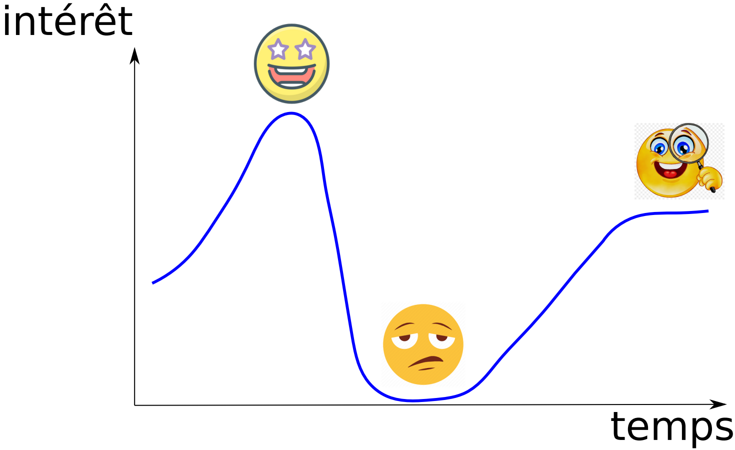
Comment développer des outils d'apprentissage automatique pour les géosciences ?

Un aperçu de différents usages de réseaux de neurones convolutifs.

Sophie Giffard-Roisin

Chargée de Recherche IRD, ISTerre (Grenoble)
sophie.giffard@univ-grenoble-alpes.fr

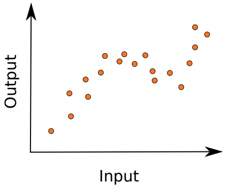
Intérêt des géoscientifiques pour l'IA / machine learning



Machine learning: what is it?

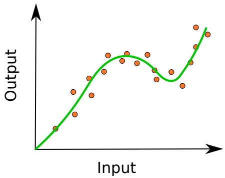
1) database collection

Use information from existing examples ...



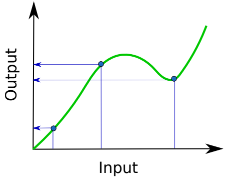
2) training

... to learn from them patterns or relations ...

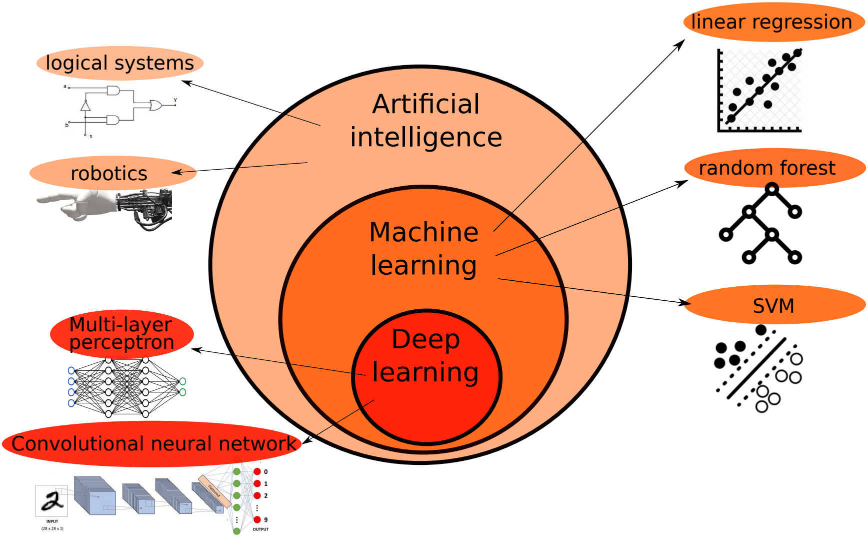


3) prediction

... that can then be applied to new examples



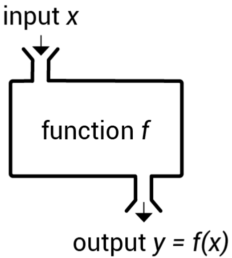
Artificial Intelligence and Machine learning



Machine learning: what for?

- *Big data*: data is often easy to collect
- But we don't always know the relations between variables (ex. physics)
- ML = algorithm that will learn from the data directly

Machine learning model = function approximation



Where is it used today?

Natural language processing: translation

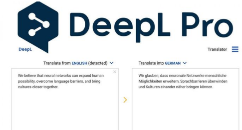


Image detection

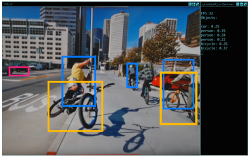
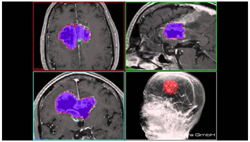
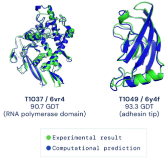


Image segmentation



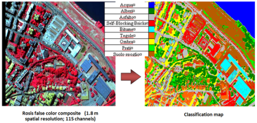
Personalised advertising

Protein unfolding

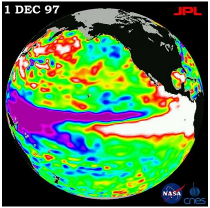


And in geosciences?

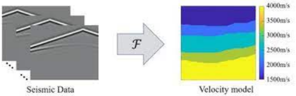
Remote sensing processing



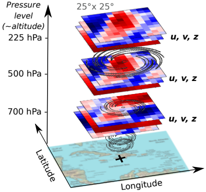
El nino prediction



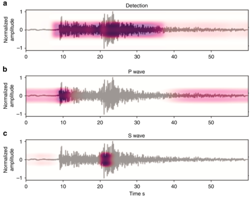
Seismic inversion



Tropical cyclone trajectory prediction



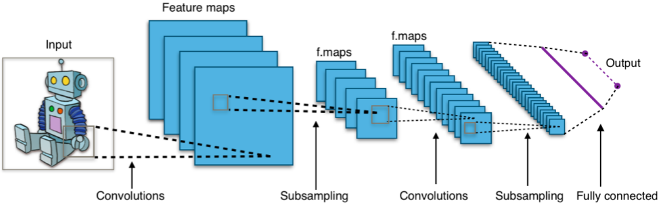
Earthquake detection



Outline

Convolutional neural network

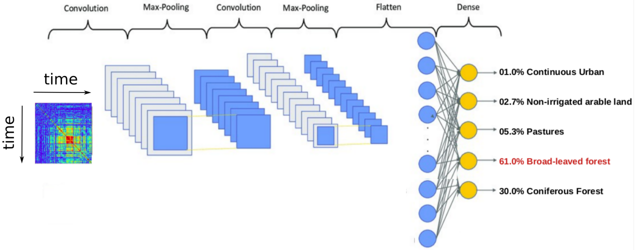
$$f_\theta(img) = output$$



- Suited for image-like data
- Non-linear/complex relationships

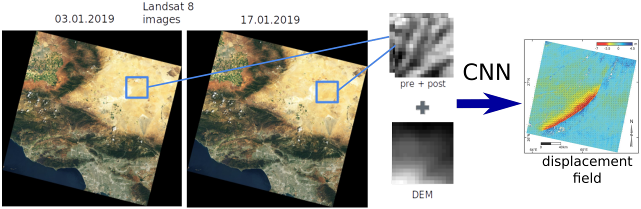
Temporal coherence InSAR matrix classification

Marie-Pierre Doin, Salah-Eddine Boudaour (M2)



Optical image correlation for deformation estimation

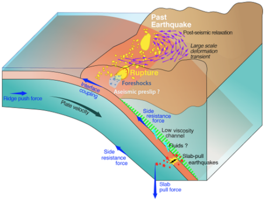
James Hollingsworth, Tristan Montagnon



Outline

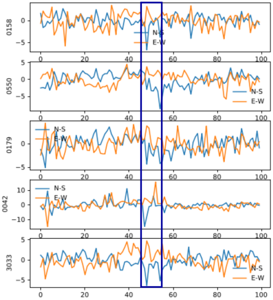
GPS time series for characterizing earthquakes (and SSEs)

ERC Deep-Trigger - Anne Socquet

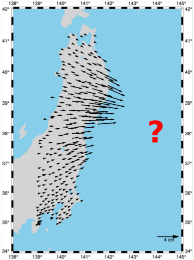


Giuseppe Costantino (PhD)

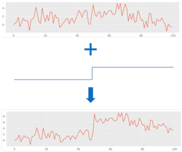
GPS time series for characterizing earthquakes (and SSEs)



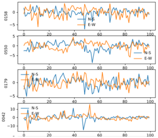
- location
- magnitude



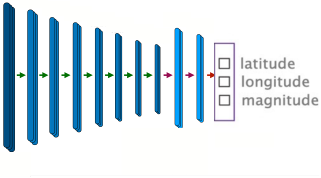
- Learn from synthetic database (20 K samples)
- Okada dislocation model + realistic noise



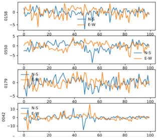
Data: list of time series



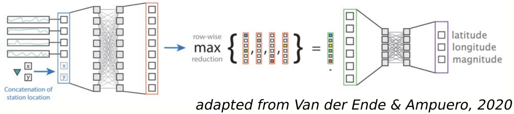
–> 1D convolutional n. network



Data: list of time series

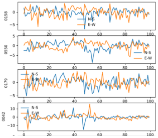


–> 1D convolutional n. network

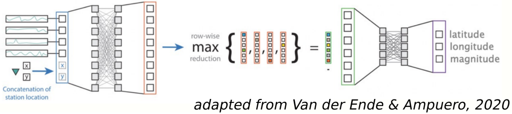


adapted from Van der Ende & Ampuero, 2020

Data: list of time series

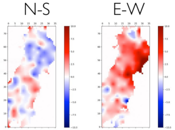


–> 1D convolutional n. network

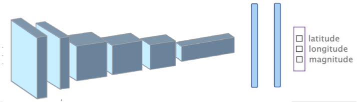


adapted from Van der Ende & Ampuero, 2020

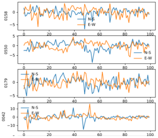
interpolated images



–> 2D convolutional n. network



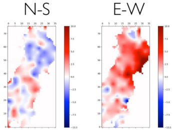
Data: list of time series



–> 1D convolutional n. network

adapted from Van der Ende & Ampuero, 2020

interpolated images



–> 2D convolutional n. network

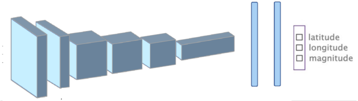
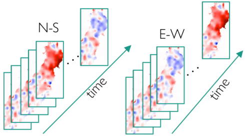
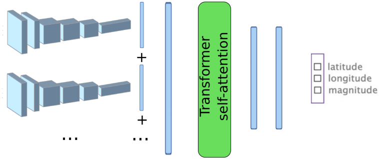


image time series



–> 2D CNN + transformer



Suite de la session:

- Introduction et machine learning pour problèmes 2D

- **Detection de séismes (1D CNN) et interprétabilité**

  Josipa Majstorović

- **Détection de signaux sismologiques précoces par IA**

  Kevin Juhel

CAMBRIDGE UNIVERSITY PRESS

Open access

OPEN FOR SUBMISSIONS: www.cambridge.org/eds

ENVIRONMENTAL DATA SCIENCE

An interdisciplinary, open access journal dedicated to the potential of artificial intelligence and data science to enhance our understanding of the environment, and to address climate change.

Editorial Team

Data and methodological scope, includes:

Machine learning; Artificial intelligence; Statistics; Data mining
Computer vision; Econometrics Data science, broadly defined.

Environmental scope, includes:

- Water cycle, atmospheric science (including air quality, climatology, meteorology, atmospheric chemistry & physics, paleoclimatology)
- Climate change (including carbon cycle, transportation, energy, and policy)
- Sustainability and renewable energy (the interaction between human processes and ecosystems, including resource management, transportation, land use, agriculture and food)
- Biosphere (including ecology, hydrology, oceanography, glaciology, soil science)
- Societal impacts (including forecasting, mitigation, and adaptation, for environmental extremes and hazards)
- Environmental policy and economics

5 Reasons to submit to *EDS:*

1. Gain **quality peer review** feedback on your work from editors and reviewers who have expertise in the use of data science in environmental disciplines.
2. Publish, if accepted, your work under **open licensing** to make it freely available to read, distribute and re-use in a venue that also provides you with options for making pre-prints, data and code openly available.
3. Reach a wider audience through **impact statements** published with articles, conveying the significance of your work.
4. **Align your conference or workshop** with *Environmental Data Science*, as a venue that can make peer-reviewed outputs open and discoverable.
5. Help us **build a community** of authors, reviewers and editors advocating for the transformative potential of data science for a better understanding of the environmen

More details at @envdatascience cambridge.org/eds @envdatascience