



Fondements du Machine Learning

Premier cours:
INTRODUCTION

Mme BEDDAD Fatima
fatima.bedad@univ-temouchent.edu.dz

2025-2026

Au programme aujourd'hui

- **Faire connaissance.**
- **Présentation des objectifs, du plan de cours, des modalités d'évaluation,**
- **Présentation informelle du domaine de l'apprentissage automatique.**

PREMIÈRE PARTIE

Plan de Cours
et autres informations pratiques

Tonalité du cours

Semestre : 01

Intitulé de l'UE : Unité fondamentale

Intitulé de la matière : Fondements du machine Learning

Crédits : 6

Coefficients : 3

Objectifs de l'enseignement :

- . Passer en revue les outils Mathématiques requis pour le Machine Learning.
- . Apprendre la programmation avec le langage Python.
- . Préparer l'étudiant pour les modules : Machine Learning et Deep Learning.

Connaissances préalables recommandées

- . **Algorithmique** et **structures de données**
- . **Mathématiques** : Algèbre linéaire(matrice, vecteur, norme, produit scalaire, . . .), Analyse(dérivée, dérivée partielle, gradient, intégrale, . . .), Probabilités
- . **Du formalisme logique** : savoir lire une formule mathématique :

$$\forall x \in \mathcal{E} \cup \mathcal{F}, \exists Y \subset \mathcal{Y}, p(x) \in Y \Rightarrow \neg q(Y)$$

- . Le cours nécessite d'être assez à l'aise en **mathématiques** & **informatique**.
- . Le cours sera donné en **Français**, mais beaucoup de matériel est en **Anglais**(lectures, et matériel parfois utilisé comme support de cours).

Contenu de la matière

- 1. Mathématiques pour le Machine Learning:** Calcul matriciel - probabilité discrète,
- 2. Python:** Programmation et bibliothèques pour le machine Learning (NumPy, pandas, scikit-learn...)
- 3. Ingénierie des caractéristiques (Features) :** Sources de données (data sets) - extraction des caractéristiques - prétraitement et nettoyage de données (valeurs aberrantes, bruit et anomalies) - manipulation des données textuelles (Tokenisation, Bag-of-words, expression régulières, ...)
- 4. Visualisation des données.**

DEUXIÈME PARTIE

Présentation informelle du domaine de l'apprentissage automatique

Préambule

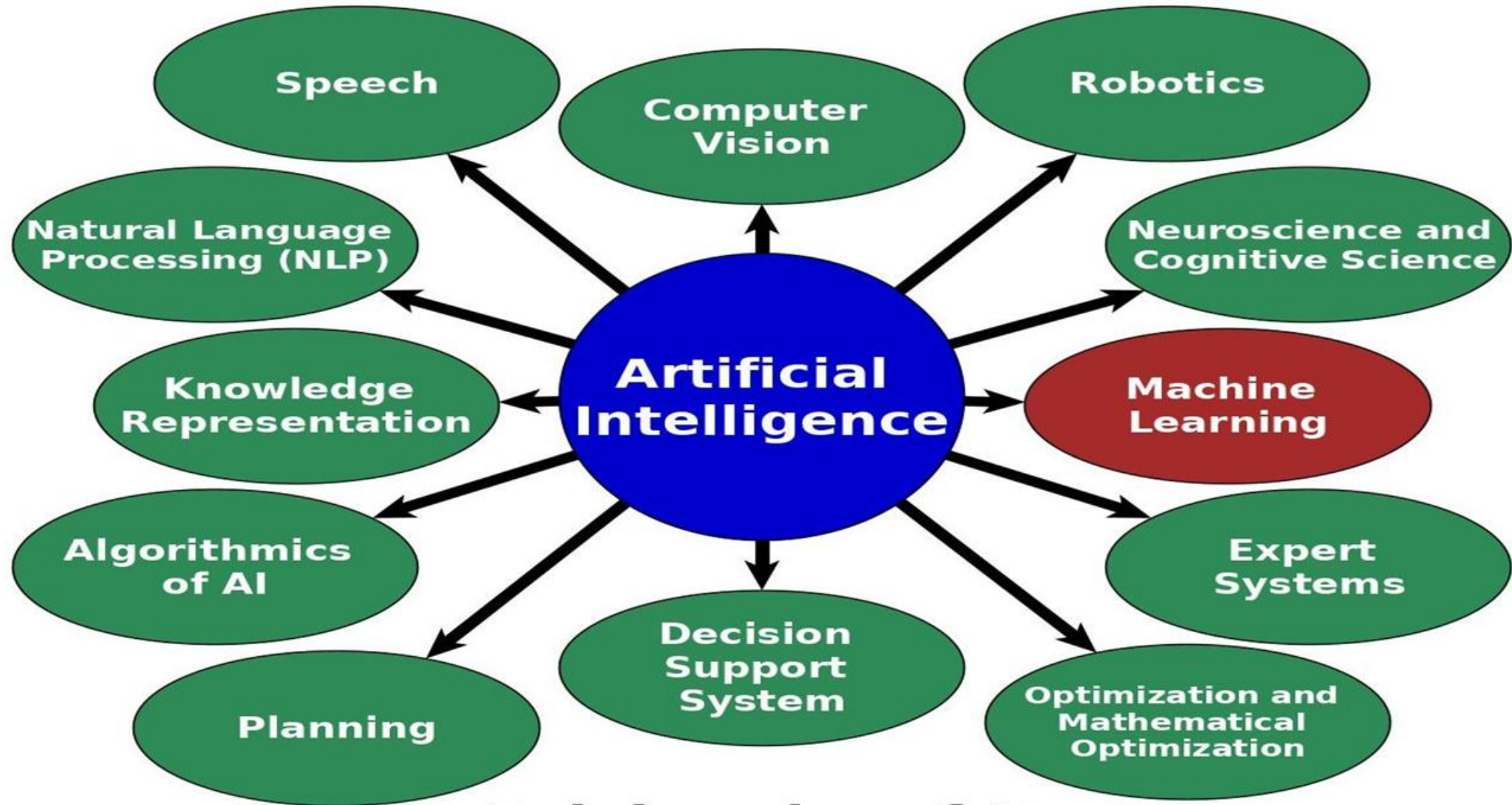
- De manière générale, un programme informatique tente de résoudre un problème pour lequel nous avons la solution. Par exemple : calculer la moyenne générale des étudiants, classer les étudiants selon leur moyenne. . .
- Pour certains problèmes, nous ne connaissons pas de solution exacte et donc nous ne pouvons pas écrire de programme informatique. Par exemple : reconnaître automatiquement des chiffres écrits à la main à partir d'une image scannée, déterminer automatiquement une typologie des clients d'une banque, jouer automatiquement aux échecs contre un humain ou un autre programme. . .
- En revanche, pour ces problèmes il est facile d'avoir une base de données regroupant de nombreuses instances du problème considéré.

D'une manière générale

L'apprentissage automatique consiste alors à programmer des algorithmes permettant d'apprendre automatiquement des données et d'expériences passées, un algorithme cherchant à résoudre au mieux un problème considéré.

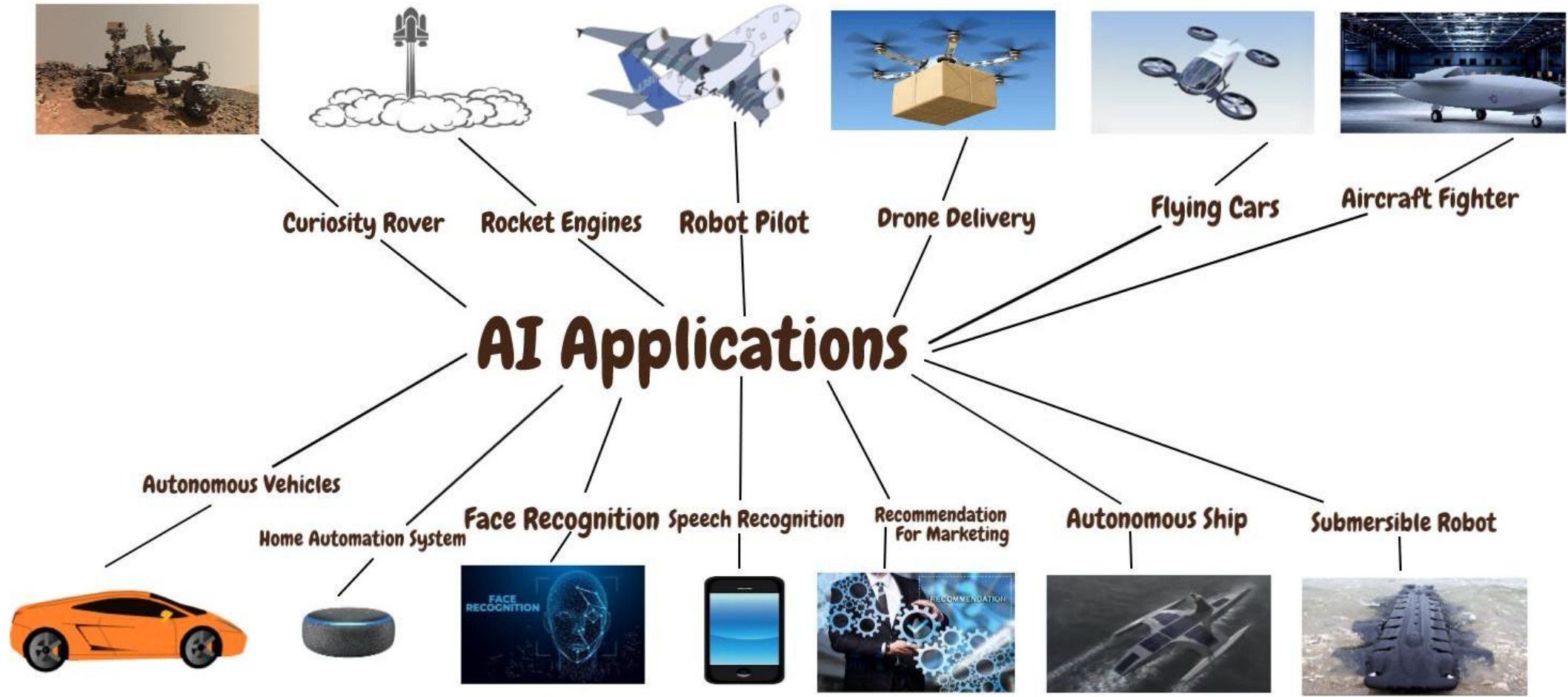
• « How can we build *computer systems* that automatically improve with experience, and what are the *fundamental laws* that govern *all learning processes*? » Tom Mitchell, 2006

Sous domaines de l'intelligence artificielle



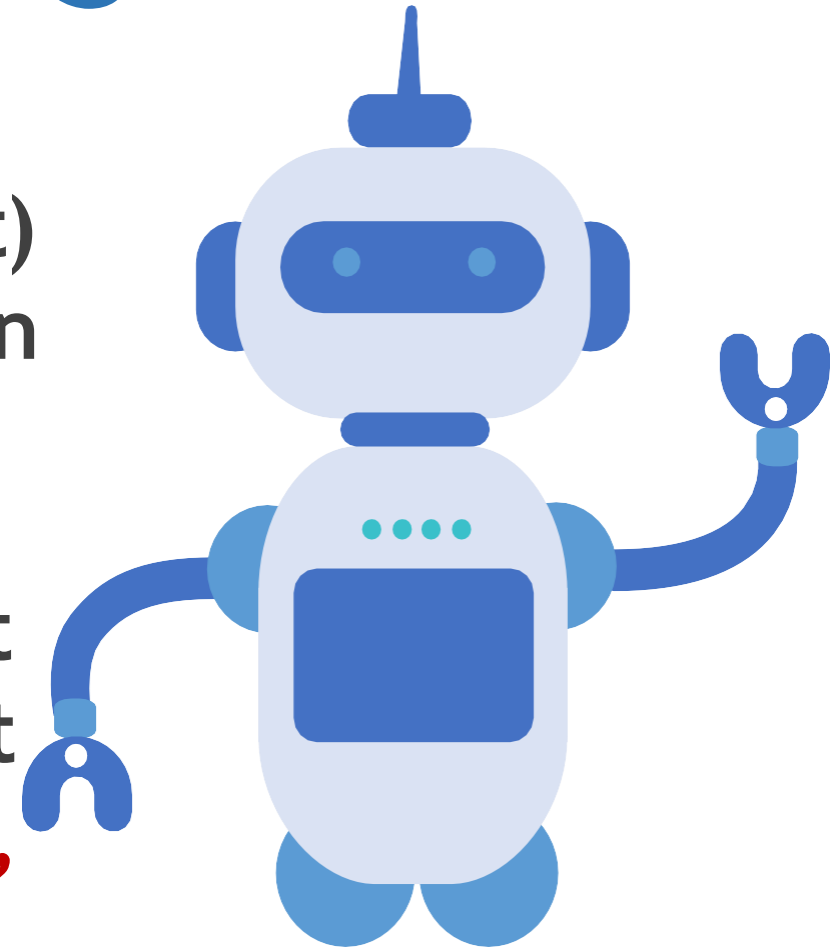
Subdomains of AI

Artificial Intelligence Applications



What Is Machine Learning?

- Machine Learning is the science (and art) of programming computers so they can learn from data.
- Machine Learning is the field of study that gives computers the ability to learn without being explicitly programmed. - **Arthur Samuel, 1959.**

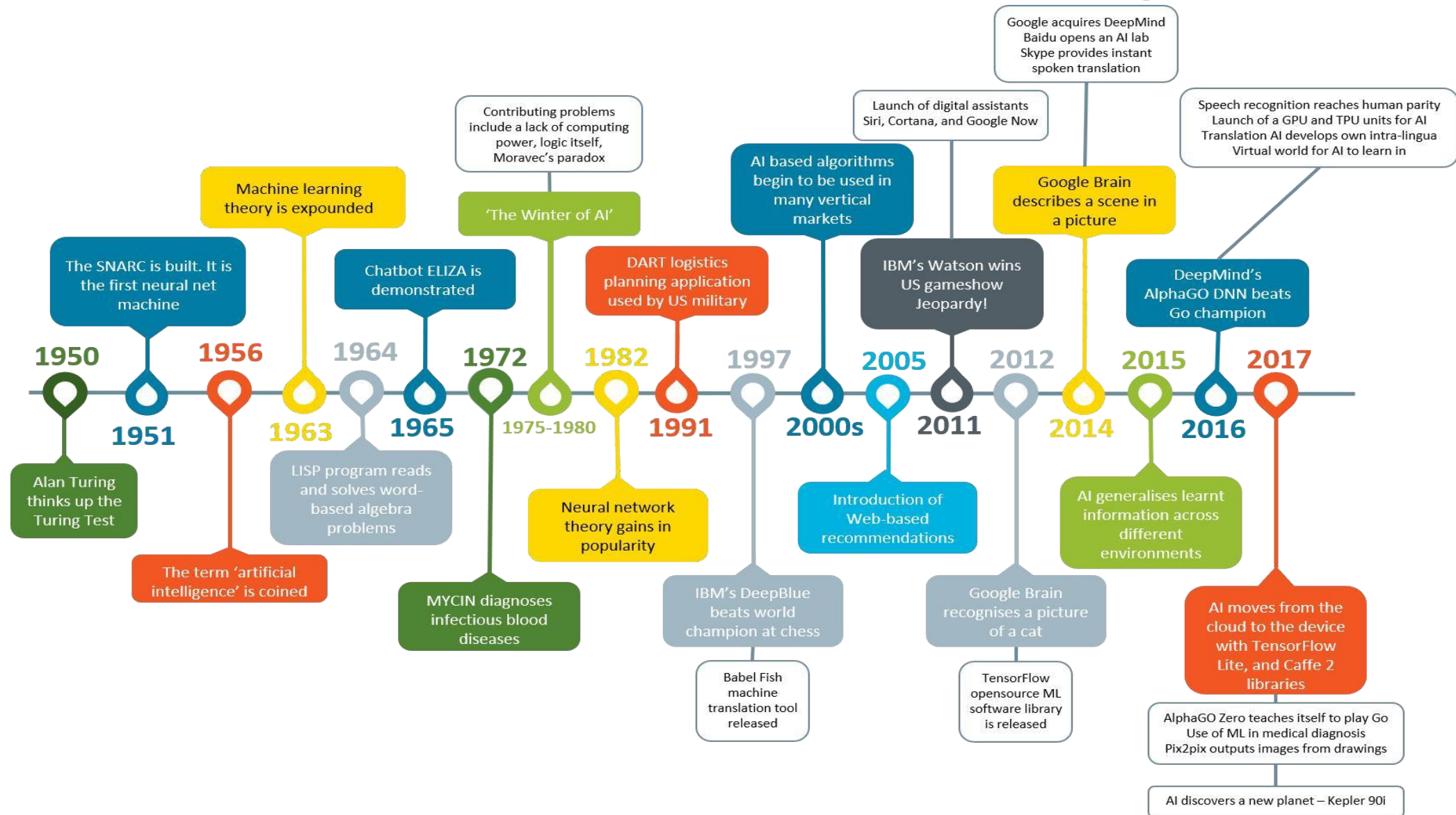


What Does Learning Mean?

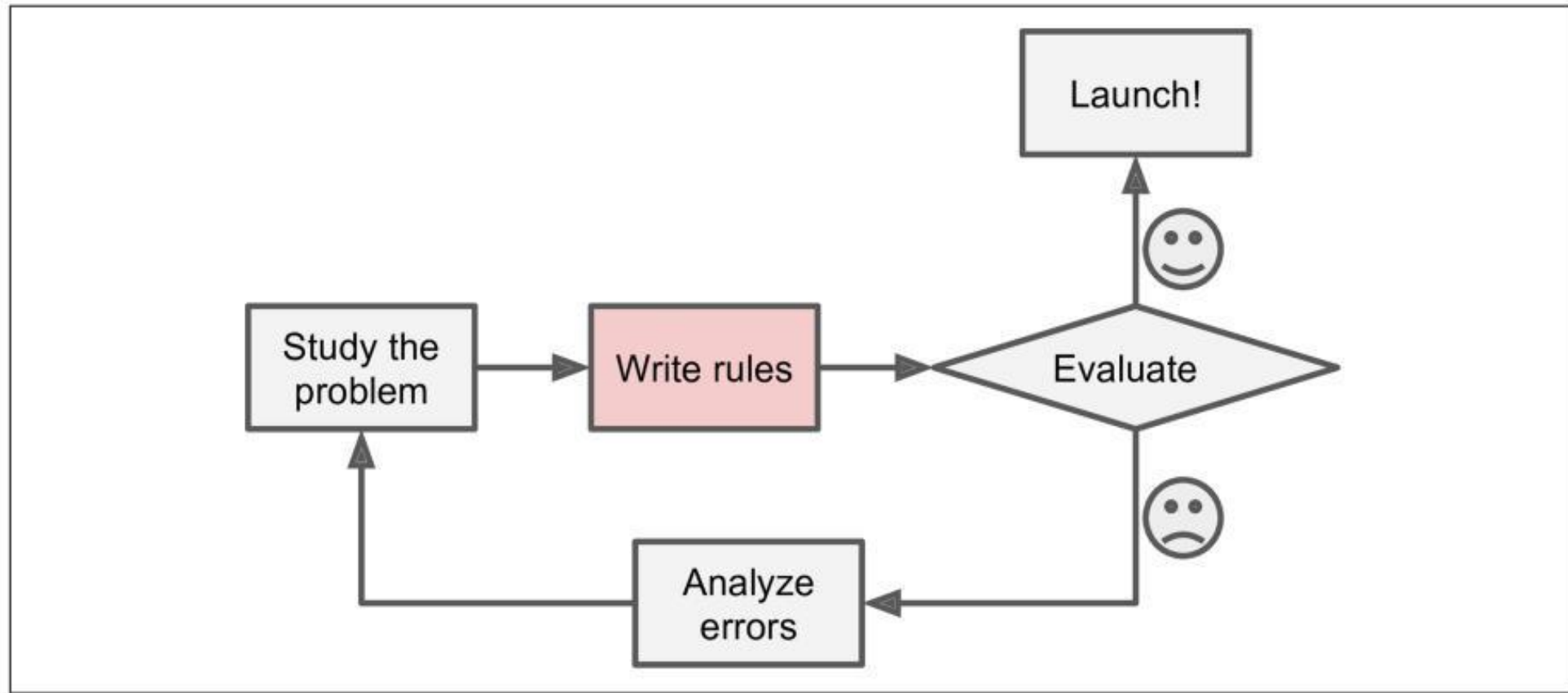
A computer program is said to learn from **experience E** with respect to some **task T** and some **performance measure P**, if its performance on **T**, as measured by **P**, improves with experience **E**. — **Tom Mitchell, 1997**



Timeline of Machine Learning

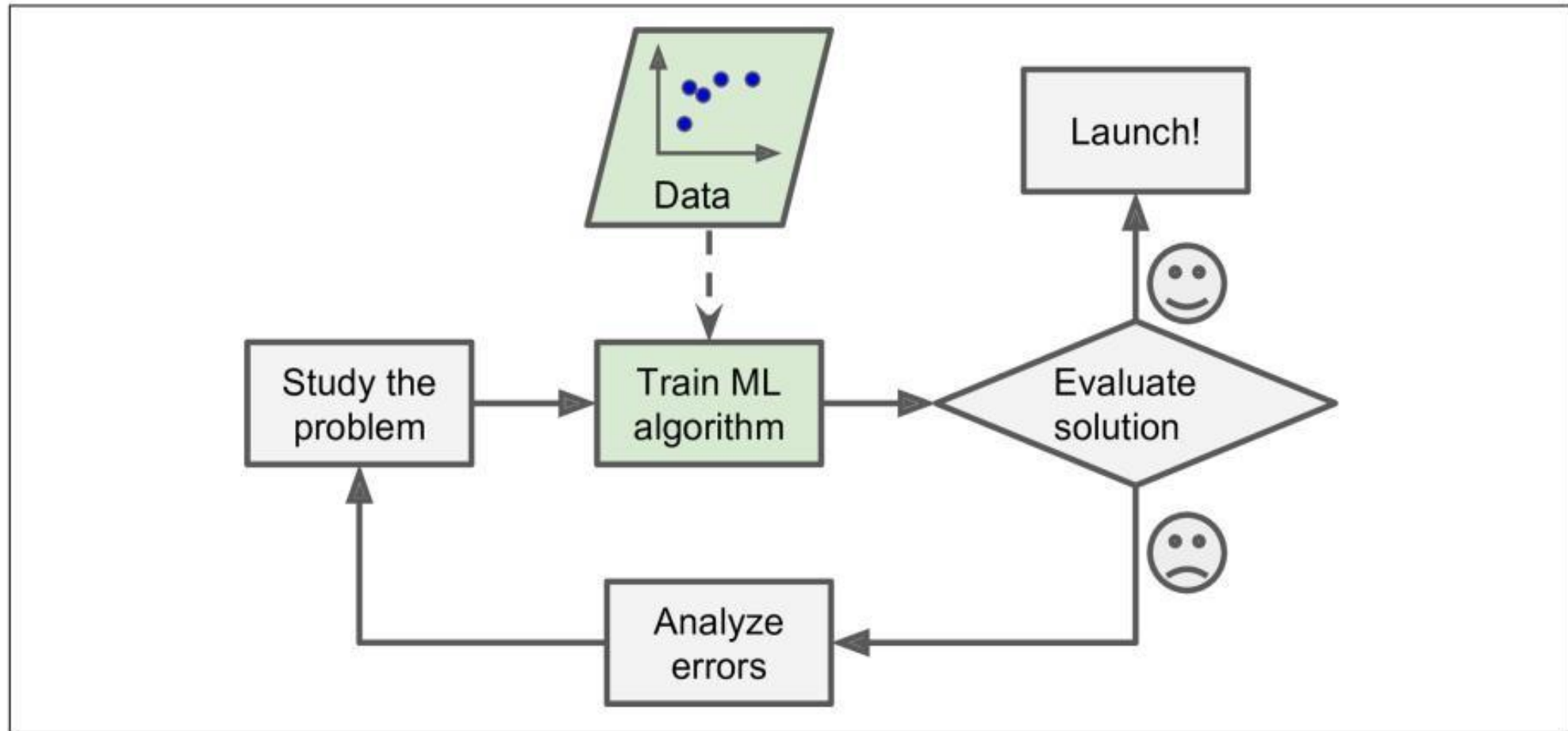


Why Use Machine Learning?



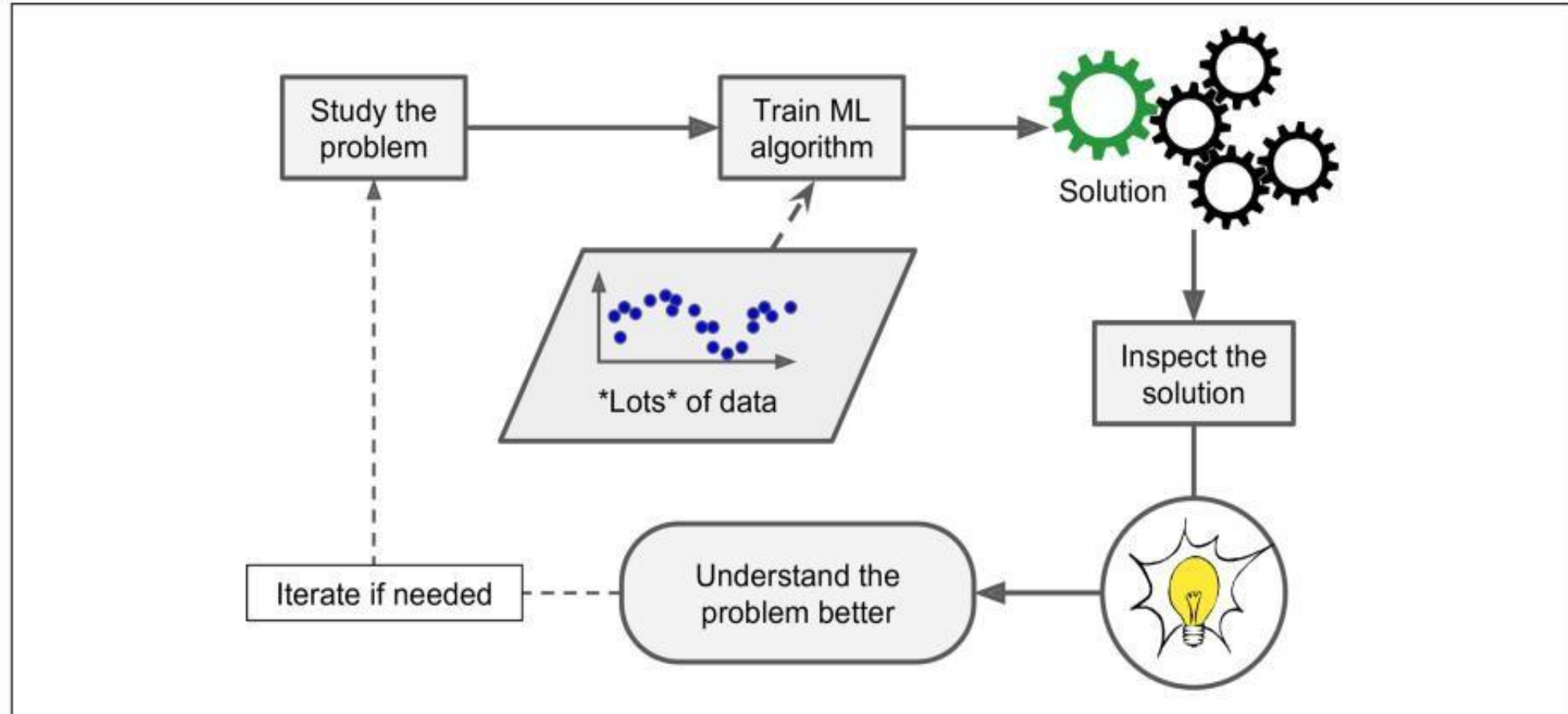
The traditional approach. *If the problem is not trivial, your program will likely become a long list of complex rules pretty hard to maintain.*

Why Use Machine Learning?



Machine Learning approach. *The program is much shorter, easier to maintain, and most likely more accurate.*

Why Use Machine Learning?

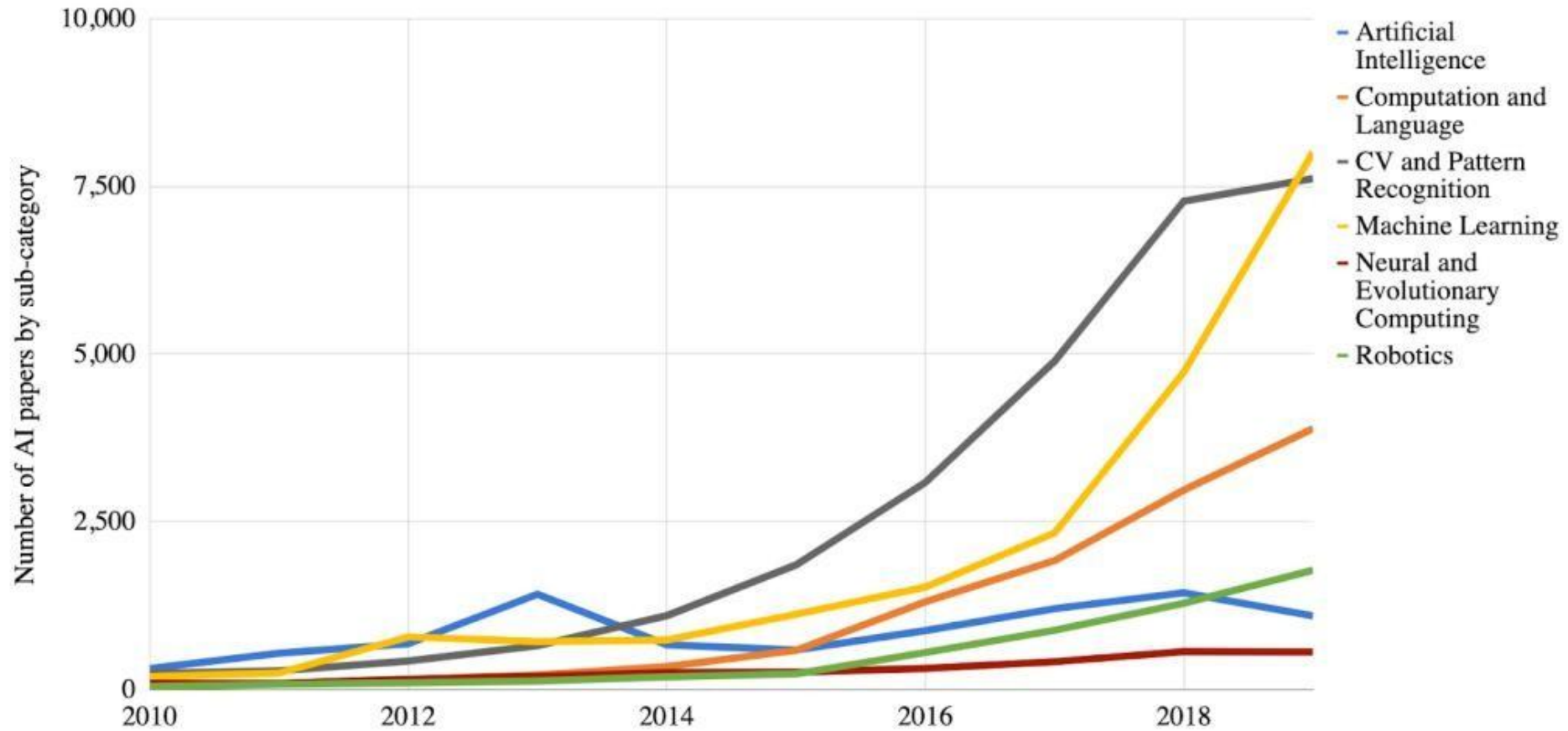


Machine Learning can help humans learn.

Why Use Machine Learning?

Number of AI papers on arXiv, 2010-2019

Source: arXiv, 2019.



AI Index 2019 Annual Report.

Applications of Machine Learning



A votre avis, ça regroupe quoi ?

En texte

Classification de documents

E-mails en spam, shopping, travail, ...

Supprimer tous les spams maintenant (les messages se trouvant dans le dossier Spam depuis plus de 30 jours sont automatiquement supprimés)

<input type="checkbox"/>	<input type="star"/>	<input type="checkbox"/>	Tatiana	Re: Paris en homero - Val the interesting info!	01:50
<input type="checkbox"/>	<input type="star"/>	<input checked="" type="checkbox"/>	cometray	Téléphones les plus compétitifs de Cometray	22:38
<input type="checkbox"/>	<input type="star"/>	<input checked="" type="checkbox"/>	Francois	188 raisons de jouer sur Majestic	27 janv.
<input type="checkbox"/>	<input type="star"/>	<input type="checkbox"/>	Fund Investigation Bureau	TREAT AS URGENT RIGHT AWAY	27 janv.
<input type="checkbox"/>	<input type="star"/>	<input type="checkbox"/>	Mrs Elizabeth Johnson	Hello My Beloved One,	27 janv.
<input type="checkbox"/>	<input type="star"/>	<input checked="" type="checkbox"/>	Evelyn	Re: Amigo, não está satisfeito com o tamanho? Isto pode te ajudar!	27 janv.
<input type="checkbox"/>	<input type="star"/>	<input type="checkbox"/>	Aranda, Aranda (2)	Re: Amigo, o que vc faria com 10km a mais?	28 janv.
<input type="checkbox"/>	<input type="star"/>	<input checked="" type="checkbox"/>	Groupe Paroche	Et encore un gagnant au Megapari !	28 janv.
<input type="checkbox"/>	<input type="star"/>	<input type="checkbox"/>	Carl, Joshua Daniel	MIA	28 janv.
<input type="checkbox"/>	<input type="star"/>	<input checked="" type="checkbox"/>	RCH Tournai	Votre Semaine avec 100000 en Tout	28 janv.
<input type="checkbox"/>	<input type="star"/>	<input checked="" type="checkbox"/>	Jimmy Kismet	Nicolas Beakotts F-E-L-L-I-M-G-_-H-O-R-N-Y-?-_-G-E-T_-L_A_I_D-_-M_O_W!	28 janv.
<input type="checkbox"/>	<input type="star"/>	<input checked="" type="checkbox"/>	Jean-Pierre	Les meilleurs casinos pour les joueurs français	28 janv.

☒ Principale

☒ Réseau social

☒ Promotions

+

<input type="star"/>	CollierPrenom	Aranda	Spécial St Valentin - 3 Jours Seulement - 15% de Réduction !	X	
<input type="star"/>	SoftLayer.com	Aranda	Get a Secure Cloud - We've secured the public cloud with private servers, private networks, and full private clouds.	X	
<input type="checkbox"/>	<input type="star"/>	<input checked="" type="checkbox"/>	Booking.com	Last-minute deals for Montréal and London. Get them before they're gone!	28/12/2014
<input type="checkbox"/>	<input type="star"/>	<input checked="" type="checkbox"/>	Voyages-enc.com	DERNIERE MINUTE NOUVEL AN : profitez des meilleurs prix !	28/12/2014
<input type="checkbox"/>	<input type="star"/>	<input checked="" type="checkbox"/>	Impossible	Year's End Clearance - Up to 20% off Film and Accessories	28/12/2014
<input type="checkbox"/>	<input type="star"/>	<input checked="" type="checkbox"/>	Booking.com	Nicolas - you qualify for at least 20% off places to stay	28/12/2014
<input type="checkbox"/>	<input type="star"/>	<input checked="" type="checkbox"/>	Communauté d'entraide Or.	Nicolas, des questions sur vos produits ?	29/12/2014

gmail.com

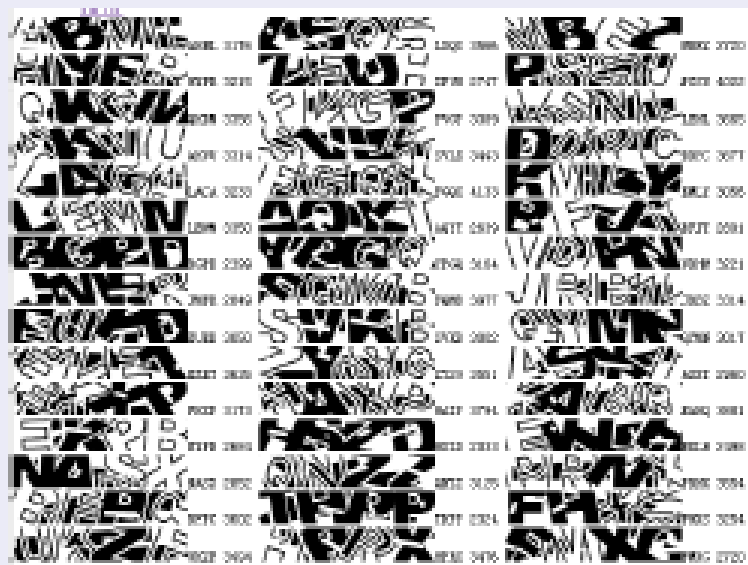
En texte toujours

Reconnaissance de chiffres

8 2 9 4 4 6 4 9 7 0 9 2 9 5 1 5 9 1 0 3
 1 3 5 9 1 7 6 2 8 2 2 5 0 7 4 9 7 8 3 2
 1 1 8 3 6 1 0 3 1 0 0 1 1 2 7 3 0 4 6 5
 2 6 4 7 1 8 9 9 3 0 7 1 0 2 0 3 5 4 6 5

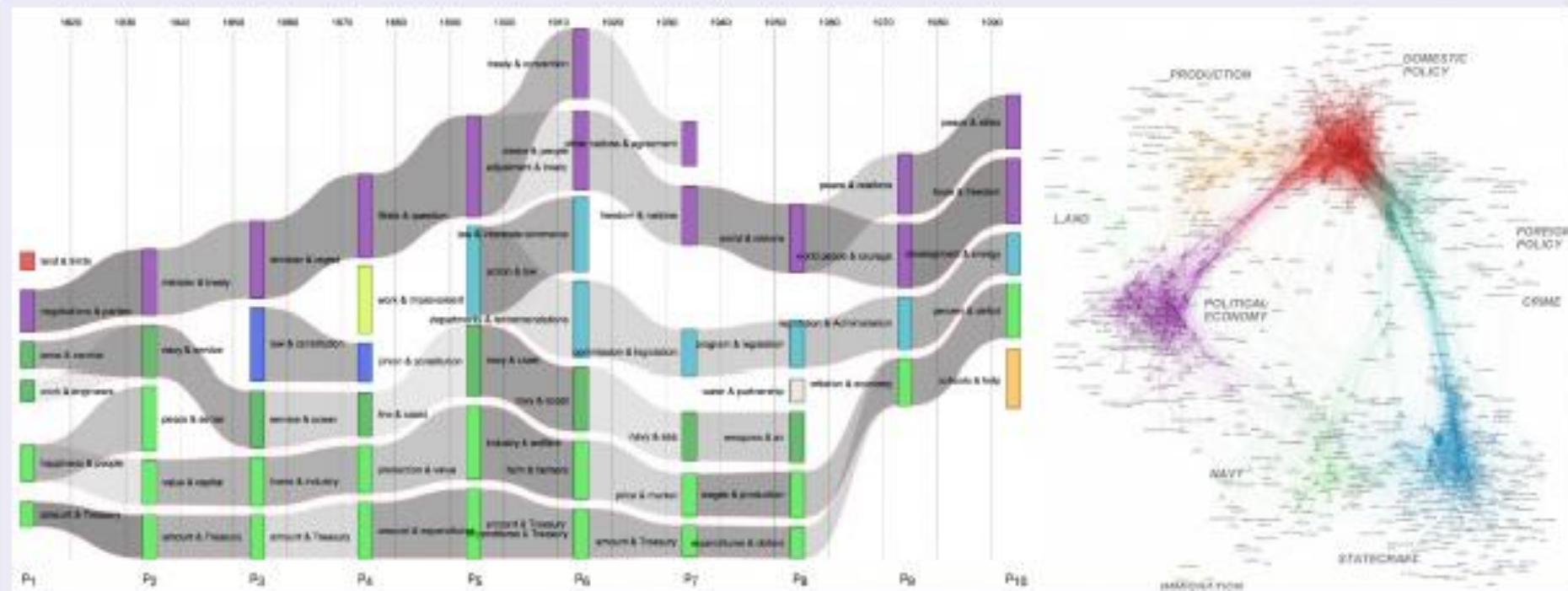
Ou de captcha

[Yann et al. 08], Newcastle University

Characters under typical distortions					Recognition rate
					~100%
					96+%
					100%
					98%
					~100%
					95+%

Sur des documents

Détection de thèmes (topic detection)



Analyse de 255 discours de l'état de l'union, États-Unis

[Rule et al, 2014]

Et plein d'autres applications : traduction, détection de plagiat, résumé automatique, ...

⇒ U.E. Traitement Automatique du Langage

En image

Détection de visages

(opencv)



Mais aussi ...

(betafaceapi.com)



Score: 0.42
X: 398.67
Y: 29.66
Width: 26.79
Height: 26.79
Angle: -5.45

age : 37 (16%), gender : male, race : white, chin size : average, color background : 4c5042 (15%), color clothes middle : 3295eb (40%), color clothes sides : 38a9b5 (96%), color eyes : ac8066, color hair : fb2ea (90%), color mustache : a56855 (65%), color skin : dbb5a1, eyebrows corners : extra low, eyebrows position : average, eyebrows size : extra thin, eyes corners : low, eyes distance : average, eyes position : average, eyes shape : extra round, glasses rim : no, hair beard : none, hair color type : blond (80%), hair forehead : yes, hair length : none, hair mustache : thick, hair sides : very thin, hair top : short, head shape : average, head width : extra narrow, mouth corners : low, mouth height : extra thin, mouth width : extra small, nose shape : extra straight, nose width : wide, teeth visible : no [collapse]



Score: 0.57
X: 216.66
Y: 155.08
Width: 28.34
Height: 28.34
Angle: 0.95

age : 46 (23%), gender : male, race : white, chin size : extra small, color background : 0e0e0d (36%), color beard : 4a2617 (50%), color clothes middle : a22e55 (82%), color clothes sides : a54031 (74%), color eyes : 966a58, color hair : 655348 (77%), color skin : b98f78, eyebrows corners : average, eyebrows position : extra high, eyebrows size : extra thin, eyes corners : average, eyes distance : close, eyes position : extra low, eyes shape : extra thin, glasses rim : no, hair beard : short, hair color type : brown light (77%), hair forehead : no, hair length : short, hair mustache : none, hair sides : thin, hair top : short, head shape : rect, head width : extra wide, mouth corners : average, mouth height : extra thin, mouth width : average, nose shape : average, nose width : extra narrow, teeth visible : no [collapse]

En image

Catégorisation et organisation automatique



En image

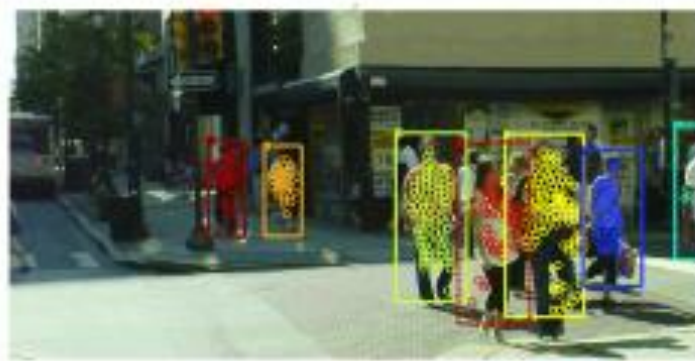
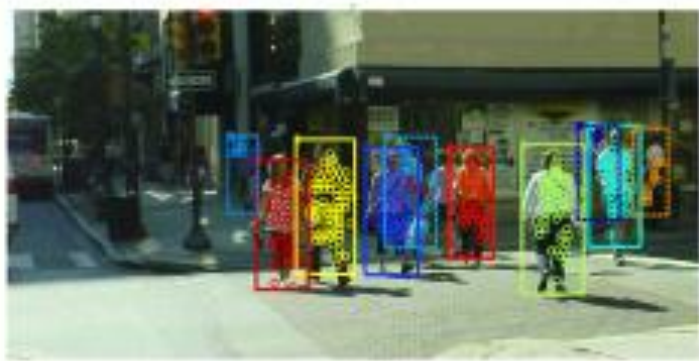
Détection d'objets

teradeep.com, Purdue University



Tracking

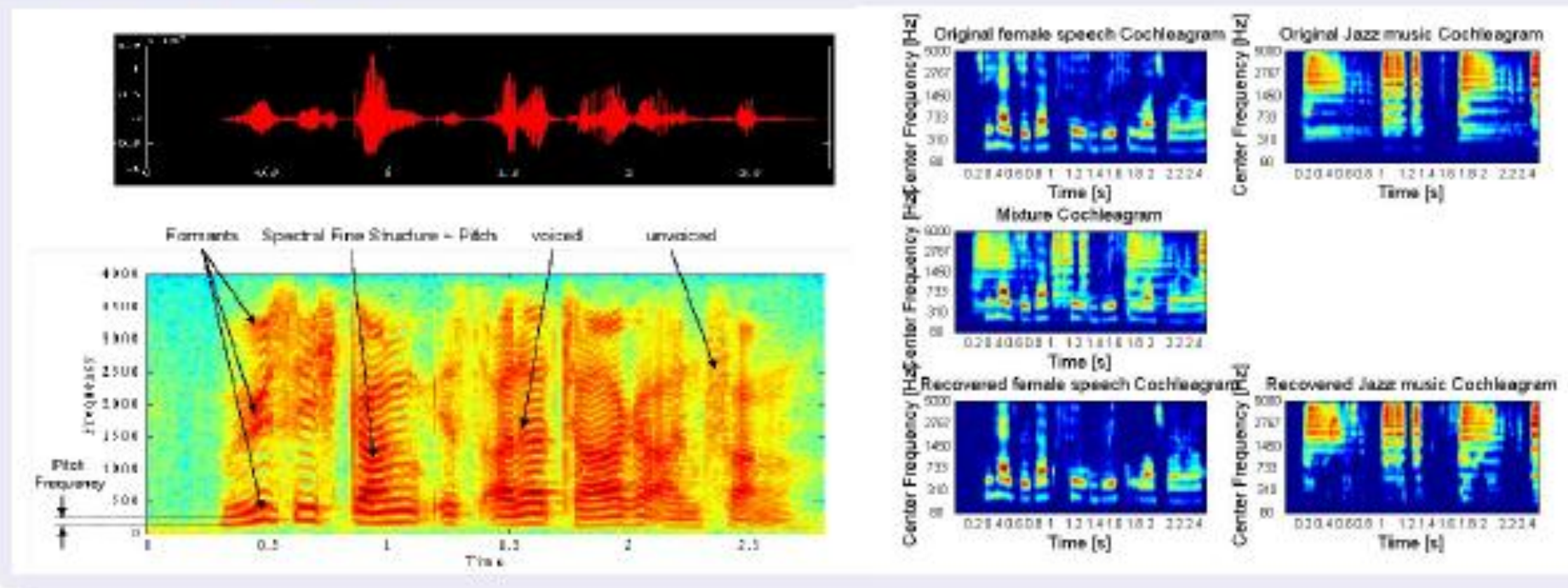
[Fragkiadaki et al. 12], Pennsylvania University



Et l'audio ...

Reconnaissance de la parole, séparation de sources

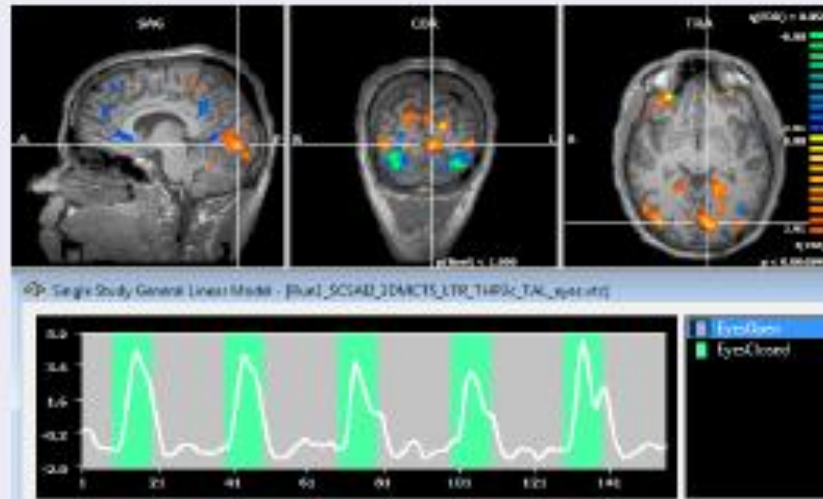
<http://markus-hauenstein.de>



Mais aussi débruitage, transcription musicale, reconnaissance du locuteur, classification/identification de musiques...

Interface cerveau-machine (BCI)

Classification d'actions, de pensées



Contrôle



Objets connectés

Traqueurs d'activité



Surveillance vidéo, monitoring consommation électrique, sécurité réseau



Réseaux sociaux

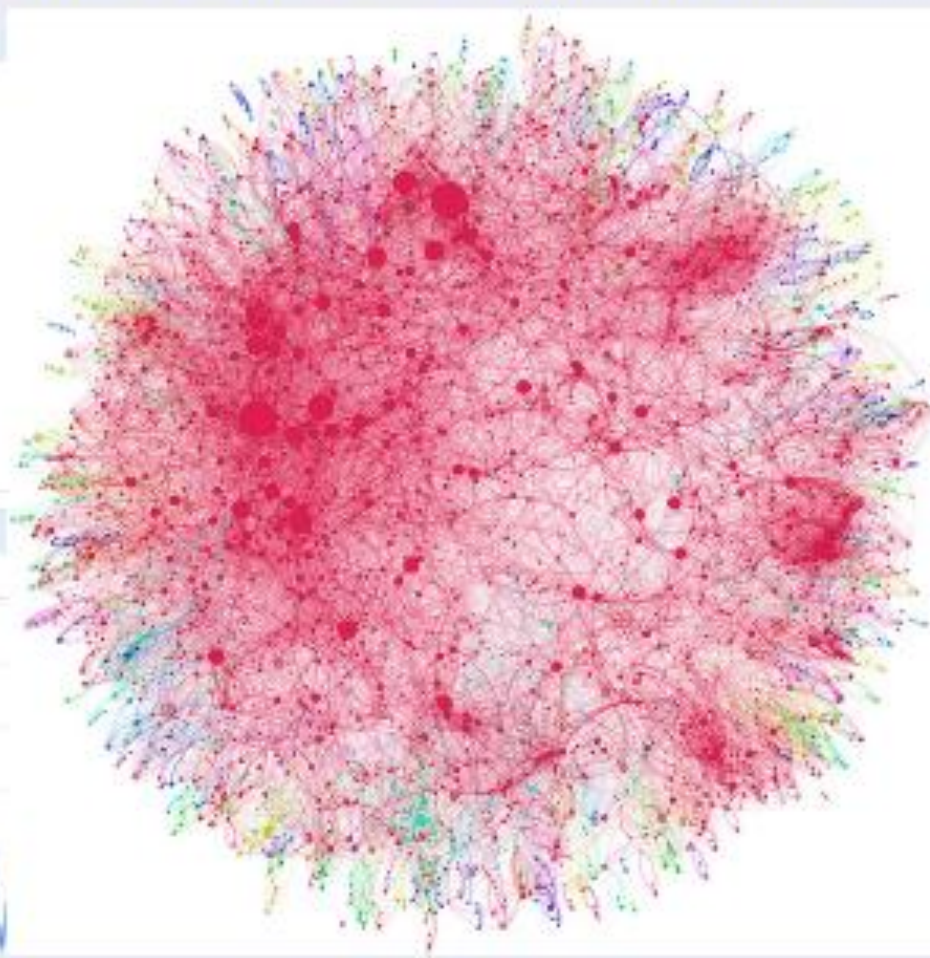
Détection de communauté, phénomènes de diffusions, classification

Largest Diffusion Network



Meme Activity

Zoom: 1' 1' 1h 1d 1w 1m 1y Max



Matchmaking

de profils, sites de rencontre



Experts, CV - Emplois, Jeux



Linked in
viadeo



Systèmes de recommandation

De musiques, de films, de produits, d'amis

Similar Artists



- 1 Bob Dylan
- 2 Radiohead
- 3 Led Zeppelin
- 3 The Rolling Stones
- 5 Pink Floyd
- 6 David Bowie
- 7 The Who
- 8 John Lennon



Recommendation Engine

Search for

Books Music Artists Shows

Umbrellas of Cherbourg, The
Jocelyne Dero
Dolby: 1963 (DVD)

Grounds for
Palace
Jonathan Kaplan
Drama (1990) PG-13

West Beirut
Ziad Doueiri
Drama (1998) PG-13

Suspect
Peter Yates
Thriller (1987) G

Heights
Jennifer Kapst
Drama (2008) R

Batman A.D.
William Riker
Action (2008) PG-13

amazon.com

Recommended for You

Amazon.com has new recommendations for you based on items you purchased or told us you own.



The Little Big Things: 152 Ways to Pursue EXCELLENCE



Fascinate! Your 7 Triggers to Persuasion and Captivation



Sherlock Holmes [Blu-ray]



Alice in Wonderland [Blu-ray]

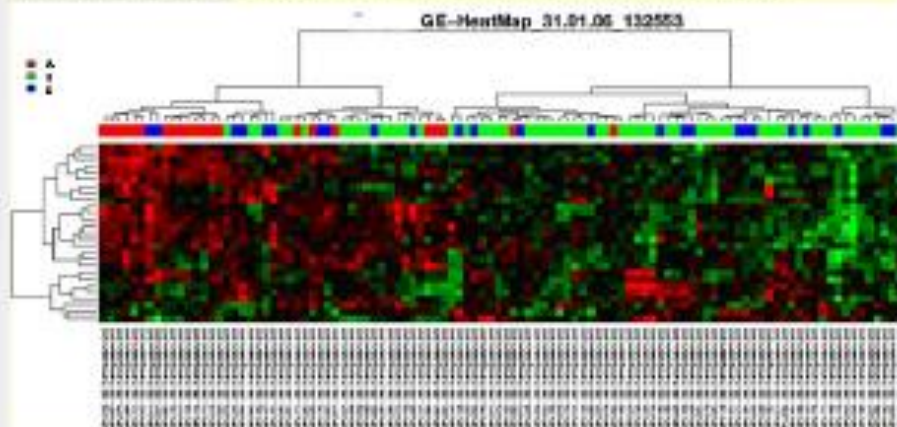


Et dans d'autres sciences

Biologie IZBI, Leipzig University

Gene Signal Value Visualization - Gene Expression Heatmap

This form draws the heatmap of Gene Expression on signals determined by a selected Experiment Group and a selected Gene Group



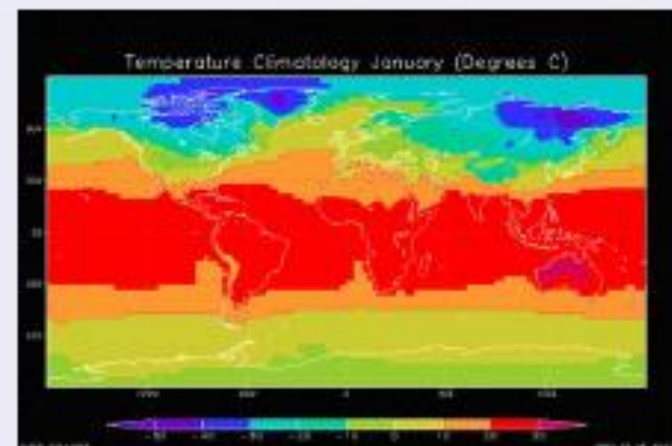
Économie



Astronomie



Climatologie (complétion données)



Dans les jeux et la robotique



Applications of Machine Learning

To summarize, Machine Learning is great for:

- Problems for which existing solutions require a lot of hand-tuning or long lists of rules: one Machine Learning algorithm can often simplify code and perform better.
- Complex problems for which there is no good solution at all using a traditional approach: the best Machine Learning techniques can find a solution.

Applications of Machine Learning

Le machine learning peut servir à résoudre des problèmes

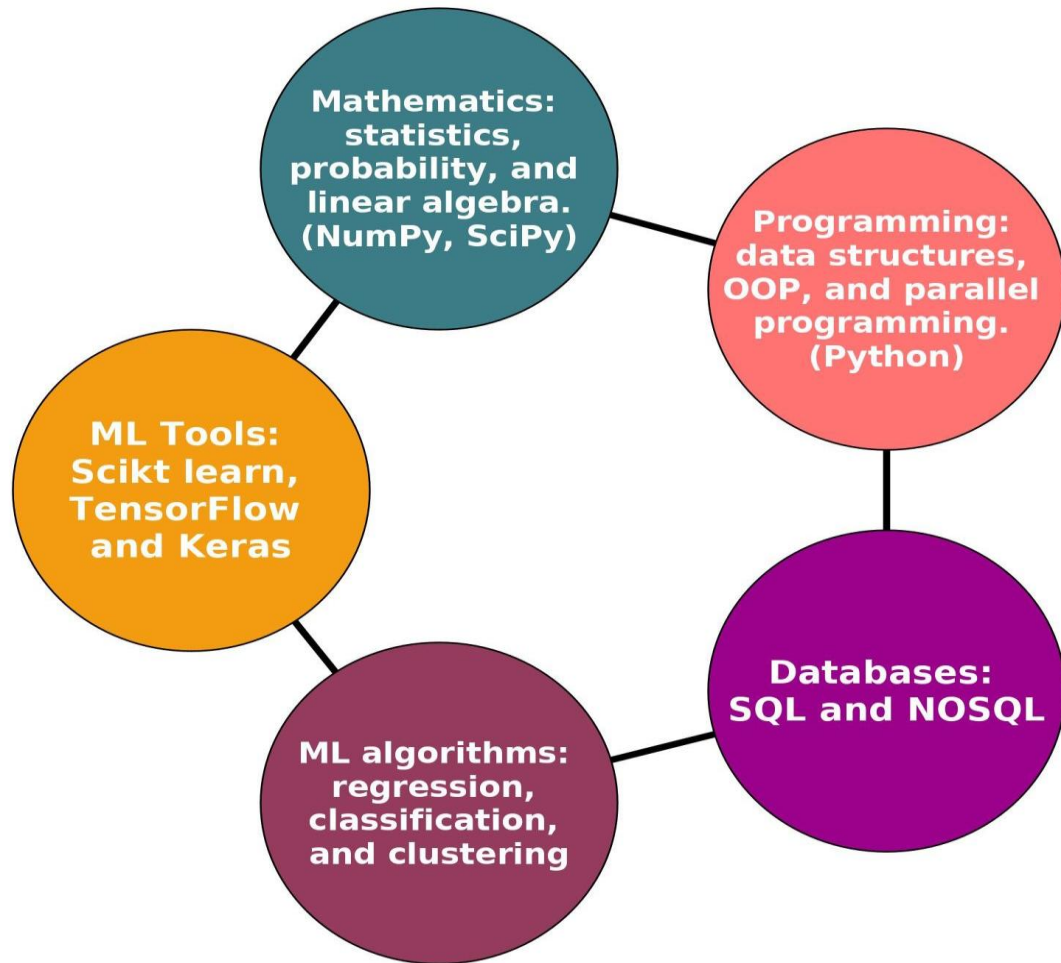
- **que l'on ne sait pas résoudre** (comme dans l'exemple de la prédiction d'achats ci-dessus) ;
- **que l'on sait résoudre, mais** dont on ne sait formaliser en termes algorithmiques comment nous les résolvons (c'est le cas par exemple de la reconnaissance d'images ou de la compréhension du langage naturel) ;
- **que l'on sait résoudre, mais** avec des procédures beaucoup trop gourmandes en ressources informatiques (c'est le cas par exemple de la prédiction d'interactions entre molécules de grande taille, pour lesquelles les simulations sont très lourdes).

Ingrédients du machine learning

Le machine learning repose sur deux piliers fondamentaux :

- d'une part, **les données**, qui sont les exemples à partir duquel l'algorithme va apprendre ;
- d'autre part, **l'algorithme d'apprentissage**, qui est la procédure que l'on fait tourner sur ces données pour produire un modèle. On appelle entraînement le fait de faire tourner un algorithme d'apprentissage sur un jeu de données.

How to get started with ML



How to get started with Machine Learning

- 1) **Mathematics:** statistics, probability, and linear algebra. (NumPy, SciPy)
- 2) **Programming:** data structures, OOP, and parallel programming. (Python)
- 3) **Databases:** SQL and NOSQL.
- 4) **ML algorithms:** regression, classification, and clustering.
- 5) **ML Tools:** Scikit learn, TensorFlow and Keras.

How to get started with ML (WHY)

L'apprentissage automatique (Machine Learning (ML)) est à la croisée de plusieurs disciplines :

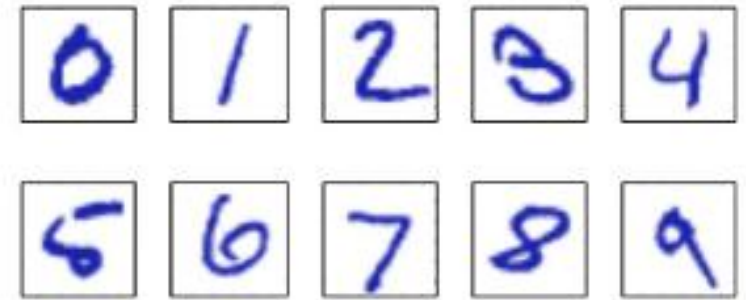
- **Les statistiques** : pour l'inférence de modèles à partir de données.
- **Les probabilités** : pour modéliser l'aspect aléatoire inhérent aux données et au problème d'apprentissage.
- **L'intelligence artificielle** : pour étudier les tâches simples de reconnaissance de formes que font les humains (comme la reconnaissance de chiffres par exemple), et parce qu'elle fonde une branche du ML dite symbolique qui repose sur la logique et la représentation des connaissances.
- **L'optimisation** : pour optimiser un critère de performance afin, soit d'estimer des paramètres d'un modèle, soit de déterminer la meilleure décision à prendre étant donné une instance d'un problème.
- **L'informatique** : puisqu'il s'agit de programmer des algorithmes et qu'en ML ceux-ci peuvent être de grande complexité et gourmands en termes de ressources de calcul et de mémoire.

Exemple 1: prédiction d'achat

- Supposons qu'une entreprise veuille connaître le montant total dépensé par un client ou une cliente à partir de ses factures.
- Il suffit d'appliquer un algorithme classique, à savoir une simple addition : un algorithme d'apprentissage n'est pas nécessaire.
- Supposons maintenant que l'on veuille utiliser ces factures pour déterminer quels produits le client est le plus susceptible d'acheter dans un mois.
- Bien que cela soit vraisemblablement lié, nous n'avons manifestement pas toutes les informations nécessaires pour ce faire.
- Cependant, si nous disposons de l'historique d'achat d'un grand nombre d'individus, il devient possible d'utiliser un algorithme de machine learning pour qu'il en tire un modèle prédictif nous permettant d'apporter une réponse à notre question.

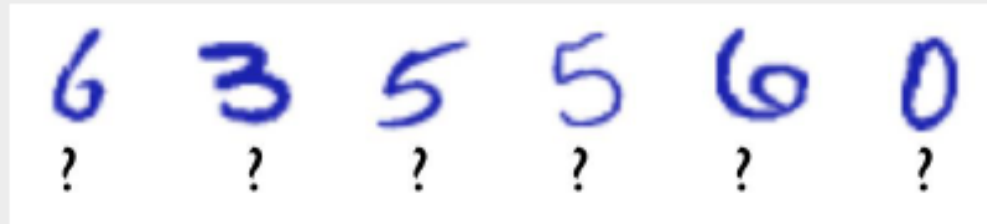
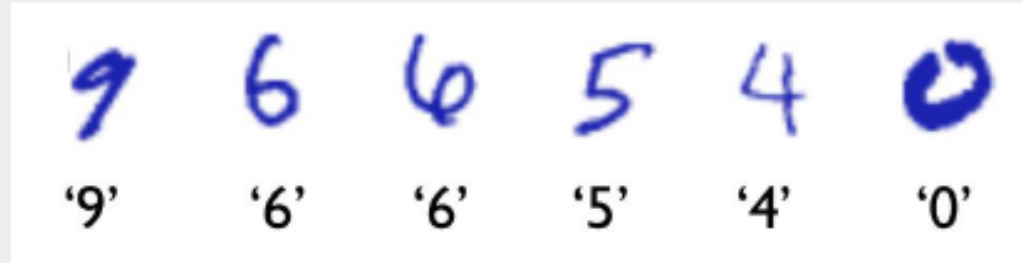
Exemple 2: Reconnaissance de caractère

- Comment reconnaître des Caractères manuscrits ?
- **par énumération de règles**
 - si intensité pixel à la position ...alors c'est un "3"
 - long et fastidieux, difficile de couvrir tous les cas
- **en demandant à la machine d'apprendre**
 - lui laisser faire des essais et apprendre de ses erreurs
 - apprentissage machine (machine-learning)



Exemple 2: Reconnaissance de caractère

- **Comment ça marche ?**
 - On donne à l'algorithme des données d'entraînement.
 - l'algorithme d'apprentissage machine apprend un modèle capable de généraliser à de nouvelles données.



On commence !

OK, Allons-y !