

# Machine Learning (ML) Workshop



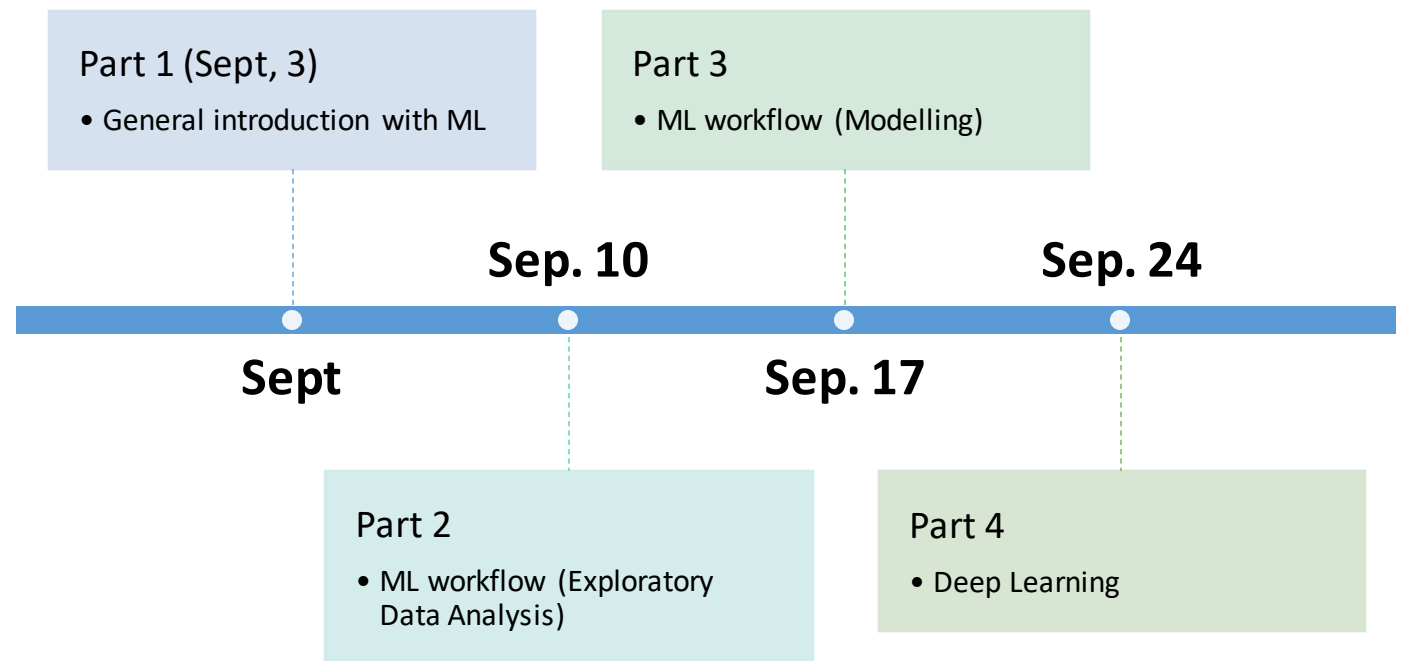
Dr Sara Soltaninejad  
Fall 2021

# Who am I?

- Sara Soltaninejad
- PhD in MRC, Computing Science Department, UofA, 2016-2020
  - PhD Thesis: Intelligent Parkinson's Disease Classification and Progress Monitoring Using Non-Invasive Techniques
  - Supervisors:
    - Dr Anup Basu
    - Dr Irene Cheng
- ML Developer AltaML, 2020-Now

# ML Workshop Outline

---





# To get you started on ML - Part 1

# What is Machine Learning (ML)?

*Machine Learning is the science of getting computers to learn and act like humans do, and improve their learning over time in autonomous fashion, by feeding them data and information in the form of observations and real-world interactions.*

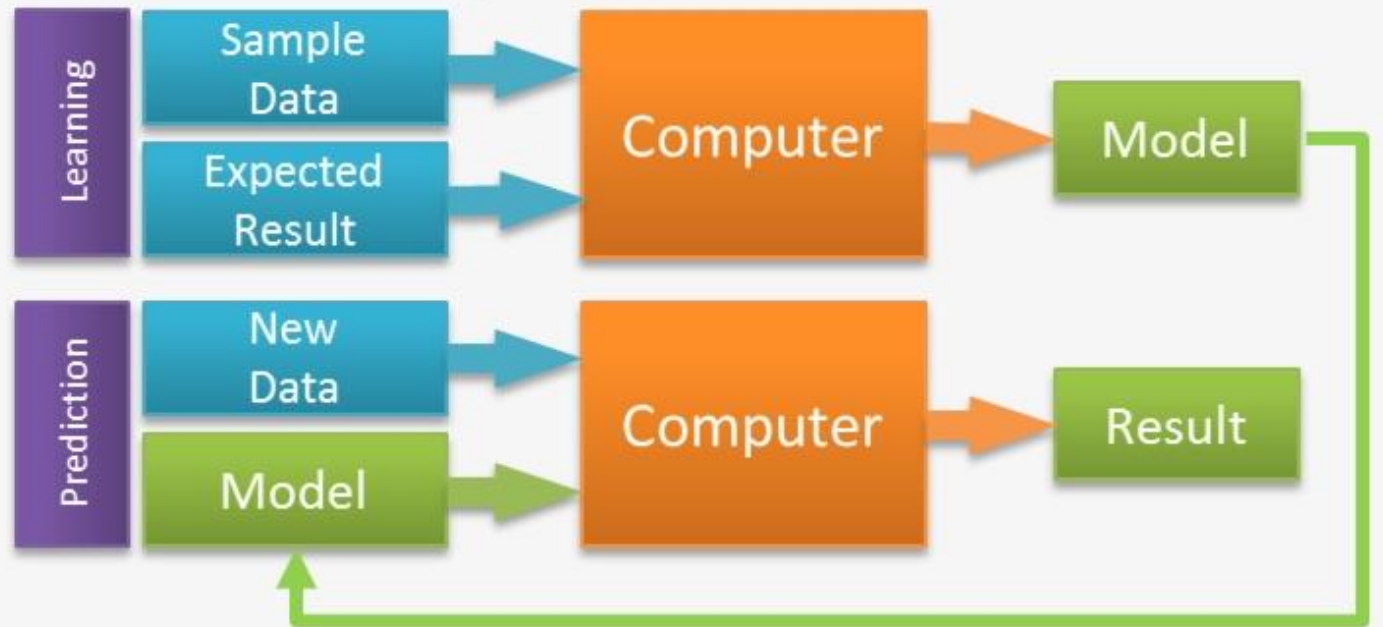


# ML vs Traditional Programming

## Traditional modeling:

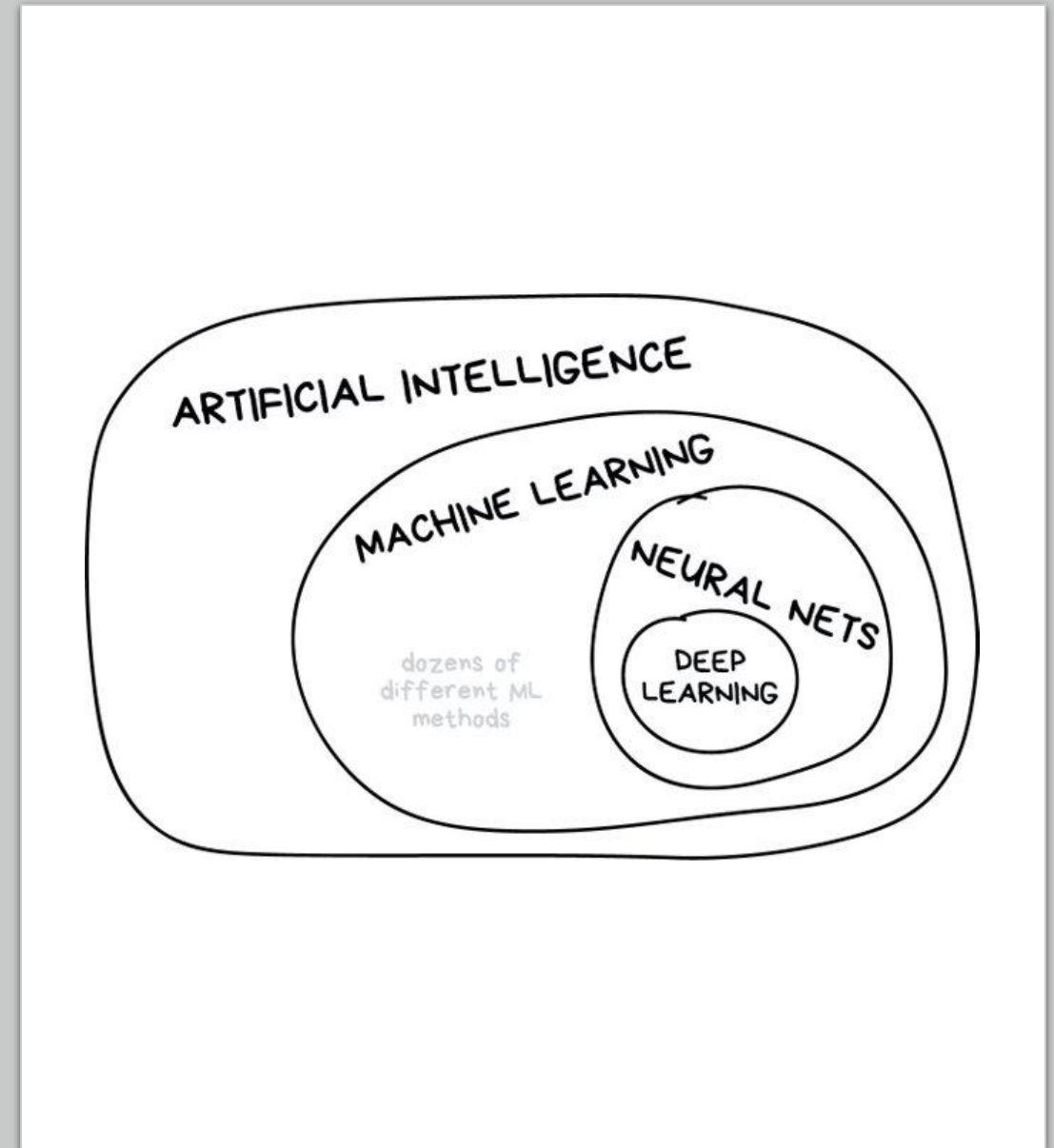


## Machine Learning:



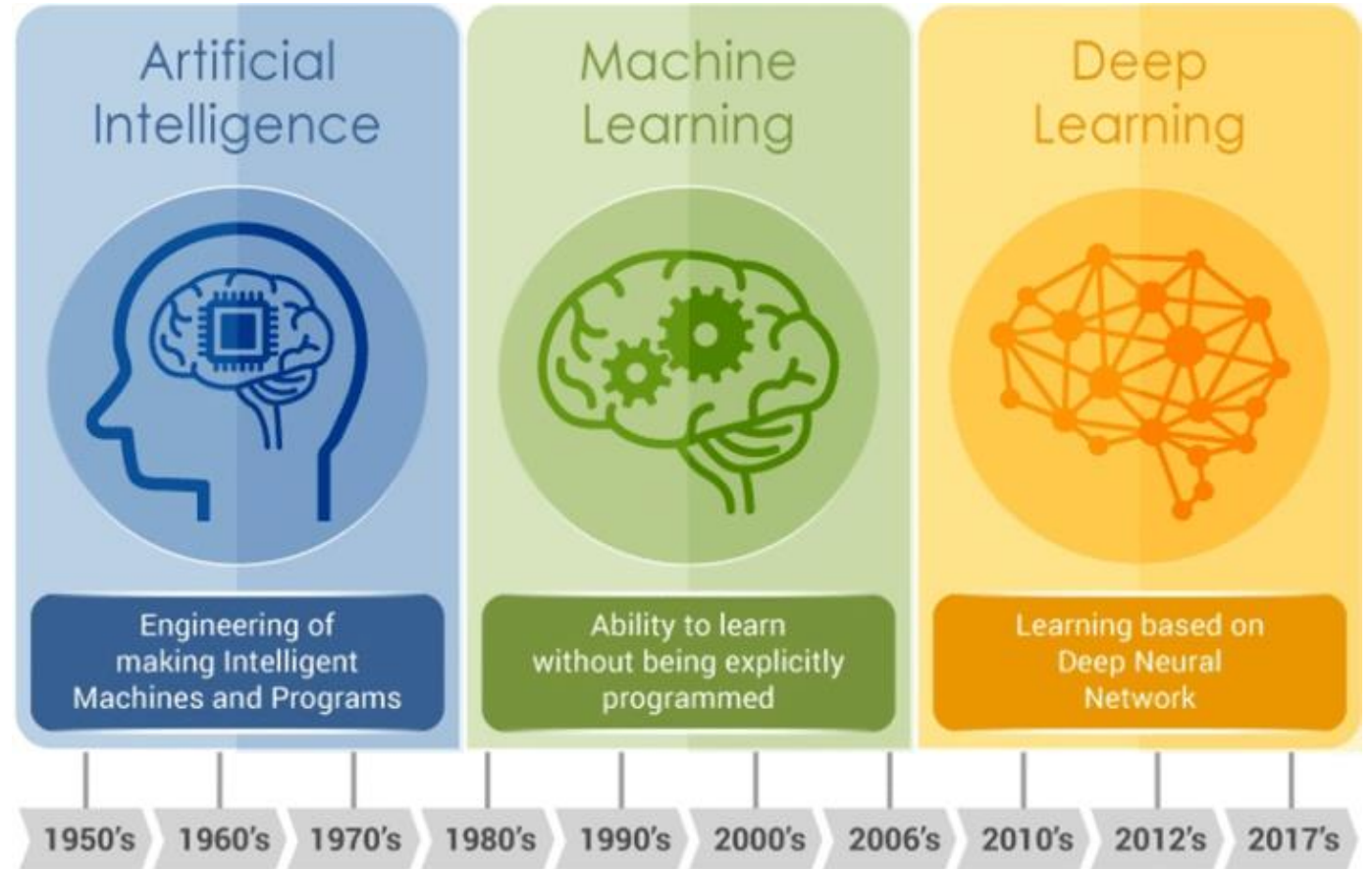
# ML and Artificial Intelligent (AI)

- While machine learning is based on the idea that machines should be able to learn and adapt through experience, AI refers to a broader idea where machines can execute tasks "smartly."
- Artificial Intelligence applies machine learning, deep learning and other techniques to solve actual problems.
- The general rule is to compare things on the same level. That's why the phrase "will neural nets replace machine learning" sounds like "will the wheels replace cars".



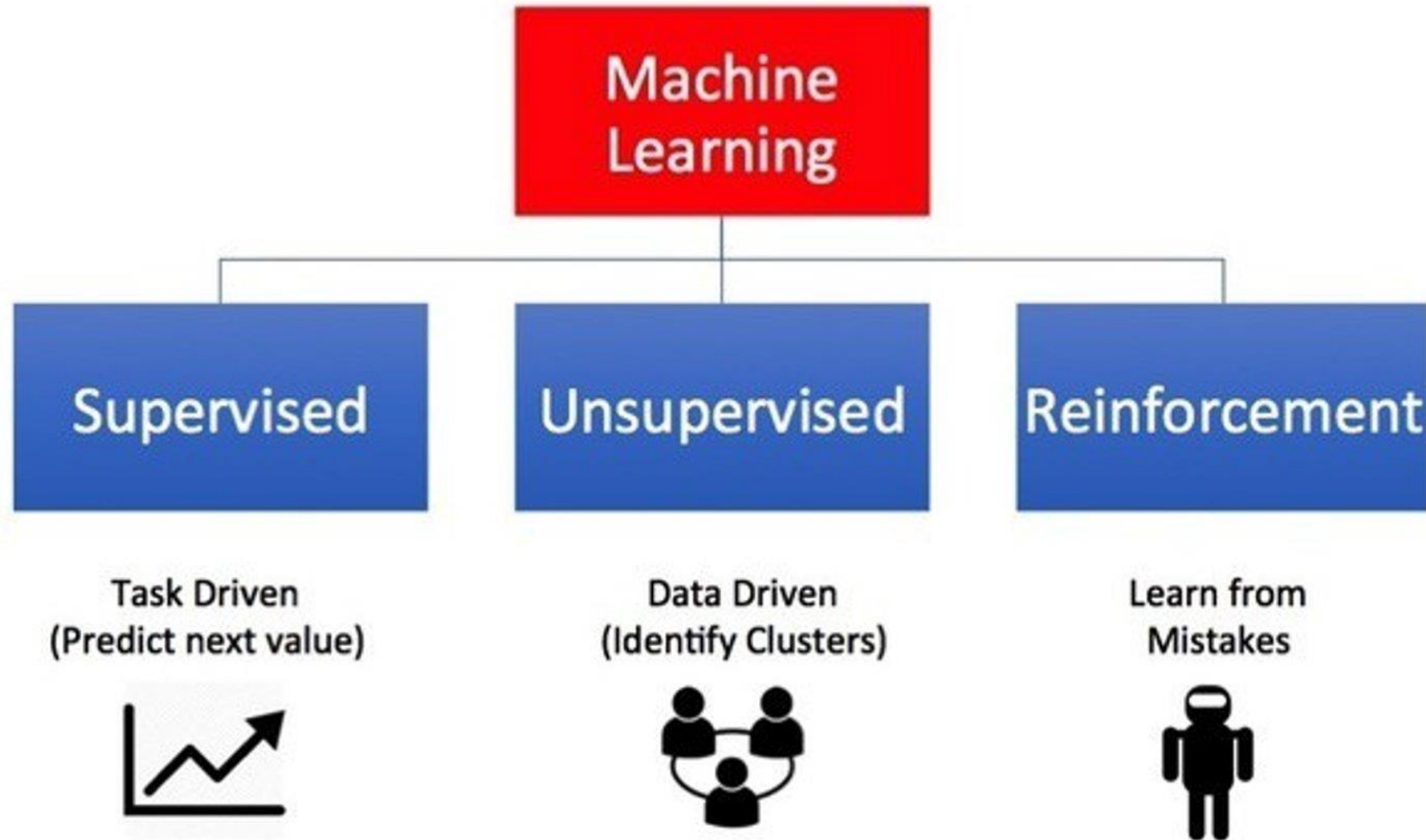


# ML History



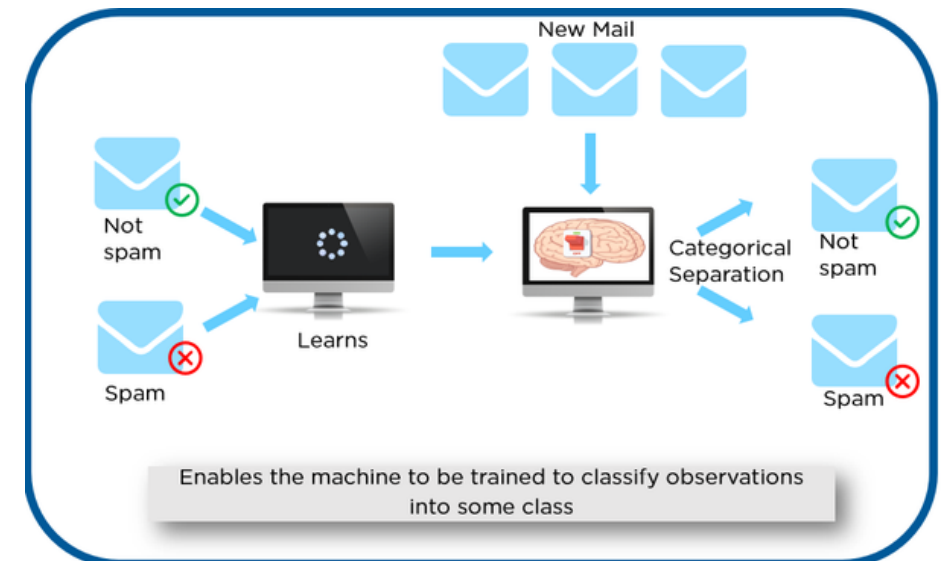


# ML Types



# ML: Supervised

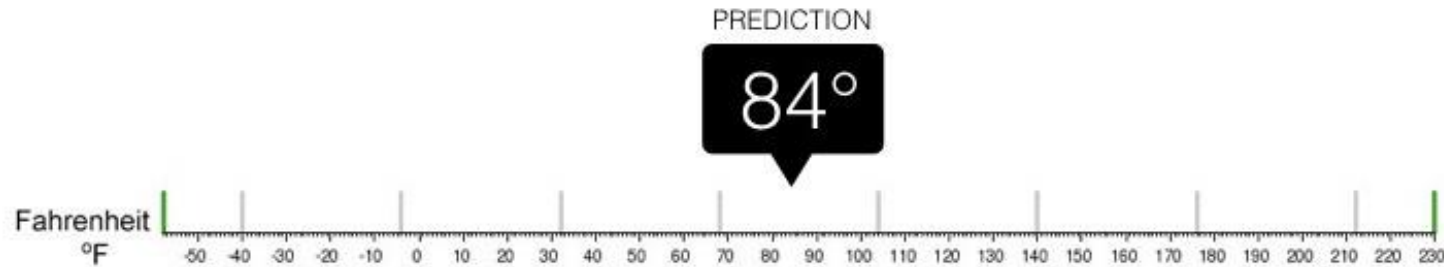
- Supervised learning is the most popular paradigm for machine learning.
- It is very similar to teaching a child with the use of flash cards.
- Supervised learning starts with training data that are tagged with the correct answers (target values). The result of training is a model with a tuned set of weights, which can predict answers for similar data that haven't already been tagged.





## Regression

What is the temperature going to be tomorrow?



## Classification

Will it be Cold or Hot tomorrow?

COLD

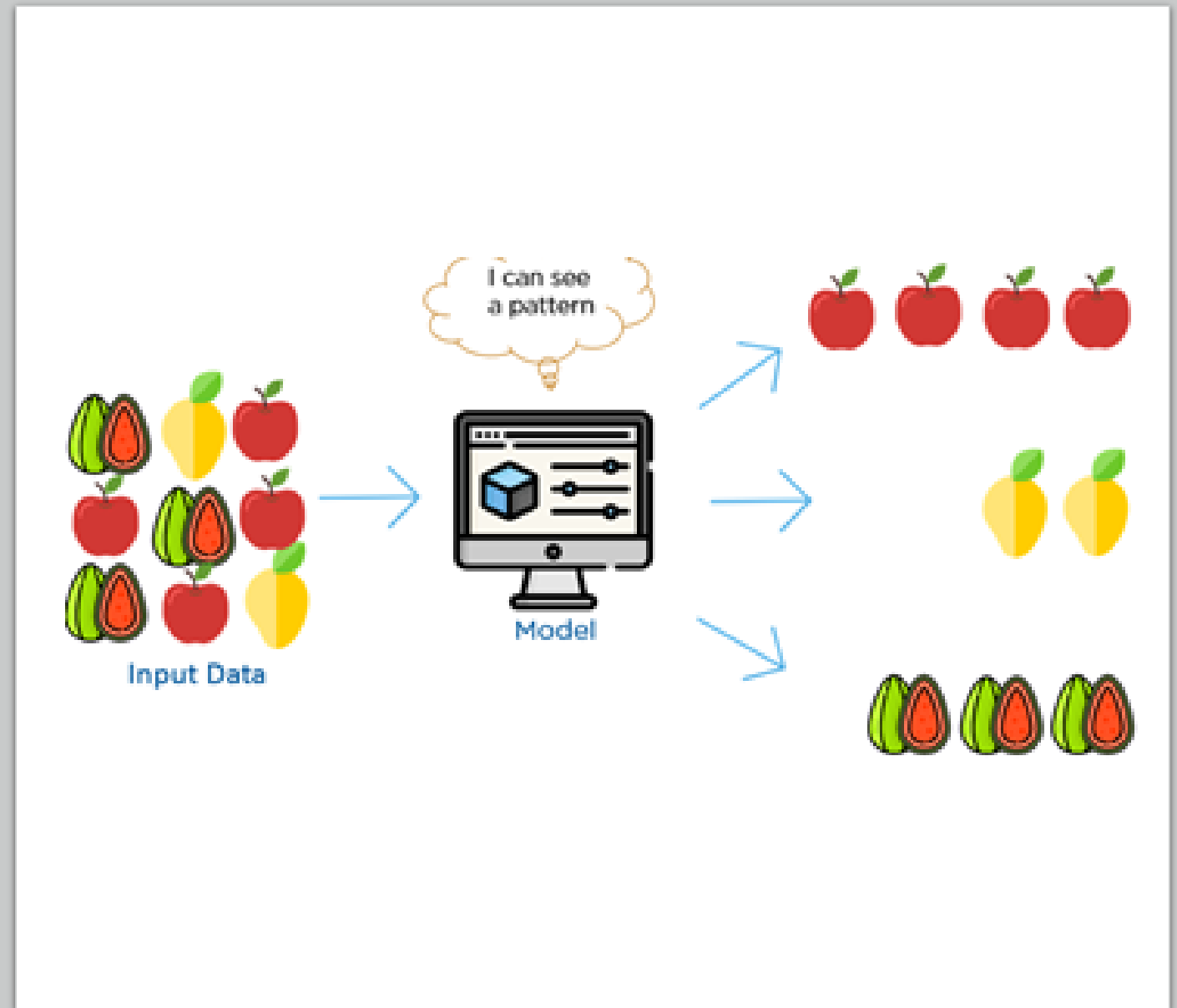
PREDICTION

HOT

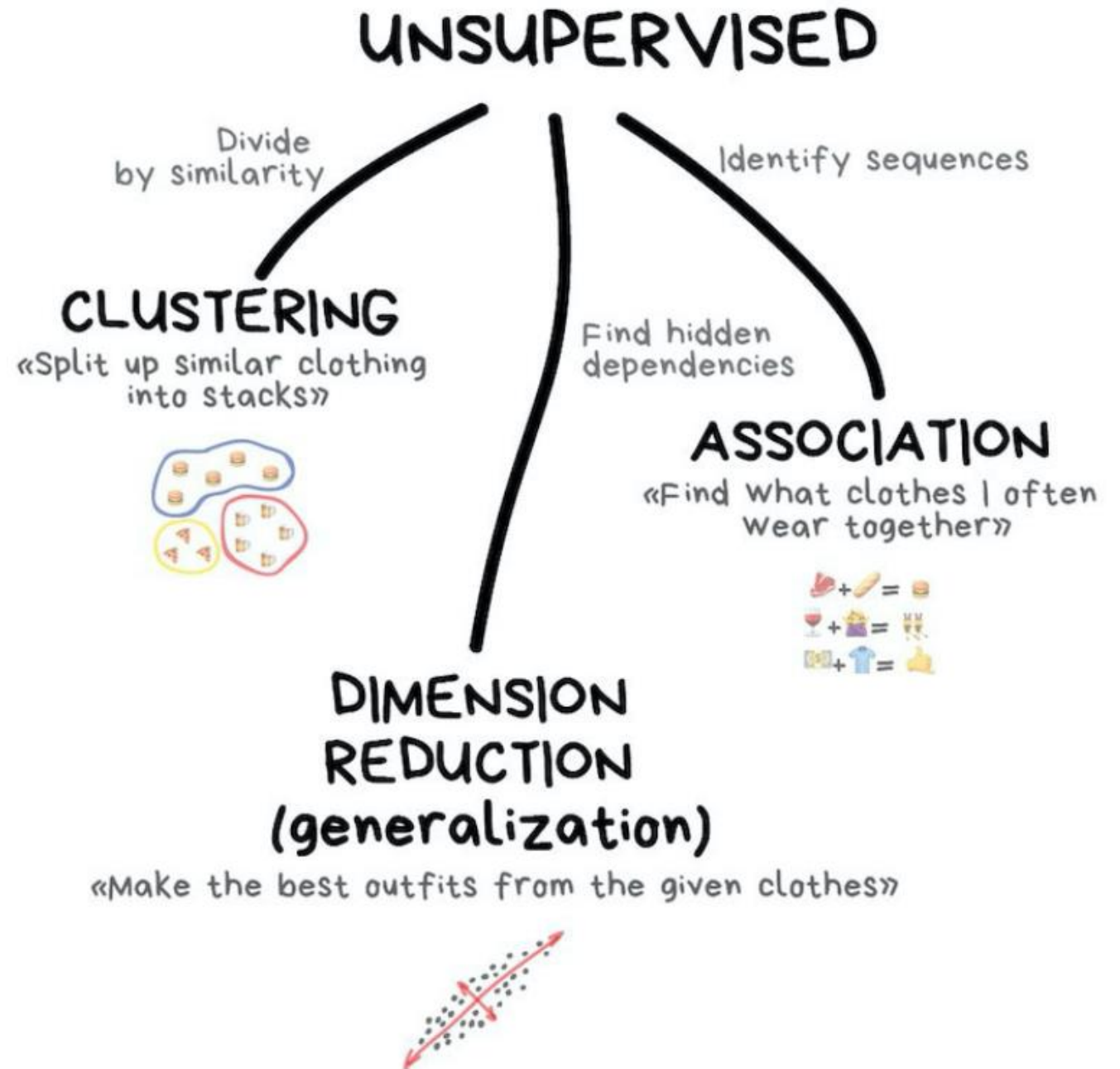
ML:  
Supervised

# ML: Unsupervised

- Overwhelming majority of data in this world is unlabeled.
- Unsupervised learning is very much the opposite of supervised learning. It features no labels.
- Instead, our algorithm would be fed a lot of data and given the tools to understand the properties of the data.

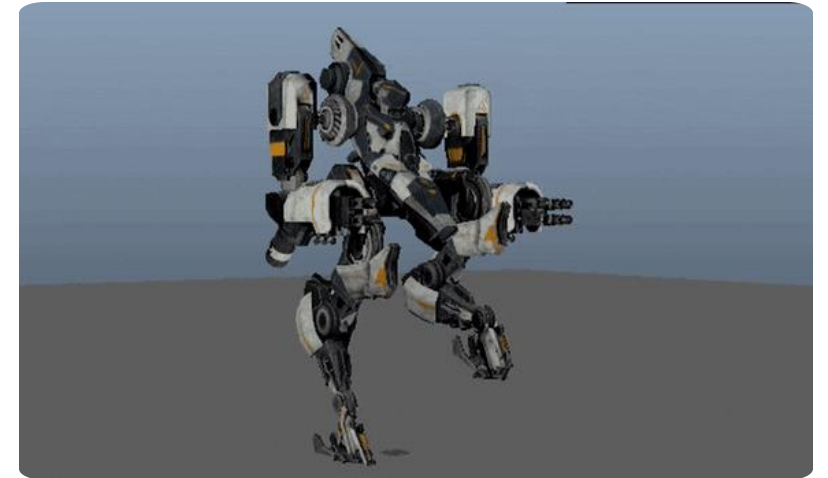


# ML: Unsupervised



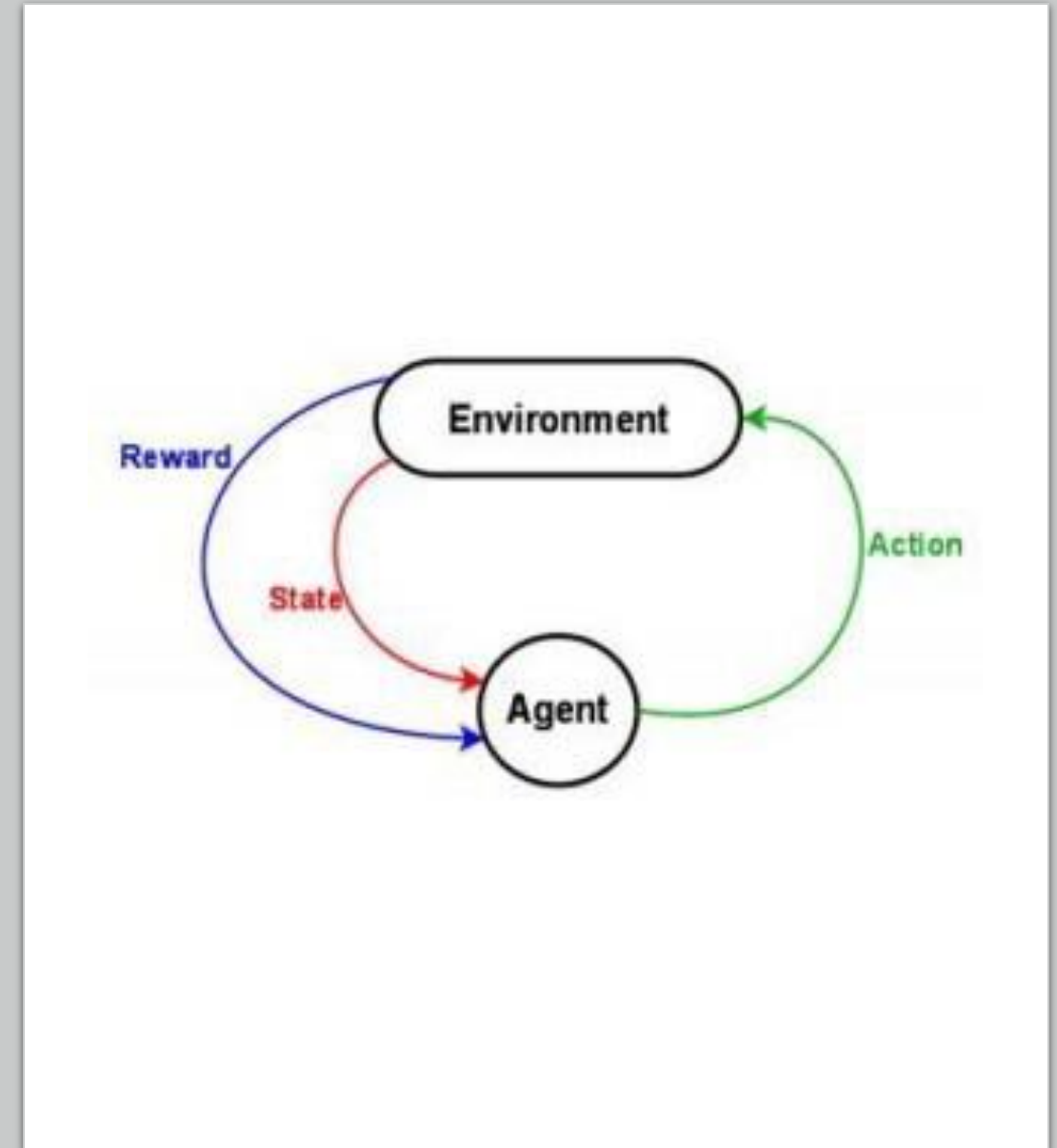
# ML: Reinforcement Learning (RL)

- As we provide some sort of signal to the algorithm that associates good behaviors with a positive signal and bad behaviors with a negative one, we can reinforce our algorithm to prefer good behaviors over bad ones. Over time, our learning algorithm learns to make less mistakes than it used to.
- Place a reinforcement learning algorithm into any environment and it will make a lot of mistakes in the beginning.
- Reinforcement learning is very behavior driven. It has influences from the fields of neuroscience and psychology.



# RL vs Supervised and Unsupervised Learning

- RL enables an agent to learn in an interactive environment by trial and error using feedback from its own actions and experiences.
- Though both supervised and RL use mapping between input and output, unlike supervised learning where feedback provided to the agent is correct set of actions for performing a task, reinforcement learning uses rewards and punishment as signals for positive and negative behavior.
- As compared to unsupervised learning, RL is different in terms of goals. While the goal in unsupervised learning is to find similarities and differences between data points, in reinforcement learning the goal is to find a suitable action model that would maximize the total cumulative reward of the agent.

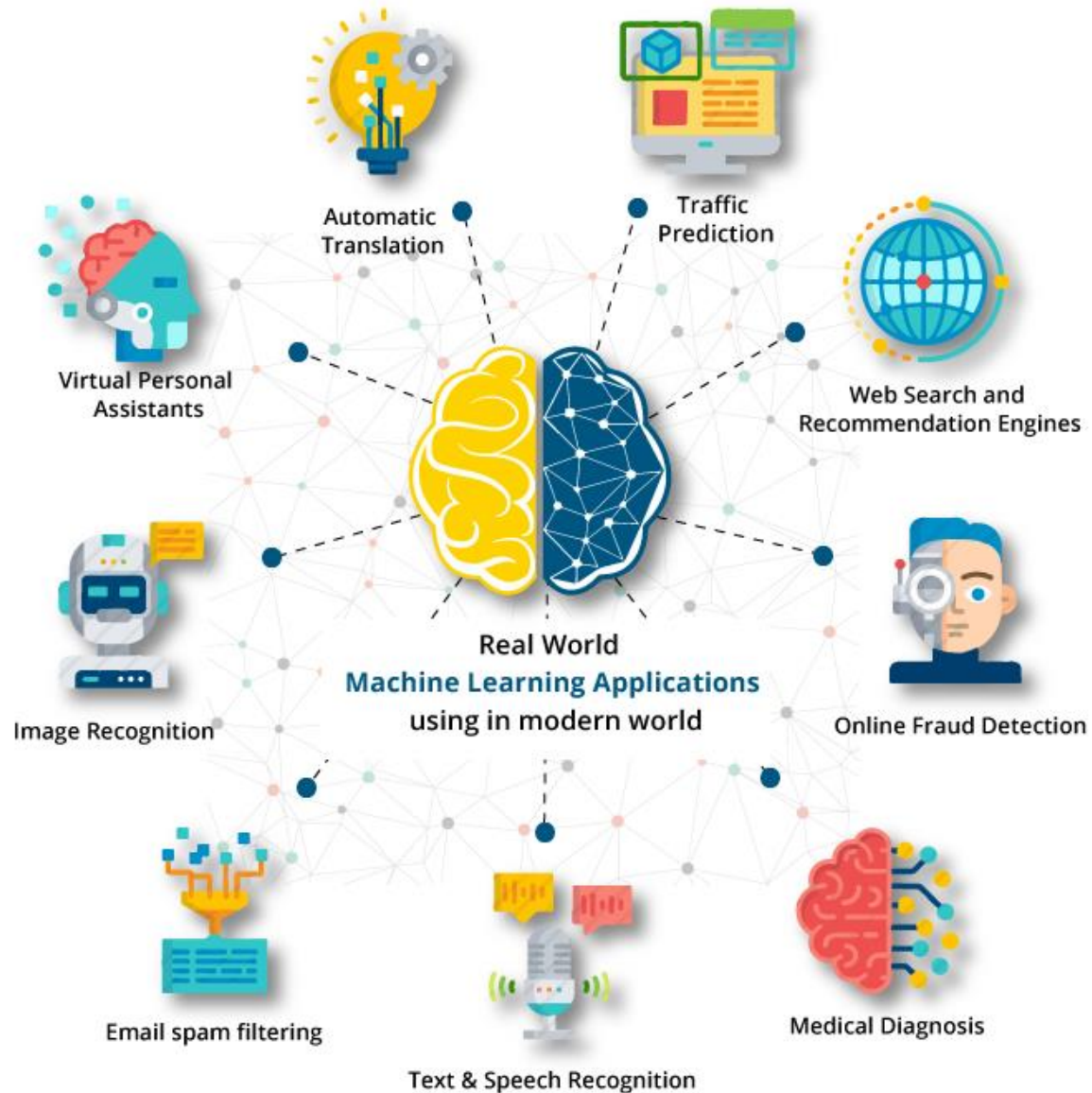




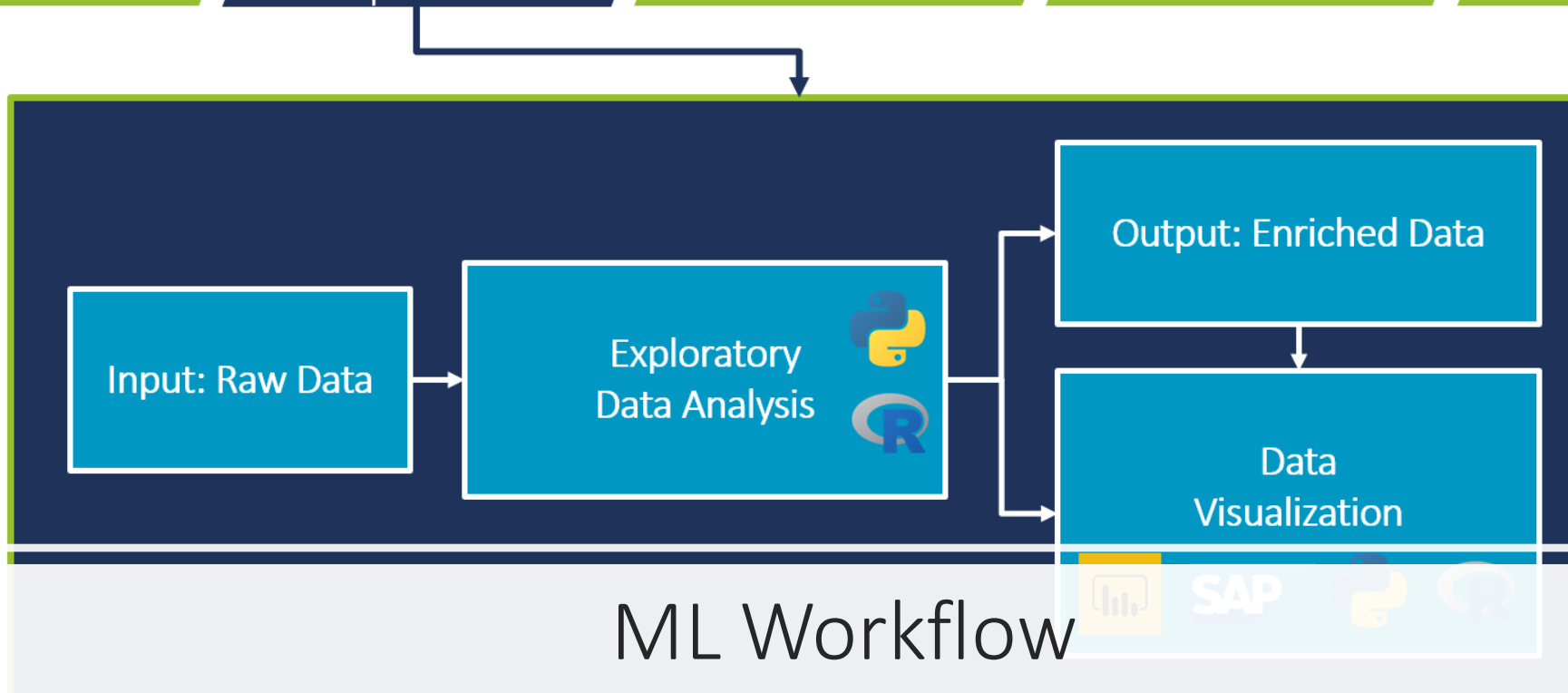


# RL Example

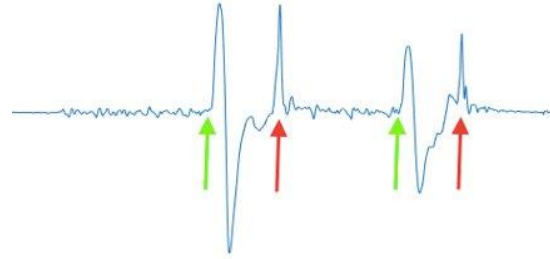
Reinforcement learning is proving to be capable of solving dynamic digital marketing problems so that high-quality recommendations can be delivered that resonate with customers' specific preferences, needs, and behavior.



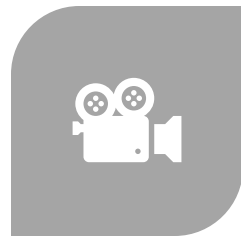
# ML Applications



# ML: Input Data



IMAGE



VIDEO



SIGNAL



TEXT



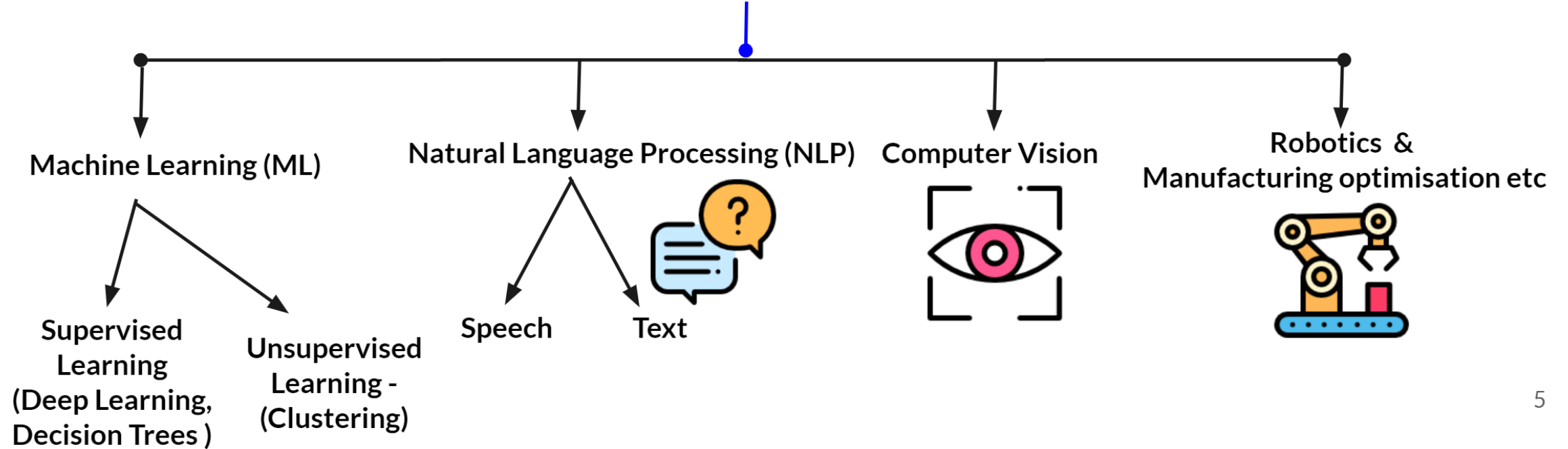
DATA FRAME



...shehasnootherneurologicsymptomsnonumbnessortinglingshedeniesanyvisualcha  
ngespastmedicalhistorygallbladderremovalpastssurgicalhistorydiabetesrheumatoida  
rthritishypertensiongerdandhypothyroidismmedicationsadvaairalbuterolallopurinol  
aspirinclobetasolfolicacidfosamaxdevoxyllisinoprilmetforminomeprazoleplaquenilp  
rednisonetestosteroneverapamilallergiesnoknowndrugallergiessocialhistorythepati  
entismarriedwithchildshedoesnotsmokeshe doesnotdrinkshe doesnotuserecreation  
aldrugsshe weighspoundsandisinchestallfamilyhistorynegativeforbrainaneurysmo  
theraneurysmitwasalsonegativeforheartdiseasehighcholesterolandhypertensionand  
negativefordiabetesreviewofsystemsthepatientispositiveforhypertensionswellint  
behandsorfeetlegpainwhilewalkingasthma pneumonia shortnessofbreathgastritisulc  
ersdiabetesthyroiddiseaseurinarytractinfectionsandthosesymptomsrelatedtothe  
pre sentillnessthe detailsofthereviewofsystemswere reviewedwiththepatientandareinclu  
dedintheneurosurgicalhealthhistoryquestionnairepainthepatienthasepisodicjointpa  
inthatistreatedwithtylenolthepatientdoesnothave anynutritionalconcernsshe doesno  
thave anysafetyconcernsphysical examination...



## Artificial Intelligence





# ML: Data Frame

**DataFrame object**

Label index  
(country code)

Column names

	Country	Popu	Percent
IT	Italy	61	0.83
ES	Spain	46	0.63
GR	Greece	11	0.15
FR	France	65	0.88
PO	Portugal	10	0.14

Data  
(different type in each column)

The screenshot shows a Microsoft Excel spreadsheet with a data table. The table has multiple columns, including 'Year', 'Country', 'Population', and 'Percentage'. The data is organized in rows, with each row representing a specific country and year. The spreadsheet is displayed in a standard Excel interface with a ribbon at the top and a grid of cells below.

# Exploratory Data Analysis (EDA)

EDA helps us understand various facets of our data. In this step, we analyze different attributes of data, uncover interesting insights, and even visualize data on different dimensions to get a better understanding.

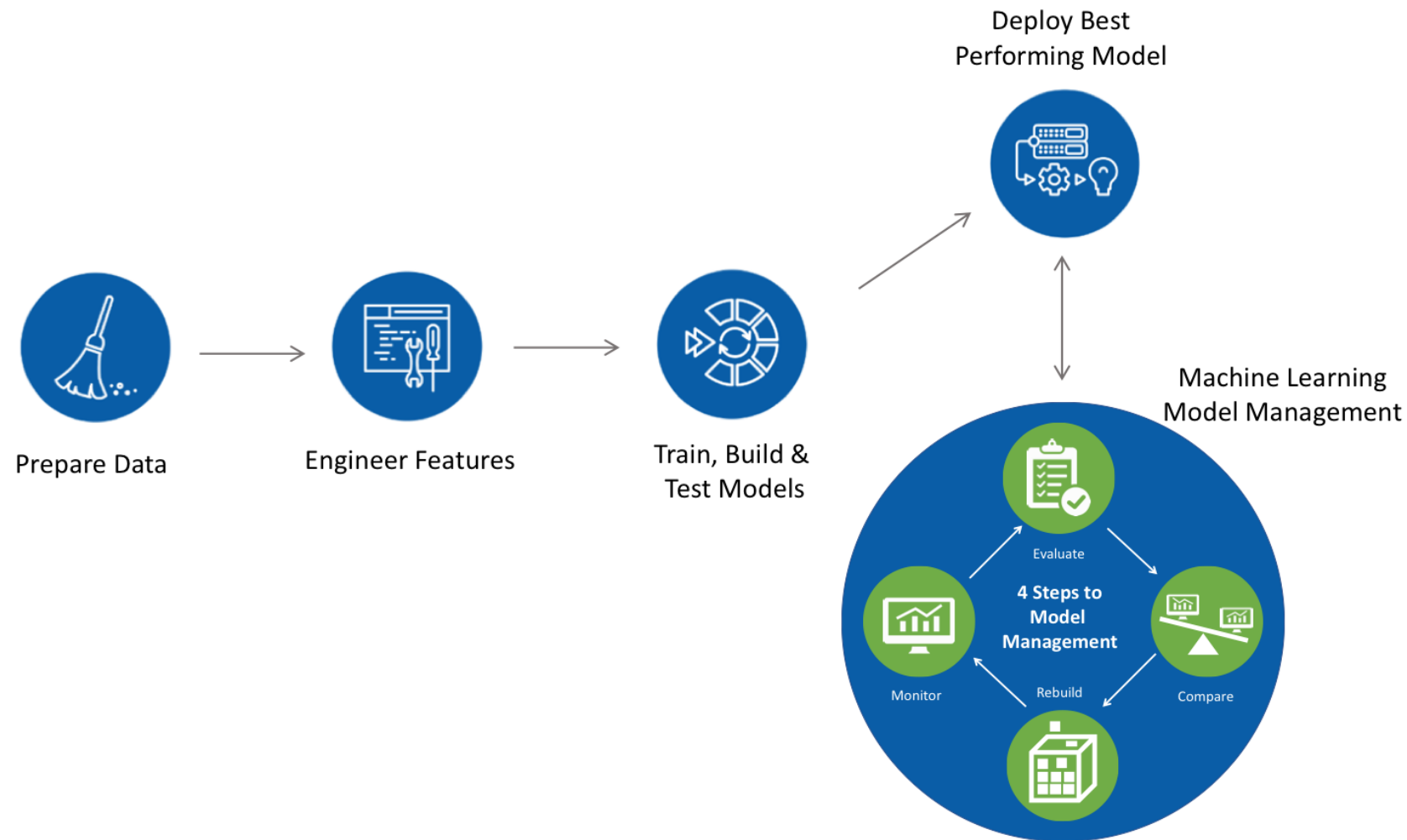


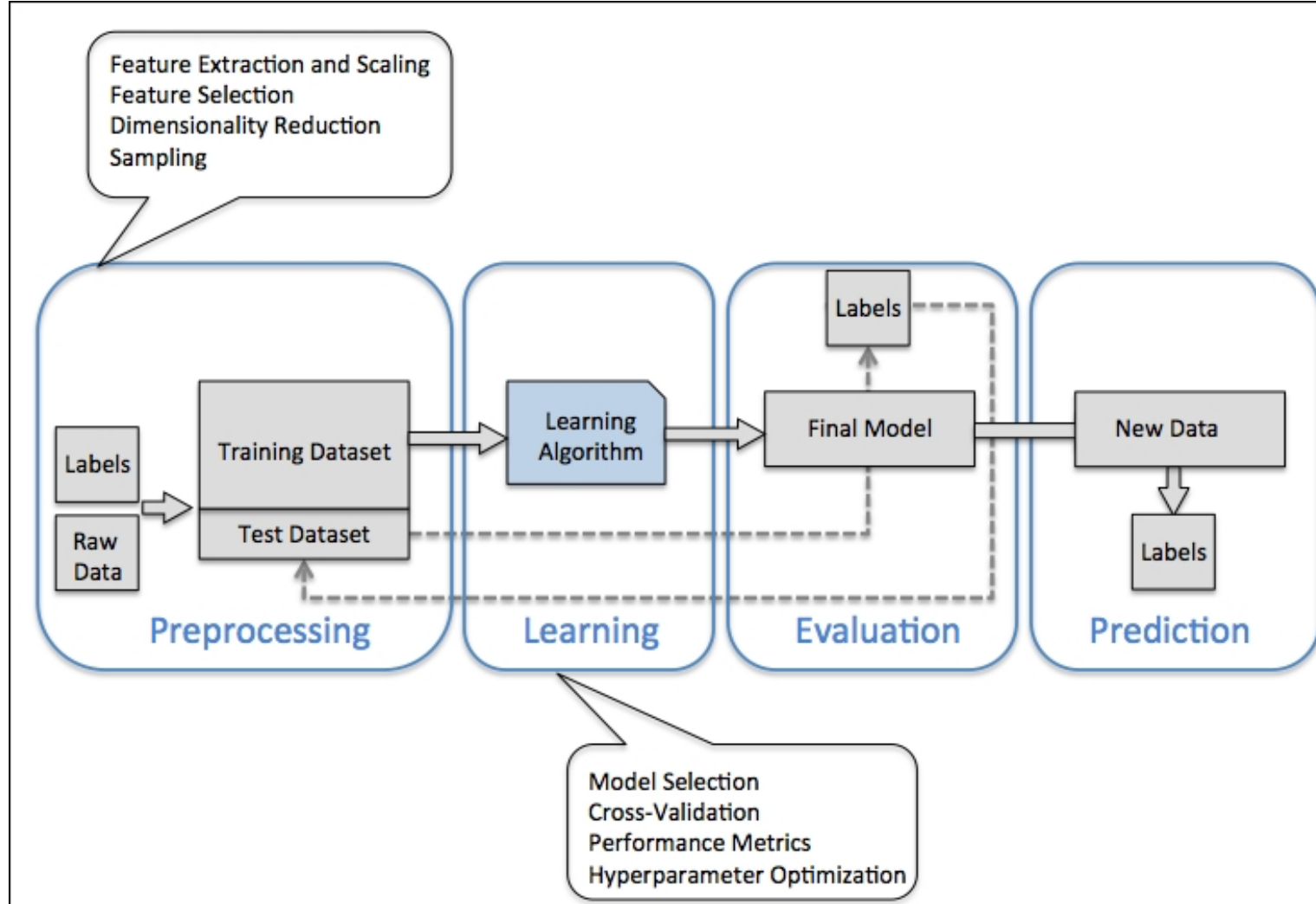


# ML: Data Preprocessing



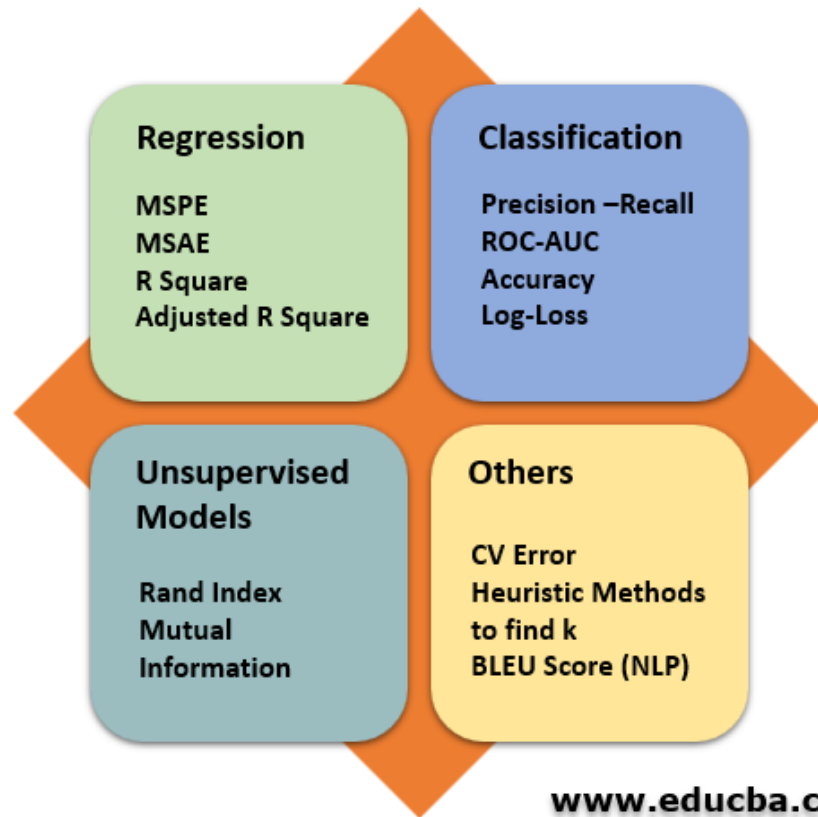
# Modelling





How to  
choose the  
best  
model?

# ML: Evaluation



		"Golden Standard" (Real Truth Values)		
		Positive	Negative	
Observed	Predicted positive	True Positive	False Positive (Type 1 error)	Precision
	Predicted Negative	False Negative (Type 2 error)	True Negative	
		Recall/ Sensitivity	(Specificity)	

$$\text{Accuracy} = \frac{TP + TN}{TN + FP + FN + TP}$$

$$\text{Precision} = \frac{TP}{FP + TP}$$

$$\text{Recall} = \frac{TP}{FN + TP}$$

$$\text{F1-score} = 2 * \frac{\text{precision} * \text{recall}}{\text{precision} + \text{recall}}$$



# Quiz



- What is the Difference Between Supervised and Unsupervised Machine Learning?
- Can you think of any RL application which we did not already mentioned?
- Specify the types of these machine learning problems:
  - Soccer player tracking (data is a group of soccer video games with no tagged player)
  - Heart Disease Diagnosis (data is a group of data frames including different medical features for different subjects who can either be healthy or having heart disease)
  - Traffic sign detection (data is a group of image of streets with traffic signs which the signs are tagged on the images)
- When should we use classification over regression?

# References

- <https://emerj.com/ai-glossary-terms/what-is-machine-learning/>
- <https://www.slideshare.net/kongqk/machine-learning-with-scikitlearn-77235584>
- <https://machinelearningmastery.com/types-of-learning-in-machine-learning/>
- <https://searchenterpriseai.techtarget.com/definition/supervised-learning>
- <https://www.oreilly.com/radar/reinforcement-learning-explained/>
- <https://towardsdatascience.com/reinforcement-learning-an-introduction-to-the-concepts-applications-and-code-ced6fbfd882d>
- <https://medium.com/app-affairs/9-applications-of-machine-learning-from-day-to-day-life-112a47a429d0>
- [https://subscription.packtpub.com/book/big\\_data\\_and\\_business\\_intelligence/9781788831307/1/ch01lvl1sec13/standard-ml-workflow](https://subscription.packtpub.com/book/big_data_and_business_intelligence/9781788831307/1/ch01lvl1sec13/standard-ml-workflow)
- <https://medium.com/free-code-camp/how-to-get-a-grip-on-cross-validations-bb0ba779e21c>
- <https://towardsdatascience.com/cross-validation-in-machine-learning-72924a69872f>
- <https://towardsdatascience.com/overfitting-underfitting-and-the-bias-variance-tradeoff-83b42fb11efb>