

Machine Learning

Erick Cuenca

Research professor at Yachay Tech University

<https://erickedu85.github.io>



Introduction

Overview

- Introduction to Artificial Intelligence AI
 - Common questions of AI
 - Characteristic of AI
 - Problems
 - Implications
- Introduction to Machine Learning ML
 - ML types
 - Supervised learning
 - Unsupervised learning
 - Reinforced learning

Common questions about AI

Will a robot take my job?

- chatbot
- Virtual assistance
- ...



Supporting transactions 2017: <https://www.youtube.com/watch?v=fAjbPdoQjVk>

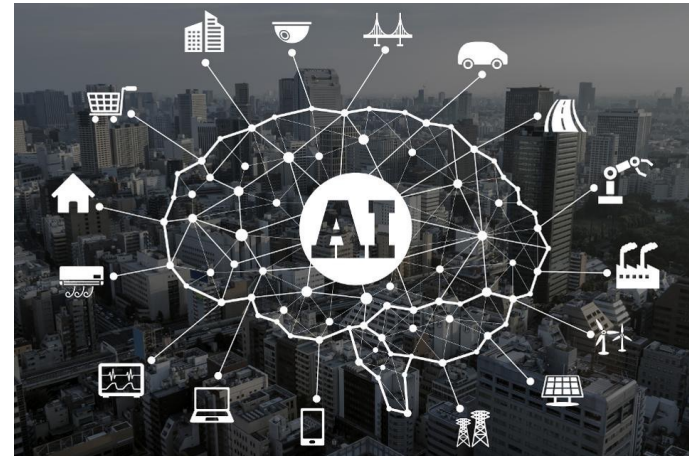
AI assistant calls businesses 2018: <https://www.youtube.com/watch?v=D5VN56jQMWM>

The rise of robots 2023: <https://www.youtube.com/watch?v=I1w6xzBN7mA>

Common questions about AI

How is AI changing our lives?

- Quality of living: health care, ...
- Time-saving: automate processes, ...
- Entertained: Spotify, Netflix, games, ...
- Transportation: autonomous vehicles



<https://www.youtube.com/watch?v=Cc4KZ8KUxCE>

What is AI?

- It is a scientific discipline: concepts + problems + methods
- AI definition is constantly redefined when some topics emerge
- What seems **easy** (*for humans*) is actually **hard** (*for machines*)?
 - **Grabbing objects:** Look around and pick up and object
 1. Use your eyes to scan your surrounding
 2. Figured ou where are some suitable objects for picking up
 3. Chose one of them and planned a trajectory for your hand reach the object
 4. Move your hand by contractic many muscles
 5. Squeeze the object with just the enough force to hold it
 - **Grabbing objects** is easy for humans (childhood practice)

Robots grabbing objects 2016: <https://www.youtube.com/watch?v=V05SuCSRAtg>

Google's AI robot 2023: https://www.youtube.com/watch?v=j6O_uePUKKI

What is AI?

- What seems **hard** (for humans) is actually **easy** (for machines)
 - Playing chess and solving mathematical exercises can seem very difficult
 - Computers follow rules and compute many alternative move sequences at a rate of billions of computations a second.
 - In 1997, Deep Blue beat the reigning human champions in chess Kasparov (https://en.wikipedia.org/wiki/Deep_Blue_versus_Garry_Kasparov)

What is AI?

Artificial *live*-forms that can surpass human intelligence through their ability to process data



AI characteristics

- Is a chess-playing algorithm more intelligent than a spam filter?
- Is a recommendation system more intelligent than a self-driving car?

Autonomy

The ability to perform tasks in complex environments without constant guidance by a user

Adaptivity

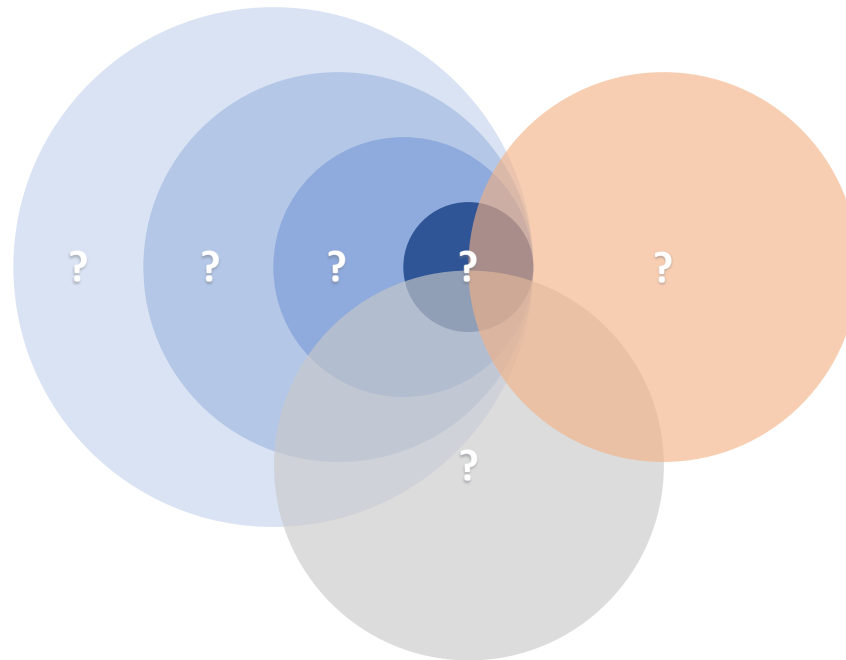
The ability to improve performance by learning from experience

“Intelligence”

- AI systems cannot be compared on a single dimension of intelligence
- AI is narrow: solving one problem tells us nothing about the ability to solve a different one

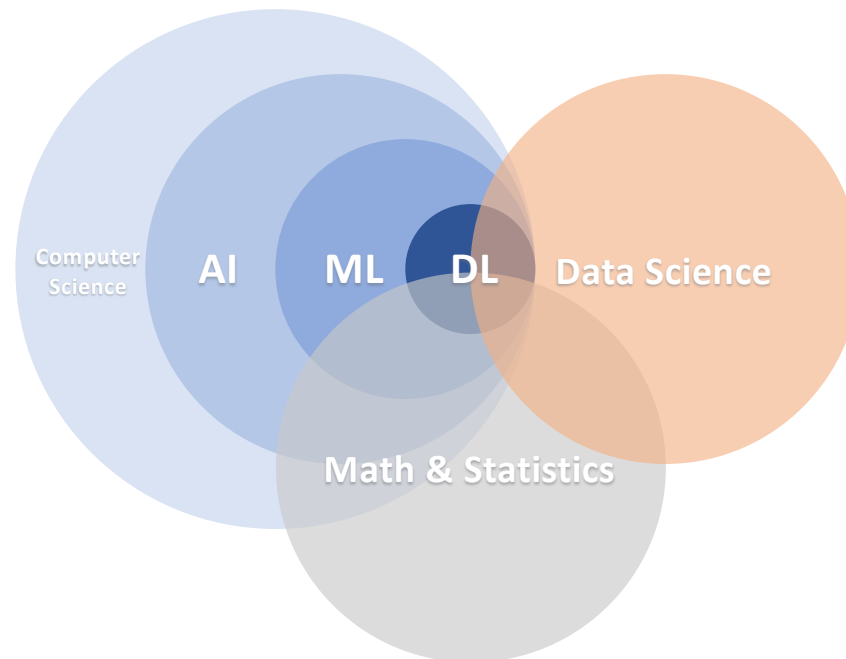
AI subfields

Deep Learning | Machine Learning | Artificial Intelligence |
Data Science | Computer Science | Math & Statistics



AI subfields

Deep Learning | Machine Learning | Artificial Intelligence |
Data Science | Computer Science | Math & Statistics



AI subfields

Artificial Intelligence AI

Machine Learning ML is a subfield of AI

ML systems improve their performance on a given task through experience or data

Deep Learning DL is a subfield of ML

It involves complex mathematical models and leverages the increased computing power of modern hardware

Data Science

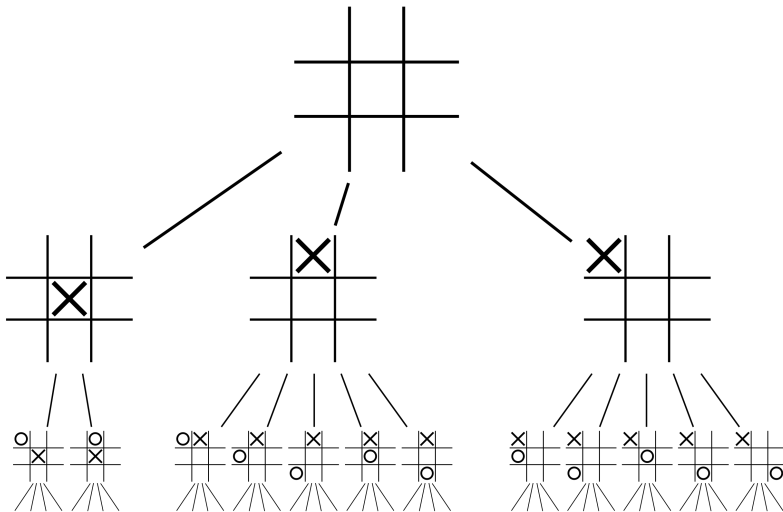
Statistic + mathematics + computer Science:

- algorithms, data storage, processing, web application development, etc.

It is a practical discipline: business or science

AI and games

The first two plies of the **game tree** for tic-tac-toe

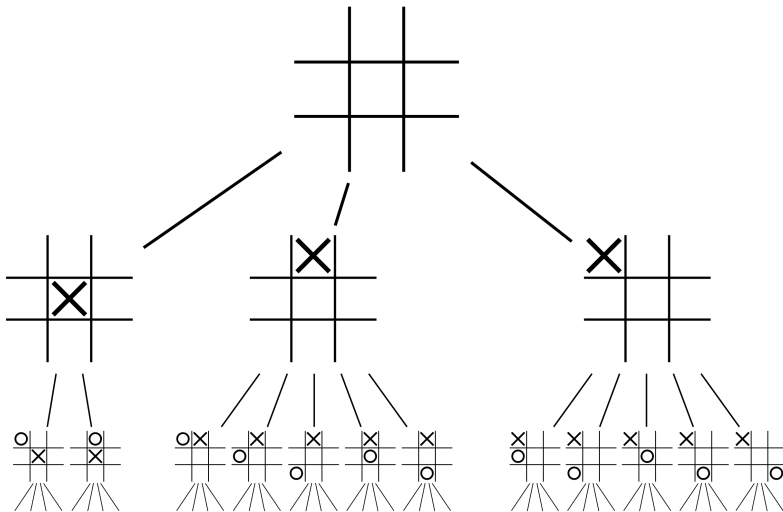


Game tree

- Nodes are positions in a game
- Edges are moves
- Combining with heuristics
 - Minimax algorithm: a recursive algorithm for choosing the optimal move in a two-player game (tic tac toe, chess, Go)

AI and games

The first two plies of the **game tree** for tic-tac-toe



Deterministic games

- There is not randomness in the game
 - E.g., Tic-Tac-Toe, Chess, Go
- Based on the assumption that both players always choose the best mover for themselves, and what is best for one is worst for the other

<https://www.google.com/search?q=tic+tac+toe>

Machine Learning ML

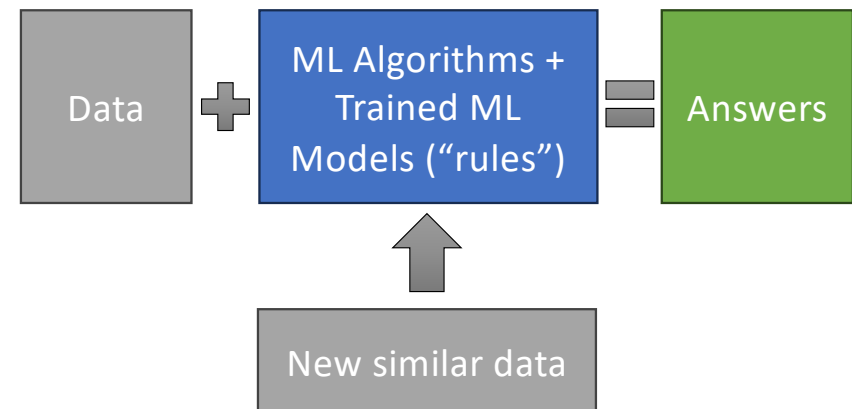
“Programming computers to **learn from experience** should eventually **eliminate the need for much of this detailed programming effort**”

Arthur Samuel (1959), Computer Scientist

Classical programming



Machine Learning



Machine Learning ML

ML = Statistics + Probability Theory + Mathematics + Computer Science

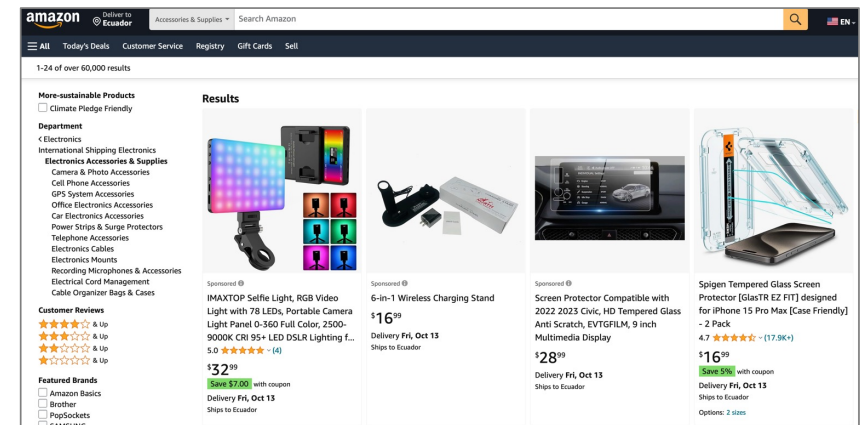
Roles of sciences:

- Statistics & Mathematics: ML uses statistical theory to building mathematical model
- Computer science:
 - Data:
 - Acquire large volumes of data
 - Store and process large-scale datasets
 - Algorithms:
 - In training:
 - Efficient algorithms for solving optimization problems
 - Once a model is learned:
 - Efficient inference algorithms (with focus on time complexity)
 - Compute:
 - GPUs, high-performance processing units, hardware acceleration, and parallelization

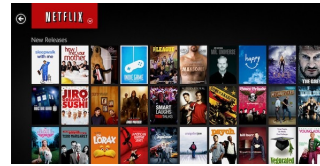


The core task of the ML is to make inferences from a sample of data

ML Applications


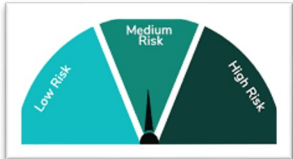
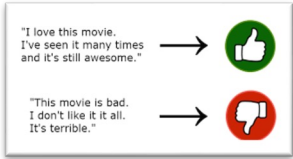
Business/ML Problem	Description	Examples
Ranking	Helping users find the most relevant thing	
Recommendation	Giving users the thing they may be most interested in	
Classification	Figuring out what kind of thing something is	
Regression	Predicting a numerical value of a thing	
Clustering	Putting similar things together	
Pattern recognition	Automated recognition of patterns and regularities in data	



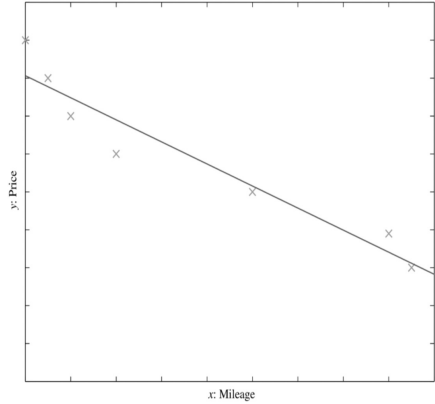
ML Applications

Business/ML Problem	Description	Examples
Ranking	Helping users find the most relevant thing	
Recommendation	Giving users the thing they may be most interested in	 <p>Based on their past history, recommend movies that customers are likely to enjoy</p>
Classification	Figuring out what kind of thing something is	
Regression	Predicting a numerical value of a thing	 <p>People who bought this product also bought ...</p>
Clustering	Putting similar things together	
Pattern recognition	Automated recognition of patterns and regularities in data	 <p>Determine who you “may know”</p>

ML Applications

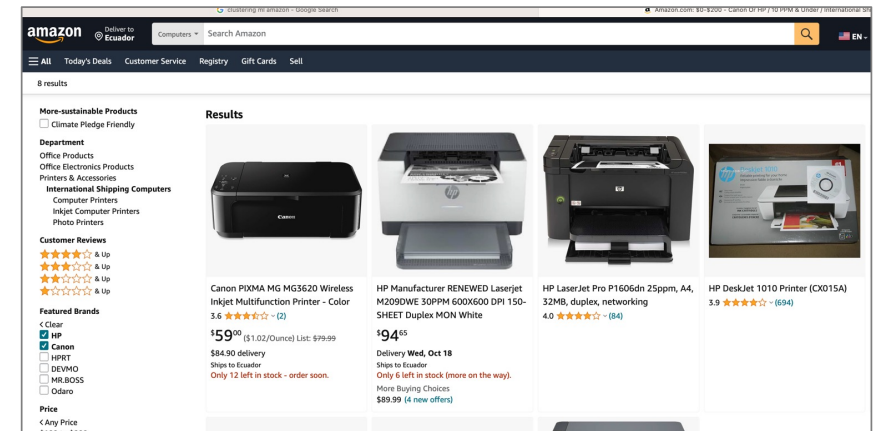
Business/ML Problem	Description	Examples
Ranking	Helping users find the most relevant thing	
Recommendation	Giving users the thing they may be most interested in	
Classification	Figuring out what kind of thing something is	
Regression	Predicting a numerical value of a thing	
Clustering	Putting similar things together	
Pattern recognition	Automated recognition of patterns and regularities in data	

ML Applications

Business/ML Problem	Description	Examples
Ranking	Helping users find the most relevant thing	 <p>We want to predict the price of a used car. Input: the car attributes: brand, year, engine capacity, mileage, ... Output: the price of the car (numeric value)</p>
Recommendation	Giving users the thing they may be most interested in	
Classification	Figuring out what kind of thing something is	
Regression	Predicting a numerical value of a thing	
Clustering	Putting similar things together	
Pattern recognition	Automated recognition of patterns and regularities in data	

ML Applications

Business/ML Problem	Description	Examples
Ranking	Helping users find the most relevant thing	
Recommendation	Giving users the thing they may be most interested in	
Classification	Figuring out what kind of thing something is	
Regression	Predicting a numerical value of a thing	
Clustering	Putting similar things together	
Pattern recognition	Automated recognition of patterns and regularities in data	



ML Applications

Business/ML Problem	Description	Examples
Ranking	Helping users find the most relevant thing	<i>Character recognition:</i> recognize characters codes from their images <ul style="list-style-type: none">Multiple classes, as many as there are characters we would like to recognize<ul style="list-style-type: none">Handwritten recognitionRead zip-codes on envelopes
Recommendation	Giving users the thing they may be most interested in	<i>Face recognition:</i> recognize people from images <ul style="list-style-type: none">More complex than character recognition since there are more classes and a face is three-dimensional, so differences in pose and light.
Classification	Figuring out what kind of thing something is	<i>Speech recognition:</i> recognize words from acoustic signal (ES, EN, FR, ...) <ul style="list-style-type: none">Complex task because different people pronouns the same word differently
Regression	Predicting a numerical value of a thing	<i>Natural Language Processing domain:</i> <ul style="list-style-type: none">Trending topics in social media: improve advertiseSentiment prediction: customer is happy or not with the product
Clustering	Putting similar things together	
Pattern recognition	Automated recognition of patterns and regularities in data	



Erick Cuenca - twitter: @erickedu85

<https://quickdraw.withgoogle.com/>

ML Applications

Why this example? Digit recognizer

Handwritten digits

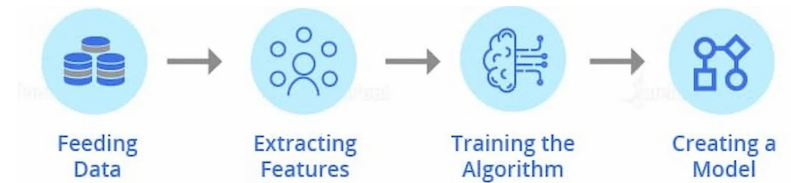


MNIST dataset. Modified National Institute of Standards and Technology

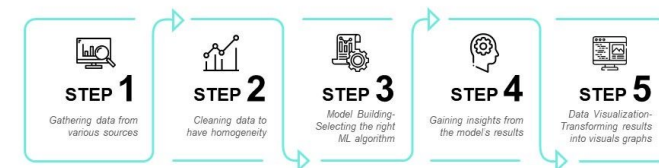
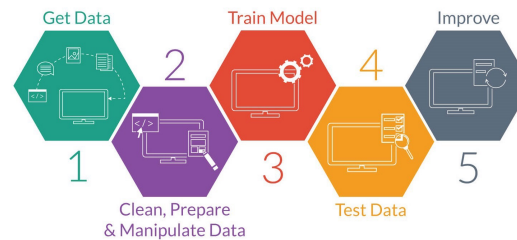
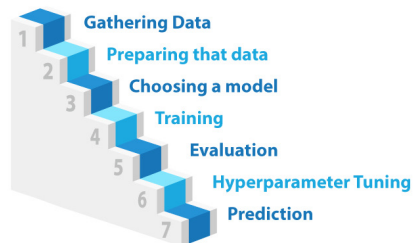
What problems to solve?

- Each image belongs to one class (0-9), but some digits are difficult to distinguish
- Build an AI model that can take an image as input and predict the correct digit label automatically

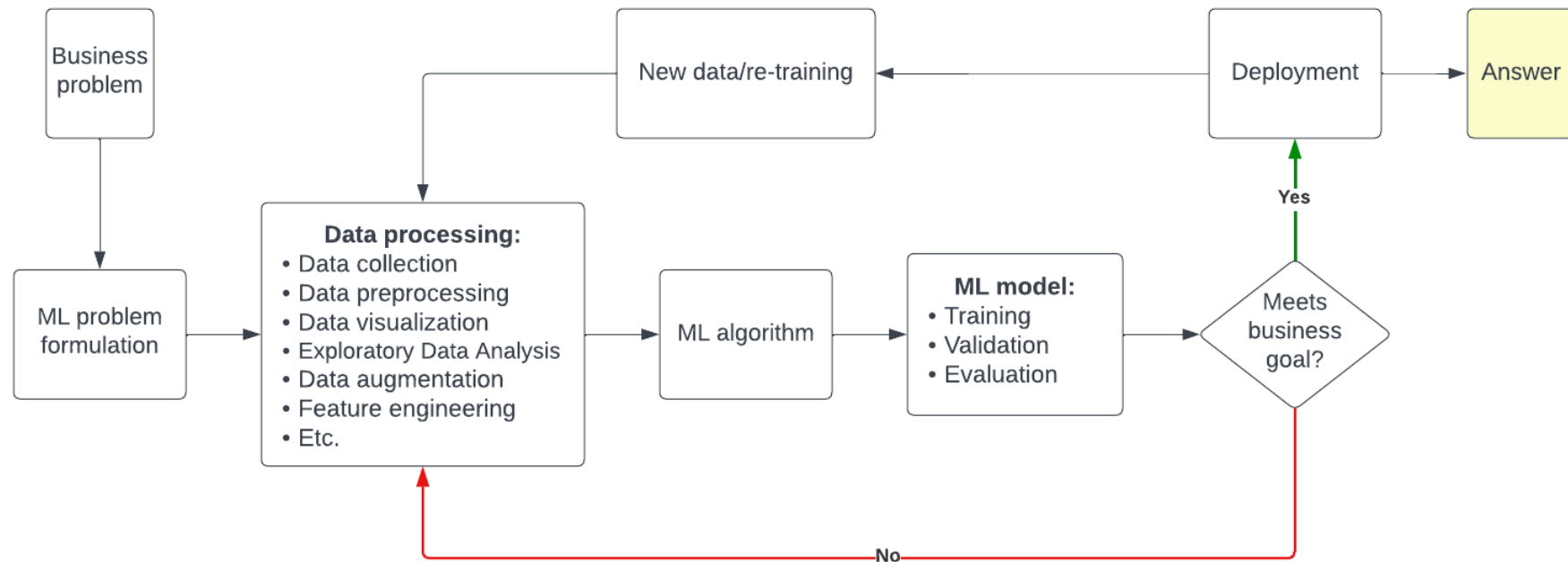
Machine learning process



7 steps of Machine Learning



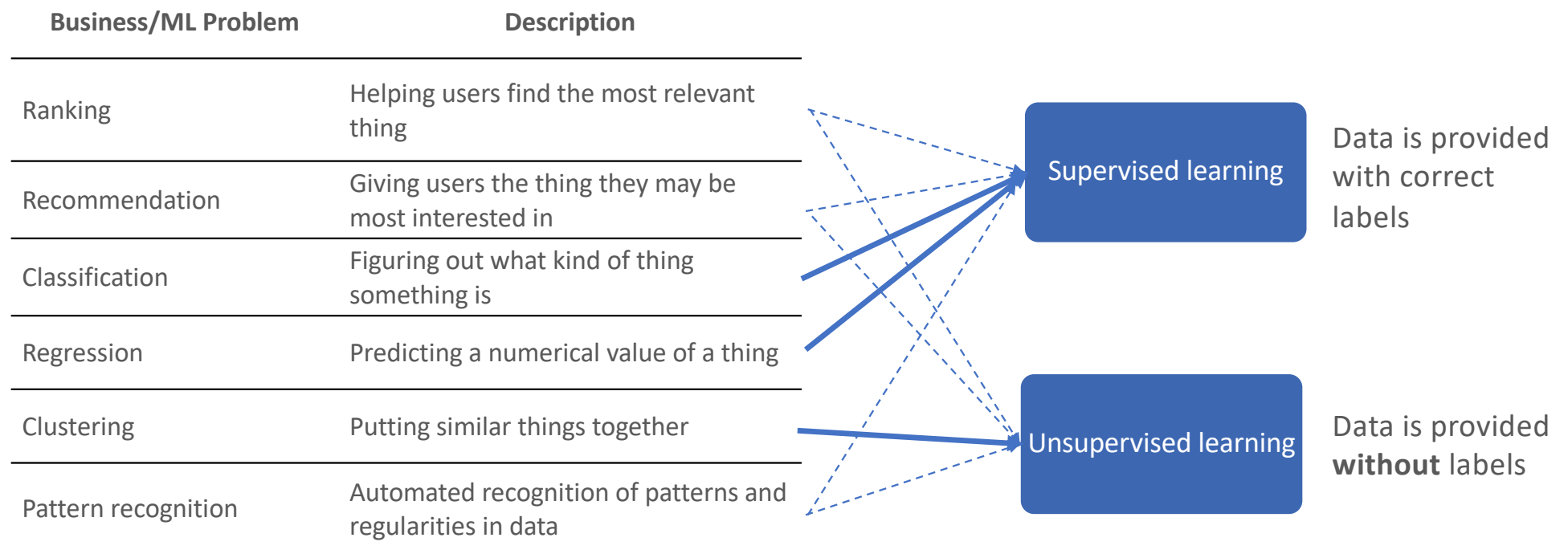
Machine learning process



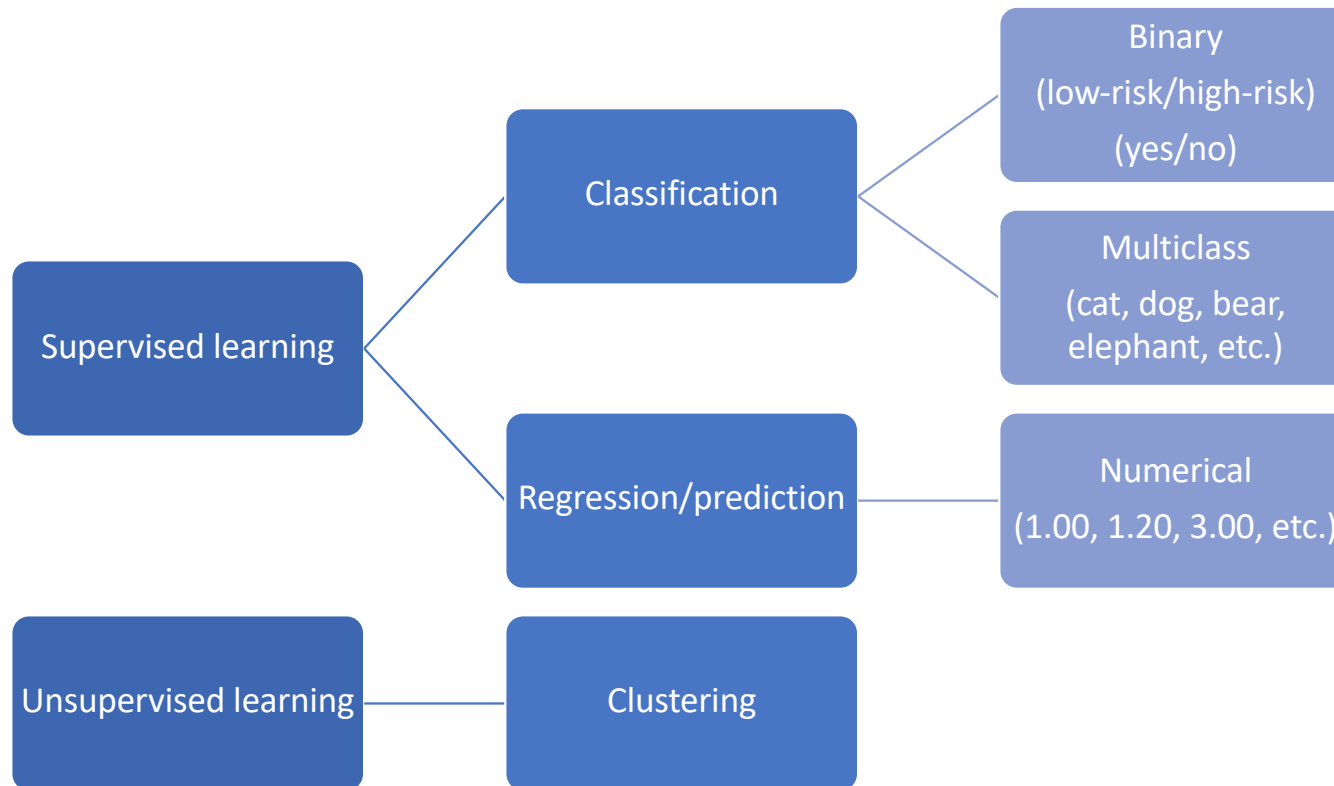
ML terms

ML	Statistics/Math	Example
Target/label/y	Dependent/response/output variable	The thing you're trying to predict
Feature/x	Independent/explanatory/input variable	Data that help us to make predictions
Feature Engineering	Transformation	Reshaping data to get more value
1D, 2D, nD	Dimensionality	Number of features
Model weights	Parameters of the model	A set of numbers embedded in a model that can predict the label

Supervised & unsupervised learning



Machine learning tasks

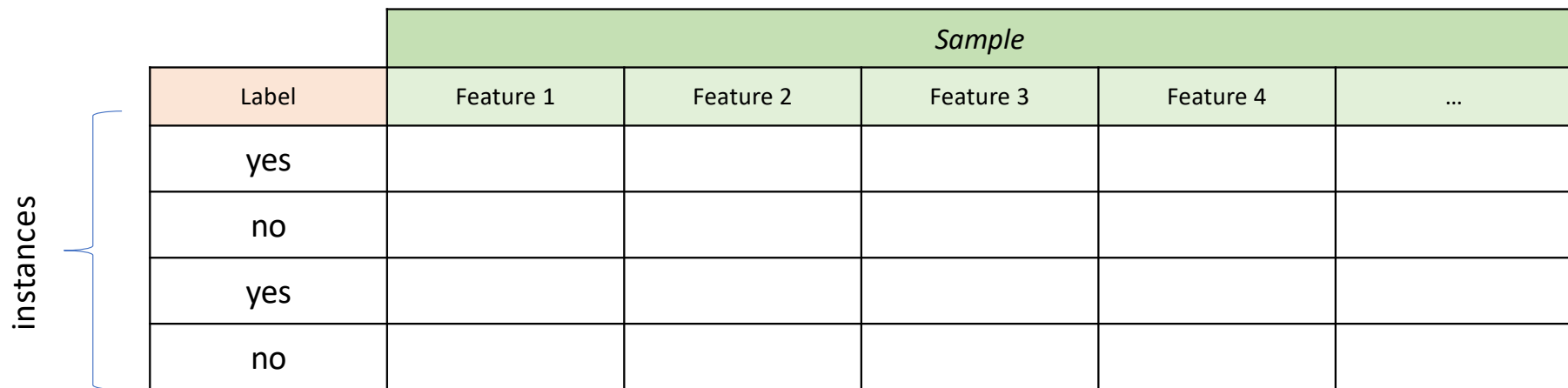


Supervised learning: dataset

Input: labeled/tagged dataset

Task: predict/classify the correct label

You have the correct label for every instance/observation you give the system



Sample					
Label	Feature 1	Feature 2	Feature 3	Feature 4	...
yes					
no					
yes					
no					

Instance: observation (row of a tabular structure)

Label: answer (target, label, y, ground truth, output, response, the desired output of the model)

Features: field values (column of a tabular structure, something we know about the observation)

Sample: collection of features

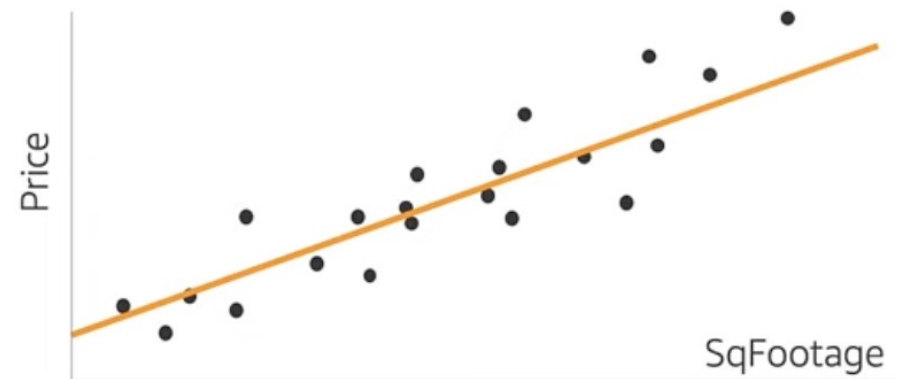
Supervised learning: classification

Label	Features		
	Points	Edges	Size
Star			
0	5	10<	750
1	2	>9	150
...



Supervised learning: regression

Label	Features		
	Bedrooms	SqFootage	Age
Price			
280000	3	3292	14
210000	2	2465	6
...



Unsupervised learning: clustering

Input: No labeled dataset

Task: discover the structure of the dataset

How:

- grouping similar items to form 'clusters', or associating them
- Reducing dimensions of data

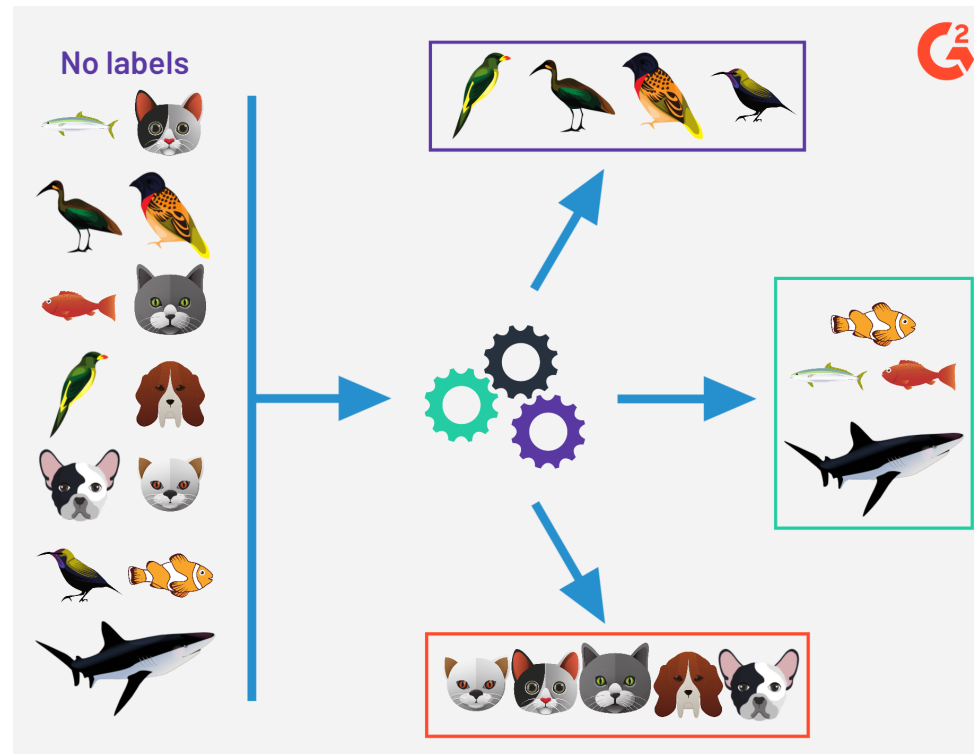
Features		
Age	Music	Books
21	Classical	Practical Magic
40	Jazz	Harry Potter
...



Input samples are grouped into clusters based on the underlying patterns.

Unsupervised learning: clustering

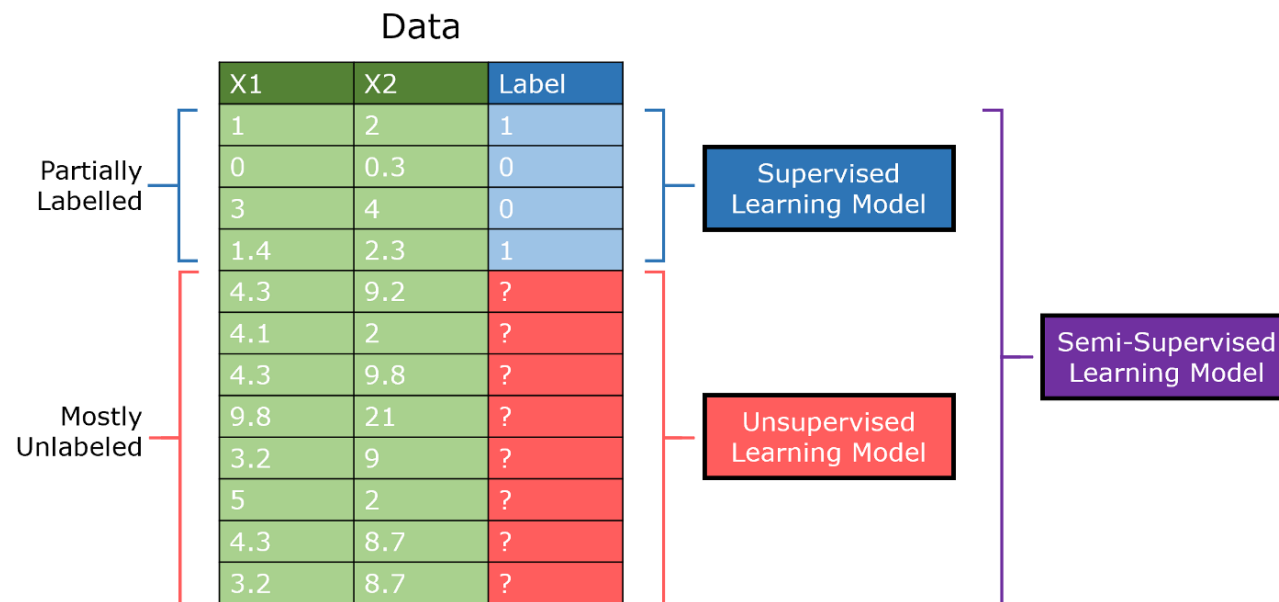
The search for patterns. You do not have any labels



Erick Cuenca - twitter: @erickedu85

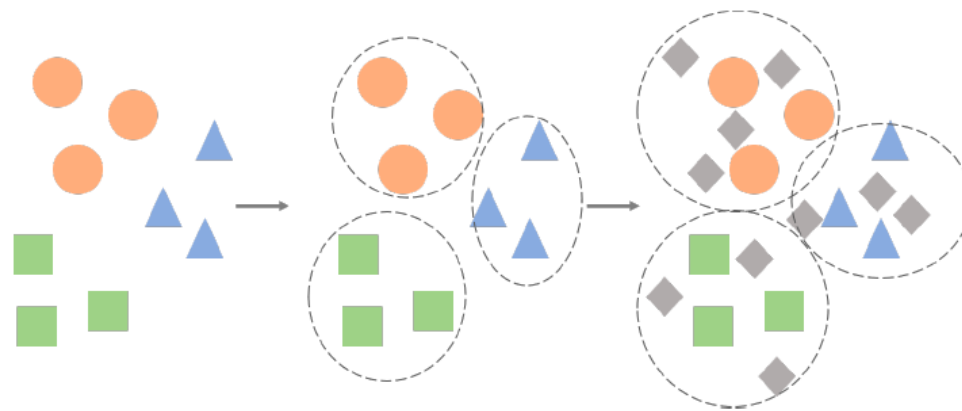
Semi-supervised learning

A blend of supervised and unsupervised learning. Some of the data are labeled, some not



Semi-supervised learning

A blend of supervised and unsupervised learning. Some of the data are labeled, some not



The clusters formed by a large amount of unlabeled data are used to classify a limited amount of labeled data.