



Mardi 18 juillet 2023

L'intelligence Artificielle pour les Juristes

Histoire de l'IA : De Turing à aujourd'hui

Issa Traoré (PhD)
Malik Koné (PhD)

Sommaire

1. Préambule

Objectif de l'informatique

Les applications de l'IA

2. Avant le XXème siècle

Algorithmes et automates pour de justes partages

Machines et Algorithmes

Alan Turing

3. Milieu du XXème siècle

Les débuts de l'Intelligence Artificielle

L'apprentissage machine

4. Fin du XXème siècle

Réseaux de neurones

La rétro-propagation (1974 - 1986)

5. Début du XXIème siècle

Le Big Data

Etat de l'art

Les erreurs et les biais

Objectif de l'informatique



- ▶ Remplacer l'homme pour certaines tâches

Oui, mais lesquelles ?

Objectif de l'informatique

- ▶ Remplacer l'homme pour certaines tâches

Oui, mais lesquelles ?

Remplacer les Artistes ?



Remplacer les Chirurgiens ?



Objectif de l'informatique

- ▶ Remplacer l'homme pour certaines tâches

Oui, mais lesquelles ?

Remplacer les Artistes ?



Remplacer les Chirurgiens ?

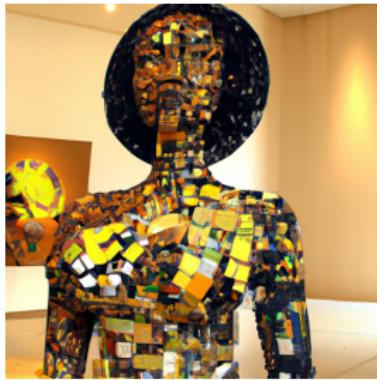


Objectif de l'informatique

- ▶ Remplacer l'homme pour certaines tâches

Oui, mais lesquelles ?

Remplacer les Artistes ?



Remplacer les Chirurgiens ?



Sommaire

1. Préambule

Objectif de l'informatique

Les applications de l'IA

2. Avant le XXème siècle

Algorithmes et automates pour de justes partages

Machines et Algorithmes

Alan Turing

3. Milieu du XXème siècle

Les débuts de l'Intelligence Artificielle

L'apprentissage machine

4. Fin du XXème siècle

Réseaux de neurones

La rétro-propagation (1974 - 1986)

5. Début du XXIème siècle

Le Big Data

Etat de l'art

Les erreurs et les biais

Les applications de l'IA

Domaine du Text

- ▶ Classification de textes
- ▶ Extraction et résumer des informations
- ▶ Résoudre ou répondre à des questions
- ▶ Traduire
- ▶ Générer du texte en plus de 100 langues

Domaine des Images

- ▶ Classification d'images,
- ▶ Détection et segmentation des objets
- ▶ Génération d'images à partir de

Domaine de Audio

- ▶ Reconnaissance de la parole
- ▶ Génération de parole
- ▶ Classification audio

Les applications de l'IA

Domaine du Text

- ▶ Classification de textes
- ▶ Extraction et résumer des informations
- ▶ Résoudre ou répondre à des questions
- ▶ Traduire
- ▶ Générer du texte en plus de 100 langues

Domaine des Images

- ▶ Classification d'images,
- ▶ Détection et segmentation des objets
- ▶ Génération d'images à partir de

Domaine de Audio

- ▶ Reconnaissance de la parole
- ▶ Génération de parole
- ▶ Classification audio

Les applications de l'IA

Domaine du Text

- ▶ Classification de textes
- ▶ Extraction et résumer des informations
- ▶ Résoudre ou répondre à des questions
- ▶ Traduire
- ▶ Générer du texte en plus de 100 langues

Domaine des Images

- ▶ Classification d'images,
- ▶ Détection et segmentation des objets
- ▶ Génération d'images à partir de

Domaine de Audio

- ▶ Reconnaissance de la parole
- ▶ Génération de parole
- ▶ Classification audio

Sommaire

1. Préambule

Objectif de l'informatique

Les applications de l'IA

2. Avant le XXème siècle

Algorithmes et automates pour de justes partages

Machines et Algorithmes

Alan Turing

3. Milieu du XXème siècle

Les débuts de l'Intelligence Artificielle

L'apprentissage machine

4. Fin du XXème siècle

Réseaux de neurones

La rétro-propagation (1974 - 1986)

5. Début du XXIème siècle

Le Big Data

Etat de l'art

Les erreurs et les biais

Automates, Ordinateurs Analogiques : Calculer le temps de la justice.



Antikythera 200 BCE (video)

L'informaticien du sultan au services des kadis

Al-Khwârizmî (origine du mot "algorithme"), mathématicien de Bagdad.



Sages et Juges à la court de Bagdad vers 700 AD

La logique : "Tout n'est qu'absence ou présence" 0 ou 1.

Le transistor



hier



aujourd'hui (vidéo intel)

Sommaire

1. Préambule

Objectif de l'informatique

Les applications de l'IA

2. Avant le XXème siècle

Algorithmes et automates pour de justes partages

Machines et Algorithmes

Alan Turing

3. Milieu du XXème siècle

Les débuts de l'Intelligence Artificielle

L'apprentissage machine

4. Fin du XXème siècle

Réseaux de neurones

La rétro-propagation (1974 - 1986)

5. Début du XXIème siècle

Le Big Data

Etat de l'art

Les erreurs et les biais

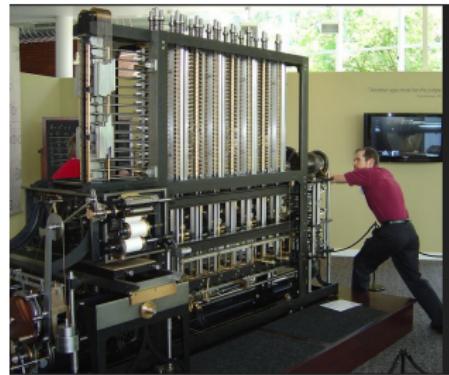
Automatiser les calculs

La Pascaline



au environs de 1650

La machine de Charles Babbage et Ada Lovelace



au environs 1830 (video)

Sommaire

1. Préambule

Objectif de l'informatique

Les applications de l'IA

2. Avant le XXème siècle

Algorithmes et automates pour de justes partages

Machines et Algorithmes

Alan Turing

3. Milieu du XXème siècle

Les débuts de l'Intelligence Artificielle

L'apprentissage machine

4. Fin du XXème siècle

Réseaux de neurones

La rétro-propagation (1974 - 1986)

5. Début du XXIème siècle

Le Big Data

Etat de l'art

Les erreurs et les biais

Alan Turing



Sommaire

1. Préambule

Objectif de l'informatique

Les applications de l'IA

2. Avant le XXème siècle

Algorithmes et automates pour de justes partages

Machines et Algorithmes

Alan Turing

3. Milieu du XXème siècle

Les débuts de l'Intelligence Artificielle

L'apprentissage machine

4. Fin du XXème siècle

Réseaux de neurones

La rétro-propagation (1974 - 1986)

5. Début du XXIème siècle

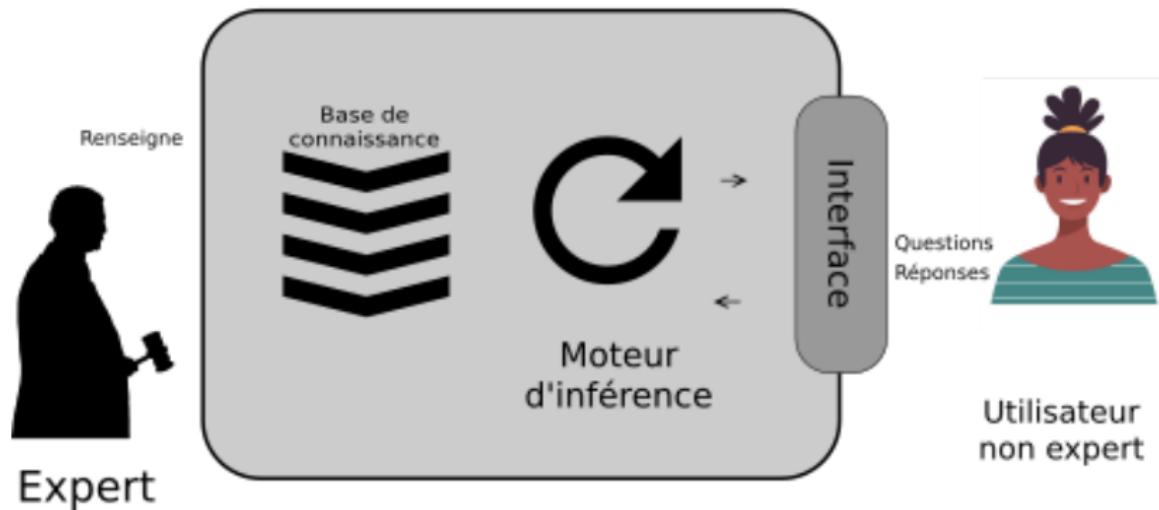
Le Big Data

Etat de l'art

Les erreurs et les biais

Formaliser les connaissances

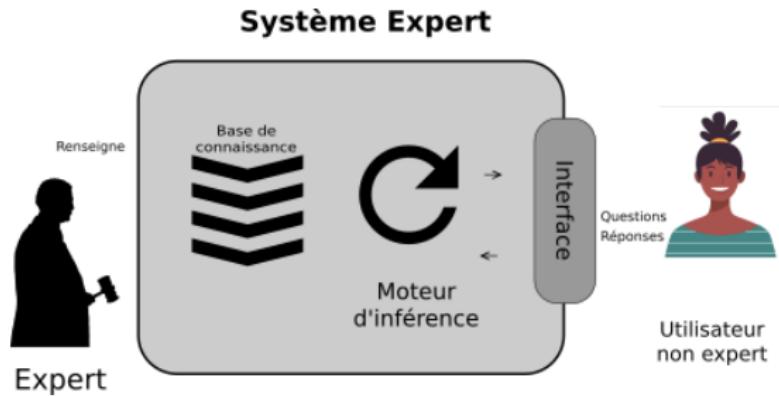
Système Expert



Problèmes

- ▶ Volume important de connaissances à formaliser (projet ontologies).
- ▶ Beaucoup de raisonnements inconscients donc non formalisable

Formaliser les connaissances



Problèmes

- ▶ Volume important de connaissances à formaliser (projet ontologies).
- ▶ Beaucoup de raisonnements inconscients donc non formalisable

Sommaire

1. Préambule

Objectif de l'informatique

Les applications de l'IA

2. Avant le XXème siècle

Algorithmes et automates pour de justes partages

Machines et Algorithmes

Alan Turing

3. Milieu du XXème siècle

Les débuts de l'Intelligence Artificielle

L'apprentissage machine

4. Fin du XXème siècle

Réseaux de neurones

La rétro-propagation (1974 - 1986)

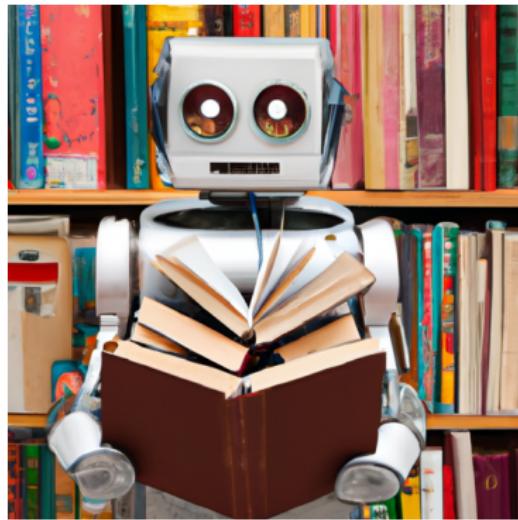
5. Début du XXIème siècle

Le Big Data

Etat de l'art

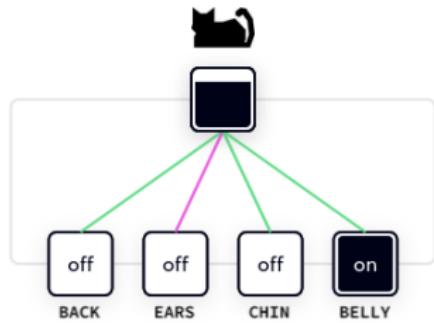
Les erreurs et les biais

C'est peut-être plus simple de faire apprendre



Comment fonctionne les réseaux de neurones

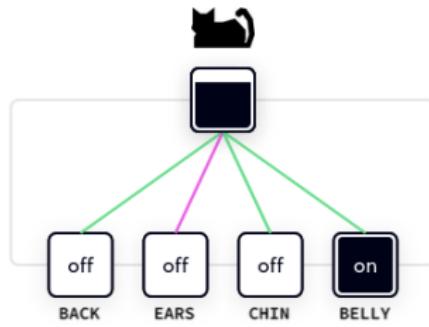
5 neurones, 1 couche



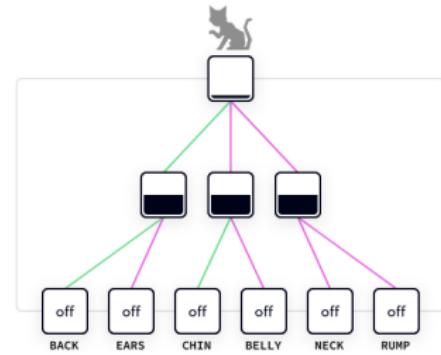
10 neurones, 2 couches

Comment fonctionne les réseaux de neurones

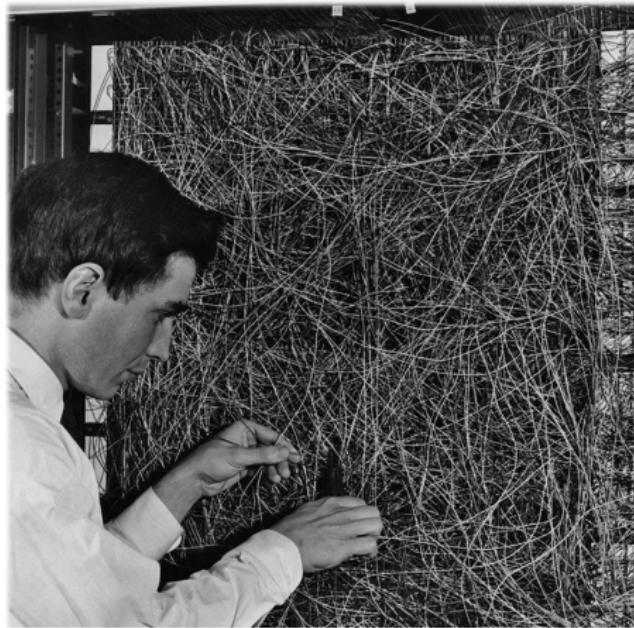
5 neurones, 1 couche



10 neurones, 2 couches



Le Perceptron



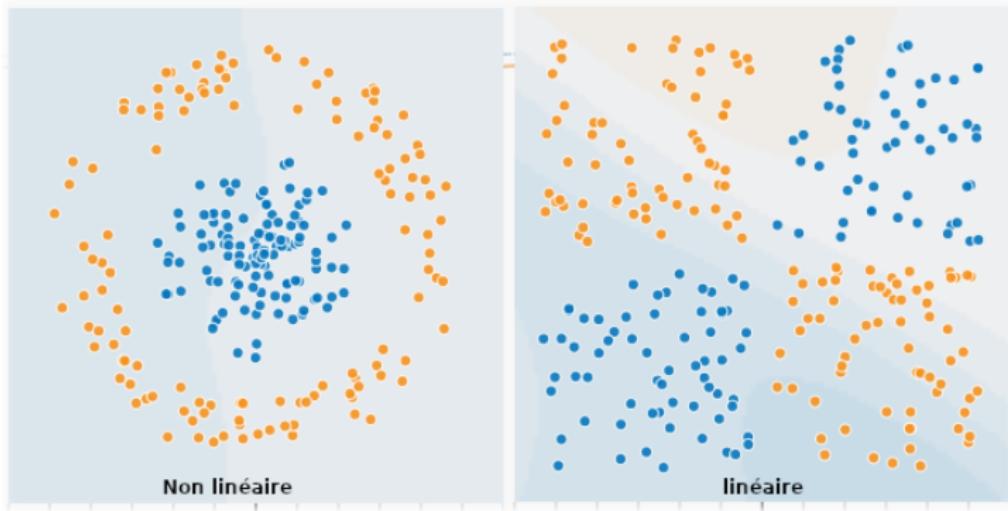
1958

- ▶ une couche, beaucoup de neurone + apprentissage automatique

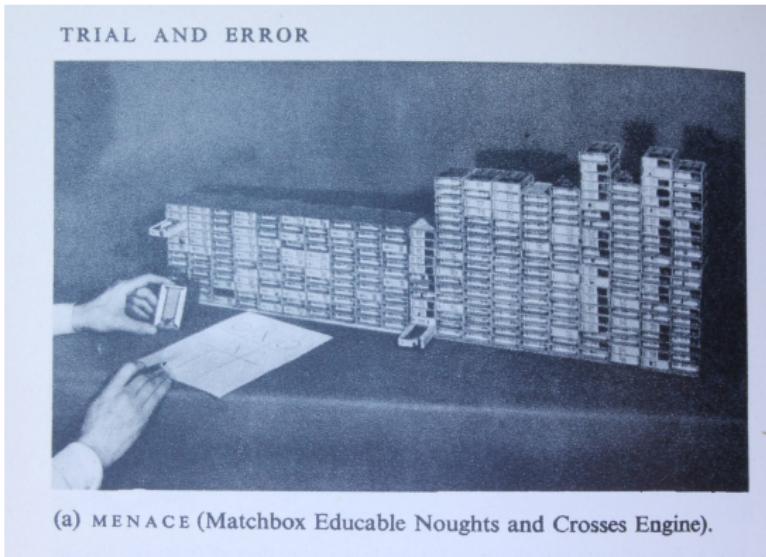
Le Perceptron

- ▶ une couche, beaucoup de neurone + apprentissage automatique

Problème Ne traite pas tous les cas.



La MENACE



Matchbox Educable Noughts and Crosses Engine

- ▶ Apprentissage par renforcement
- ▶ + perceptron multicouche -> Apprentissage profond.

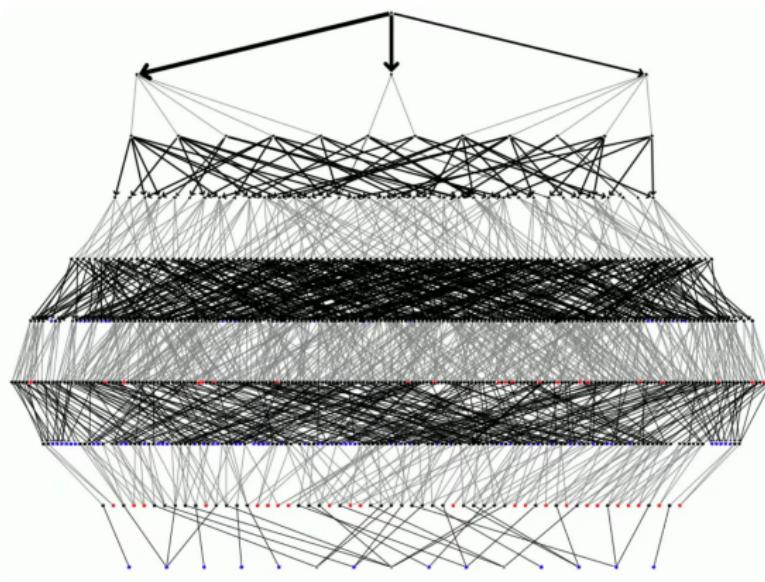
La MENACE



Apprendre à jouer à Tic-Tac-Toe

- ▶ Apprentissage par renforcement
- ▶ + perceptron multicouche -> Apprentissage profond.

La MENACE



L'arbre de probabilité

- ▶ Apprentissage par renforcement
- ▶ + perceptron multicouche -> Apprentissage profond.

Sommaire

1. Préambule

Objectif de l'informatique

Les applications de l'IA

2. Avant le XXème siècle

Algorithmes et automates pour de justes partages

Machines et Algorithmes

Alan Turing

3. Milieu du XXème siècle

Les débuts de l'Intelligence Artificielle

L'apprentissage machine

4. Fin du XXème siècle

Réseaux de neurones

La rétro-propagation (1974 - 1986)

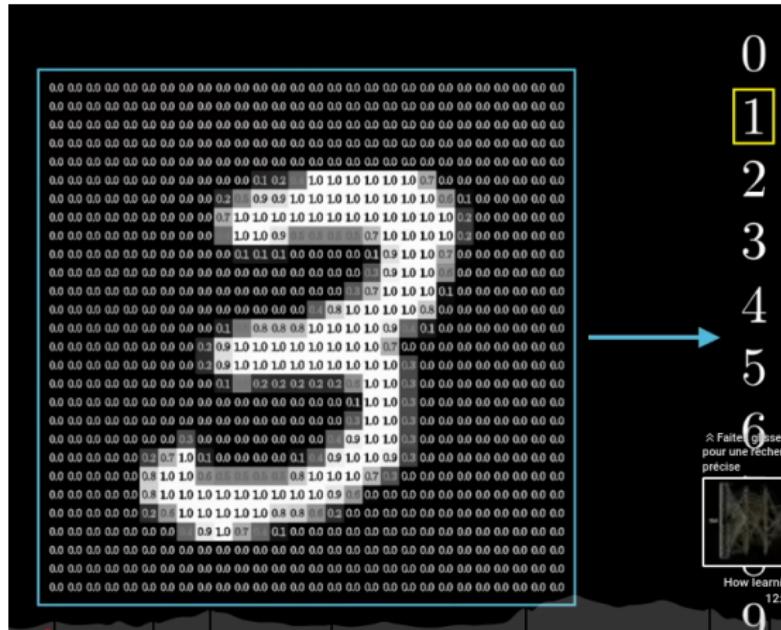
5. Début du XXIème siècle

Le Big Data

Etat de l'art

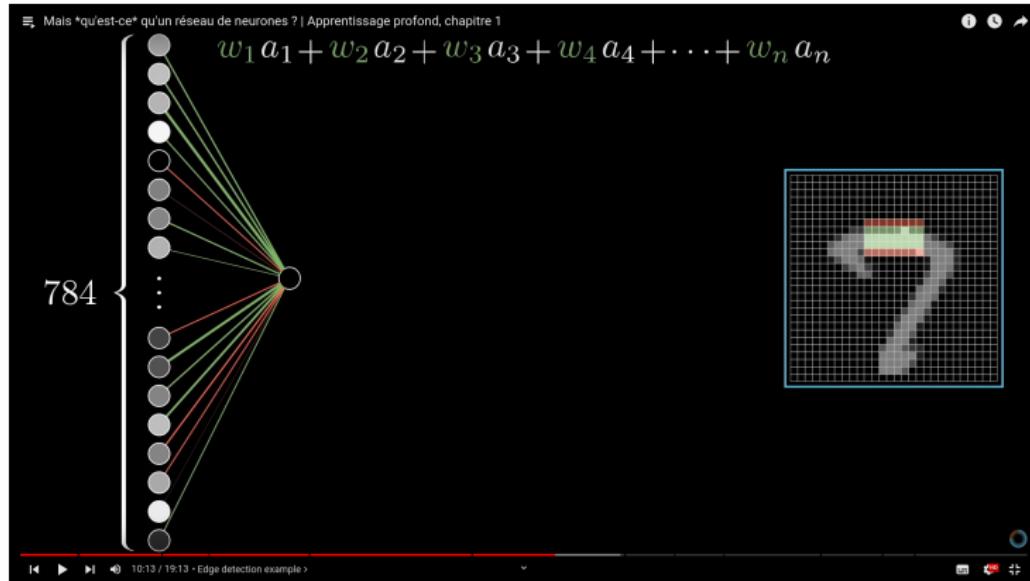
Les erreurs et les biais

Réseaux multicouches : une image



- ▶ Une image est un tableau de chiffre

Réseaux multicouches : une couche de neurones



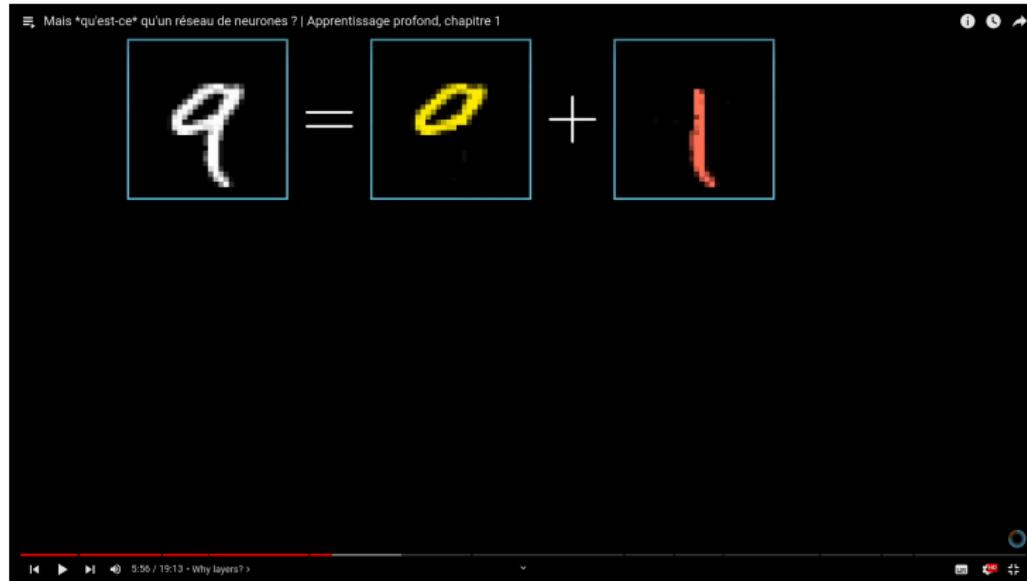
- ▶ Les neurones reconnaissent des parties de l'image

Reconnaitre les caractéristiques de l'image



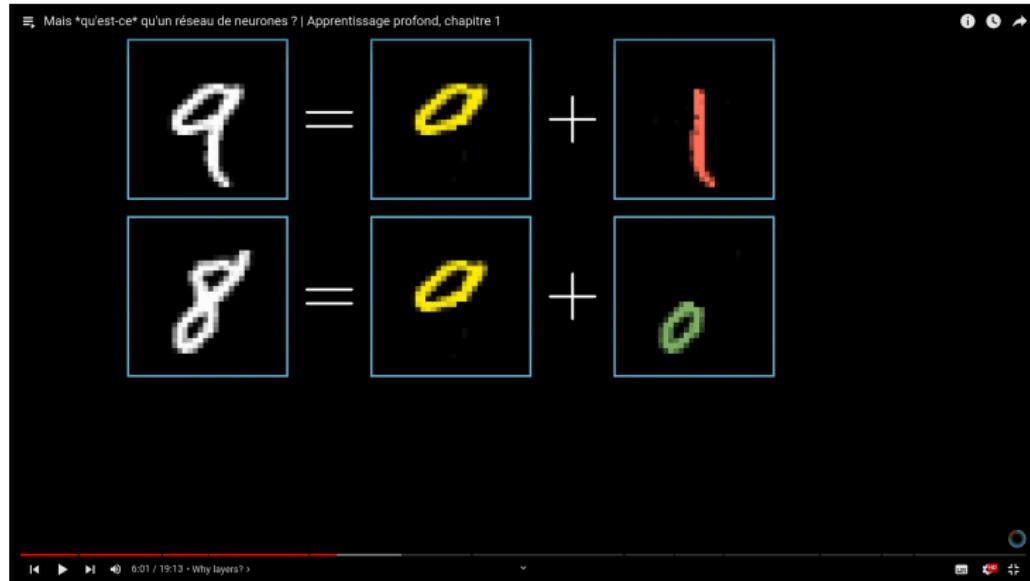
- ▶ Les chiffres sont reconnus partie par partie, pixel par pixel

Reconnaitre les caractéristiques de l'image



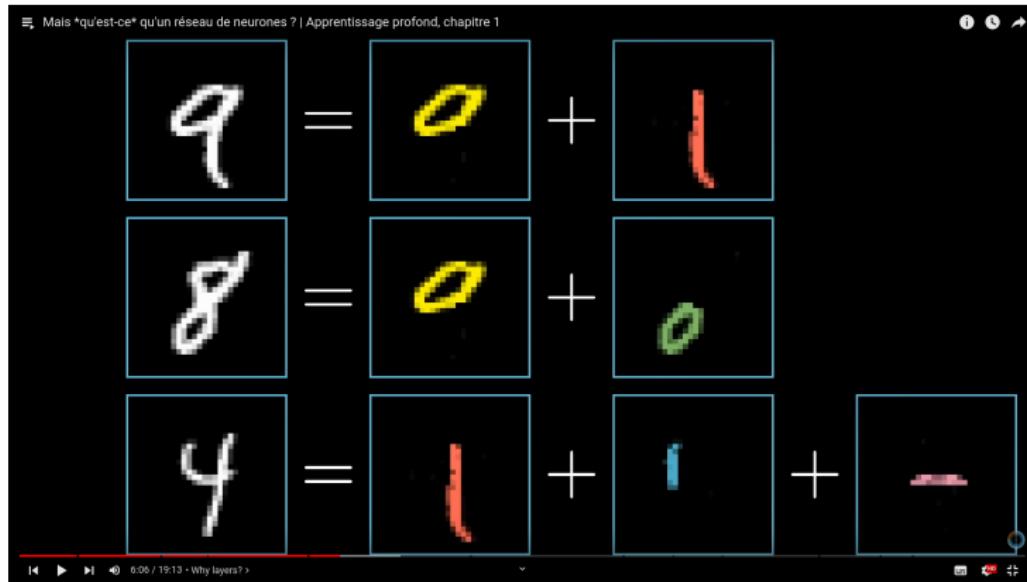
- ▶ Les chiffres sont reconnus partie par partie, pixel par pixel

Reconnaitre les caractéristiques de l'image



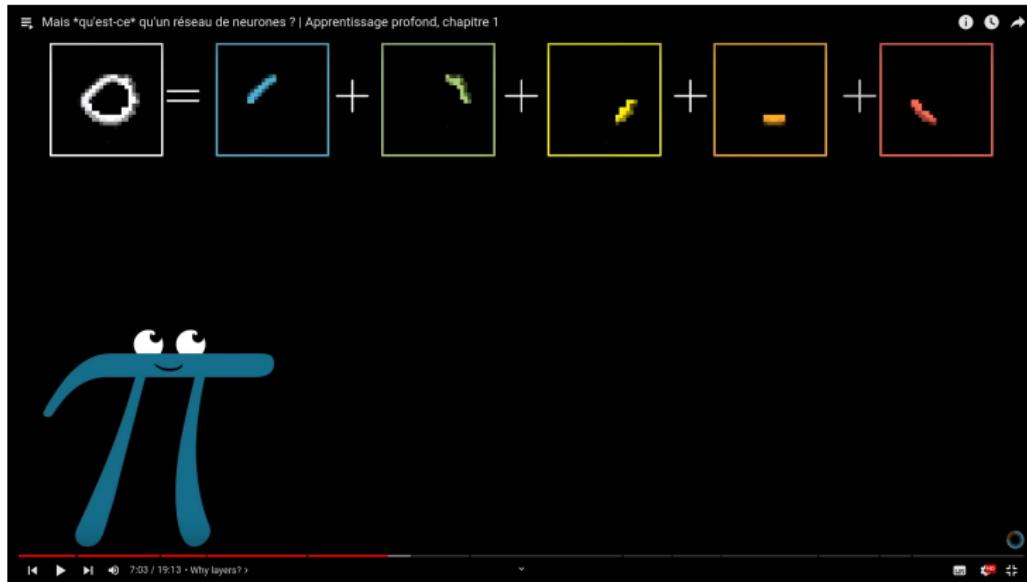
- ▶ Les chiffres sont reconnus partie par partie, pixel par pixel

Reconnaitre les caractéristiques de l'image



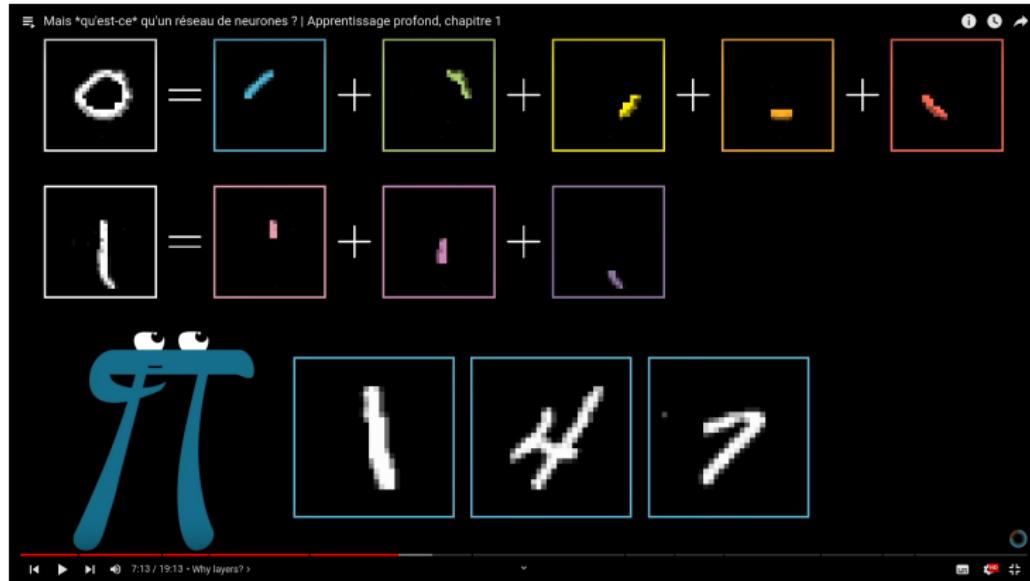
- ▶ Les chiffres sont reconnus partie par partie, pixel par pixel

Reconnaitre les caractéristiques de l'image



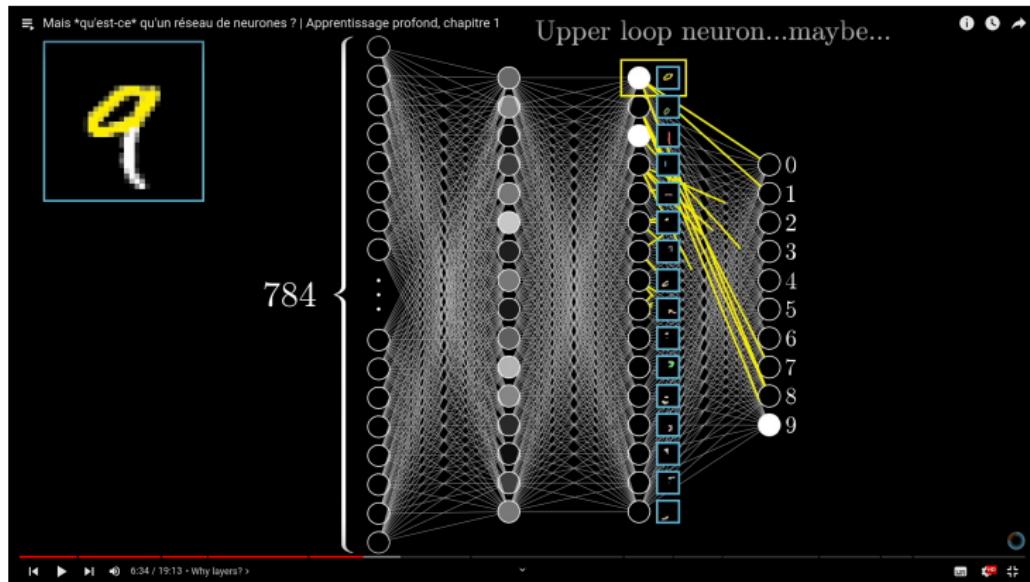
- ▶ Les chiffres sont reconnus partie par partie, pixel par pixel

Reconnaitre les caractéristiques de l'image



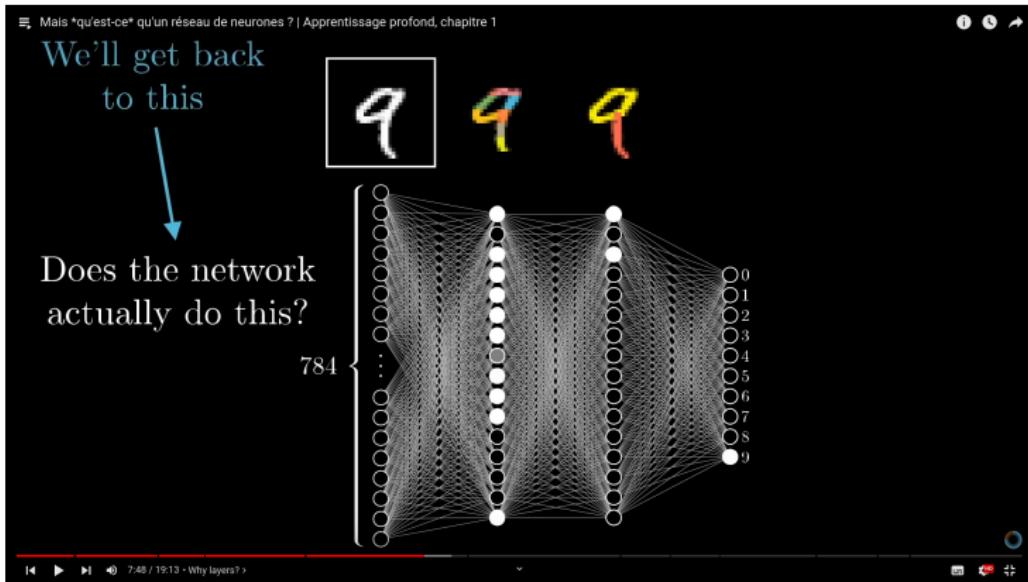
- ▶ Les chiffres sont reconnus partie par partie, pixel par pixel

Reconnaitre les caractéristiques de l'image



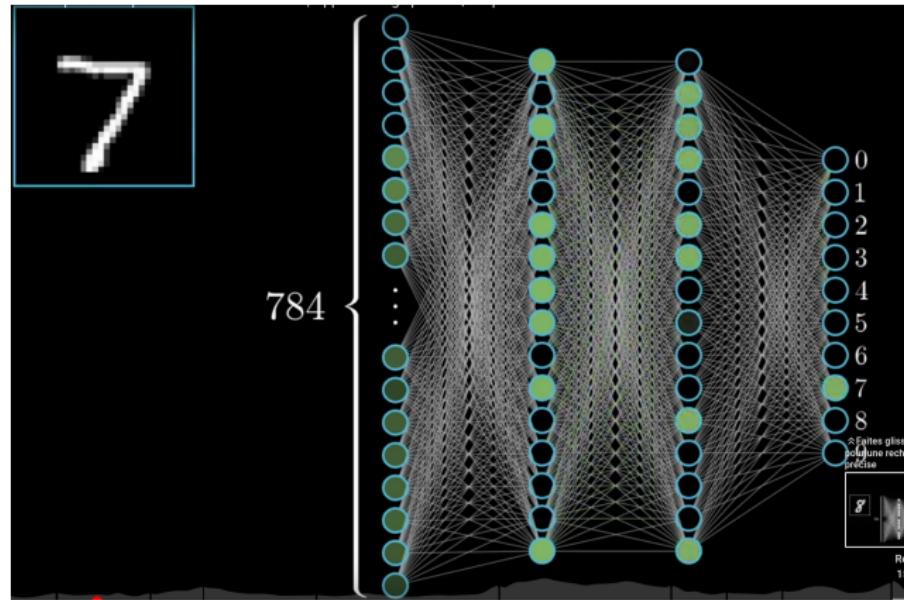
- L'assemblage de couches -> reconnaissance forme complexes

Reconnaitre les caractéristiques de l'image



- ▶ L'assemblage de couches -> reconnaissance forme complexes

Reconnaitre les caractéristiques de l'image



- ▶ L'assemblage de couches -> reconnaissance forme complexes

Sommaire

1. Préambule

Objectif de l'informatique

Les applications de l'IA

2. Avant le XXème siècle

Algorithmes et automates pour de justes partages

Machines et Algorithmes

Alan Turing

3. Milieu du XXème siècle

Les débuts de l'Intelligence Artificielle

L'apprentissage machine

4. Fin du XXème siècle

Réseaux de neurones

La rétro-propagation (1974 - 1986)

5. Début du XXIème siècle

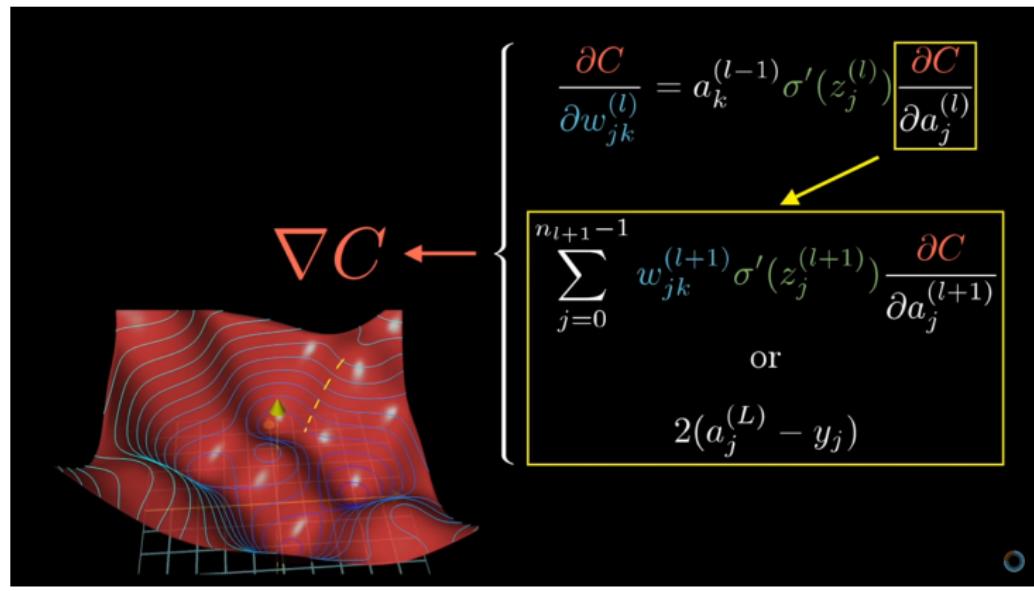
Le Big Data

Etat de l'art

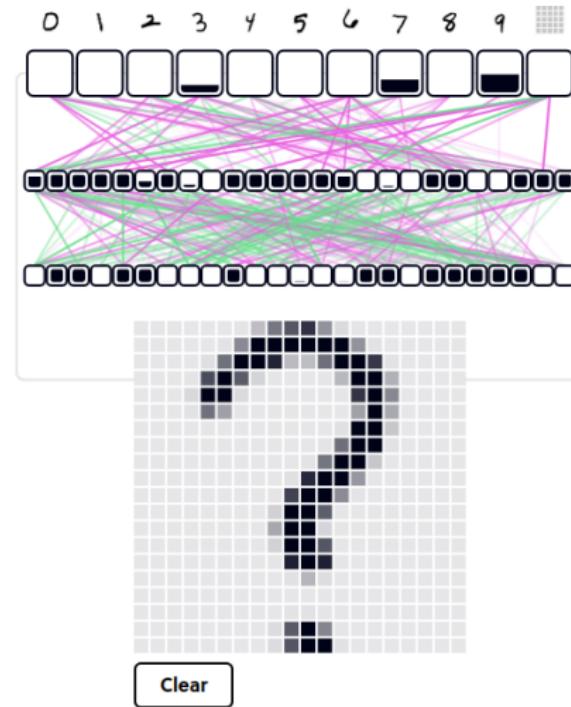
Les erreurs et les biais

La rétro-propagation (1974 - 1986)

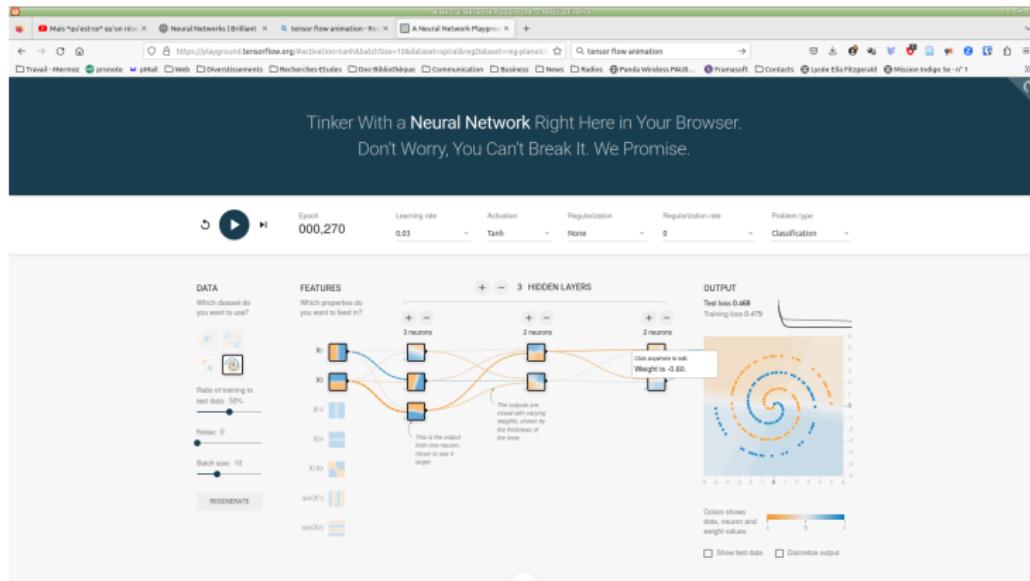
- ▶ Les "poids" ne sont plus définis a priori, mais appris automatiquement
- ▶ Les features (caractéristiques) s'apprennent seules...
- ▶ En résolvant une fonction avec **beaucoup** de variables



Quel est ce chiffre ?



Simulateur de réseau de neurone



Les débuts du succès : réseau convolutionnel 1989

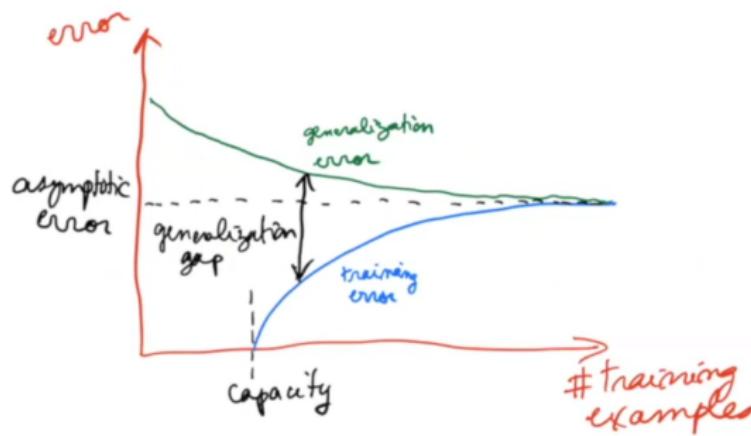


Limites : la malédiction de la dimensionnalité

- ▶ chaque couche ajoutée accroît de beaucoup :
 - ▶ la complexité des calculs
 - ▶ et le nombre de caractéristiques apprissent et à manipuler
- ▶ Plus il y a de caractéristiques à apprendre
 - ▶ Plus il devient difficile de généraliser.

Limites : la malédiction de la dimensionnalité

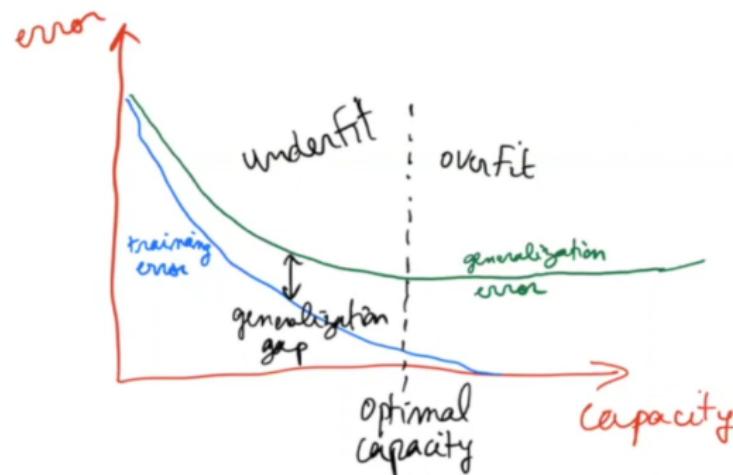
- ▶ chaque couche ajoutée accroît de beaucoup :
 - ▶ la complexité des calculs
 - ▶ et le nombre de caractéristiques apprises et à manipuler



- ▶ Plus il y a de caractéristiques à apprendre
 - ▶ Plus il devient difficile de généraliser.

Limites : la malédiction de la dimensionnalité

- ▶ chaque couche ajoutée accroît de beaucoup :
 - ▶ la complexité des calculs
 - ▶ et le nombre de caractéristiques apprises et à manipuler



- ▶ Plus il y a de caractéristiques à apprendre
 - ▶ Plus il devient difficile de généraliser.

Sommaire

1. Préambule

Objectif de l'informatique

Les applications de l'IA

2. Avant le XXème siècle

Algorithmes et automates pour de justes partages

Machines et Algorithmes

Alan Turing

3. Milieu du XXème siècle

Les débuts de l'Intelligence Artificielle

L'apprentissage machine

4. Fin du XXème siècle

Réseaux de neurones

La rétro-propagation (1974 - 1986)

5. Début du XXIème siècle

Le Big Data

Etat de l'art

Les erreurs et les biais

BigData : La révolution d'ImagNet 2012

- ▶ Performance accrue,
 - ▶ AlexNet 15% "d'erreur" contre 25% l'année précédente
- ▶ Importance reconnue du big-data
- ▶ Émergence et la popularité accrue pour l'apprentissage profond.



Sommaire

1. Préambule

Objectif de l'informatique

Les applications de l'IA

2. Avant le XX^{ème} siècle

Algorithmes et automates pour de justes partages

Machines et Algorithmes

Alan Turing

3. Milieu du XX^{ème} siècle

Les débuts de l'Intelligence Artificielle

L'apprentissage machine

4. Fin du XX^{ème} siècle

Réseaux de neurones

La rétro-propagation (1974 - 1986)

5. Début du XXI^{ème} siècle

Le Big Data

Etat de l'art

Les erreurs et les biais

Techniques

- ▶ Architectures plus profondes et complexes avec transfert d'apprentissage

Applications

- ▶ Retouche d'images et Vidéos
- ▶ Génération de text, d'images et d'audio et de vidéos

Les données

- ▶ Chat GPT-3, 175 milliards de paramètres.
- ▶ ImageNet 14 millions d'images taguées
- ▶ Facebook 350 millions d'images par jour.
- ▶ twitter 10 millions de tweet par jour
- ▶ Youtube about 300 000 h de vidéo par jour

Techniques

- ▶ Architectures plus profondes et complexes avec transfert d'apprentissage

Applications

- ▶ Retouche d'images et Vidéos
- ▶ Génération de text, d'images et d'audio et de vidéos

Les données

- ▶ Chat GPT-3, 175 milliards de paramètres.
- ▶ ImageNet 14 millions d'images taguées
- ▶ Facebook 350 millions d'images par jour.
- ▶ twitter 10 millions de tweet par jour
- ▶ Youtube about 300 000 h de vidéo par jour

Techniques

- ▶ Architectures plus profondes et complexes avec transfert d'apprentissage

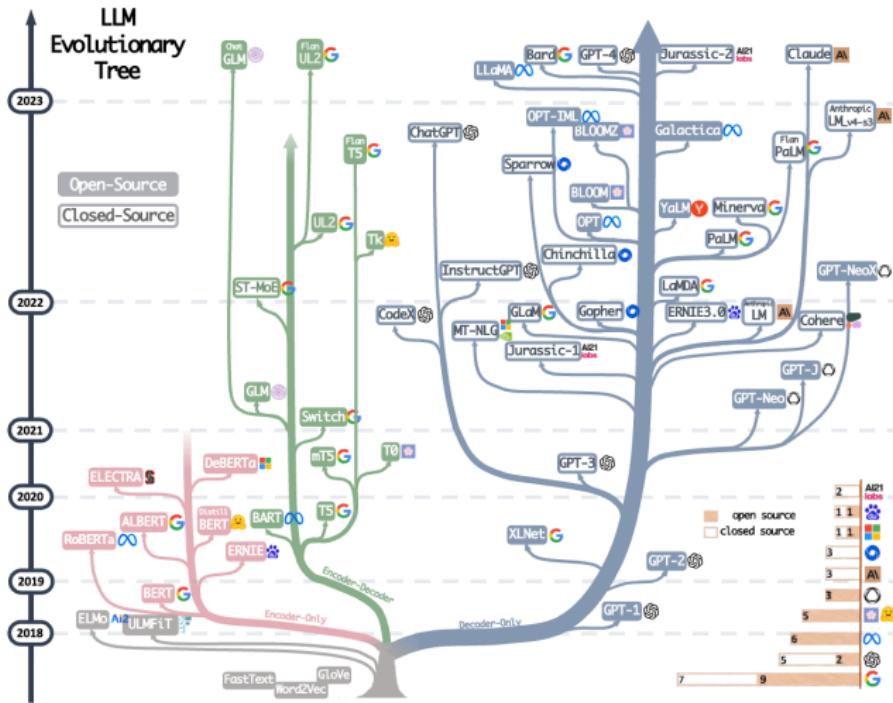
Applications

- ▶ Retouche d'images et Vidéos
- ▶ Génération de text, d'images et d'audio et de vidéos

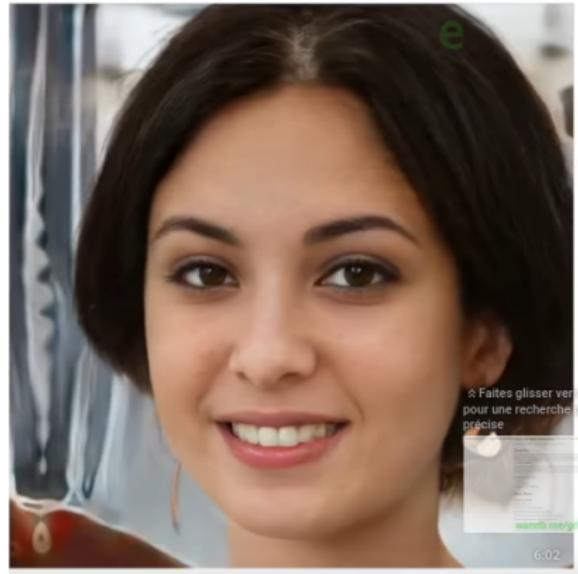
Les données

- ▶ Chat GPT-3, 175 milliards de paramètres.
- ▶ ImageNet 14 millions d'images taguées
- ▶ Facebook 350 millions d'images par jour.
- ▶ twitter 10 millions de tweet par jour
- ▶ Youtube about 300 000 h de vidéo par jour

ChatGPT et similaires



Méta-humains, Nvidia, Unreal, Microsoft, Google...



Méta-humains, Nvidia, Unreal, Microsoft, Google. . .



Deep fake



Morgan Freeman

Deep fake



Real or Fake?

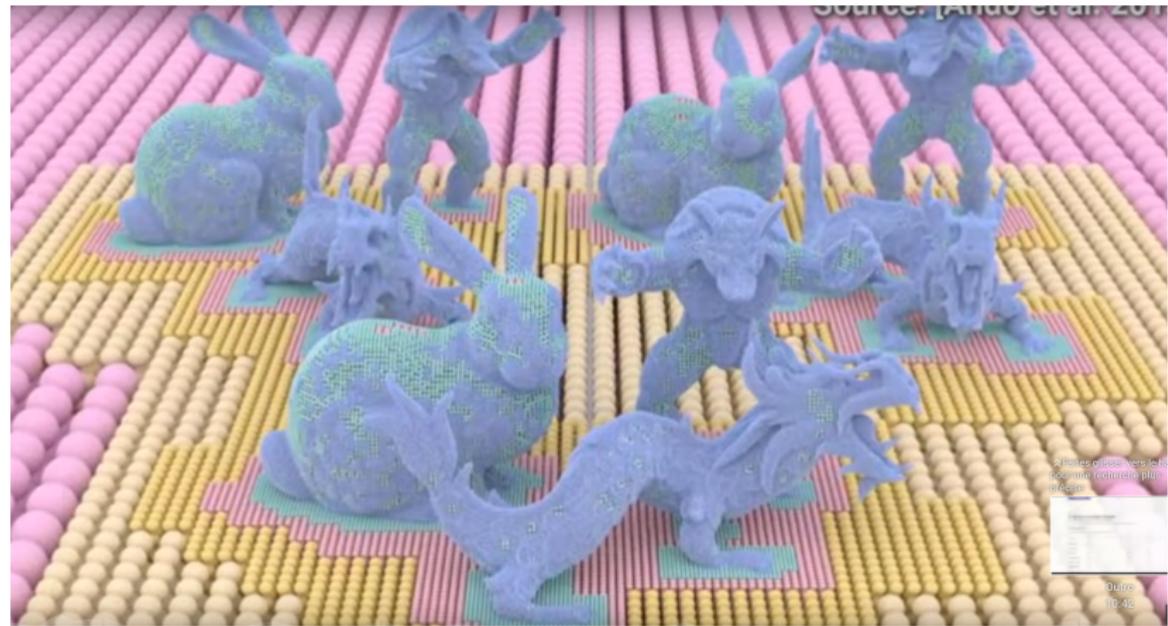


Adobe Photoshop AI generative fill firefly

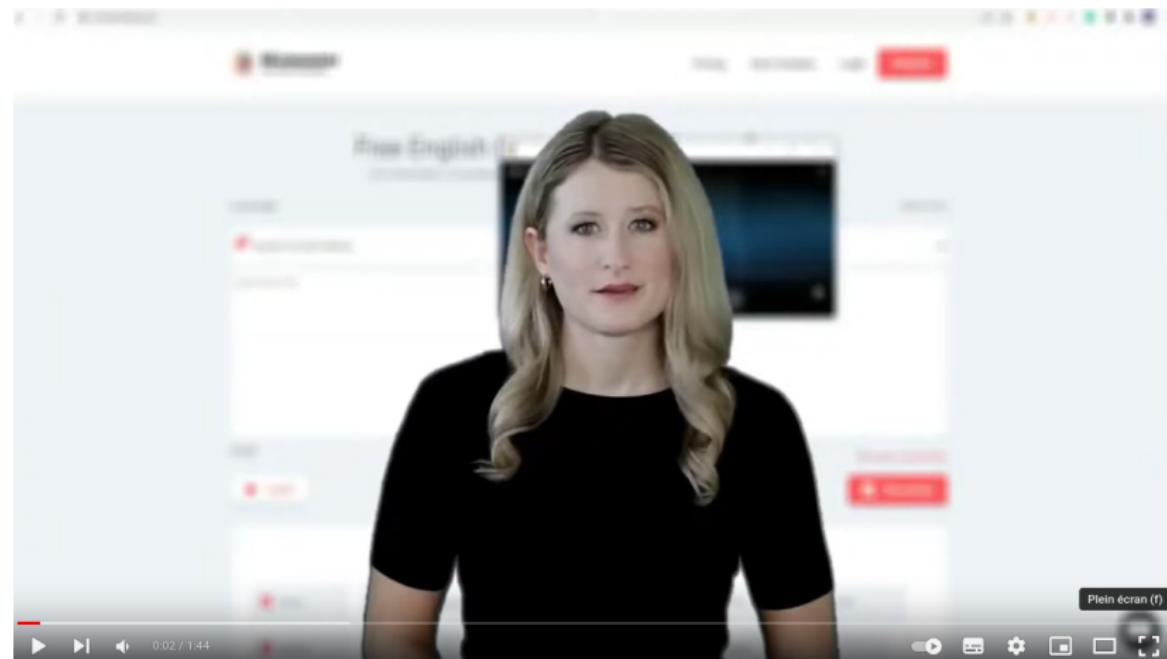


Où s'arrête le vrai ?

Mondes virtuels



Du Texte à la parole



Du texte à la vidéo



52

Sommaire

1. Préambule

Objectif de l'informatique

Les applications de l'IA

2. Avant le XX^{ème} siècle

Algorithmes et automates pour de justes partages

Machines et Algorithmes

Alan Turing

3. Milieu du XX^{ème} siècle

Les débuts de l'Intelligence Artificielle

L'apprentissage machine

4. Fin du XX^{ème} siècle

Réseaux de neurones

La rétro-propagation (1974 - 1986)

5. Début du XXI^{ème} siècle

Le Big Data

Etat de l'art

Les erreurs et les biais

Les erreurs et les biais

Matrice de confusion

	Positif en Vrai	Negatif en Vrai
Prédit Positif	Vrai Positif	Faux Positif
Prédit Negatif	Faux Negatif	Vrai Negatif

Les biais d'apprentissage

- ▶ dans le recrutement en faveur des hommes (Facebook, Amazon).
- ▶ dans la détection des composants sanguins à travers la peau en défaveur des peaux noires (Apple).
- ▶ dans la détection des visages
- ▶ biais raciaux et de genre (prêts bancaires, reconnaissance de la parole)

Les erreurs et les biais

Matrice de confusion

	Positif en Vrai	Negatif en Vrai
Prédit Positif	Vrai Positif	Faux Positif
Prédit Negatif	Faux Negatif	Vrai Negatif

Les biais d'apprentissage

- ▶ dans le recrutement en faveur des hommes (Facebook, Amazon).
- ▶ dans la détection des composants sanguins à travers la peau en défaveur des peaux noires (Apple).
- ▶ dans la détection des visages
- ▶ biais raciaux et de genre (prêts bancaires, reconnaissance de la parole)

Moratoire sur l'IA

"Pause Giant AI Experiments" : An Open Letter

- ▶ Les IA super humaine posent des risques majeurs à l'humanité
- ▶ Des IA de plus en plus puissantes sont déployées, mais sans contrôle fiable
- ▶ *Pause de 6 mois* sur l'entraînement des IA les plus puissants pour élaborer :
 - ▶ des protocoles de sécurité partagés et audités
 - ▶ des IA plus précises, robustes, interprétables, digne de confiance
 - ▶ des systèmes de gouvernance de l'IA.

L'IA peut permettre à l'humanité de prospérer, à condition de gérer prudemment son développement

Signé (entre autres) par le Prix Turing Yoshua Bengio

Moratoire sur l'IA

"Pause Giant AI Experiments" : An Open Letter

- ▶ Les IA super humaine posent des risques majeurs à l'humanité
- ▶ Des IA de plus en plus puissantes sont déployées, mais sans contrôle fiable
- ▶ *Pause de 6 mois* sur l'entraînement des IA les plus puissants pour élaborer :
 - ▶ des protocoles de sécurité partagés et audités
 - ▶ des IA plus précises, robustes, interprétables, digne de confiance
 - ▶ des systèmes de gouvernance de l'IA.

L'IA peut permettre à l'humanité de prospérer, à condition de gérer prudemment son développement

Signé (entre autres) par le Prix Turing Yoshua Bengio

Moratoire sur l'IA

"Pause Giant AI Experiments" : An Open Letter

- ▶ Les IA super humaine posent des risques majeurs à l'humanité
- ▶ Des IA de plus en plus puissantes sont déployées, mais sans contrôle fiable
- ▶ *Pause de 6 mois* sur l'entraînement des IA les plus puissants pour élaborer :
 - ▶ des protocoles de sécurité partagés et audités
 - ▶ des IA plus précises, robustes, interprétables, digne de confiance
 - ▶ des systèmes de gouvernance de l'IA.

L'IA peut permettre à l'humanité de prospérer, à condition de gérer prudemment son développement

Signé (entre autres) par le Prix Turing Yoshua Bengio

Moratoire sur l'IA

"Pause Giant AI Experiments" : An Open Letter

- ▶ Les IA super humaine posent des risques majeurs à l'humanité
- ▶ Des IA de plus en plus puissantes sont déployées, mais sans contrôle fiable
- ▶ *Pause de 6 mois* sur l'entraînement des IA les plus puissants pour élaborer :
 - ▶ des protocoles de sécurité partagés et audités
 - ▶ des IA plus précises, robustes, interprétables, digne de confiance
 - ▶ des systèmes de gouvernance de l'IA.

L'IA peut permettre à l'humanité de prospérer, à condition de gérer prudemment son développement

Signé (entre autres) par le Prix Turing Yoshua Bengio

Moratoire sur l'IA

"Pause Giant AI Experiments" : An Open Letter

- ▶ Les IA super humaine posent des risques majeurs à l'humanité
- ▶ Des IA de plus en plus puissantes sont déployées, mais sans contrôle fiable
- ▶ *Pause de 6 mois* sur l'entraînement des IA les plus puissants pour élaborer :
 - ▶ des protocoles de sécurité partagés et audités
 - ▶ des IA plus précises, robustes, interprétables, digne de confiance
 - ▶ des systèmes de gouvernance de l'IA.

L'IA peut permettre à l'humanité de prospérer, à condition de gérer prudemment son développement

Signé (entre autres) par le Prix Turing Yoshua Bengio

IA super humaine

Elle sera bientôt là

- ▶ AlphaGo, qui a battu un joueur professionnel de Go en 2016.
- ▶ AlphaFold de DeepMind, qui a réussi à prédire la structure des protéines avec une précision sans précédent.
- ▶ Five d'OpenAI, qui est une équipe de cinq agents IA capables de battre des joueurs professionnels de Dota 2.
- ▶ Waymo, qui développe des voitures autonomes qui ont déjà parcouru des millions de kilomètres sur les routes publiques.

IA super humaine

Elle sera bientôt là

- ▶ AlphaGo, qui a battu un joueur professionnel de Go en 2016.
- ▶ AlphaFold de DeepMind, qui a réussi à prédire la structure des protéines avec une précision sans précédent.
- ▶ Five d'OpenAI, qui est une équipe de cinq agents IA capables de battre des joueurs professionnels de Dota 2.
- ▶ Waymo, qui développe des voitures autonomes qui ont déjà parcouru des millions de kilomètres sur les routes publiques.

IA super humaine

Elle sera bientôt là

- ▶ AlphaGo, qui a battu un joueur professionnel de Go en 2016.
- ▶ AlphaFold de DeepMind, qui a réussi à prédire la structure des protéines avec une précision sans précédent.
- ▶ Five d'OpenAI, qui est une équipe de cinq agents IA capables de battre des joueurs professionnels de Dota 2.
- ▶ Waymo, qui développe des voitures autonomes qui ont déjà parcouru des millions de kilomètres sur les routes publiques.

IA super humaine

Elle sera bientôt là

- ▶ AlphaGo, qui a battu un joueur professionnel de Go en 2016.
- ▶ AlphaFold de DeepMind, qui a réussi à prédire la structure des protéines avec une précision sans précédent.
- ▶ Five d'OpenAI, qui est une équipe de cinq agents IA capables de battre des joueurs professionnels de Dota 2.
- ▶ Waymo, qui développe des voitures autonomes qui ont déjà parcouru des millions de kilomètres sur les routes publiques.



Mardi 18 juillet 2023

L'intelligence Artificielle pour les Juristes

Merci ! Continuons les échanges ...

Issa Traoré (PhD)
Malik Koné (PhD)