Artificial intelligence (AI) for medicine and oncology: some concepts and applications

Julien VIBERT
MD Medical Oncology - PhD Bioinformatics
DITEP - Gustave Roussy

IFSBM Module 12
Gustave Roussy
20 January 2022

Outline of presentation

http://www.unehistoiredeplumes.fr/a-vol-doiseau/

- Simple intro to AI
- AI in medicine
- AI in oncology
- AI in cancer genomics and transcriptomics
- *Use case*: AI-based classification of cancers of unknown primary -> Later session
- Conclusions and perspectives

Outline of presentation

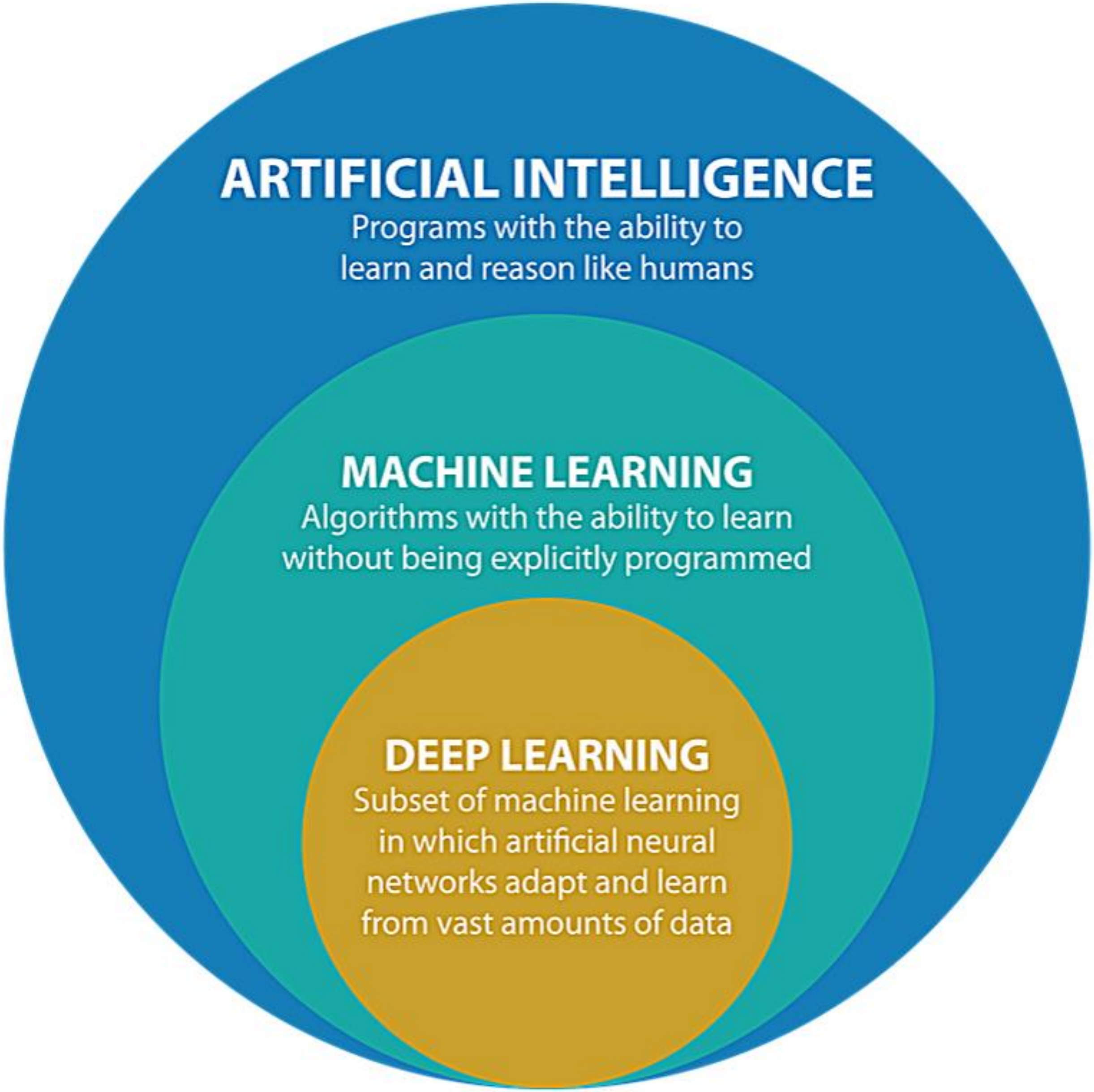
http://www.unehistoiredeplumes.fr/a-vol-doiseau/

- Simple intro to AI
- AI in medicine
- AI in oncology
- AI in cancer genomics and transcriptomics
- *Use case*: AI-based classification of cancers of unknown primary -> Later session
- Conclusions and perspectives

Don't hesitate to interrupt and ask questions!

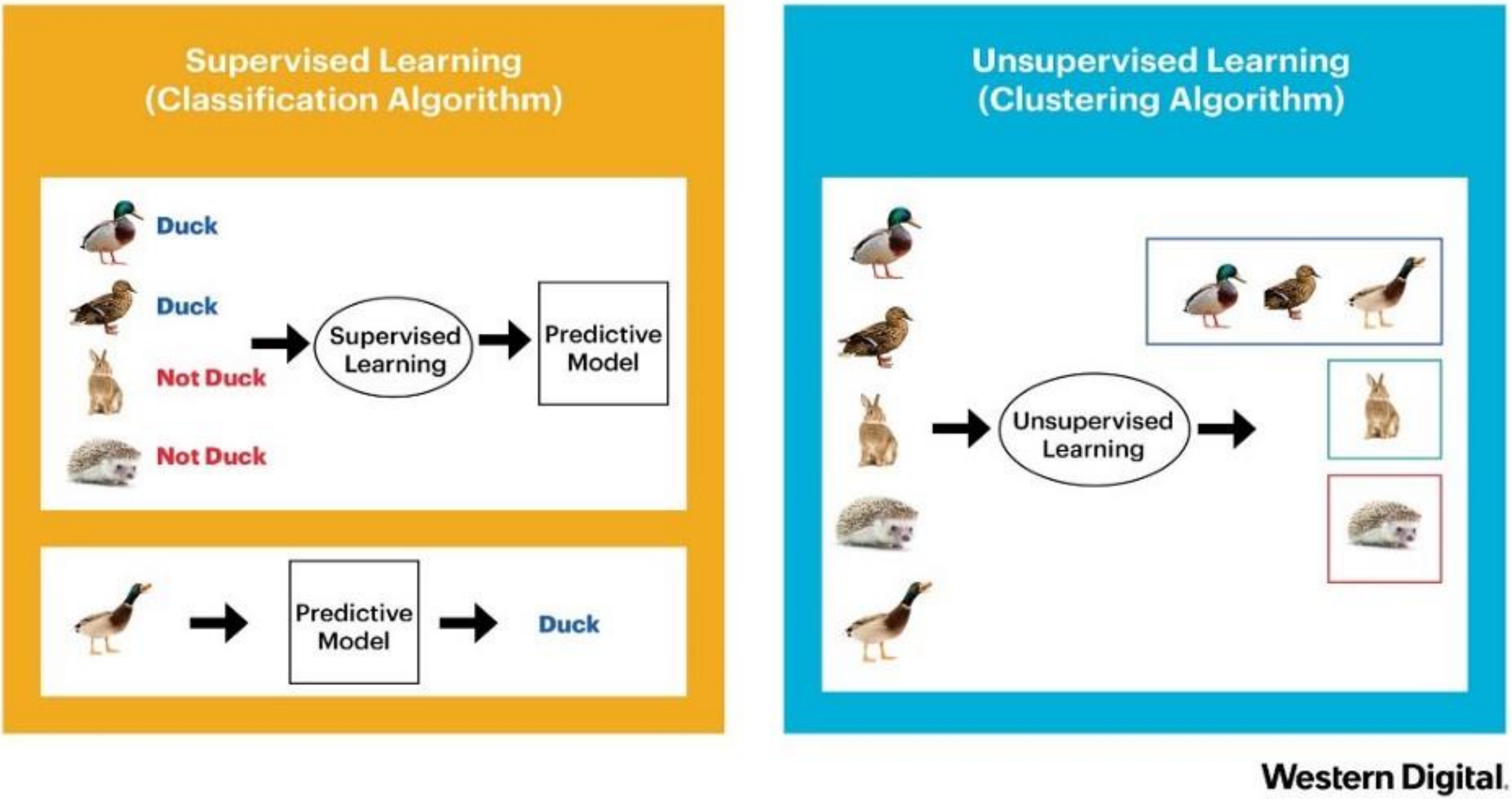
Artificial intelligence (AI): some concepts

Definitions



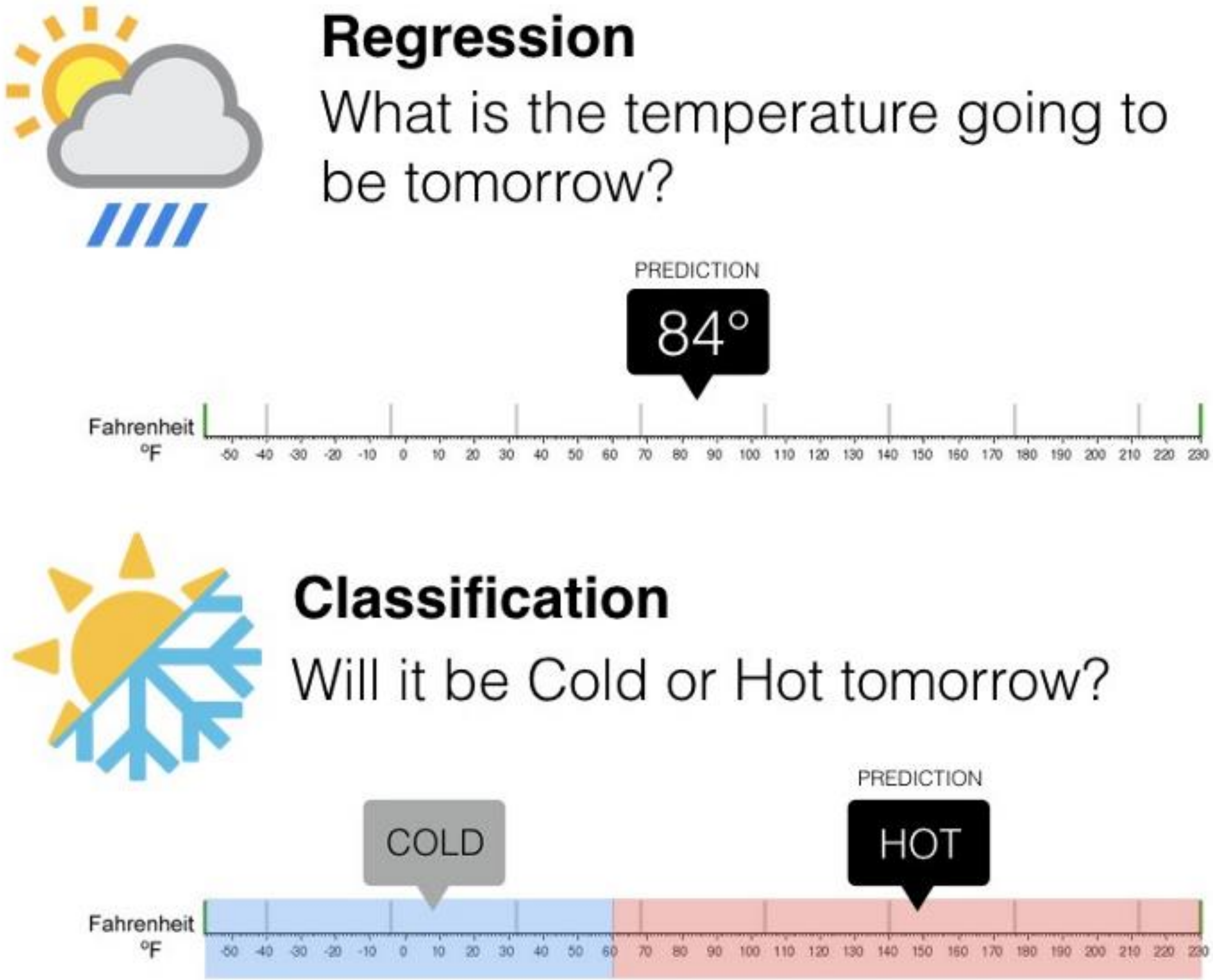
https://blog.digitalogy.co/the-difference-between-artificial-intelligence-machine-learning-and-deep-learning/

Supervised/unsupervised learning



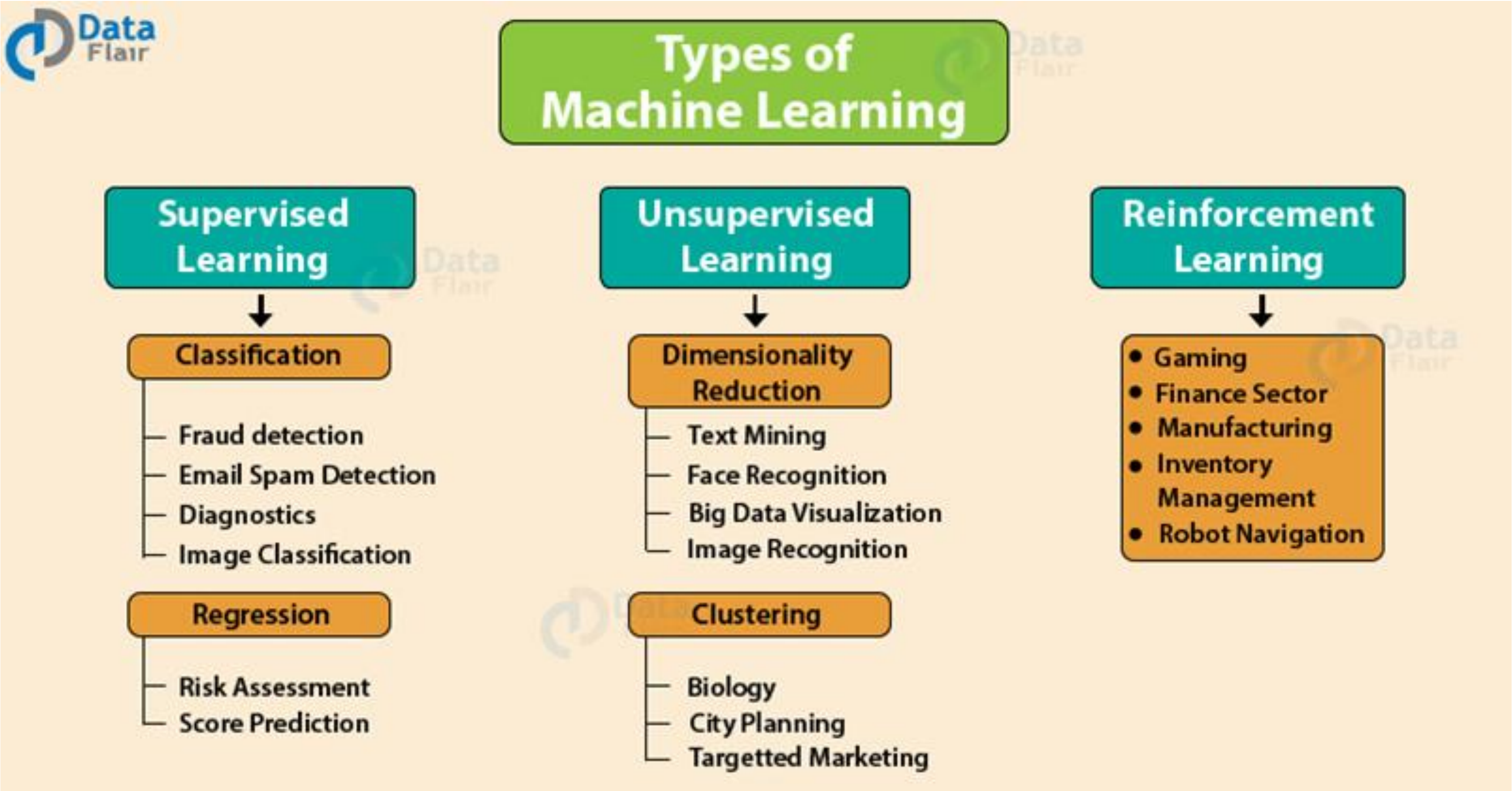
https://laptrinhx.com/supervised-vs-unsupervised-learning-1861950659/

Regression or classification



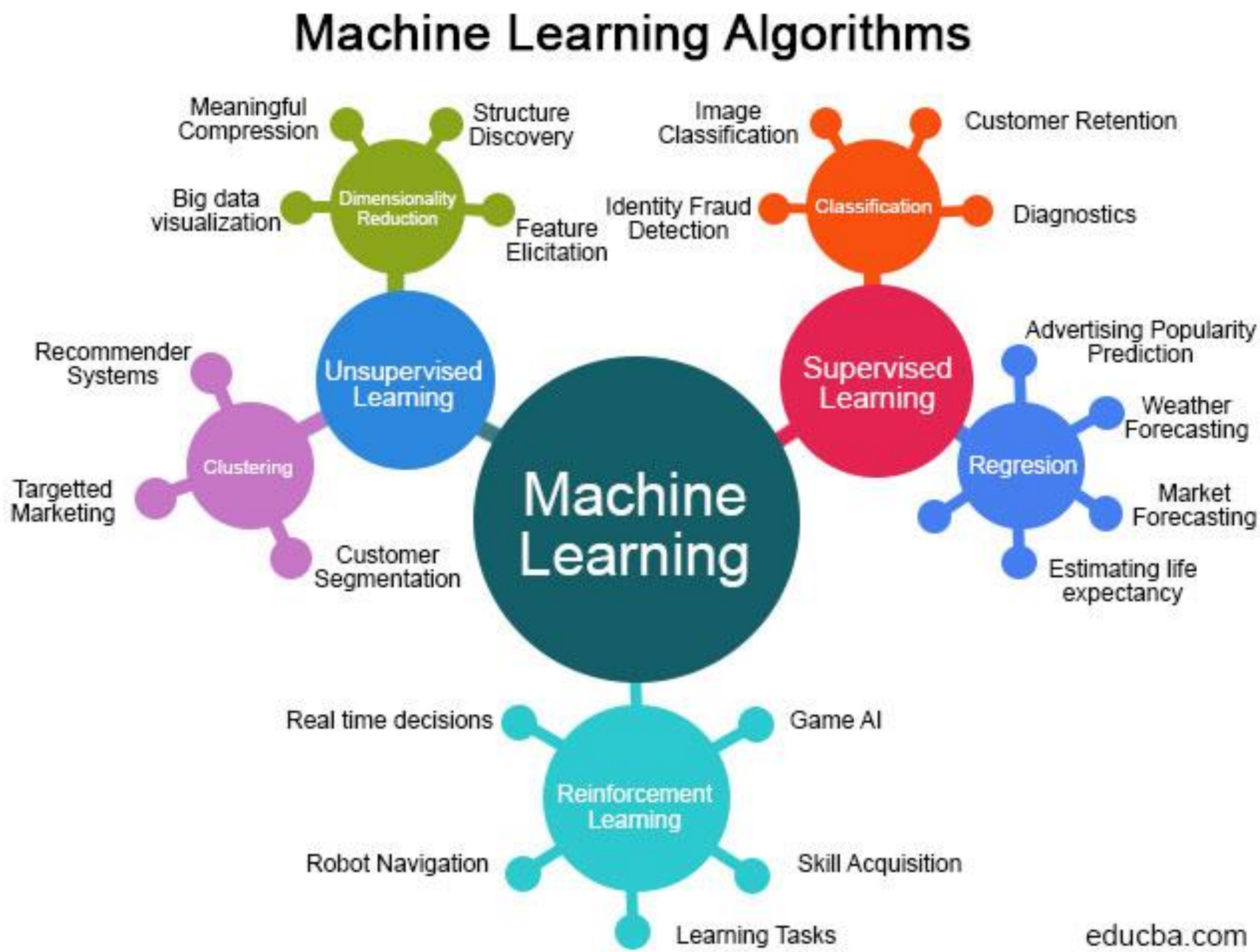
https://towardsdatascience.com/regression-or-classification-linear-or-logistic-f093e8757b9c

Main types of machine learning



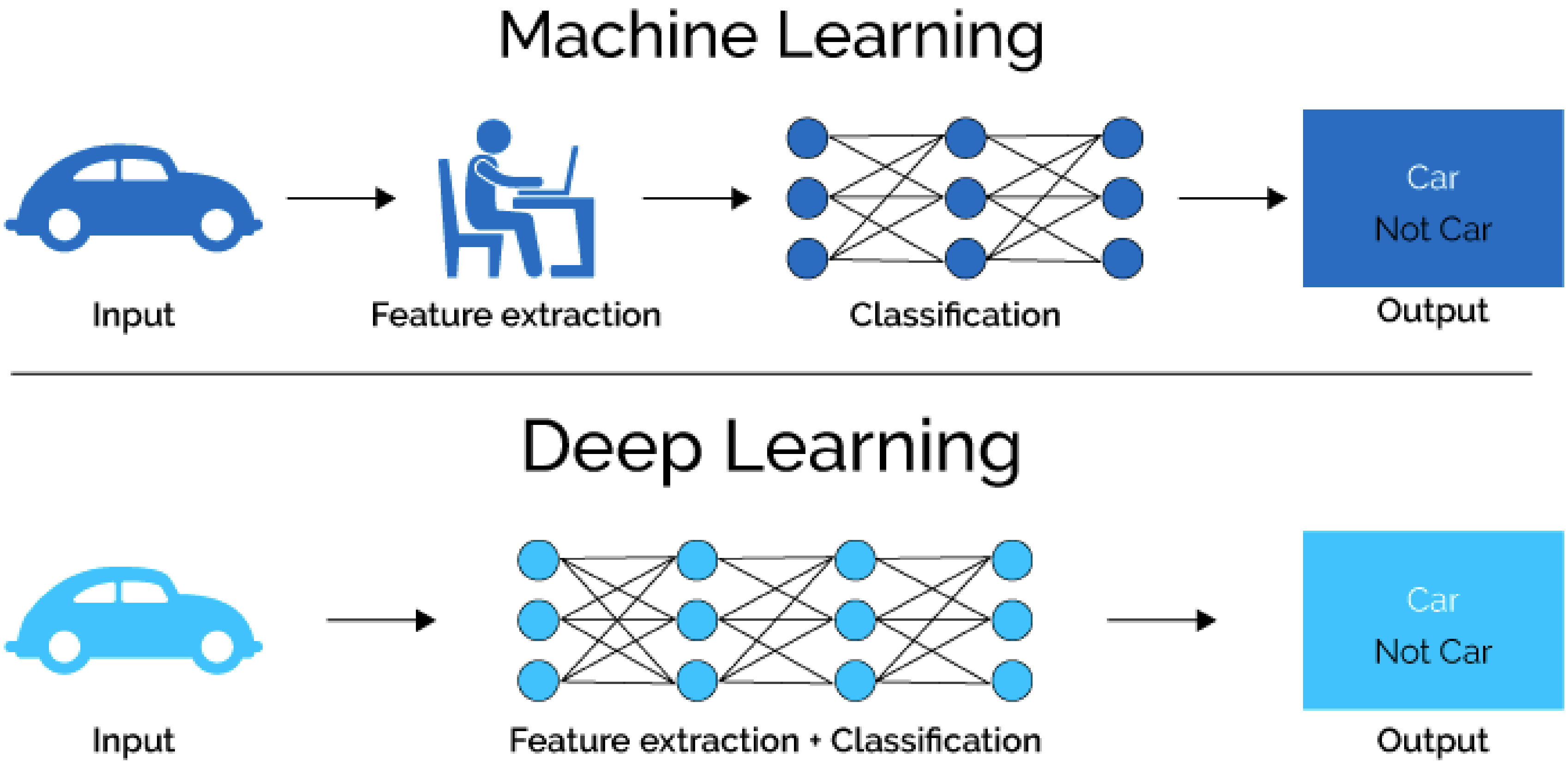
https://data-flair.training/blogs/types-of-machine-learning-algorithms/

Algorithms are proliferating…



https://www.educba.com/machine-learning-algorithms/

Machine learning vs deep learning



https://lawtomated.com/tag/unsupervised-learning/

What algorithms are eating: « Big data »

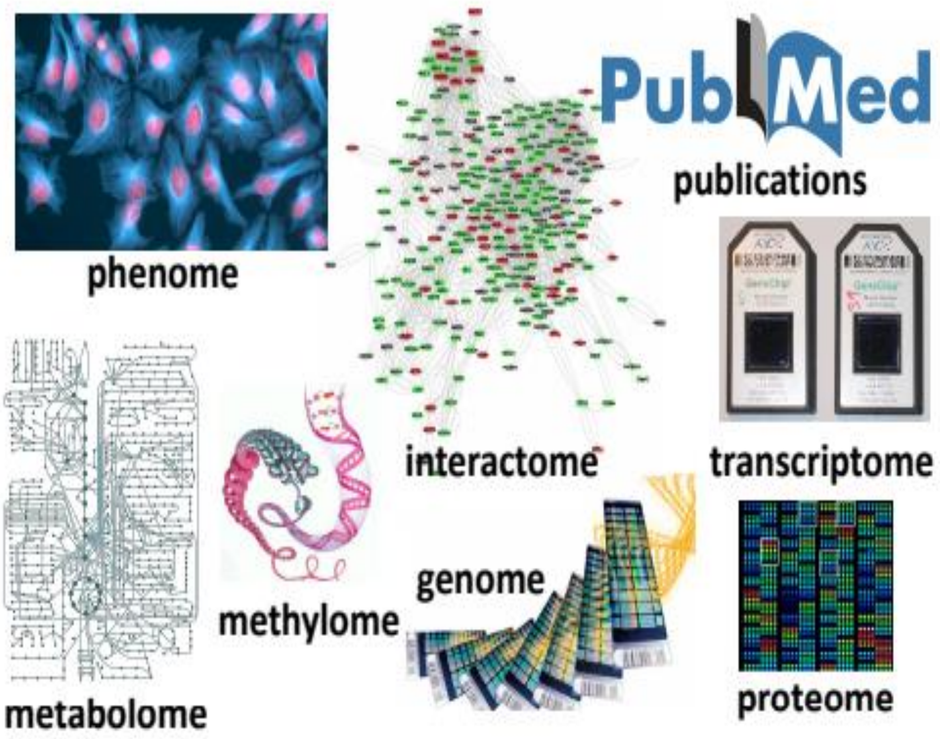
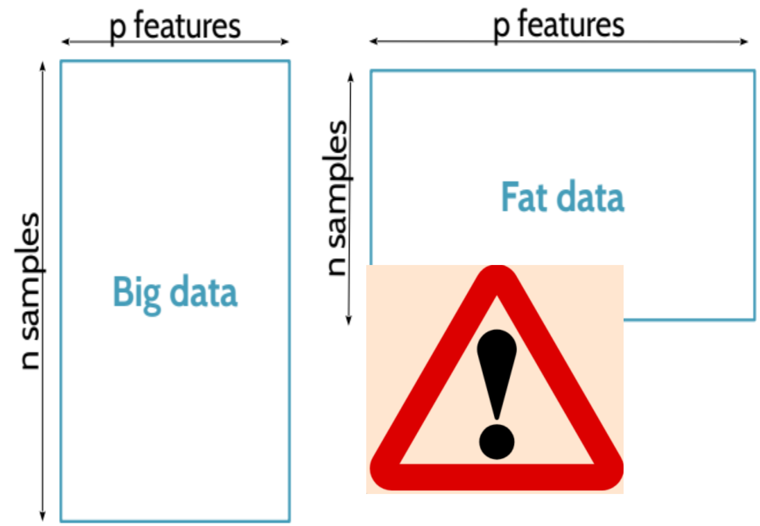


Image sources: ajc1@ flickr; Zlir'a@wikimedia

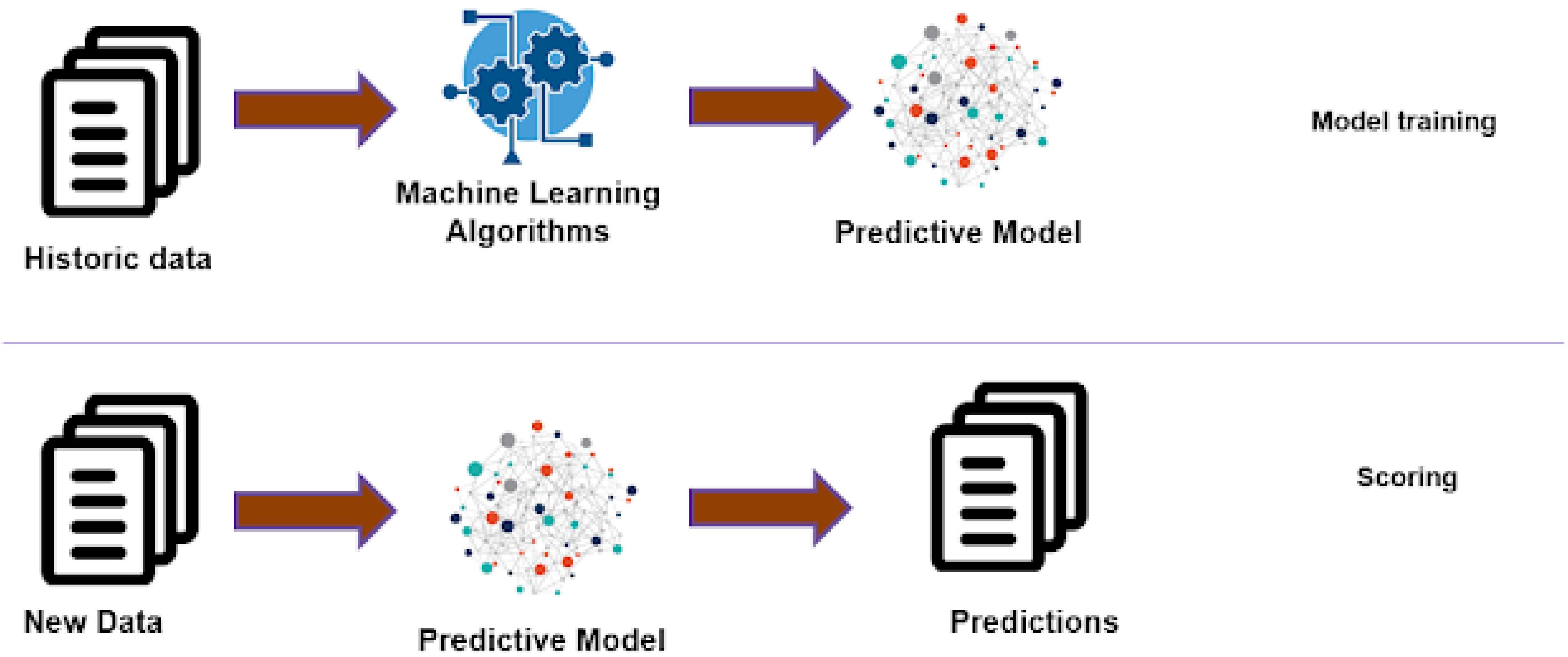


E.g. **G**enome-**W**ide **A**ssociation **S**tudies (GWAS):

- $p = 10^5 - 10^7$ **S**ingle **N**ucleotide **P**olymorphisms (SNPs)
- $n = 10^2 - 10^4$ samples.

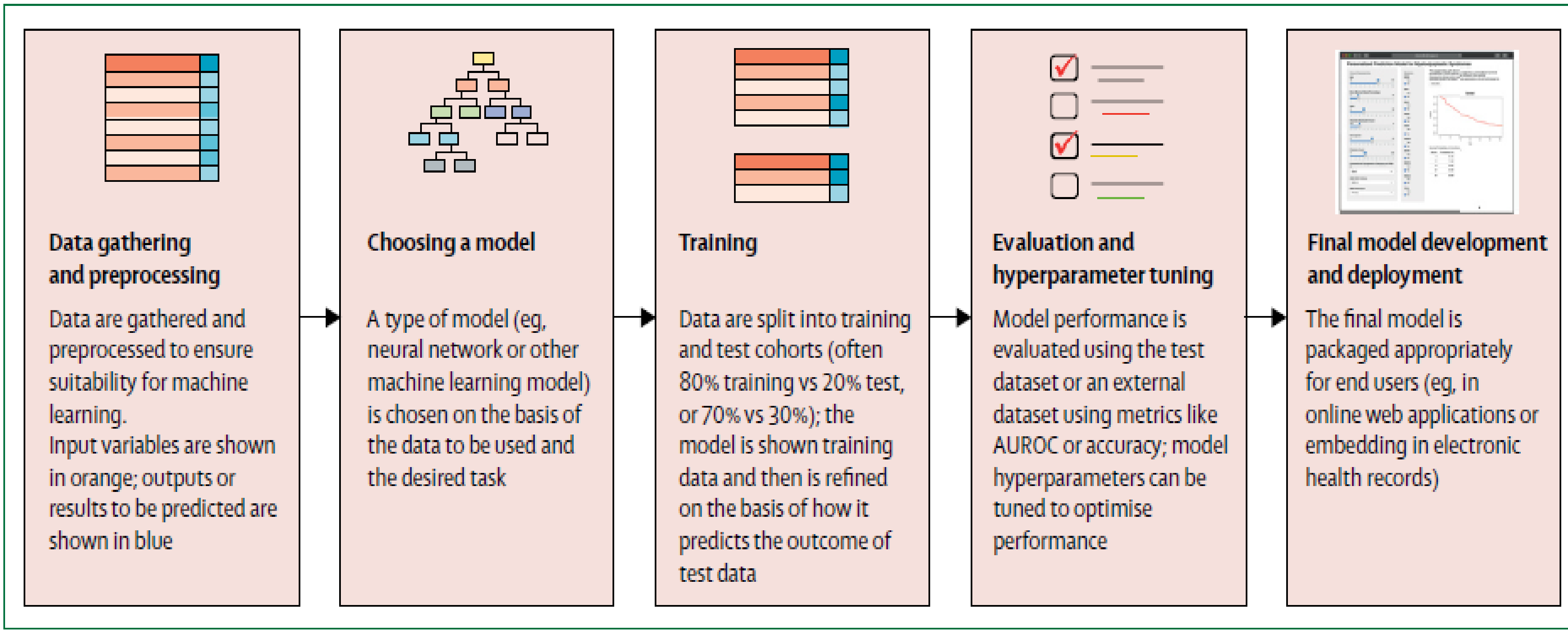
http://cazencott.info/dotclear/public/publications/2016-01-13-azencott.pdf

Typical machine learning workflow



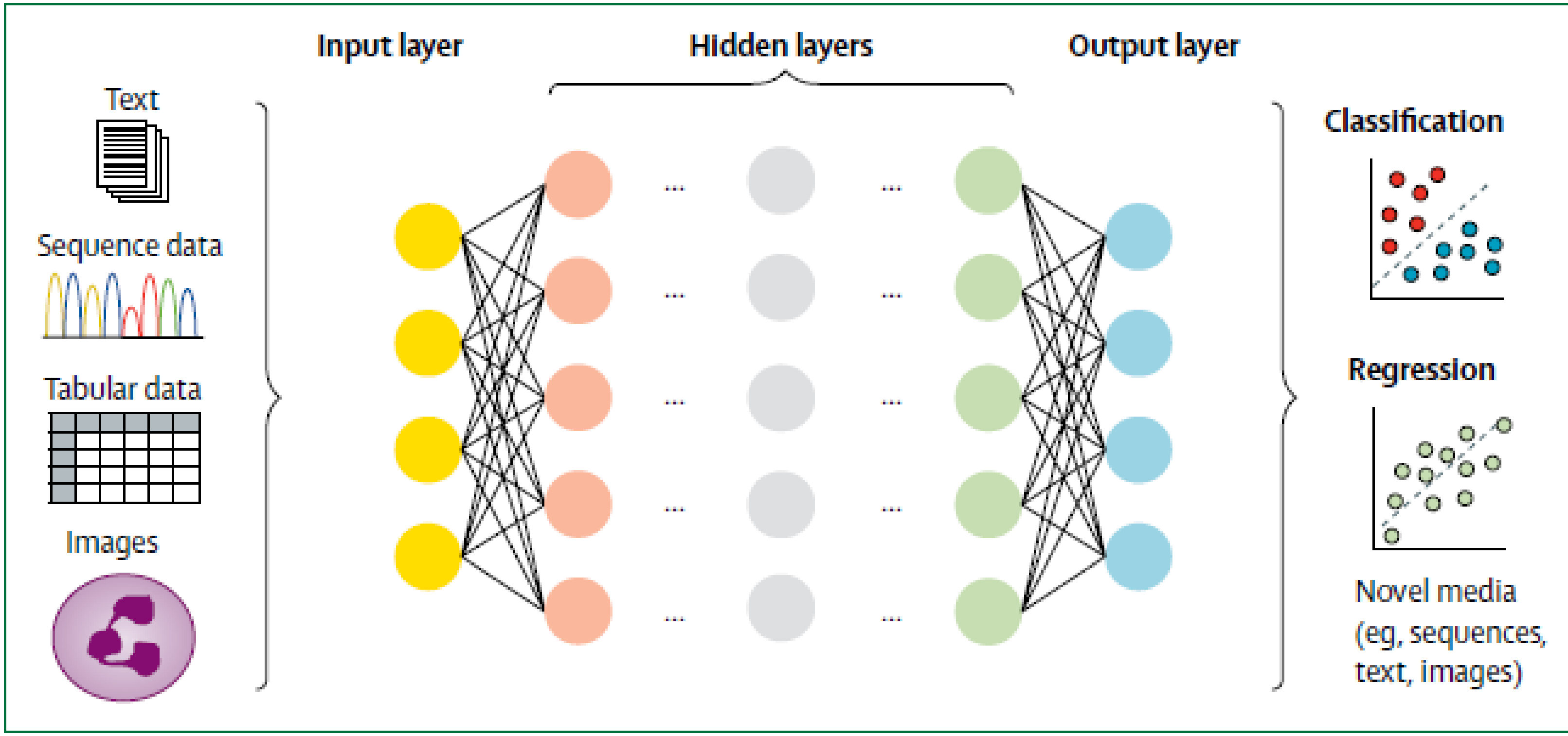
http://singaporebusinessintelligence.blogspot.com/2018/10/what-is-automated-machine-learning.html

Typical machine learning workflow (2)



Radakovich et al., Lancet Haem 2020

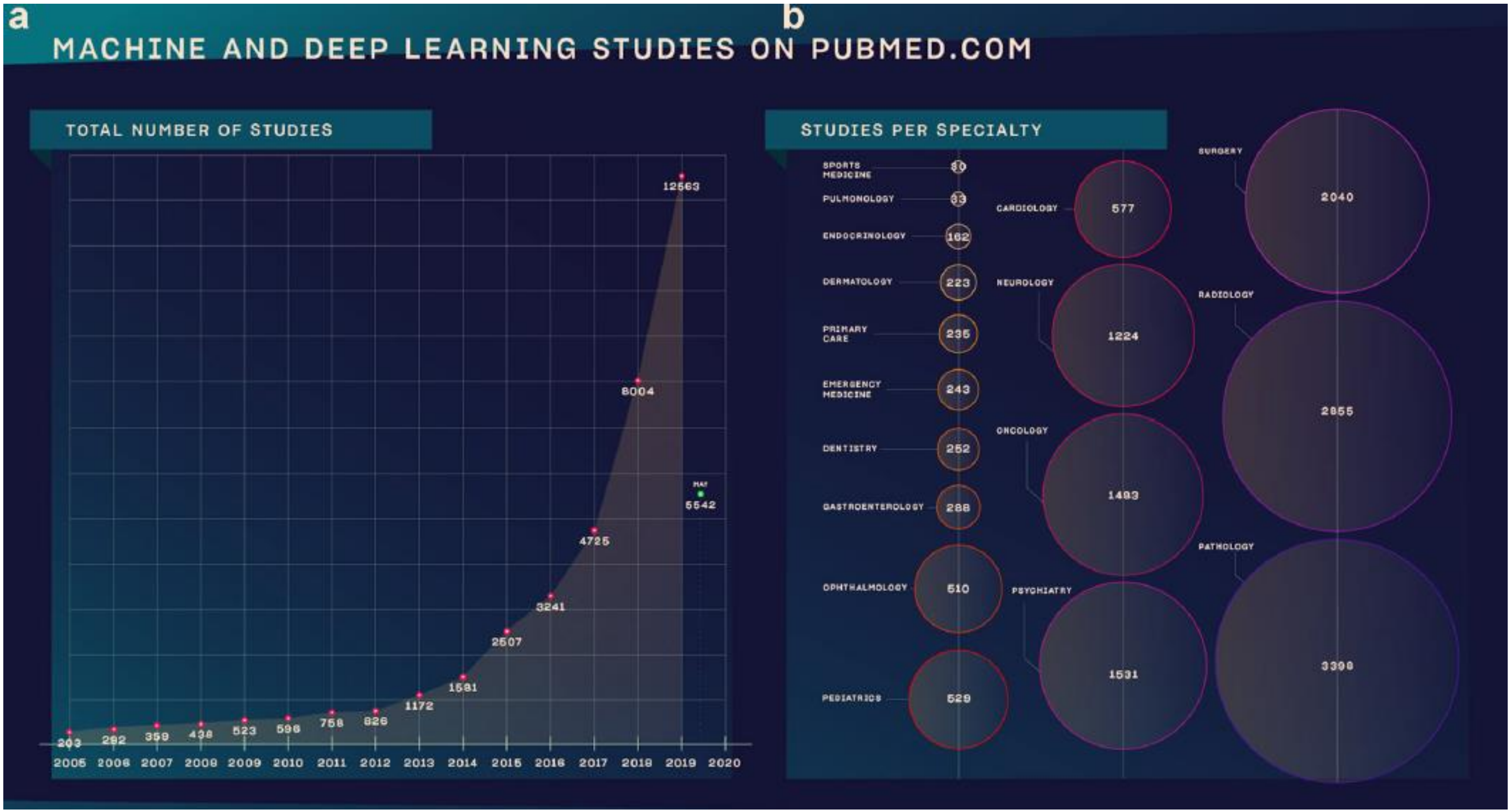
Typical scheme of deep learning



Radakovich et al., Lancet Haem 2020

AI in medicine:
what, why, how?

AI and medicine: articles are proliferating…



Mesko et al., Npj Digital Med, 2020

Potential applications in medicine (non-exhaustive list)

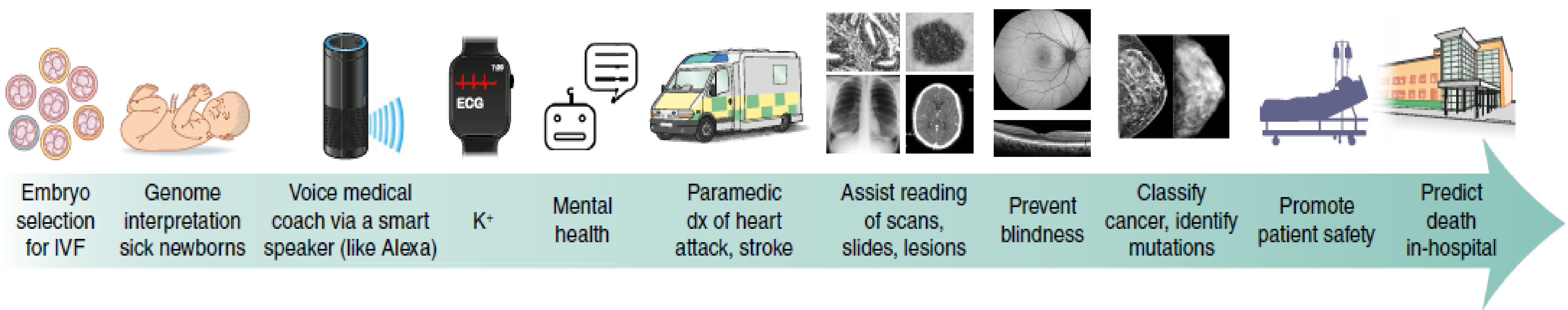


Fig. 2 | Examples of AI applications across the human lifespan. dx, diagnosis; IVF, in vitro fertilization K^+, potassium blood level. Credit: Debbie Maizels/ Springer Nature

Topol, Nat Med 2019

Potential applications in medicine

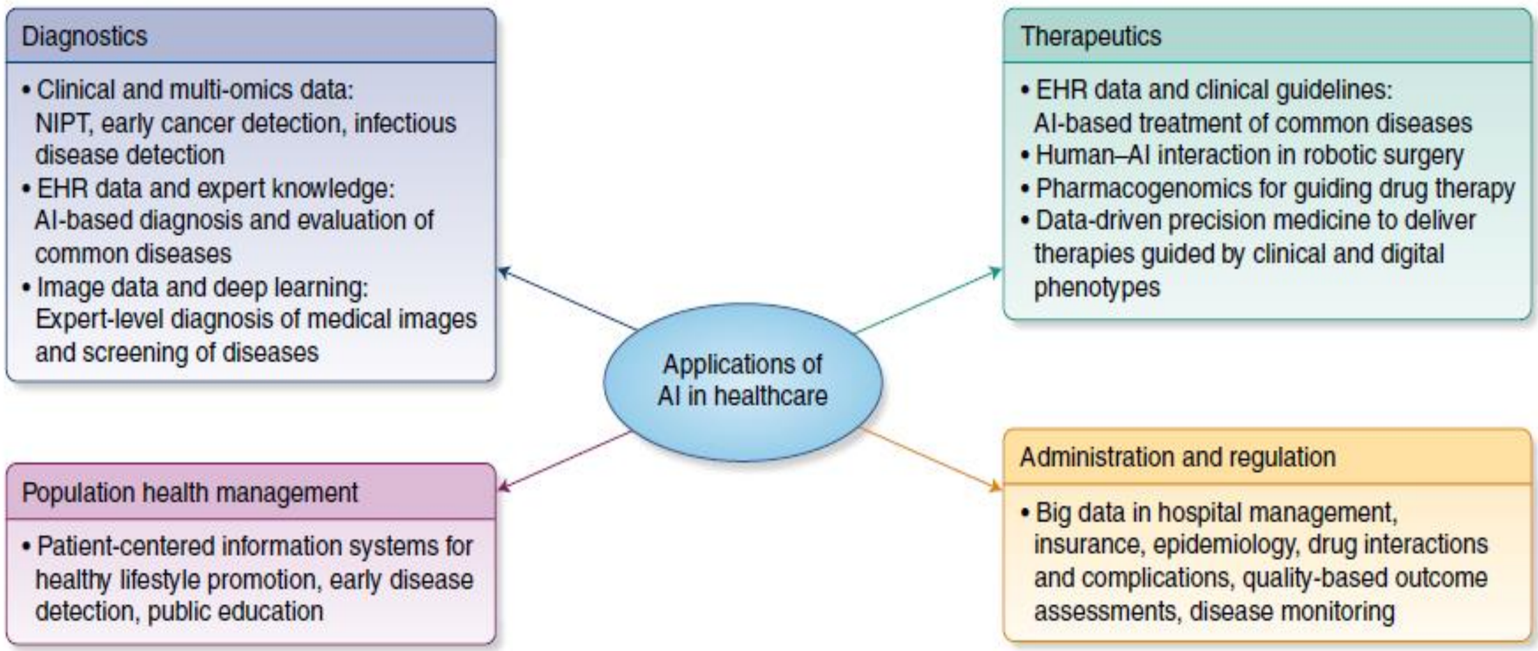
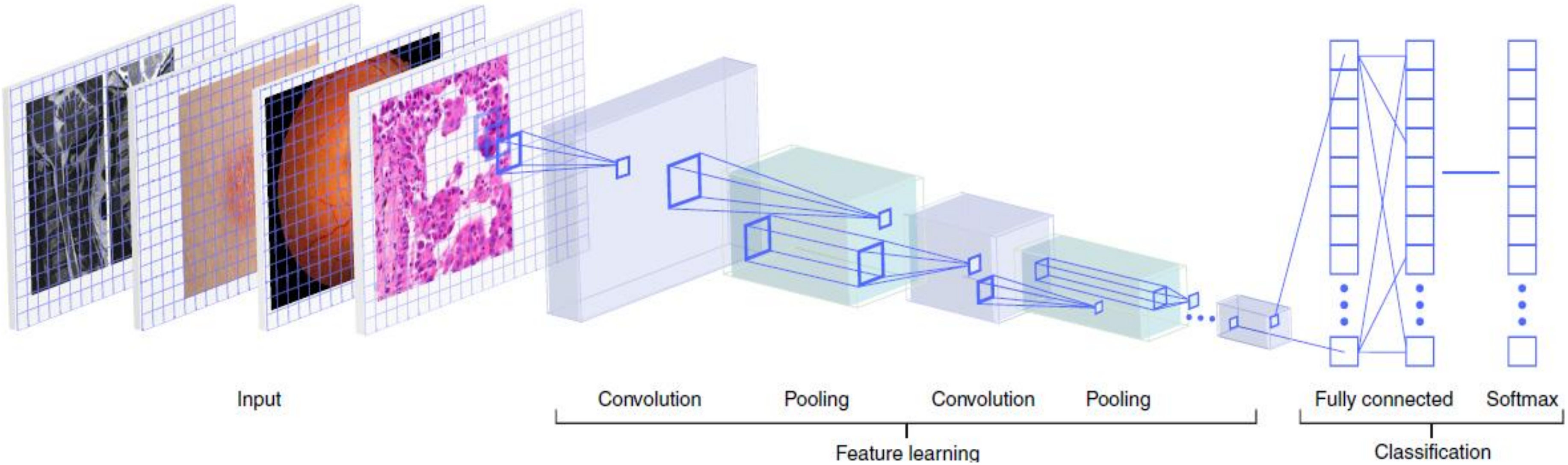


Fig. 1 | Potential roles of AI-based technologies in healthcare. In the healthcare space, AI is poised to play major roles across a spectrum of application domains, including diagnostics, therapeutics, population health management, administration, and regulation. NIPT, noninvasive prenatal test. Credit: Debbie Maizels/Springer Nature

He et al., Nat Med 2019

Pioneer applications: medical image analysis



Detecting **lung cancer** from CT Scans

Assess **cardiac health** from electrocardiograms

Classify **skin lesions** from images of the skin

Identify **retinopathy** from eye images

https://www.datarevenue.com/en-blog/artificial-intelligence-in-medicine

Esteva et al., Nat Med 2019

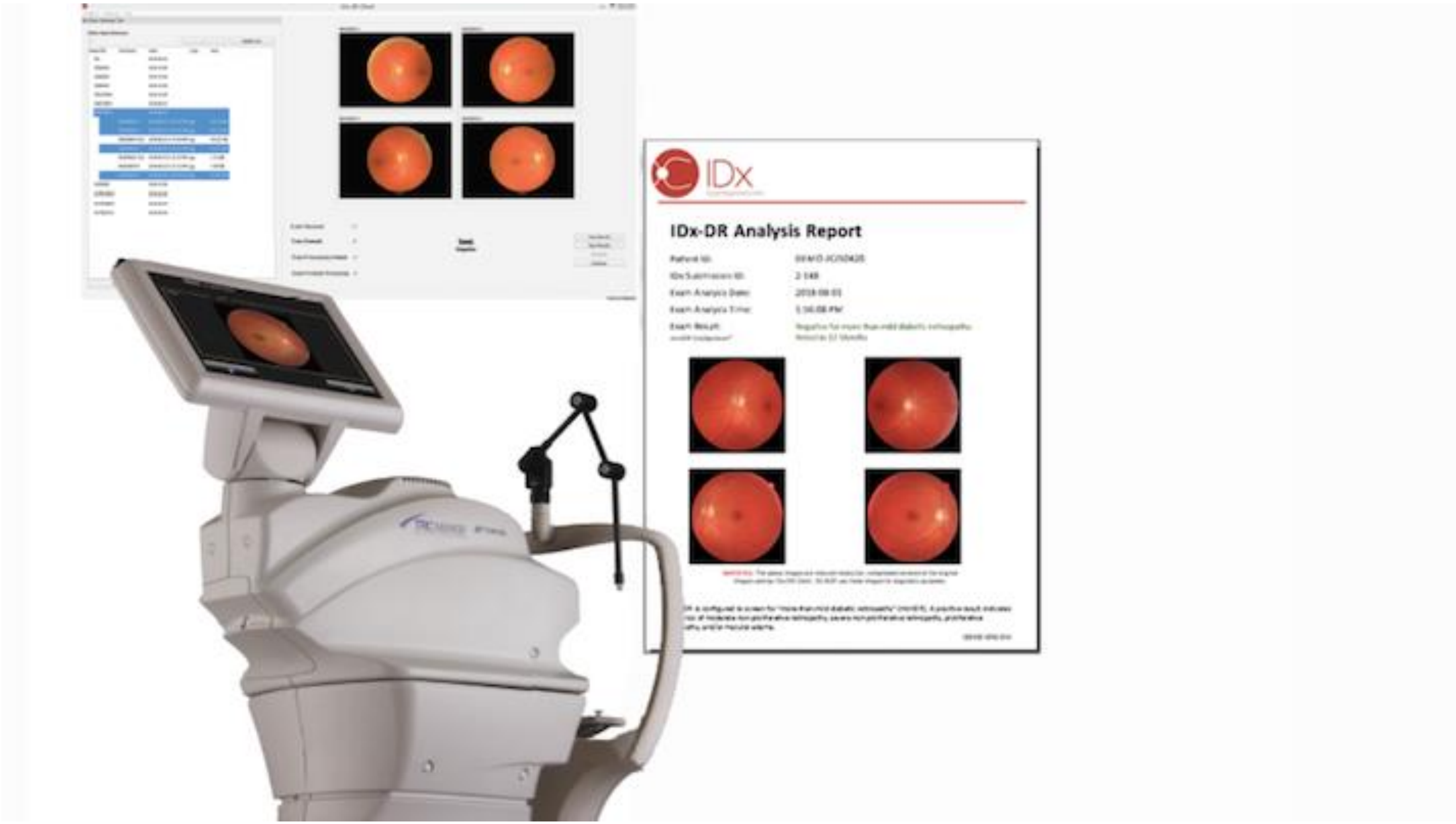
Some AI tools are already used in clinical practice

Table 2 | FDA AI approvals are accelerating

Company	FDA Approval	Indication
Apple	September 2018	Atrial fibrillation detection
Aidoc	August 2018	CT brain bleed diagnosis
iCAD	August 2018	Breast density via mammography
Zebra Medical	July 2018	Coronary calcium scoring
Bay Labs	June 2018	Echocardiogram EF determination
Neural Analytics	May 2018	Device for paramedic stroke diagnosis
IDx	April 2018	Diabetic retinopathy diagnosis
Icometrix	April 2018	MRI brain interpretation
Imagen	March 2018	X-ray wrist fracture diagnosis
Viz.ai	February 2018	CT stroke diagnosis
Arterys	February 2018	Liver and lung cancer (MRI, CT) diagnosis
MaxQ-AI	January 2018	CT brain bleed diagnosis
Alivecor	November 2017	Atrial fibrillation detection via Apple Watch
Arterys	January 2017	MRI heart interpretation

Topol, Nat Med 2019

Screening for diabetic retinopathy



https://dxs.ai/

Perspective: the « virtual medical coach »

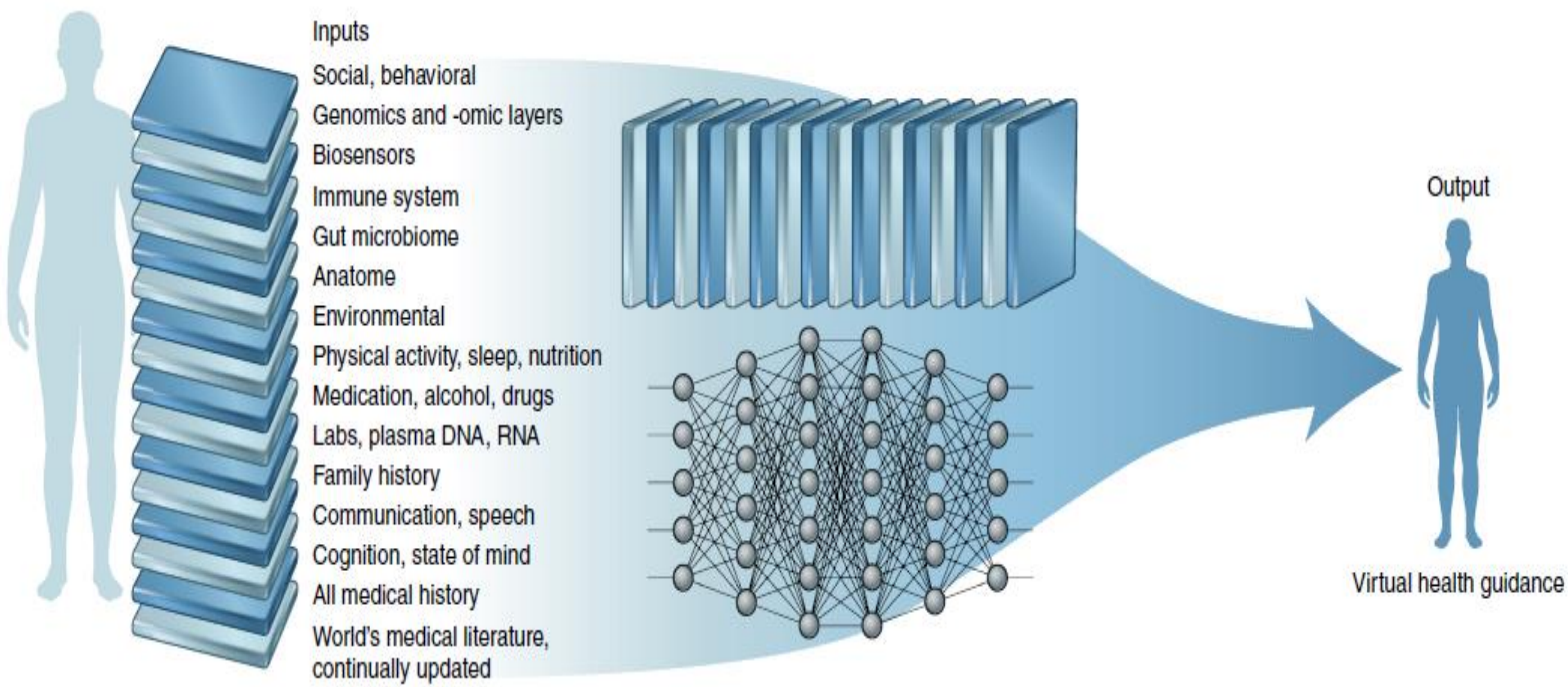


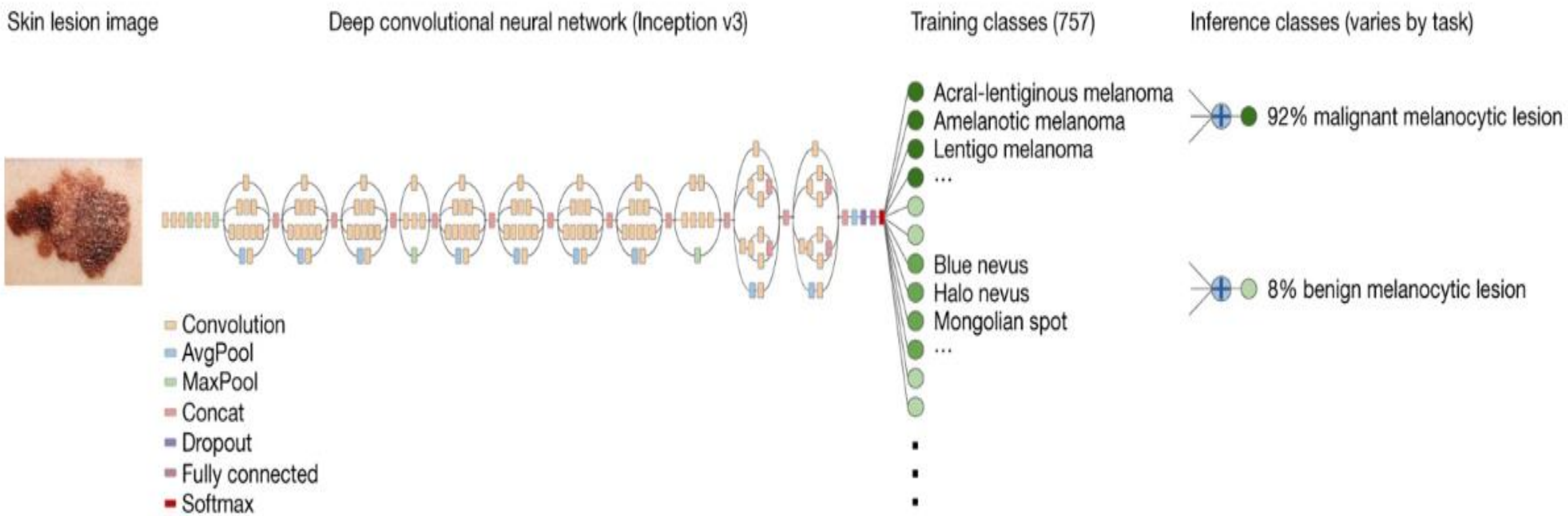
Fig. 3 | The virtual medical coach model with multi-modal data inputs and algorithms to provide individualized guidance. A virtual medical coach that uses comprehensive input from an individual that is deep learned to provide recommendations for preserving the person's health. Credit: Debbie Maizels/ Springer Nature

Topol, Nat Med 2019

AI in oncology:
some applications

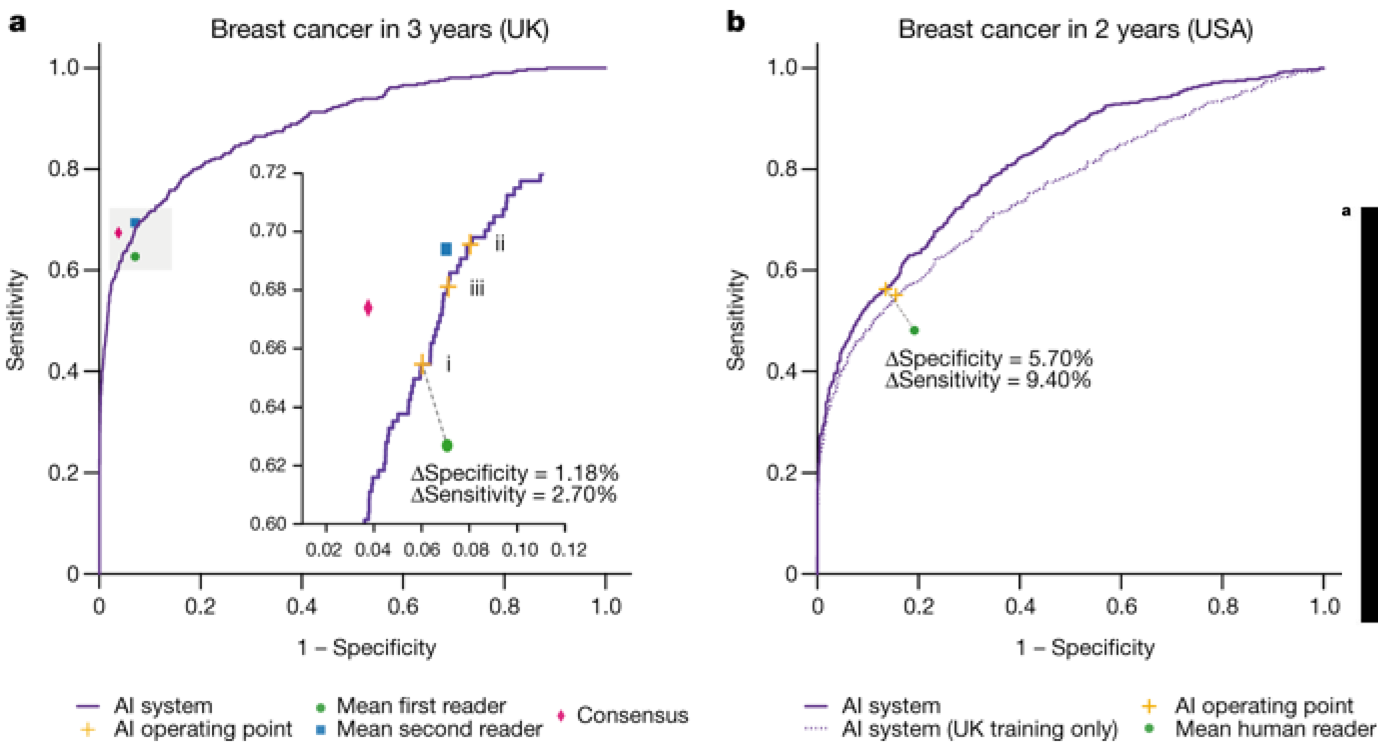
Skin cancer classification

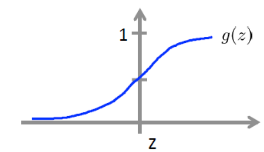
From: Dermatologist-level classification of skin cancer with deep neural networks

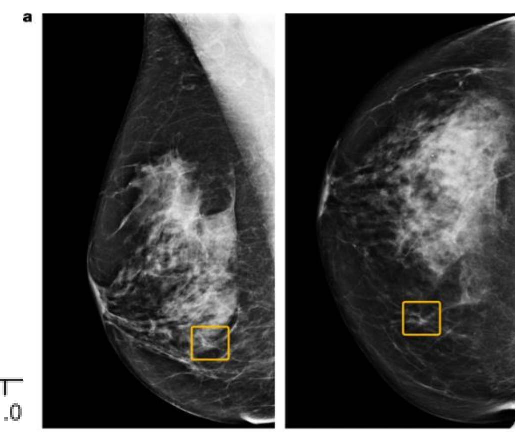


Esteva et al., Nature 2017

Breast cancer screening



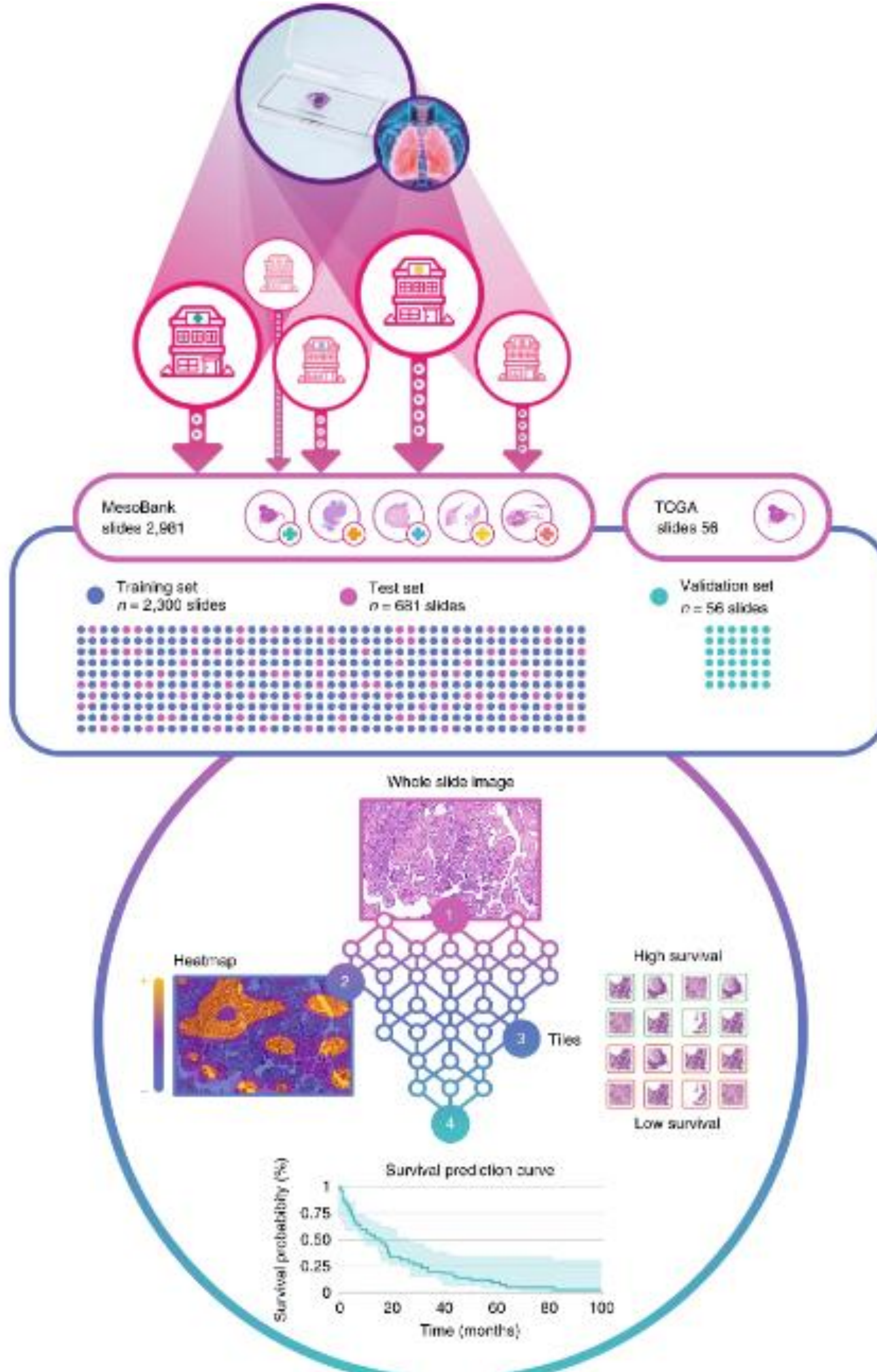




International evaluation of an AI system for breast cancer screening, Scott Mayer McKinney, Marcin Sieniek, […]Shravya Shetty, Nature volume 577, pages89–94(2020)

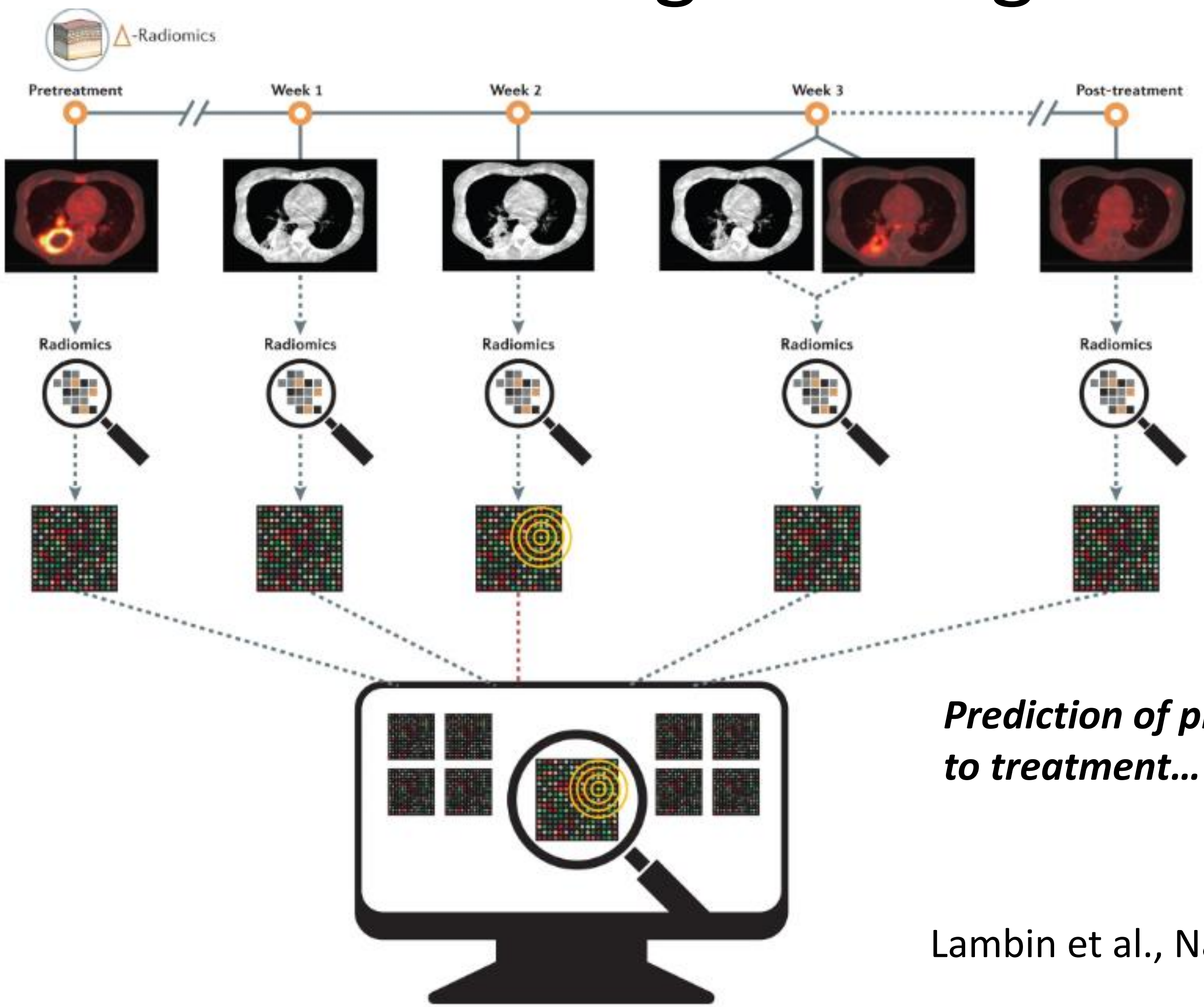
Loïc Verlingue

Prediction of mesothelioma prognosis based on pathological slides



Courtiol et al.,
Nature Med 2019

Radiomics: « texture » analysis of radiological images



Prediction of prognosis, response to treatment…

Lambin et al., Nat Rev Clin Onc 2017

AI in hematology

	Sample size	Application	Method	Results
Kimura et al (2019)[21]	3261 peripheral smears	Leucocyte classification; distinguishing aplastic anaemia and myelodysplastic syndrome	CNN, gradient boosting	Sensitivity vs specificity was 93·5% vs 96·0% for leucocyte detection, and 96·2% vs 100% for aplastic anaemia vs myelodysplastic syndrome differentiation
Achi et al (2019)[22]	128 patients	Differentiating diffuse large B-cell lymphoma, small lymphocytic lymphoma, Burkitt lymphoma, and normal lymph nodes	CNN	95% accuracy per slide; 100% accuracy per patient
Li (2019)[23]	41 patients	Detection of acute myeloid leukaemia bone marrow involvement via PET-CT	Manual feature engineering (PyRadiomics)	Sensitivity vs specificity was 87·5% vs 89·5% for bone marrow involvement; outperformed visual inspection of scans
Milgrom et al (2019)[24]	251 patients	Predicting refractory Hodgkin lymphoma from PET-CT	CNN	AUROC of 0·95 for model vs 0·78 for tumour volume; 0·65 for standardised uptake value
Moraes et al (2019)[25]	283 patients	Differential diagnosis of chronic lymphocytic leukaemia and B cell lymphomas via flow cytometry	Decision tree	95% inclusion of correct diagnosis in differential diagnosis; 66% definitive diagnosis
Ni et al (2016)[26]	51 patients	Detection of MRD in acute myeloid leukaemia via flow cytometry	Support vector machine	Similar performance with manual flow analysis (concordance=0·986)
Fuse (2019)[27]	217 patients	Prediction of acute leukaemia relapse after allogeneic stem cell transplantation	Decision tree	0·75 AUROC for relapse after transplantation
Goswami et al (2019)[28]	347 patients	Risk stratification for autologous stem cell transplantation in multiple myeloma	Decision tree	Significant risk-stratification and identification of high-risk features
Nazha (2019)[29]	433 and 113 patients	Predicting resistance to hypomethylating agents in patients with myelodysplastic syndrome on the basis of NGS myeloid malignancy panel	Recommender algorithm	Improved stratification of patients by risk of resistance to a hypomethylating agent
Gal et al (2019)[30]	493 patients	Response to induction therapy in paediatric acute myeloid leukaemia	K-nearest neighbours	0·84 AUROC for response to induction
Candia (2015)[31]	60 patients	Unsupervised analysis of microRNA expression signatures in acute myeloid leukaemia and acute lymphocytic leukaemia	Dimension reduction, network analysis	Identification of novel microRNA signatures in acute myeloid leukaemia and acute lymphocytic leukaemia

AUROC=area under the receiver operating characteristic curve. CNN=convolutional neural network. MRD=minimal residual disease. NGS=next-generation sequencing.

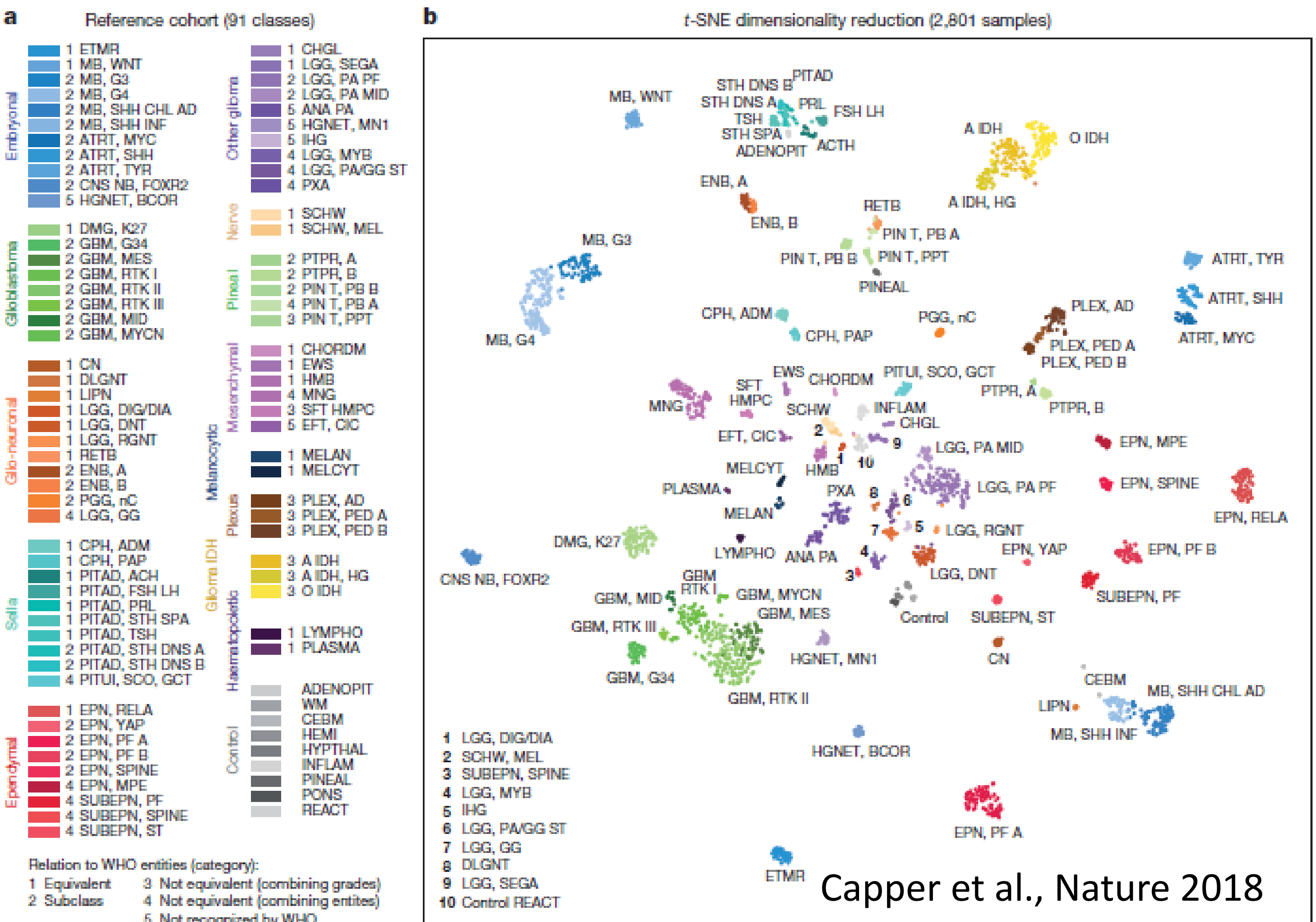
***Table:* Representative artificial intelligence in malignant haematology publications**

AI in cancer genomics and transcriptomics

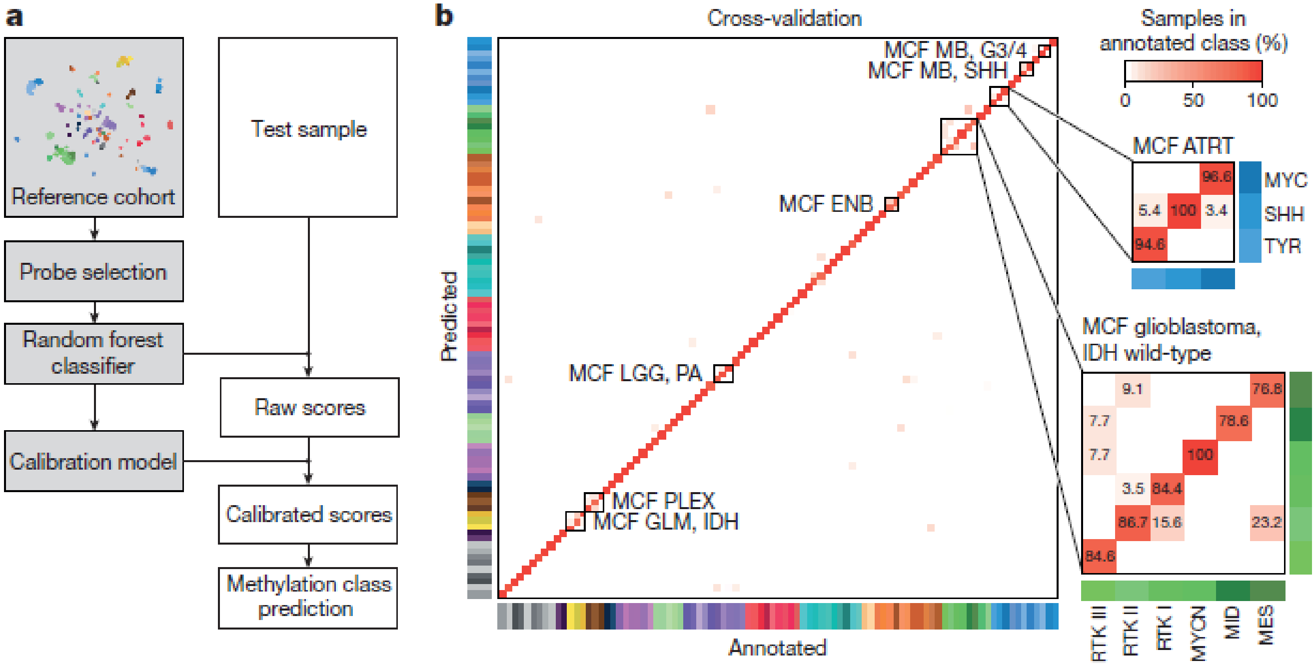
ARTICLE

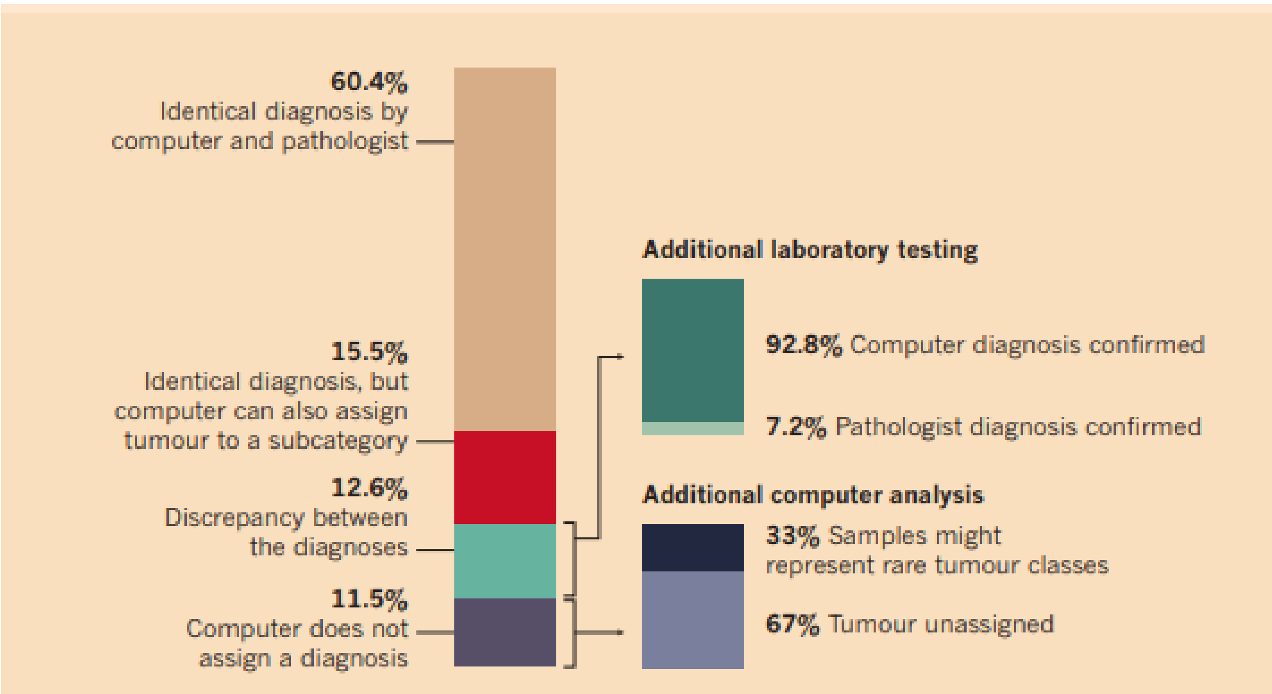
doi:10.1038/nature26000

DNA methylation-based classification of central nervous system tumours



Capper et al., Nature 2018

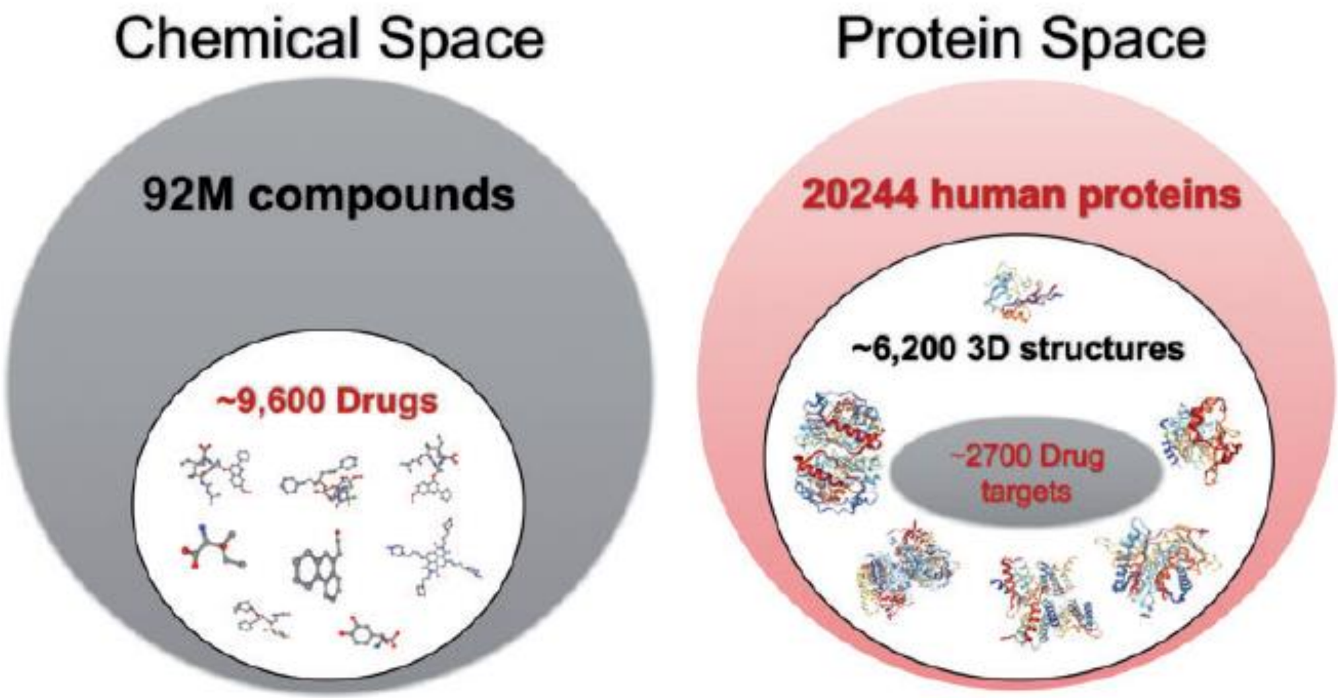




Capper et al., Nature 2018

- *Cf later session: prediction of primary tumor using transcriptomics*

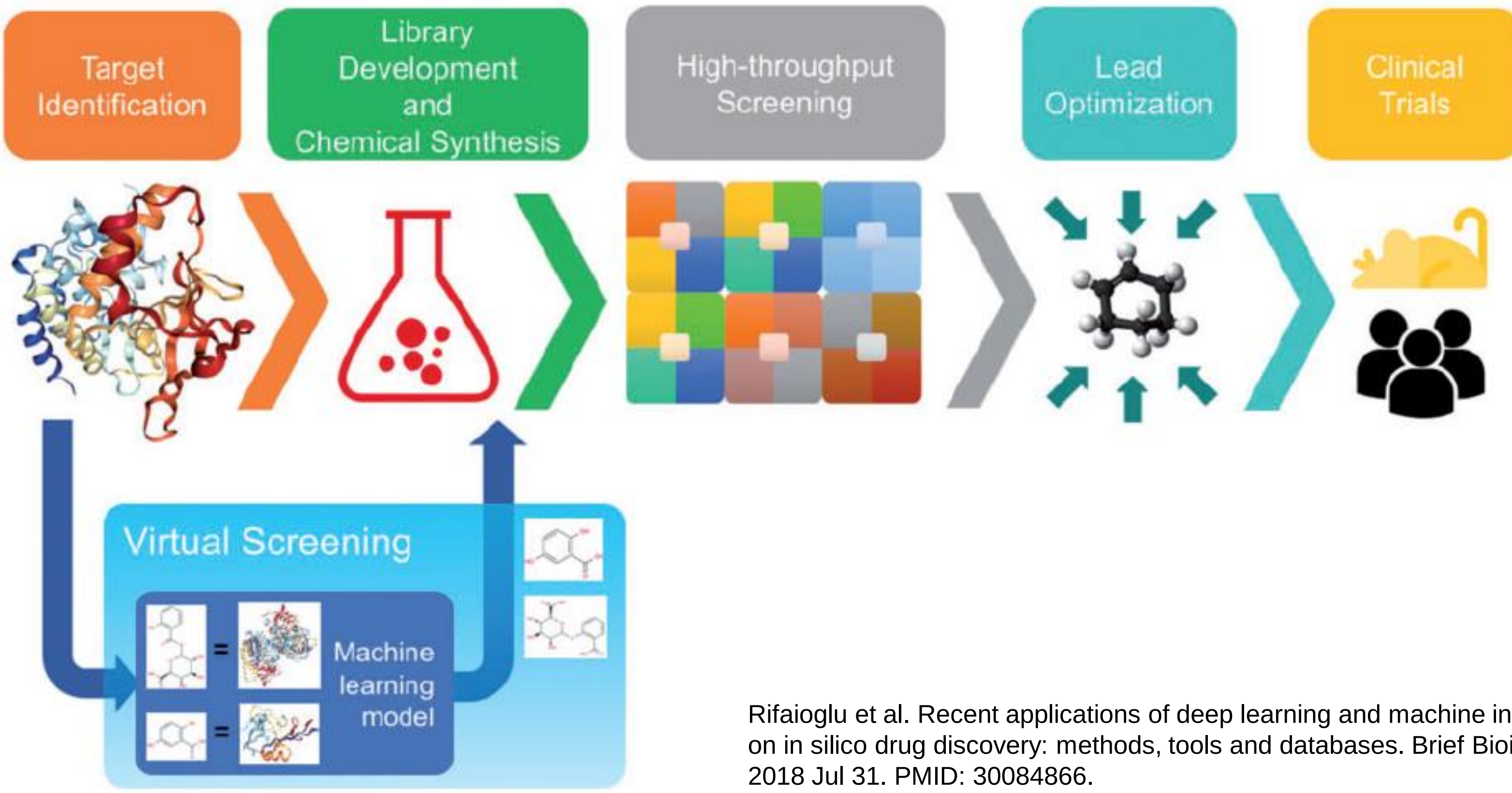
Drug development



Rifaioglu et al. Recent applications of deep learning and machine intelligence on in silico drug discovery: methods, tools and databases. Brief Bioinform. 2018 Jul 31. PMID: 30084866.

Loïc Verlingue

Drug discovery in details



Rifaioglu et al. Recent applications of deep learning and machine intelligence on in silico drug discovery: methods, tools and databases. Brief Bioinform. 2018 Jul 31. PMID: 30084866.

Loïc Verlingue

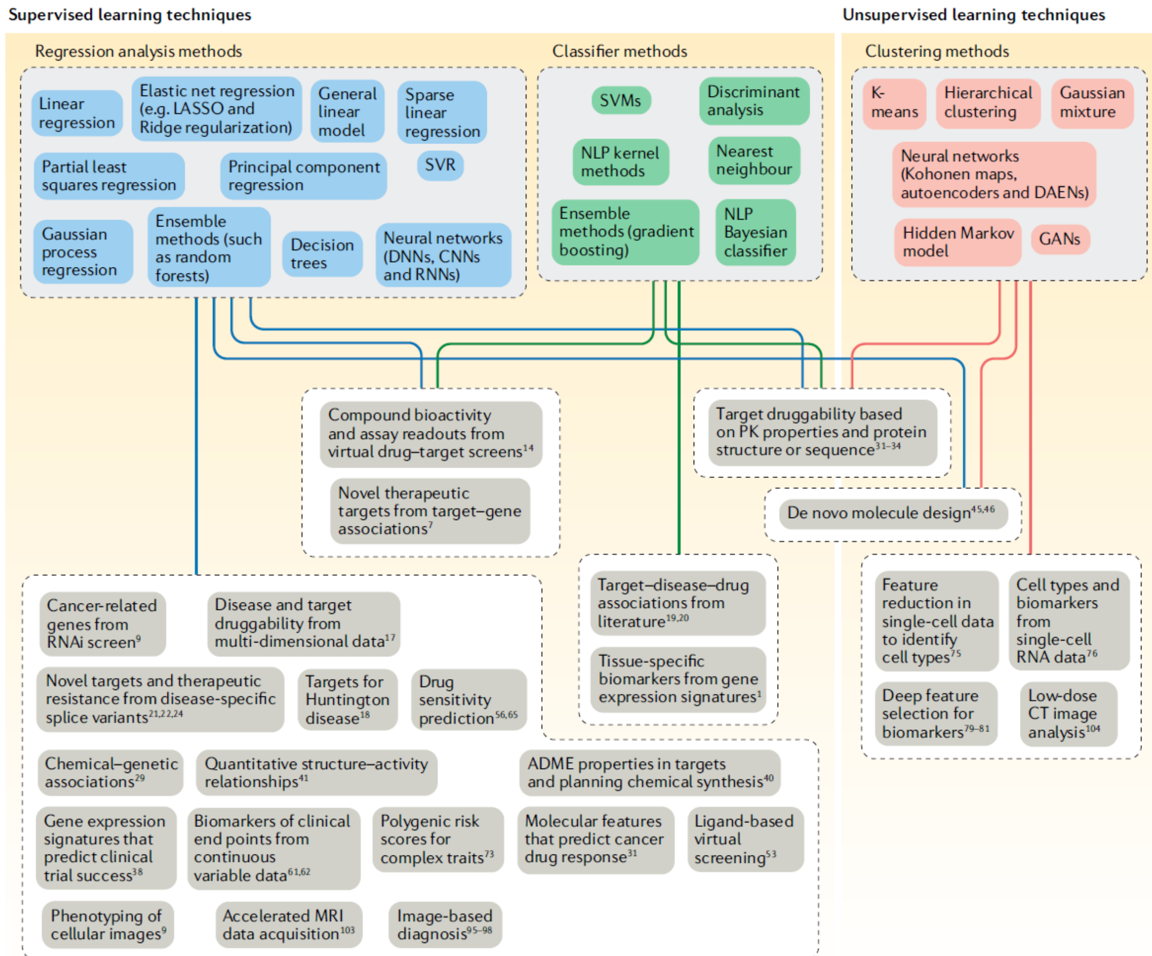
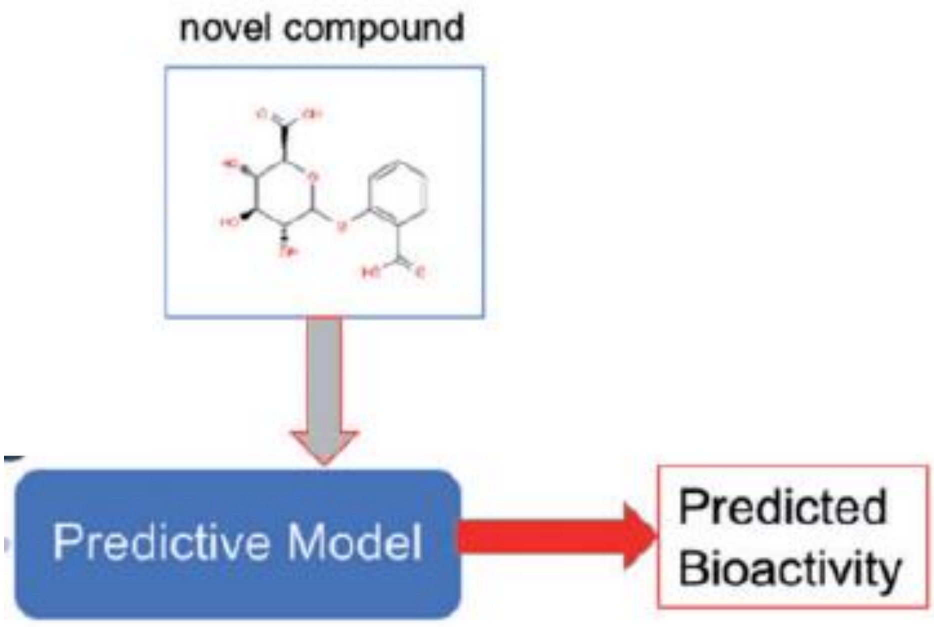


Fig. 2 | **Machine learning tools and their drug discovery applications.** This figure gives an overview of the machine

Vamathevan J,. Applications of machine learning in drug discovery and development. Nat Rev Drug Discov. 2019 Apr 11. doi: 10.1038/s41573-019-0024-5. [Epub ahead of print] Review. PubMed PMID: 30976107

Loïc Verlingue

Predictions



Rifaioglu et al. Recent applications of deep learning and machine intelligence on in silico drug discovery: methods, tools and databases. Brief Bioinform. 2018 Jul 31. PMID: 30084866.

Loïc Verlingue

Tasks with EHR

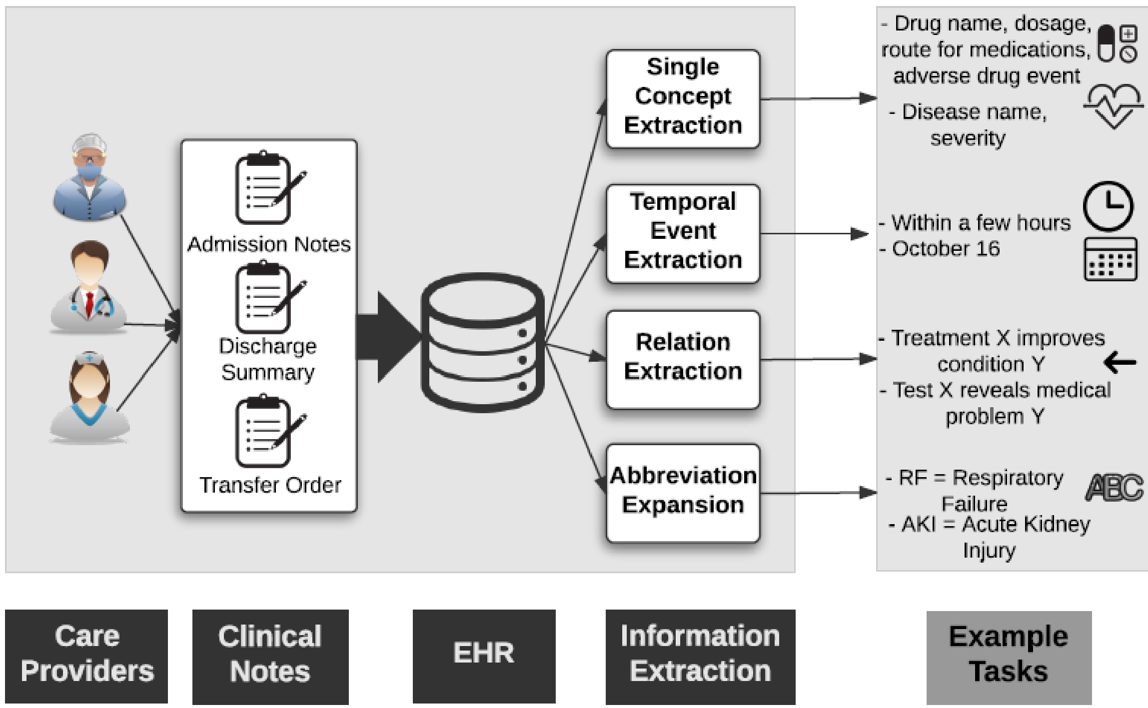


Fig. 7. EHR Information Extraction (IE) and example tasks.

Deep EHR: A Survey of Recent Advances in Deep Learning Techniques for Electronic Health Record (EHR) Analysis. Benjamin Shickel, Patrick J. Tighe, Azra Bihorac, and Parisa Rashidi. arXiv:1706.03446v2

Loïc Verlingue

Some comments

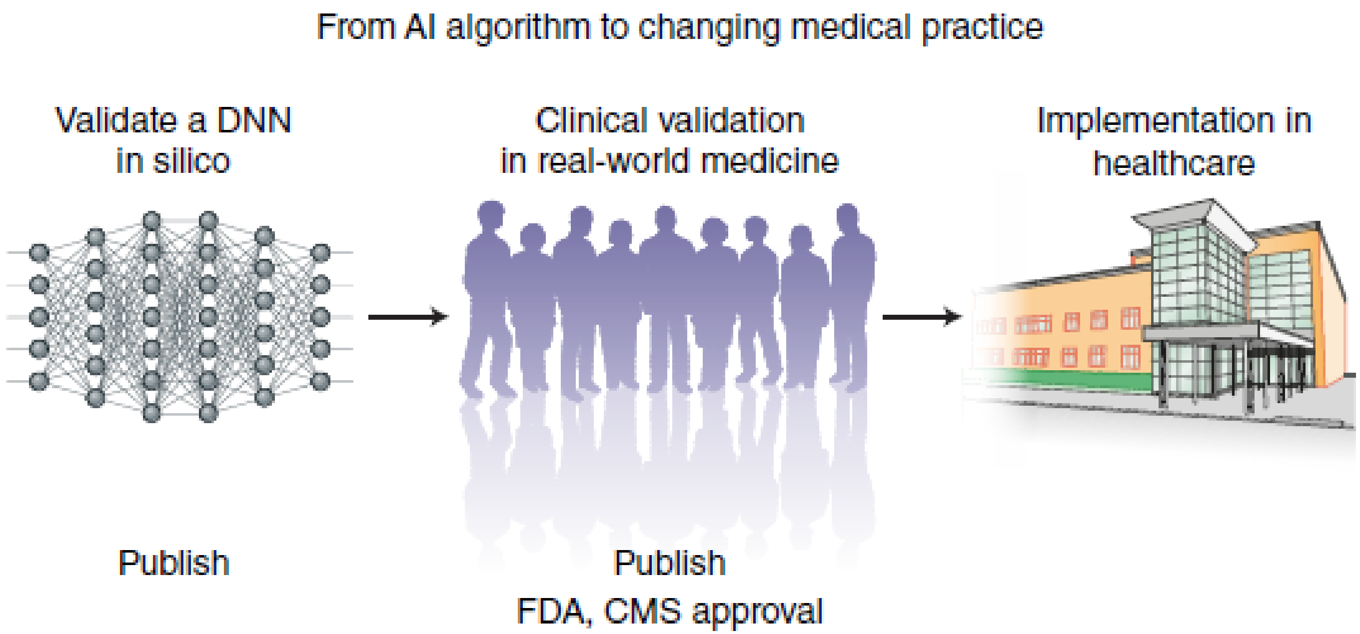
- AI can potentially be used in **every question in medicine that can be answered with data**: prediction of diagnosis, prognosis, response to treatment…
- As soon as there are **enough good data** to train an algorithm
- **Genomics and transcriptomics are well-suited to AI** (high number of features, « big data »), especially with recent techniques (**single-cell**)
- ***But may suffer from low number of samples (« fat data »)***

Conclusions and perspectives

- AI is **essential** to help the human brain in dealing with « big data »
- It has the potential to change practice in almost **all areas of medicine**
- However **AI does not replace** the doctor
- Algorithms have to be validated and show benefit in **real life**
- **Interpretability** issue

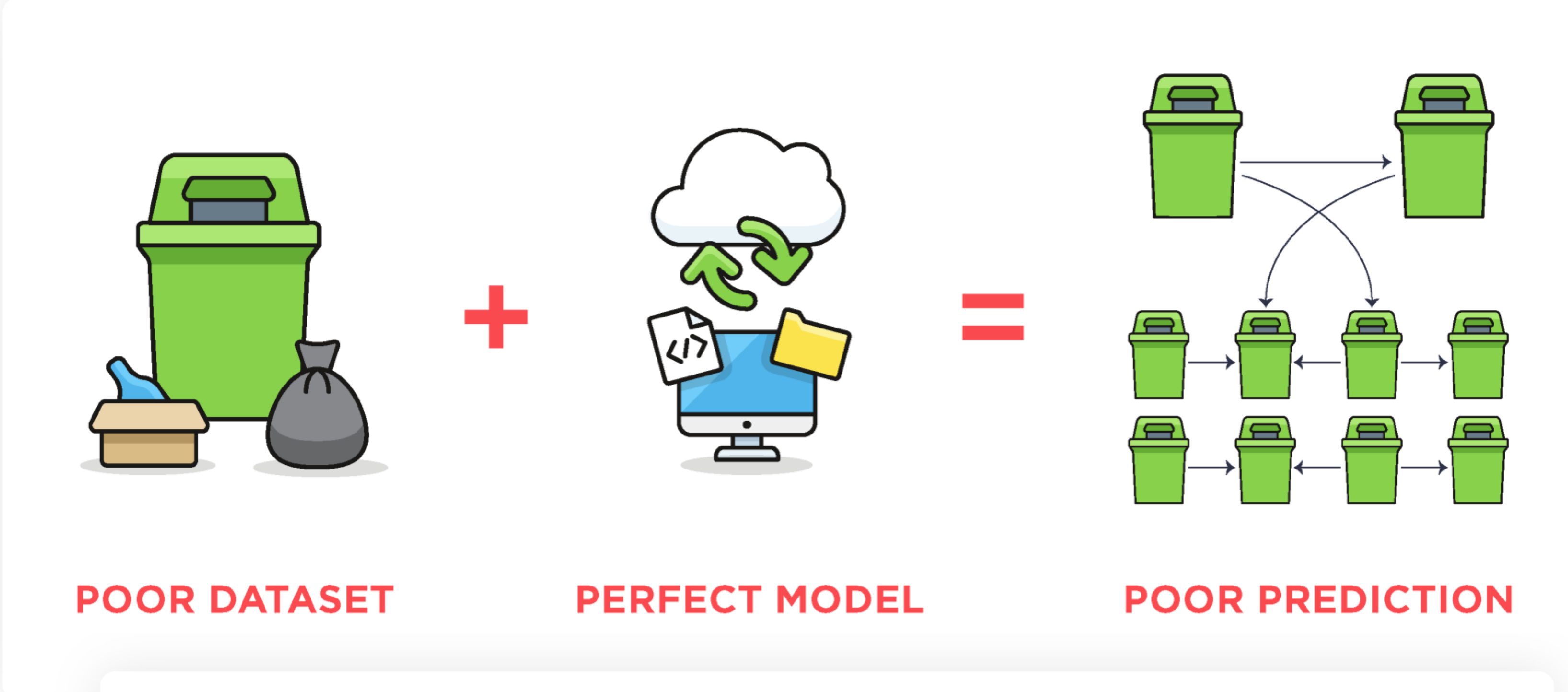
This subject is being updated every day…

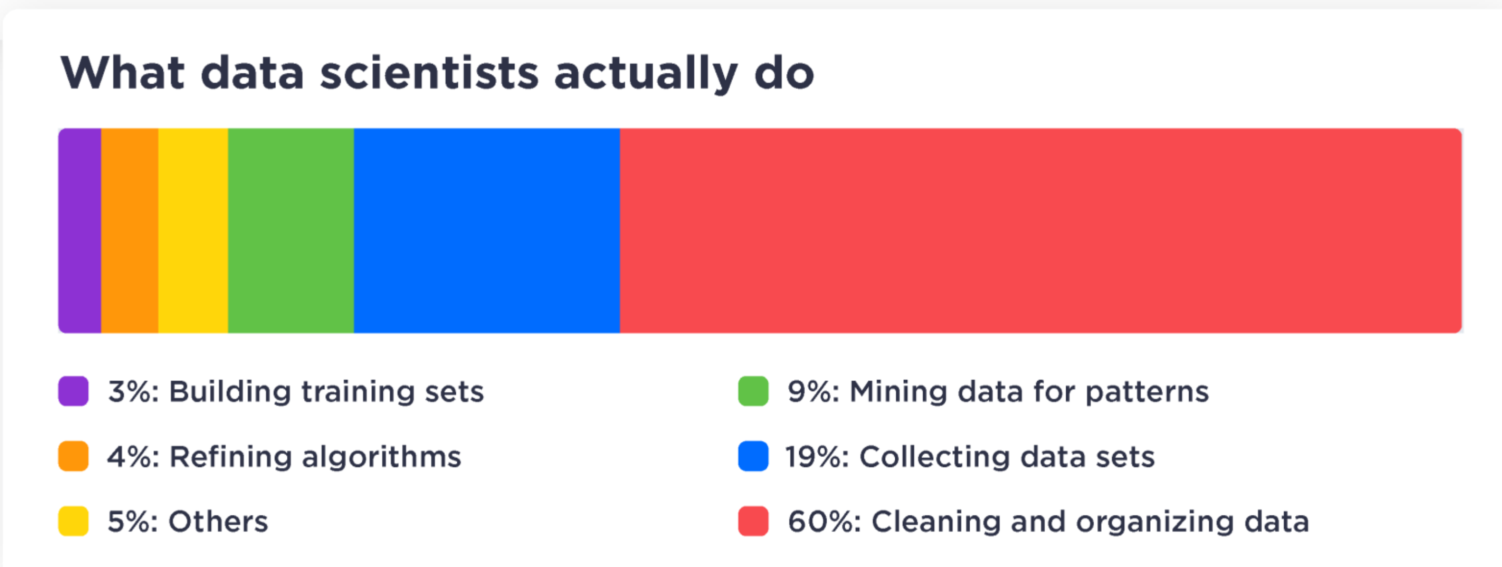
Be careful not to skip steps before use in real life



External validity!

Topol, Nat Med 2019





https://cldcvr.com/news-and-media/blog/clean-data-the-foundation-of-effective-machine-learning/

Acknowledgments

Dr Stéphane Champiat (DITEP)
Dr Loïc Verlingue (Léon Bérard)

THANK YOU FOR YOUR ATTENTION!