



INSTITUT
de MATHÉMATIQUES
de TOULOUSE



Big Data, Machine Learning : qu'est-ce que la science des données ?

Journée de l'IREM, Bordeaux

Aurélien Garivier

18 janvier 2017

Institut de Mathématiques de Toulouse

LabeX CIMI

Université Paul Sabatier

1. Un monde numérique
2. Qu'est-ce que le machine learning ?
3. Le phénomène Big Data
4. Quel Machine Learning pour les Big Data?
5. Les Big Data, le Machine learning et la loi
6. Enseignement des sciences des données

Un monde numérique

Nous vivons aujourd'hui dans une "ère du numérique":

- Les données sont partout : sondages, indices de popularité, scores électoraux, statistiques économiques voire sportives...
- Nous produisons nous-mêmes des données de notre plein gré : photos, interactions sur les réseaux sociaux, objets connectés...
- sans le vouloir vraiment : recherches sur Internet, traces laissées par nos actions (achats, réseaux sociaux ...),
- ou en échange de services (pas si) gratuits.

Notre vie est de plus en plus dirigée par ce monde numérique :

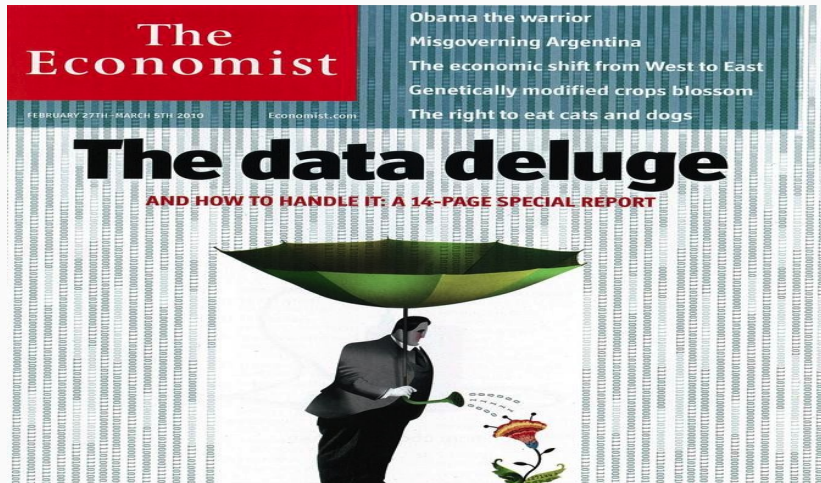
- Société de consommation (prédiction des goûts et des achats)
- Notre relation avec les banques, assurances (scoring pour déterminer crédit, avantages)
- Algorithmes pour la carte scolaire, pour l'orientation scolaire...
- Surveillance policière...
- Choix d'un conjoint par sites de rencontre...
- Voitures qui roulent seule et réagissent à leur environnement ...

Tout cela grâce avec des algorithmes...

Un monde numérique : le mythe du Big Data

Des chiffres tous les jours dans tous les media, chiffres sur lesquels s'appuie la pensée :

Economie quantitative, psychologie quantitative mais aussi décisions politiques.





- Une donnée quantitative (un nombre) ça ne se discute pas
- c'est sérieux
- c'est facile à interpréter
- c'est synthétique ..



Bref ... **c'est efficace !**

- Une donnée quantitative (un nombre) ça ne se discute pas
- c'est sérieux
- c'est facile à interpréter
- c'est synthétique ..

Peut-on faire mentir les nombres ?

- Disraeli : Il y a les mensonges, les gros mensonges, et les statistiques.

Par manque de temps , de connaissances, on s'expose à des erreurs d'interprétation.

Nécessité de comprendre l'usage de la statistique et leur signification : la vérité mathématique est-elle vérité au sens usuel ?

- HG. Wells : Le jugement statistique sera un jour aussi nécessaire à l'exercice de base des fonctions du citoyen que la capacité de lire et d'écrire.

- La statistique est l'étude de la collecte de données, leur analyse, leur traitement, l'interprétation des résultats et leur présentation afin de rendre les données compréhensibles par tous. C'est à la fois une science, une méthode et un ensemble de techniques.

Source : Wikipedia

- La statistique est utilisée dans presque tous les domaines de l'activité humaine : sciences sociales, économie, médecine, biologie, industrie ...
- Mais aujourd'hui, Open Data, Big Data, Big Science... changent les règles du jeu.

Qu'est-ce que le machine learning ?

1. Un monde numérique
2. Qu'est-ce que le machine learning ?
 - Intelligence artificielle
 - Machine Learning
3. Le phénomène Big Data
4. Quel Machine Learning pour les Big Data?
5. Les Big Data, le Machine learning et la loi
6. Enseignement des sciences des données

Intelligence Artificielle (IA) : définition

Intelligence des machines

- simuler les capacités cognitives des humains (big data: les humains apprennent en utilisant des sources de données très abondantes et diverses).
- une machine mime les fonctions cognitives que les humains associent à l'esprit humain, tels que *apprendre* ou *résoudre un problème*.

Machine intelligente idéale =

agent rationnel flexible qui *perçoit* son environnement et qui *prend des décisions* qui maximisent ses chances de succès pour un but donné.

Fondé sur le postulat que l'intelligence humaine

peut être décrite si précisément qu'on peut construire une machine la simulant.

Buts opérationnels

- Robots autonomes pour réaliser des tâches pas trop spécialisées
- En particulier, vision + compréhension et production de langage (naturel)

Tension entre les objectifs opérationnels et les buts philosophiques

- Au fur et à mesure que les machines accomplissent de plus en plus de tâches, des compétences qu'on pensait relever de l'*intelligence* sont progressivement retirées de la liste. Par exemple, la reconnaissance de caractères n'est plus considérée comme relevant de l'IA, mais comme une technologie de routine.
- Parmi les compétences encore classées en IA, il y a le jeu de go ou les voitures autonomes...

AI: principaux thèmes

Principaux objectifs de l'IA:

- raisonnement
- connaissance
- planification
- apprentissage
- traitement des langues naturelles
- perception
- intelligence "générale"

Approches centrales de l'IA:

- approche symbolique traditionnelle (cf. logique)
- méthodes d'inspiration statistique
- soft computing

S'appuie sur:

- informatique
- mathématiques
- linguistique
- philosophie
- neurosciences
- psychologie (artificielle)

Outils:

- optimisation mathématique
- logique
- algorithmes d'inspiration probabiliste

1. Un monde numérique
2. Qu'est-ce que le machine learning ?
 - Intelligence artificielle
 - Machine Learning
3. Le phénomène Big Data
4. Quel Machine Learning pour les Big Data?
5. Les Big Data, le Machine learning et la loi
6. Enseignement des sciences des données

Machine Learning (ML): Définition

Arthur Samuel (1959)

Champ d'étude qui donne aux ordinateurs la capacité d'apprendre sans avoir été programmés explicitement

Tom M. Mitchell (1997)

On dit qu'un programme *apprend d'une expérience E* par rapport à une classe de tâches T et à une mesure de performance P si sa performance sur T, mesurée par P, augmente avec l'expérience E.

ML: Apprendre des données et faire des prédictions

- Les algorithmes construisent un modèle à partir d'*exemples* donnés en entrée, dans but de faire des *prédictions* ou de prendre des *décisions*...
- ...plutôt que de suivre strictement une suite statique d'instructions : c'est utile quand il serait impossible ou inefficace de concevoir et de programmer de tels algorithmes.

Analyse de données (Data Analytics)

- Le Machine Learning est utilisé pour concevoir des modèles complexes et des algorithmes qui conduisent eux-même à des prédictions - le mot commercial est souvent *predictive analytics*.
- www.sas.com: "Produce reliable, repeatable decisions and results" and uncover "hidden insights" through learning from historical **relationships and trends** in the data.
- évolution à partir de la reconnaissance de motifs (pattern recognition) de la computational learning theory en IA.

Machine Learning: problèmes-types

- filtrage de spams, classification de textes
- reconnaissance de caractères (OCR)
- moteurs de recherche
- plateformes de recommandation
- outils de reconnaissance de la parole
- vision par ordinateur
- bio-informatique, analyse du génome, médecine (prédictive)

Pour chacune de ces tâches, il est possible mais inefficace d'écrire des programmes explicitement destinés à résoudre les buts recherchés.

Il apparaît beaucoup plus fécond d'apprendre à des machines à inférer elles-même les bonnes règles de décision.

- **Statistique computationnelle**: centré sur la prédiction obtenue par l'usage de modèles statistiques nécessitant des calculs numériques intensifs (ex: méthodes bayésiennes)
- **Apprentissage statistique**: ML basé sur des méthodes statistiques, avec un point de vue statistique (garanties probabilistes: consistance, inégalités oracles, minimax...) → plus axés sur la *corrélacion*, et moins sur la *causalité*
- **Data Mining** (apprentissage non supervisé) centré plutôt sur l'analyse exploratoire des données et la découverte de propriétés inconnues des données.
- Importance des méthodes basées sur les **probabilités** et les **statistiques** → **Data Science** (Michael Jordan)
- Liens très forts avec l'**optimisation mathématique**, qui fournit des méthodes, des concepts et des applications au ML.

Classification supervisée : cadre statistique

Définition (terme anglais)	ex: reconnaissance de chiffres
Input space \mathcal{X}	64×64 images
Output space \mathcal{Y}	$\{0, 1, \dots, 9\}$
Joint distribution $P(x, y)$?
Prediction function $h \in \mathcal{H}$	
Risk $R(h) = P(h(X) \neq Y)$	
Sample $\{(x_i, y_i)\}_{i=1}^n$ Empirical risk $\hat{R}_n(h) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}\{h(x_i) \neq y_i\}$	MNIST dataset
Learning algorithm $\phi_n : (\mathcal{X} \times \mathcal{Y})^n \rightarrow \mathcal{H}$ Expected risk $R_n(\phi) = \mathbb{E}_n[R(\phi_n)]$	NN,boosting...
Empirical risk minimizer $\hat{h}_n = \arg \min_{h \in \mathcal{H}} \hat{R}_n(h)$ Regularized empirical risk minimizer $\hat{h}_n = \arg \min_{h \in \mathcal{H}} \hat{R}_n(h) + \lambda C(h)$	

Inégalité de Hoeffding: avec probabilité au moins $1 - \eta$,

$$|R(h) - \hat{R}_n(h)| \leq \sqrt{\frac{1}{2n} \log \left(\frac{2}{\eta} \right)}.$$

Problème: vrai pour *chaque* h fixé mais *pas* pour \hat{h}_n !

Ex: Prédiction of 10 lancers de Pile ou Face

Ex: régression polynomiale \rightarrow sur-apprentissage

Fléau de la dimension

Minimisation structurelle du risque

→ loi des grands nombres *uniforme* — inégalité de Vapnik-Chervonenkis :
si \mathcal{H} a une *dimension de VC* $d_{\mathcal{H}}$, alors

$$\sup_{h \in \mathcal{H}} |R(h) - \hat{R}_n(h)| \leq O \left(\sqrt{\frac{1}{2n} \log \left(\frac{2}{\eta} \right) + \frac{d_{\mathcal{H}}}{n} \log \left(\frac{n}{d_{\mathcal{H}}} \right)} \right) .$$

Structure:

$$\mathcal{H} = \bigcup_m \mathcal{H}_m$$

Ex: polynômes/splines de degré m , arbres de décision de profondeur m, \dots

Décomposition du risque en biais-variance

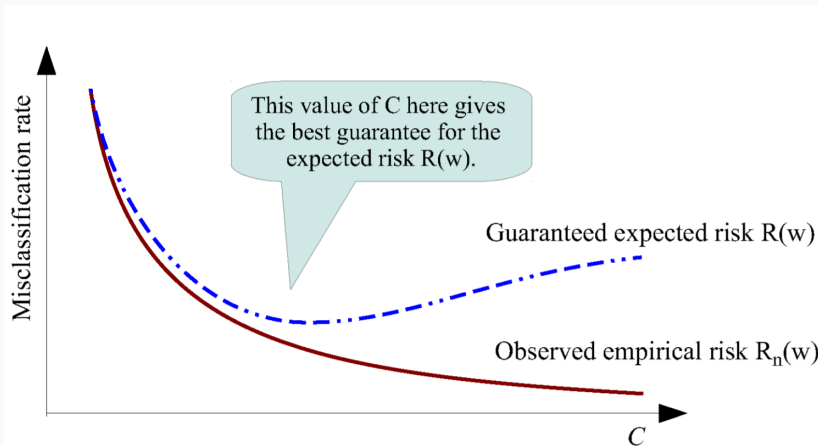
Minimisation structurelle du risque :

$$\hat{h}_n = \arg \min_{h \in \mathcal{H}} \hat{R}_n(h) + \lambda K(h)$$

ou

$$\hat{h}_n = \arg \min_{K(h) \leq C} \hat{R}_n(h)$$

Structural Risk Minimization Tradeoff

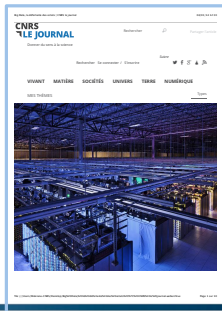


Source: Bottou et al. tutorial on optimization

- L'analyse de données (inférence, description) est le but des statistiques depuis longtemps.
- Le Machine Learning a des buts plus **opérationnels** (ex: la consistance est importante en statistique mais moins en ML). Les modèles (quand il y en a) sont *instrumentaux*.
Ex: modèle linéaires (jolie théorie mathématique) vs Random Forests (utilisation massive de modèles pauvres et sans signification propre).
- Machine Learning pour les big data: plus de séparation entre modélisation stochastique et optimisation (contrairement aux statistiques classiques).
- En ML, les données sont souvent là a priori (malheureusement).
- Pas de frontière infranchissable (la statistique aussi évolue).

Le phénomène Big Data

Sur le site du CNRS



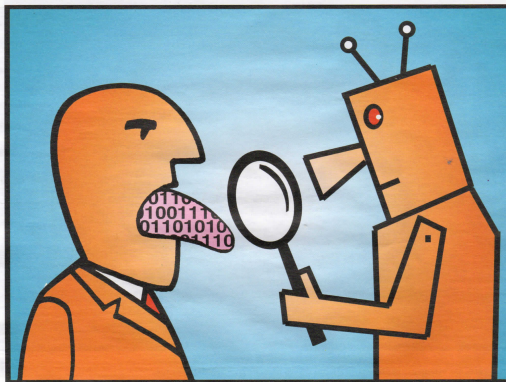
Et ailleurs



En couverture

Nouvelles technologies

Le partage des données nous sauvera-t-il ?



↑ Dessin de Kopelnitsky, Etats-Unis.



P. 33

1. Un monde numérique
2. Qu'est-ce que le machine learning ?
3. Le phénomène Big Data
 - Les Big Data dans la presse
 - Qu'est-ce que les Big Data ?
 - Éviter les fausses découvertes
4. Quel Machine Learning pour les Big Data?
5. Les Big Data, le Machine learning et la loi
6. Enseignement des sciences des données

Aucun domaine n'échappe à l'avalanche des données



- **Commerce et les affaires**
 - SI d'entreprise, Banques, transactions commerciales, systèmes de réservation, ...
- **Gouvernements et organisations**
 - Lois, réglementations, standards, infrastructures,
- **Loisirs**
 - Musique, vidéo, jeux, réseaux sociaux...
- **Sciences fondamentales**
 - Astronomie, physique et énergie, génome, ...
- **Santé**
 - Dossier médical, sécurité sociale,...
- **Environnement**
 - Climat, dév durable, pollution, alimentation,...
- **Humanités et Sciences Sociales**
 - Numérisation du savoir (littérature, histoire, art, architectures), données archéologiques...

Qu'est-ce qu'une (très grande) masse de données ?



VLDB

XLDB

Massive Data

Big Data

Very Big Data

Data Deluge

Data Masses

Data inflation

2

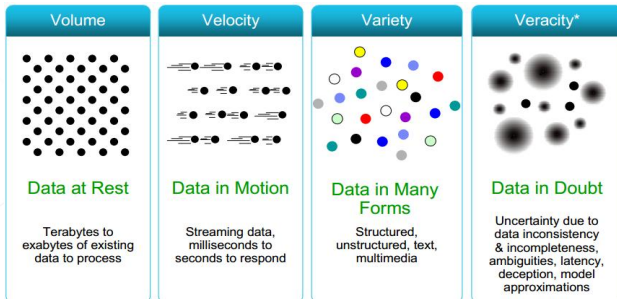
Unit	Size	What it means
Bit (b)	1 or 0	Short for "binary digit", after the binary code (1 or 0) computers use to store and process data
Byte (B)	8 bits	Enough information to create an English letter or number in computer code. It is the basic unit of computing
Kilobyte (KB)	1,000, or 2^{10} , bytes	From "thousand" in Greek. One page of typed text is 2KB
Megabyte (MB)	1,000KB; 2^{20} bytes	From "large" in Greek. The complete works of Shakespeare total 5MB. A typical pop song is about 4MB
Gigabyte (GB)	1,000MB; 2^{30} bytes	From "giant" in Greek. A two-hour film can be compressed into 1-2GB
Terabyte (TB)	1,000GB; 2^{40} bytes	From "monster" in Greek. All the catalogued books in America's Library of Congress total 15TB
Petabyte (PB)	1,000TB; 2^{50} bytes	All letters delivered by America's postal service this year will amount to around 5PB. Google processes around 1PB every hour
Exabyte (EB)	1,000PB; 2^{60} bytes	Equivalent to 10 billion copies of <i>The Economist</i>
Zettabyte (ZB)	1,000EB; 2^{70} bytes	The total amount of information in existence this year is forecast to be around 1.2ZB
Yottabyte (YB)	1,000ZB; 2^{80} bytes	Currently too big to imagine

Source: *The Economist*

The prefixes are set by an inter-governmental group, the International Bureau of Weights and Measures. Yotta and Zetta were added in 1991; terms for larger amounts have yet to be established.

Grandes Conf du domaine: VLDB, XLDB, ICDE, EDBT, ...

Complexité multidimensionnelle des Big Data



• Nouvelles archi. de stockage

• Nouvelles archi. d'interopérabilité

• Défi pour les réseaux de communication

• Nouveaux modèles de calcul sur des flux

• Nettoyage et transformation

• Fusion de données

Nouveaux modèles de qualité (données & processus de traitement)

<http://www.datasciencecentral.com/profiles/blogs/data-veracity>

1. Un monde numérique
2. Qu'est-ce que le machine learning ?
3. Le phénomène Big Data
 - Les Big Data dans la presse
 - Qu'est-ce que les Big Data ?
 - Éviter les fausses découvertes
4. Quel Machine Learning pour les Big Data?
5. Les Big Data, le Machine learning et la loi
6. Enseignement des sciences des données

De nombreux journaux parlent d'études scientifiques s'appuyant sur des études et des sondages



The image shows a screenshot of a web browser displaying a news article. The browser's address bar shows the URL: www.sudouest.fr/2012/09/27/plus-un-homme-aide-a-la-maison-plus-il-risque-le-divorce-833503-4776.php. The article title is "Plus un homme aide à la maison, plus il risque le divorce". Below the title, it says "A LA UNE / REDACTION / INSOLITE / Publié le 27/09/2012 . Mis à jour le 27/01/2016 à 16h23 par SudOuest.fr". There are social media sharing buttons for Facebook, Twitter, Google+, LinkedIn, and Email, along with a yellow button that says "S'ABONNER À PARTIR DE 1€" and a blue button that says "117 COMMENTAIRES". The main image of the article shows a man and a woman standing on a beach, looking out at the ocean. The man is on the left, wearing a light blue shirt and dark shorts, and the woman is on the right, wearing a white dress. They are standing with their backs to each other, looking out at the sea. The sky is overcast and the water is a pale blue-grey color.

De nombreux journaux parlent d'études scientifiques s'appuyant sur des études et des sondages

The screenshot shows a web browser displaying a news article on the Le Point website. The article title is "Plus un pays mange de chocolat, plus il a de prix Nobel, révèle une étude". The article text discusses a study by Dr. Franz Messerli from Columbia University, suggesting that countries with high chocolate consumption also have a higher number of Nobel laureates. The article includes social media sharing icons and a quote from the author. To the right of the article, there is a "EN CONTINU" section with a list of other news items, a "BOURSE" section with market data for CAC 40, BEL 20, AEX-INDEX, and PSI 20, and a small image of the US Capitol building.

Plus un pays mange de chocolat, plus il a de prix Nobel, révèle une étude

Publié le 11/10/2012 à 16:41 | AFP

ABONNEZ-VOUS À PARTIR DE 1€

Plus la population d'un pays mange de chocolat, plus il compte de prix Nobel, une indication que le cacao operait les capacités mentales, selon une étude publiée jeudi dans la très sérieuse revue médicale américaine New England Journal of Medicine.

Les flavonoïdes, de puissants antioxydants qu'on trouve en grande quantité dans les fèves de cacao, le thé vert et le vin rouge, ont montré qu'ils réduisaient le risque de démence et amélioreraient les fonctions mentales chez les personnes âgées, note le Dr Franz Messerli, de l'Université Columbia à New York et auteur de cet essai.

"Puisque le chocolat peut hypothétiquement améliorer les fonctions cognitives chez les individus et à fortiori dans l'ensemble d'une population, je me suis demandé s'il pouvait y avoir une corrélation entre la consommation de cacao dans un pays et les capacités mentales de ses habitants", explique avec un pointe d'humour le médecin.

"A ma connaissance, il n'existe pas de données disponibles mesurant les fonctions mentales de toute une nation", poursuit-il. "On peut donc concevoir que le nombre total de Nobel par tête pourrait donner une certaine idée des fonctions cognitives d'une nation".

En poursuivant votre navigation vous acceptez l'utilisation de cookies pour vous proposer des services et offres adaptés à vos centres d'intérêts. Pour en savoir plus, cliquez ici.

EN CONTINU

- 08H29 "Dul'f'èdt cru ?": 11 secrets sur les peintres et leurs...
- 14H55 Un airbus d'Eihhad pris dans des turbulences : 31 blessés !
- 12H15 Elle donne 25 millions d'euros à ses proches... Ils ne lui parlent...
- 08H28 Indonésie: des habitants ont pris un jouet sexuel pour un ange...
- 17H46 Vosges: pour sauver leur usine, des salariés s'offrent une...

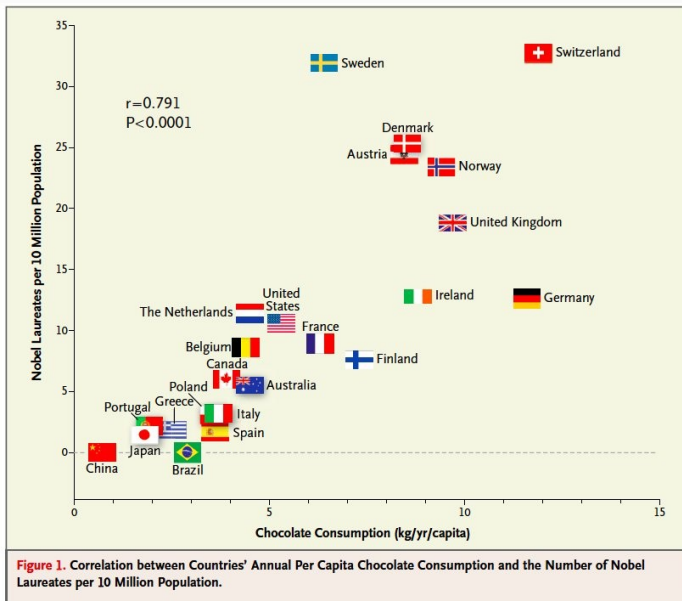
Voir toute l'actualité en continu

BOURSE → Voir la page bourse

🇫🇷 CAC 40	▼ -1,12 %	4270,880
🇸🇦 BEL 20	▼ -0,22 %	3320,660
🇳🇱 AEX-INDEX	▼ -0,07 %	428,110
🇮🇹 PSI 20	▼ -0,38 %	5017,940

28

Si les mathématiques l'affirment ...



Incompréhension autour de la notion de corrélation

- En grandes dimensions **tout est corrélé**.
- **Corrélation n'est pas causalité** :
A nombre de prix Nobel et B consommation de chocolat

$$A \mapsto B \quad \text{ou} \quad B \mapsto A$$

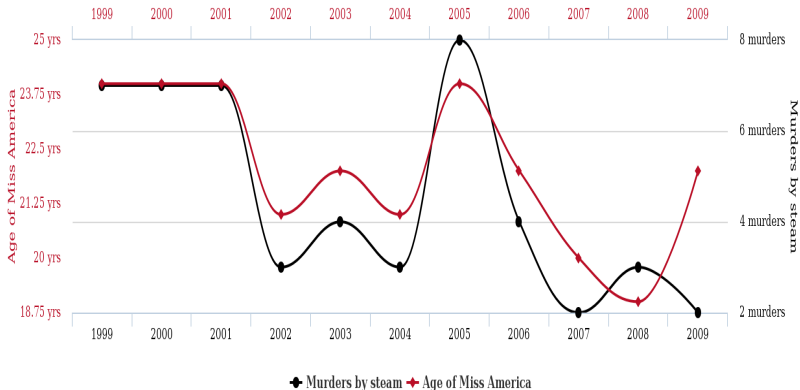
- Variables non observées : **variables confondantes**



Age of Miss America

correlates with

Murders by steam, hot vapours and hot objects



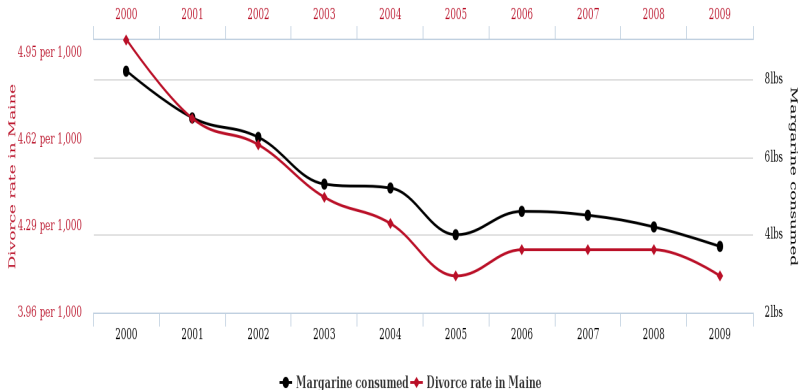
tylervigen.com

[Src: Loubes, *Petit guide de self-défense contre la statistique et ceux qui la manipulent*]

Divorce rate in Maine

correlates with

Per capita consumption of margarine

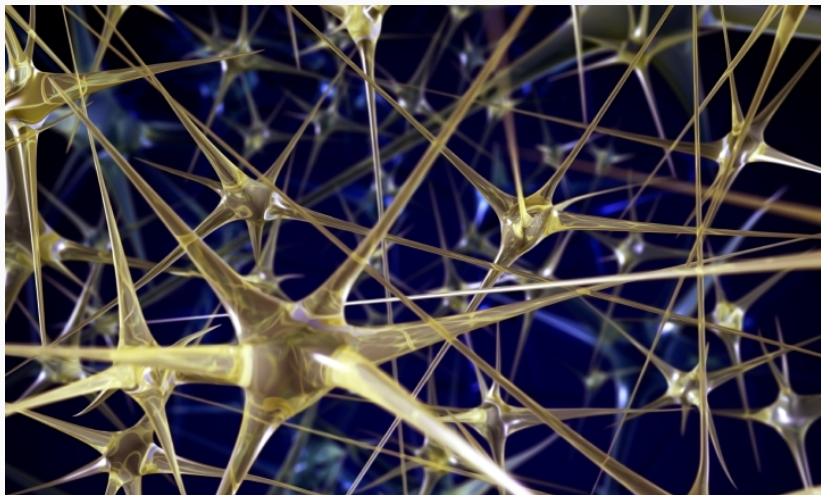


tylervigen.com

[Src: Loubes, *Petit guide de self-défense contre la statistique et ceux qui la manipulent*]

Quel Machine Learning pour les Big Data?

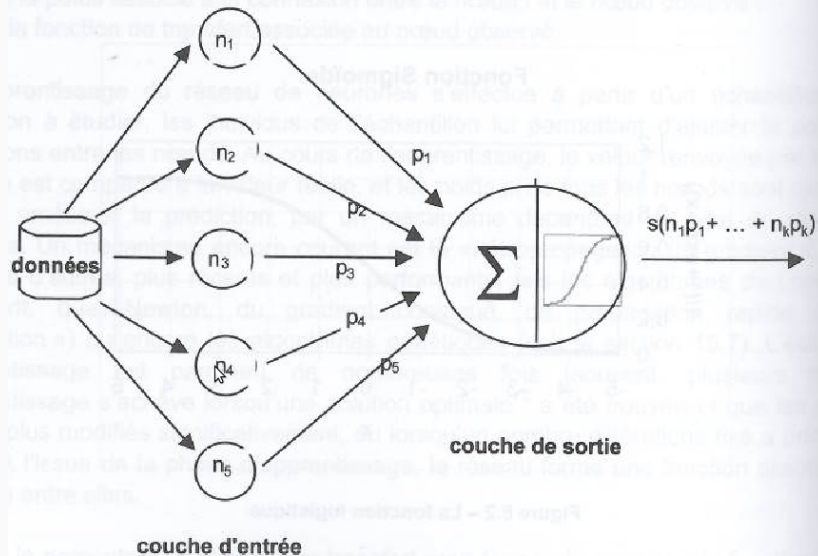
Réseaux de neurones



Source: <http://insanedevelop.co.uk/open-cranium/>

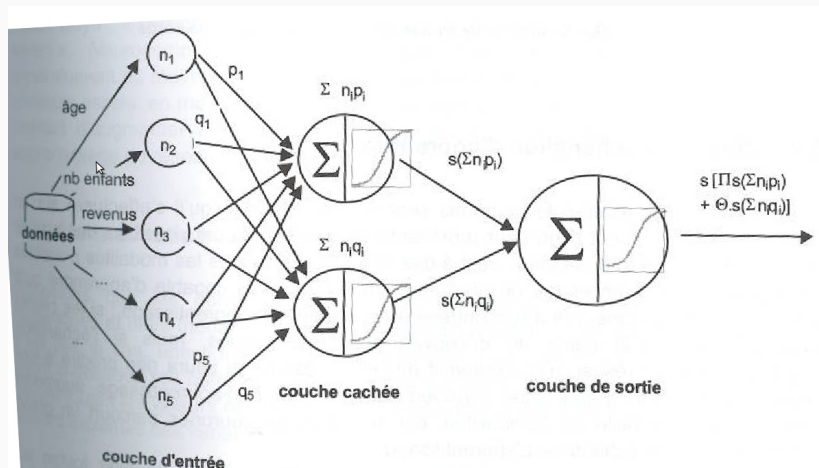
Réseau mono-couche

Source: [Tufféry, Data Mining et Informatique Décisionnelle]

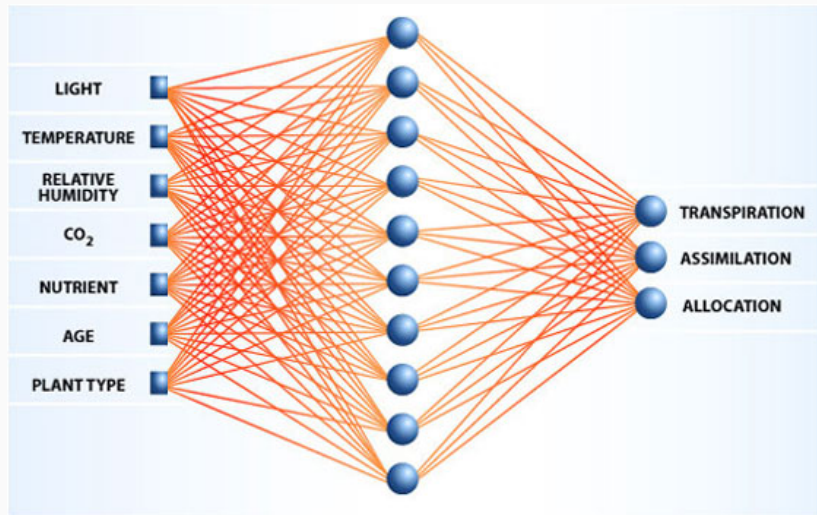


Réseau avec couche intermédiaire

Source: [Tufféry, Data Mining et Informatique Décisionnelle]



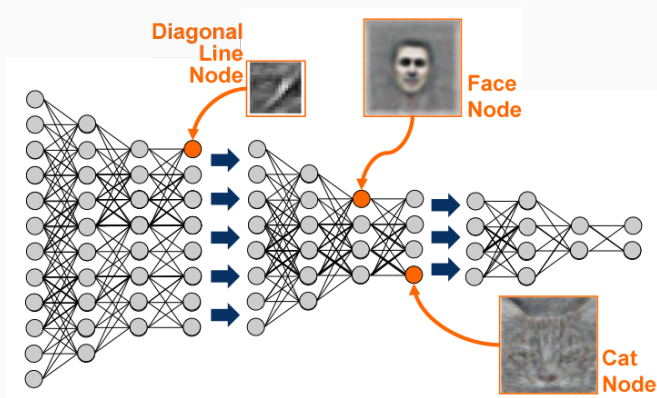
Réseau à une couche cachée



Des réseaux de neurones au deep learning (apprentissage profond)

Deep learning = réseaux de neurones + 3 améliorations:

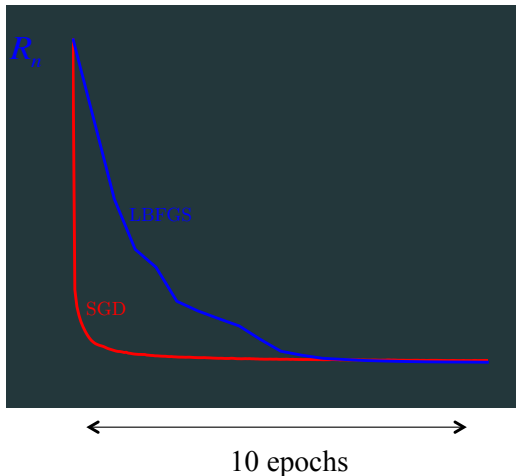
- extensions (nouvelles fonction d'activation, convolution, récursivité)
- régularisation (dropout, pooling)
- calcul (GPU, jeux de données massifs)



1. Un monde numérique
2. Qu'est-ce que le machine learning ?
3. Le phénomène Big Data
4. Quel Machine Learning pour les Big Data?
 - Deep Learning
 - Optimisation
5. Les Big Data, le Machine learning et la loi
6. Enseignement des sciences des données

- Les méthodes du second ordre sont trop coûteuses (même une seule itération)
- Même les méthodes classiques du premier ordre sont trop coûteuses avec des données vraiment massives
- Le gradient stochastique (et ses variantes) utilisent plus efficacement l'information que les algorithmes batch.

Practical Experience

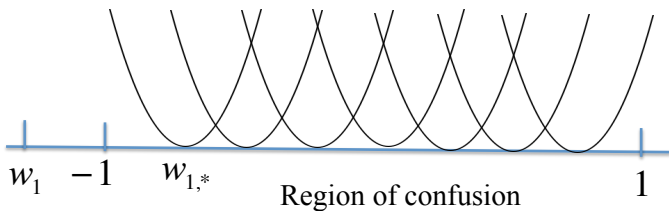


Fast initial progress
of SG followed by
drastic slowdown

Can we explain this?

Example by Bertsekas

$$R_n(w) = \frac{1}{n} \sum_{i=1}^n f_i(w)$$



Note that this is a **geographical** argument

Analysis: given w_k what is the **expected decrease** in the objective function R_n as we choose one of the quadratics randomly?

Les Big Data, le Machine learning et la loi

Deux principes et une contrainte

Principes :

- **principe de minimisation**
enlever les données qui ne sont pas utiles (contre logique big-data)
- **principe de finalité**
pas le droit de croiser des données de différents services

Contrainte :

- **contrainte de compatibilité**
ne pas empêcher les progrès en France !

(au moins) 5 challenges

- **explicabilité des décisions**
ex: pour les systèmes de référencement, la loi République numérique adoptée en oct 2016 oblige à préciser les modalités de référencement
Quel contrôle ?
- **anonymisation des données**
garantir la non-identifiabilité (contre-exemples célèbres)
- **non-discrimination** des sous-populations
création d'une plateforme de dénonciation des mauvaises expériences avec les algorithmes
- **distorsion de concurrence**
barrière à l'entrée du fait d'avoir les données
- **ouverture/transparence** versus protection du secret d'affaire

L'Apprentissage Machine comdamné

THE WALL STREET JOURNAL.

Subscribe Now | Sign In

SPECIAL OFFER: JOIN NOW

Home World U.S. Politics Economy Business Tech **Markets** Opinion Arts Life Real Estate



Oil at One-Year High on Falling Stockpiles



U.S. Stocks Rise on Oil Rally, Bank Earnings



Platinum Partners' Flagship Hedge Fund Files for Bankruptcy



MARKETS

U.S. Government Uses Race Test for \$80 Million in Payments

Checks are ready for minority borrowers allegedly discriminated against on Ally Financial auto loans

By **ANNAMARIA ANDRIOTIS** and **RACHEL LOUISE ENSIGN**

Updated Oct. 29, 2015 9:32 p.m. ET

Recommended Videos

Enseignement des sciences des données

Dosage informatique (bases de données) / mathématique (statistiques)

Bordeaux: Master Mathématiques appliquées et statistique (MAS):
parcours

- Modélisation Mathématique pour le Signal et l'Image
- Modélisation statistique et stochastique
- Image Processing and Computer Vision
- Cursus Master Ingénierie Statistique et Informatique

Toulouse:

- UPS CMI MAPI3
- UPS CMI SID-bigdata
- INSA parcours GMM
- ISAE
- TSE Master StatEco

Intérêts:

- motivation
- autonomie
- travail en équipe
- débrouillardise
- initiation recherche

Challenge 2016-2017 :
voir <http://challenge.openbikes.co/>



Accueil | FAQ | Se connecter | S'inscrire

Open Bikes CHALLENGE 2016

TELECHARGER LES DONNEES DE LA PREMIERE PARTIE

TELECHARGER LES DONNEES DE LA SECONDE PARTIE

Classement pour la première partie

Classement	Equipe	Parcours/Ecole	Score	Submission(s)	Dernière soumission
1	Dream Team	ISAE - SUPAERO	2.79442	49	2017/01/12
2	Mr Nobody	Université de Bordeaux	2.49000	14	2017/01/11
3	Oh l'équipe	Université de Bordeaux	2.49000	45	2017/01/11
4	PréSIX	Ecole Polytechnique	2.54029	47	2017/01/07
5	Louison Bobet	Toulouse School of Economics	2.61000	61	2017/01/11
6	Armstrong	INSA	2.62000	35	2017/01/09
7	OpenBikes	Université Paul Sabatier	2.62921	7	2016/12/28
8	Ravenclaw	Université de Bordeaux	2.67000	64	2017/01/11
9	LA ROUE ARRÊTÉE	Université Paul Sabatier	2.67926	6	2016/12/06
10	WeLoveTheHail	INSA	2.70333	5	2017/01/11
11	GMRendox	INSA	2.72667	20	2017/01/11
12	TEAM_DUY	Université de Bordeaux	2.73333	17	2017/01/12
13	zoomzoom	Université de Bordeaux	2.73333	21	2017/01/12
14	KAMD-HAMENHA	Université de Bordeaux	2.73333	21	2017/01/12