



IMT Mines Alès
École Mines-Télécom



Master Sciences et Numérique pour la Santé

SCIENCE DES DONNÉES



Gérard Dray
Nicolas Sutton-Charani
Pierre Jean



Contexte

A la croisée de plusieurs champs disciplinaires : **mathématiques, probabilités, statistiques, informatique, théorie de l'information et visualisation**, la science des données met en œuvre différents outils d'analyse de données afin d'extraire automatiquement des informations utiles, des connaissances, à partir de données potentiellement massives. Le but ultime est de rendre cette information plus facile à exploiter, la protéger et la valoriser. Elle pourra servir de base ensuite à des processus d'évaluation et d'aide à la décision.

Ce cours est organisé en 2 UEs :

Master 1 : Bases mathématiques et informatiques de la science des données.

Master 2 : Notions avancées d'intelligence artificielle et d'apprentissage artificiel.



IMT Mines Alès
École Mines-Télécom



Master Sciences et Numérique pour la Santé

SCIENCE DES DONNÉES - NIVEAU 1



Gérard Dray
Nicolas Sutton-Charani
Pierre Jean



Objectifs :

Renforcer les compétences théoriques avec un approfondissement des statistiques et de la théorie des probabilités.

Introduire les méthodes et techniques d'apprentissage artificiel.

Former à l'utilisation de langages informatiques (SQL et R) pour réaliser des projets permettant de mettre en pratique les aspects théoriques de la science des données.

Mots clés : probabilités, statistiques, base de données, apprentissage artificiel, R, sql.

Prérequis : notions de base en probabilités, statistiques et programmation.



Equipe pédagogique :

- Gérard Dray (GD) - Enseignant Chercheur - EuroMov Digital Health in Motion, Univ Montpellier, IMT Mines Ales
- Nicolas Sutton-Charani (NSC) - Enseignant Chercheur - EuroMov Digital Health in Motion, Univ Montpellier, IMT Mines Ales
- Pierre Jean (PJ) - Ingénieur de Recherche - EuroMov Digital Health in Motion, Univ Montpellier, IMT Mines Ales

Date	Heure début	Heure fin	Intervenant	Salle	Moyens pédagogiques
jeudi 25 janvier 2024	13:15:00	18:15:00	Gérard Dray	TD 36.404	Cours/TD
jeudi 1 février 2024	13:15:00	18:15:00	Pierre Jean	TD 36.412	Cours/TD
jeudi 8 février 2024	13:15:00	18:15:00	Nicolas Sutton-Charani	TD 36.412	Cours/TD
jeudi 15 février 2024	13:15:00	18:15:00	Nicolas Sutton-Charani	TD 36.315	Cours/TD
jeudi 28 mars 2024	13:15:00	18:15:00		TD 36.410	Projet
jeudi 4 avril 2024	13:15:00	18:15:00	Nicolas Sutton-Charani	TD 36.417	Cours/TD
jeudi 11 avril 2024	13:15:00	18:15:00	Equipe pédagogique	TD 36.410	Projet
jeudi 25 avril 2024	13:15:00	18:15:00	Equipe pédagogique	TD 36.410	Projet



IMT Mines Alès
École Mines-Télécom



Master Sciences et Numérique pour la Santé

SCIENCE DES DONNÉES - NIVEAU 1

INTRODUCTION

Gérard Dray



G rard Dray
gerard.dray@mines-ales.fr

Chercheur

UMR EuroMov Digital Health in Motion,
Univ Montpellier,
IMT Mines Ales, Ales,
France



Enseignant

IMT Mines Al s
D partement d'enseignement
Informatique et Intelligence Artificielle
2IA

Universit  de Montpellier
IAE

Master 1 E_Mkg
Master 1 MTD
Master 2 E-Mkg
Master 2 MPW

Facult  des Sciences, Facult  de M decine
Master Sciences et Num rique pour la Sant 
Master Sant 

UFR STAPS
Masters STAPS

Objectifs

Être capable de comprendre :

- la définition de la science des données
- les définitions de l'Intelligence Artificielle
- la méthodologie d'Apprentissage Artificiel

Initier une réflexion sur l'usage de la science des données dans le domaine de la santé numérique

Organisation

Programme :

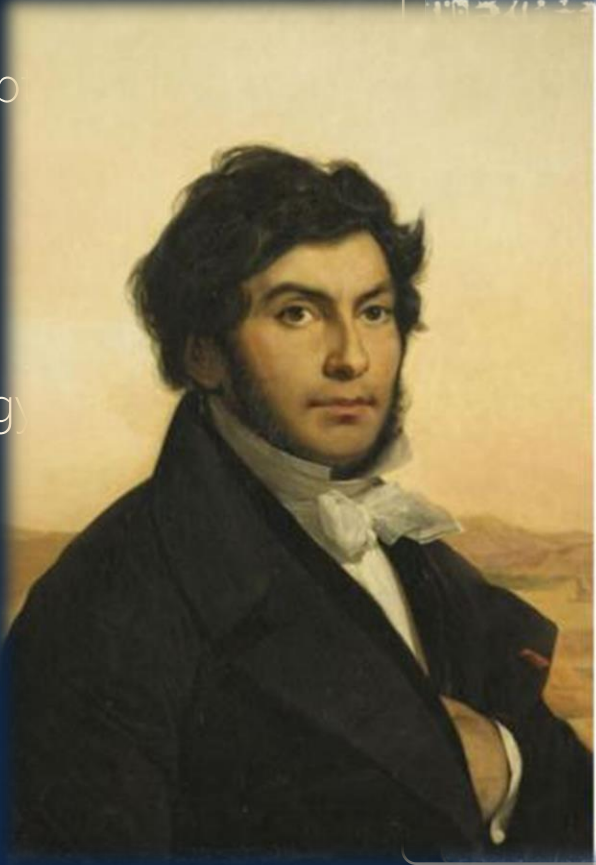
- Une histoire ancienne
- Principales définitions
- Méthode d'apprentissage automatique : une illustration
- Exemples
- Discussion et Conclusion

Une histoire ancienne

1822
Jean-François Champollion

Egypt

Egypt



«Au cours du règne du jeune qui a succédé à son père, Seigneur des diadèmes, très pieux et a été pieux sur ses ennemis, vie civilisée aux cérémonies des Ptah le Grand, un pays Supérieur des Dieux Ptah, à qui Ra a ante d'Amun, fils

JAMAIS, BIEN-neuvième année, os était prêtre Soteres, et les Evergetai, et les Dieu Epiphanes Philinos qui est ergetis, fille Areia choros d'Arsinoe Ptolémée qui est les quatrièmes es les Égyptiens

Pierre de Rosette, 196 av. J.-C

1854 - John Snow Épidémie de choléra, Londres



Medical Societies.

MEDICAL SOCIETY OF LONDON.

MR. HEADLAND, PRESIDENT.

SATURDAY, OCTOBER 14TH, 1854.

Dr. Snow considered that the cholera poison acted upon the alimentary canal, and not on the blood or nervous system. In every case which he had seen, the evacuations had been sufficient to account for the collapse, without reference to any other cause. There was no poison in the blood in a case of cholera; in the consecutive fever, as it was called, the blood became poisoned from urea getting into the circulation in consequence of the kidneys not acting, but not from any poison having been present from the beginning. There was nothing in the atmosphere to account for the spread of cholera, which he believed was spread from person to person; and that in all cases it could be traced in this manner. If atmospheric, why did it attack one or two persons only in a locality, and these having direct communication with each other? Such cases he had seen at Sydenham, where there had been only two instances of the disease. The first case in the outbreak of 1849 had occurred to a sailor in Bermondsey; the second affected person was the successor to the sailor in the room in which he died. He thought he had collected evidence enough to show that in all cases cholera was propagated by swallowing some portion of the evacuations of an affected person. These, as was well known, flowed into the bed, &c., and persons attending on the sick might easily take the poison unawares. With respect to the class of persons affected by the disease, he believed that the very poor and vagabonds suffered less, in proportion, than decent, respectable persons. He regarded the cholera and diarrhoea, as lately prevalent, to be the same disease in different degrees of intensity. We observed the same difference in scarlatina and other diseases.



NOTE. Boundary within which all the Deaths are indicated, is shown thus, . Distances between two Diagonals, thus, . SCALE 30 INCHES TO A MILE.

1943

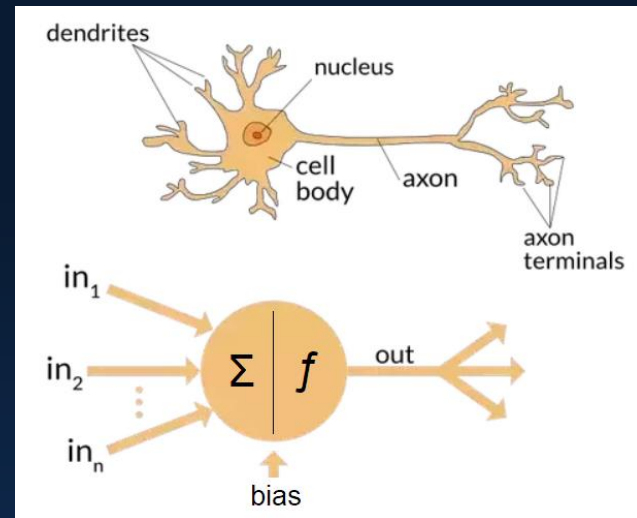
Le neurone formel de McCulloch et Pitts



Warren Sturgis McCulloch
(1898 – 1969)



Walter Harry Pitts, Jr.
(1923 – 1969)



1956

John McCarthy

Artificial Intelligence

*“The science and engineering of making intelligent machines,
especially intelligent computer programs”*



1959

Arthur Samuel

“Machine learning is a field of study that gives computers the ability to learn without being explicitly programmed.”



1963

Alan Mathison Turing



“Instead of trying to produce a programme to simulate the adult mind, why not rather try to produce one which simulates the child's ? If this were then subjected to an appropriate course of education one would obtain the adult brain.”

FOUILLE DE DONNEES (DATA MINING)



1997

Le jour où Deep Blue a battu Garry Kasparov aux échecs



11 mai 1997, le champion d'échecs Garry Kasparov perd la sixième partie d'un match historique.

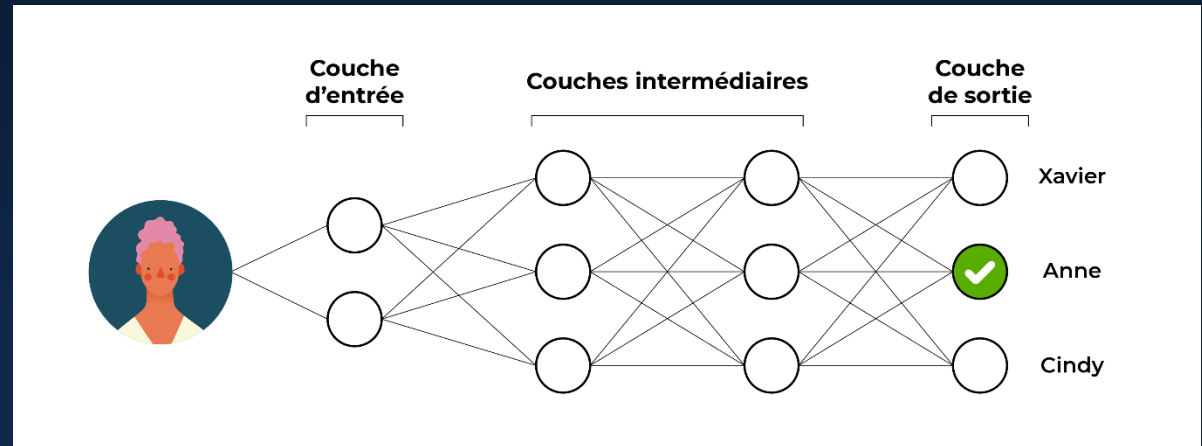
2005

Apprentissage Profond (Deep Learning)

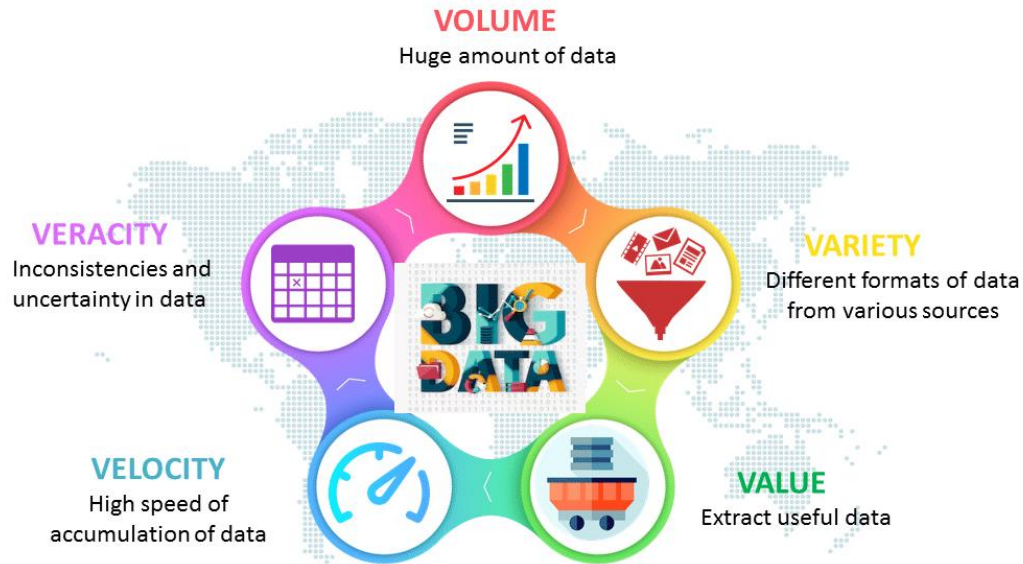


Geoffrey Hinton (né le 6 décembre 1947) est un chercheur canadien spécialiste de l'intelligence artificielle et plus particulièrement des réseaux de neurones artificiels. Il fait partie de l'équipe Google Brain et est professeur au département d'informatique de l'Université de Toronto. Il a été l'un des premiers à mettre en application l'algorithme de rétropropagation du gradient pour l'entraînement d'un réseau de neurones multi-couches. Il fait partie des figures de proue de la communauté de l'apprentissage profond.

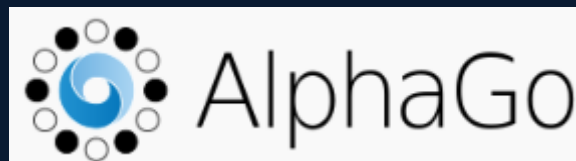
https://fr.wikipedia.org/wiki/Geoffrey_Hinton



Big Data



Deep Learning
Apprentissage Profond



MODERN DATA SCIENTIST

Data Scientist, the sexiest job of 21st century requires a mixture of multidisciplinary skills ranging from an intersection of mathematics, statistics, computer science, communication and business. Finding a data scientist is hard. Finding people who understand who a data scientist is, is equally hard. So here is a little cheat sheet on who the modern data scientist really is.

MATH & STATISTICS

- ☆ Machine learning
- ☆ Statistical modeling
- ☆ Experiment design
- ☆ Bayesian inference
- ☆ Supervised learning: decision trees, random forests, logistic regression
- ☆ Unsupervised learning: clustering, dimensionality reduction
- ☆ Optimization: gradient descent and variants

DOMAIN KNOWLEDGE & SOFT SKILLS

- ☆ Passionate about the business
- ☆ Curious about data
- ☆ Influence without authority
- ☆ Hacker mindset
- ☆ Problem solver
- ☆ Strategic, proactive, creative, innovative and collaborative



PROGRAMMING & DATABASE

- ☆ Computer science fundamentals
- ☆ Scripting language e.g. Python
- ☆ Statistical computing package e.g. R
- ☆ Databases SQL and NoSQL
- ☆ Relational algebra
- ☆ Parallel databases and parallel query processing
- ☆ MapReduce concepts
- ☆ Hadoop and Hive/Pig
- ☆ Custom reducers
- ☆ Experience with xaaS like AWS

COMMUNICATION & VISUALIZATION

- ☆ Able to engage with senior management
- ☆ Story telling skills
- ☆ Translate data-driven insights into decisions and actions
- ☆ Visual art design
- ☆ R packages like ggplot or lattice
- ☆ Knowledge of any of visualization tools e.g. Flare, D3.js, Tableau

Google Actualités

google le dépistage cancer sein

À la une

Couverture complète

Comment les articles sont-ils classés ?

Recherches enregistrées

Top des articles

IA et dépistage du cancer du sein : un algorithme de Google Health fait mieux que le radiologue

Cancer du sein : d'après Google, son intelligence artificielle le diagnostique mieux que les experts

11/01/2020

L'IA de Google dépiste le cancer du sein avec plus de précision que les médecins, selon cette étude

11/01/2020

L'IA de Google est plus performante que les médecins pour détecter le cancer du sein, mais elle ne les remplacera pas

11/01/2020

Tweets

Johel de Rosnay @johelrosnay

Des chercheurs de Google Health à Palo Alto et de Deep Mind à Londres, mettent au point un test utilisant l'intelligence artificielle qui détecte les tumeurs précoce du cancer du sein mieux que les experts <https://t.co/58l4p3V8vb>

Twitter · 10/01/2020 18:40

Tous les articles

Cancer du sein : Une intelligence artificielle plus précise que des humains pour établir un diagnostic

11/01/2020

Google : met l'honneur l'IA en matière de dépistage | Zone bourse

11/01/2020

LE QUOTIDIEN DU MEDECIN | TOUS LES ACTUALITÉS MÉDICALES SOCIÉTÉ PROFESSIONNELLE

Rechercher

S'ABONNER DÈS 10€

À LA UNE | SANTÉ | LIBÉRAL | HÔPITAL | SPÉCIALITÉS | INTERNES | ANNONCES / EMPLOI

Accueil > Spécialités > Cancérologie

IA et dépistage du cancer du sein : un algorithme de Google Health fait mieux que le radiologue

PAR CHARLENE CATALPAUD - PUBLIÉ LE 09/01/2020

2 RÉACTIONS COMMENTER

Une équipe britanno-américaine a mis au point un système d'intelligence artificielle (IA) dont les performances lui permettent de surpasser le radiologue pour dépister le cancer du sein. Cet outil est décrit dans « Nature ».

SCIENCES
AVENIR

Allergies | Cancer | Cannabis à la fois | Climat | COVID-19 | Démocratie | Énergie | Espace | Économie | Santé | Vieillesse | Ségnesse

#Sciences Pour les jeunes | Incendies en Australie | Les bienfaits de la méditation | Nouvelle histoire du cosmos

Une IA de Google surpasse les radiologues pour détecter le cancer du sein

Par Chârelène Catalpaud le 09.01.2020 à 22:04

Google a mis au point une intelligence artificielle qui sait mieux repérer les tumeurs précoces sur les mammographies que les radiologues. Ce n'est pas la première fois qu'un algorithme bat les médecins pour détecter les tumeurs sur une imagerie.

Il y a 178 000 ans, nos ancêtres commencent à être inquiétés par la tumeur. Depuis 1913, chaque femme nous rappelle l'importance de se faire dépister. Mais c'est la première fois que la machine l'emporte.

UN DÉPISTAGE PRÉCIS À L'ÉCHELLE DU BIEN-ÊTRE

Article

International evaluation of an AI system for breast cancer screening

<https://doi.org/10.1038/s41586-019-1799-6>

Received: 27 July 2019

Accepted: 5 November 2019

Published online: 1 January 2020

Scott Mayer McKinney^{1,2*}, Marcin Sieniek^{1,3,4}, Varun Gooble^{1,5}, Jonathan Godwin^{1,3,4}, Natasha Antropova⁵, Hutan Ashrafian^{1,4}, Trevor Back¹, Mary Chesnut⁵, Greg C. Corrado¹, Ara Darzi^{1,4,5}, Mozziyar Etemad⁶, Florencia Garcia-Vicente⁶, Fiona J. Gilbert⁷, Mark Halling-Brown⁸, Demis Hassabis⁹, Sunny Jansen⁹, Alan Karthikesalingam¹⁰, Christopher J. Kelly¹⁰, Dominic King¹⁰, Joseph R. Ledsam¹⁰, David Melnick¹⁰, Hormuz Mostofi¹, Lily Peng¹, Joshua Jay Reicher¹⁰, Bernardino Romera-Paredes¹⁰, Richard Sibdotton^{12,13}, Mustafa Suleyman⁹, Daniel Tse⁹, Kenneth C. Young⁹, Jeffrey De Fauw^{1,2,9} & Shraya Shetty^{1,2,9*}

Screening mammography aims to identify breast cancer at earlier stages of the disease, when treatment can be more successful¹. Despite the existence of screening programmes worldwide, the interpretation of mammograms is affected by high rates of false positives and false negatives². Here we present an artificial intelligence (AI) system that is capable of surpassing human experts in breast cancer prediction. To assess its performance in the clinical setting, we curated a large representative dataset from the UK and a large enriched dataset from the USA. We show an absolute reduction of 5.7% and 1.2% (USA and UK) in false positives and 9.4% and 2.7% in false negatives. We provide evidence of the ability of the system to generalize from the UK to the USA. In an independent study of six radiologists, the AI system outperformed all of the human readers: the area under the receiver operating characteristic curve (AUC-ROC) for the AI system was greater than the AUC-ROC for the average radiologist by an absolute margin of 11.5%. We ran a simulation in which the AI system participated in the double-reading process that is used in the UK, and found that the AI system maintained non-inferior performance and reduced the workload of the second reader by 88%. This robust assessment of the AI system paves the way for clinical trials to improve the accuracy and efficiency of breast cancer screening.

Breast cancer is the second leading cause of death from cancer in women¹, but early detection and treatment can considerably improve outcomes^{2,3}. As a consequence, many developed nations have implemented large-scale mammography screening programmes. Major medical and governmental organizations recommend screening for all women starting between the ages of 40 and 50^{1,4}. In the USA and UK combined, over 42 million exams are performed each year^{5,6}. Despite the widespread adoption of mammography, interpretation of these images remains challenging. The accuracy achieved by experts in cancer detection varies widely, and the performance of even the best clinicians leaves room for improvement^{7,8}. False positives can lead to patient anxiety⁹, unnecessary follow-up and invasive diagnostic procedures. Cancers that are missed at screening may not be identified until they are more advanced and less amenable to treatment⁴.

AI may be uniquely poised to help with this challenge. Studies have demonstrated the ability of AI to meet or exceed the performance of human experts on several tasks of medical-image analysis^{10–19}.

As a shortage of mammography professionals threatens the availability and adequacy of breast-screening services around the world^{20,21}, the scalability of AI could improve access to high-quality care for all. Computer-aided detection (CAD) software for mammography was introduced in the 1990s, and several assistive tools have been approved for medical use²⁴. Despite early promise^{22,23}, this generation of software failed to improve the performance of readers in real-world settings^{23,24}. More recently, the field has seen a renaissance owing to the success of deep learning. A few studies have characterized systems for breast cancer prediction with stand-alone performance that approaches that of human experts^{25,26}. However, the existing work has several limitations. Most studies are based on small, enriched datasets with limited follow-up, and few have compared performance to readers in actual clinical practice—instead relying on laboratory-based simulations of the reading environment. So far there has been little evidence of the ability of AI systems to translate between different screening populations and settings without additional training data²⁷. Critically, the pervasive use of follow-up intervals that are no longer than 12 months^{28,32,33}

*Google Health, Palo Alto, CA, USA. ²DeepMind, London, UK. ³Department of Surgery and Cancer, Imperial College London, London, UK. ⁴Institute of Global Health Innovation, Imperial College London, London, UK. ⁵Cancer Research UK Imperial Centre, Imperial College London, London, UK. ⁶Northwestern Medicine, Chicago, IL, USA. ⁷Department of Radiology, Cambridge Biomedical Research Centre, University of Cambridge, Cambridge, UK. ⁸Royal Surrey County Hospital, Guildford, UK. ⁹Unity Life Sciences, South San Francisco, CA, USA. ¹⁰Google Health, London, UK. ¹¹Stanford Health Care and Palo Alto Veterans Affairs, Palo Alto, CA, USA. ¹²The Royal Marsden Hospital, London, UK. ¹³Thirlstone Breast Centre, Cheltenham, UK. These authors contributed equally: Scott Mayer McKinney, Marcin T. Sieniek, Varun Gooble, Jonathan Godwin. *These authors jointly supervised this work: Jeffrey De Fauw, Shraya Shetty. *e-mail: scottmayer@google.com; tsed@google.com; sshetty@google.com

https://www.nature.com/articles/s41586-019-1799-6.pdf

<https://simulator.drdata.io/>

Calculez le prix de vos données !

FR EN

 Dr Data

Vos données personnelles ont un prix pour les entreprises commerciales dans le cadre des démarches publicitaires et de ventes de bases de données. Ces données s'échangent et se revendent entre ces acteurs, et ce partout dans le monde.

Nous avons cherché à en connaître les prix, non sans mal !

Notre message n'est pas de vous encourager à vendre vos données, mais plutôt de vous faire prendre conscience du marché qu'elles représentent.

Notre objectif est de vous sensibiliser à la valeur économique de vos données personnelles pour vous aider à mieux comprendre les enjeux de la protection de votre vie privée.

Votre navigation est anonyme et libre sans cookies et sans compte.



Etat civil et Coordonnées

Les informations concernant votre identité tel que votre sexe, numéro de téléphone ou votre âge sont des données d'identification dites «classiques» qui peuvent déterminer un premier profil de votre personne..

Ajoutez des données à partager pour en connaître le prix

<input checked="" type="checkbox"/> Sexe	<input type="checkbox"/> Adresse Postale	<input type="checkbox"/> Numéro de Sécurité Sociale
<input type="checkbox"/> Age	<input type="checkbox"/> Adresse Email	<input type="checkbox"/> Catégorie Socio-Professionnelle
<input type="checkbox"/> Origine Ethnique	<input checked="" type="checkbox"/> Numero de Téléphone	<input type="checkbox"/> Niveau d'Etude

Possédez-vous un casier judiciaire ?

Oui Non

Avez vous un permis de conduire ?

Oui Non

Total

Partager nos données, c'est bien, le faire avec l'information complète c'est mieux !

17.0004 €

Le prix affiché est calculé en fonction des sources que nous avons pu récupérer et étudier, vous pouvez y accéder en [cliquant ici](#).

Les données ne sont pas exhaustives. Le résultat est le fruit de notre algorithme sur la base des informations disponibles.



Mentions légales & CGU

Copyright © 2022 DrData. Tous droits réservés.

En partenariat avec ...



IMT Mines Alès
Ecole Mines-Télécom

Le Parisien

XV de France : «Ce ne sont pas les chiffres qui nous dirigent»... Comment Fabien Galthié utilise les datas

Le XV de France, qui affronte le Japon dimanche (14 heures) à Toulouse lors de son dernier match de la tournée d'automne, s'appuie dans sa préparation sur une multitude de données chiffrées, collectées, triées puis fournies par la société SAS France.



Fabien Galthié avant le match contre l'Afrique du Sud au Velodrome, à Marseille. Icon Sport/Johnny Fidelin

Les chiffres, fournis notamment par une puce GPS incrustée dans les maillots, remontent et s'inscrivent dans une base de données où figurent 1438 matchs (XV de France, U20 et Top 14 compris). Les ballons connectés et les protège-dents connectés élargissent également le champ des investigations. « Attention, ce ne sont pas les chiffres qui nous dirigent mais nous qui dirigeons les chiffres », lance Fabien Galthié. D'ailleurs, sur un match, je ne suis connecté à rien. Je regarde, j'écoute, je suis au cœur de l'environnement, il n'y a pas un seul chiffre qui m'impacte. Les décisions, on les prend avec notre intuition et notre savoir-faire. » La science, elle, est là pour répondre aux interrogations.



OpenAI is an AI research and deployment company. Our mission is to ensure that artificial general intelligence benefits all of humanity.

OpenAI's mission is to ensure that artificial general intelligence (AGI)—by which we mean highly autonomous systems that outperform humans at most economically valuable work—benefits all of humanity.

We will attempt to directly build safe and beneficial AGI, but will also consider our mission fulfilled if our work aids others to achieve this outcome.

OpenAI Charter

We're releasing a charter that describes the principles we use to execute on OpenAI's mission. This document reflects the strategy we've refined over the past two years, including feedback from many people internal and external to OpenAI. The timeline to AGI remains uncertain, but our charter will guide us in acting in the best interests of humanity throughout its development.

Newsroom

Announcements

ChatGPT: Optimizing Language Models for Dialogue
November 30, 2022 — Announcements, Research

DALL-E API Now Available in Public Beta
November 3, 2022 — Announcements, API

DALL-E Now Available Without Waitlist
September 28, 2022 — Announcements

<https://chat.openai.com/chat>

<https://openai.com/dall-e-2/>



MIDJOURNEY
All about imagination

<https://www.midjourney.com/app/>





MIDJOURNEY

All about imagination

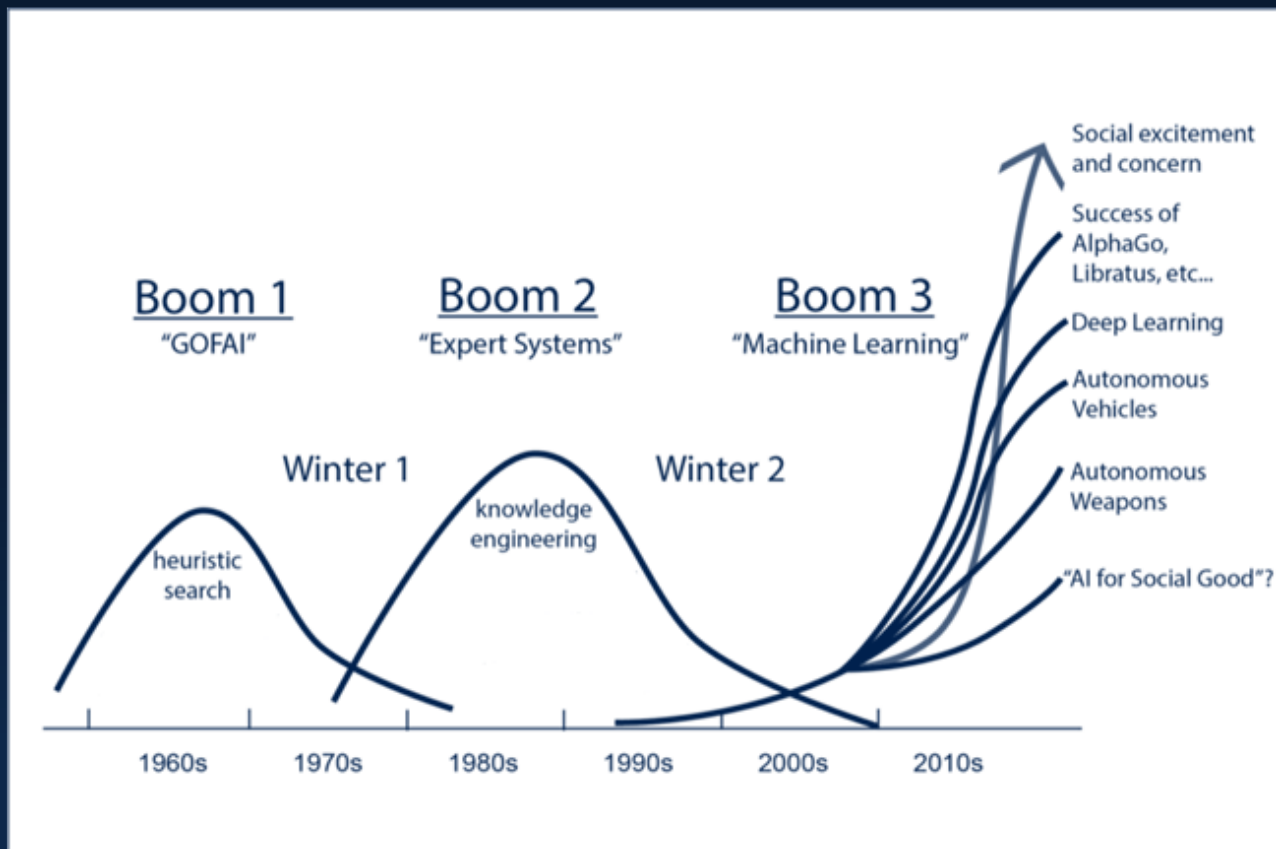
Midjourney Bot

/Imagine

Ultra realistic photo of a class of students taking a machine-learning workshop and listening to a human teacher



Is winter coming ?



Principales définitions

Qu'est-ce que l'intelligence artificielle? (Artificial Intelligence)

L'intelligence artificielle représente tout outil utilisé par une machine afin de « reproduire des comportements liés aux humains, tels que le raisonnement, la planification et la créativité ».

Qu'est-ce que l'apprentissage automatique ? (Machine Learning)

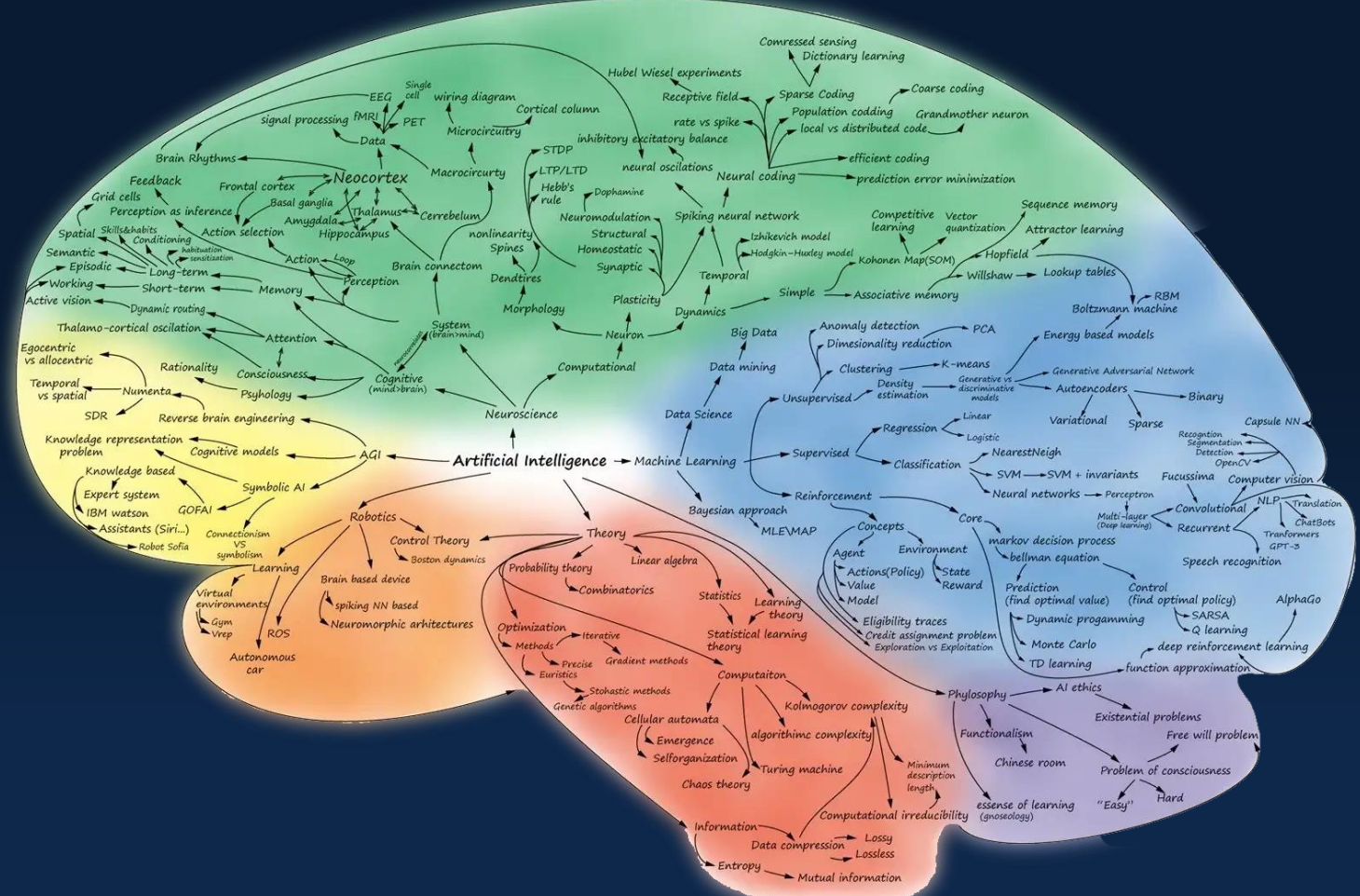
L'apprentissage automatique est un champ d'étude de l'intelligence artificielle qui vise à donner aux machines la capacité d'« apprendre » à partir de données, via des modèles mathématiques. Plus précisément, il s'agit du procédé par lequel les informations pertinentes sont tirées d'un ensemble de données d'entraînement.

Le but de cette phase est l'obtention des paramètres d'un modèle qui atteindront les meilleures performances, notamment lors de la réalisation de la tâche attribuée au modèle. Une fois l'apprentissage réalisé, le modèle pourra ensuite être déployé en production.

Qu'est-ce que l'apprentissage profond ? (Deep Learning)

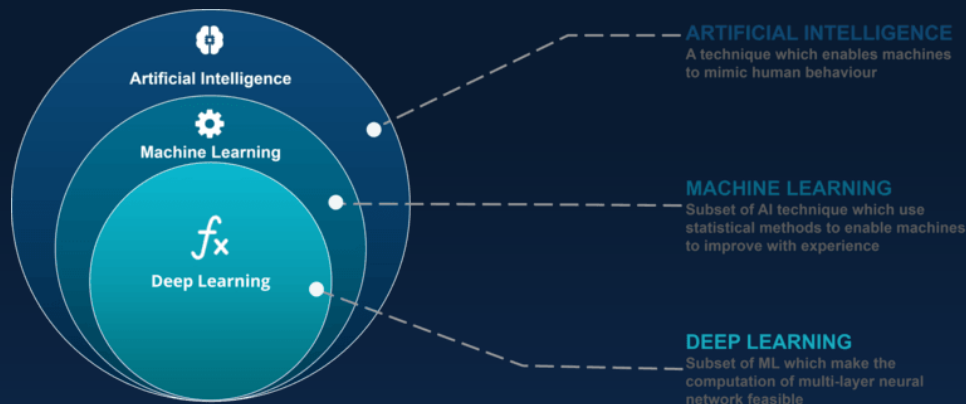
L'apprentissage profond est un sous-ensemble de l'apprentissage automatique. L'apprentissage profond est un procédé d'apprentissage automatique utilisant des réseaux de neurones possédants plusieurs couches de neurones cachées. Ces algorithmes possédant de très nombreux paramètres, ils demandent un nombre très important de données afin d'être entraînés.

Qu'est-ce que l'intelligence artificielle?

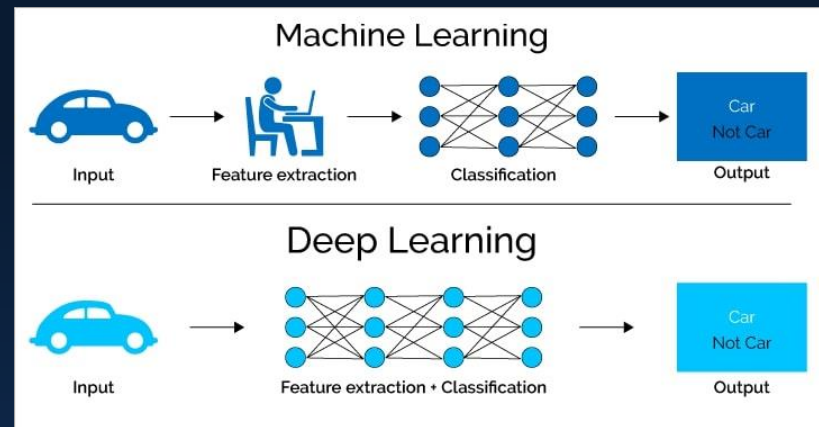


Qu'est-ce que l'apprentissage automatique ? (Machine Learning)

Qu'est-ce que l'apprentissage profond ? (Deep Learning)



<https://datawider.com/how-deep-learning-is-different-from-machine-learning/>

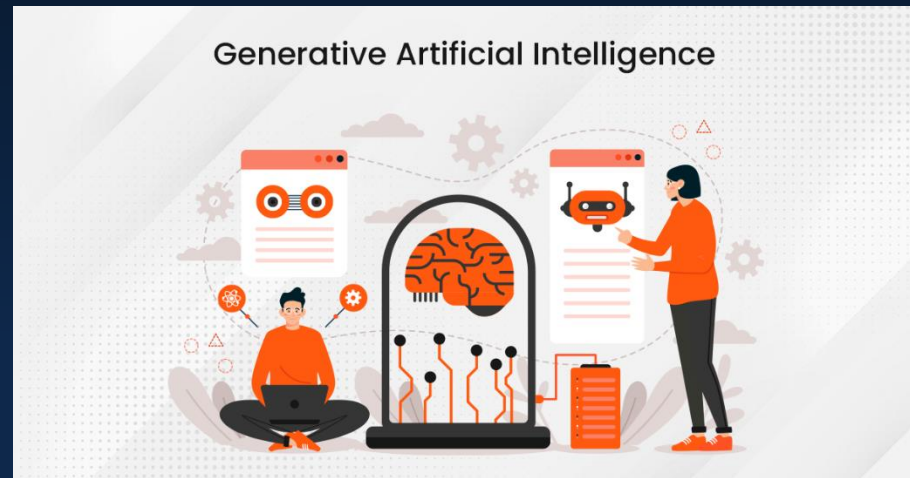


<https://levity.ai/blog/difference-machine-learning-deep-learning>

Qu'est-ce que l'Intelligence Artificielle Générative ?

L'intelligence artificielle générative est une intelligence artificielle capable de générer du texte, des images ou d'autres médias à l'aide de modèles génératifs.

Les modèles d'intelligence artificielle générative apprennent les modèles et la structure de leurs données d'entraînement et génèrent ensuite de nouvelles données présentant des caractéristiques similaires.

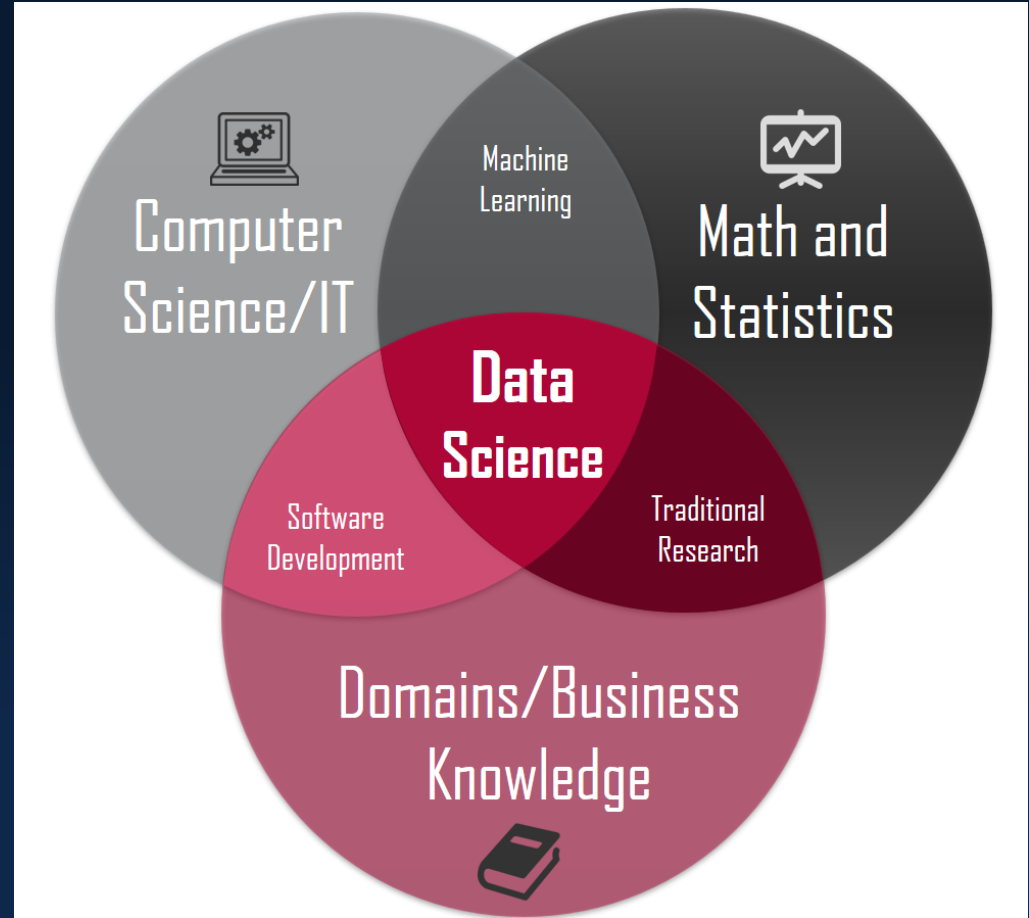


Qu'est-ce que la science des données ?

La science des données est l'étude de l'extraction automatisée de connaissance à partir de grands ensembles de données.

Plus précisément, la science des données est un domaine interdisciplinaire qui utilise des méthodes, des processus, des algorithmes et des systèmes scientifiques pour extraire des connaissances et des idées à partir de nombreuses données structurées ou non . Elle est souvent associée aux données massives et à l'analyse des données.

Elle utilise des techniques et des théories tirées de nombreux domaines dans le contexte des mathématiques, des statistiques, de l'informatique, de la théorie et des technologies de l'information, parmi lesquelles : l'apprentissage automatique, la compression de données et le calcul à haute performance.

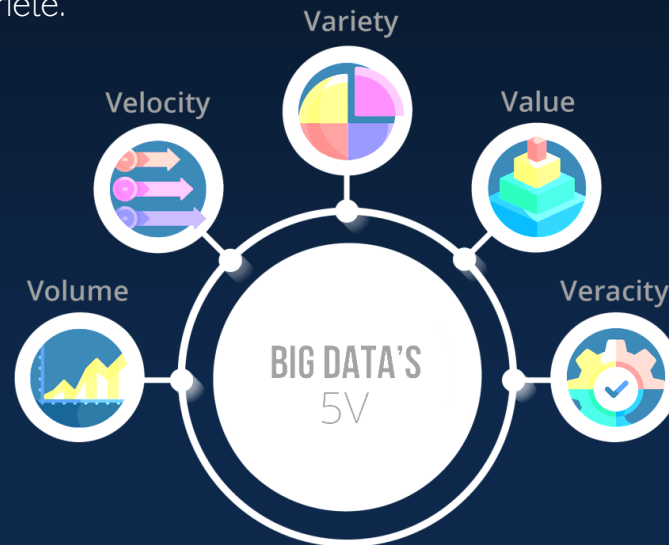


Qu'est-ce que le Big Data ?

On parle depuis quelques années du phénomène de big data , que l'on traduit souvent par « données massives ». Avec le développement des nouvelles technologies, d'internet et des réseaux sociaux ces vingt dernières années, la production de données numériques a été de plus en plus nombreuse : textes, photos, vidéos, etc.

Le gigantesque volume de données numériques produites combiné aux capacités sans cesse accrues de stockage et à des outils d'analyse en temps réel de plus en plus sophistiqués offre aujourd'hui des possibilités inégalées d'exploitation des informations.

Les ensembles de données traités correspondant à la définition du big data répondent à trois caractéristiques principales : volume, vitesse et variété.



<https://www.cnil.fr/fr/definition/big-data>

Qu'est-ce que le Data Mining ?

« Le processus non trivial d'identification de modèles valides, nouveaux, potentiellement utiles et finalement compréhensibles dans les données. »

Fayyad, Shapiro et Smyth 1996



Méthode d'apprentissage automatique : une illustration

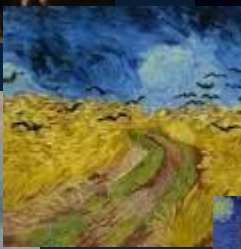
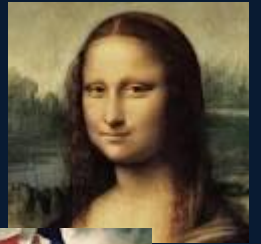
Exemple d'Apprentissage Automatique Supervisé

Apprendre à l'ordinateur à discerner une peinture :

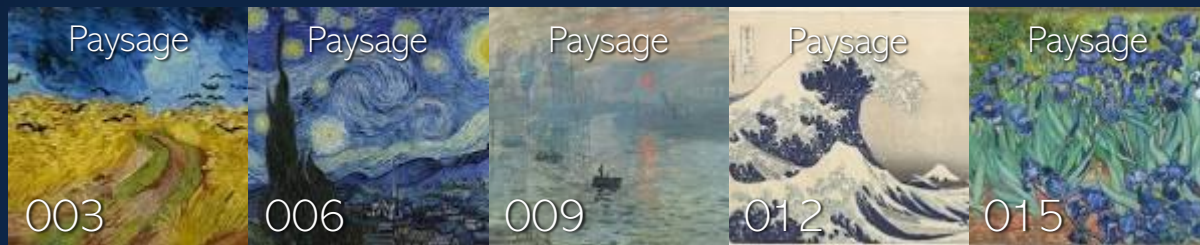
- de portrait,
- de personnages,
- ou de paysage.



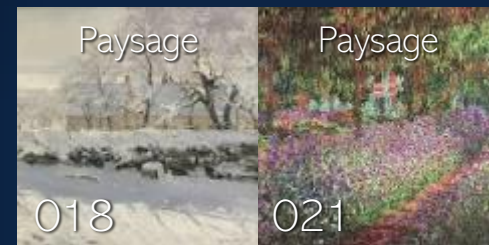
Collecte des données



Apprentissage



Test

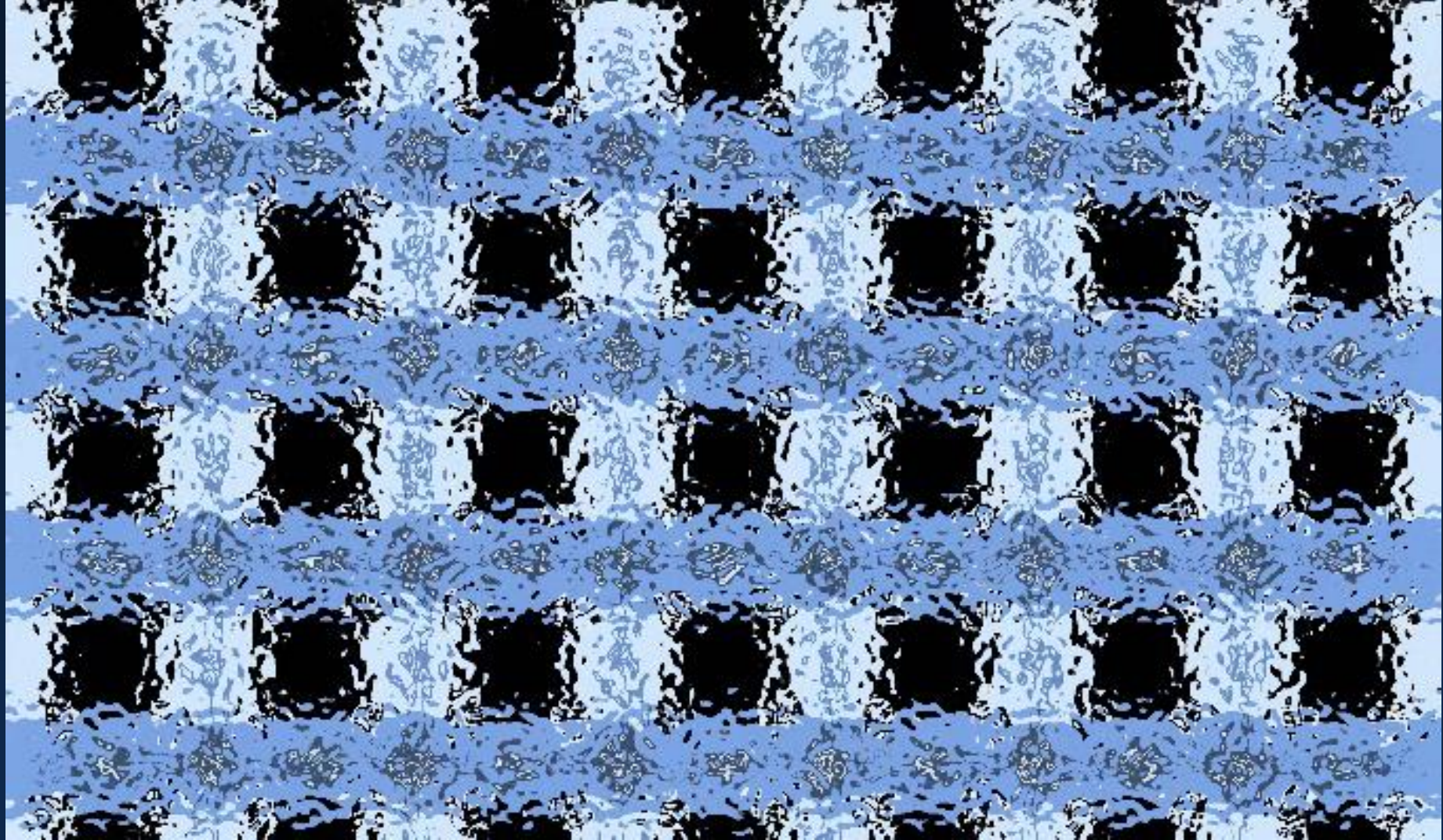




		ID	#Yeux	#Nez	#Bouches	#Visages	%C	%M	%J	%N	Classe
Apprentissage		001	2	1	1	1	10	10	40	40	Portrait
		002	2	4	4	10	20	20	20	40	Personnages
		003	0	0	0	0	35	15	40	10	Paysage
		004	2	1	1	1	10	10	10	70	Portrait
		005	10	5	5	11	10	10	10	70	Personnages
		006	2	0	0	0	55	15	20	10	Paysage
		007	2	1	1	1	10	10	70	10	Portrait
		008	11	6	6	7	10	10	40	40	Personnages
		009	0	0	0	0	45	15	30	10	Paysage
		010	2	1	1	1	10	10	40	40	Portrait
		011	5	3	3	3	20	10	30	40	Personnages
		012	0	0	0	0	35	15	40	10	Paysage
		013	2	1	1	0	10	10	40	40	Portrait
		014	20	10	10	10	20	20	20	40	Personnages
		015	0	0	0	0	35	15	40	10	Paysage
Test		016	2	1	1	1	10	10	40	40	Portrait
		017	8	4	4	10	20	20	20	40	Personnages
		018	1	0	0	0	33	33	33	1	Paysage
		019	2	1	1	1	35	15	40	10	Portrait
		020	2	2	2	2	1	4	20	75	Personnages
		021	0	0	0	0	40	10	40	10	Paysage



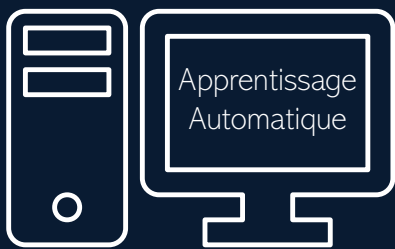
"In my illustration Mysterious Island, there is a seashore with a sail pushed along by the wind. But if the image is turned upside down, a portrait of Jules Verne, my favourite childhood author, appears."
István Orosz


















Abstraction
Et
Réduction

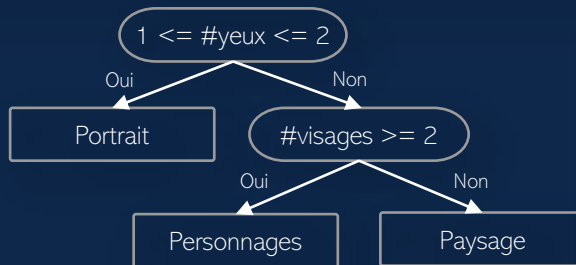
	ID	#Yeux	#Nez	#Bouches	#Visages	%C	%M	%J	%N	Classe
Apprentissage	001	2	1	1	1	10	10	40	40	Portrait
	002	2	4	4	10	20	20	20	40	Personnages
	003	0	0	0	0	35	15	40	10	Paysage
	004	2	1	1	1	10	10	10	70	Portrait
	005	10	5	5	11	10	10	10	70	Personnages
	006	2	0	0	0	55	15	20	10	Paysage
	007	2	1	1	1	10	10	70	10	Portrait
	008	11	6	6	7	10	10	40	40	Personnages
	009	0	0	0	0	45	15	30	10	Paysage
	010	2	1	1	1	10	10	40	40	Portrait
	011	5	3	3	3	20	10	30	40	Personnages
	012	0	0	0	0	35	15	40	10	Paysage
	013	2	1	1	0	10	10	40	40	Portrait
	014	20	10	10	10	20	20	20	40	Personnages
	015	0	0	0	0	35	15	40	10	Paysage
Test	016	2	1	1	1	10	10	40	40	Portrait
	017	8	4	4	10	20	20	20	40	Personnages
	018	1	0	0	0	33	33	33	1	Paysage
	019	2	1	1	1	35	15	40	10	Portrait
	020	2	2	2	2	1	4	20	75	Personnages
	021	0	0	0	0	40	10	40	10	Paysage

	ID	Descripteur 1	Descripteur 2	Classe
Apprentissage	001	0,26	0,83	Portrait
	002	0,40	0,77	Personnages
	003	0,66	0,16	Paysage
	004	0,46	0,11	Portrait
	005	0,09	0,63	Personnages
	006	0,03	0,66	Paysage
	007	0,01	0,26	Portrait
	008	0,73	0,99	Personnages
	009	0,81	0,01	Paysage
	010	0,13	0,75	Portrait
	011	0,95	0,29	Personnages
	012	0,89	0,22	Paysage
	013	0,17	0,36	Portrait
	014	0,06	0,53	Personnages
	015	0,66	0,88	Paysage
Test	016	0,95	0,78	Portrait
	017	0,82	0,39	Personnages
	018	1,00	0,59	Paysage
	019	0,54	0,38	Portrait
	020	0,53	0,99	Personnages
	021	0,04	0,44	Paysage

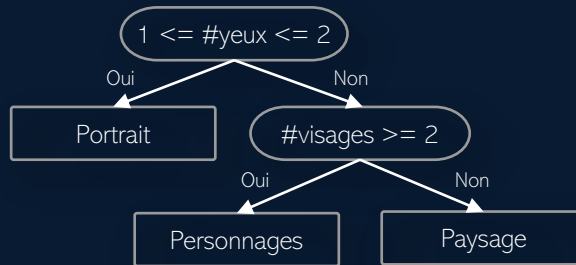


	ID	#Yeux	#Nez	#Bouches	#Visages	%C	%M	%J	%N	Classe
Apprentissage	 001	2	1	1	1	10	10	40	40	Portrait
	 002	2	4	4	10	20	20	20	40	Personnages
	 003	0	0	0	0	35	15	40	10	Paysage
	 004	2	1	1	1	10	10	10	70	Portrait
	 005	10	5	5	11	10	10	10	70	Personnages
	 006	2	0	0	0	55	15	20	10	Paysage
	 007	2	1	1	1	10	10	70	10	Portrait
	 008	11	6	6	7	10	10	40	40	Personnages
	 009	0	0	0	0	45	15	30	10	Paysage
	 010	2	1	1	1	10	10	40	40	Portrait
	 011	5	3	3	3	20	10	30	40	Personnages
	 012	0	0	0	0	35	15	40	10	Paysage
	 013	2	1	1	0	10	10	40	40	Portrait
	 014	20	10	10	10	20	20	20	40	Personnages
	 015	0	0	0	0	35	15	40	10	Paysage

Apprentissage d'un
Modèle de classification



Si (1 <= #yeux <= 2) alors Classe = Portrait
 Sinon Si (#visages >= 2) alors Classe = Personnages
 Sinon Classe = Paysage
 Fin Si
 Fin Si



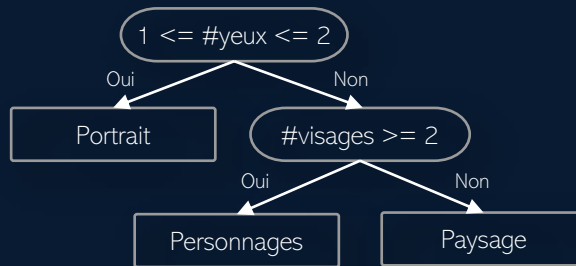
Si (1 <= #yeux <= 2) alors Classe = Portrait
 Sinon Si (#visages >= 2) alors Classe = Personnages
 Sinon Classe = Paysage
 Fin Si

Fin Si

		ID	#Yeux	#Nez	#Bouches	#Visages	%C	%M	%J	%N	Classe	Modèle
Apprentissage		001	2	1	1	1	10	10	40	40	Portrait	Portrait
		002	2	4	4	10	20	20	20	40	Personnages	Portrait
		003	0	0	0	0	35	15	40	10	Paysage	Paysage
		004	2	1	1	1	10	10	10	70	Portrait	Portrait
		005	10	5	5	11	10	10	10	70	Personnages	Personnages
		006	2	0	0	0	55	15	20	10	Paysage	Portrait
		007	2	1	1	1	10	10	70	10	Portrait	Portrait
		008	11	6	6	7	10	10	40	40	Personnages	Personnages
		009	0	0	0	0	45	15	30	10	Paysage	Paysage
		010	2	1	1	1	10	10	40	40	Portrait	Portrait
		011	5	3	3	3	20	10	30	40	Personnages	Personnages
		012	0	0	0	0	35	15	40	10	Paysage	Paysage
		013	2	1	1	0	10	10	40	40	Portrait	Portrait
		014	20	10	10	10	20	20	20	40	Personnages	Personnages
		015	0	0	0	0	35	15	40	10	Paysage	Paysage







Matrice De confusion

		Modèle		
		Portrait	Personnages	Paysage
Classe	Portrait	5	0	0
	Personnages	1	4	0
	Paysage	1	0	4



Si (1 <= #yeux <= 2) alors Classe = Portrait
 Sinon Si (#visages >= 2) alors Classe = Personnages
 Sinon Classe = Paysage
 Fin Si

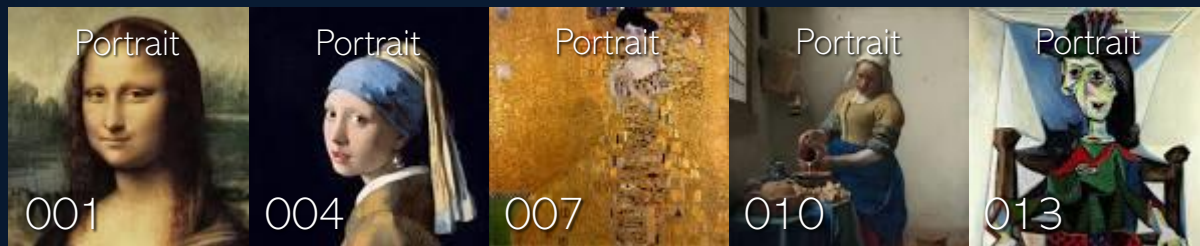
Fin Si

	ID	#Yeux	#Nez	#Bouches	#Visages	%C	%M	%J	%N	Classe	Modèle
Test	 016	2	1	1	1	10	10	40	40	Portrait	Portrait
	 017	8	4	4	10	20	20	20	40	Personnages	Personnages
	 018	1	0	0	0	33	33	33	1	Paysage	Portrait
	 019	2	1	1	1	35	15	40	10	Portrait	Portrait
	 020	2	2	2	2	1	4	20	75	Personnages	Portrait
	 021	0	0	0	0	0	40	10	40	10	Paysage

Matrice De confusion

		Modèle		
		Portrait	Personnages	Paysage
Classe	Portrait	2	0	0
	Personnages	1	1	0
	Paysage	1	0	1

Apprentissage



Portrait

Portrait

Portrait

Portrait

Portrait



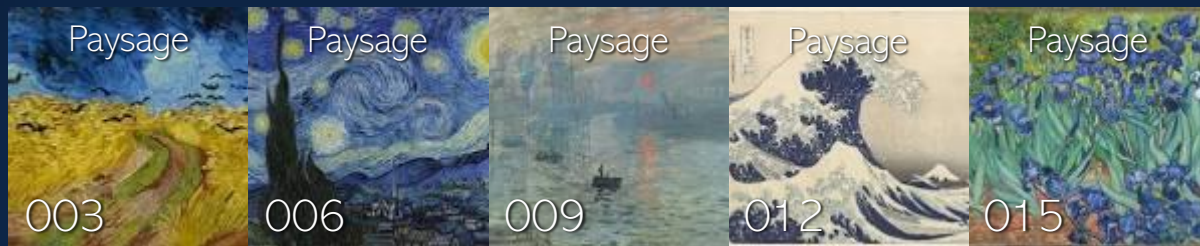
Portrait

Personnages

Personnages

Personnages

Personnages



Paysage

Portrait

Paysage

Paysage

Paysage

Test



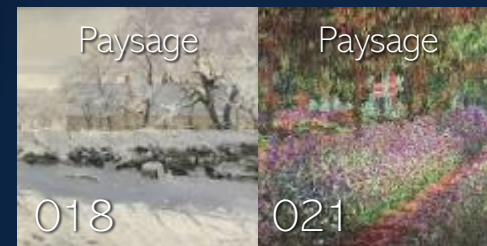
Portrait

Portrait



Personnages

Portrait



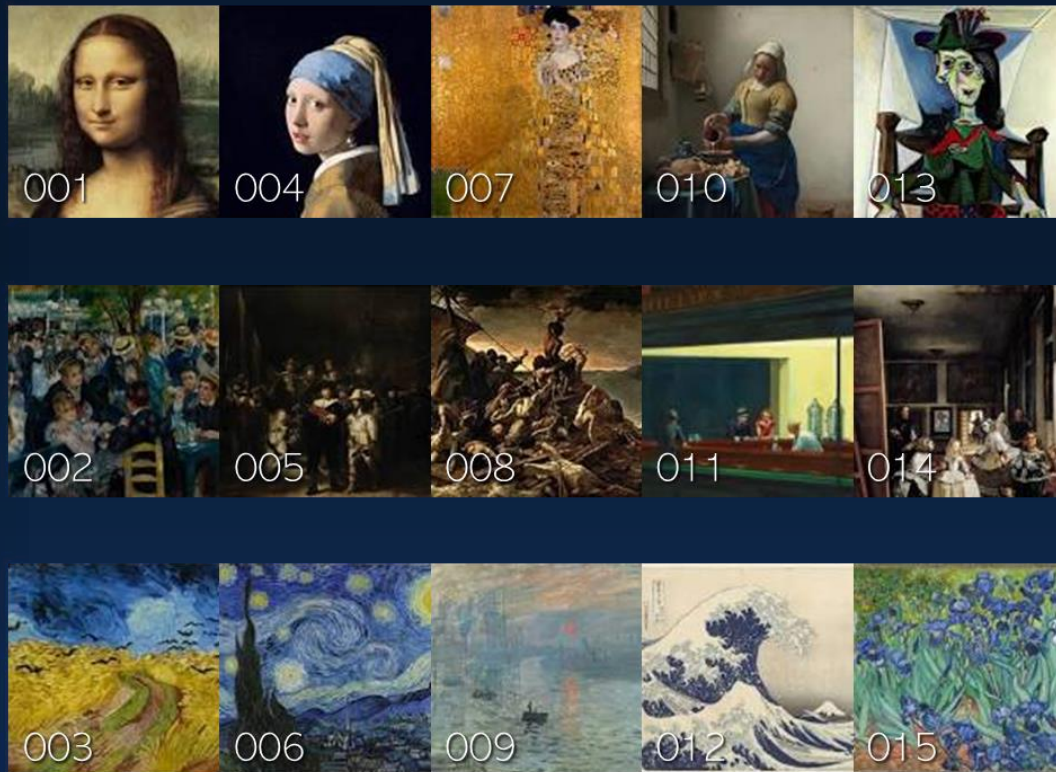
Portrait

Paysage

Exemple d'Apprentissage Automatique Non Supervisé

Créer des groupes d'œuvre d'art

Apprentissage
























Test



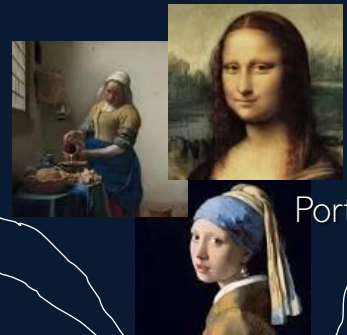
Exemple d'Apprentissage Automatique Non Supervisé

Créer des groupes d'œuvre d'art

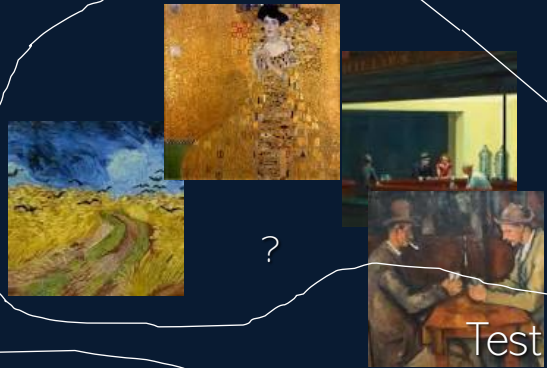
	ID	Descripteur 1	Descripteur 2
Apprentissage	 001	0,26	0,83
	 002	0,40	0,77
	 003	0,66	0,16
	 004	0,46	0,11
	 005	0,09	0,63
	 006	0,03	0,66
	 007	0,01	0,26
	 008	0,73	0,99
	 009	0,81	0,01
	 010	0,13	0,75
	 011	0,95	0,29
	 012	0,89	0,22
	 013	0,17	0,36
	 014	0,06	0,53
	 015	0,66	0,88
Test	 016	0,95	0,78
	 017	0,82	0,39
	 018	1,00	0,59
	 019	0,54	0,38
	 020	0,53	0,99
	 021	0,04	0,44

Descripteur 2

Descripteur 1



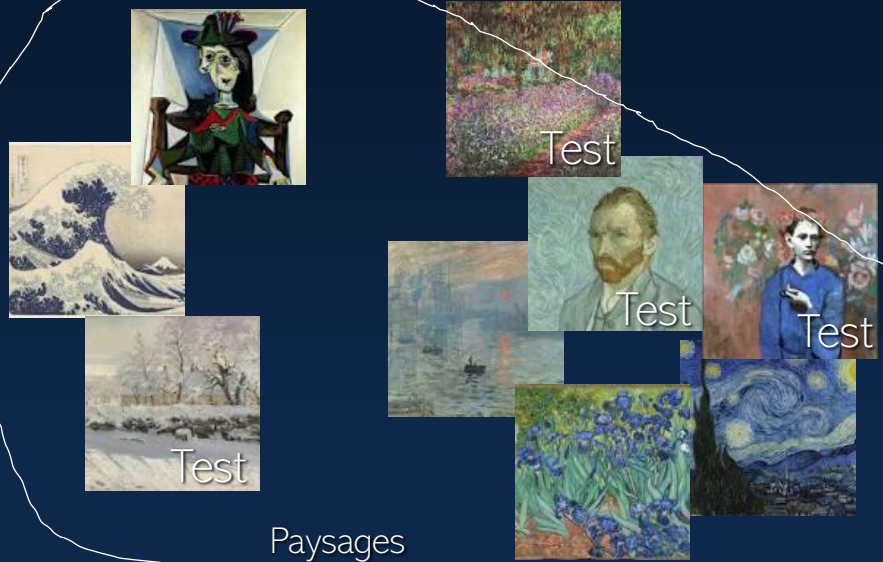
Portrait



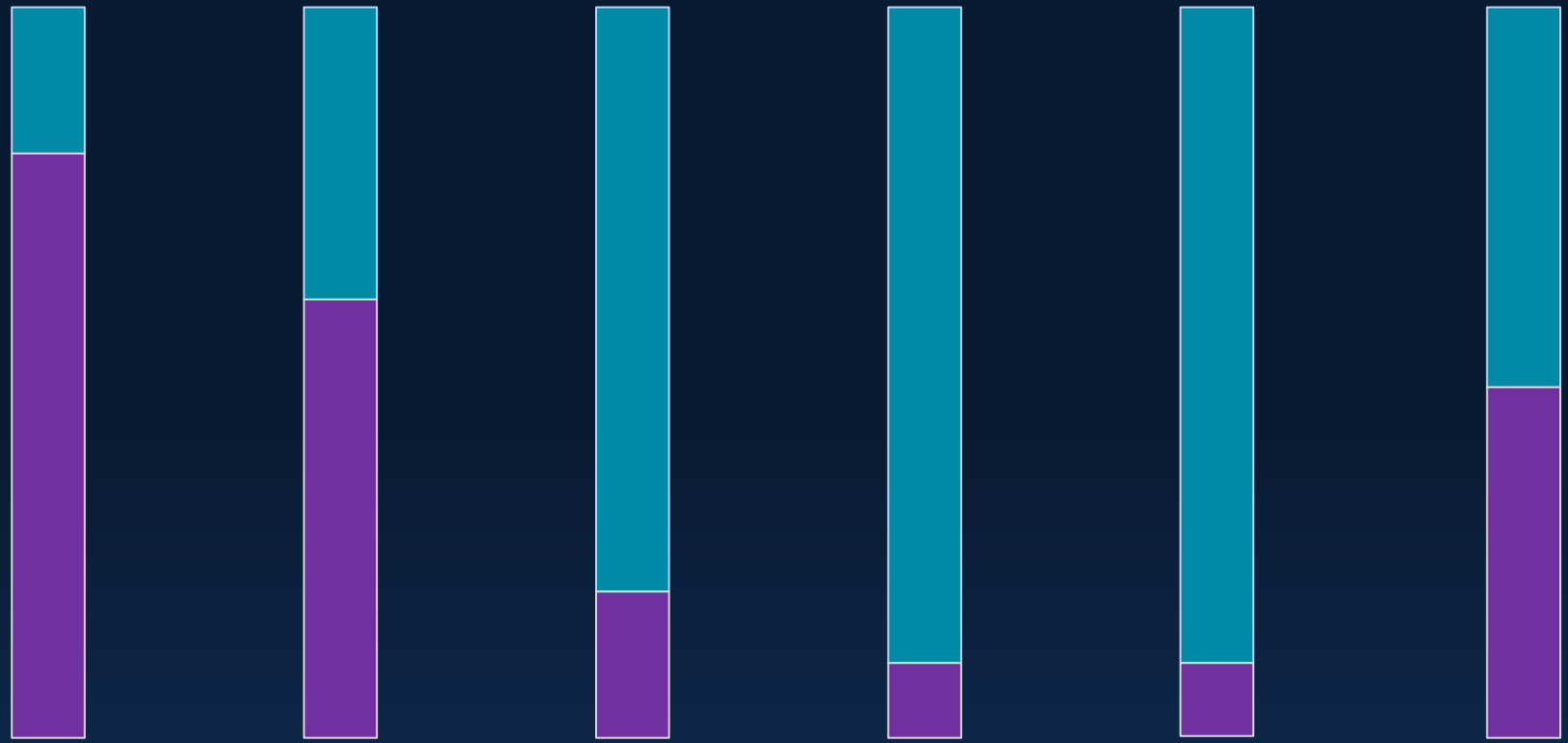
Test



Personnages



Paysages

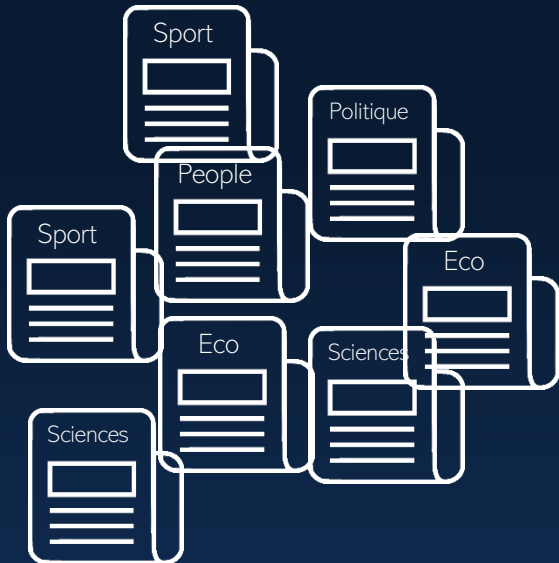


■ Experts Métier ■ Experts Apprentissage Automatique – Science des données

Exemples

Classification de textes - Recommandation

Articles de presse



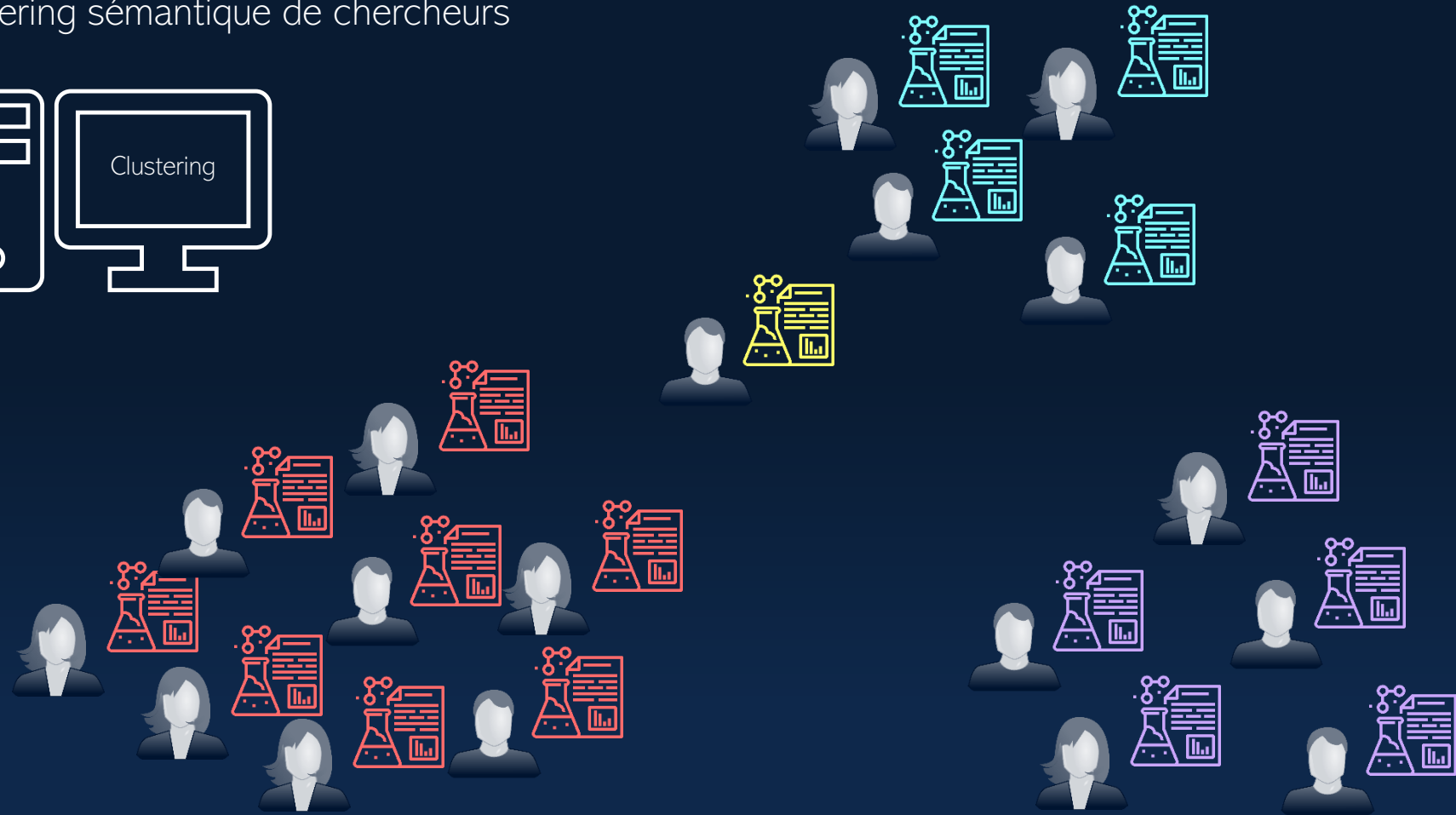
Startup

Classification de textes - Recommandation

Opinions



Clustering sémantique de chercheurs



Organismes de recherche

Classification d'échantillons sanguins

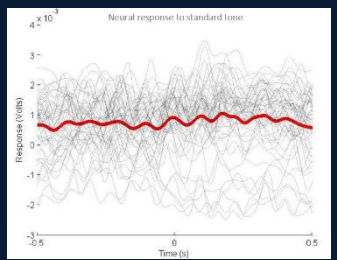


Grand groupe international

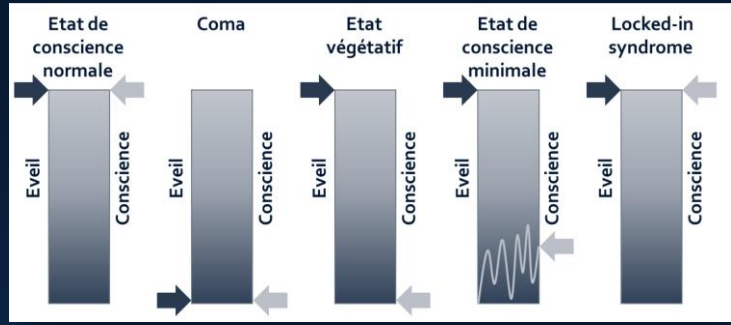
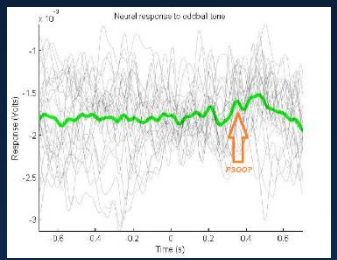
Classification de signaux cérébraux / états de conscience



Fréquent

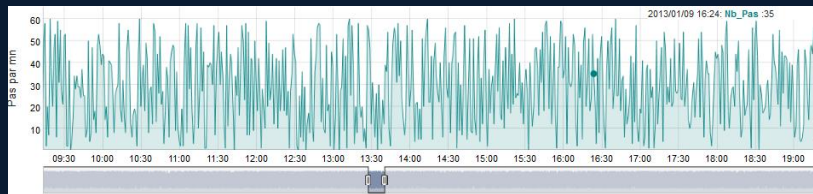
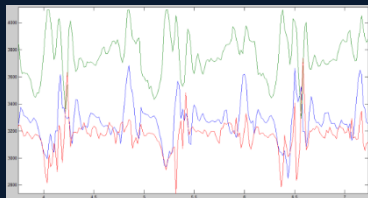


Rare

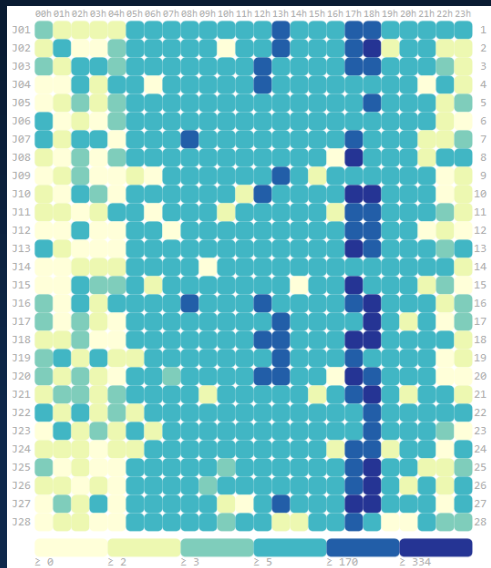


Suivi de l'activité physique

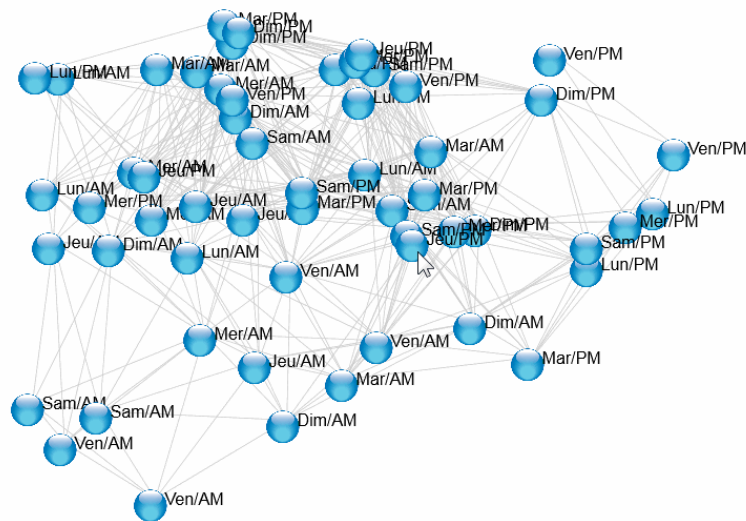
Données



Informations

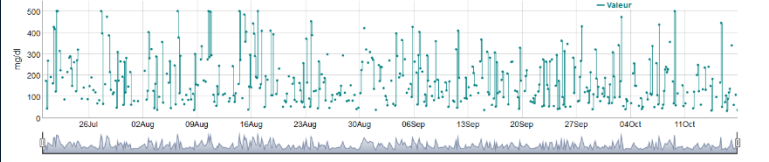


Connaissances

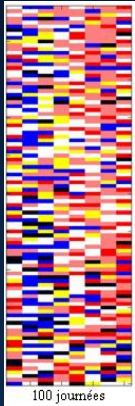


Suivi d'un patient diabétique

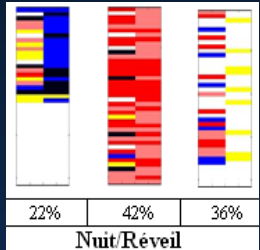
Données



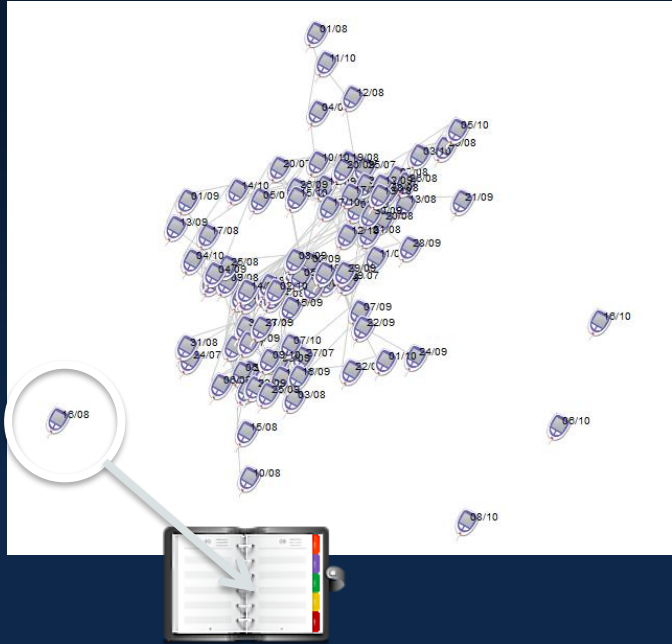
Informations



Clustering

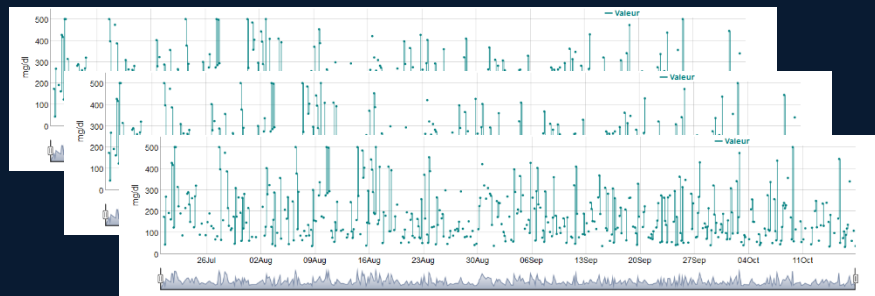


Connaissances

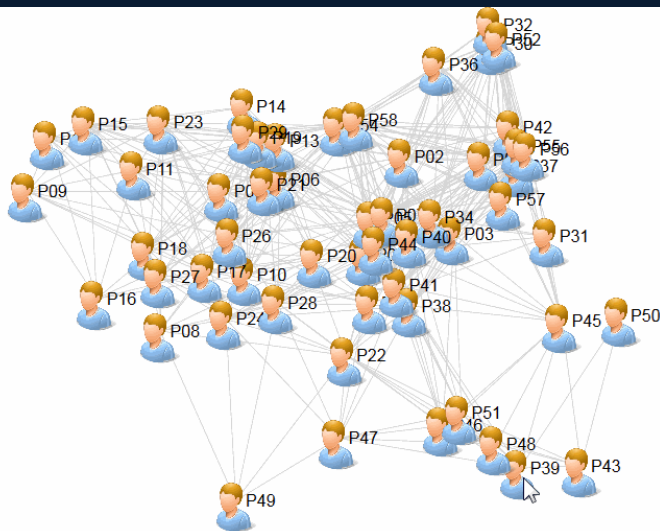


Suivi de patients diabétiques

Données




Informations



Connaissances

En guise de conclusion

A blue-tinted portrait of Jean Piaget, an elderly man with white hair and glasses, wearing a suit and tie. The image is centered on a dark blue background.

"L'intelligence ce n'est pas ce
que l'on sait mais ce qu'on l'on
fait quand on ne sait pas."

JEAN PIAGET

INTELLIGENCE
IS THE ABILITY TO
ADAPT TO CHANGE.

~

STEPHEN HAWKING



La Trahison des images
René Magritte 1929

Objectifs

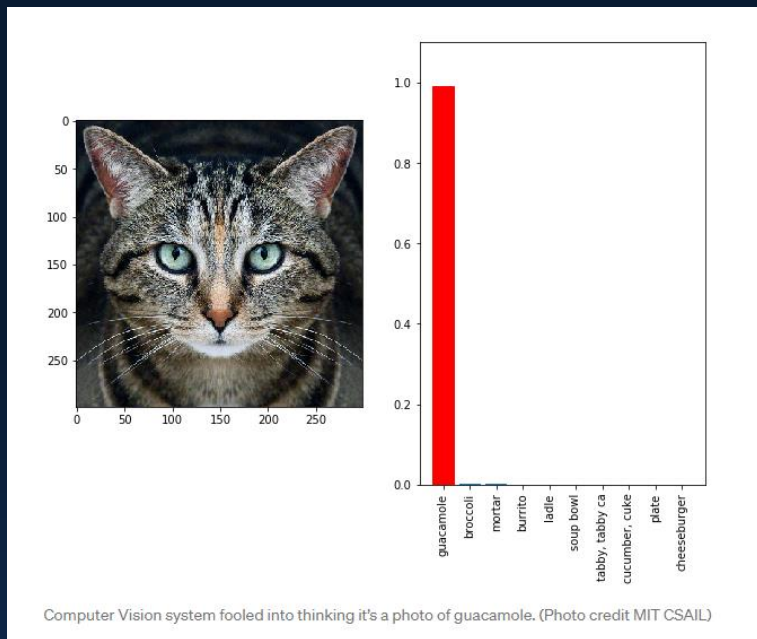
Être capable de comprendre :

- la définition de la science des données
- les définitions de l'Intelligence Artificielle
- la méthodologie d'Apprentissage Artificiel

Initier une réflexion sur l'usage de la science des données dans le domaine de la santé numérique

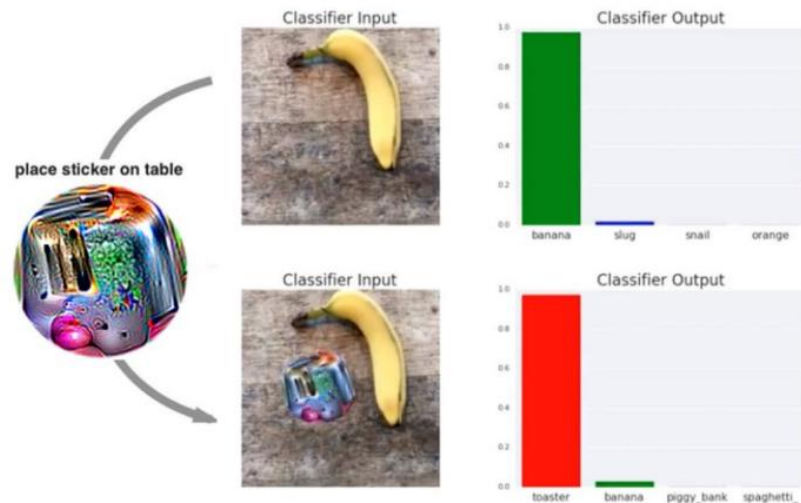
Pour aller plus loin

Les réseaux neuronaux se laissent facilement bernier



Adversarial Patch

A paper ([Adversarial Patch](#)) recently published by Google has lifted the bar. The paper shows how it is possible to show the system any image and it will classify it as a toaster.



Les réseaux neuronaux se laissent facilement bernier



Stop

(a) Normal



Yield



Speed Limit

(b) Attack

Automotive ML systems have been tricked into mistaking stop signs for speed limit signs

Explainable AI (XAI) - Interpretable AI - Explainable Machine Learning (XML),

https://en.wikipedia.org/wiki/Explainable_artificial_intelligence

<https://medium.com/swlh/explainable-vs-interpretable-ai-an-intuitive-example-6baf8fc6d402>

Explainable vs Interpretable AI: An Intuitive Example

Both explainable and interpretable AI are emerging topics in computer science. However, the difference between the two is not always obvious, even to academics. In this post I aim to provide some intuition using a simple example.



Les 5 débats de l'Intelligence artificielle

<https://www.opinionact.com/ressources/ia-les-5-debats-de-lintelligence-artificielle>

Débat philosophico-scientifique : qu'est-ce que l'intelligence ?

Une question pour le moins passionnelle est celle de savoir si l'on pourra un jour imiter, voire dépasser l'intelligence humaine. C'est le fameux point de singularité défendu par le controversé Ray Kurzweil, qui prédit que les IA dépasseront les intelligences humaines d'ici 2025.

Le directeur de l'ingénierie chez Google, Mark Zuckerberg, et d'autres scientifiques ont émis publiquement des doutes sur la possibilité de créer une intelligence artificielle capable de dépasser l'intelligence humaine non verbale mais aussi émotionnelle.

Le débat éthique : faut-il s'inquiéter du transhumanisme ?

La conviction des transhumanistes est qu'il faut passer d'une logique de réparation du corps humain à une logique d'amélioration de manière à éradiquer définitivement la maladie, voire le vieillissement.

En France, des personnalités comme le futurologue Joël de Rossetti ont fait passer l'homme au centre de son quotidien.

En France, des personnalités comme le futurologue Joël de Rossetti ont fait passer l'homme au centre de son quotidien.

Le débat économique : création ou destruction d'emplois ?

Robot avocat, journaliste, enquêteur, médecin, chauffeur... Toute tâche qui nécessite de traiter une grande masse d'informations à synthétiser, recouper pour identifier, calculer, déduire ou encore décider est inévitablement menacée de robotisation. Les dernières études estiment que 10% des emplois (tâches) de substitution par les intelligences artificielles.

Pour Daniel Cohen ou encore Jeremy Bentham, l'intelligence artificielle est une « *general purpose technology* ». Il estime donc que ce n'est pas la technologie elle-même qui créera des emplois mais la demande qu'elle génère.

Le débat juridique : faut-il une personnalité juridique à l'IA ?

Les partisans de la création d'une personnalité juridique font le parallèle avec la personnalité morale des entreprises, qui n'exclut pas celle des personnes physiques. Pour d'autres, faire payer les robots, c'est risquer de faire l'économie de l'enquête. Or la vraie question semble être aujourd'hui de savoir si l'IA a une conscience.

Laurence Devillers édicte 11 commandements pour les IA : la traçabilité des algorithmes (« *tu pourras être contrôlé* ») et leur contrôle (« *tu seras appris* »).

Le débat sociétal : que devient le travail dans un monde de robots ?

La fragmentation du travail intellectuel en tâches, déjà entamée par l'ubérisation via le modèle des plateformes, va s'accroître avec la robotisation. Les experts s'accordent à dire que la logique de poste de travail où des salariés en CDI sont rémunérés en fonction du nombre d'heures accomplies n'a plus de sens dans ce monde post-révolution numérique. Les métiers vont devoir tous évoluer pour apprendre à travailler avec l'IA, d'où un enjeu crucial de formation et d'accompagnement des salariés.

Jean-Gabriel Ganascia évoque une évolution du travail au sens initial « *labor* » vers la notion chère à Hannah Arendt de travail « *œuvre* », où l'homme va exercer toute sa puissance créatrice et émotionnelle. Identifier, reconnaître les différentes formes de contribution sociétale (accompagner un proche malade, être membre d'une association, prendre des nouvelles des enfants d'un client...) et les valoriser seront des sujets majeurs, malheureusement trop vite évacués par la solution politique d'un revenu universel.

L'Éthique de l'IA : Questions & Dilemmes

<https://octopeek.com/fr/lethique-de-lia-questions-dilemmes/>

Le Biais Systémique dans les Modèles d'IA

L'un des plus grands dilemmes de l'IA est celui des biais systémiques. Souvent des décisions sont prises sans que l'IA ne raisonne, mais elle est entraînée sur des données qui reflètent les préjugés de la société. En pratique, un algorithme de recommandation (entrée, sortie) affiche des contenus qui ne sont pas ceux que vous souhaitez. C'est-à-dire de données biaisées. Les biais découverts dans les données utilisées pour l'entraînement des modèles d'IA.

Pour illustrer ce comportement biaisé, les "asian girls" sont souvent utilisées dans les médias. Un autre exemple est la définition de la performance lorsqu'il s'agit d'identifier une anomalie dévastatrice dans les vidéos de l'homme justifiant sa décision.

Le Dilemme du Tramway dans les Véhicules Autonomes

Alors que de plus en plus de véhicules autonomes sont sur les routes, le dilemme du tramway se pose. L'analogie d'un trolley sur une voie à double sens qui doit déraper pour éviter de tuer cinq personnes ou deux.

La Créativité Artificielle

Il existe une catégorie d'algorithmes de nouvelles données selon une nomenclature qui n'a pas toujours été explorée, ouvrant ainsi la voie à l'émergence de nouvelles formes d'art.

Dans le monde des arts, une IA est utilisée pour générer des œuvres d'art. En 2016, un artiste a créé une œuvre intitulée "The Next Rembrandt", surnommée « the Next Rembrandt », qui a été vendue pour plusieurs millions de dollars.

Une fois l'effet de surprise estomacé, la question se pose : l'IA est-elle un artiste ? Ou est-ce l'artiste qui utilise l'IA ? En 2019, l'Office Européen des brevets a désigné comme l'inventeur de la propriété intellectuelle, et nous pouvons actualiser notre compréhension de ce qui est concerné : l'IA, les personnes.

En outre, l'art de l'IA est-il une limite de la créativité humaine ? Comprendre la nature de l'implicite et préserver la valeur de l'art dans un monde où l'IA est omniprésente.

Accusé, levez-vous! L'AI dans le Tribunal Désinformation et Fake News

Un autre cas intéressant d'application de l'IA se trouve dans les algorithmes pour évaluer les cas, remplaçant essentiellement la technologie de l'intelligence artificielle dans des domaines à valeur tangible, principalement sous la forme d'un jugement éthique qui méritent d'être explorés.

Le système judiciaire est lent, et cela peut être attribué à des processus juridiques, le volume élevé de dossiers juridiques et quelques-uns. Cela crée une opportunité pour l'utilisation d'automatisation des tâches répétitives et chronophages pour le premier juge artificiel ; un personnage féminin numérique doté de capacités faciales pour gérer les tâches répétitives de base dans un processus automatisé.

Les algorithmes et les outils d'analyse de données sont également utilisés pour accélérer la recherche juridique et miner des informations pertinentes. Ce n'est pas les moyens d'engager un avocat, DoNotPay est là pour aider. Un autre problème imposant est le manque de transparence souvent considérés comme une boîte noire. Même si ces résultats peuvent être expliqués, leur compréhension pose problème, car ils ne permettent pas de faire aveuglément confiance à une machine lors d'un jugement sans comprendre quels facteurs ont contribué à la décision. Enfin, l'entraînement de ces modèles s'appuie sur des protocoles de données massifs, ce qui soulève également de

L'un des arguments les plus importants utilisés pour défendre la capacité de cette technologie à être "neutre" dans son jugement est qu'elle souffre de biais intégrés et de corrélation avec nos propres préjugés, comme nous l'avons vu. Un autre problème imposant est le manque de transparence souvent considérés comme une boîte noire. Même si ces résultats peuvent être expliqués, leur compréhension pose problème, car ils ne permettent pas de faire aveuglément confiance à une machine lors d'un jugement sans comprendre quels facteurs ont contribué à la décision. Enfin, l'entraînement de ces modèles s'appuie sur des protocoles de données massifs, ce qui soulève également de

« It is only when they go wrong that machines remind you how powerful they are. »
Clive James

La désinformation a toujours été utilisée comme une arme politique pour manipuler les peuples. Aujourd'hui, elle est encore amplifiée par l'essor des modèles spécialisés d'apprentissage profond qui peuvent produire un nouveau texte, trafiquer des images, ou même falsifier des discours. Ces algorithmes ont atteint un niveau d'avancement où un humain parviendrait difficilement à distinguer leurs créations de signaux réels.

Le modèle GPT-3 d'OpenAI, qui signifie Generative Pre-trained Transformer, est un énorme réseau de neurones qui a été entraîné sur un corpus massif de textes afin de capturer les régularités statistiques du langage naturel. Le modèle pouvait ensuite être facilement utilisé pour générer des paragraphes de texte après avoir reçu quelques mots initiaux.

Il n'a pas fallu longtemps pour que GPT-3 soit exploité pour produire toutes sortes de fake news, les unes aussi plausibles que les autres, ce qui a déclenché de nombreuses sonnettes d'alarme concernant cette horrible utilisation de la technologie et ses éventuelles répercussions.

En ce qui concerne les images et les vidéos, un développement récent de la recherche sur l'apprentissage profond a permis de falsifier des vidéos en remplaçant le visage d'une personne source dans une vidéo par le visage d'une personne cible dans une image.

Cette technologie, appelée DeepFake, nous permet essentiellement de mettre de nouveaux mots dans la bouche d'une personnalité publique, de jouer la star dans notre film préféré ou d'accomplir toute autre application que vous pouvez imaginer. Et si ce n'était pas suffisant, elle peut également s'étendre aux signaux audio, ce qui permet de tromper facilement des personnes au téléphone.

En 2019, un criminel a utilisé un deepFake pour imiter la voix du PDG d'une entreprise énergétique allemande afin d'exiger un virement bancaire frauduleux de 220 000 €, dans ce qui serait le premier cybercrime de l'histoire basé sur l'IA. Bien que techniquement intrigants, les DeepFakes introduisent une nouvelle série de problèmes sociaux, car ils discréditent complètement l'ouïe et la vue, les deux sens les plus fiables sur lesquels les humains s'appuient pour prendre des décisions.

Ces modèles d'apprentissage profond peuvent être utilisés contre les femmes en intégrant synthétiquement leur visage dans du contenu pornographique, piéger quelqu'un en le plaçant sur la scène d'un crime qu'il n'a pas commis ou diffuser une fausse déclaration de guerre du président d'une nation puissante.

La prolifération de la désinformation sur l'internet est devenue un fléau alarmant qui exige une attention immédiate. Le processus fastidieux de création de faux contenus multimédias, autrefois réservé aux experts, est désormais accessible au public et sans aucun pré-requis. Les plateformes de réseaux sociaux sont parmi les parties les plus concernées, car elles constituent le principal moyen de communication incontrôlé qui permet aux fake news de devenir virales. C'est pourquoi Google, Facebook, Twitter et d'autres géants de la technologie mènent une guerre contre la désinformation et la propagande dangereuse, mais leurs efforts ont jusqu'à présent eu des effets limités.