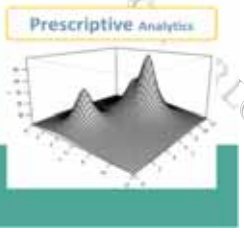
Machine Learning
Introduction

1

A taxonomy of analytics



2

Applications of Machine Learning & Artificial Intelligence

Smarter Healthcare

Multi-channel

Finance

Log Analysis

Homeland Security

Traffic Control

Telecom

Search Quality

Manufacturing

Trading Analytics

Fraud and Risk

Retail: Churn, NBO

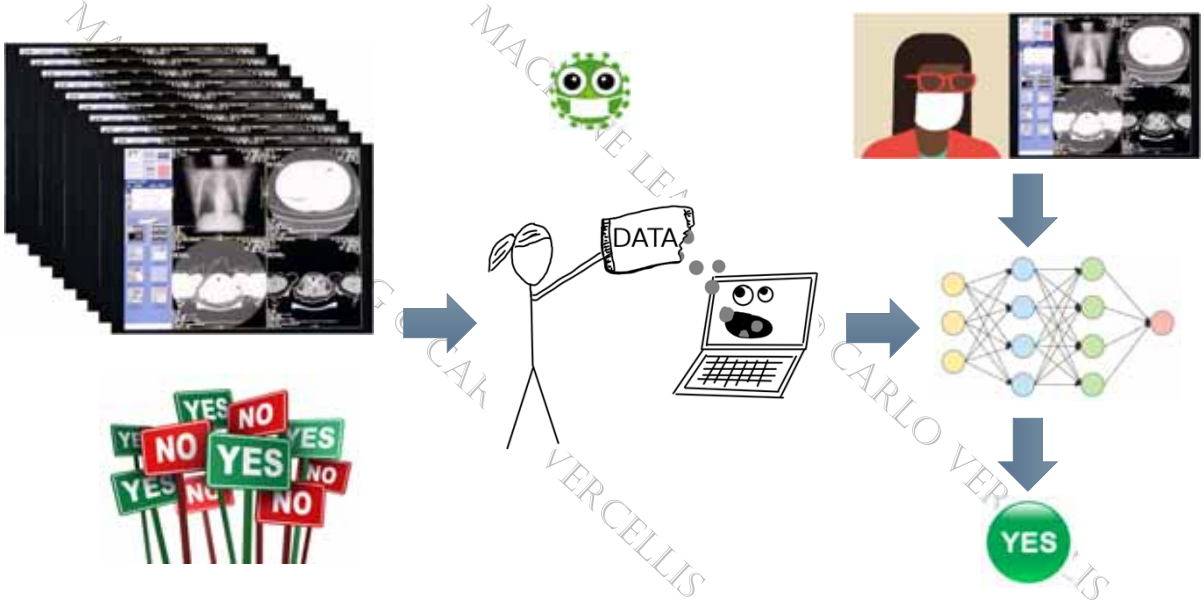
3

Artificial Intelligence & Machine Learning help in healthcare

COVID-19 Outbreak: Alibaba Claims Its AI-Based System Can Detect Virus With 96% Accuracy

4

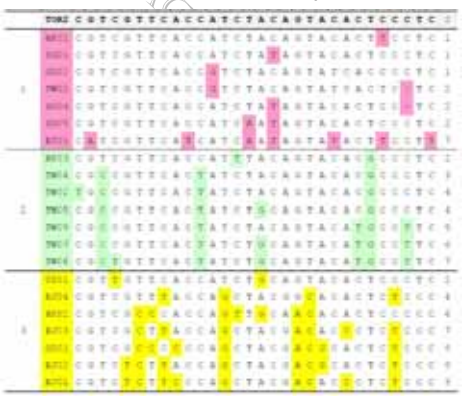
AI / ML what they can really do: learn from past labeled data



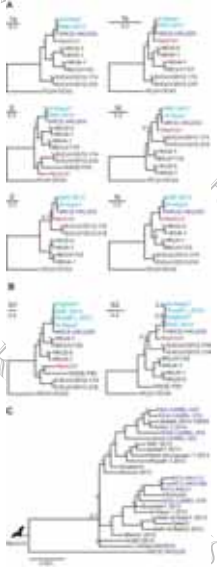
5

AI / ML what they can really do: clustering to get phylogeny

SARS-Covid DNA nucleotide sequences



Adenine
Guanine
Cytosine
Thymine (present in DNA only)
Uracil (present in RNA only)



6

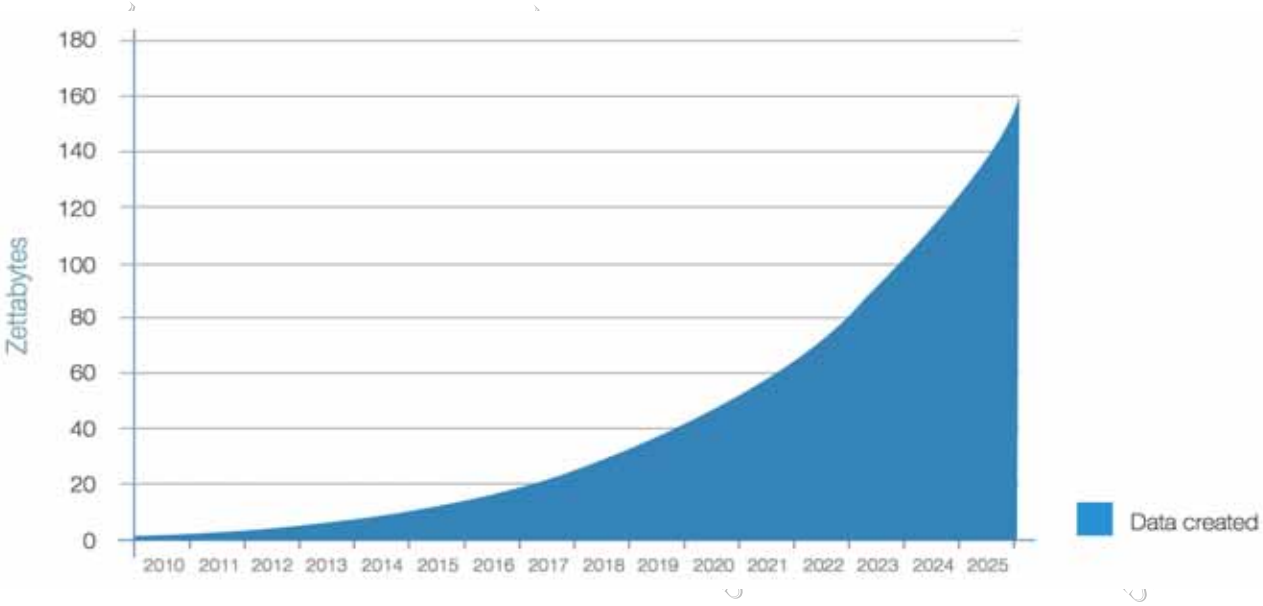
Internet of people – Velocity and Variety

7

Internet of things - Velocity

8

Big Data have high Volume – exponential growth



9

Artificial Intelligence

Emulate human behavior

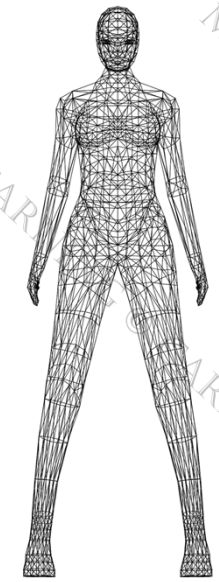


Image & Video & Sound & Smell & Speech Processing

Machine Learning & Optimization

Interacting Phisically & Socially

10

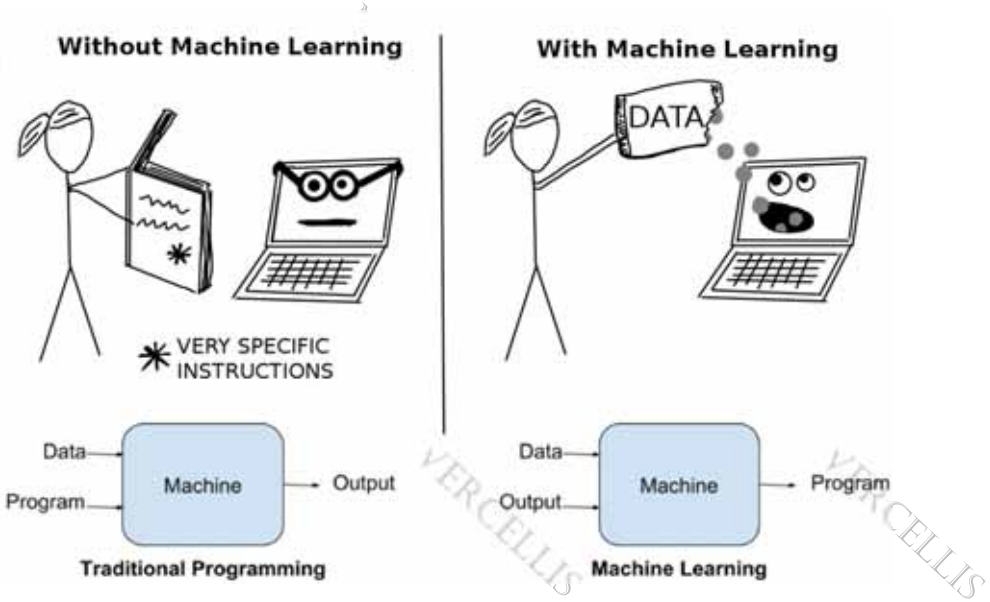
In search of hidden patterns

area	numin	timein	numout	Pothers	Pmob	Pland	numsms	numserv	numcall	dropt	churner
3	32	8093	45	0.14	0.75	0.12	18	1	0	0	0
3	277	157842	450	0.26	0.35	0.38	9	3	0	1	0
1	17	15023	20	0.37	0.23	0.40	1	1	0	0	0
1	46	22459	69	0.10	0.39	0.51	33	1	0	0	0
1	19	8640	9	0.00	0.00	1.00	0	0	0	0	0
2	17	7652	66	0.16	0.42	0.43	1	3	0	1	0
3	47	17768	11	0.45	0.00	0.55	0	0	0	0	0
3	19	9492	42	0.18	0.34	0.48	3	1	0	0	1
1	1	84	9	0.09	0.54	0.37	0	0	0	0	1
2	119	87605	126	0.84	0.02	0.14	12	1	0	0	0
4	24	6902	47	0.25	0.26	0.48	4	1	0	0	0
1	32	28072	43	0.28	0.66	0.06	0	1	0	0	0
3	103	112120	24	0.61	0.28	0.11	24	2	0	0	0
3	45	21921	94	0.34	0.47	0.19	45	2	0	1	0
1	8	25117	89	0.02	0.89	0.09	189	1	3	0	0
3	4	945	16	0.00	0.00	1.00	0	0	0	0	1
2	83	44263	83	0.00	0.00	0.67	0	0	0	0	1
2	22	15979	59	0.05	0.53	0.41	5	2	0	1	1
2	0	0	57	0.00	1.00	0.00	15	1	1	0	1
4	162	114108	273	0.18	0.15	0.41	2	3	0	1	1
4	21	4141	70	0.14	0.58	0.28	0	1	0	1	1
4	33	10066	45	0.12	0.21	0.67	0	0	0	0	1
4	5	965	40	0.41	0.27	0.32	64	1	0	0	1

Target

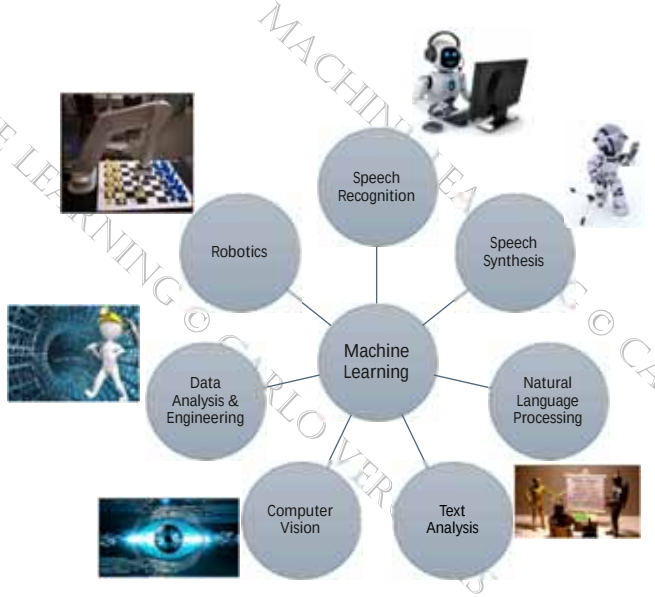
11

How Machine Learning works: learn from past data



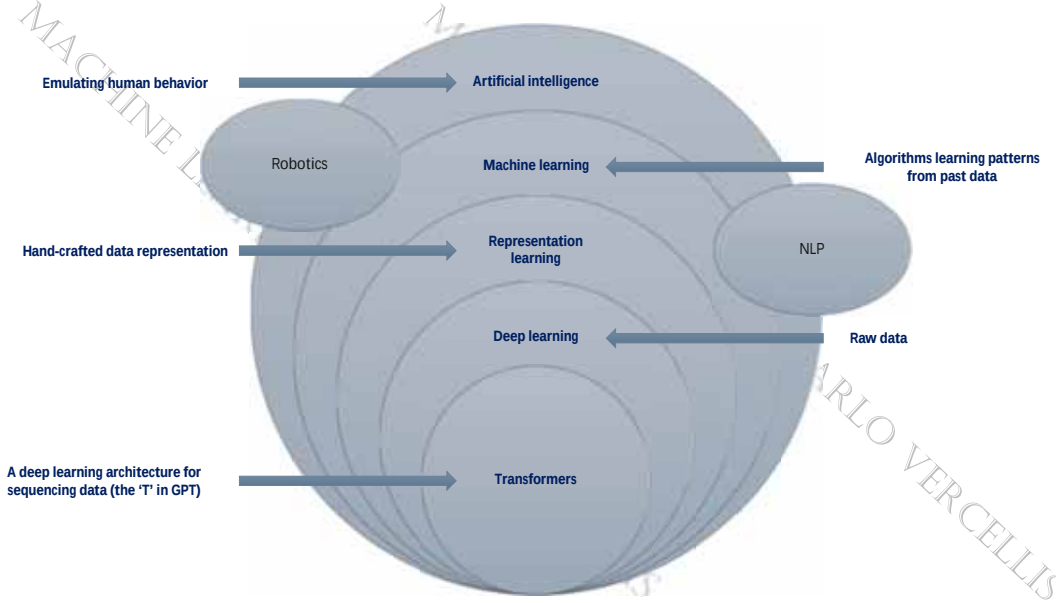
12

Machine Learning is central in Artificial Intelligence



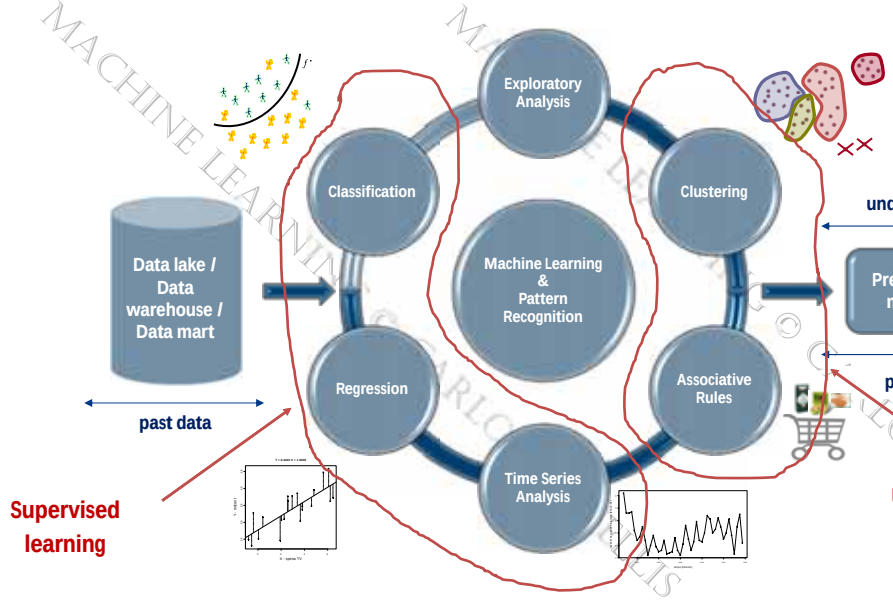
13

Machine learning is the major component in AI



14

Machine Learning tasks

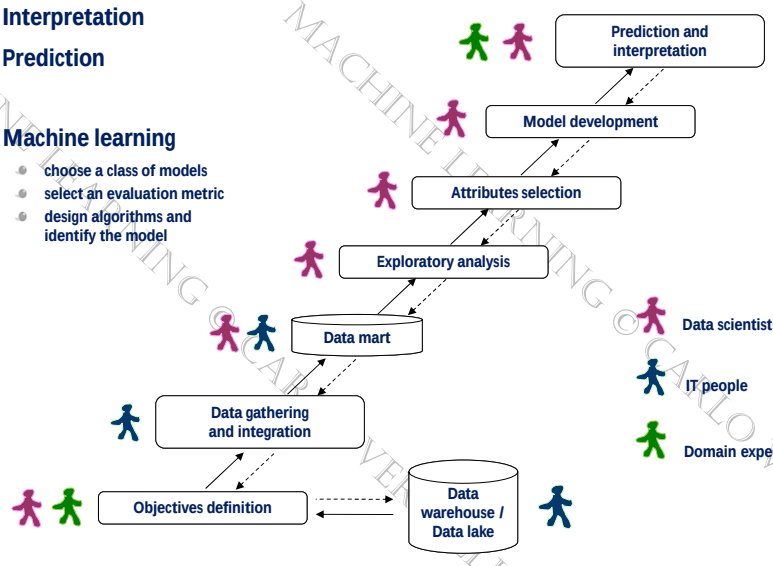


- Semi-supervised learning
- Reinforcement learning

15

Machine learning process

- Interpretation
- Prediction
- Machine learning
  - choose a class of models
  - select an evaluation metric
  - design algorithms and identify the model



16

Machine learning – input data

- Dataset: Two dimensional tables
- In a dataset:
  - Rows are the observations
  - Columns are the data that characterize each observation: features, attributes variables
- Attribute types:
  - Categorical attributes
    - Binary - True/False or 0/1
    - Nominal attribute without a natural sorting
    - Ordinal attribute with a natural sorting
  - Numerical attributes
    - Discrete attribute
    - Continuous attribute

17

Dataset

- A dataset $\mathcal{D}$ is represented by a $m \times n$ matrix

$$\mathbf{X} = [x_{ij}], \quad i \in \mathcal{M} = \{1, 2, \ldots, m\}, \quad j \in \mathcal{N} = \{1, 2, \ldots, n\}.$$

$$\mathbf{x}_i = (x_{i1}, x_{i2}, \ldots, x_{in}),$$
$$\mathbf{a}_j = (x_{1j}, x_{2j}, \ldots, x_{mj}),$$

18