AI and emerging technologies in translational and clinical research, what are the potentials and challenges?

AFPT 2022

June 2022

Frederic Commo, PhD
Director & lead data science
Connected Tech, KPMG
fcommo@kpmg.fr

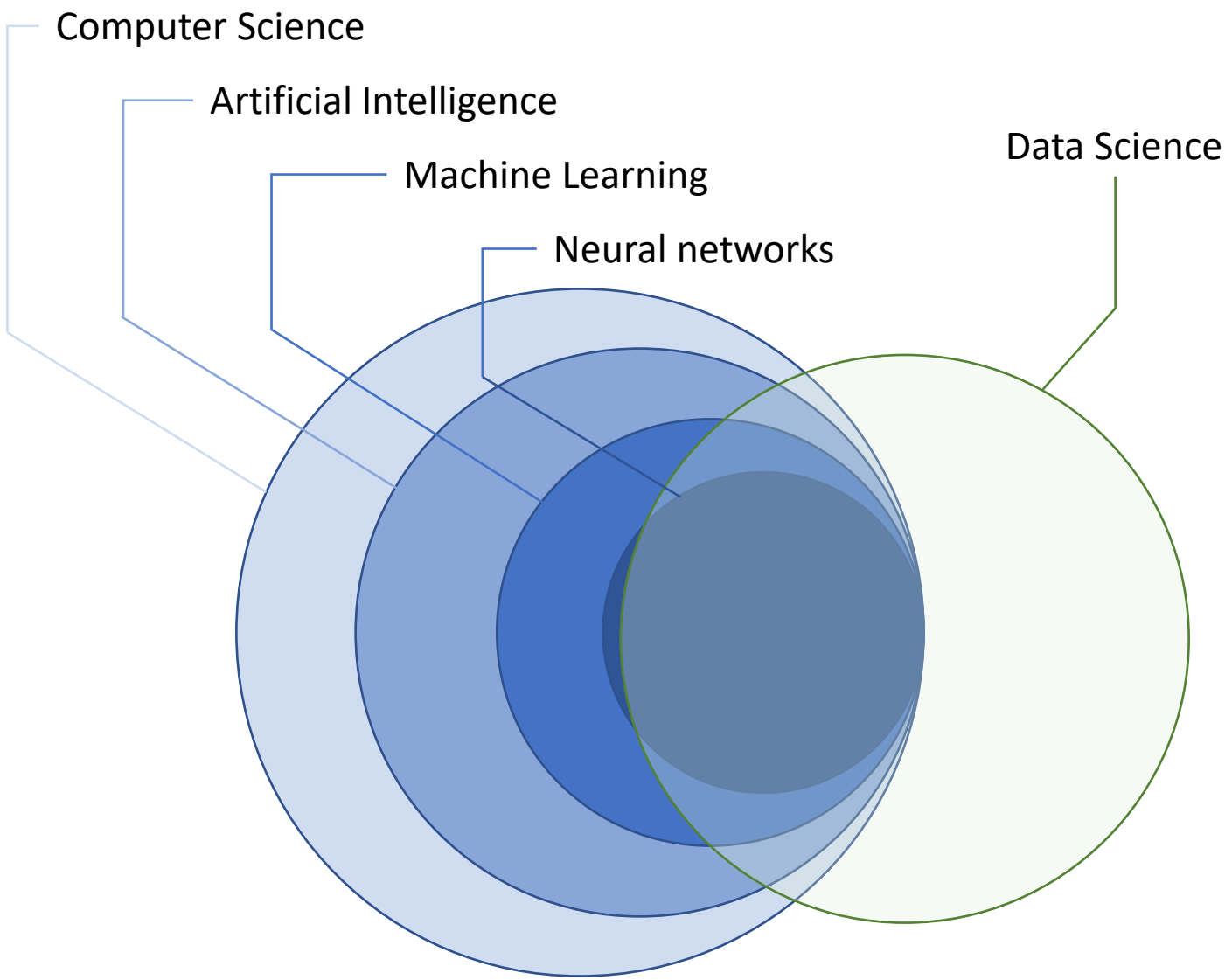
kpmg.fr

What is AI?

Artificial Intelligence:
Compute and perform complex tasks

Machine Learning:
Train an algorithm to identify patterns and reproduce decisions

Data Science:
Combine technical and field expertise



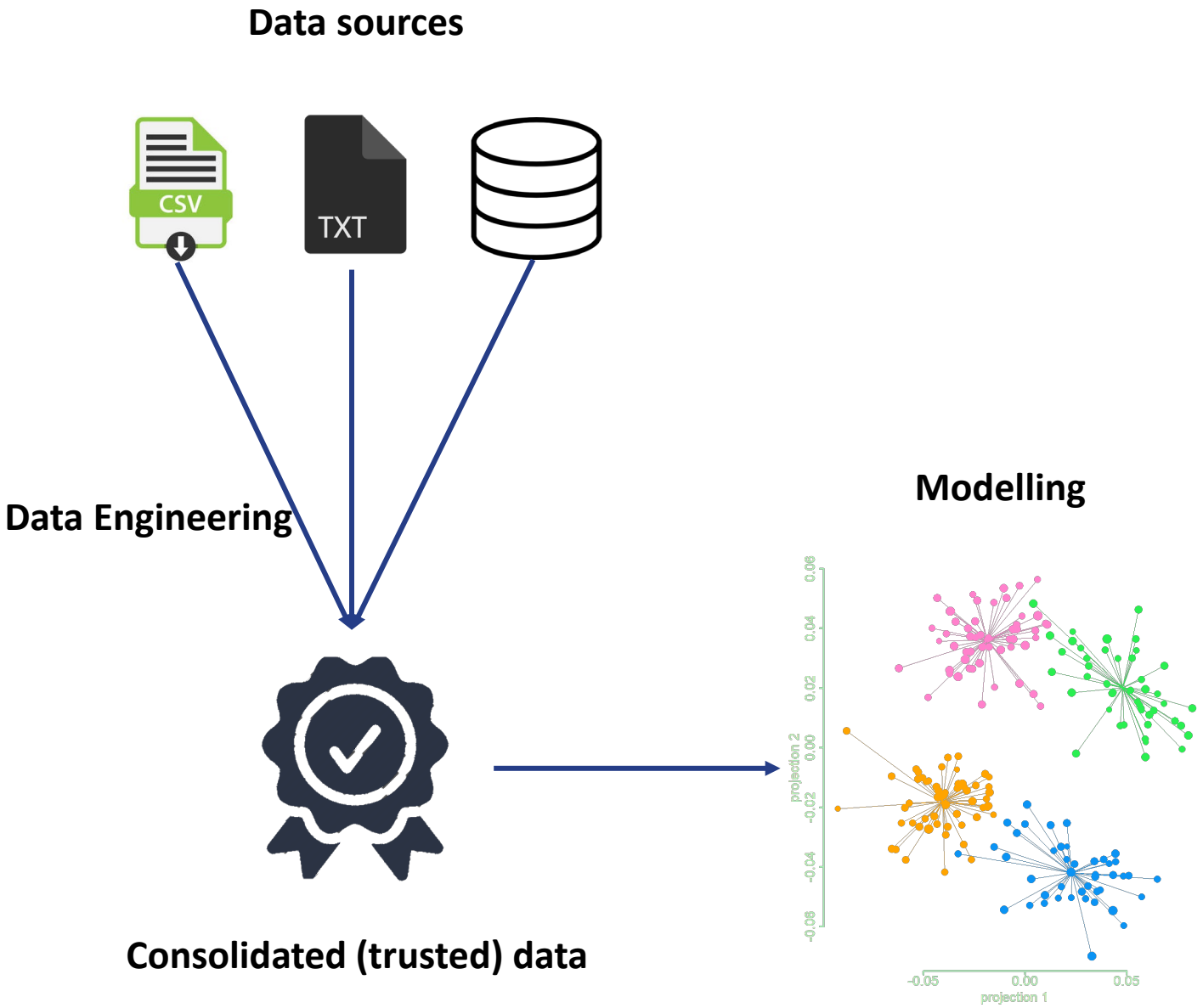
KPMG

What is Machine Learning?

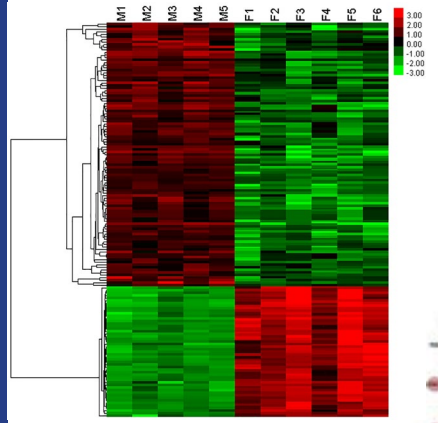
Artificial Intelligence:
Compute and perform complex tasks

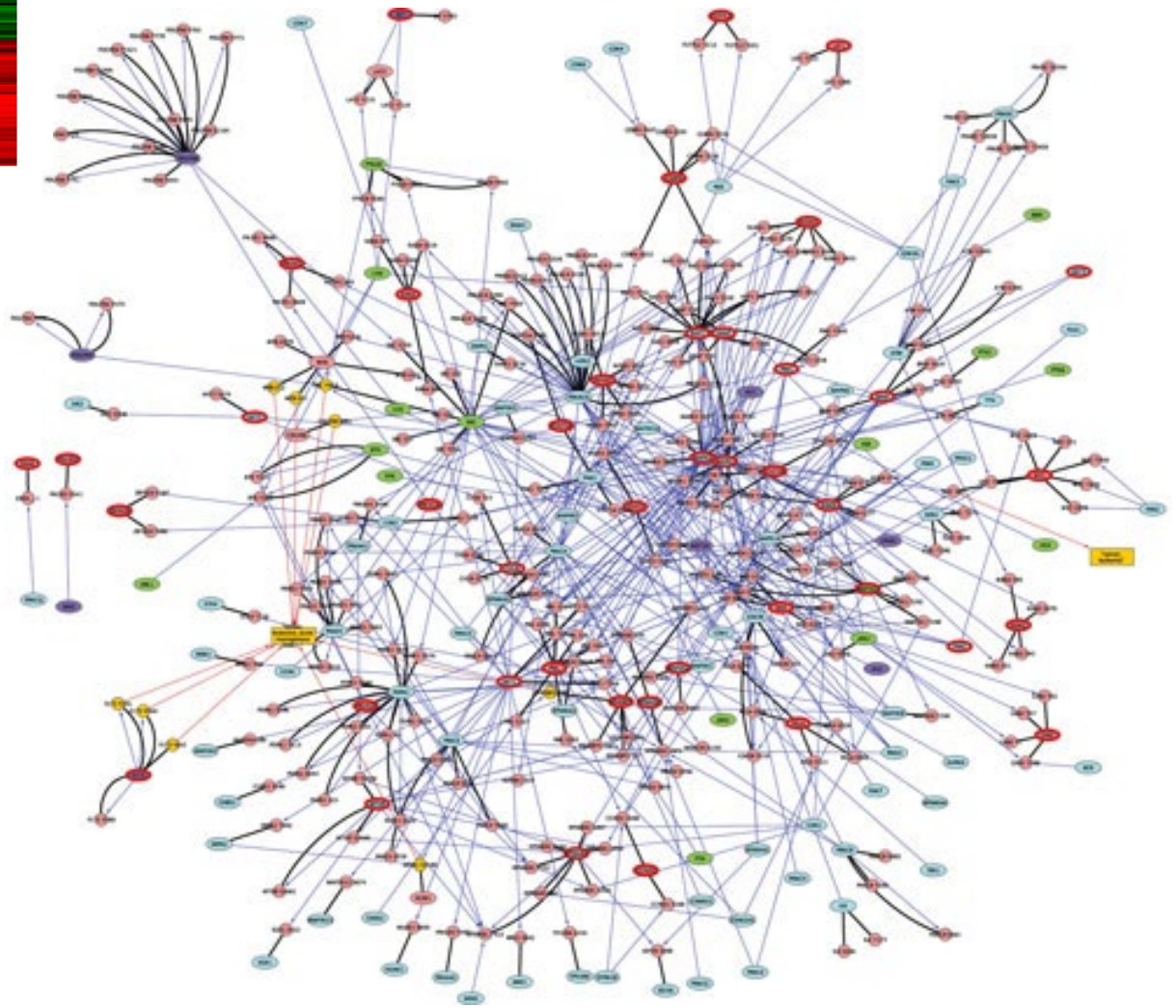
Machine Learning:
Train an algorithm to identify patterns and reproduce decisions

Data Science:
Combine technical and field expertise



What is Data Science?





Artificial Intelligence:
Compute and perform complex tasks

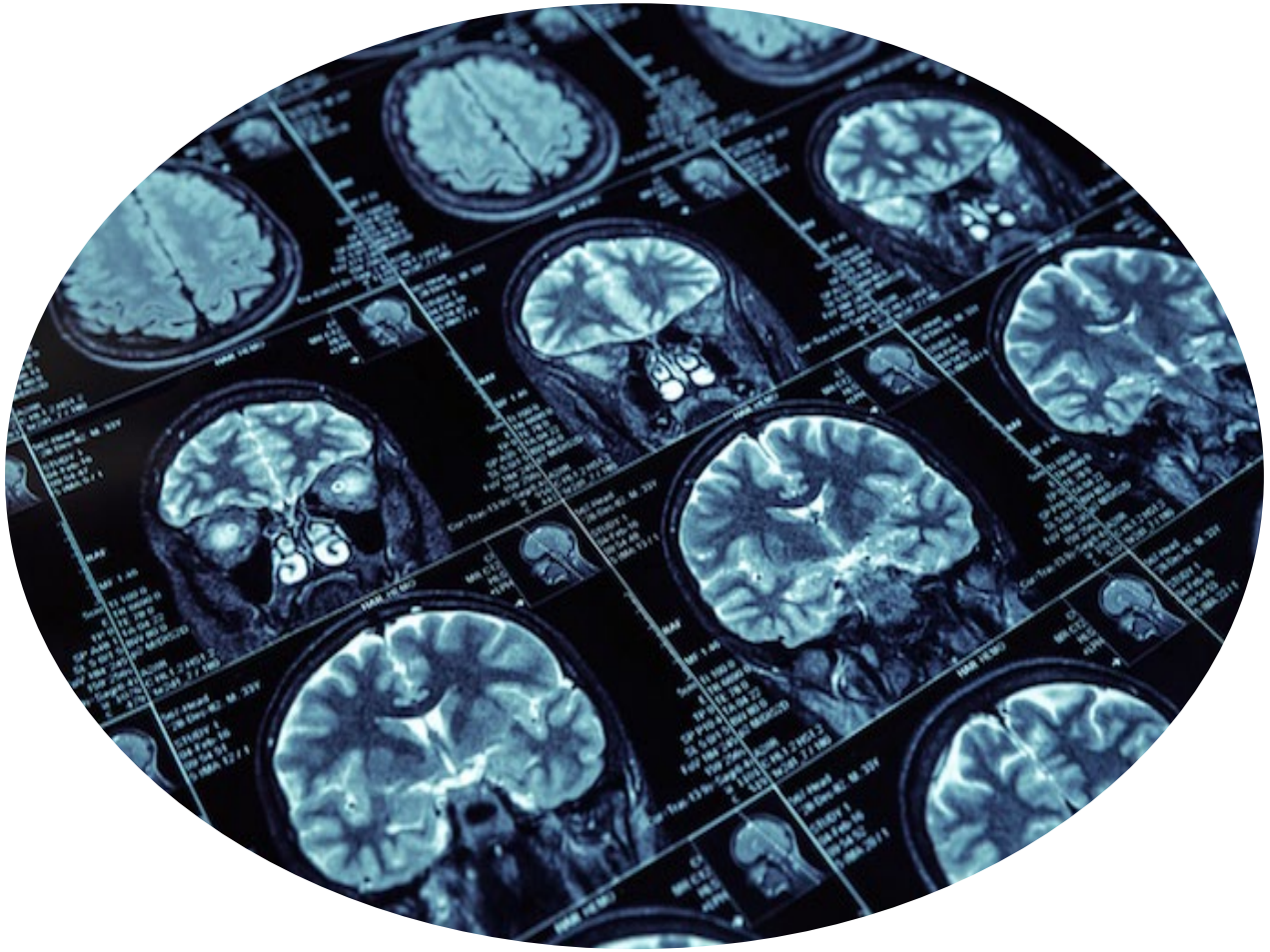
Machine Learning:
Train an algorithm to identify patterns and reproduce decisions

Data Science:
Combine technical and field expertise

KPMG

AI success

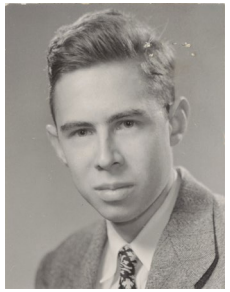
Automated image analysis



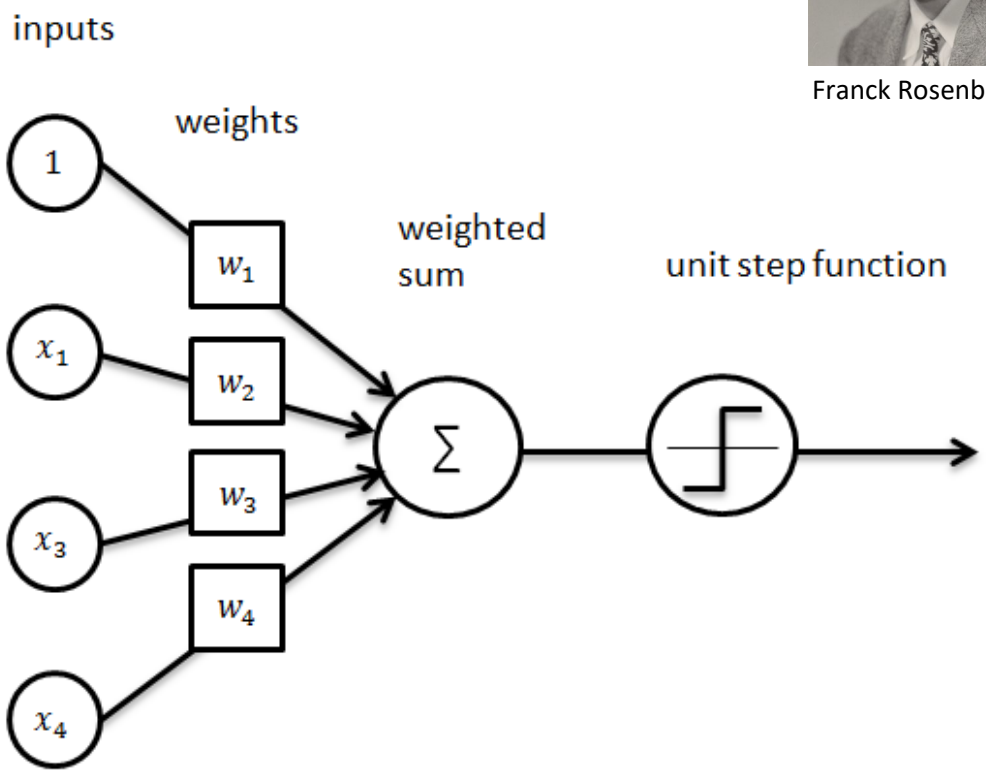
- Cardiovascular abnormalies
- Musculoskeletal injuries
- Neurological diseases, e.g. amyotrophic lateral sclerosis (ALS)
- Cancers, e.g. lung and breast

Neural Networks

The Rosenblatt's perceptron, 1958



Franck Rosenblatt

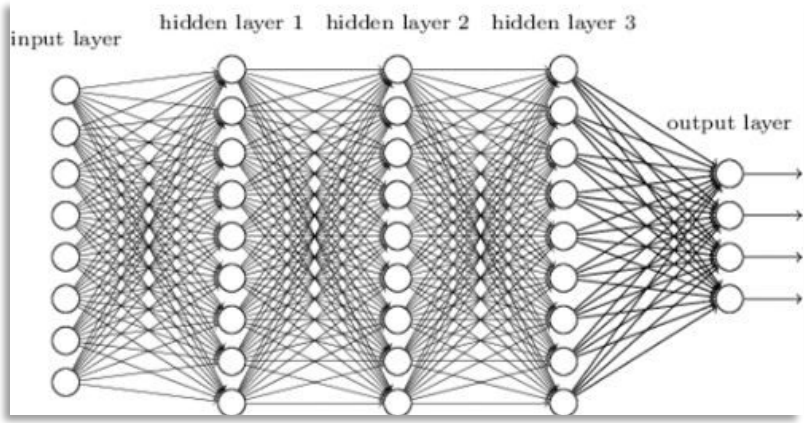


Derived from McCulloch & Pitts biological neuron concept, 1943

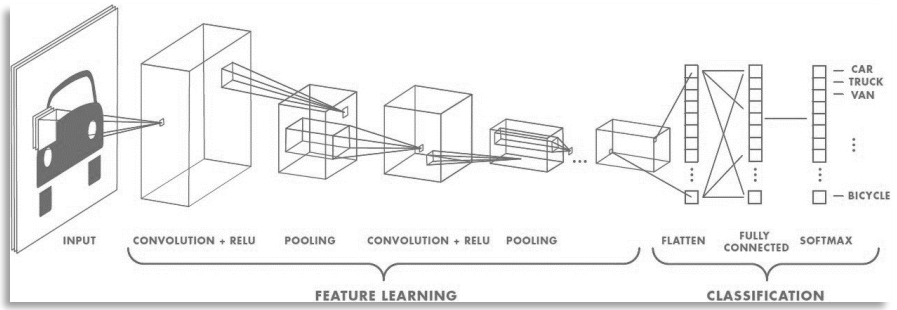
Neural Networks

More complex architectures

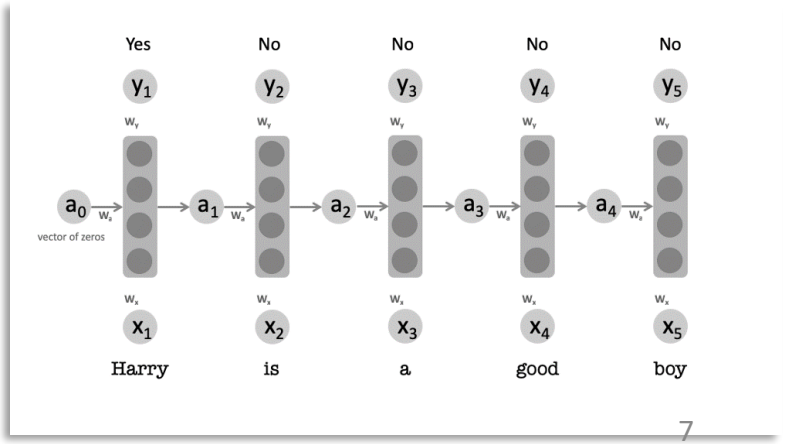
Deep (dense)
neural network



Convolutional
neural network (CNN)



Recurrent
neural network (RNN)



Still some problems to solve

Patient data is sometimes partial and unreliable

Risk of failures :

- **Wrong diagnoses**
- **Wrong proposed treatments**

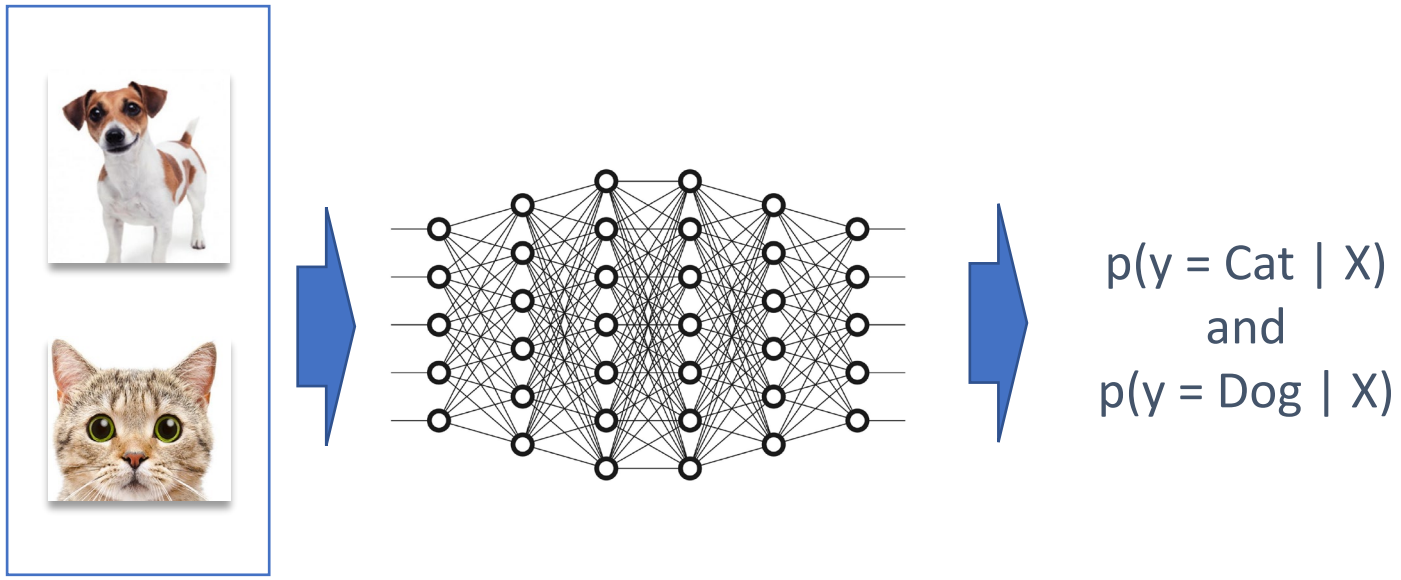
Why :

- **Amount of information was too low (incomplete)**
- **Data complexity and inconsistency**
- **Learning biases**

Back to basics

Almost all is about training

Cat/Dog image Classifier



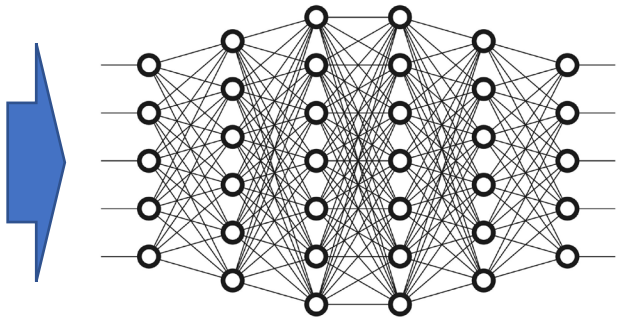
Back to basics

Train a model with valid data

(garbage in, garbage out)

Training Data

Cat/Dog image Classifier



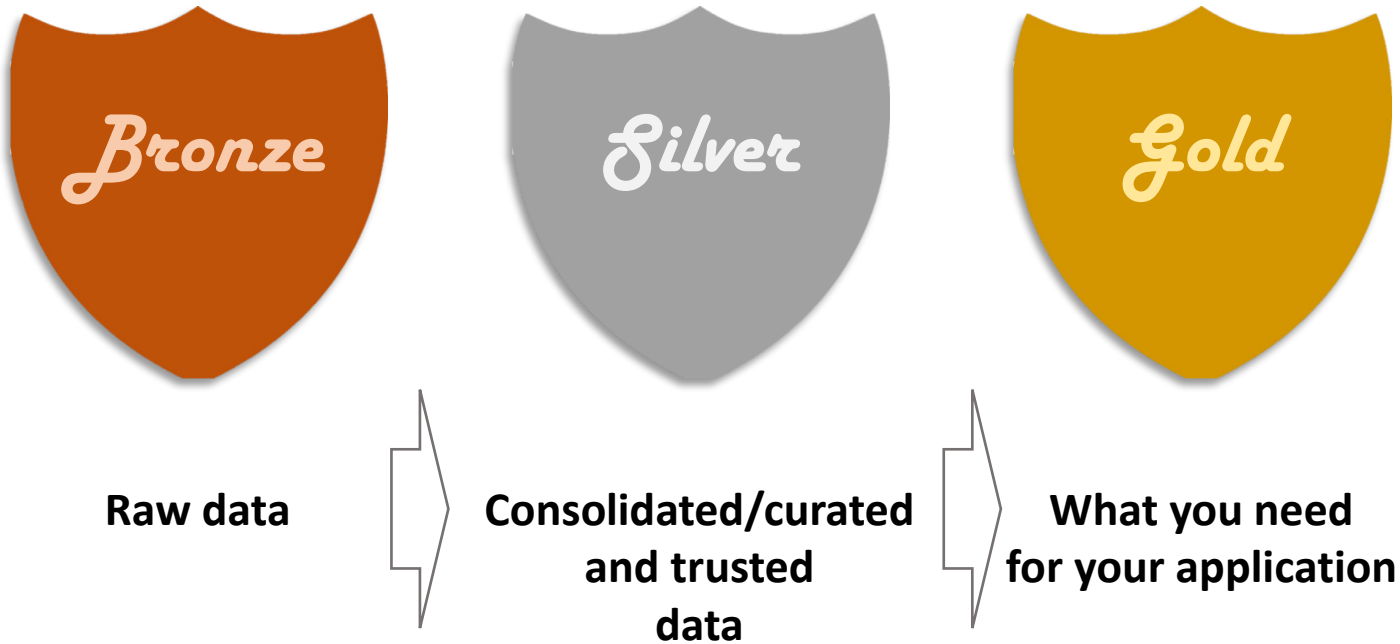
Algorithm

p(y = Cat | X)
and
p(y = Dog | X)

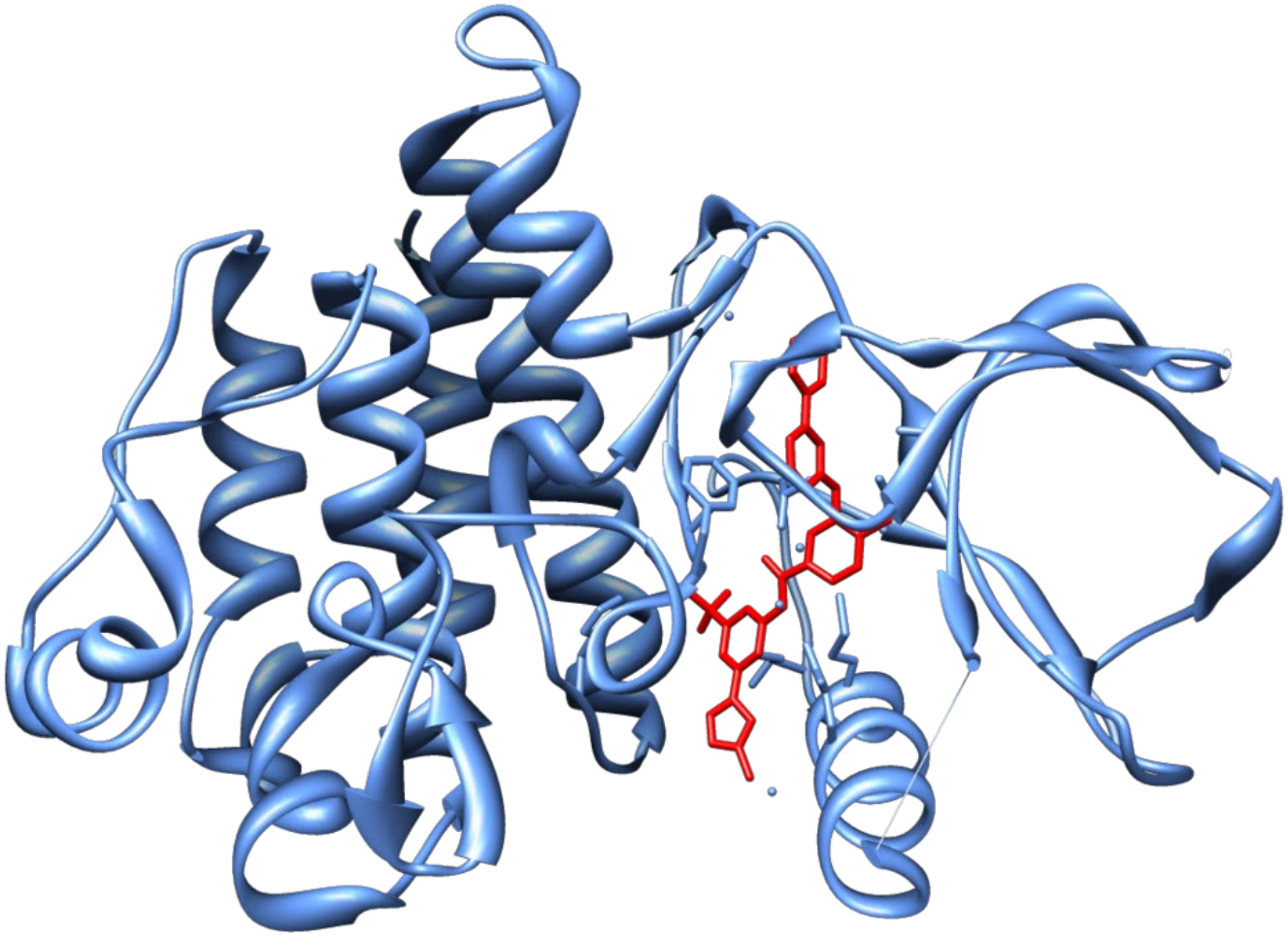
Output

Takeaway #1

Start with consistent and trusted data



Drug discovery



Crystal structure of the second generation Bcr-Abl tyrosine-kinase inhibitor nilotinib (red) in complex with an Abl kinase domain (blue). Nilotinib is used to treat chronic myelogenous leukemia (CML), a hematological malignancy.

Credit: SocratesJedi - Own work; Rendering of PDB 3CS9, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=17161180

Drug discovery & development

Time:	**10 to 15 years on average**
Cost:	**up to $2.8 billion**
Success:	**10 to 20%**

Source: Wouters *et al.*, JAMA, 2020

How AI can help

Objective:

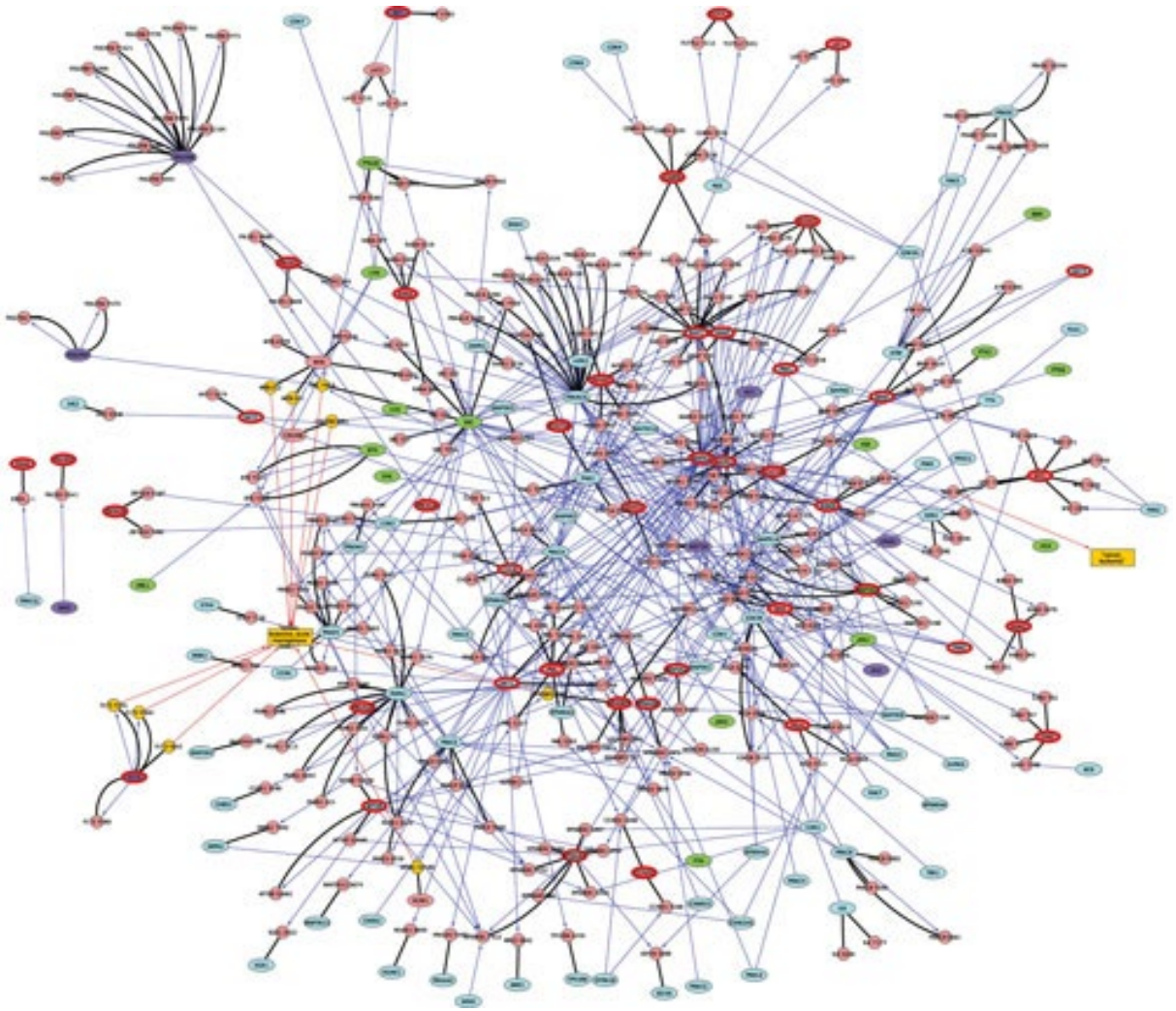
Significantly reduce the time and budget spent on developing and bringing new drugs to market

Challenges :

- Identify new targets
- Design new drugs
- Predict their efficacy and side effects
- Identify new applications for existing drugs
- Identify the right models for preclinical trials
- Identify the right cohorts for clinical trials

Identifying new targets

Computational biology



TargetDB

Consolidated sources of information combined with machine learning

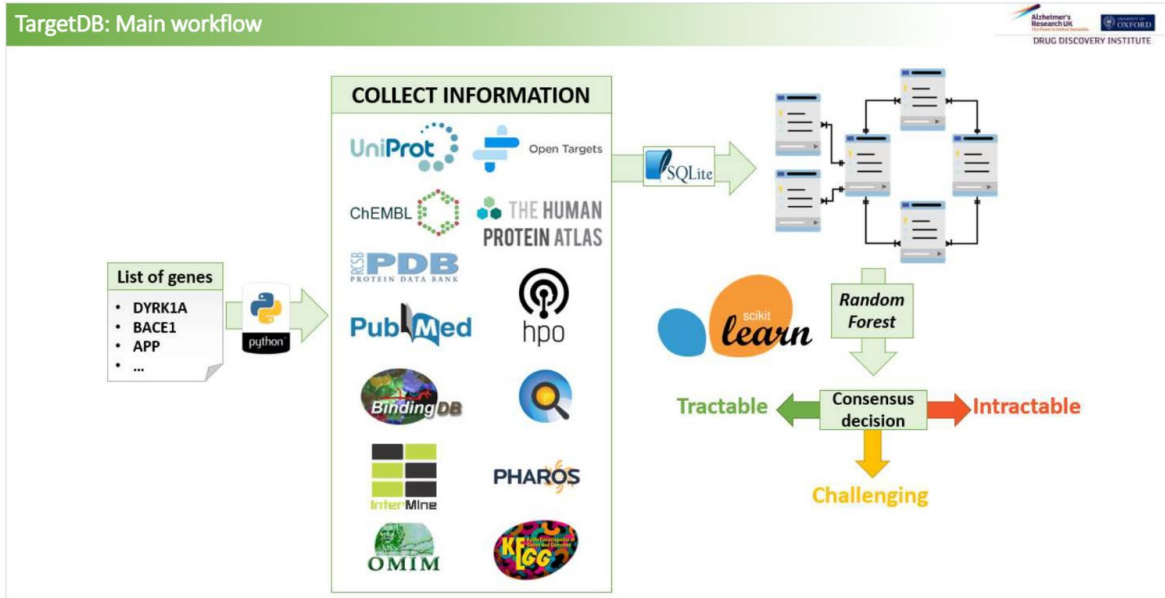
RESEARCH ARTICLE

TargetDB: A target information aggregation tool and tractability predictor

Stephane De Cesco*, John B. Davis, Paul E. Brennan*

Nuffield Department of Medicine, ARUK Oxford Drug Discovery Institute, Target Discovery Institute, University of Oxford, Oxford, United-Kingdom

PLOS ONE | https://doi.org/10.1371/journal.pone.0232644 September 2, 2020



AlfaFold (DeepMind)

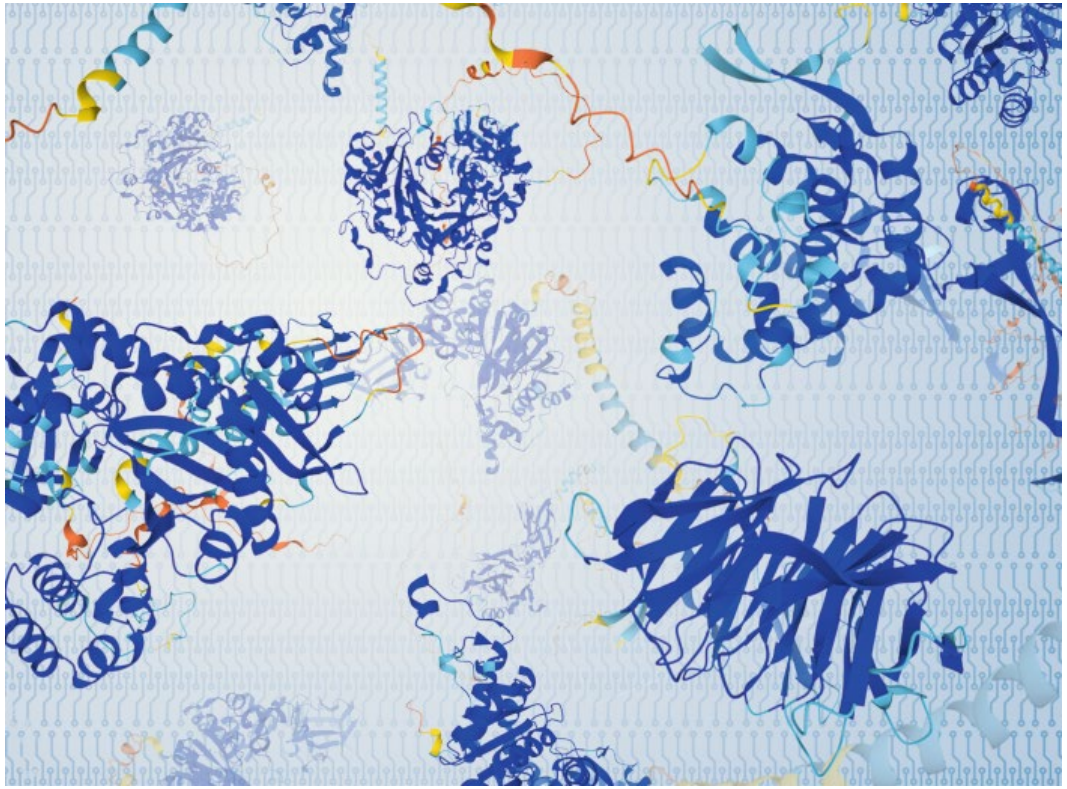
3-D Protein structure prediction

Article | Open Access | Published: 15 July 2021

Highly accurate protein structure prediction with AlphaFold

John Jumper, Richard Evans, ... Demis Hassabis + Show authors

Nature **596**, 583–589 (2021) | Cite this article

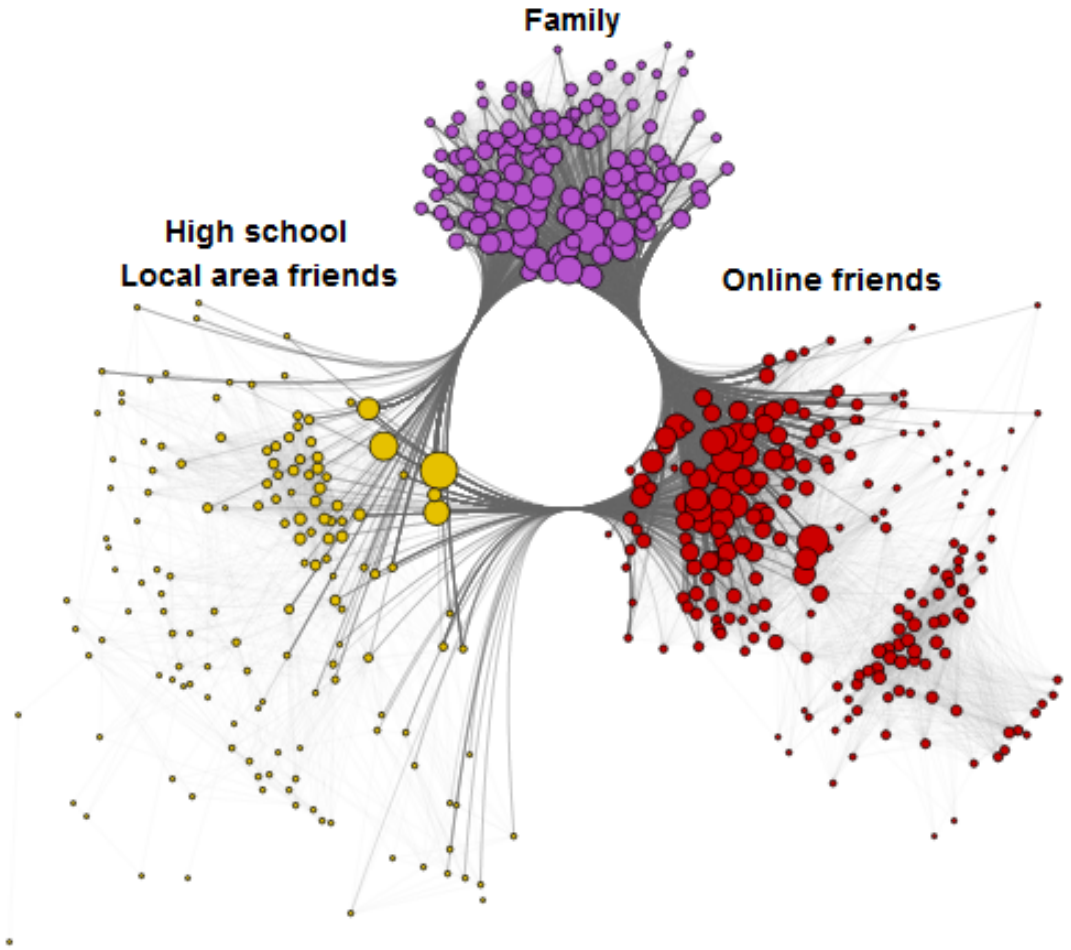


Current: ~1M proteins

2022 target: 100M proteins

Predict drug efficacy and side-effects

We need to model interactions



Credit: https://distill.pub/2021/gnn-intro/

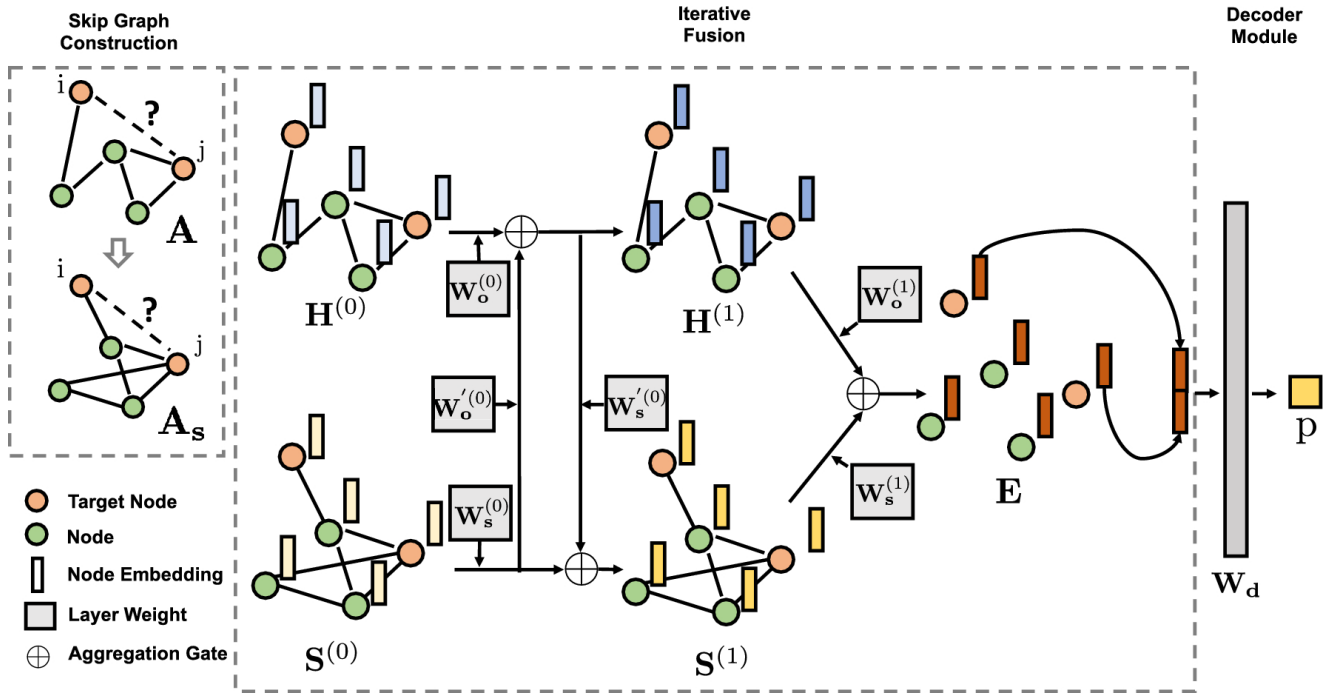
Predict molecular interactions

Graph Neural Networks (GNN)

Article | Open Access | Published: 03 December 2020

SkipGNN: predicting molecular interactions with skip-graph networks

Kexin Huang, Cao Xiao, Lucas M. Glass, Marinka Zitnik & Jimeng Sun



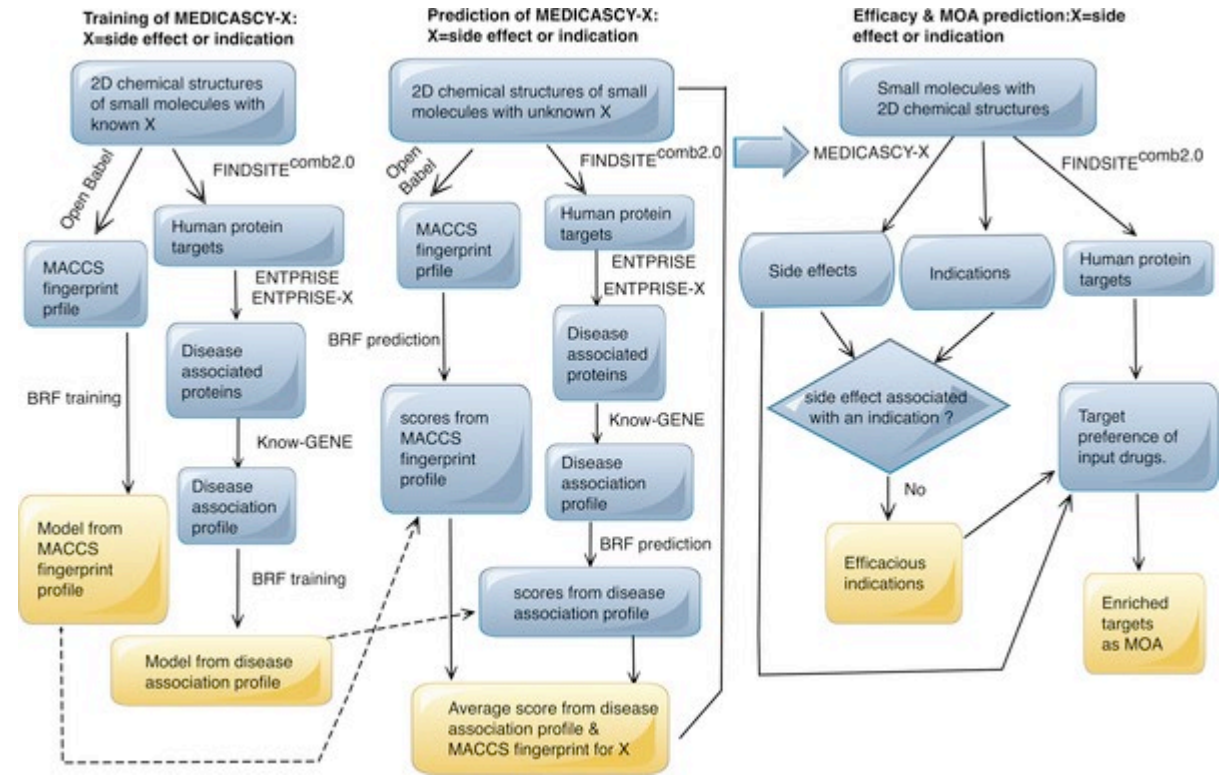
Predict drug efficacy and side-effects

Published in final edited form as:
Mol Pharm. 2020 May 04; 17(5): 1558–1574. doi:10.1021/acs.molpharmaceut.9b01248.

MEDICASCY: A Machine Learning Approach for Predicting Small Molecule Drug Side Effects, Indications, Efficacy and Mode of Action

Hongyi Zhou[1], **Hongnan Cao**[1], **Lilya Matyunina**[2], **Madelyn Shelby**[2], **Lauren Cassels**[2], **John F. McDonald**[2], **Jeffrey Skolnick**[1,*]

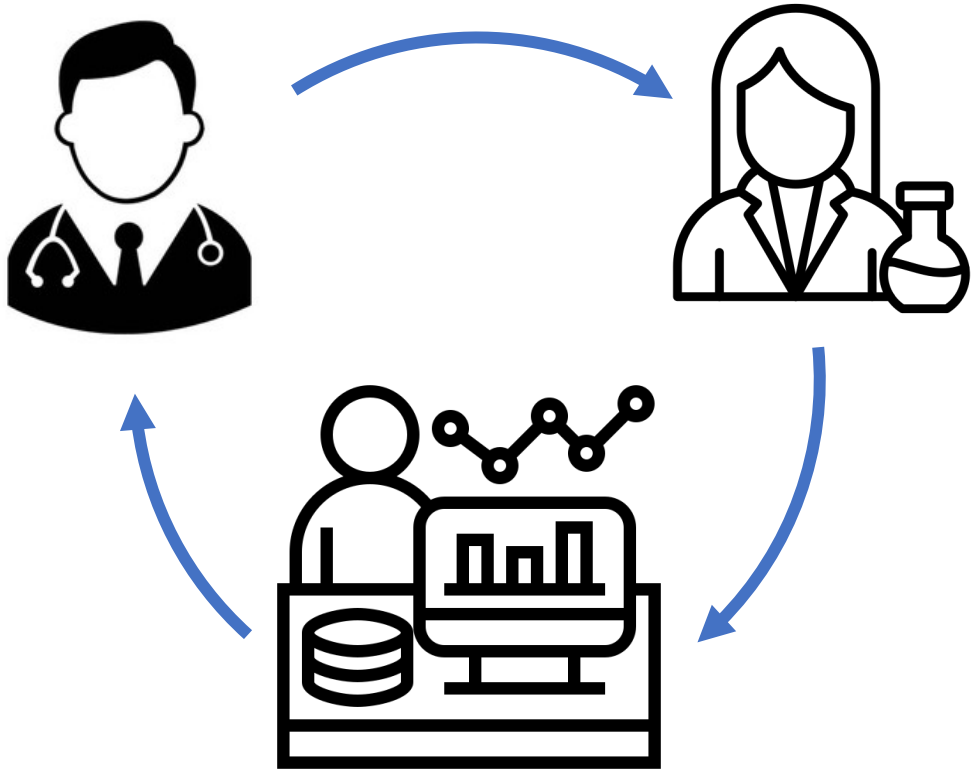
Boosted random forest



Takeaway #2

AI projects are not just about (Math) Nnets

Do not neglect interdisciplinary collaborations



The future of computers

Quantum computing

0/1

3D render of a quantum computer. Credit: Bartlomiej K. Wroblewski/ Shutterstock Image

Identify the right cohorts for clinical trials

The curse of medical databases :

- Still a lot of paper out there
- Missing information or errors
- Lack of standardized information
  - Biological data
  - Medical Imaging
  - Genomics
  - Pathology reference systems
  - Treatment history
  - …

Medical databases initiatives

DRIM France IA
Intelligence artificielle radiologie

CONSORE
continuum soins recherche
unicancer

Consore
BIG DATA
21 OCTOBRE 2021

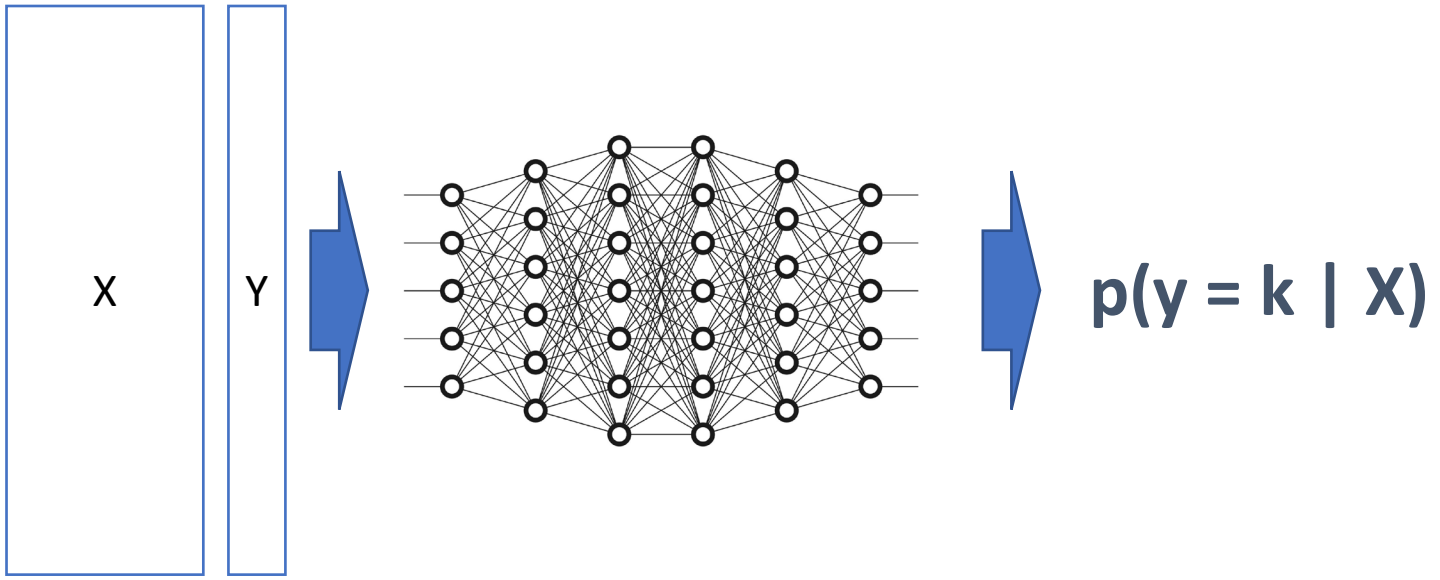
Onco Data Hub | **OncoBiome** | **CANTO**

KPMG

Takeaway #3

Always keep human in the loop

Classifier



Training Data

Algorithm

The output is a probability, not a decision

Thank you

Frederic Commo, PhD
Director & lead data science
Connected Tech, KPMG
fcommo@kpmg.fr