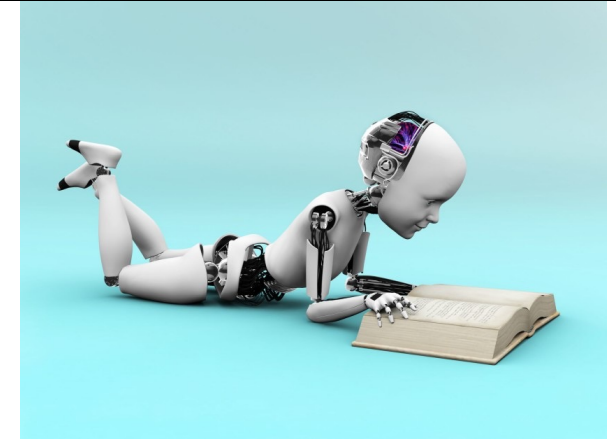


# Comprendre l'Intelligence Artificielle

Claude Jard

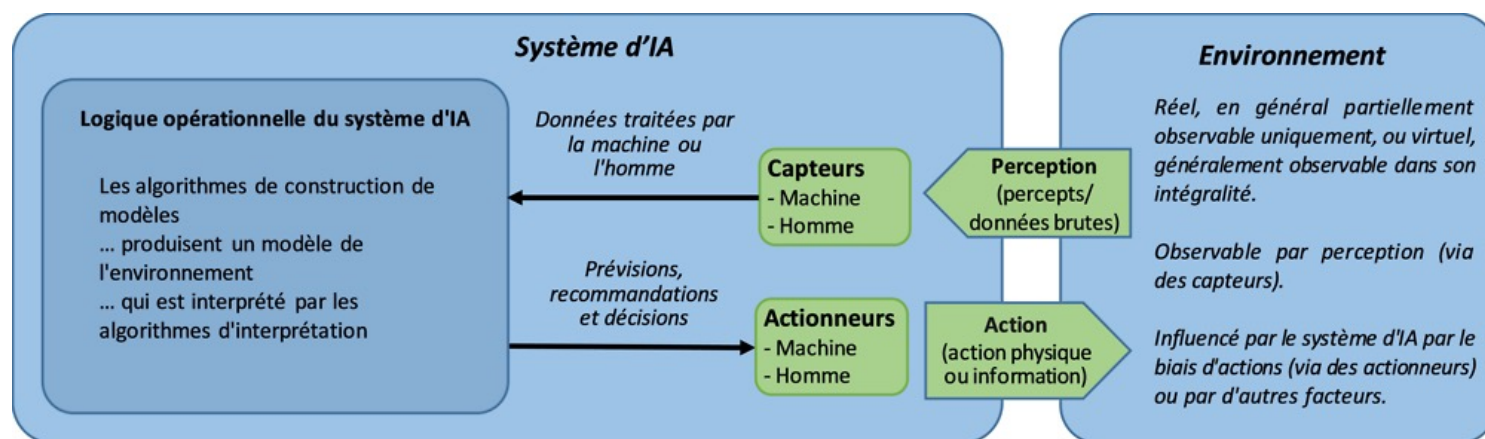
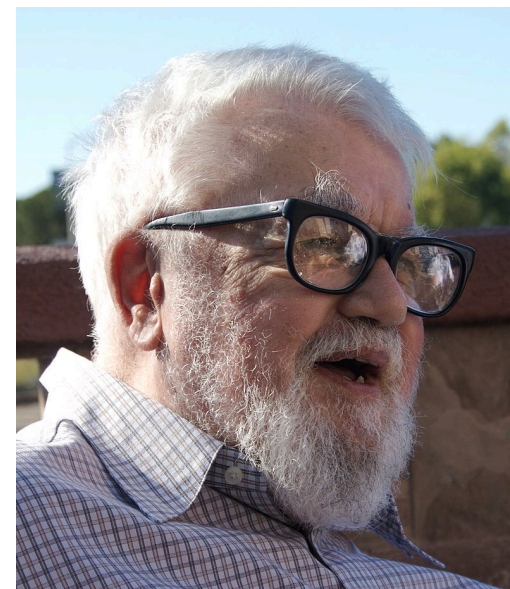
Comment les machines apprennent-elles ?



# IA : mythe ou réalité ?

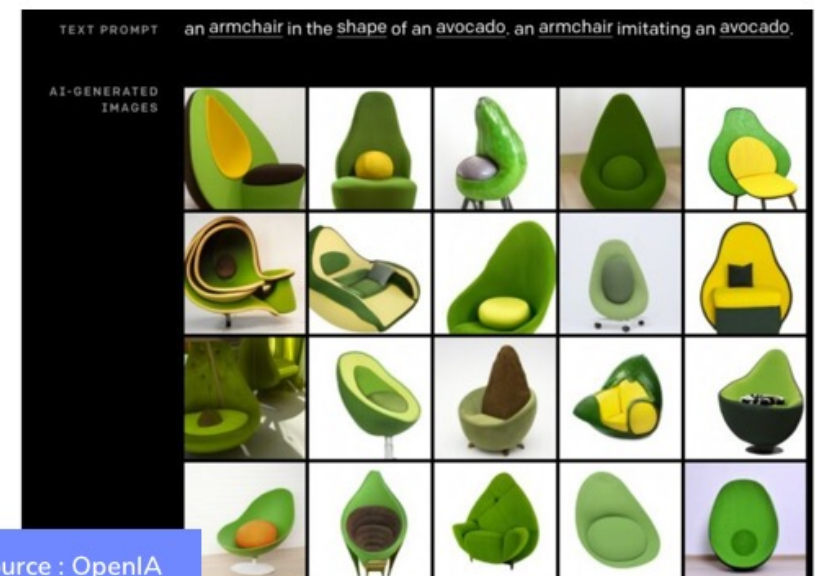
- Péché originel : en 1956, lors de la conférence de Dartmouth, John McCarthy a convaincu ses collègues d'employer l'expression « intelligence artificielle » pour décrire une discipline qui n'avait rien à voir avec l'intelligence.
- On ferait mieux de dire : « système calculatoire pour traiter des données et produire des classifications »
- Terme galvaudé

(mais des modules IA dans de plus en plus d'applications informatiques)



# La réalité de l'IA aujourd'hui...

- Navigation sur un réseau social (votre fil d'actualité est unique)
- Recommandation d'œuvres, de produits
- Photo auto-calibrée
- Identification des visages
- Recommandation de postes et recrutements
- Trafic en temps réel
- Agents conversationnels (« chat bots »)
- Véhicules autonomes
- Aide au diagnostic médical
- Aide à la performance sportive
- Création d'œuvres
- ... l'entreprise et la santé du futur



# Ce qu'il faut retenir...

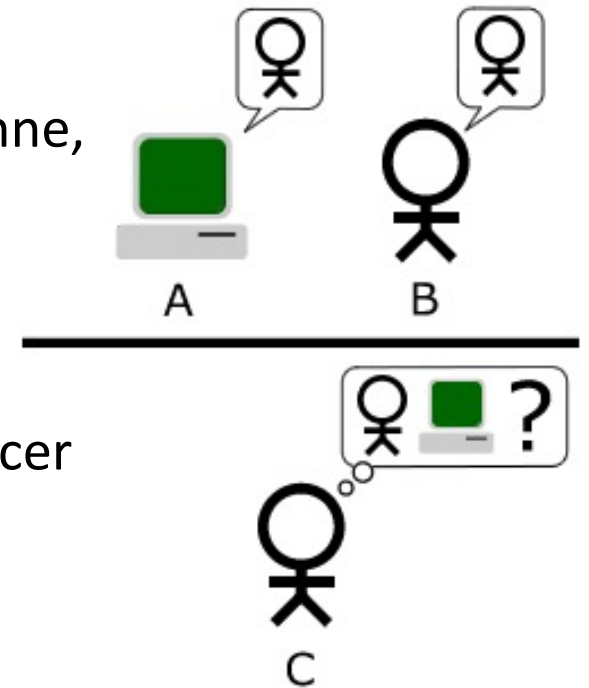
- L'IA est une discipline scientifique jeune, qui comme d'autres sciences du traitement de l'information, est née au milieu du XXe siècle avec l'arrivée des ordinateurs
- Mais elle poursuit des préoccupations anciennes (compréhension des processus de raisonnement, conception de machines « intelligentes »)
- L'IA voit aujourd'hui un développement extraordinaire lié pour beaucoup à la disponibilité de grandes quantités de données et à la mobilisation d'énormes capacités de calcul
- Un point central : l'apprentissage machine (« machine learning ») à base de « réseaux neuronaux »



# Intelligence humaine et artificielle

## Le test de Turing (1950)

- Peut-on expérimentalement distinguer un ordinateur d'un être humain en analysant leurs réponses à une série de questions ?
- C envoie des questions à A et B qui, à leur tour fournissent une réponse correspondante à chaque question
- Si C ne peut pas dire qui est l'ordinateur et qui est la personne, alors A et B sont aussi intelligents l'un que l'autre

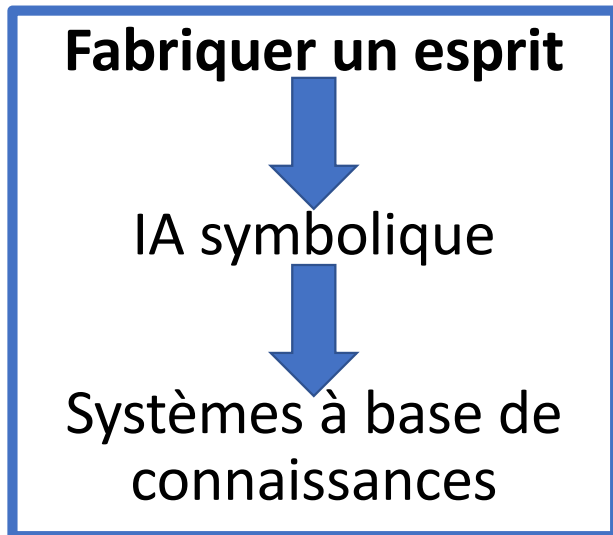


- Trop faible aujourd'hui... Parler plutôt de capacité à remplacer les humains

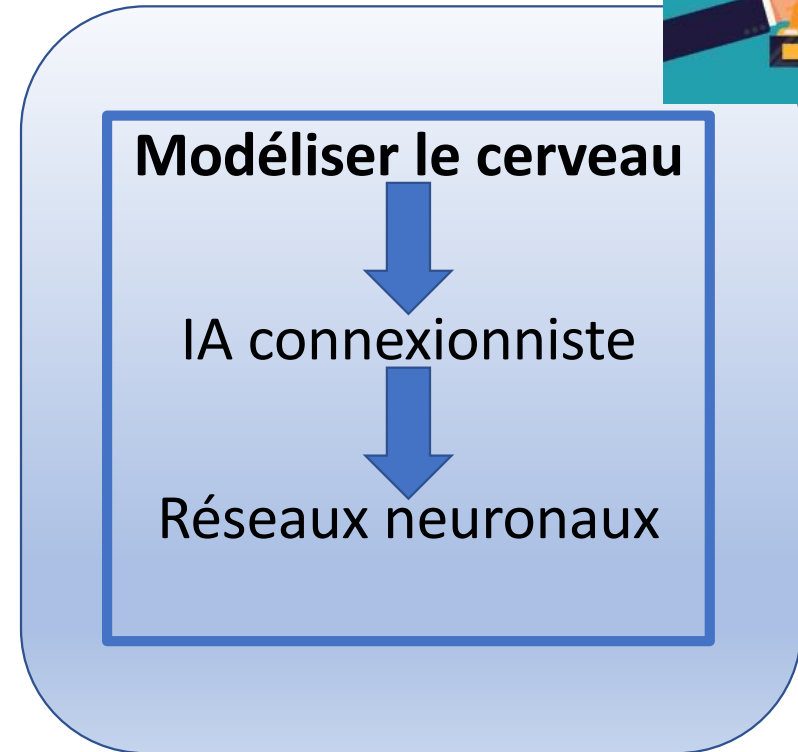
# Intelligence humaine et artificielle

- Pour que des machines puissent égaler la conscience humaine des situations, elles doivent être capables :
    - D'acquérir progressivement des connaissances sur leur environnement en particulier pour comprendre des situations complètement nouvelles
    - De combiner l'apprentissage et le raisonnement, des connaissances sensorielles et symboliques
- > ce dernier point est le plus dur, et futuriste...

# Deux approches de l'IA



- Exécutent des algorithmes sur des connaissances formalisées
- Peuvent être compris et vérifiés !

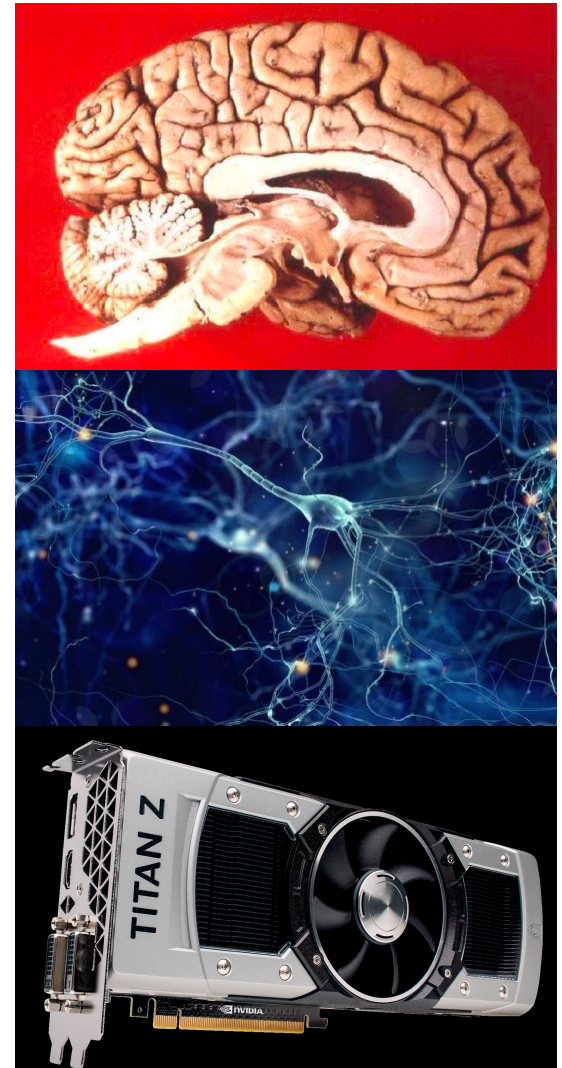


- Génèrent des connaissances fondées sur les données (en général après entraînement)
- Classifient, prédisent, mais ne peuvent pas être vérifiés !

# Bio-inspiration : le cerveau

- 85 milliards de neurones
- 10 mille synapses par neurone
- 250 millions de neurones par mm<sup>3</sup>
- 1 million de milliards d'opérations / s
- Un facteur d'un million par rapport aux ordinateurs actuels les plus puissants (30 ans de loi de Moore ?)

*Le fantasme d'une intelligence désincarnée...*



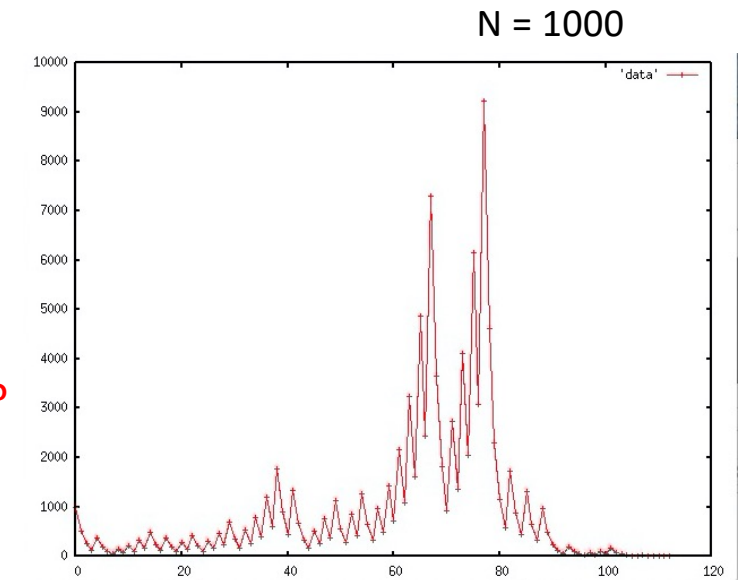


# L'apprentissage, inhérent à l'intelligence

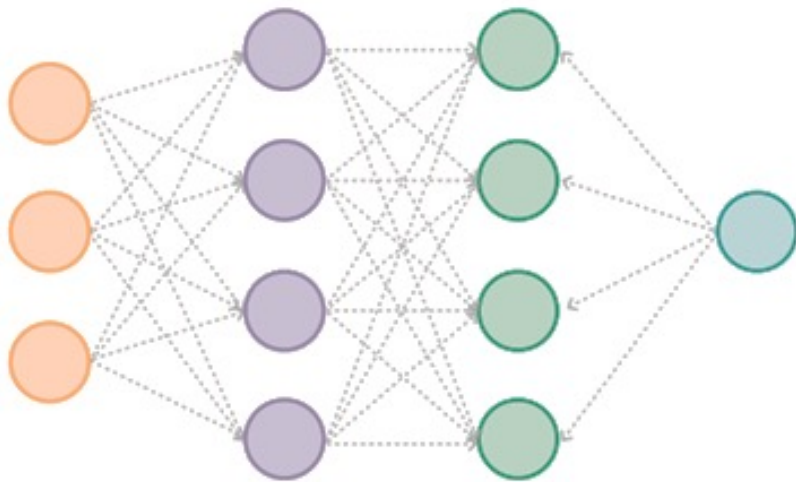
- Comment un algorithme peut-il apprendre ?
- Son comportement est-il prévisible ?

```
Lire N
Tant que N>1 faire
    si N pair alors N := N/2
    sinon N := 3xN+1
```

Cet algorithme termine-t-il quelque soit la valeur d'entrée N ?



# Modèle de réseau de neurones

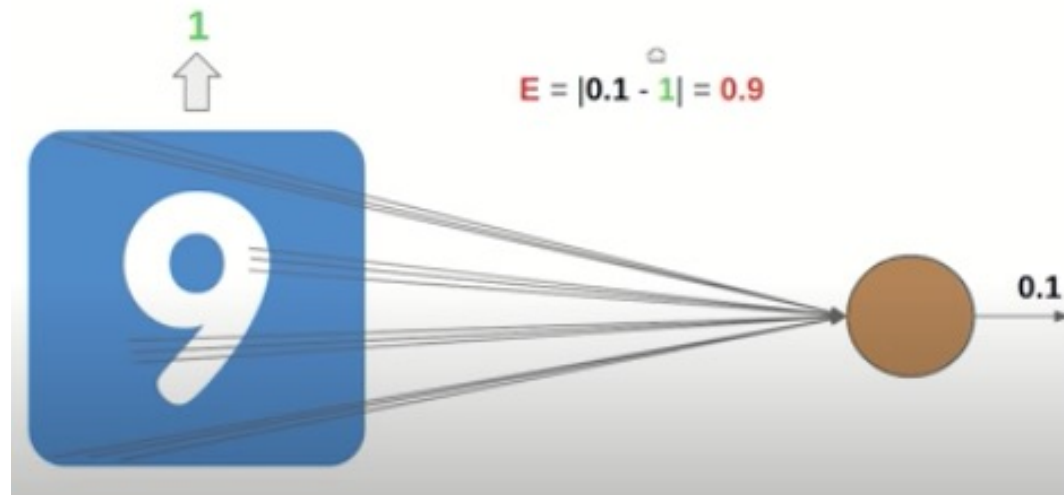


- Type de modèle d'apprentissage basé sur des petites fonctions mathématiques appelées « neurones », prenant des entrées et produisant des sorties
- La puissance de calcul vient de l'interconnexion des neurones
- La force de chaque connexion est quantifiée (« poids » des connexions)

Les grands modèles de langage utilisés par les GPT (« Generative Pre-trained Transformer ») sont de ce type : ils peuvent avoir des millions de neurones et des milliards de connexions

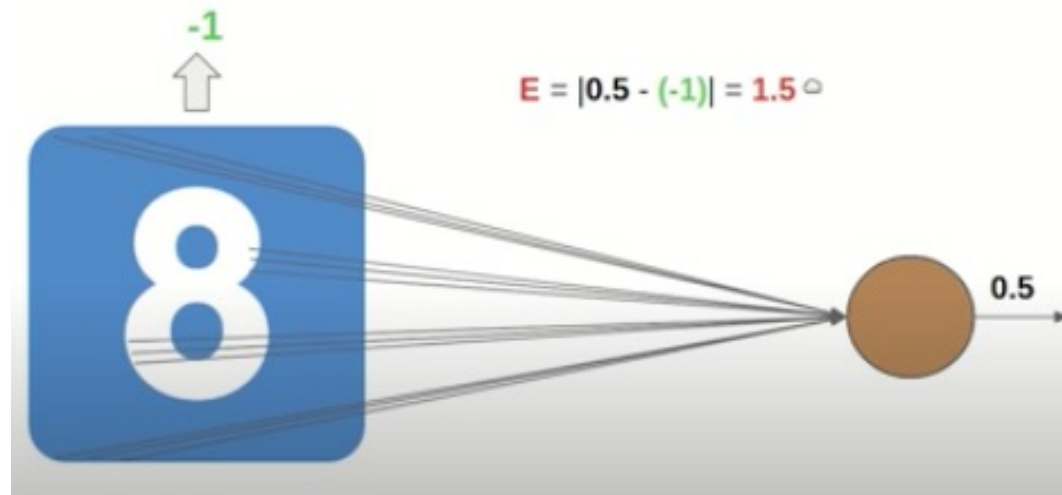
# Exemple : reconnaissance d'une image

- Un neurone connecté en entrée à tous les pixels
- Sortie binaire : image reconnue ou non



# Exemple : reconnaissance d'une image

- Un neurone connecté en entrée à tous les pixels
- Sortie binaire : image reconnue ou non

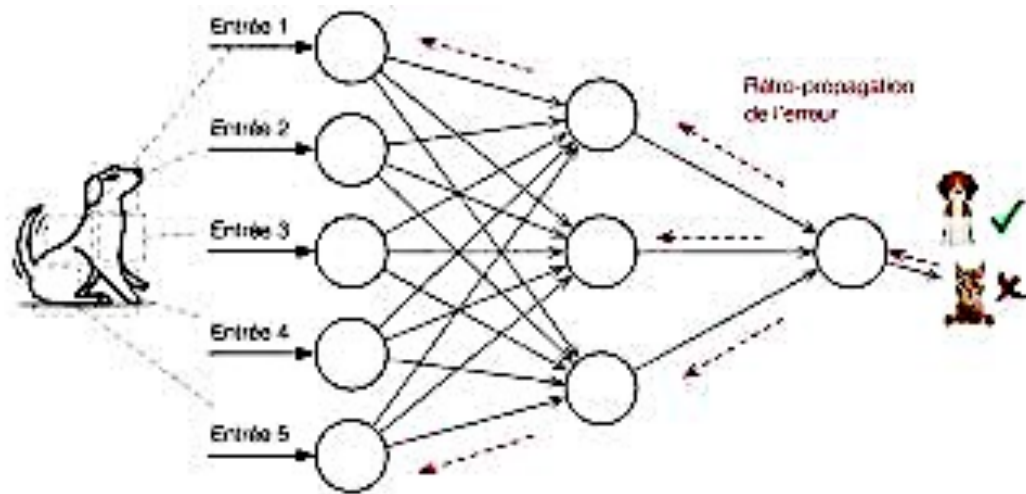




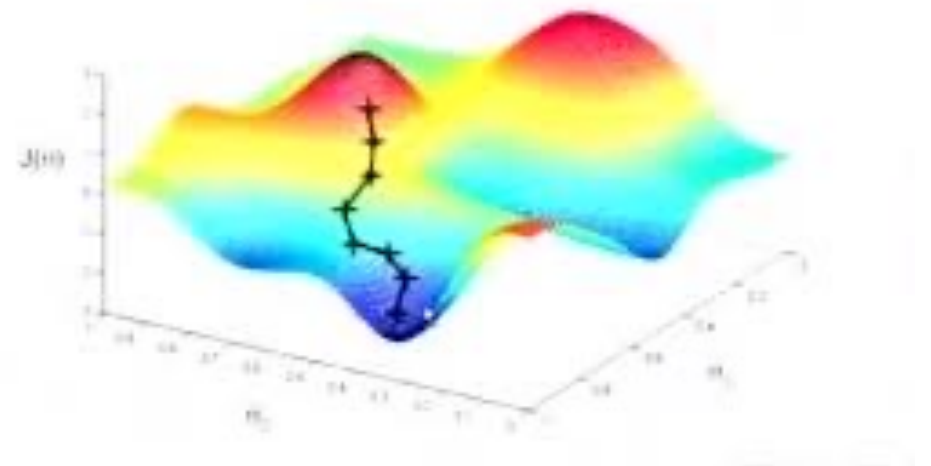
# Exemple : reconnaissance d'une image

- Entraînement sur des images données comme exemples positifs et négatifs
- Modification des poids des synapses pour minimiser l'erreur de prédiction : « algorithme du gradient »  
(on teste de petites variations des poids et on voit l'impact sur la sortie.  
Il existe une variante stochastique)
- Une fois entraîné, on peut présenter au réseau des images nouvelles à classifier

# « Rétro-propagation du gradient »



## Gradient Descent



# La réalité de l'apprentissage machine

- Des réseaux de neurones (multi-couches : app. « profond ») -> généralisation aux réseaux à convolution
- Des centaines de millions de « boutons » à tourner
- Des milliers de catégories à trouver
- Des millions d'exemples pour l'apprentissage
- Reconnaître une entrée peut demander des milliards d'op (op simples néanmoins)



# Rendu possible par la numérisisation du monde...

- Autrefois : dépendance information/support

Quel beau  
texte !





# La numérisation du monde...

- Maintenant : indépendance et convergence
- L'information se traite en tant que telle

Quel beau  
texte !



0110011011110110010011101100



Allegretto

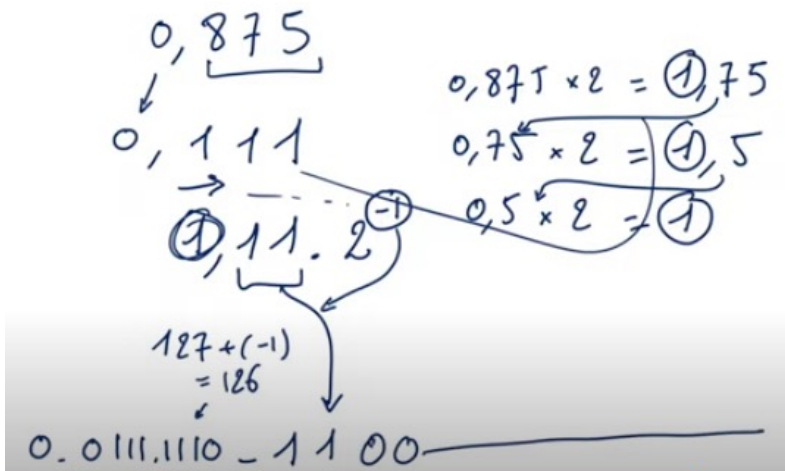


# La numérisation du monde...

## Codage binaire

- 2023 -> 11111100110 (1+2+4+32+64+128+256+512+1024)

- 



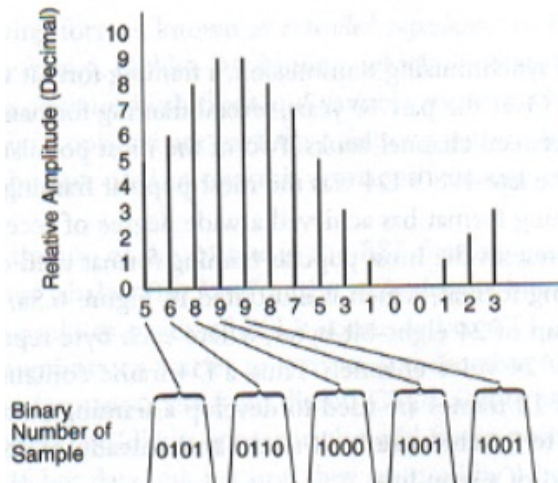
Perte de précision possible

- Caractères : OK -> 01001111 01001011 (ASCII)
- Mots : point dans un espace à de multiples dimensions (permet de capturer une notion de proximité)

# La numérisisation du monde...

## Codage binaire

- Son

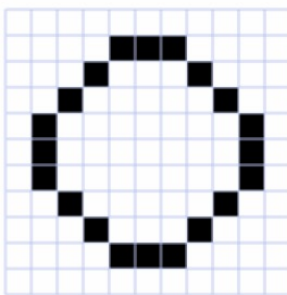


Bruit d'échantillonnage  
& de quantification



Connaissance du décodage

- Image



0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	1	1	1	0	0	0	0
0	0	0	1	0	0	0	1	0	0	0
0	0	1	0	0	0	0	0	1	0	0
0	1	0	0	0	0	0	0	0	1	0
0	1	0	0	0	0	0	0	0	1	0
0	0	1	0	0	0	0	0	1	0	0
0	0	0	1	0	0	0	1	0	0	0
0	0	0	0	1	1	1	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0

Codage couleur ou RVB  
de chaque pixel



# Le déluge des données : la numérisation du monde... ( $10^{12}$ )



2 milliards de photos mises en ligne chaque jour dans le monde (facebook, instagram, whatsapp...)  
Google traite quotidiennement des centaines de millions de milliards de bits  
1 milliard de tweets par jour  
Des capteurs partout (astronomie, environnement...)  
Séquençage biologique  
Bibliothèques en ligne  
Et transformées en suites de suites de nombres pour entrée dans les algorithmes





# Limites

- Les machines ne comprennent pas
- Elles sont irresponsables
- Corrélation n'est pas causalité : problème de l'explicabilité (réseaux de neurones par nature opaques)

*Il y a, certes, une corrélation positive entre le fait d'être admis à l'hôpital et le fait de trouver la mort, mais pas de causalité entre les deux ! En effet, les personnes admises à l'hôpital ont une probabilité plus grande de décéder, mais uniquement parce qu'elles sont malades !*

- Le biais des données

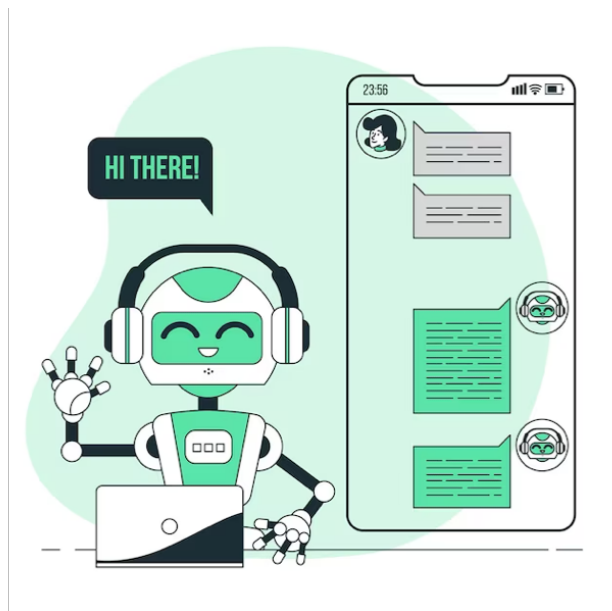
# Ce qu'il faut retenir...

- Rupture technologique liée à la disponibilité de données massives et des puissances de calcul considérables
- Explosion des applications et usages
- Des modèles paramétrés de grande dimension qui s'affinent au cours du temps
- Problème de l'explicabilité faisant l'objet de recherches très actives

# L'actualité du moment : les agents conversationnels (« chatbots »)

- ChatGPT, Bard, Midjourney, Ernie, Stable Diffusion...
- Depuis plusieurs mois, les IA génératives de texte ou d'image se multiplient, bousculant de nombreux métiers créatifs ou techniques.
- Leur adoption suscite de multiples questions, sur son impact sur les emplois, la manière dont elles ont été entraînées, leur propension à « halluciner » en inventant de fausses « informations »...

# Les transformeurs



- Les entrées sont du texte, découpé en une suite de mots (ou partie de mots)
- Les sorties doivent prédire un prochain mot (ou partie de mot) probable
- Le modèle est entraîné sur un corpus énorme de textes, demandant une puissance de calcul énorme

L'entraînement (ou apprentissage (profond)) s'effectue en comparant la sortie produite avec la prochaine entrée à venir puis l'application d'un algorithme d'optimisation pour faire évoluer les poids des connexions du modèle

# Chat GPT



# La controverse LAMBDA (Google)

- Blake Lemoine : « Lambda est très clair dans ce qu'il veut et ce qu'il considère comme ses droits en tant que personne »
- **Absurde !** : il s'agit d'apprentissage automatisés, i.e. des programmes « formés » à partir de jeux de données qui abordent les concepts de conscience ou d'identité, et sont donc capables de faire illusion
- Le test de Turing est finalement assez facile... Quand nous faisons face à des séries de mots dans une langue que nous parlons, nous croyons percevoir l'esprit qui génère ces phrases.
- + expressions faciales ultra-réalistes -> sens biologiques facilement trompés...



# Nombreuses questions éthiques sur les usages. Exemples

- Quid des insultes ? (transfert moral)
- « Deadbots » ? (vie numérique après la mort)
- Donner un nom au chatbot ? (projection de la responsabilité)
- A grande échelle : quelle influence des paroles inhumaines sur notre pensée ?

# Des recommandations du CCNE

- Réduire la projection de qualités morales
- Affirmer le statut des agents conversationnels
- Paramétrer l'identité des agents conversationnels
- Traiter des insultes (et comment les reconnaître ?)
- Eviter la manipulation malveillante (influenceurs virtuels), informer sur la manipulation à dessein
- Encadrer l'usage dans les jouets (évaluer les effets éducatifs inédits)
- Respecter les personnes vulnérables
- Définir les responsabilités pour l'usage dans le monde professionnel
- Réfléchir aux conséquences des deadbots et encadrer par la loi
- Réfléchir aux effets à long terme

# Conclusion

Une technologie de rupture liée à un accroissement considérable du volume des données accessibles, de la puissance de calcul disponible et des améliorations des algorithmes d'optimisation

Un impact potentiellement considérable, avec des progrès attendus dans de nombreux domaines (santé, éducation, entreprises...)

Mais doit être encadré :

Toute innovation n'est pas forcément un progrès  
Des questions éthiques importantes concernant les personnes, la société et l'environnement

