Aprendizagem Aplicada à Segurança

Mário Antunes

October 14, 2023

Universidade de Aveiro

Context

It is becoming difficult to identify Cybersecurity attacks. These attacks can originate internally due to malicious intent or negligent actions or externally by malware, target attacks, and APT (Advanced Persistent Threats).

But insider threats are more challenging and can cause more damage than external threats because they have already entered the network.

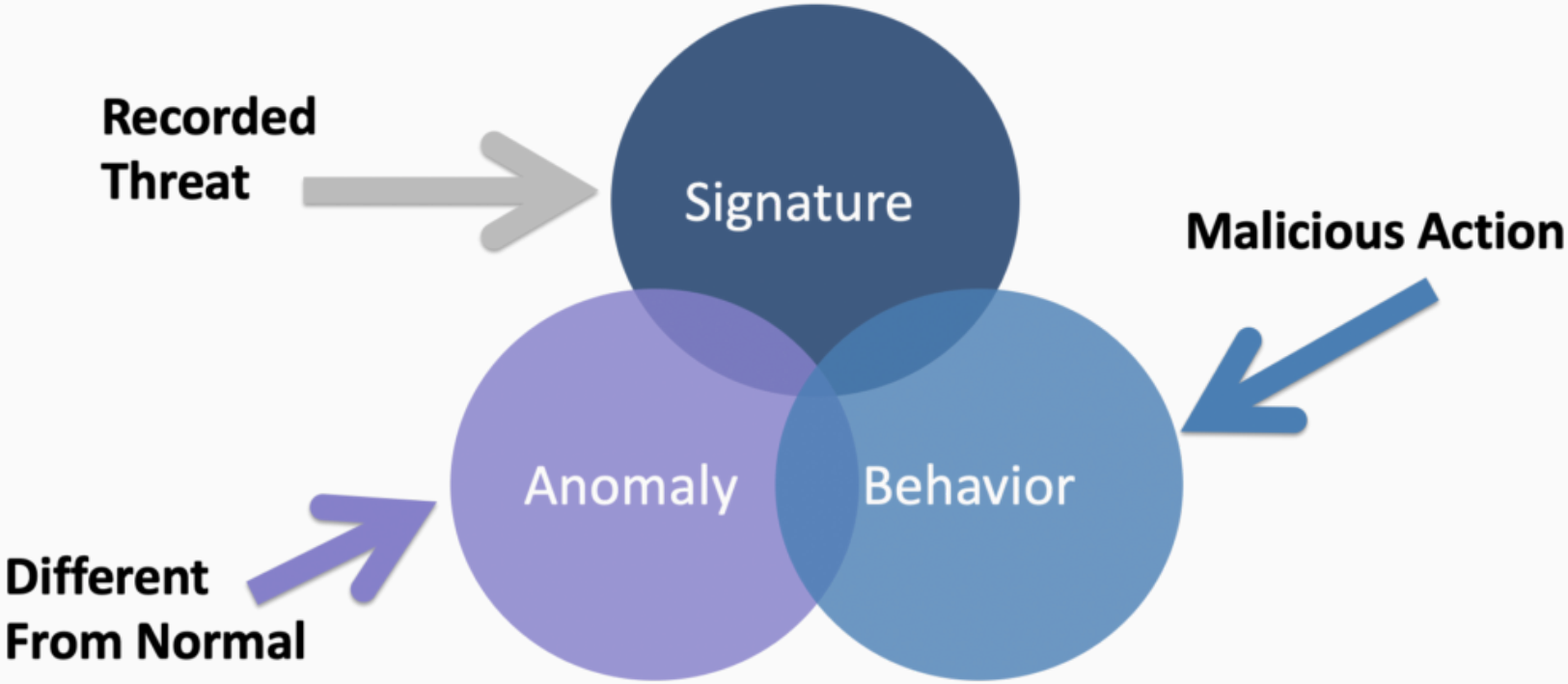
These activities present unknown threats and can steal, destroy or alter the assets.

Context

Earlier firewalls, web gateways, and some other intrusion prevention tools are enough to be secure, but now hackers and cyber attackers can bypass approximately all these defense systems.

Therefore with making these prevention systems strong, it is also equally essential to use detection. So that if hackers get into the network, the system should be able to detect their presence.

Context


Recorded
Threat
Signature
Malicious Action
Anomaly
Behavior
Different
From Normal

Context

Signature detection requires knowing what to look for and comparing hashes or other strings to identify a match. Signature detection is a common feature found within antivirus and IPS/IDS products.

Behavior detection looks for malicious or other known behavior characteristics and alarms the SOC when a match is made. An example is identifying port scanning or a file attempting to encrypt your hard drive, which is an indication of ransomware behavior. Antimalware and sandboxes are examples of tools that heavily leverage behavior detection capabilities.

Anomay detection it takes into consideration hot topics including big data, threat intelligence, and “zero-day” detection.

Anomaly Detection

Anomaly detection, also called outlier detection, is the identification of unexpected events, observations, or items that differ significantly from the norm:

- Anomalies in data occur only very rarely
- The features of data anomalies are significantly different from those of normal instances

What is an anomaly?

Generally speaking, an **anomaly** is something that differs from a norm: a deviation, an exception. In software engineering, by anomaly we understand a rare occurrence or event that doesn't fit into the pattern, and, therefore, seems suspicious. Some examples are:

- sudden burst or decrease in activity;
- error in the text logs;
- sudden rapid drop or increase in temperature.

What is an anomaly?

Common reasons for outliers are:

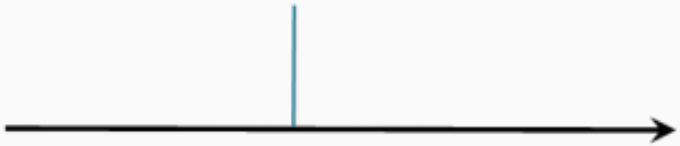
- data preprocessing errors;
- noise;
- fraud;
- attacks.

Types of Anomalies

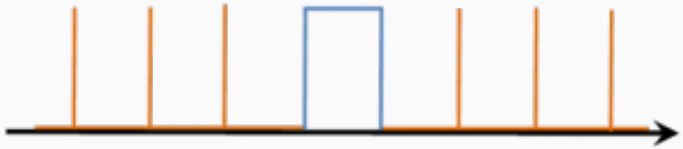
Anomalies can be broadly categorized as:

- Point anomalies: A single instance of data is anomalous if it's too far off from the rest.
- Contextual anomalies: The abnormality is context specific. This type of anomaly is common in time-series data.
- Collective anomalies: A set of data instances collectively helps in detecting anomalies.

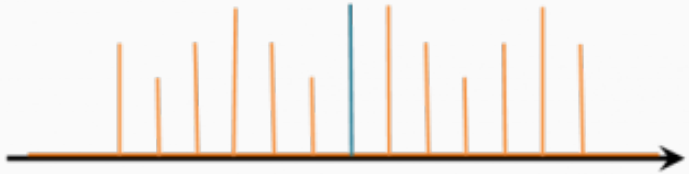
Types of Anomalies



(a) Point Anomaly



(b) Collective Anomaly



(c) Contextual Anomaly

Anomaly Detection - Example #1

Network anomalies: Anomalies in network behavior deviate from what is normal, standard, or expected. To detect network anomalies, network owners must have a concept of expected or normal behavior. Detection of anomalies in network behavior demands the continuous monitoring of a network for unexpected trends or events.

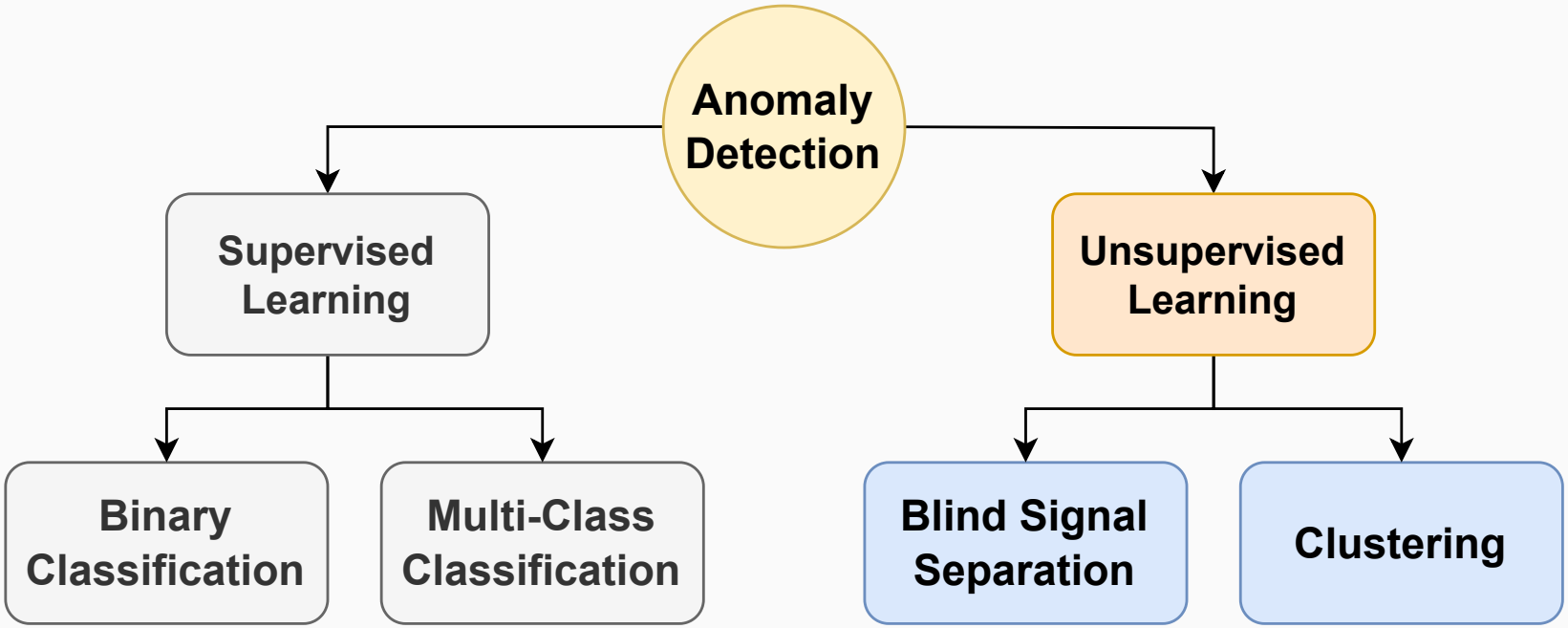
Anomaly Detection - Example #2

Application performance anomalies: These are simply anomalies detected by end-to-end application performance monitoring. These systems observe application function, collecting data on all problems, including supporting infrastructure and app dependencies. When anomalies are detected, rate limiting is triggered and admins are notified about the source of the issue with the problematic data.

Anomaly Detection - Example #3

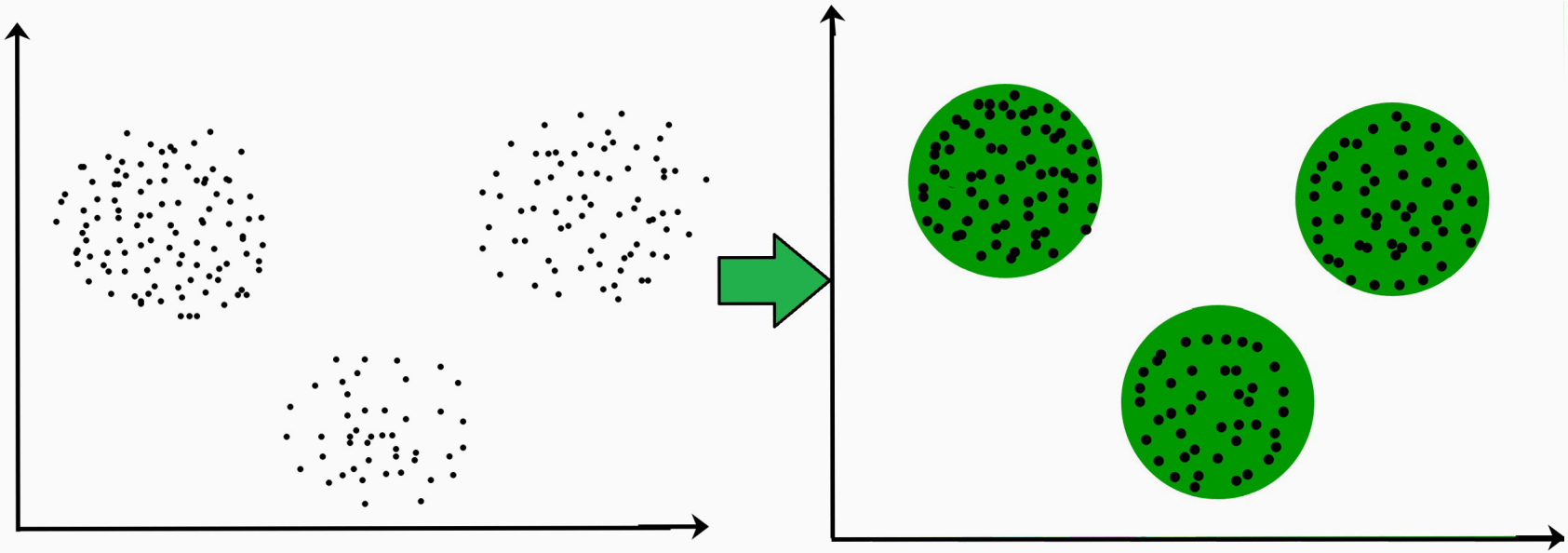
Web application security anomalies: These include any other anomalous or suspicious web application behavior that might impact security such as CSS attacks or DDOS attacks.

Anomaly Detection



Clustering

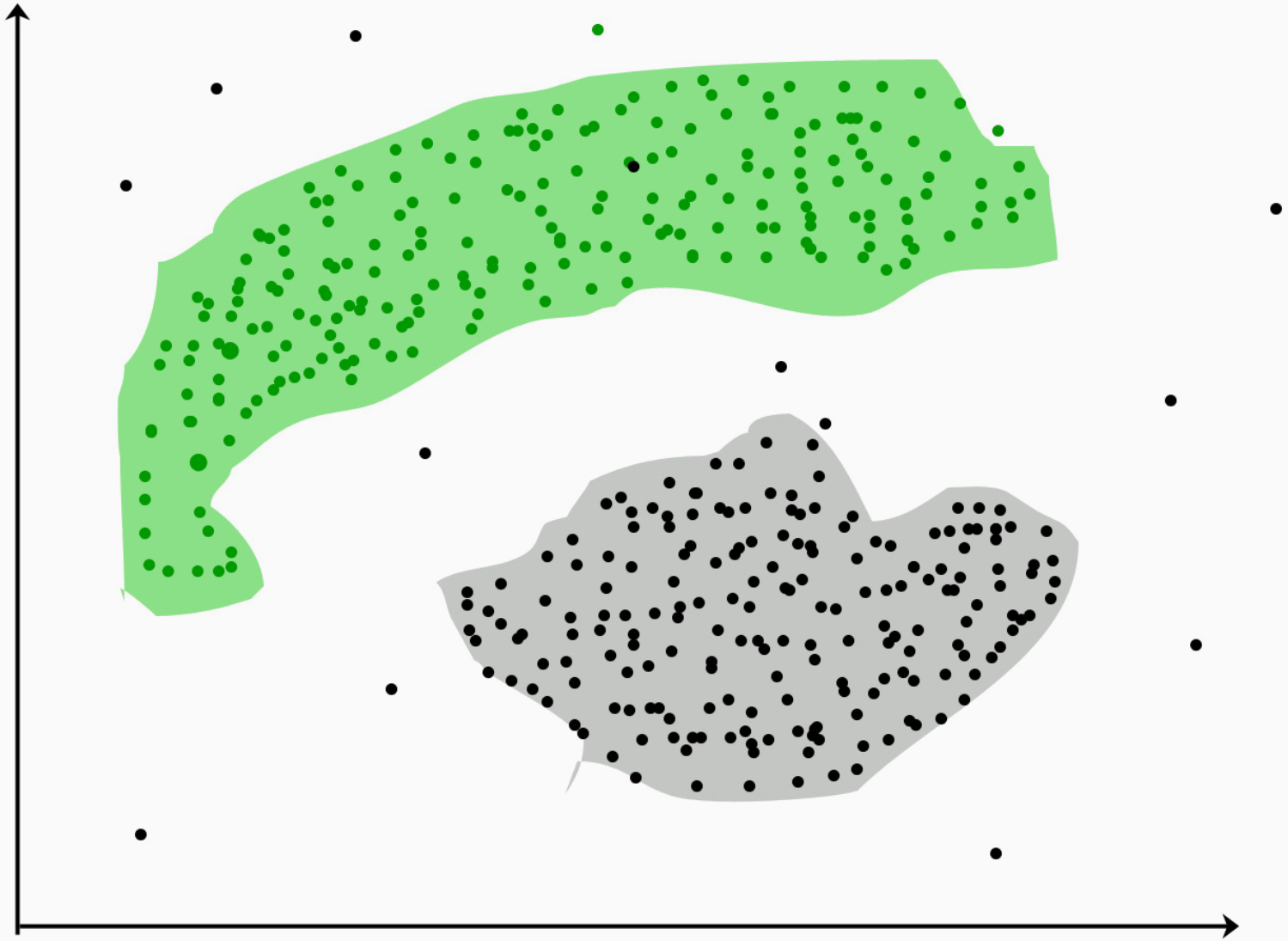
Type of **unsupervised learning method**. Generally, it is used as a process to find meaningful structure, explanatory underlying processes, generative features, and groupings inherent in a set of examples.



Clustering Methods

- **Density-Based Methods:** These methods consider the clusters as the dense region having some similarities and differences from the lower dense region of the space. These methods have good accuracy and the ability to merge two clusters.
- **Hierarchical Based Methods:** The clusters formed in this method form a tree-type structure based on the hierarchy. New clusters are formed using the previously formed one.
- **Partitioning Methods:** These methods partition the objects into k clusters and each partition forms one cluster. This method is used to optimize an objective criterion similarity function such as when the distance is a major parameter.

Clustering: Anomaly Detection



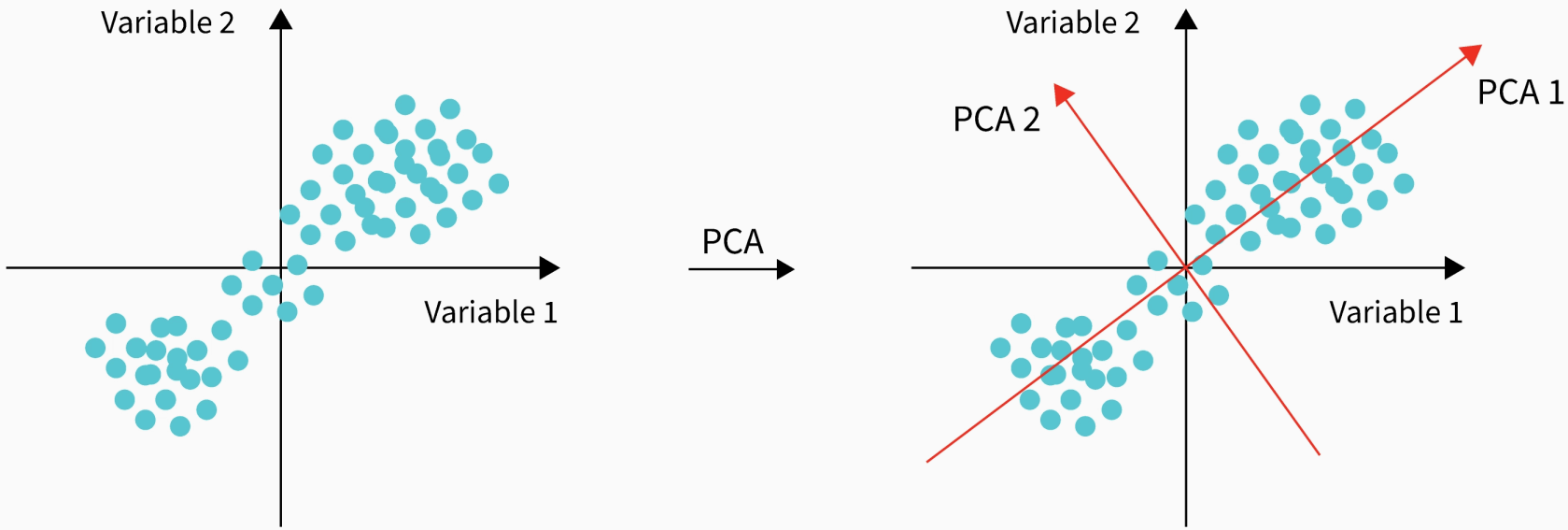
Blind Source Separation

Blind Source Separation (BSS) refers to a problem where both the sources and the mixing methodology are unknown, only mixture signals are available for further separation process.

In several situations it is desirable to recover all individual sources from the mixed signal, or at least to segregate a particular source.

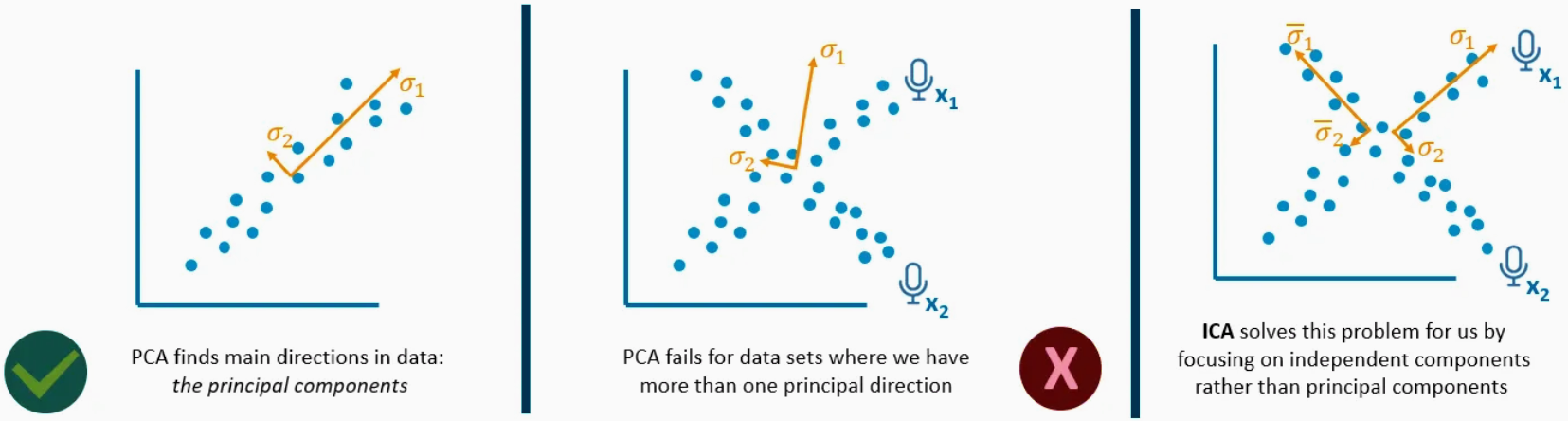
Blind Source Separation: PCA

** Principal component analysis**, or PCA, is a statistical procedure that allows you to summarize the information content in large data tables by means of a smaller set of “summary indices” that can be more easily visualized and analyzed.



Blind Source Separation: ICA

Independent Component Analysis (ICA) is a powerful technique in the field of data analysis that allows you to separate and identify the underlying independent sources in a multivariate data set.

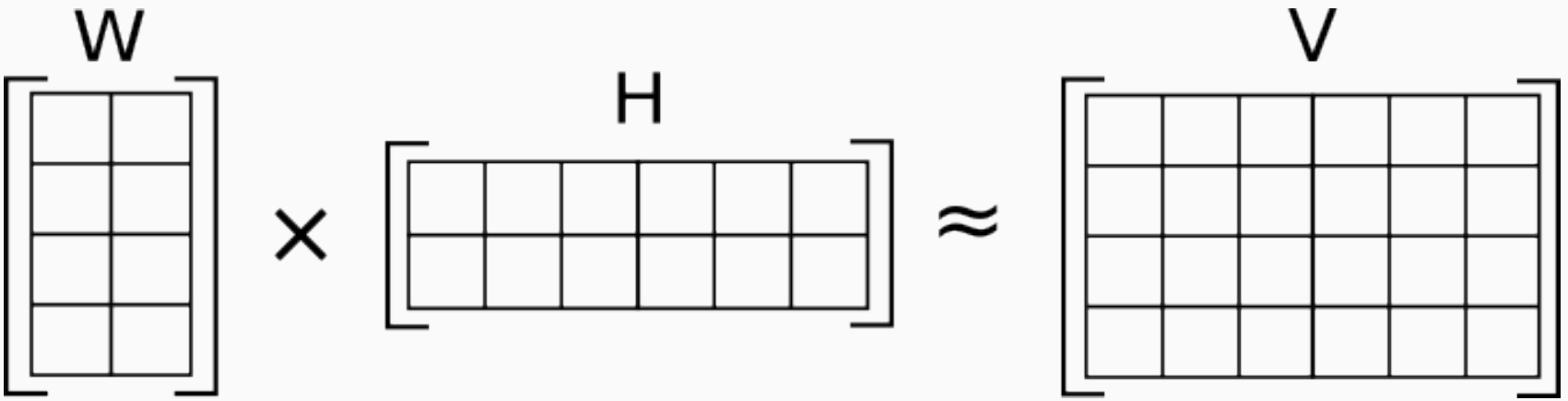


Blind Source Separation: NNMF

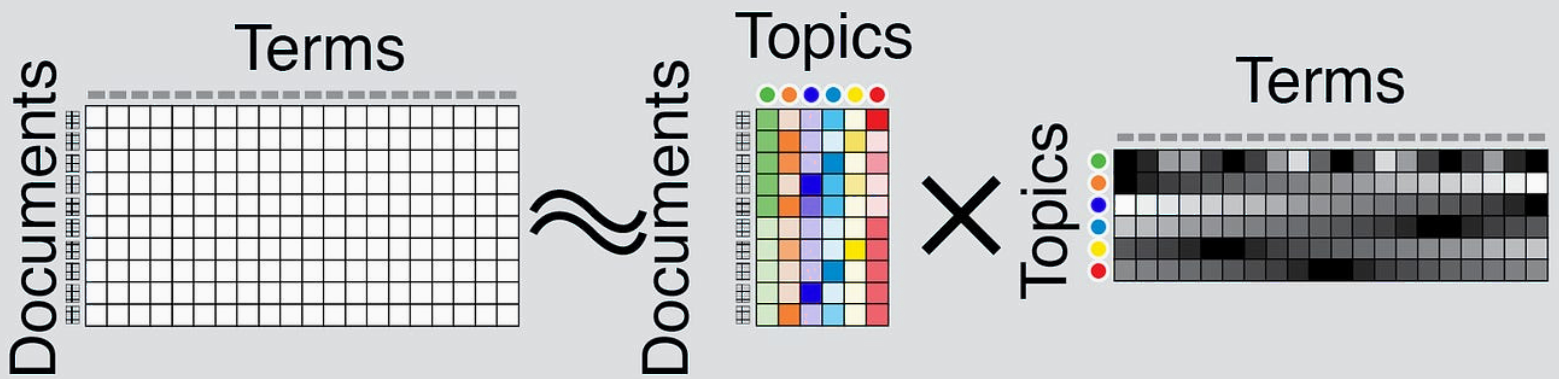
Non-negative matrix factorization (NNMF) is a group of algorithms in multivariate analysis and linear algebra where a matrix V is factorized into two matrices W and H, with the property that all three matrices have no negative elements.

This non-negativity makes the resulting matrices easier to inspect. Also, in applications such as processing of audio spectrograms or muscular activity, non-negativity is inherent to the data being considered.

Since the problem is not exactly solvable in general, it is commonly approximated numerically.



Non-Negative Matrix Factorization Diagram - Example



V ≈ **W** × **H**

Visible Variables
Input
Document x Term Matrix
n x m
10 x 20

Weights
Feature Set
Document x Topic Matrix
n x p
10 x 6

Hidden Variables
Coefficients
Topic x Term Matrix
p x m
6 x 20

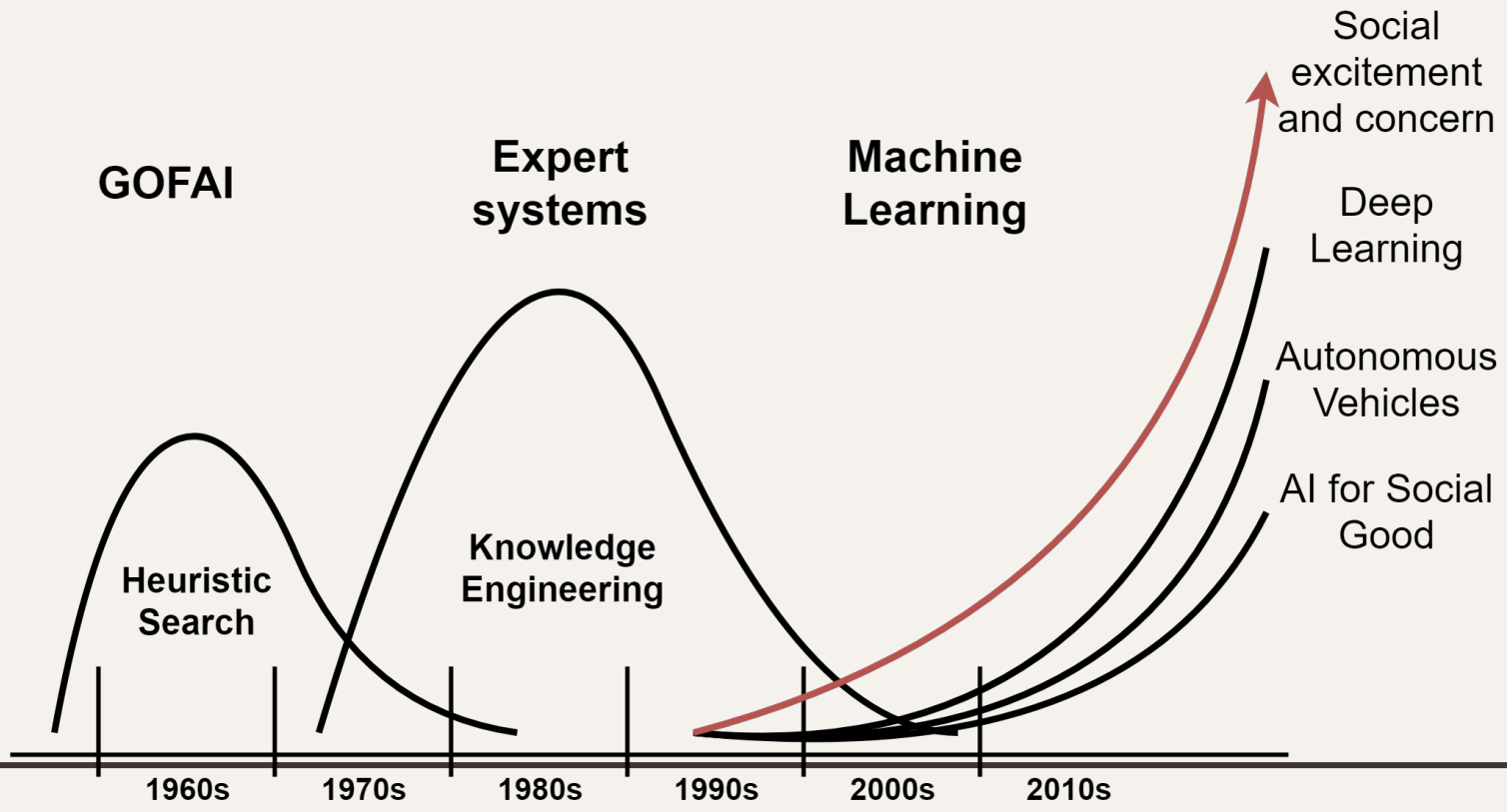
Machine Learning

Refresher

About Me

- Masters and PhD on Artificial Intelligence and Machine Learning
- Researcher at IT Aveiro
- Areas of interest: Artificial Intelligence, Machine Learning, text mining, stream mining, IoT, M2M

AI & ML

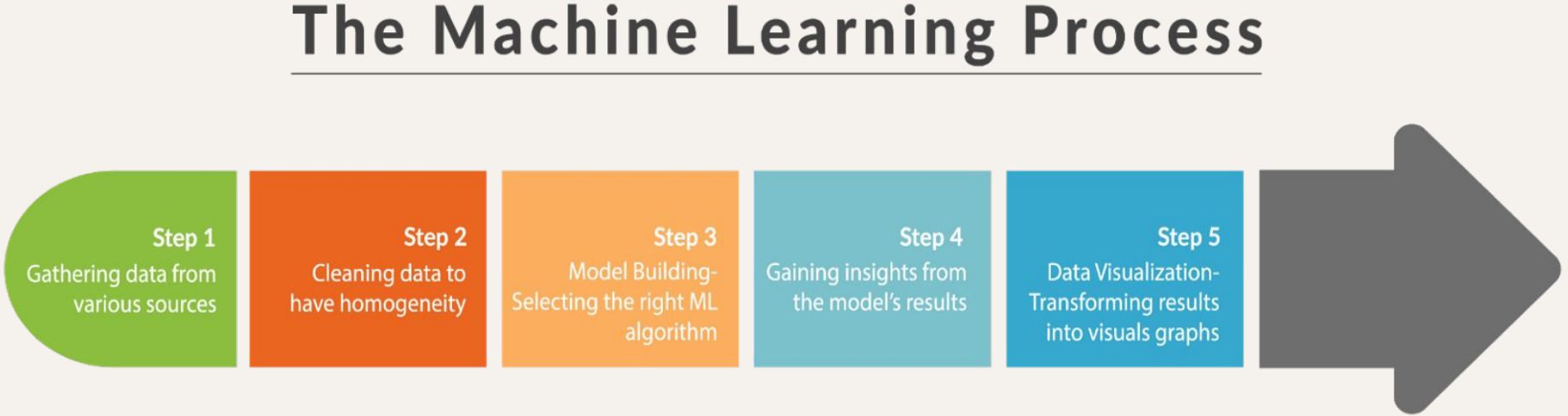

GOFAI
Expert systems
Machine Learning
Social excitement and concern
Deep Learning
Autonomous Vehicles
AI for Social Good
Heuristic Search
Knowledge Engineering
1960s
1970s
1980s
1990s
2000s
2010s

What is ML (Why should i Care)?

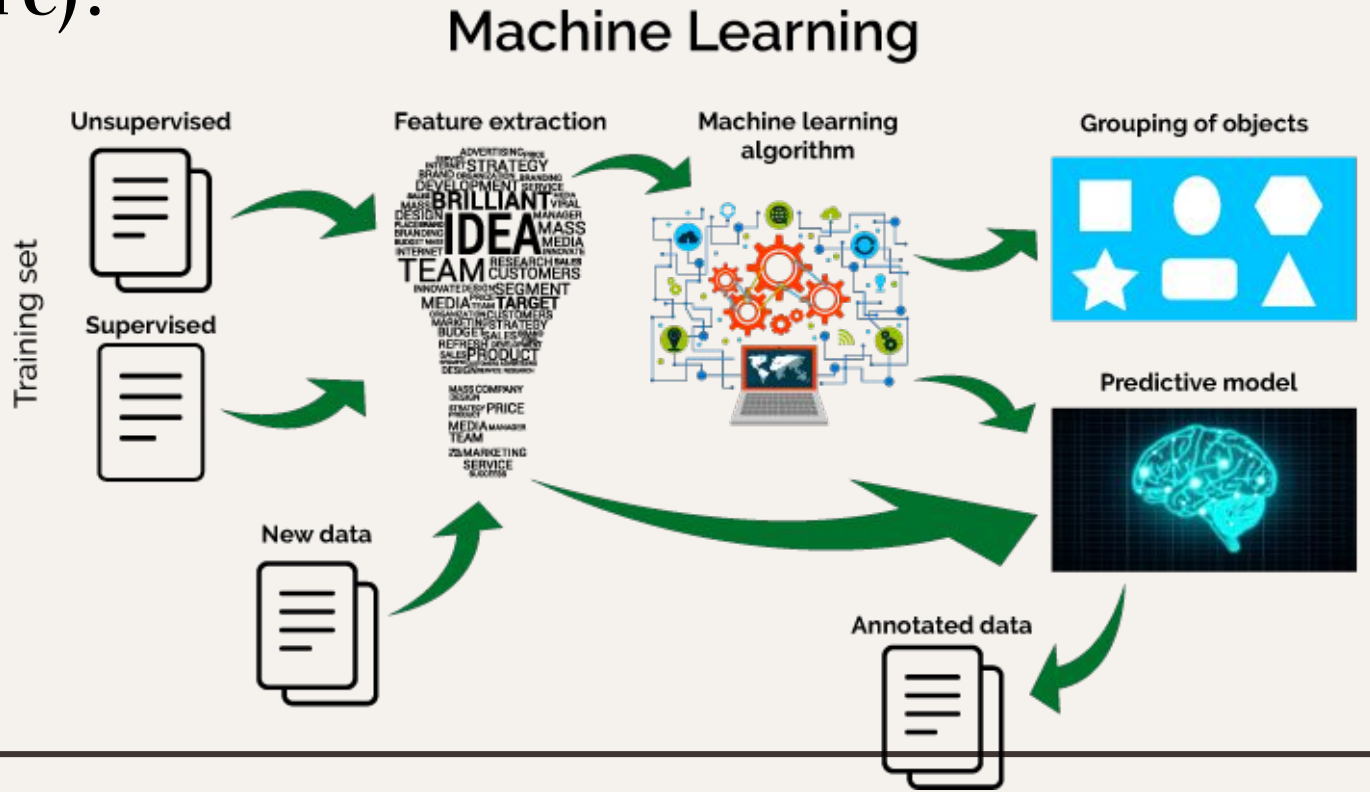
What does machine learning mean?

The term machine learning (abbreviated ML) refers to the capability of a machine to improve its own performance. It does so by using a statistical model to make decisions and incorporating the result of each new trial into that model. In essence, the machine is programmed to learn through trial and error.

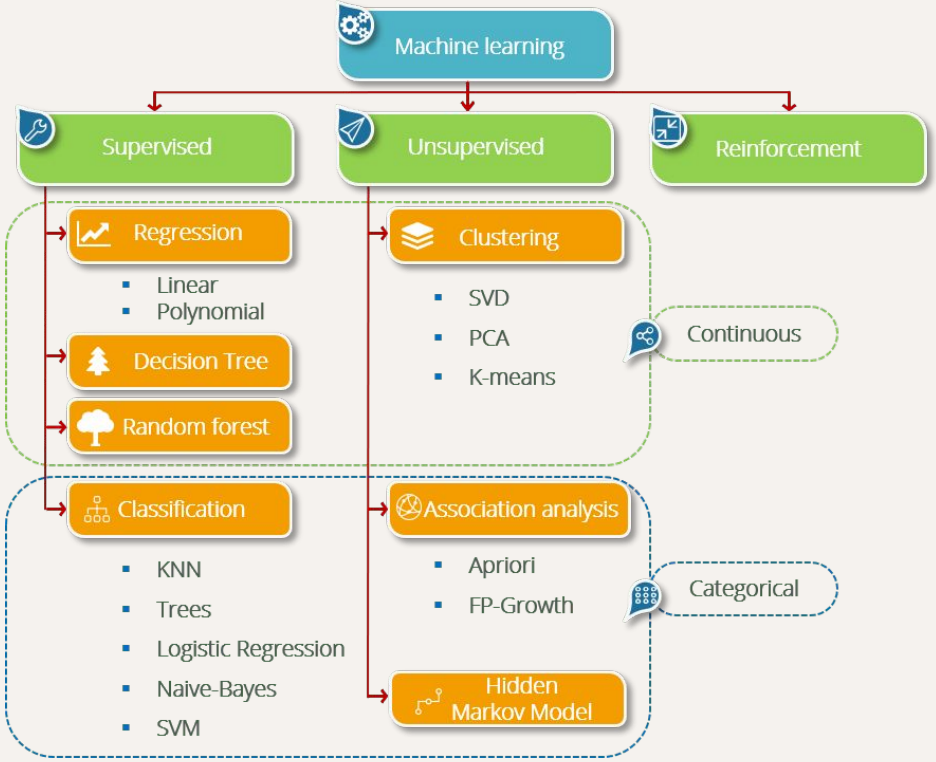
What is ML (Why should i Care)?



What is ML (Why should i Care)?

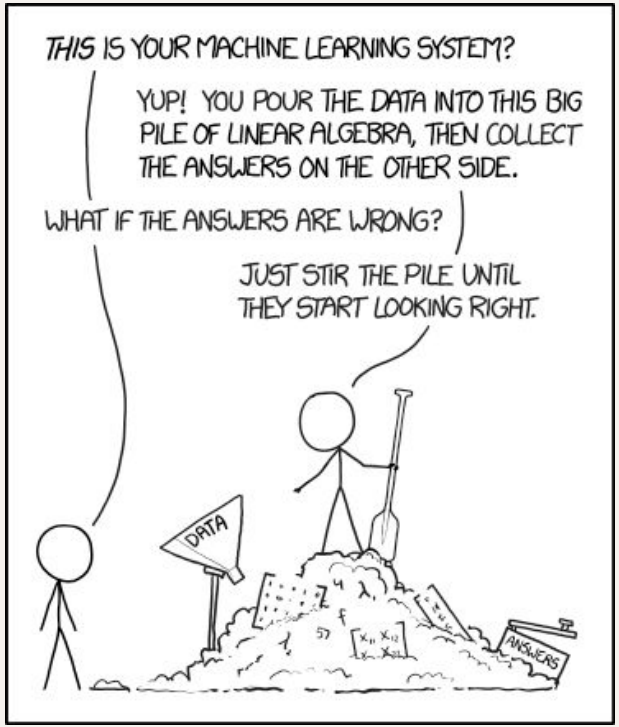


What is ML (Why should i Care)?

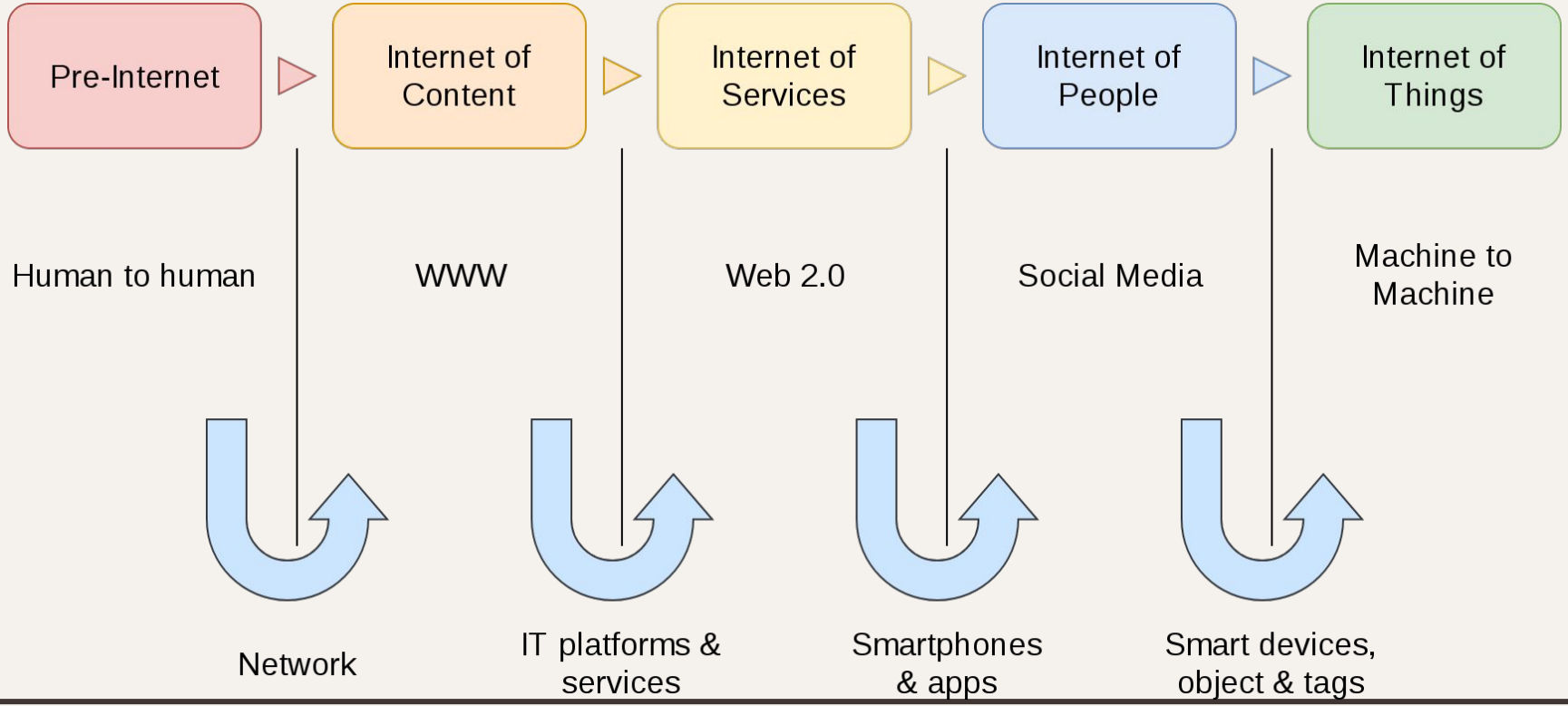


What is ML?

- A body of knowledge related with learning methods for machines (computers)
- Research area
- Opportunities for something useful

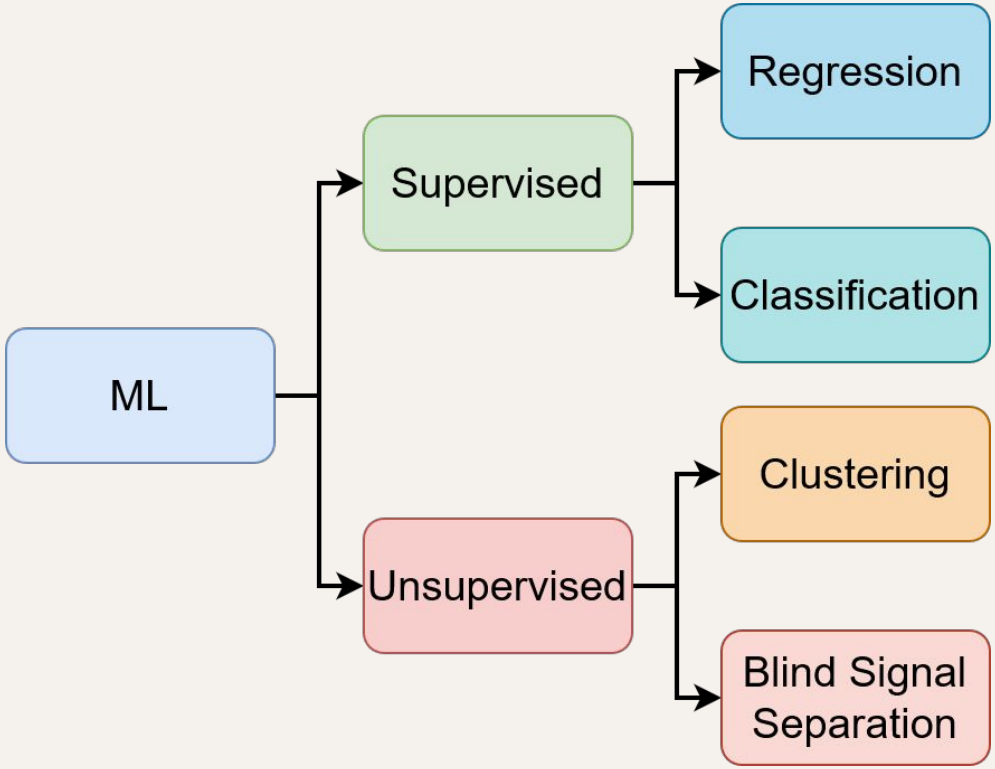


Why Should You Care?

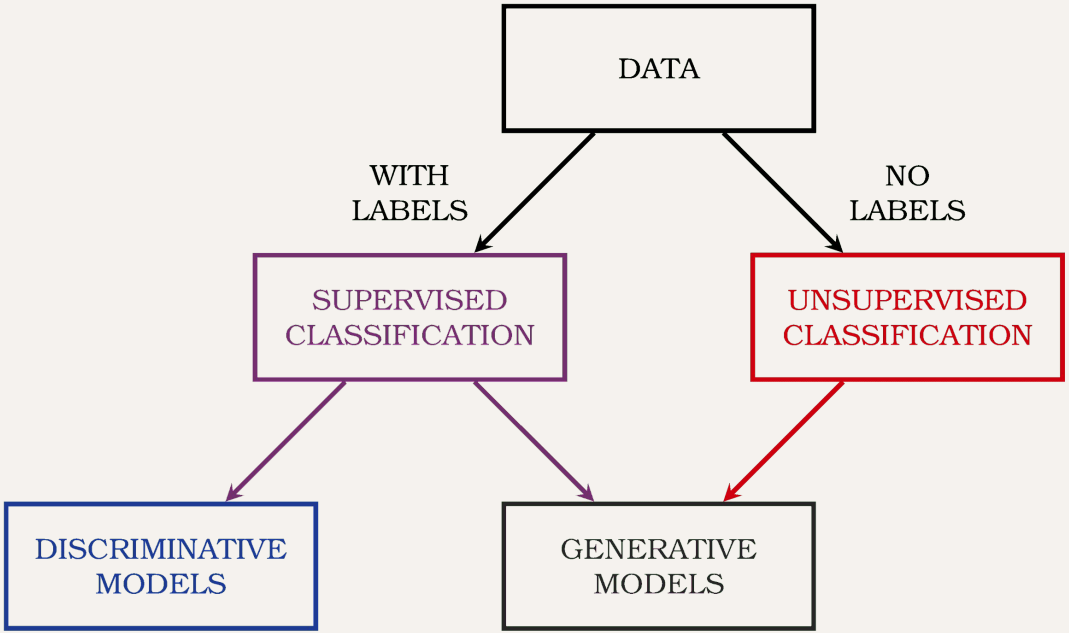


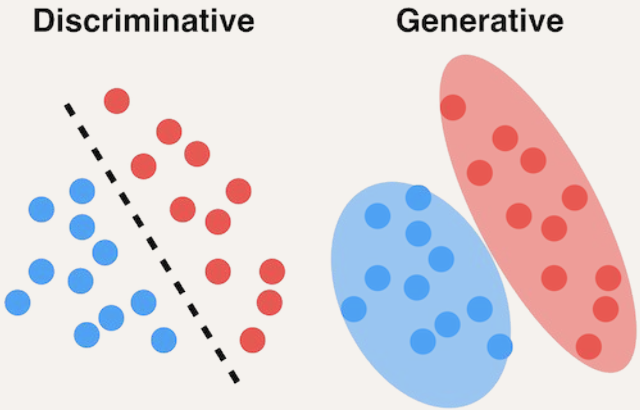
Taxonomy

Taxonomies...


ML
Supervised
Unsupervised
Regression
Classification
Clustering
Blind Signal Separation

Taxonomies...


DATA
WITH LABELS
NO LABELS
SUPERVISED CLASSIFICATION
UNSUPERVISED CLASSIFICATION
DISCRIMINATIVE MODELS
GENERATIVE MODELS


Discriminative
Generative

Taxonomies...

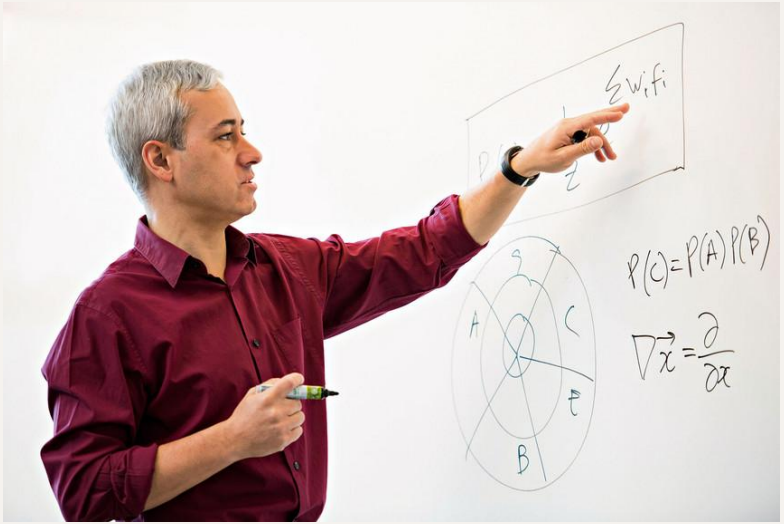
Induction symbolic reasoning

Neural Networks connections modelled on brain's neurons

Evolutionary algorithms learn from random generations (genetic algorithm)

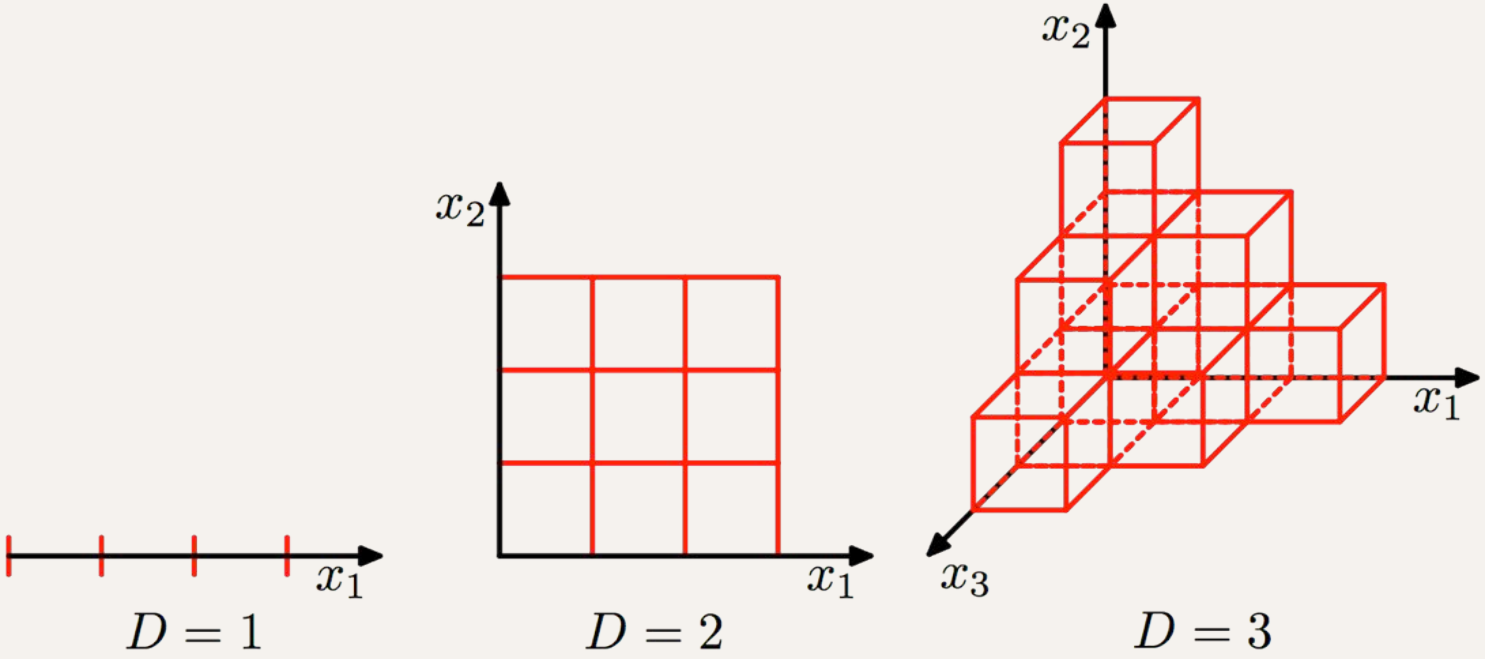
Bayesian inference probabilistic models based on bayes' theorem

Analogy learns by finding similar examples



Limitations

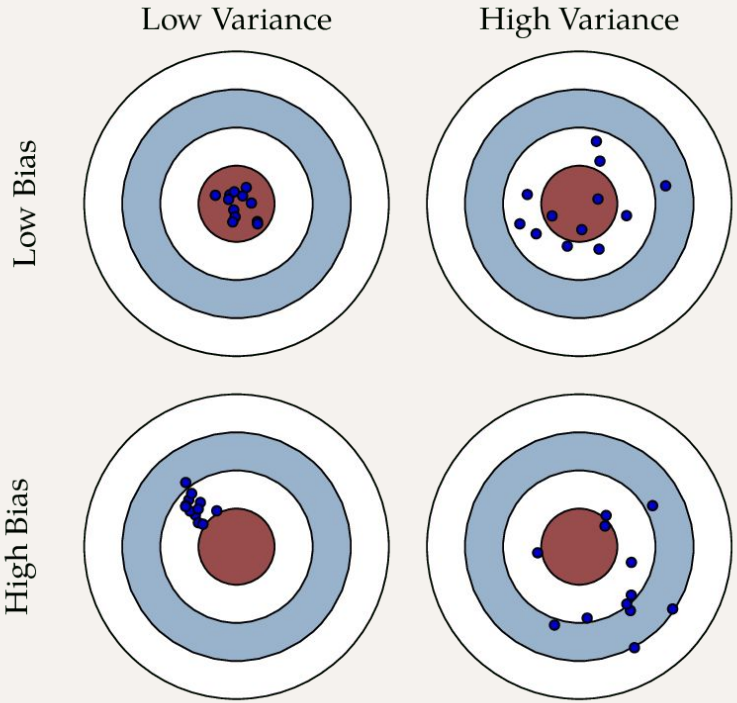
Limitations...



Limitations...

- Our model is a simplification of reality
- Simplification is based on assumptions (model bias)
- Assumptions fail in certain situations

Bias and Variance



Terminology

Terminology

Dataset: organized set of examples, typically composed of features and labels

Feature: single property of an example (input variable)

Label: classification category of an example (output variable)

Example: single instance of a dataset