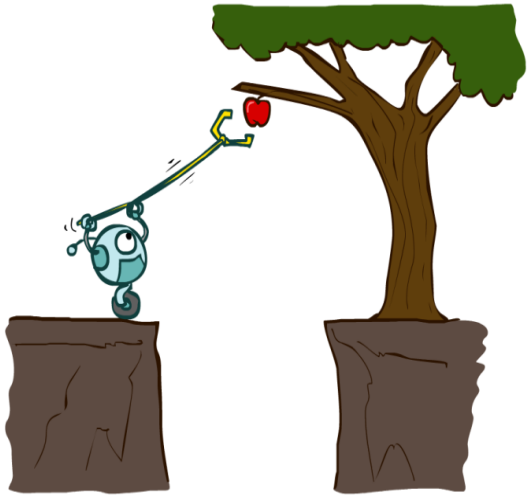
Unsupervised Learning

- Sometimes we don’t have output labels available in our dataset and we wish to group our data into clusters
  - e.g., having a movies dataset with lots of movie features such as their genre, rating, target audience, etc, and then grouping movies that are similar to each other

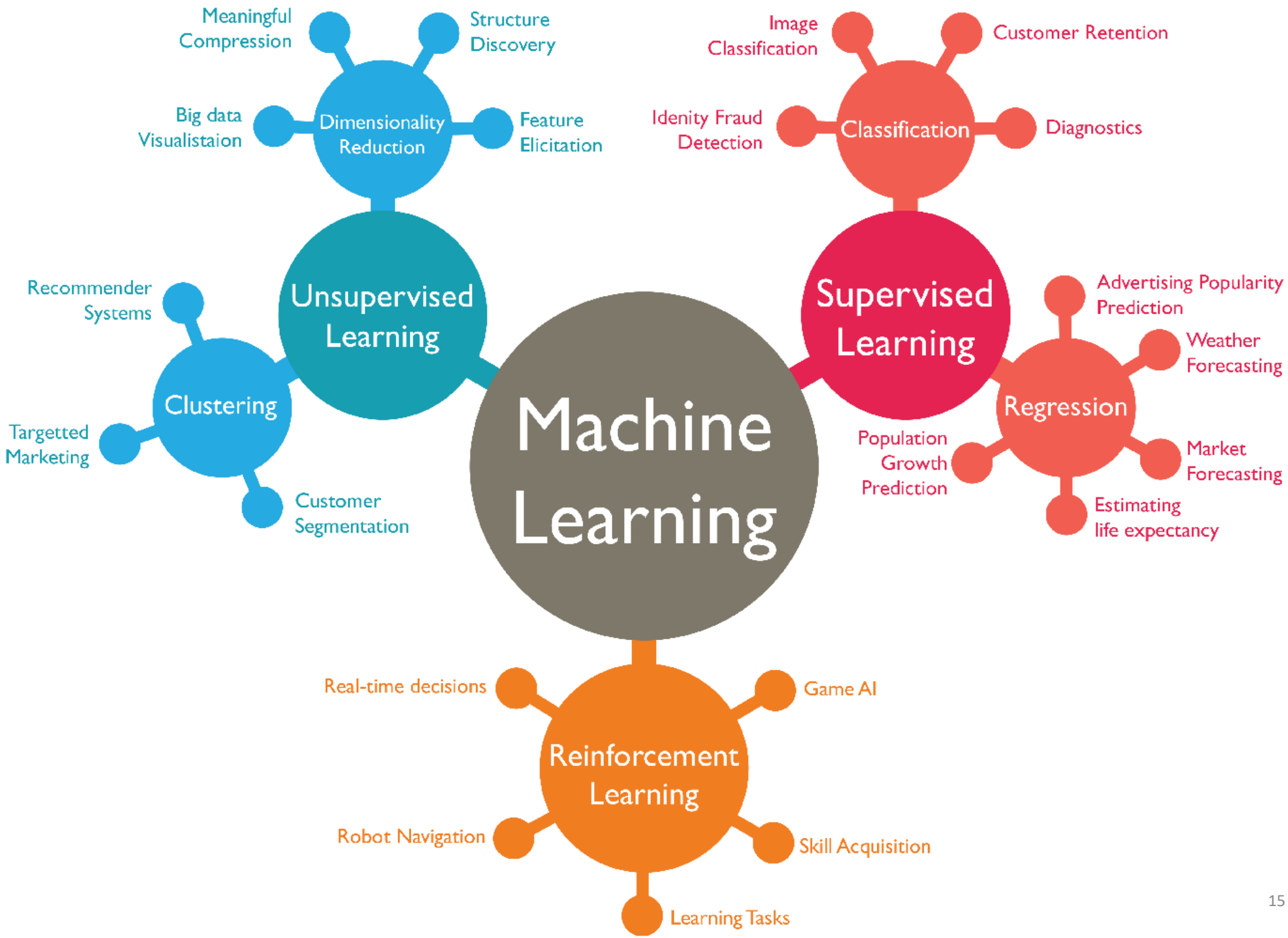
Reinforcement Learning

- These work similar to how a user may play a video game. The model is set up with a system of awards and penalties wherein, if the ‘player’ performs well they get an award and if they don’t they are penalized
  - Learns from mistakes. Over time, the model gets better and better at playing the game

  e.g., beat the world’s best players at the game of Go and Chess



ML Application Areas



[Link to source]

Common Classic ML Algorithms

[Link to source]

Linear		Linear Regression	**The "best fit" line** through all data points. Predictions are numerical.	Easy to understand — you clearly see what the biggest drivers of the model are.	Sometimes too simple to capture complex relationships between variables. Does poorly with correlated features.
		Logistic Regression	The adaptation of **linear regression** to problems of classification (e.g., yes/no questions, groups, etc.)	Also easy to understand.	Sometimes too simple to capture complex relationships between variables. Does poorly with correlated features.
Tree-Based		Decision Tree	**A series of yes/no rules based on the features,** forming a tree, to match all possible outcomes of a decision.	Easy to understand.	Not often used on its own for prediction because it's also often too simple and not powerful enough for complex data.
		Random Forest	Takes advantage of many decision trees, with rules created from subsamples of features. Each tree is weaker than a full decision tree, but **by combining them we get better overall performance.**	A sort of "wisdom of the crowd." Tends to result in very high quality models. Fast to train.	Models can get very large. Not easy to understasnd predictions.

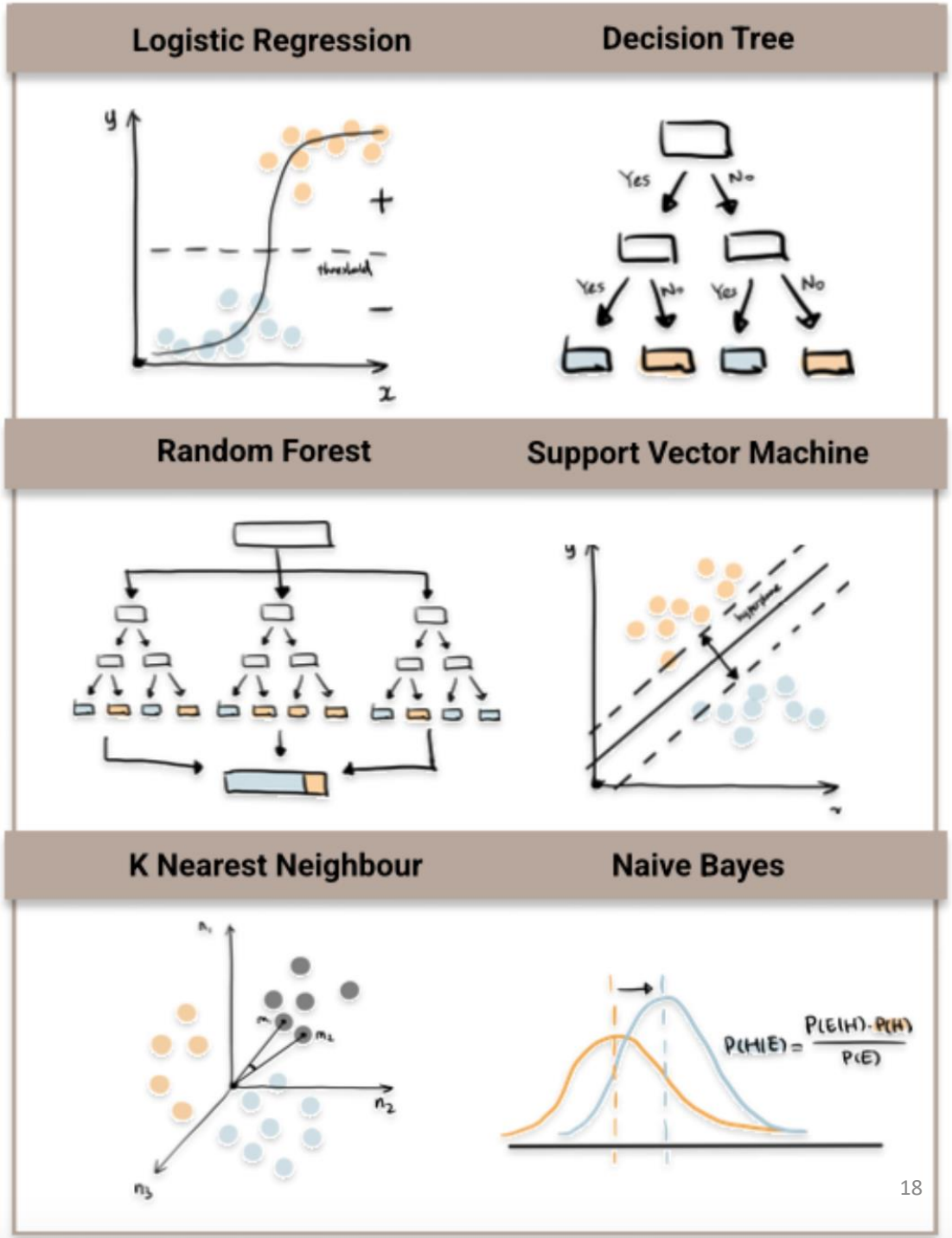
Common ANN ML Algorithms

Link to source

Neural Networks

	**CNN (Convolutional Neural Network)**	Neurons connected in a grid fashion to be spatially aware.	Best in class for image recognition use cases.	Vulnerable to noise & examples which are out of the training data context.
	**RNN (Recurrent Neural Network)**	Neurons connected sequentially for temporal awareness.	Performant for time series use cases.	Weak for long-term memory — struggle to remember patterns from much earlier time steps.
	**LSTM (Long Short-Term Memory)**	RNNs capable of both long- & short-term memory.	Can selectively remember and recall history from the recent & distant past.	Slow to compute and difficult to optimize.
	**Transformer**	Networks with the ability to focus attention on history more selectively.	Best in class for natural language understanding, processing, and generation use cases.	Very resource intensive, even for inference. Very difficult to interpret.

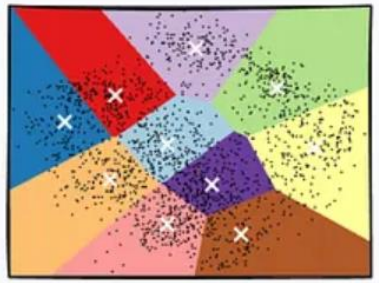
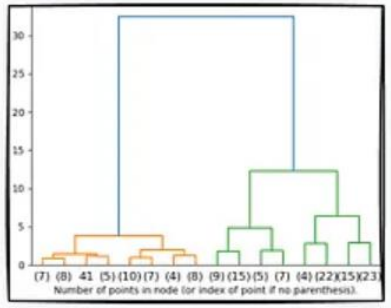
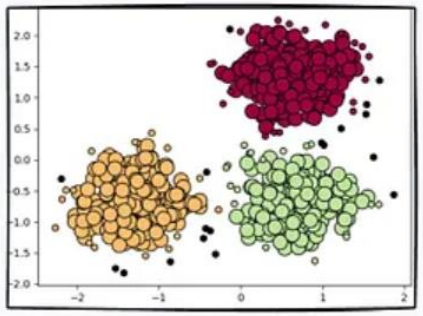
Common Classification Algorithms



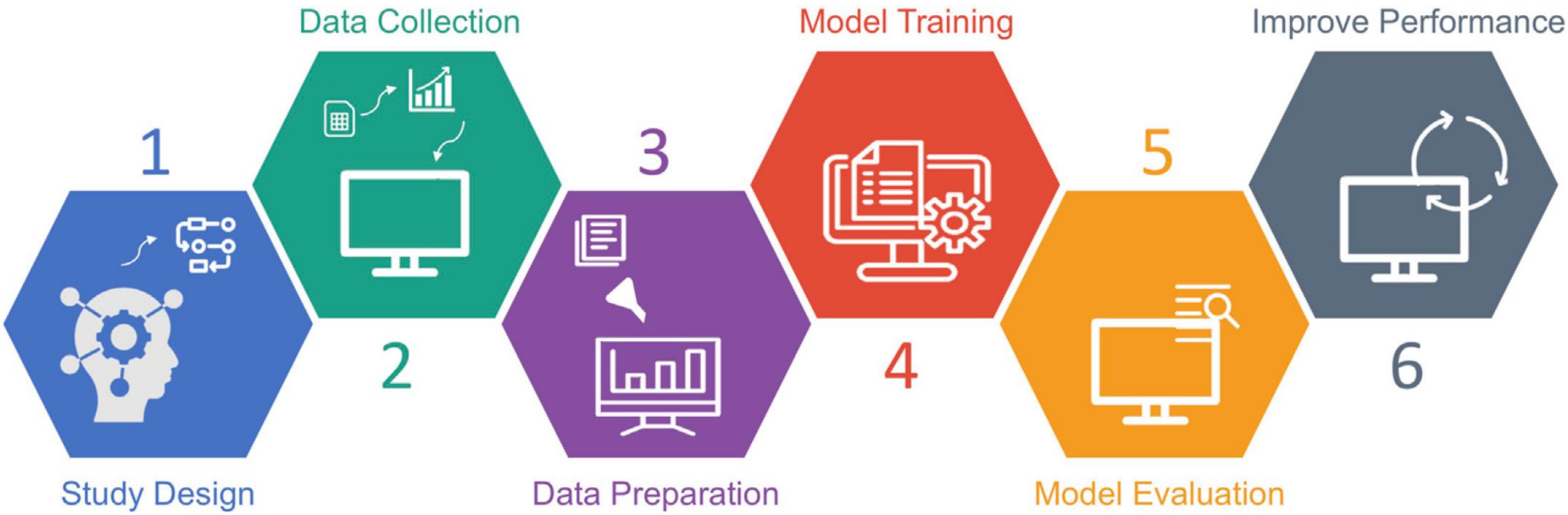
[Link to source]

Common Clustering Algorithms

Link to source

Clustering Algorithm Type		Clustering Methodology	Algorithm(s)
	Centroid-based	Cluster points based on proximity to centroid	KMeans KMeans++ KMedoids
	Connectivity-based	Cluster points based on proximity between clusters	Hierarchical Clustering (Agglomerative and Divisive)
	Density-based	Cluster points based on their density instead of proximity	DBSCAN OPTICS HDBSCAN

ML Workflow



Link to source

Resources

- Cheat Sheets of Machine Learning

https://medium.com/machine-learning-in-practice/cheat-sheet-of-machine-learning-and-python-and-math-cheat-sheets-a4afe4e791b6