

L'INTELLIGENCE ARTIFICIELLE CHALLENGE POUR LES LABORATOIRES DANS LES PROCHAINES ANNEES ?

L'I.A : Matrice du Laboratoire du Futur ...



Yvon Gervaise

E.S.C. (ExpertScience – Consulting)

animateur GT SECF #LaboratoireduFutur

<http://www.expertscience.fr>

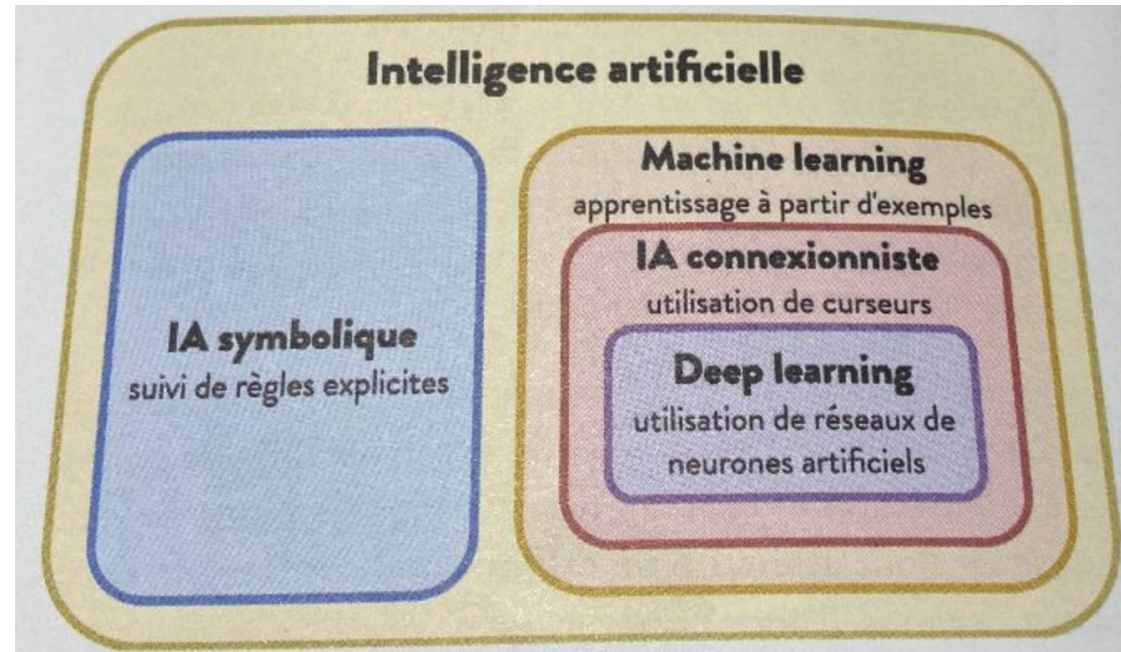
L'Intelligence Artificielle (I.A) ?

L'intelligence artificielle (I.A.)

« c'est l'ensemble des techniques permettant à des machines d'accomplir des tâches et de résoudre des problèmes normalement réservés aux humains et à certains animaux »

L'I.A comprend 3 catégories :

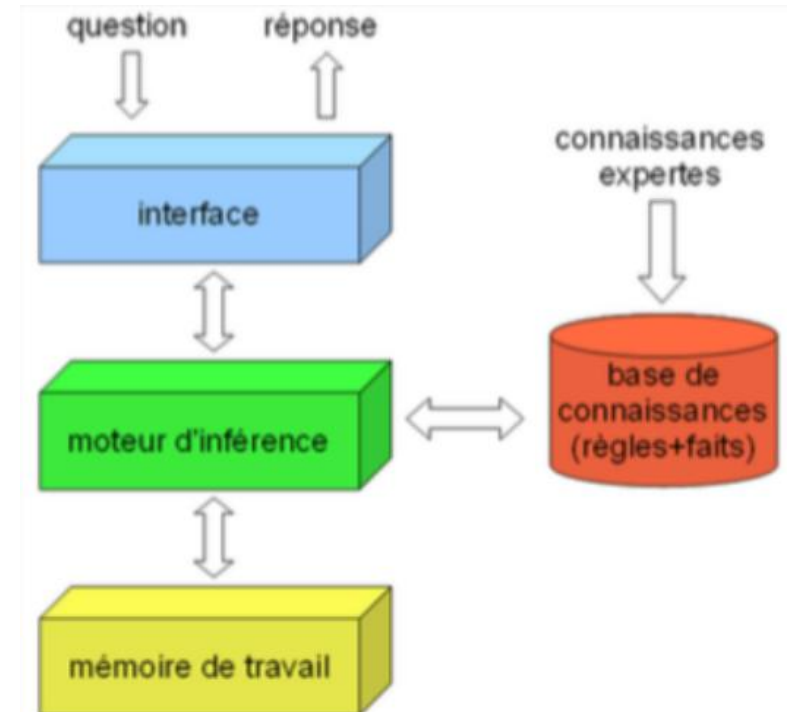
- **I.A. Symbolique** : les algorithmes à base de règles (**Systemes Experts**)
- **I.A. Connexionniste** : les les algorithmes d'apprentissage avec **réseau neurones**
- **I.A. Générative (GPT)** réseaux de neurones avec variante **Transformer**



Les systèmes Expert (S.E) :

- Encapsulent les **connaissances des experts** sous forme de règles (si X alors Y)
- Disposent d'un mécanisme d'**inférence**
- Utilisent ces connaissances pour résoudre des problèmes et **automatiser le traitement de processus**
- Ont évolué suivant différents **modèles de représentation** sémantique , graphe conceptuel , arbre de décisions , langages informatique

(Exemples de S.E : prospector Mycin , Dendral)



Le Système Expert Meta-Dendral énigmatique pour les chimistes ...depuis longtemps dans les LABORATOIRES !

Permet de identifier des constituants chimiques d'un matériau à partir de son spectre de masse et de RMN.

Utilise la règle : « si le spectre de la molécule présente 2 pics X1 et X2 tel que $X1 - X2 = M + 28$...alors la Molécule contient un groupe cétone ».

Ce S.E est le premier algorithme qui **modélise la connaissance** d'un expert chimiste analyste.

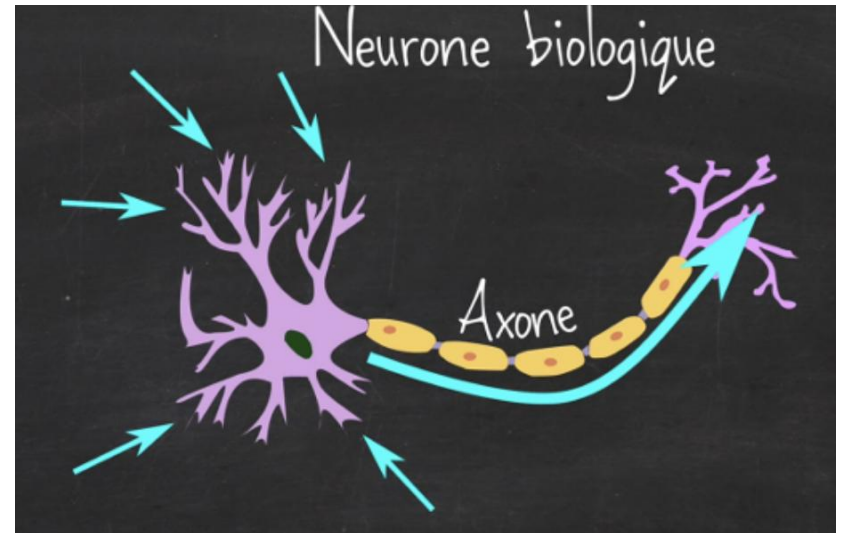
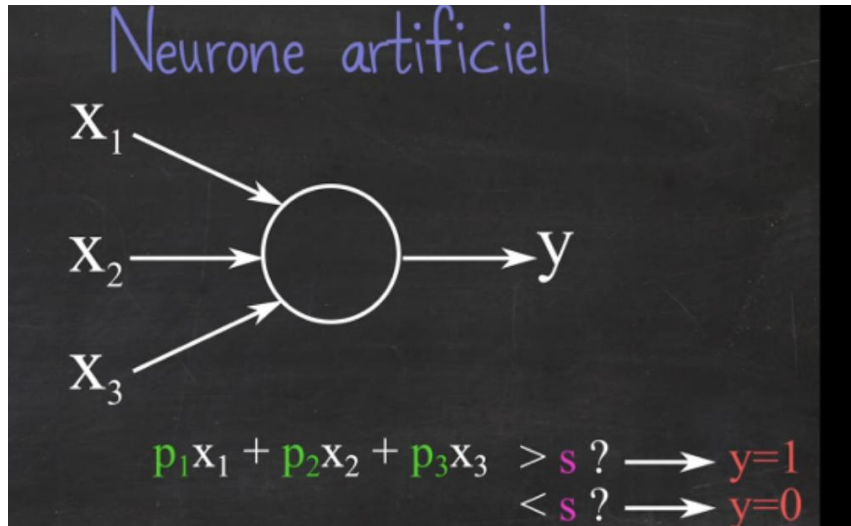


 ResearchGate

1. A Typical Meta-DENDRAL Rule ...



- **Neurone artificiel est une fonction mathématique qui mime le neurone biologique** avec son hardware

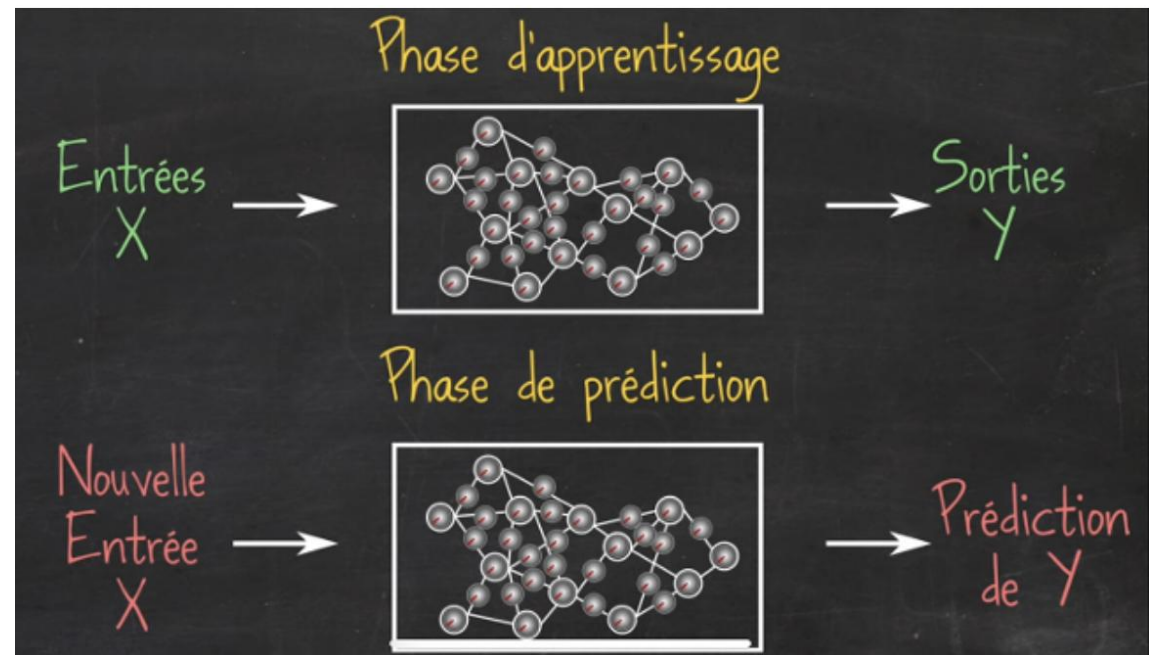
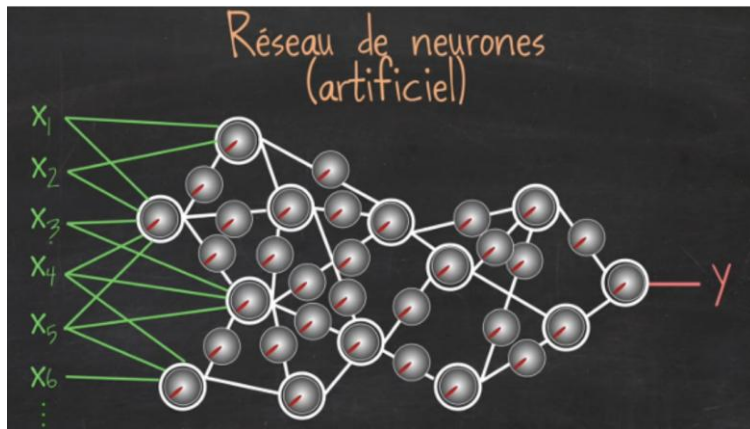


- ✓ L'I.A connexionniste imite l'architecture de base du cerveau
- ✓ Les réseaux neurones artificiels peuvent recevoir et transmettre de l'information



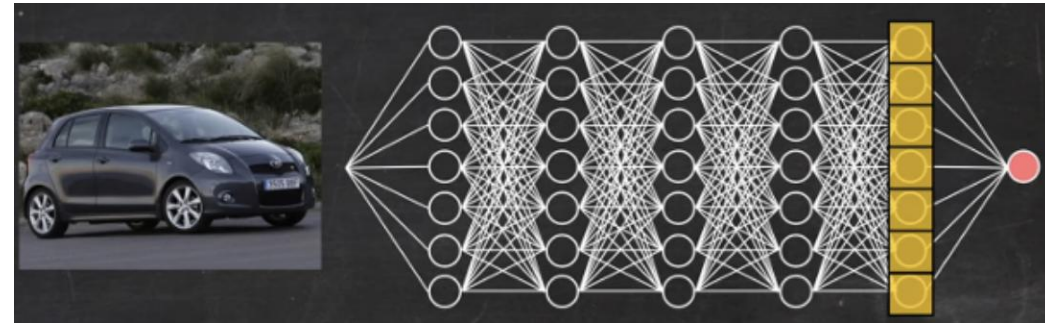
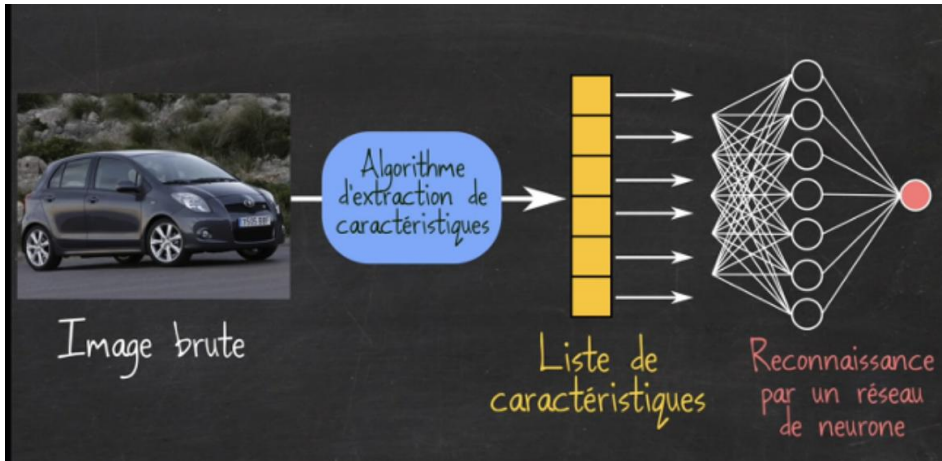
L'I.A connexionniste la machine apprend puis résout

- **Phase d'apprentissage** nourrit l'algorithme de multiples exemples caractéristiques
- **Phase de prédiction**
- **Phase de test** de la qualité réponse/ prédiction



Du machine learning au Deep learning grâce au réseau neurones profonds multicouches

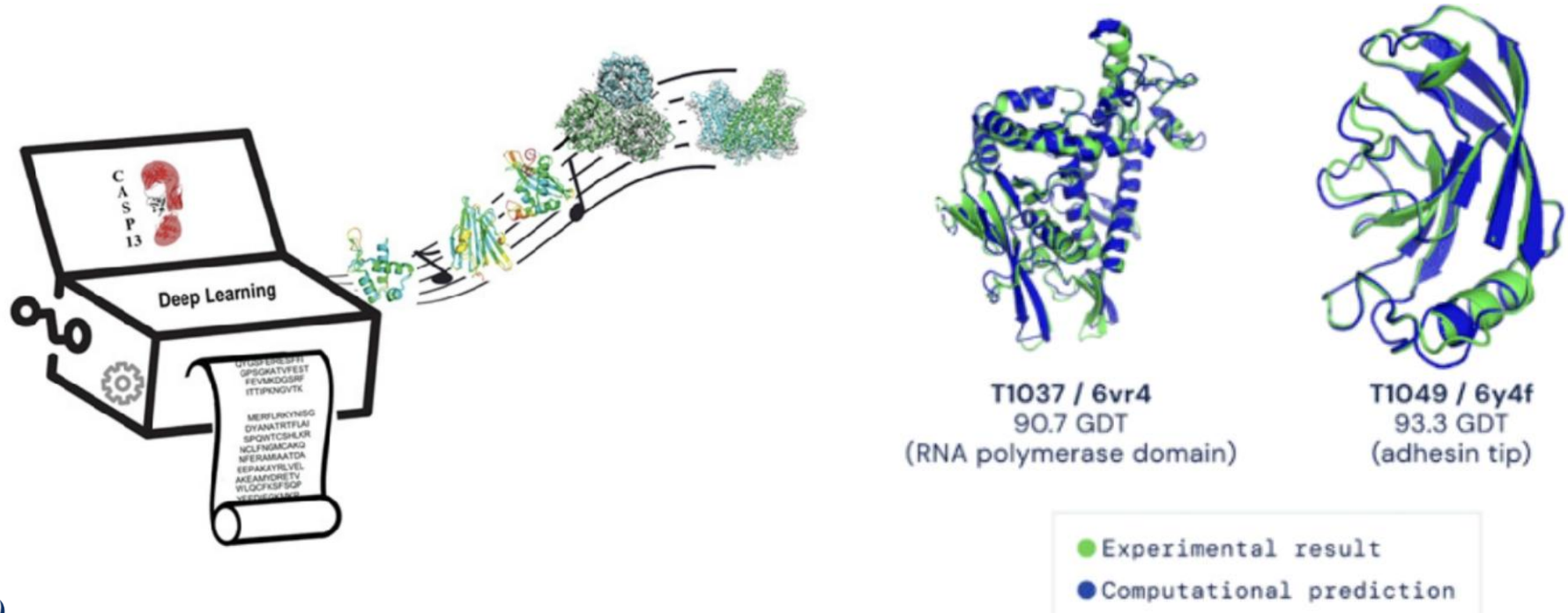
La machine découvre elle même les caractéristiques et est ensuite apte à capturer la relation entre entrée et sortie.



- Aptitudes à la résolution problèmes complexes « Deep learning » Yann Le Cun
- Exigences de :
 - Capacités de calcul (processeurs graphiques)
 - Disponibilité de grandes quantités de données pour entraînement de l'algorithme
- Optimisation réseau neurone (poids synaptiques , seuils d'activation ... rétropropagation de gradient)



- **Le problème :** Impossible de lire dans une séquence d'acides aminés la structure 3D après repliement d'une protéine
- **La solution :** AlphaFold algorithm de Deep Learning remporte en 2021 avec succès le **CASP** (Critical Assessment Proteins structure Prediction) avec GDT (Global Distant Test) > 92

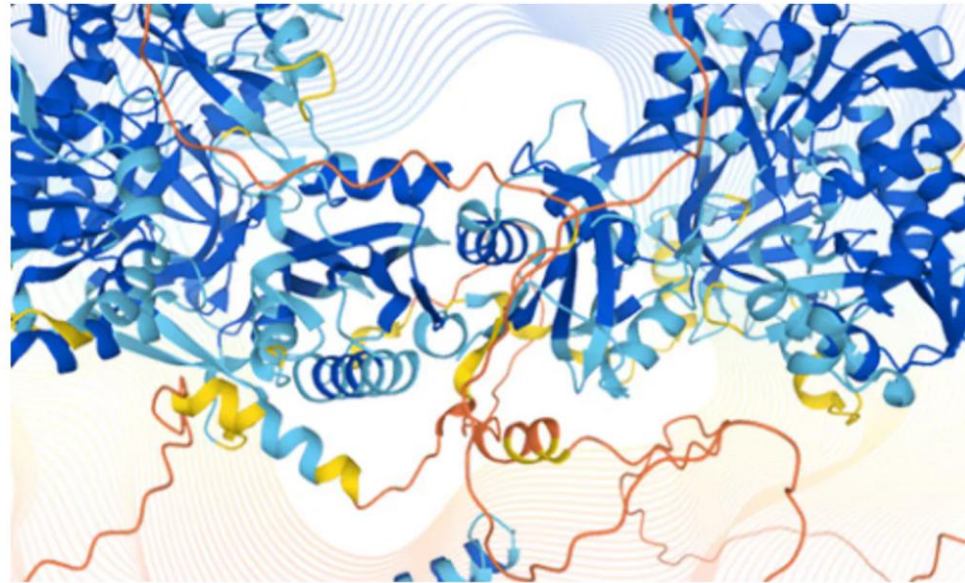


L'I.A. un nouveau challenge pour les Laboratoires : la révolution AlphaFold

Google **Deep-Mind** en partenariat avec **EMBL** (Laboratoire Européen de Biologie Moléculaire) a mis au point AlphaFold

De juillet 2021 à juillet 2022 la **base de données AlphaFold** de structures 3D de protéines a été multipliée par X 200

DeepMind and EMBL release the most complete database of predicted 3D structures of human proteins



Protein structures representing the data obtained via AlphaFold. Source image: AlphaFold. Design credit: Karen Arnott/EMBL-EBI



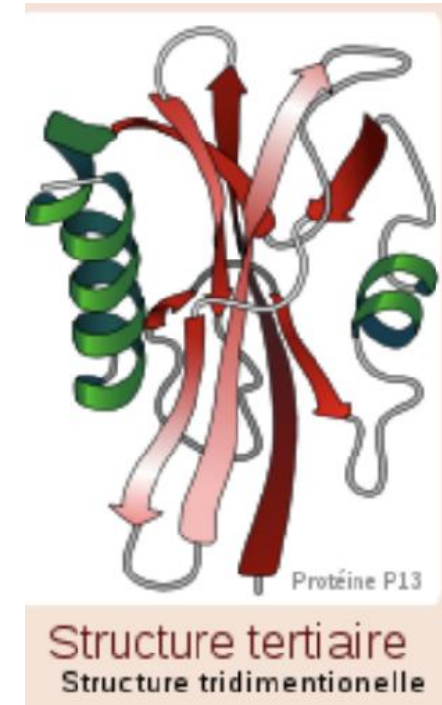
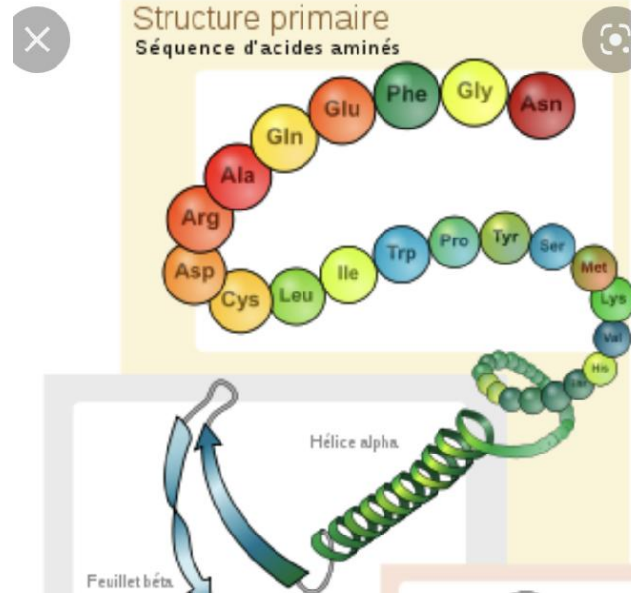
Nouvelle approche & Performances de l'IA : la révolution AlphaFold

La connaissance acquise de la structure dans l'espace de ces molécules que sont les protéines renseigne sur leur fonction dans les organismes vivants

° Ce résultat s'apprécie comme approche combinée associant :

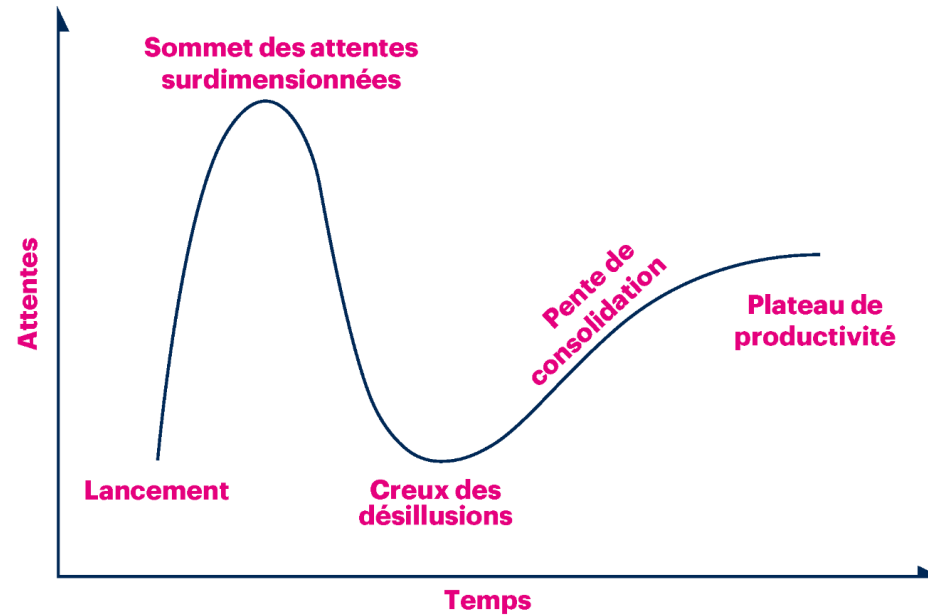
- **les données d'entraînement** : un corpus de données issus du séquençage des acides aminés et de quelques structures 3D établies par les techniques de **Cristallographie X** et **Cryo-microscopie électronique**

- **la capacité de l'IA (AlphaFold)** : à **décrypter le lien** unissant structure primaire / tertiaire



Son intégration en tant que solution pour optimiser le fonctionnement d'un nombre croissant de processus industriels et R&D

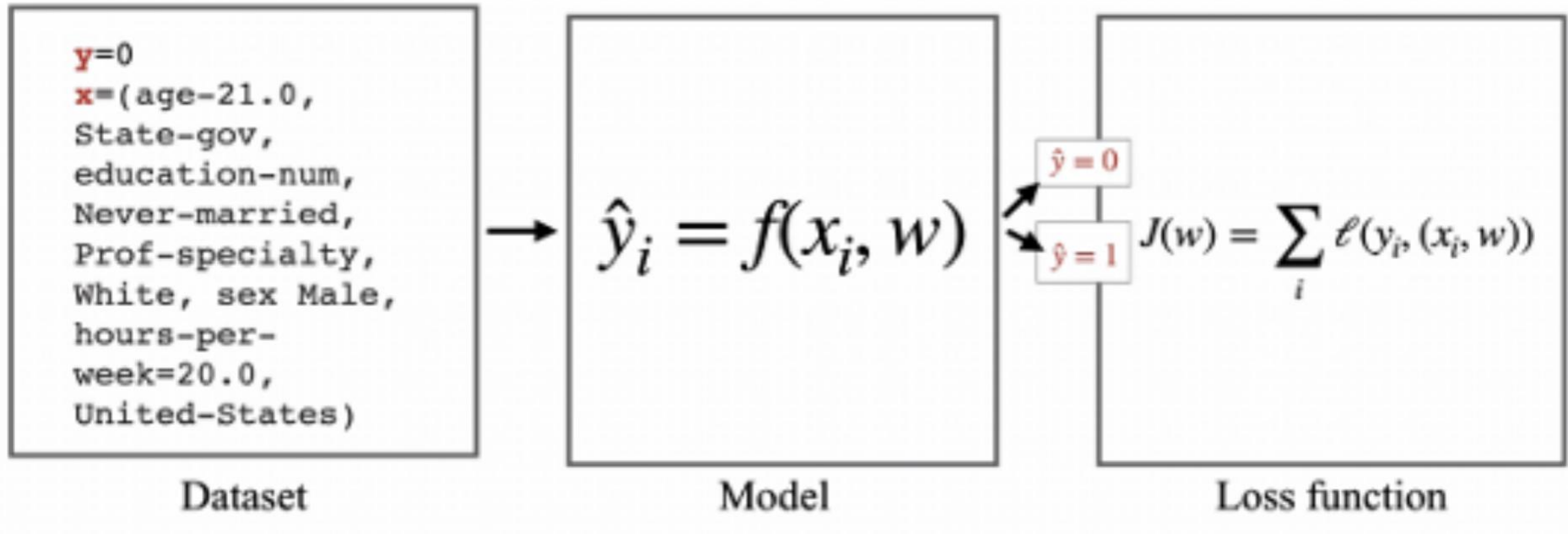
courbe de Gartner
(hype cycle)



Les technologies I.A vont s'installer de manière discrète dans les interstices où elles créent réellement de la valeur



L'IA Matrice du laboratoire du futur: Les Ingrédients...et le principe :

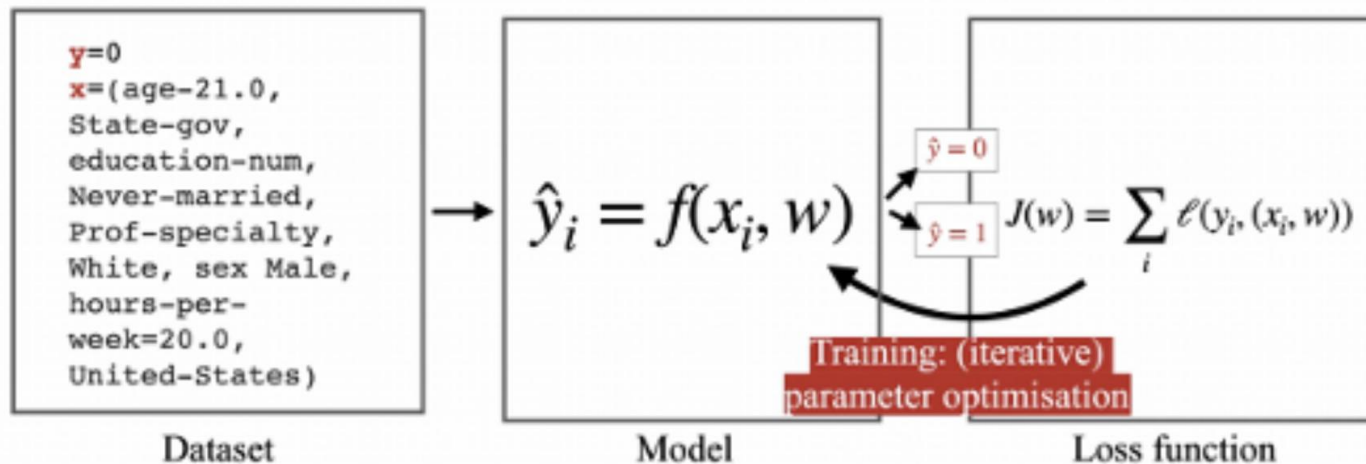
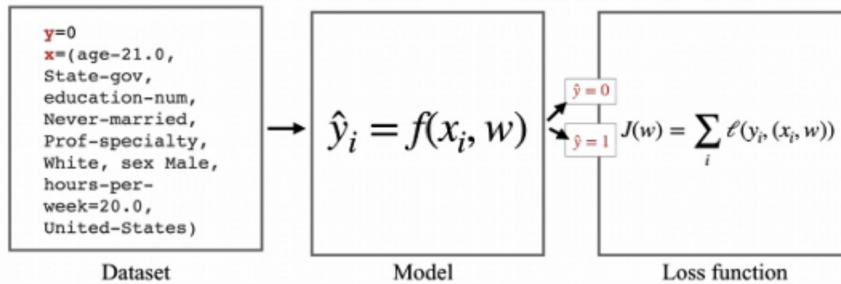
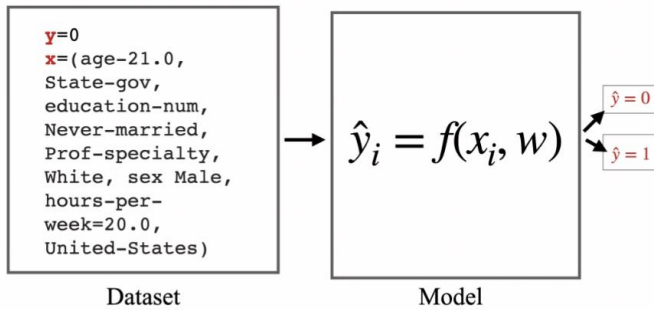


Jeux de données – Modèle – Fonction de de Coût

■ Modèles Paramétriques :

- arbre de décision
- régression
- réseau de neurones





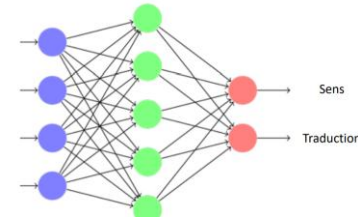
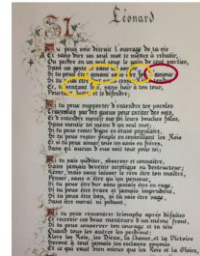
Réseau de Neurons : IA et Prédiction Propriété Chimique

Detournement analyse du langage naturel

Détourner l'Analyse du Langage Naturel (NLP)



Texte

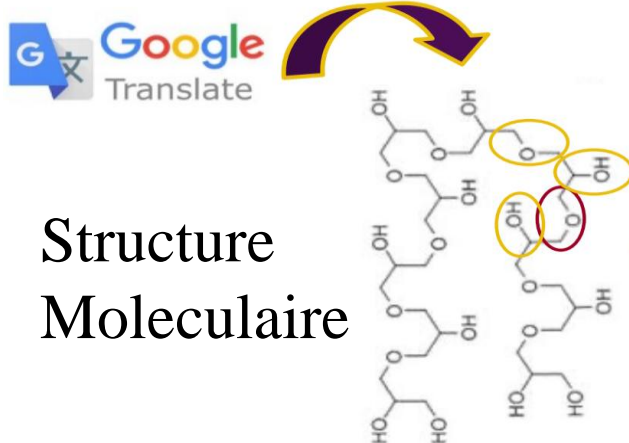


Sens
Traduction

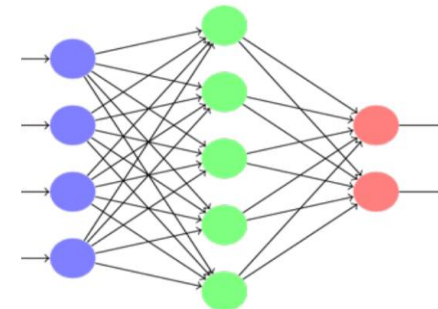
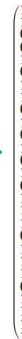
Sens
Traduction

Reseau de neurones

Détourner l'Analyse du Langage Naturel (NLP)



Structure
Moléculaire



Viscosité

...

Viscosité
Propriété

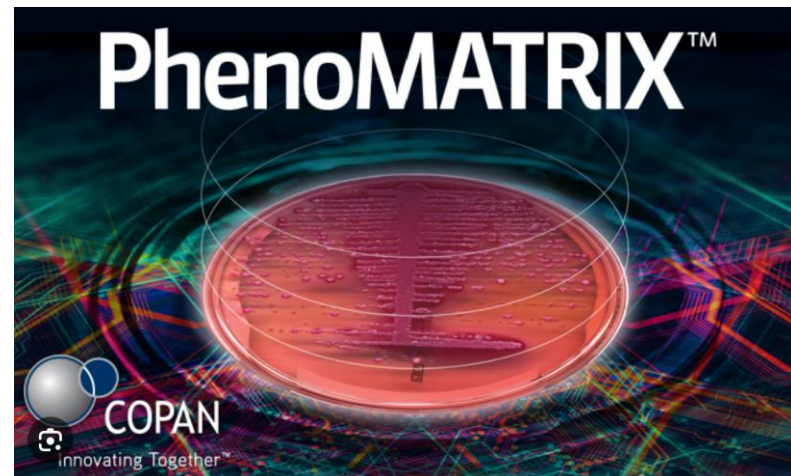


L' Intégration de L'IA dans les Laboratoires...

...Revolutionne le traitement des échantillons

- 1- l'exemple de PhénoMatrix : I.A. embarquée dans système automatisé de traitement échantillons en **Microbiologie**
- 2- un Algorithme de traitement d'images (**Deep Learning**) travaille à partir d'un panel 5000 images fournies par le **laboratoire** apprend et analyse les images mettant en place les meilleures stratégies de détection des éléments pertinents pour classer les images
- 3- Grâce à une intelligence artificielle (**IA**) avancée, le logiciel pré-évalue et pré-trie automatiquement **les plaques de culture** permettant aux laboratoires de microbiologie de **lire, d'interpréter** et de **séparer** les cultures bactériennes

Robotic Process Automation

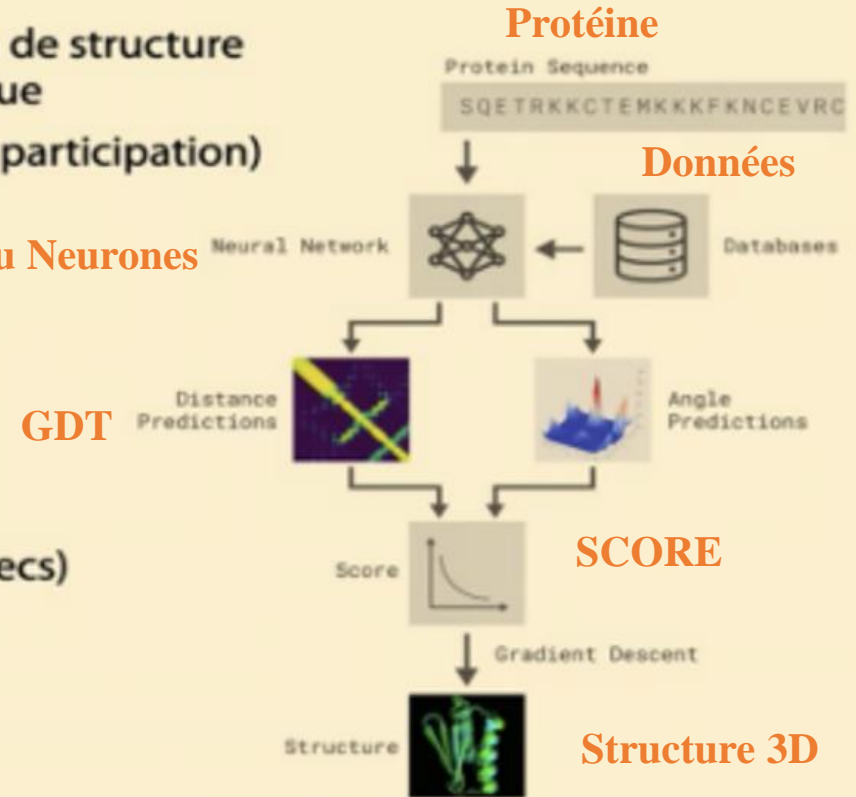


Un tournant en biochimie

- ★ Compétition CASP 13 de prédiction de structure (repliement) de protéines au Mexique
- ★ Un nouveau programme (première participation) a remporté la première place
- ★ AlphaFold gage 25 / 43 tests **Réseau Neurones**
- ★ L'équipe n° 2 a remporté 3 tests

Dans de nombreux domaines

- ★ 1997 : Deep Blue bat Kasparov (échecs)
- ★ 2011 : Watson gagne à *Jeopardy!*
- ★ 2016 : AlphaGo bat Lee Sedol (go)
- ★ 2022 : ChatGPT

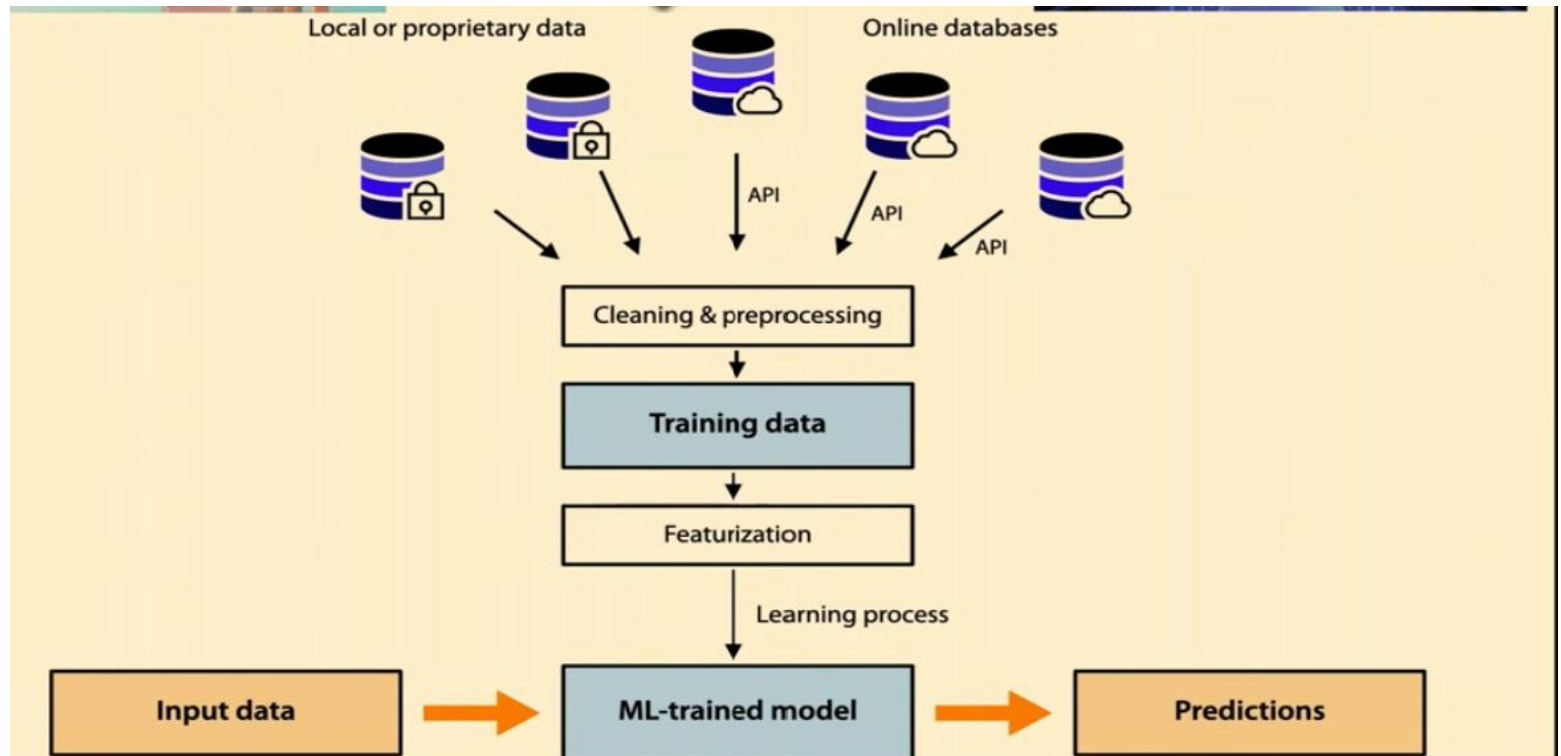


Valorisation et exploitation **des données**

Laboratoire et IA : données d'entraînement , apprentissage et exploitation

Modèle de Machine Learning ...

Bases de données Laboratoires



Apprentissage
Entraînement
du Modèle

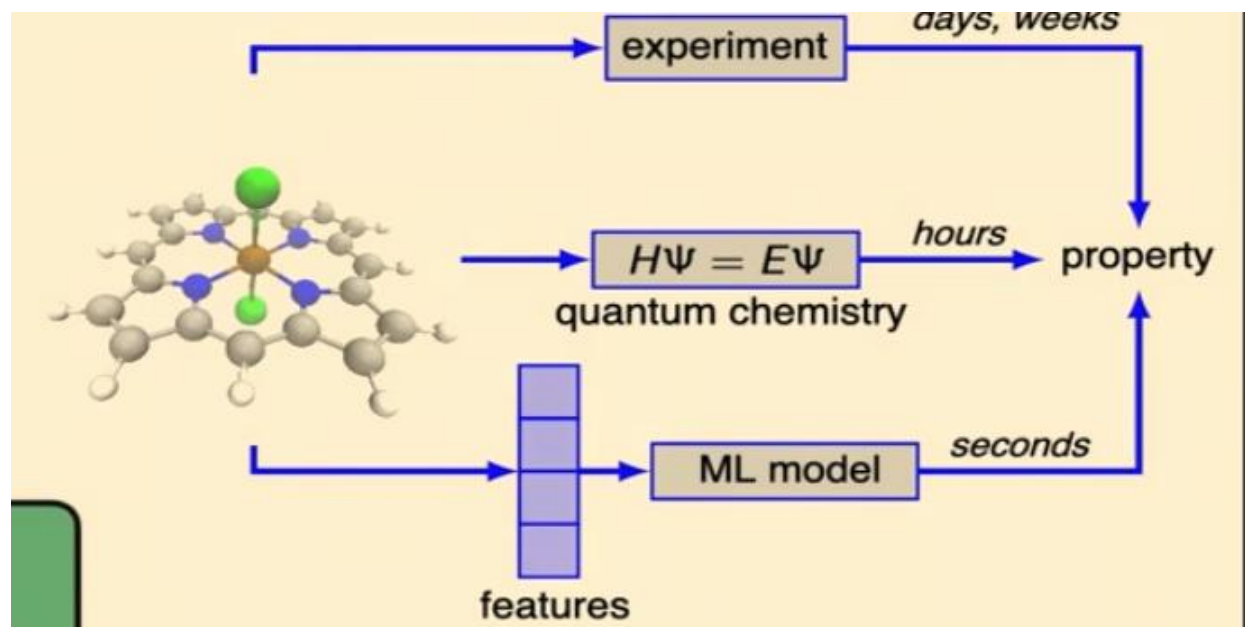


Exploitation du **Modèle** : prédiction solubilité ,enthalpie etc ...

Chimie & IA (ML apprentissage supervisé) ... Performances dans la prédiction des relations structures/ propriétés ...

3 possibilités :

- Expérience
- Chimie Théorique
- Machine Learning



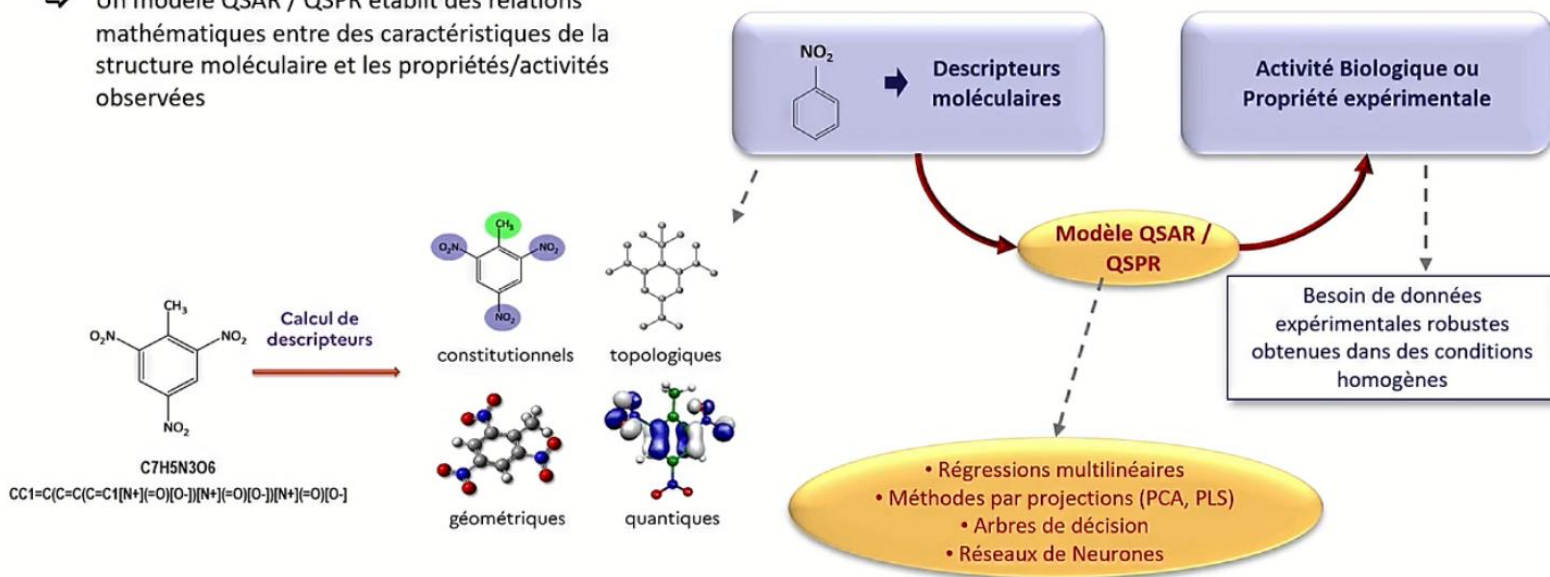
La 3^{ème} voie si vous avez assez des données pour entrainer un **Modèle accélère la prédiction** structure / propriété

Performances = Semaines / Heures/+ Secondes

IA et Nouvelles Approches Méthodologiques pour la maîtrise du Risque Chimique : méthodes alternatives QSAR / QSPR) usages reseaux de Neurones, IA Machine Learning supervisé et non supervisé

Quantitative Structure-Activity/Property Relationships (QSAR/QSPR)

⇒ Un modèle QSAR / QSPR établit des relations mathématiques entre des caractéristiques de la structure moléculaire et les propriétés/activités observées



Usage reseaux de Neurones, IA Machine Learning



APPROCHE IA et Analyse HRMS des Contaminants ...

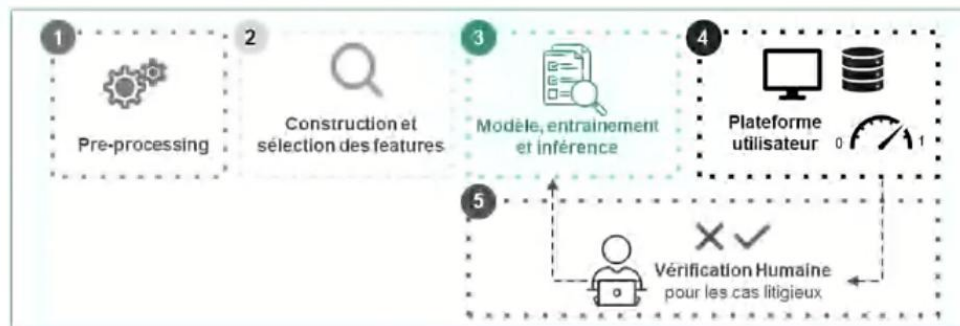
... construction d'un **MODELE** à partir identification caractéristiques particulières **dans les spectres**

pour un développement durable

Projet « Empreintes environnementales »
(AMI IA#2 2020 & France Relance 2021)



SIAPARTNERS

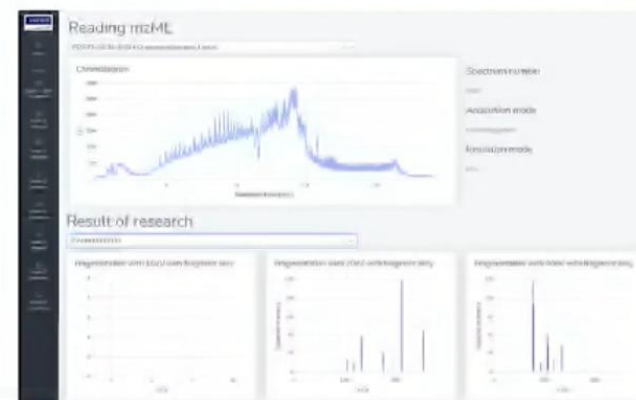
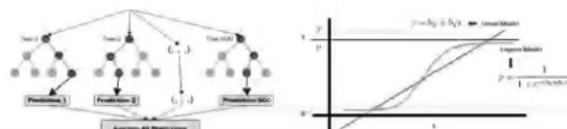


Bases de données

- Base interne Ineris
400 molécules
- Base constructeur
1200 molécules
- Base externe : MassBank
35000 molécules
(4000 avec même équipement que Ineris)

Différentes approches testées :

- MLP Classifier
- Decision Tree
- Random Forest
- KNN Classifier
- SVM Classifier
- Logistic Regression
- Bayes Classifier



GAINS : +> base de données étendue

+> travail + exhaustif & + rapide

@Guillaume Fayet
INERIS



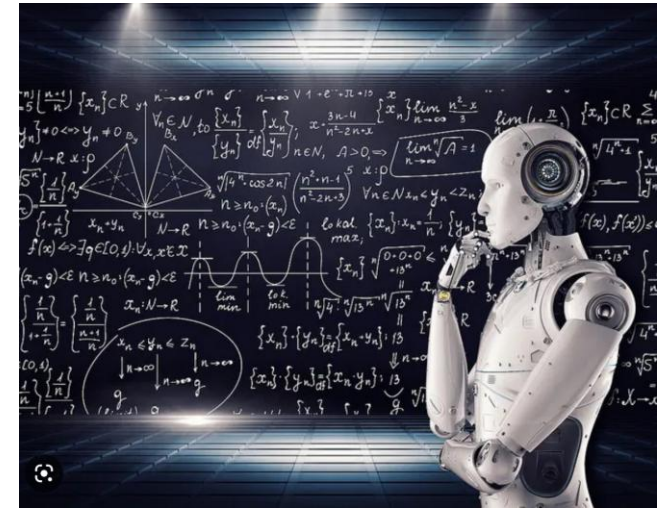
- 1- I.A. Générative** = exploite des techniques d'IA pour générer du contenu (textes , images , son et meme code)
- 2- 2000 Modèle de langue (LLM)** concept clé est formulé
 - = modèle capable de prédire le prochain mot dans une phrase
 - = représentation vectorielle des mots avec coordonnées dans espace sémantique
- 3- 2022 arrivée de ChatGPT (GPT=Generative P retained T ransformer)**
 - = Algorithme capable d'apprendre des représentations de mots contextualisés
 - = **Modèles statistiques probabilistes** qui déterminent l'agencement, la distribution des mots en s'appuyant sur des données massives ingérées lors de leur entraînement

>> **Bing GPT4** — Bibliographie
-- solver d'équation

Microsoft openAI , Anthropic Amazon , Google



L'intelligence Artificielle :
Une opportunité pour les LABORATOIRES !
une Collaboration ...Entre les Laborantins et
les SPECIALISTES de l'IA
Avènement des Laborantins augmentés ?



« La puissance des Algorithmes de Machine – Learning
ouvre une nouvelle ère pour le **laboratoire du futur !** »

Yvon Gervaise



L'I.A. Matrice du laboratoire du Futur ...

#LaboratoireduFutur Pour aller plus loin ..

les Replay vidéos de l'intégralité des #conférences [de 2 colloques >](#)



SECF - Journée "Laboratoire du futur"

<https://webtv.insa-rouen.fr/videos/laboratoire-du-futur-2-le-defi-de-lintelligence-artificielle/>



<https://www.mediachimie.org/ressource/chimie-et-intelligence-artificielle-colloque-f%C3%A9vrier-2023>



Merci !

Yvon Gervaise  <https://www.linkedin.com/in/yvon-gervaise-22a1b035/>

 <https://twitter.com/expertsience>

Mail : expertsience.ygervaise@protonmail.com

Web: <http://www.expertsience.fr>

Tél : 06 60 67 91 50

