

Initiation à l'apprentissage automatique en science des matériaux

6. Datasciences

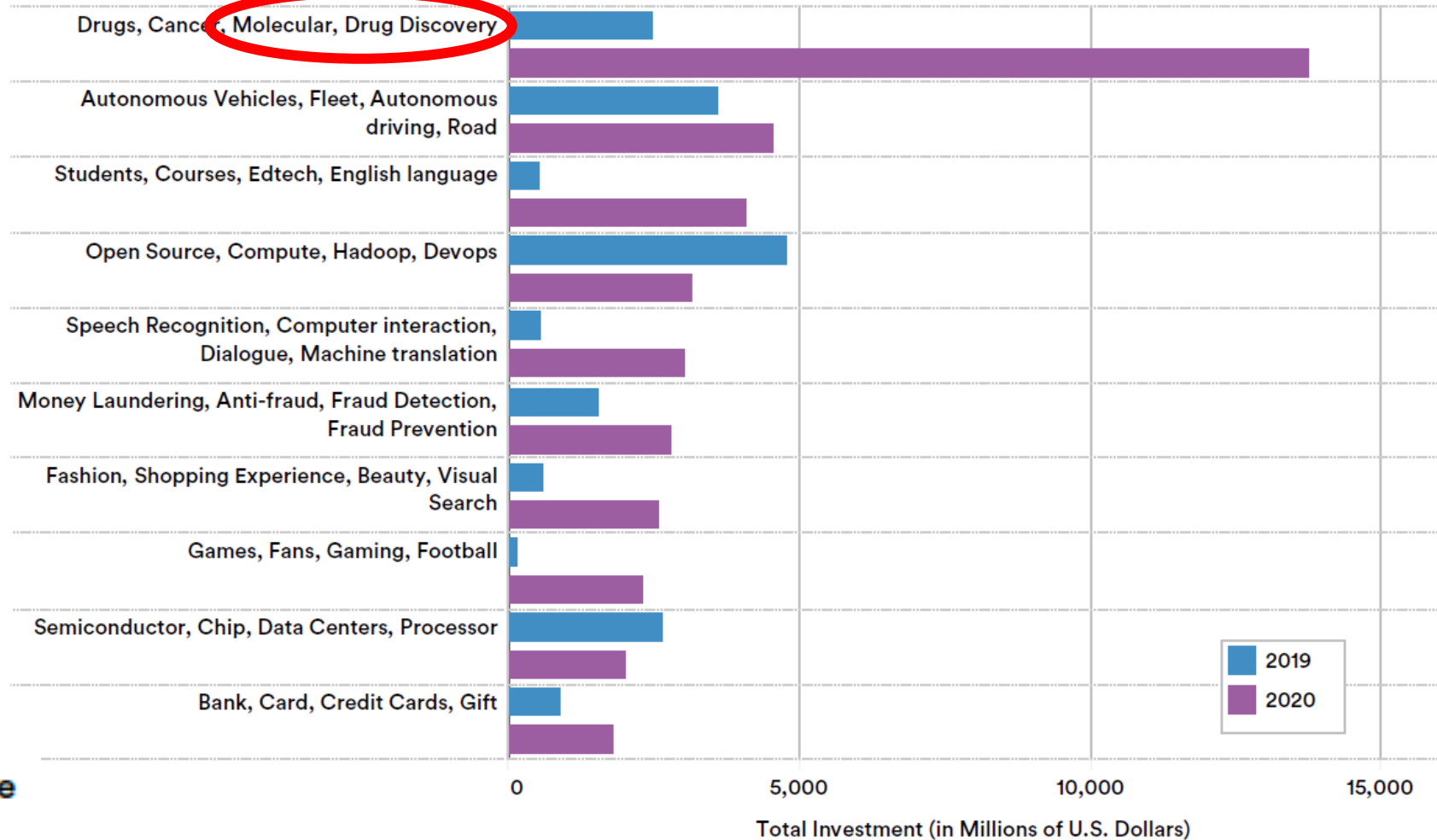
J.-C. Crivello, LINK : jean-claude.crivello@cnrs.fr

C. Barreteau, ICMPE : celine.barreteau@cnrs.fr

IA and chemistry

GLOBAL PRIVATE INVESTMENT in AI by FOCUS AREA, 2019 vs 2020

Source: CapIQ, Crunchbase, and NetBase Quid, 2020 | Chart: 2021 AI Index Report



Machine learning notions

1. Supervised / nonsupervised / reinforcement
2. Classification / Regression
3. Learning models, penalties, optimization
4. Training / Testing / CV
5. Bias / variance / overfitting → learning curves
6. Metrics / Loss functions
7. Neural networks

Machine learning resources

1. CNRS FIDLE:

<https://www.youtube.com/c/CNRSFormationFIDLE>

2. Coursera ML by Andrew NG:

<https://www.coursera.org/learn/machine-learning>

3. Google AI: <https://ai.google/education/>

4. Open Classrooms:

<https://openclassrooms.com/fr/courses/4011851-initiez-vous-au-machine-learning>

5. Scikit-learn: https://scikit-learn.org/stable/user_guide.html

...

Conclusions 1/2

Review of machine Learning applied to chemistry and material sciences :

- Large field of applications
- Golden age of Discriminative approaches ...
- Welcome to the Generative approaches age !

Conclusions 2/2

- Artificial intelligence is not able to think instead of human
- IA is just efficient for a dedicate learning
- We need data!



Human:

100 .10⁹ neurons
10³ synapses/neurons
200.10³ electric bonds

7 rules for data scientist workflow

1. Main goal? Definition of the ML problem
2. Collecting data (open data, rights,...)
3. Analyzing data, statistics, visualization
4. Cleaning data: drop missing, normalization, standardization, overlap info?
5. Preparation training/test sets ☐ a first training
6. Modelisation, for each model:
 - optimization of parameters, grid search
 - evaluation of CV score
7. Deployment

Project: based on Material Project

Home About Apps Documentation Forum API Tutorial

Search for materials information by chemistry, composition, or property

Explore Materials [Advanced Search Syntax](#)

by Elements Na-O search

1 H																	2 He																	
3 Li	4 Be															5 B	6 C	7 N	8 O	9 F	10 Ne													
11 Na	12 Mg															13 Al	14 Si	15 P	16 S	17 Cl	18 Ar													
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr																	
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe																	
55 Cs	56 Ba	57-71 La-Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn																	
87 Fr	88 Ra	89-103 Ac-Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112 Cn																							
																		57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb	71 Lu		
																		89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No	103 Lr		

of elements
e.g., 4 or >2 & <6

excluded elements
Cl Br

Submit

External Provenance
 ICSD [?]
 Exptl. ICSD [?]

Material Tags
ingreite

Band Gap (eV)
0 10

Energy Above Hull
0 6

Formation Energy
-4 4

<https://materialsproject.org/>