Exemples d'applications et actuels tendances de l'IA

Une visite guidée de l’IA pour les problèmes environnementaux et géospatiaux

Michele Volpi
Lead Data Scientist / ETHZ Senior Scientist
Swiss Data Science Center

michele.volpi@sdsc.ethz.ch
datascience.ch

Qu’est-ce que c’est l’IA?

Swiss Data Science Center

Qu'est-ce que c'est l’intelligence artificielle?

L'intelligence artificielle (IA) est un **domaine vaste et en constante évolution** qui englobe **un large éventail de technologies et d'approches**. En général, l'IA vise à créer des **systèmes informatiques capables de simuler l'intelligence humaine**. Cela peut inclure des tâches telles que la compréhension du langage naturel, l'apprentissage et la résolution de problèmes, et la prise de décisions autonomes.

gemini

Qu’est-ce que c’est l’IA?

Swiss Data Science Center

Qu'est-ce que c'est l’intelligence artificielle?

L'intelligence artificielle (IA) est un **domaine vaste et en constante évolution** qui englobe **un large éventail de technologies et d'approches**. En général, l'IA vise à créer des **systèmes informatiques capables de simuler l'intelligence humaine**. Cela peut inclure des tâches telles que la compréhension du langage naturel, l'apprentissage et la résolution de problèmes, et la prise de décisions autonomes.

gemini

- Pas de définition claire, surtout ces dernières années
- L’IA est un système complexe, pas seulement une méthode ou un «tool»
- Vise premièrement à la création de «logiciels»

Qu’est-ce que c’est l’IA?

Swiss Data Science Center

De quoi est elle composée l’intelligence artificielle?

- Apprentissage
- Perception (vision, audio, senseurs)
- Raisonnement, logique
- Représentation de la connaissance
- Langage (NLP)
- Prise de décision

Qu’est-ce que c’est l’IA?

Swiss Data Science Center

De quoi est elle composée l’intelligence artificielle?

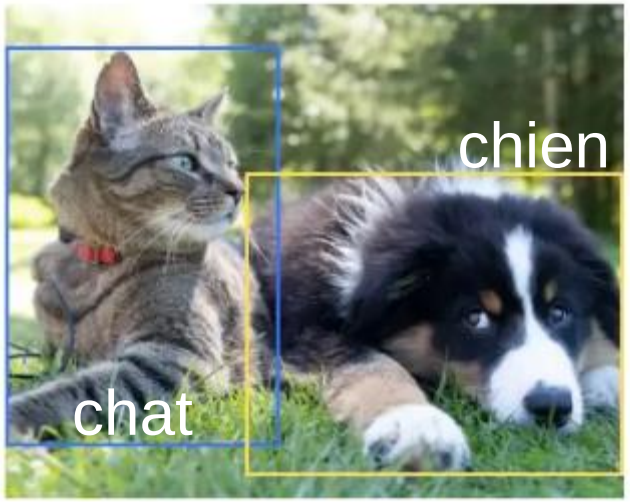
- Apprentissage
- Perception (vision, audio, senseurs)
- Raisonnement, logique
- Représentation de la connaissance
- Langage (NLP)
- Prise de décision

Qu’est-ce que c’est l’IA?

Swiss Data Science Center

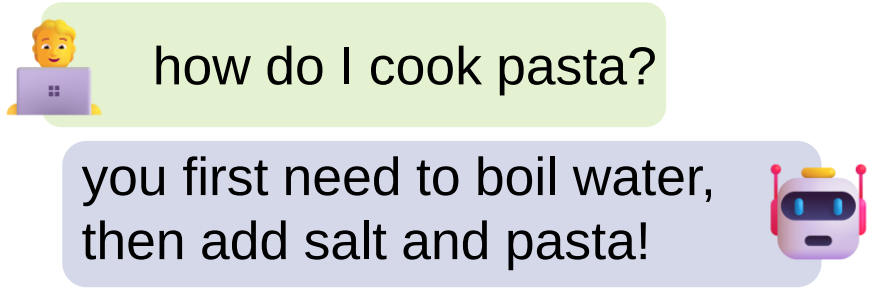
Composante principale de l’IA: apprentissage automatique à grande échelle (**machine learning**)

→l’*apprentissage automatique* est un ensemble de méthodes et d’algorithmes permettant apprendre de fonctions directement depuis les données
→Permet les ordinateurs d’apprendre des règles sans être spécifiquement programmés

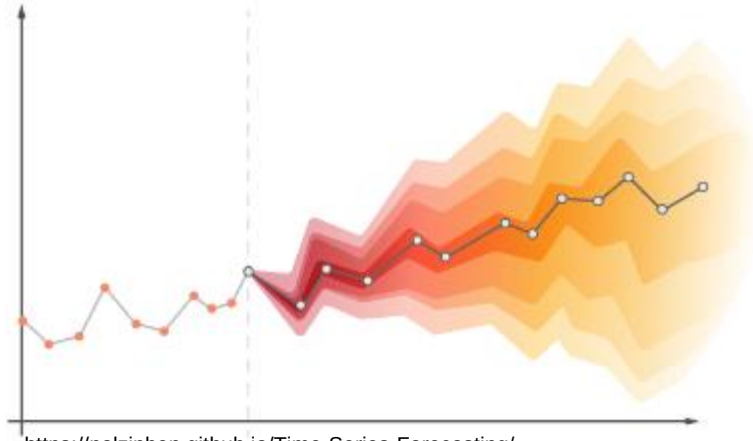


https://www.superannotate.com/blog/image-classification-basics

reconnaissance d’images



chatbots / assistants


https://polzinben.github.io/Time-Series-Forecasting/

séries temporelles / traitement des signaux

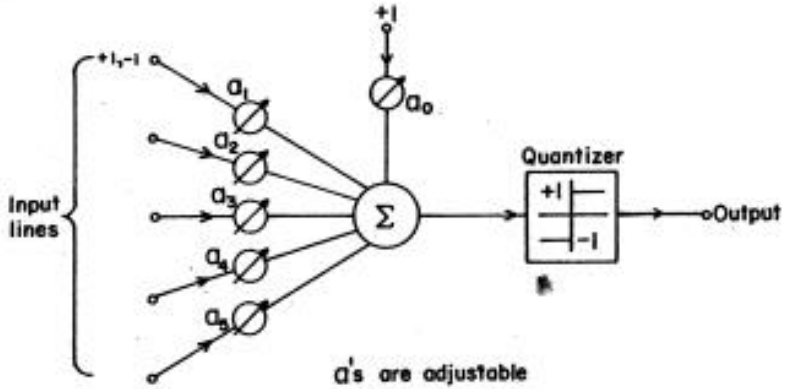


Fig. 1. An adjustable neuron.

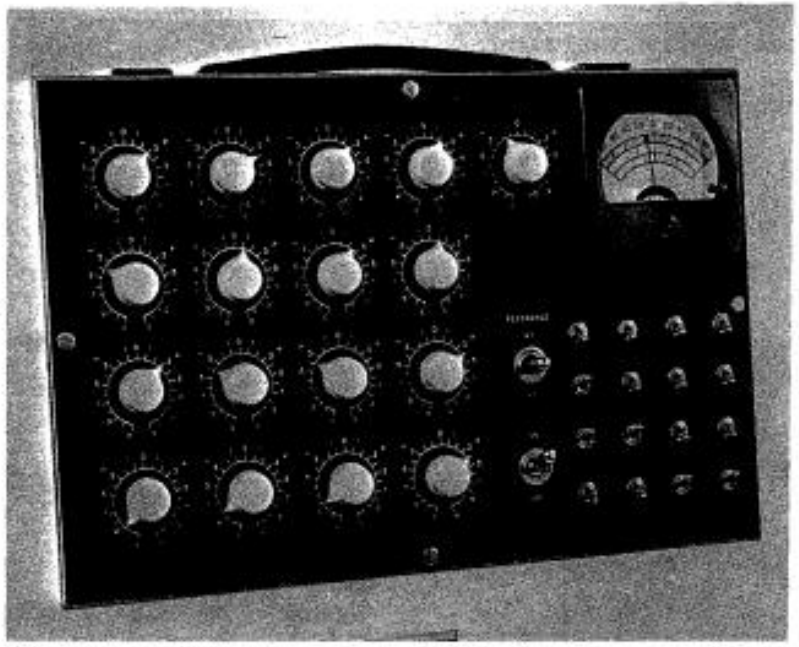


Fig. 2. Adaline.

https://www-isl.stanford.edu/~widrow/papers/c1960adaptiveswitching.pdf

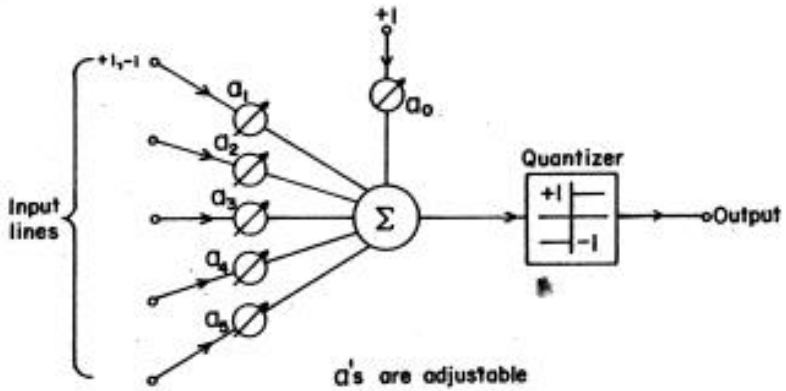


Fig. 1. An adjustable neuron.

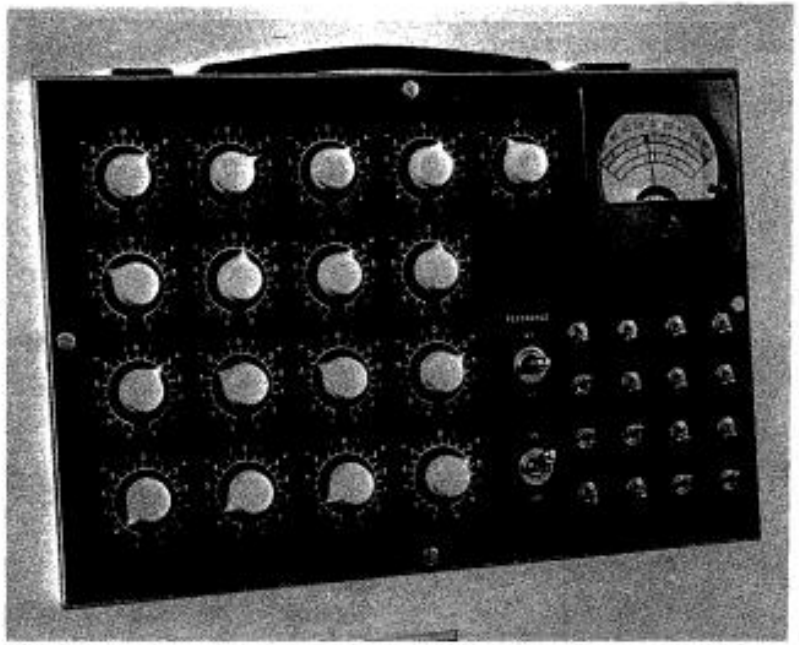
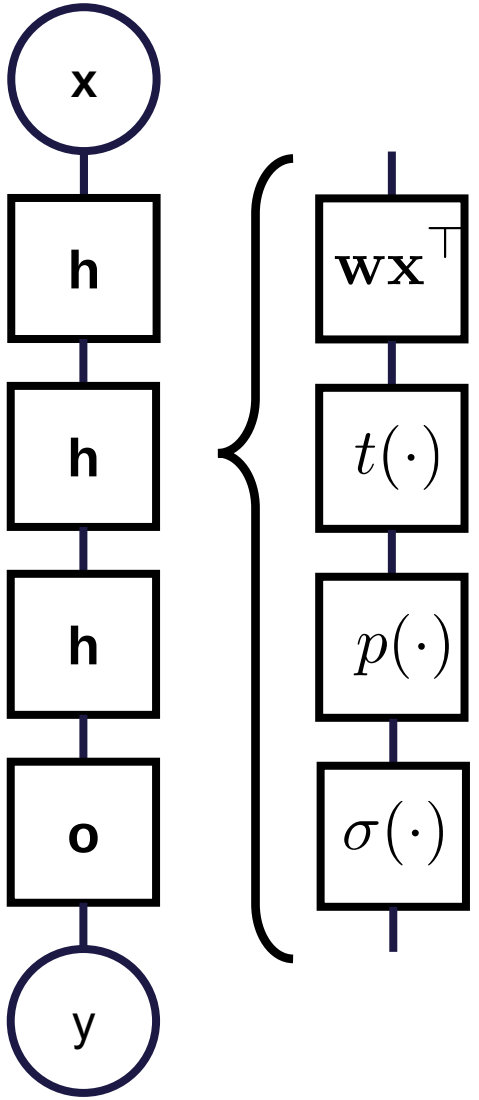


Fig. 2. Adaline.

analogique
simple

digital
complexe
automatique

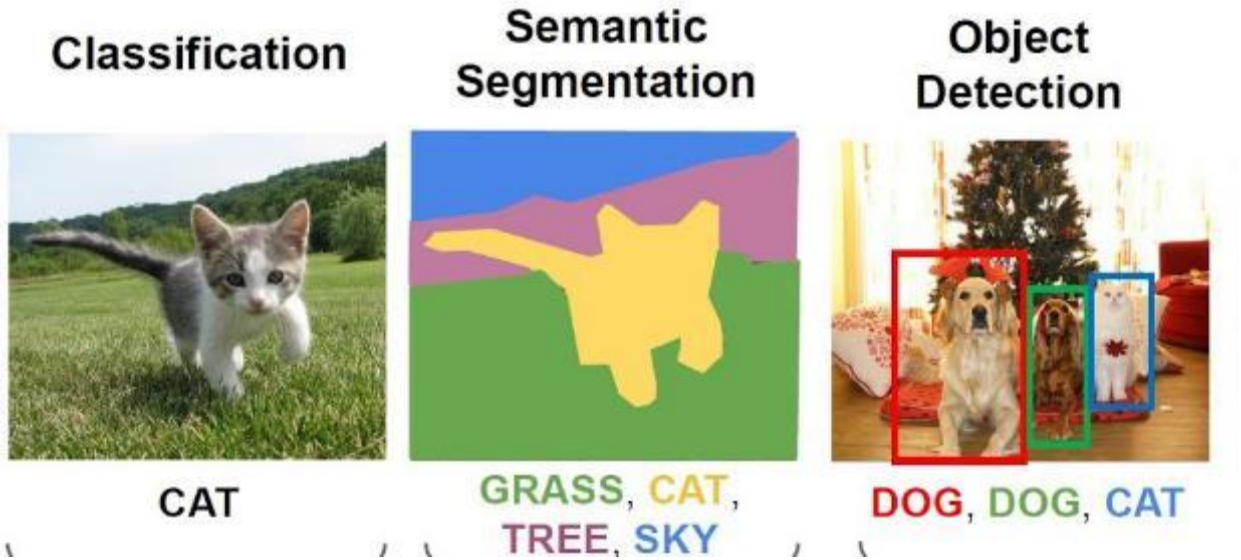


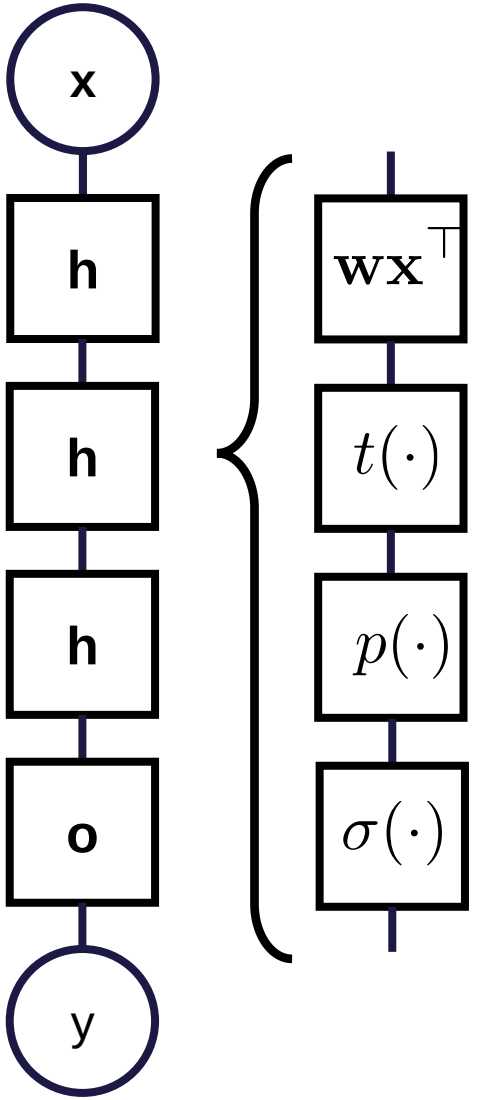
https://www-isl.stanford.edu/~widrow/papers/c1960adaptiveswitching.pdf

Réseaux de neurones (neural networks)

Input **x**: image, son, texte, série temporelle, …

Output y: étiquette (classe), valeur continue, séries, matrices, …





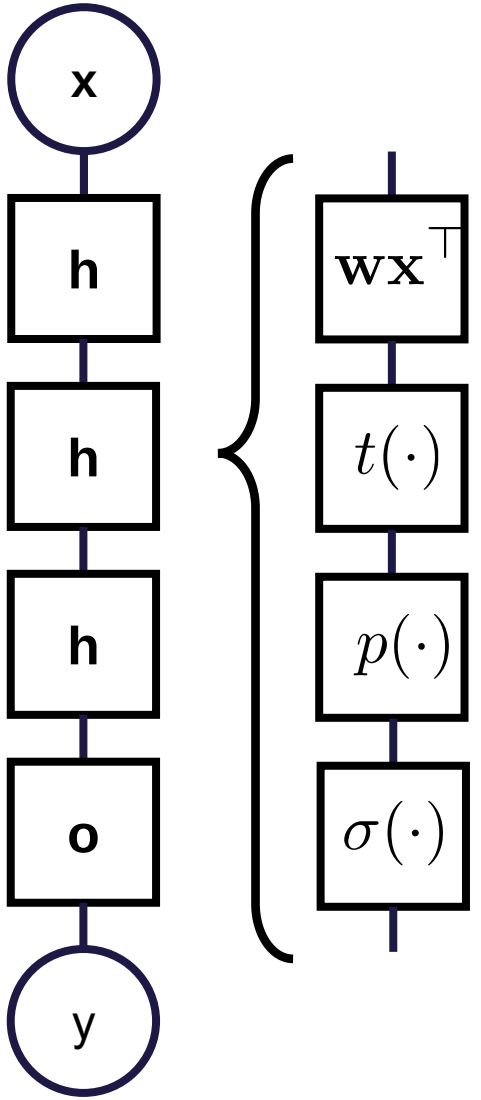
https://cs231n.stanford.edu/slides/2017/cs231n_2017_lecture11.pdf

Réseaux de neurones (neural networks)

Input **x**: image, son, texte, série temporelle, …

Output y: étiquette (classe), valeur continue, …

Tout ce qui est entre x et y est appris automatiquement depuis les données!



Réseaux de neurones (neural networks, NN)

Swiss Data Science Center

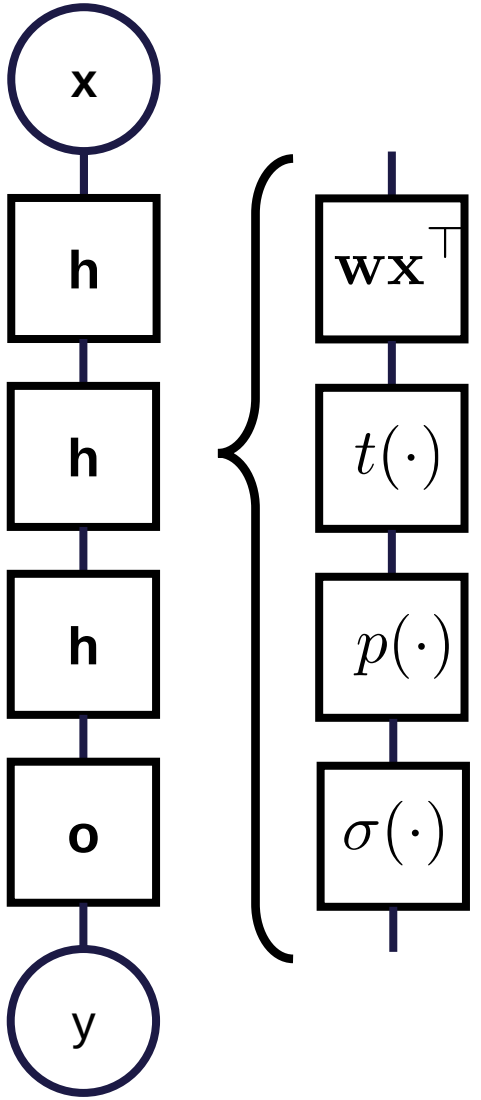
Input **x**: image, son, texte, série temporelle, …

Output y: étiquette (classe), valeur continue, …

Tout ce qui est entre x et y est appris automatiquement depuis les données!

→ NN apprennent fonctions **adaptées** pour chaque tâche et chaque jeu de donné

→ Il faut beaucoup de données pour entrainer des modèles puissants et généraux



Modèles de apprentissage automatique

Swiss Data Science Center

What if… ?

- Pas assez de données
- Pas assez de puissance de calcul

Modèles de apprentissage automatique

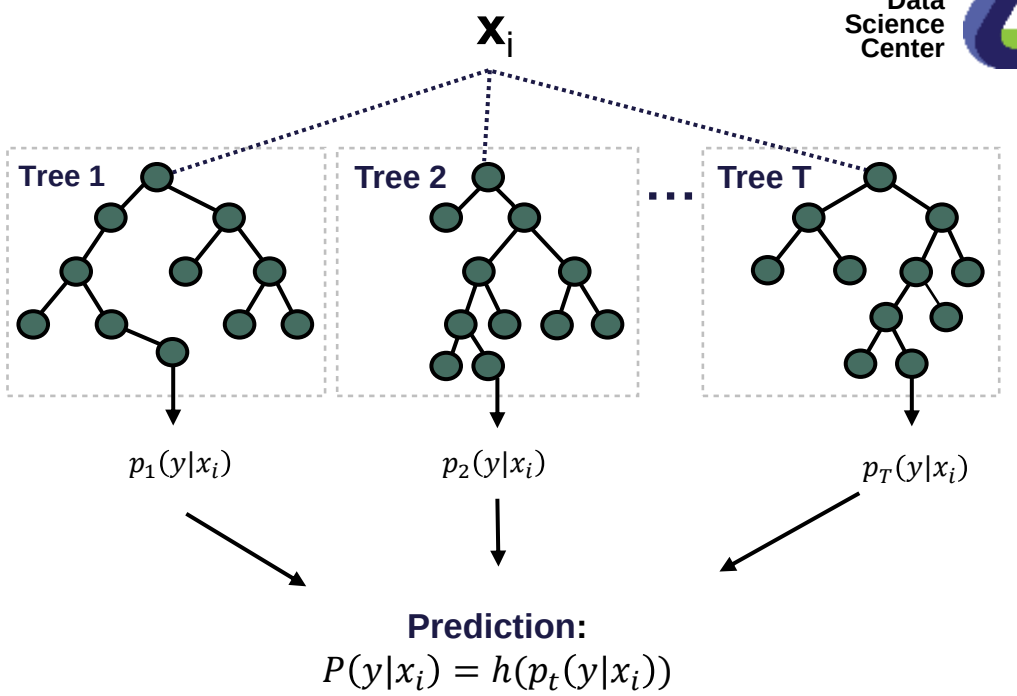
Swiss Data Science Center

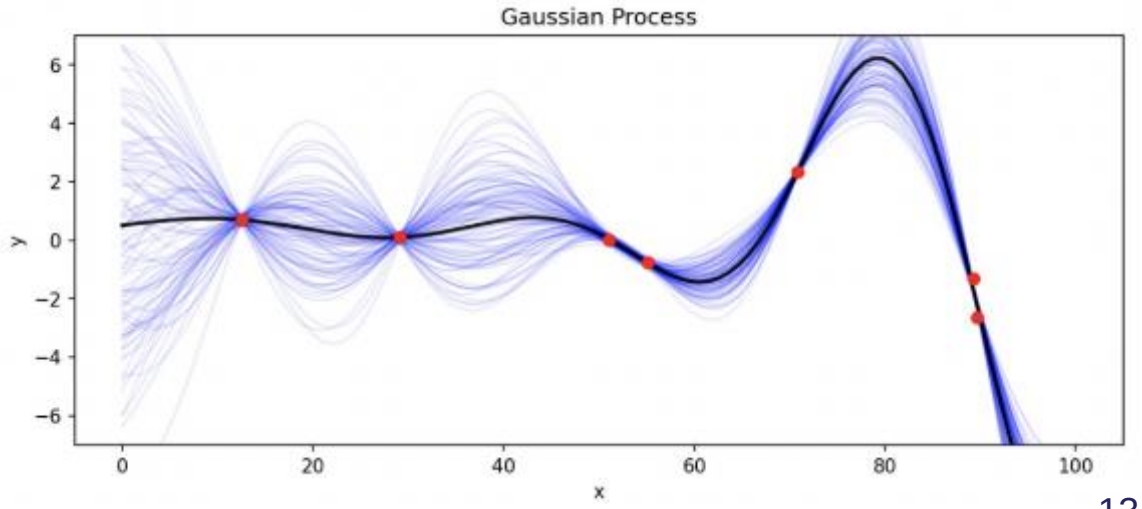
What if… ?

- Pas assez de données
- Pas assez de puissance de calcul

Beaucoup d'autre approches!

- Random forests, boosted decision trees
- Gaussian Processes
- …





Entrainement

Supervision d'experts
(étiquetage)

+ Beaucoup de données
+ Puissance de calcul
+ Software et librairies

Données entrainement

Optimisation modèle / entrainement

fonction à optimiser
(p.ex. est-ce que cela est un chat?)

Production / déploiement

Données prédiction

+ Navigateur
+ Ordi moderne

prédiction
(C'est un chien avec probabilité 0.78!)

SDSC «data science» de l'environnement

Swiss Data Science Center

Dangers naturels: cartes de susceptibilité, détection d'événements
Sciences du climat: modèles physiques → IA, prévision
Ecologie: cartographie, détection (camera traps), mise a jour inventaires
Agriculture: mapping, estimation, 3D reconstruction
Géographie: science urbaines, cartographie

SDSC «data science» de l’environnement

Swiss Data Science Center

Dangers naturels: cartes de susceptibilité, détection d’événements
Sciences du climat: modèles physiques → IA, prévision
Ecologie: cartographie, détection (camera traps), mise a jour inventaires
Agriculture: mapping, estimation, 3D reconstruction
Géographie: science urbaines, cartographie

Données:
Satellites
Modèles physiques
Stations météo (géolocalisés)
Images naturelles (géolocalisés)
…

SDSC «data science» de l'environnement

Dangers naturels: cartes de susceptibilité, détection d'événements
Sciences du climat: modèles physiques → IA, prévision
Ecologie: cartographie, détection (camera traps), mise a jour inventaires
Agriculture: mapping, estimation, 3D reconstruction
Géographie: science urbaines, cartographie

Données:
Satellites
Modèles physiques
Stations météo (géolocalisés)
Images naturelles (géolocalisés)
…

Surveillance biodiversité par pièges photo

Swiss Data Science Center

Surveillance biodiversité par pièges photo

Cas d'étude: détecter colibris dans de videos

- Amérique centrale et Am. du sud, importantes indicateurs de la santé des écosystèmes
- Pollinisateurs pour une majorité des ~7000 espèces de plantes
- Modèles doivent être accessible et portable (terrain ordis)

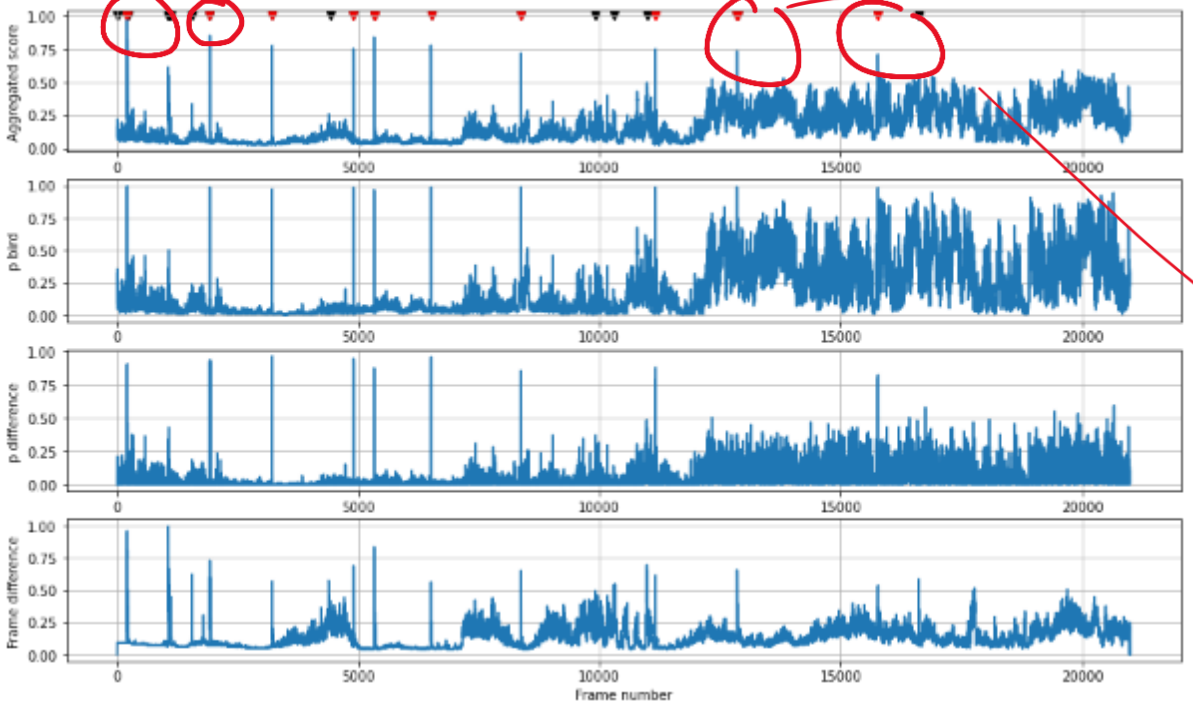
Surveillance biodiversité par pièges photo

- **Module de prédiction:** chaque frame p(y=bird|x), deep NN
- **Module nouveautés:** comparaison entre frames «anomalies»
- **Module drift classificateur**: Prise en compte illumination / bruit
- **Module localisation:** post-hoc localisation (gradCAM)

t-1 t t+1

Surveillance biodiversité par pièges photo

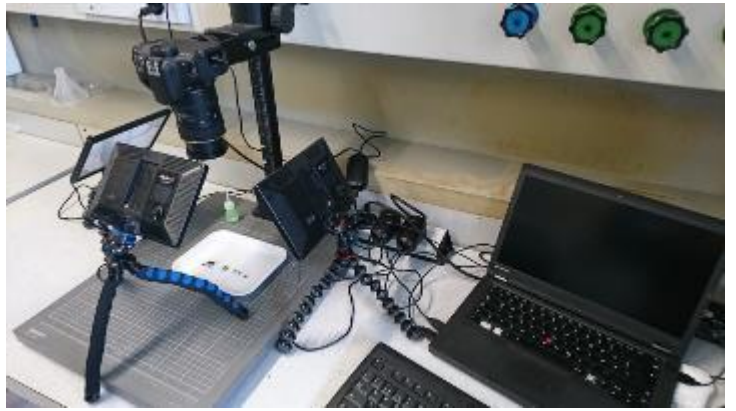
Swiss Data Science Center

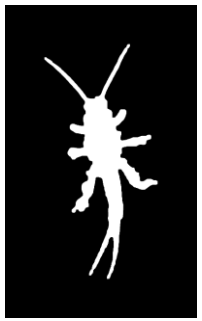


Surveillance biodiversité par échantillonnage

Phénotypage macrozoobenthos

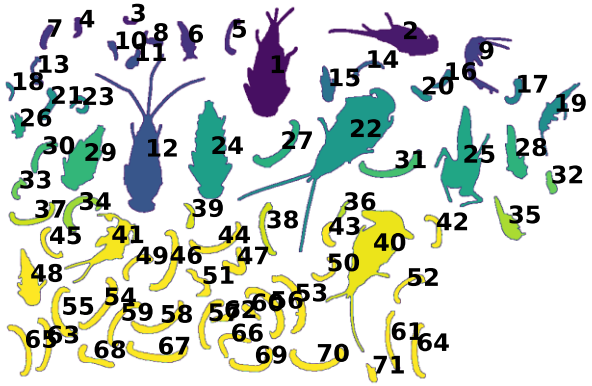
- Qualité eaux et écosystème réseaux fluviatile en Suisse (EAWAG)
- Classification et mensuration
- Traitement d'image traditionnel + IA (deep NN / Vision Transformer)





7: oligochaeta
- **l: 35mm**
- **w: 2mm**

5: plecoptera
- **l: 51mm**
- **w: 6mm**



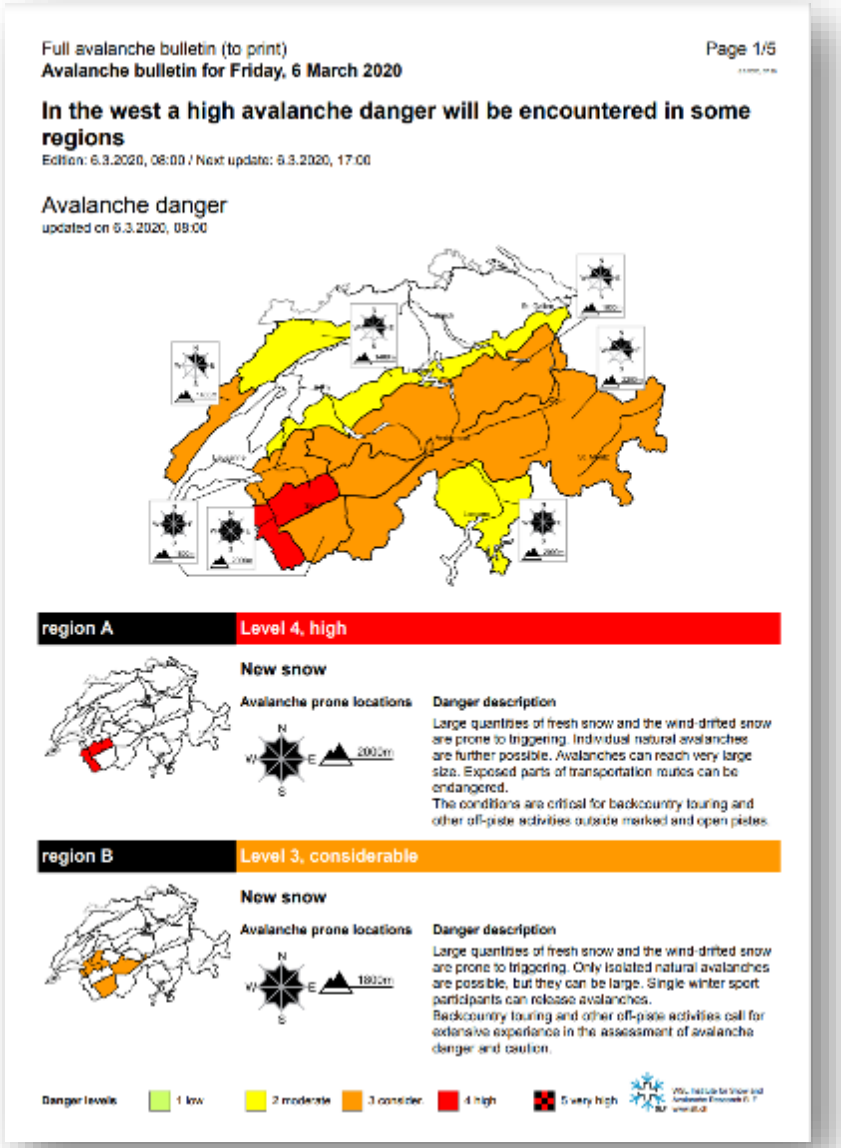
SDSC «data science» de l'environnement

Dangers naturels: cartes de susceptibilité, détection d'événements
Sciences du climat: modèles physiques → IA, prévision
Ecologie: cartographie, détection (camera traps), mise a jour inventaires
Agriculture: mapping, estimation, 3D reconstruction
Géographie: science urbaines, cartographie

Données:

Satellites
Modèles physiques
Stations météo (géolocalisés)
Images naturelles (géolocalisés)
…

Cartographie du danger d'avalanches



Full avalanche bulletin (to print)
Page 1/5

Avalanche bulletin for Friday, 6 March 2020

In the west a high avalanche danger will be encountered in some regions

Edition: 6.3.2020, 08:00 / Next update: 6.3.2020, 17:00

Avalanche danger

updated on 6.3.2020, 08:00

region A — **Level 4, high**

New snow

Avalanche prone locations

Danger description

Large quantities of fresh snow and the wind-drifted snow are prone to triggering. Individual natural avalanches are further possible. Avalanches can reach very large size. Exposed parts of transportation routes can be endangered.
The conditions are critical for backcountry touring and other off-piste activities outside marked and open pistes.

region B — **Level 3, considerable**

New snow

Avalanche prone locations

Danger description

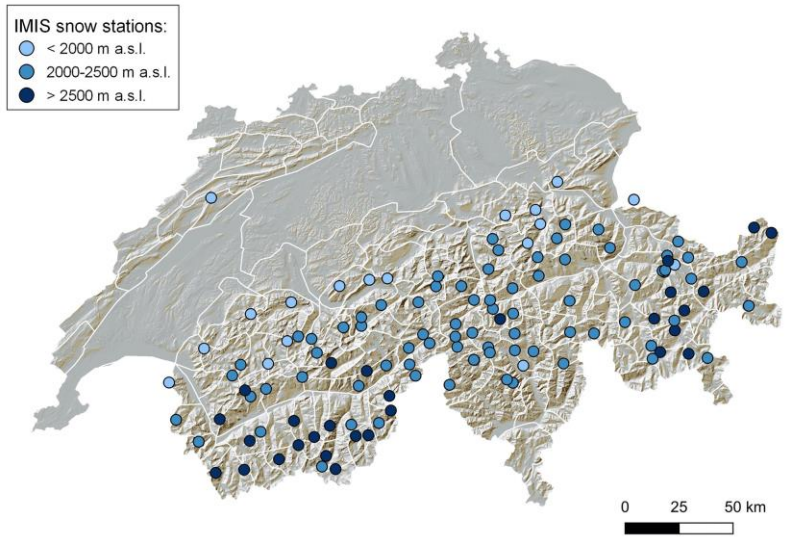
Large quantities of fresh snow and the wind-drifted snow are prone to triggering. Only isolated natural avalanches are possible, but they can be large. Single winter sport participants can release avalanches.
Backcountry touring and other off-piste activities call for extensive experience in the assessment of avalanche danger and caution.

Danger levels: 1 low | 2 moderate | 3 considerable | 4 high | 5 very high

Cartographie du danger d'avalanches

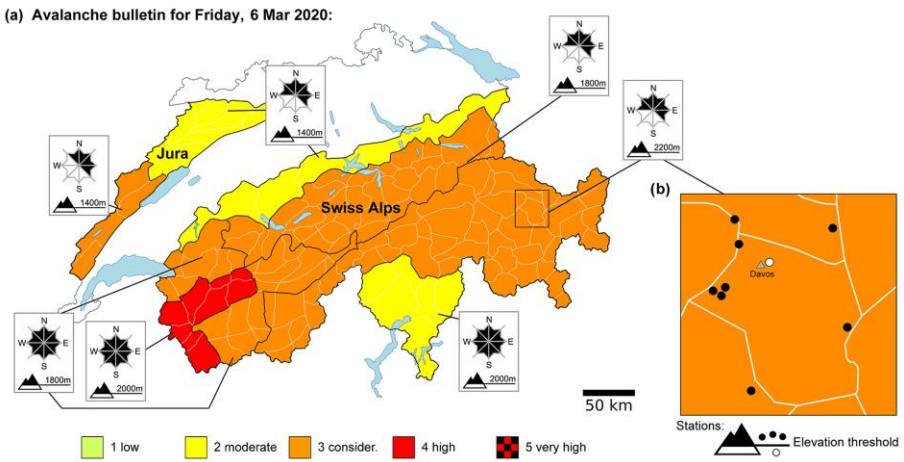
Inputs:

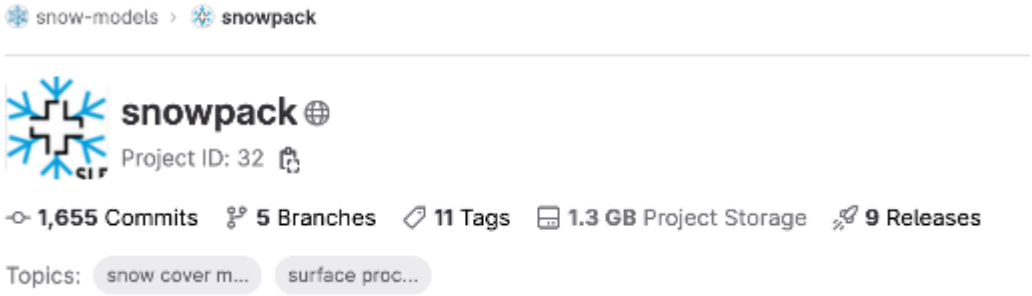
- stations météo dans les Alpes (réseau IMIS)
- Modèle numériques de la neige
- Modèle numérique du terrain



Outputs / target du modèle:

- Bulletin de la neige / avalanches





Pérez-Guillén, C., Techel, F., Hendrick, M., Volpi, M., van Herwijnen, A., Olevski, T., Obozinski, G., Pérez-Cruz, F., and Schweizer, J.: Data-driven automated predictions of the avalanche danger level for dry-snow conditions in Switzerland, Nat. Hazards Earth Syst. Sci., 22, 2031–2056, https://doi.org/10.5194/nhess-22-2031-2022, 2022.

Cartographie du danger d'avalanches

Saison 2018-2019
sur varies bulletins

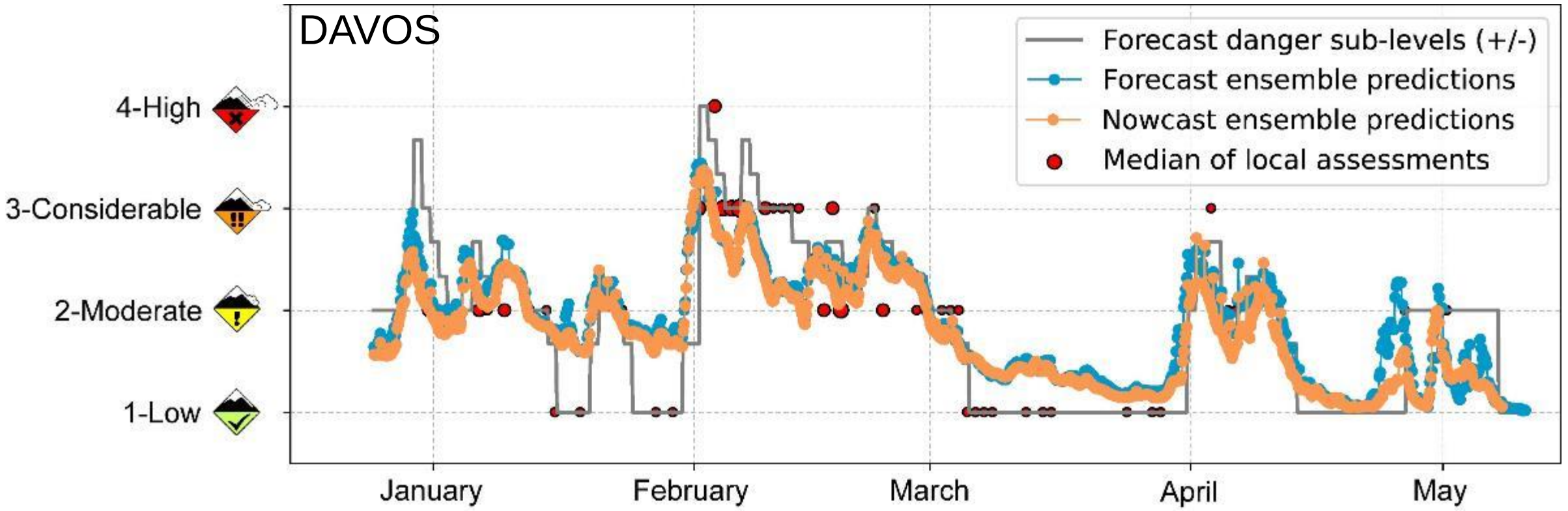
[redacted]

RF #1 danger level predictions for the winter season 2018-2019

[redacted]

Pérez-Guillén, C., Techel, F., Hendrick, M., Volpi, M., van Herwijnen, A., Olevski, T., Obozinski, G., Pérez-Cruz, F., and Schweizer, J.: Data-driven automated predictions of the avalanche danger level for dry-snow conditions in Switzerland, Nat. Hazards Earth Syst. Sci., 22, 2031–2056, https://doi.org/10.5194/nhess-22-2031-2022, 2022.

Cartographie du danger d'avalanches

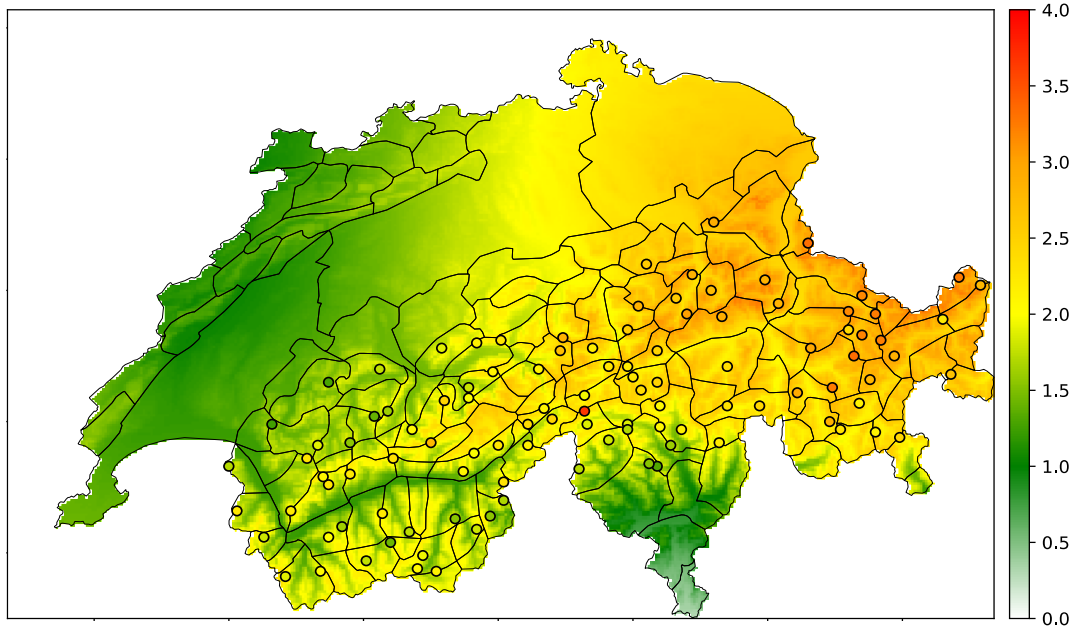


"nowcast" (jour même) vs "forecast" (24h en avance, COSMO)

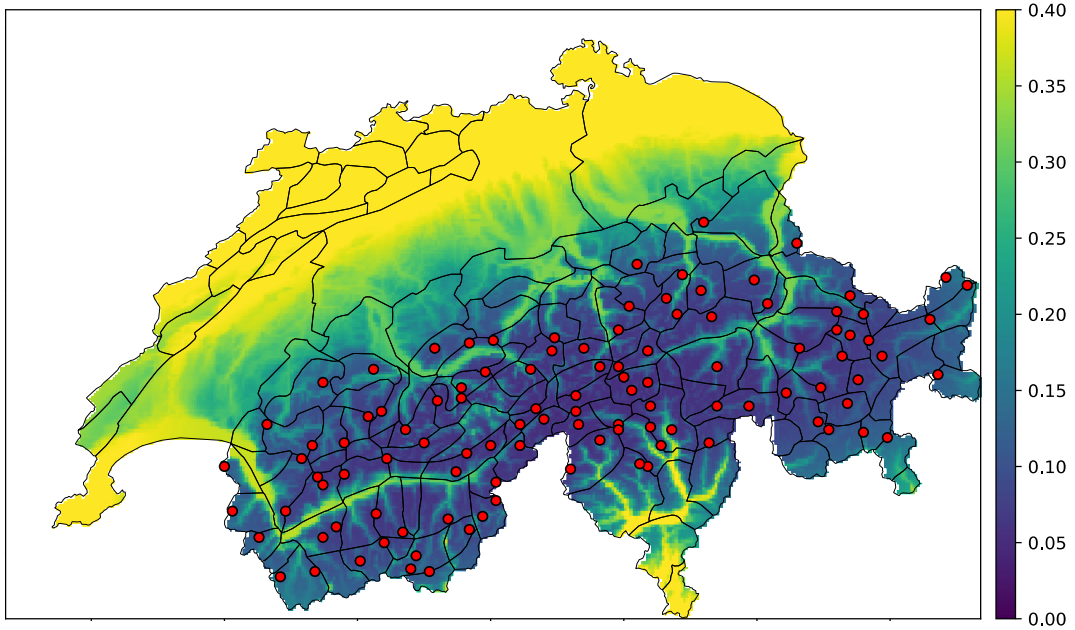
Pérez-Guillén, C., Techel, F., Hendrick, M., Volpi, M., van Herwijnen, A., Olevski, T., Obozinski, G., Pérez-Cruz, F., and Schweizer, J.: Data-driven automated predictions of the avalanche danger level for dry-snow conditions in Switzerland, Nat. Hazards Earth Syst. Sci., 22, 2031–2056, https://doi.org/10.5194/nhess-22-2031-2022, 2022.

Cartographie du danger d'avalanches

Interpolation à 1km²



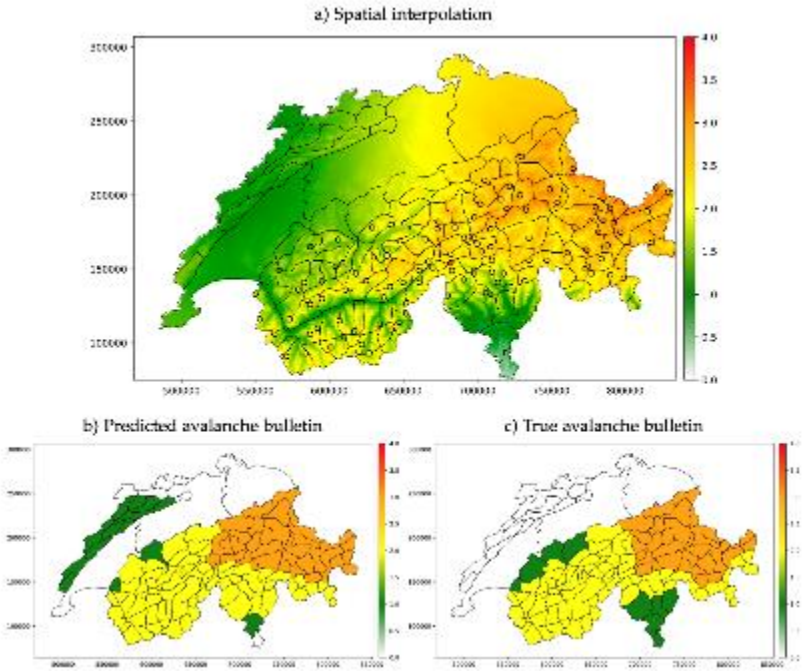
«moyenne» du niveau de danger par cellule



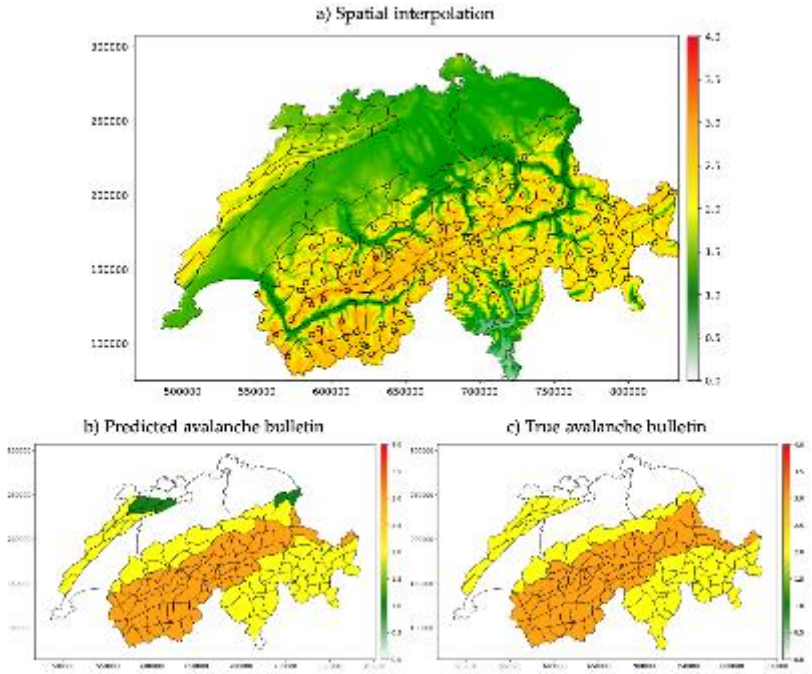
«variance» du score par cellule

Maissen, A., Techel, F., and Volpi, M.: A three-stage model pipeline predicting regional avalanche danger in Switzerland (RAvaFcast v1.0.0): a decision-support tool for operational avalanche forecasting, EGUsphere [preprint], https://doi.org/10.5194/egusphere-2023-2948, 2024.

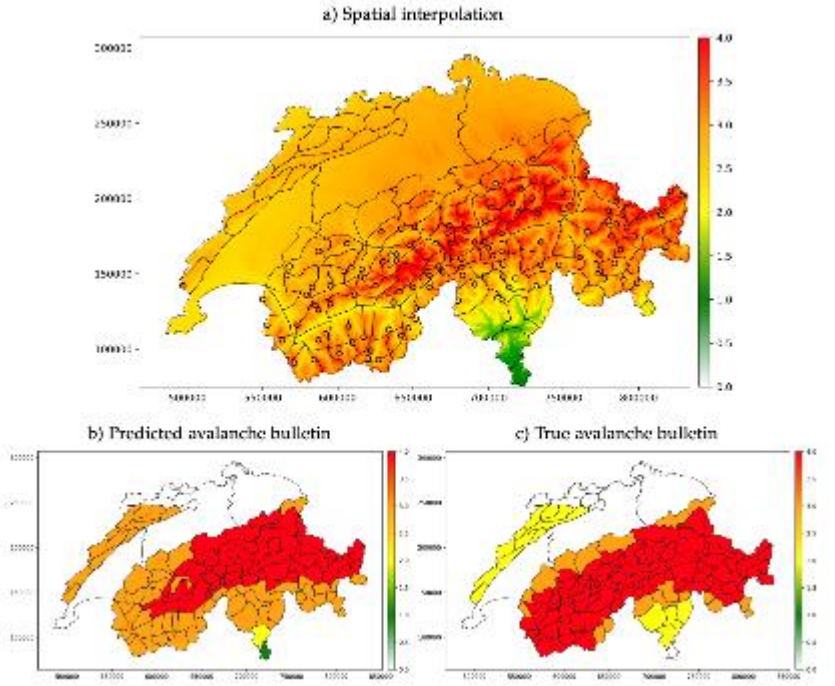
Swiss Data Science Center



7 Jan 2019



11 Mar 2019



15 Jan 2019

Maissen, A., Techel, F., and Volpi, M.: A three-stage model pipeline predicting regional avalanche danger in Switzerland (RAvaFcast v1.0.0): a decision-support tool for operational avalanche forecasting, EGUsphere [preprint], https://doi.org/10.5194/egusphere-2023-2948, 2024.

Quel avenir pour l'IA dans les sciences?

Swiss Data Science Center

Explainabilité → interpretabilité → causalité: Les NNs sont souvent des «black boxes», on ne comprend pas ce qu'ils font (trop compliqués), ni la relation entres les variables.

Réalisme: Des modèles IA purement «statistiques» ne sont pas souvent physiquement rigoureux (même si précis).

«Humans-in-the-loop»: Des experts peuvent corriger les modèles en temps réel.

Multimodalités: Rarement la nature se décrit avec un seul jeu de données

Swiss Data Science Center

Merci

Qu’est-ce que c’est l’IA?

Swiss Data Science Center

Qu’est-ce que c’est l’IA?

« C'est la science et l'ingénierie de la fabrication de machines intelligentes, en particulier de **programmes informatiques intelligents**. Elle est liée à la tâche similaire qui consiste à utiliser des ordinateurs pour comprendre l'intelligence humaine, mais l'IA ne doit pas se limiter aux méthodes qui sont biologiquement observables. »

McCarthy

[McCartthy, what is artificial intelligence, 2007]
https://www-formal.stanford.edu/jmc/whatisai.pdf

Réseaux de neurones (neural networks, NN)

Swiss Data Science Center

A bit of history...

$$\frac{\partial E_p}{\partial w_{ji}} = \frac{\partial E_p}{\partial o_{pj}} \frac{\partial o_{pj}}{\partial w_{ji}}$$

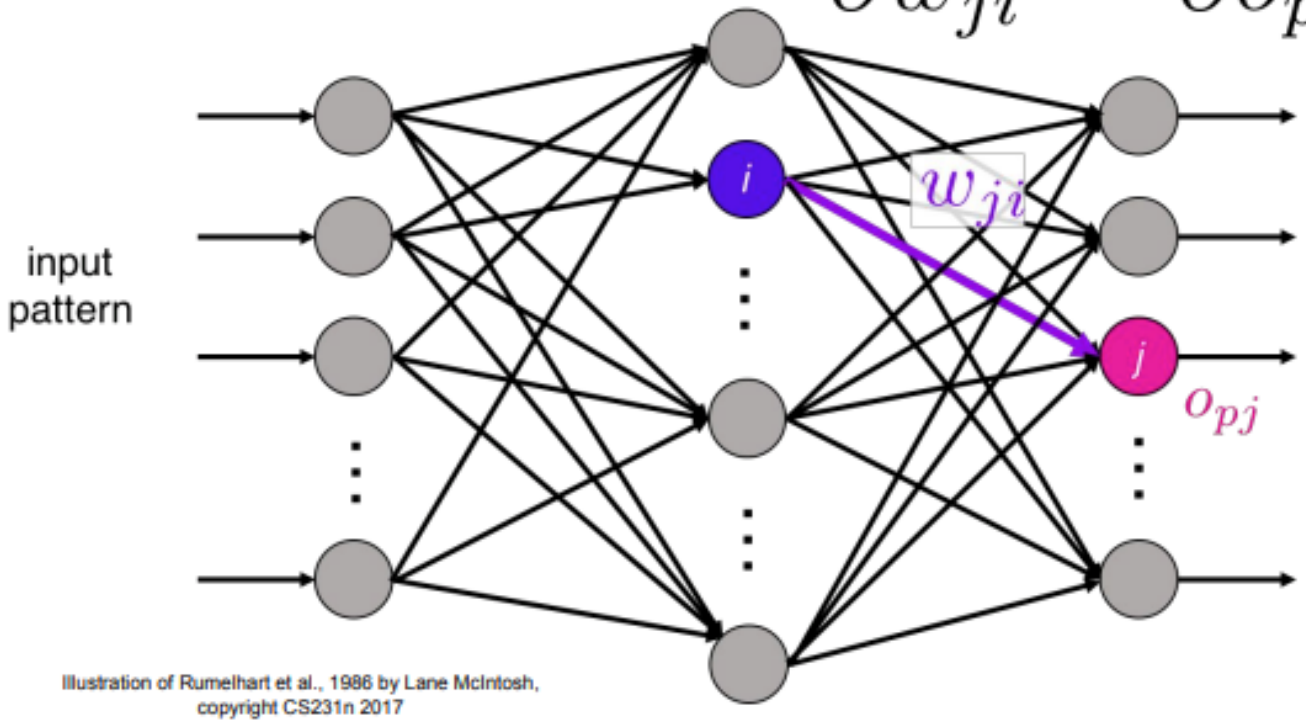


Illustration of Rumelhart et al., 1986 by Lane McIntosh, copyright CS231n 2017

E := erreur entre étiquette et prédiction du modelé

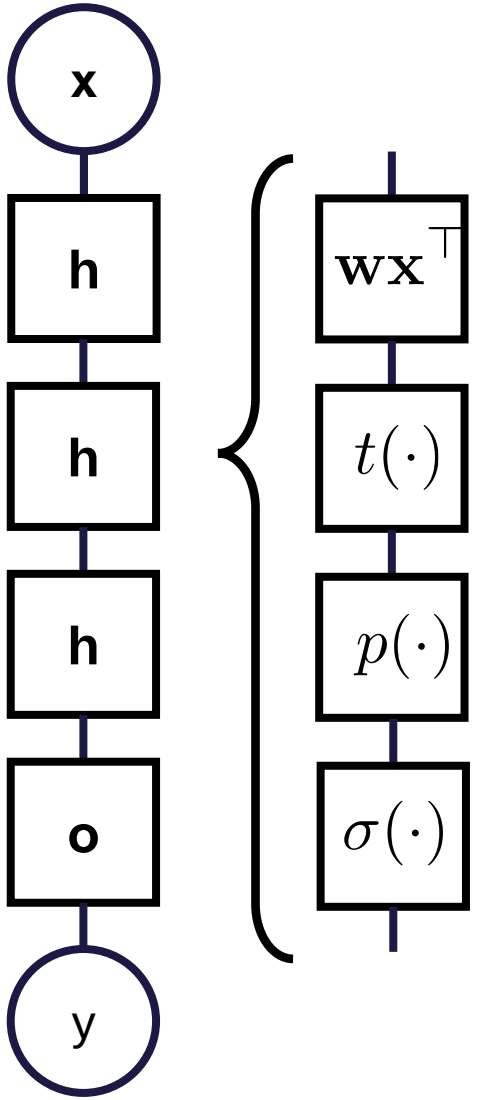
Mise a jour du modèle = Reduction de cet erreur

Réseaux de neurones (neural networks, NN)

Swiss Data Science Center

Qu’est-ce qu’il a changé depuis les années 60?

- Beaucoup de données (p.ex. Google, Meta, …)
- Puissance de calcul très bon marché
- Spécialisation (p.ex. nous à SDSC)
- Accessibilité dans le «state-of-the-art»



Qu’est-ce qu’il faut pour l’IA?

Swiss Data Science Center

Pour créer de systèmes:

- **Beaucoup de données.** Mieux si prétraitées et propre (information plus facile a extraire)
- **Puissance de calcul.** GPUs, CPUs, réseau et filesystem rapides!
- **Software, code et librairies.** Python est devenu le standard, beaucoup d’open source.

Pour utiliser les systèmes:

- Navigateur internet, connexion (IA dans le cloud)
- Ordinateur moderne (laptop souvent suffit)

SDSC data science de l'environnement

Swiss Data Science Center

Dangers naturels: cartes de susceptibilité, détection d'événements
Sciences du climat: modèles physiques → IA, prévision
Ecologie: cartographie, détection (camera traps), mise a jour inventaires
Agriculture: mapping, estimation, 3D reconstruction
Géographie: science urbaines, cartographie

Données:

Satellites
Modèles physiques
Stations météo (géolocalisés)
Images naturelles (géolocalisés)
…

Backup application additionelle: cartographie

Swiss Data Science Center

En utilisant de **réseaux de neurones profondes (deep NN):**

- **Apprendre automatiquement** changements dans le bâti à Londre
- **Images géolocalisées** comme source d'information
- Apprendre des représentations des images via self-supervised learning (même algo derrière les chatGPTs, etc)

Ex1: Cartographie des changements urbaines

- Est-ce qu'on peut extraire de l'informations sur la structure urbaine[1] depuis images géolocalisées? p.ex. google streetview?
- Est-ce qu'on peut le faire sans avoir des exemples de changements?

[1] Cette méthode est indépendante de l'application!

Ex1: Cartographie des changements urbaines

- Une IA (réseau de neurones) – seulement en connaissant où et quand l'image a été prise – apprend les correspondances entre images
- Ce qui ne corresponde pas, est un changement

Ex1: Cartographie des changements urbaines

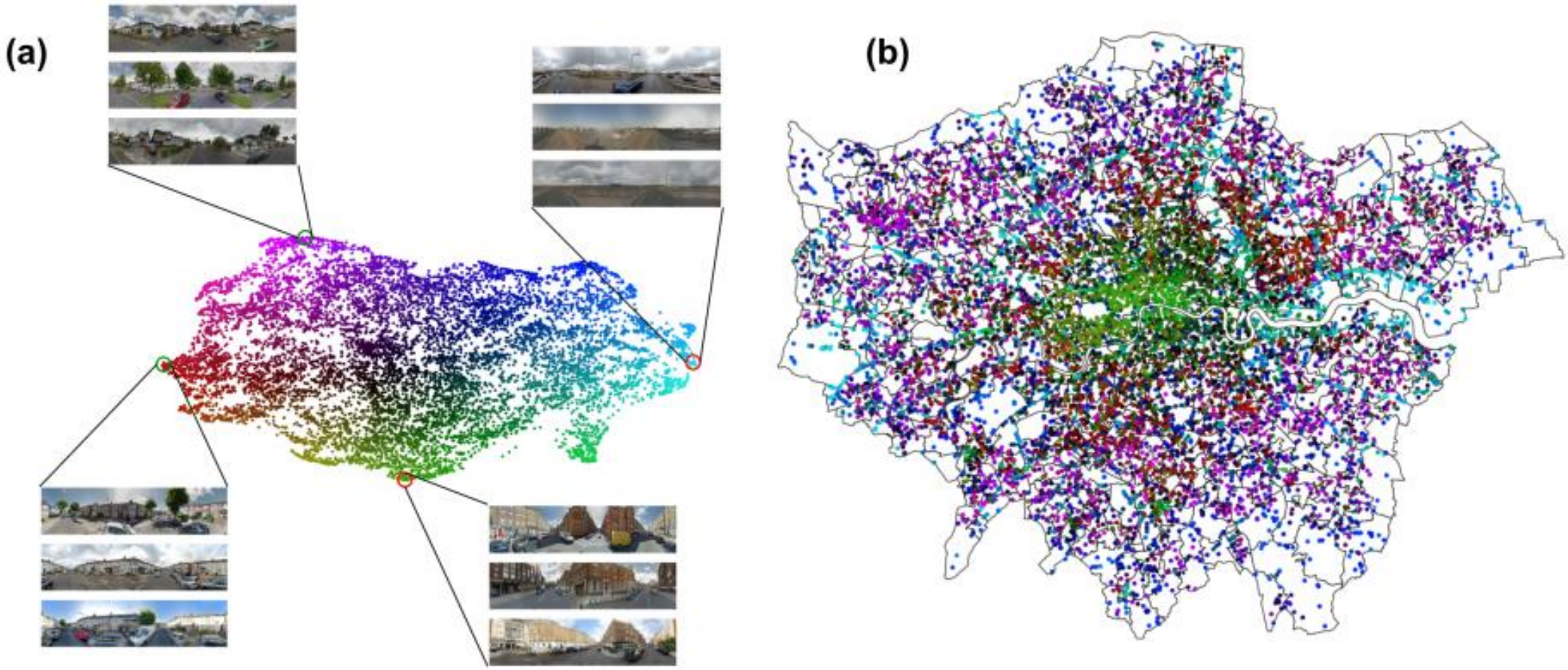
- Une IA (réseau de neurones) – seulement en connaissant où et quand l'image a été prise – apprend les correspondances entre images
- Ce qui ne corresponde pas, est un changement

2008

2018

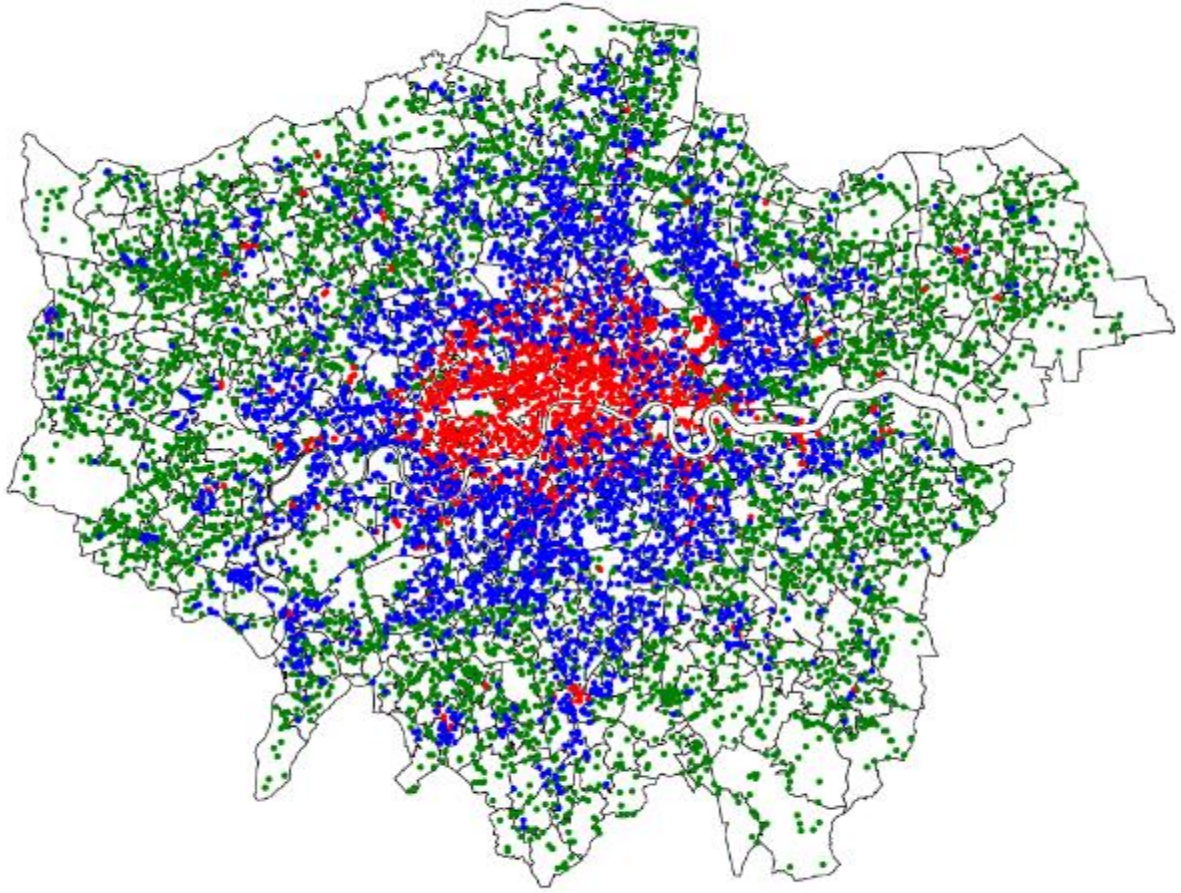
Ex1: Cartographie des changements urbaines

Visualisation de la «représentation interne» des images apprise par l'IA

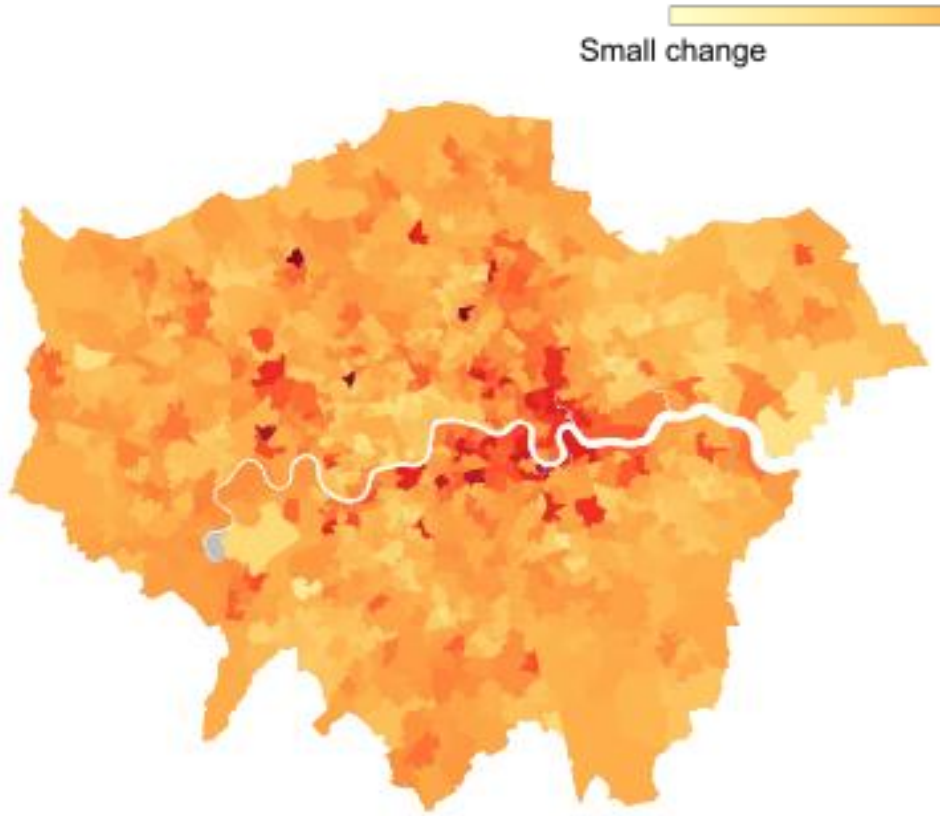


Ex1: Cartographie des changements urbaines

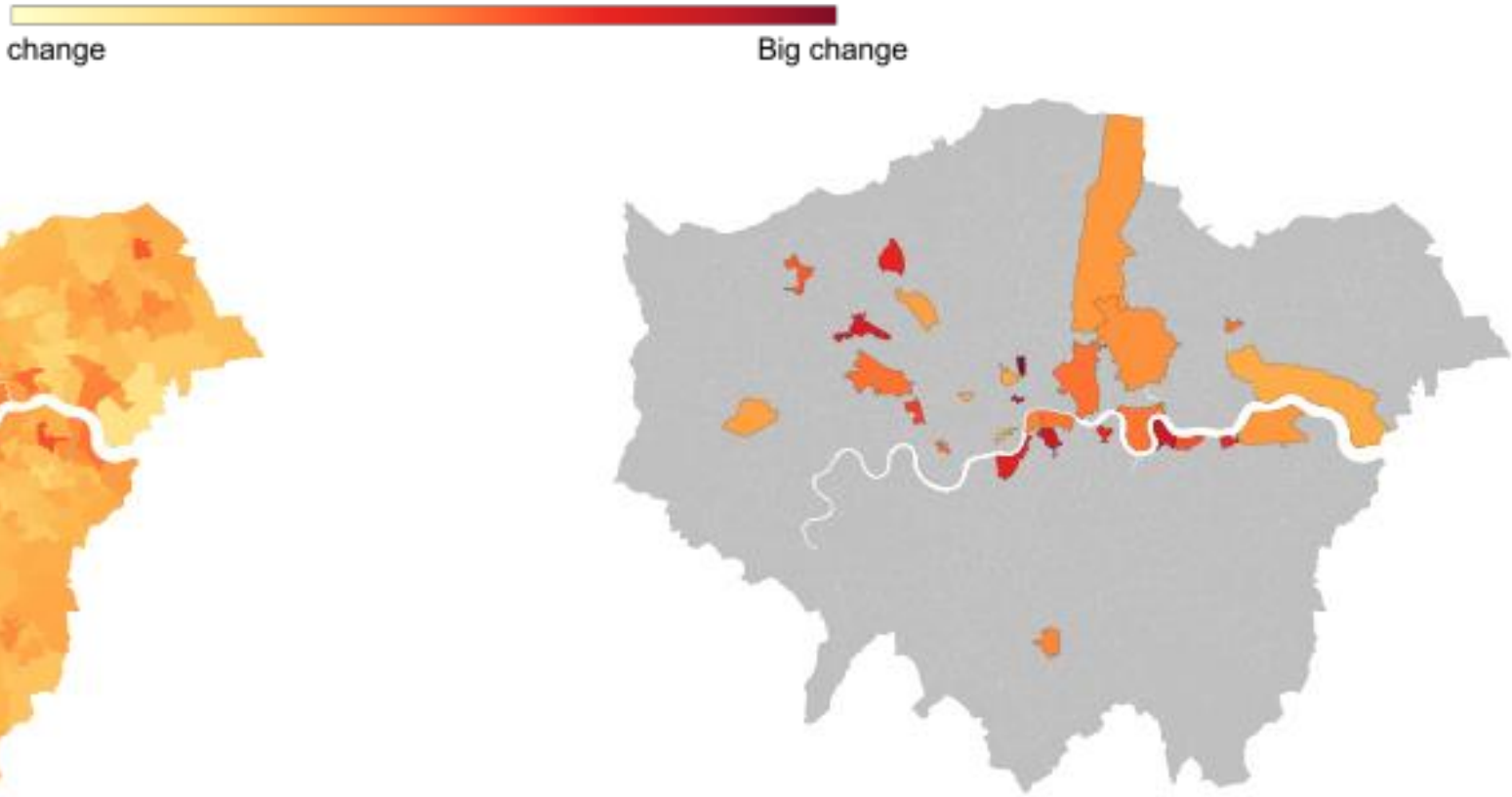
L’IA apprends plus que juste changement bâti / pas bâti



Ex1: Cartographie des changements urbaines



Prédiction de changements par l'IA



Prédiction par la ville de Londres
(potentiel de développement)