

Intelligence Artificielle

Boi Faltings
EPFL

L'Intelligence Artificielle: qu'est-ce que c'est?

- Comportement rationnel
- Apprentissage à partir des données
- Sens commun

Comportement rationnel

- Les algorithmes font toujours la même chose.
- L'IA fait ce qu'il faut:
 - selon le contexte.
 - optimisé selon l'objectif actuel.
- Programmation => définition d'objectifs.

Exemple: planification

- Planifier les mouvements au sol des avions.
- Objectifs: horaire, éviter les blocages et conflits.



- Meilleur plan dépend du contexte: très difficile d'écrire un algorithme.
- Un des premiers grand succès de l'IA.

Machine learning

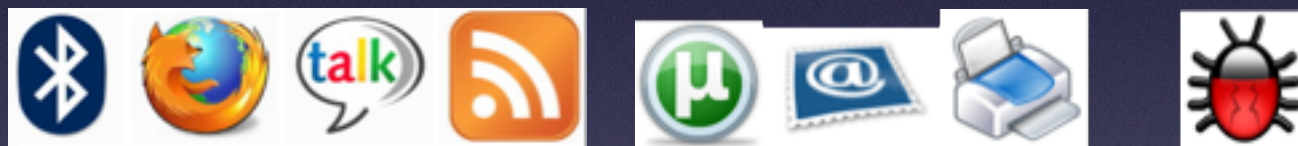
- Programmation: règles générales pour atteindre les objectifs: $z = x + y$
- Apprentissage: adapter un modèle à un ensemble d'exemples:

x	1	4	3	7	2	8	1
y	3	2	8	9	3	7	2
z	4	6	11	16	5	15	3

- meilleur modèle: $z = x + y$

NexThink (2003)

- Quelles applications et services sont installés sur les machines client?

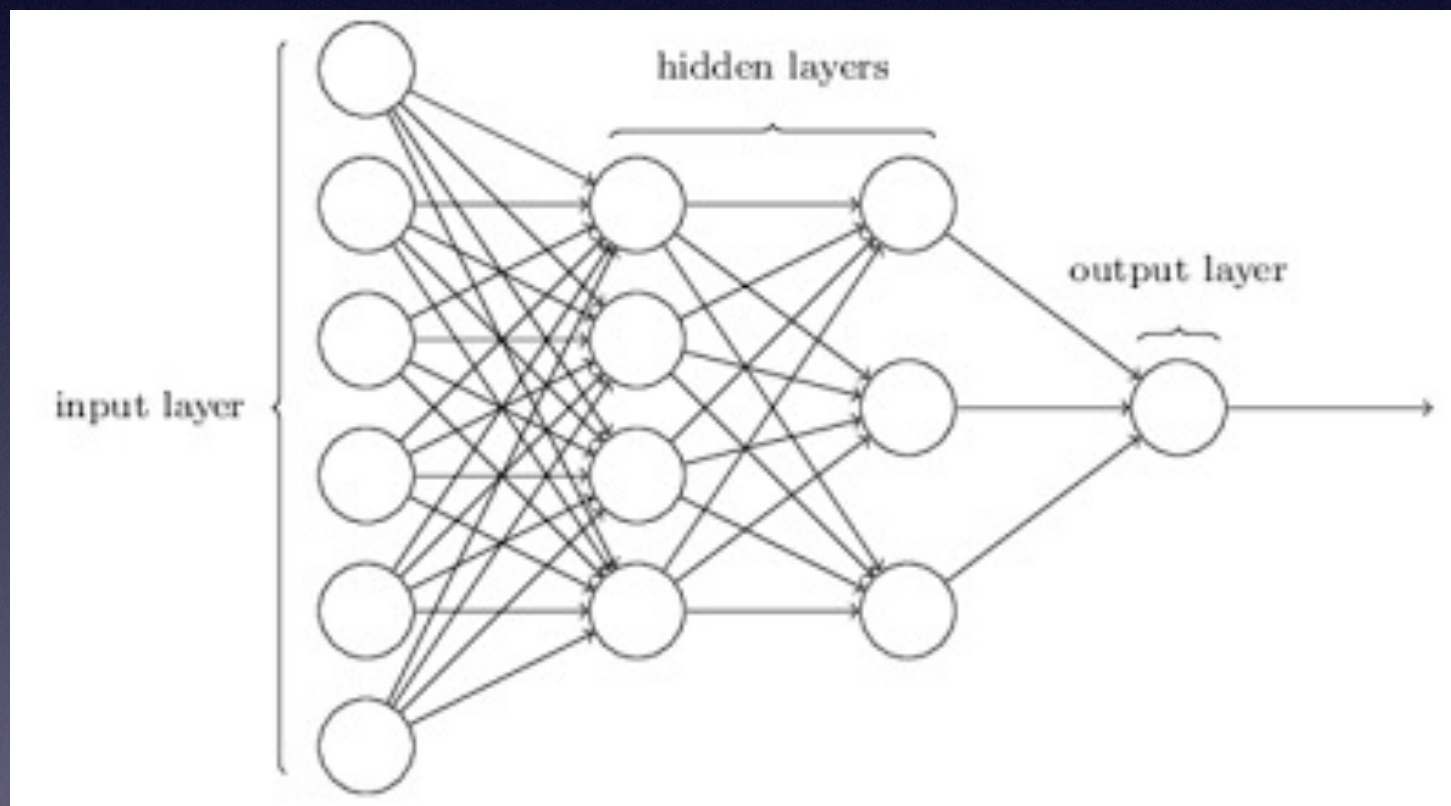


- Comment sont-ils utilisés?
- Inférer le modèle à partir des données de trafic.



Réseaux de neurones artificiels

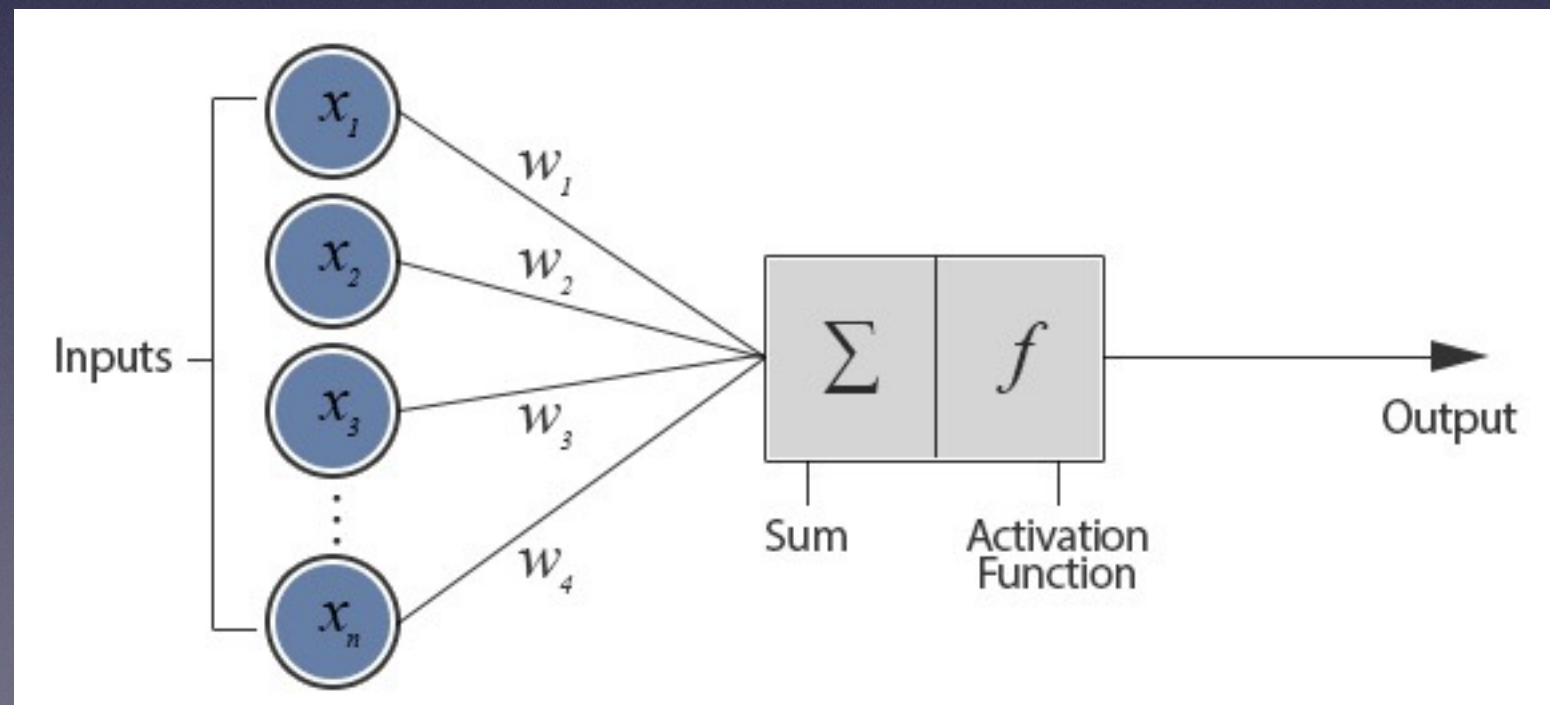
- Le cerveau est un réseau de milliards de neurones interconnectés.



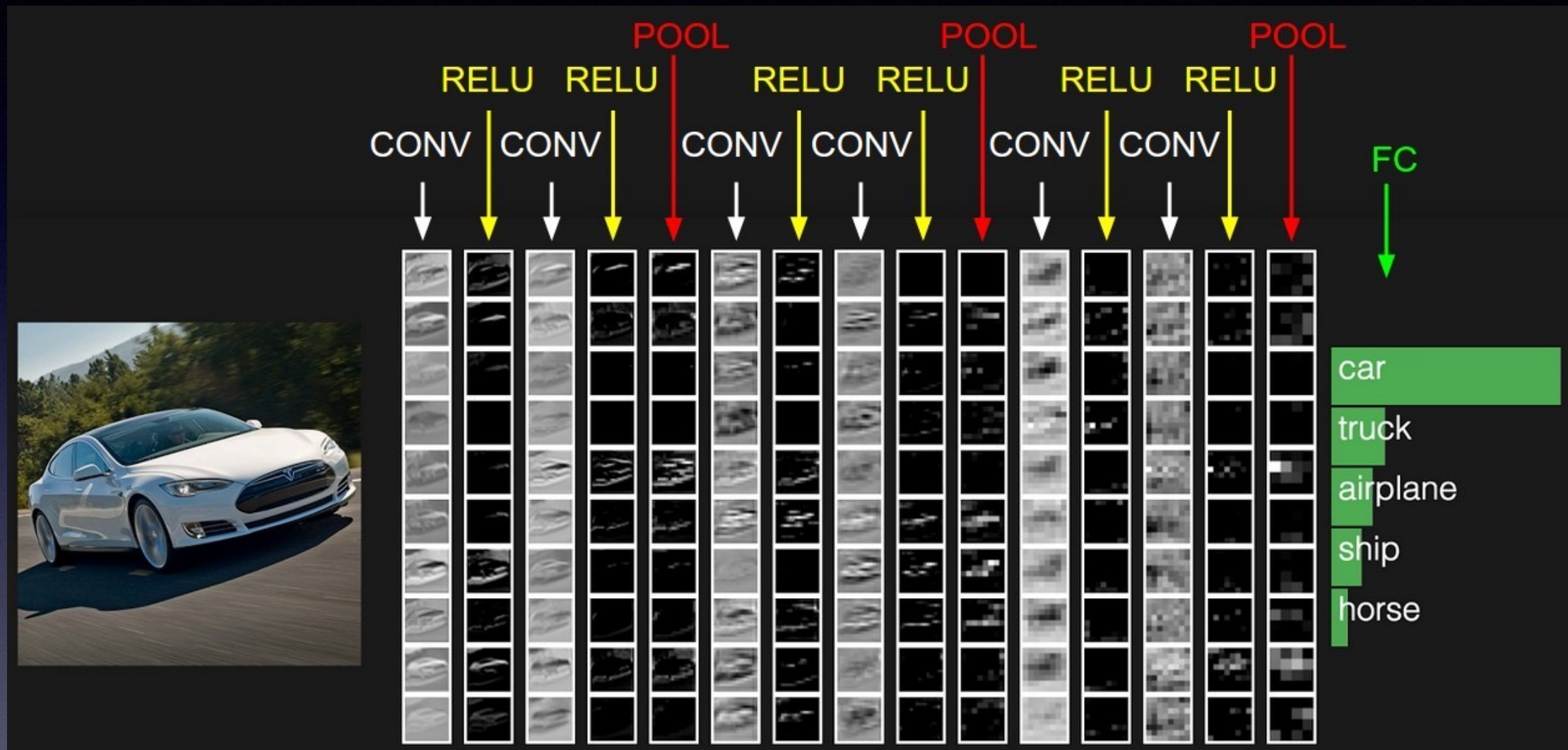
- Imiter dans un ordinateur, en particulier pour l'analyse d'image et de la parole.

Neuron artificiel

- activation = somme pondérée des entrées.
- sortie = fonction non linéaire f appliquée à l'activation: seuil / simoïde / redresseur.



Deep Learning



- <http://cs231n.github.io/assets/cnn/convnet.jpeg>
- Le réseau apprend la cascade de traits.

Apprendre à agir

- Observation de la récompense des actions.
- $\Rightarrow$ meilleures actions pour chaque contexte.
- Compromis entre:
 - Exploitation: utiliser les actions qui ont bien fonctionné dans le passé.
 - Exploration: essayer des actions avec une récompense inconnue.

Prediggo

- Observer la séquence des visites de pages et des ventes.
- Apprendre à prévoir le prochain achat probable.
- Recommandations conditionnées sur l'historique.



Sens commun

- L'intelligence repose sur des connaissances de sens commun.
- Ex: signification de *long* dépend du contexte:

Fortunately the battery life is **long** .

This camera takes a **long** time to focus.

- Aucun principe => difficile à modéliser.

Crowdsourcing

- Sens commun: tout le monde l'a.
- Recruter des travailleurs via Internet (*Mechanical Turk* et autres plates-formes)
- Chaque travailleur fournit une partie distincte.
- Travailleurs multiples par tâche assure qu'on obtienne vraiment le sens commun.

Modèles de langage

- Objectif: apprendre une langue sur la base de documents écrits.
- Difficulté: comment connaître la signification d'un mot/d'une phrase?
- Plus simple: prédire le prochain mot:

Aujourd'hui, il fait ...

Aujourd'hui, il fait ... et chaud.

Sens Commun

- Entraînement sur énormément de données (GPT4: bibliothèque de 50'000 km)
- => le modèle englobe toutes les connaissances du text
- Entraînement par renforcement pour aligner les réponses (p.ex. suivre des instructions ou répondre à des questions).

L'IA générale

- Modèle fondamental: entraîné sur des Terrabytes de données (bibliothèque de 50'000km).
- Un modèle fondamental peut fournir la base de sens commun pour toute sorte de tâche: prédiction de maladies, rédaction de documents, création d'images, etc.

Problème résolu?

Encore ouvert:

- Causalité
- Apprentissage continu
- Planification

L'optimisation est l'essentiel

- Comportement rationnel: choisir l'action qui optimise critère de performance.
- Apprendre des données: choisir le modèle qui correspond le mieux aux données.
- Apprendre à agir: choisir le modèle des actions avec les meilleures récompenses.
- Conséquence: imprévisibilité

Responsabilité légale

- L'action optimale est imprévisible.
- Souvent même inexplicable.
- Le programmeur est tenu responsable => de nombreux systèmes ne sont jamais utilisés.

Dilemmes éthiques

- Certaines situations nécessitent des choix:
 - écraser cette personne, ou
 - percuter une autre voiture?
- Décision préprogrammée toujours inacceptable pour une des parties.



Un contraste saisissant



- Analyse de trafic explicite: véhicules, bicyclettes, piétons.
- Prévoit les mouvements.
- => planifier la trajectoire.
- Garantie correct.
- Temps de développement: 40 ans.



- Un réseau neuronal de bout en bout associe les capteurs aux commandes de conduite.
- Apprend des données recueillies auprès de 100 000 conducteurs humains.
- Aucune garantie, peut ne jamais devenir correct.
- Temps de développement: 2 ans.

Méfiez-vous du hype

- 1957: "L'IA va battre les meilleurs joueurs d'échecs humains en 1967"
- 1997: IBM Deep Blue bat Kasparov.
- 2017: "Nous aurons des voitures autonomes dans les 5 ans." ?
- 2023: "ChatGPT va remplacer les créateurs de contenu." ?

Conclusions

Intelligence artificielle =

1. Action rationnelle.
2. Apprendre à partir de données.
3. Programmes avec sens commun.

Beaucoup de progrès ont été réalisés, mais l'IA reste un outil plutôt qu'un remplaçant.