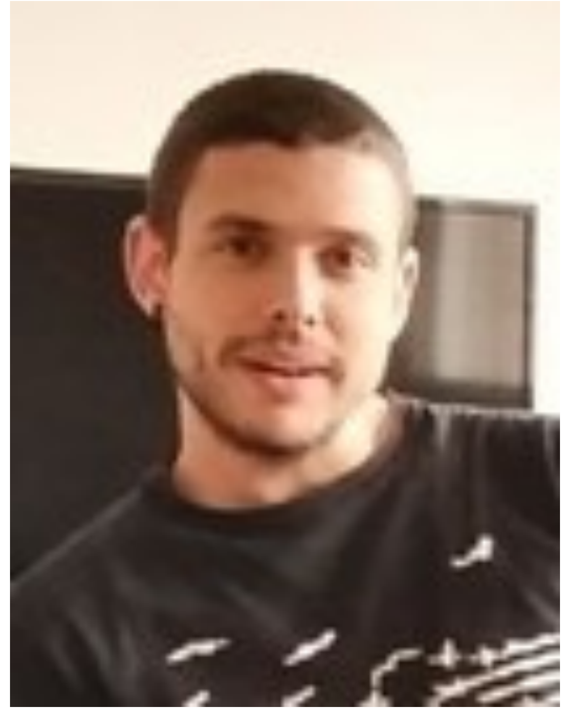
Deep Learning for Ocean and Atmosphere (OA) Dynamics
June 26-30, 2023

Brest

Ronan
Fablet

Lucas
Drumetz



Carlos Granero
Belinchon

Grenoble

Bruno
Deremble

Audrey
Monsimer

Emmanuel
Cosme

Objective

- Introduction of the main Deep learning concepts (we think relevant for OA sciences)
- Introduction to PyTorch Deep Learning “ecosystem”
- Training through practice (lab sessions and project sessions)
- ability to deploy a deep learning approach for OA-related topics

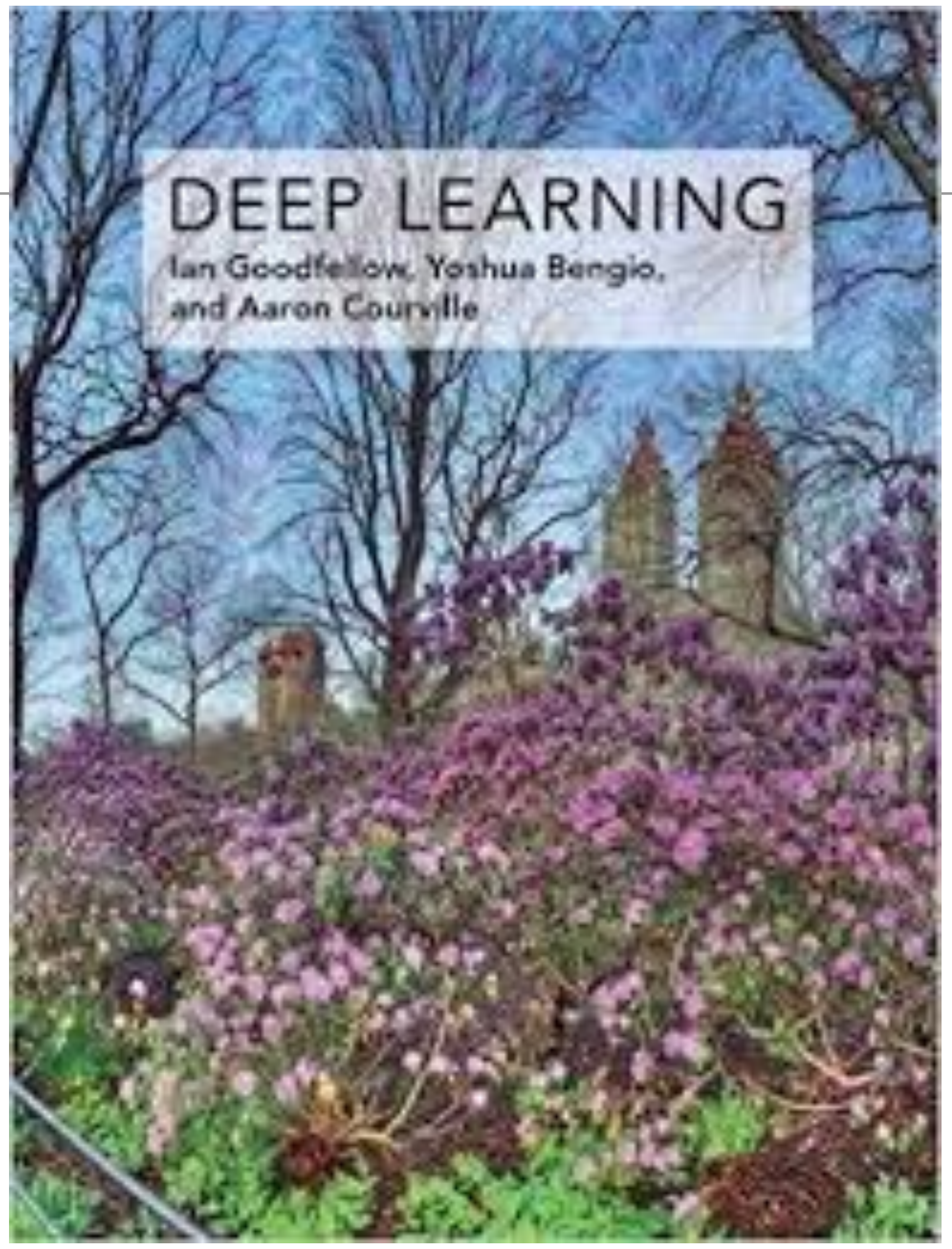
Program

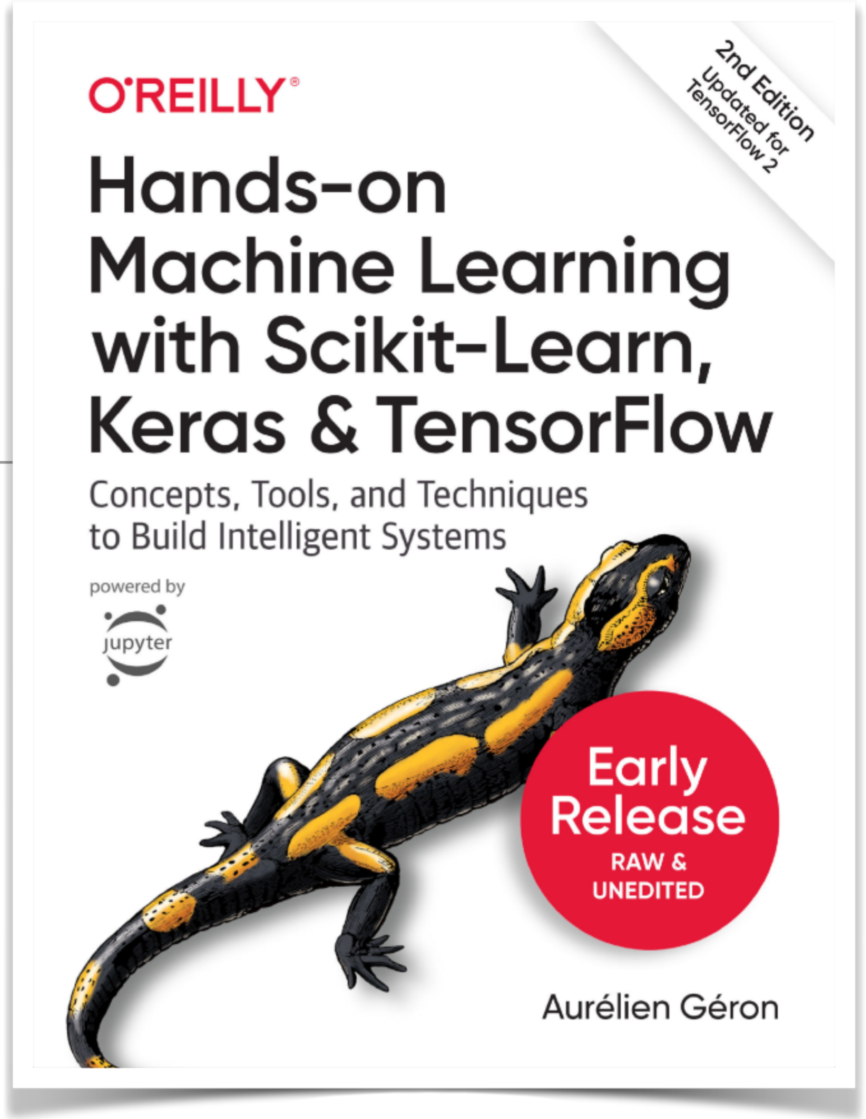
Day #1, 9.30am-12.30pm **Introduction to Deep Learning** ● What's learning ? ● MLP / backprop ● Practice on toy regression/classification examples	**Day #1, 2.00pm-5.30pm** **Project session #1:** ● Project selection ● Learning-based problem formulation ● Exploratory analysis/visualization of the considered dataset
Day #2, 9.30am-12.30pm **Convolutional Neural Networks** ● From MLP to CNN ● Deep Learning methodology ● Practice on MNIST digit classification	**Day #2, 2.00pm-5.30pm** **Project session #2:** ● Tutorial on PyTorch Lighning ● Selection of neural architectures ● Design of the training scheme
Day #3, 9.30am-12.30pm **Auto-encoders and Generative models** ● Auto-encoder architectures (Dense AEs/PCA, Convolutional AEs, U-Net) ● Practice on MNIST dataset ● Opening towards generative models: VAE, NF, GAN	**Day #3, 2.00pm-5.30pm** **Project session #3:** ● First results for a simple architecture ● Sensitivity analysis ● Updated architectures
Day #4, 9.30am-12.30pm **Recurrent Neural Networks** ● RNN / LSTM ● Neural ODE / PINN ● Practice on L63 system	**Day #4, 2.00pm-5.30pm** **Project session #4:** ● Optimization of the architecture ● Synthesis of the experiments, incl. the benchmarking of several architectures
Day #5, 9.30am-12.30pm **Deep Learning and Inverse Problems** ● DL, AutoDiff and minimization ● Deep inverse models ● Deep unfolded architectures ● Practice on L63 system	**Day #5, 2.00pm-5.30pm** **Project session #5:** ● Presentation for each project (10'+5' for each project)

- **Book: Deep Learning**

Goodfellow, Bengio, Courville, MIT Press

Online version http://www.deeplearningbook.org/





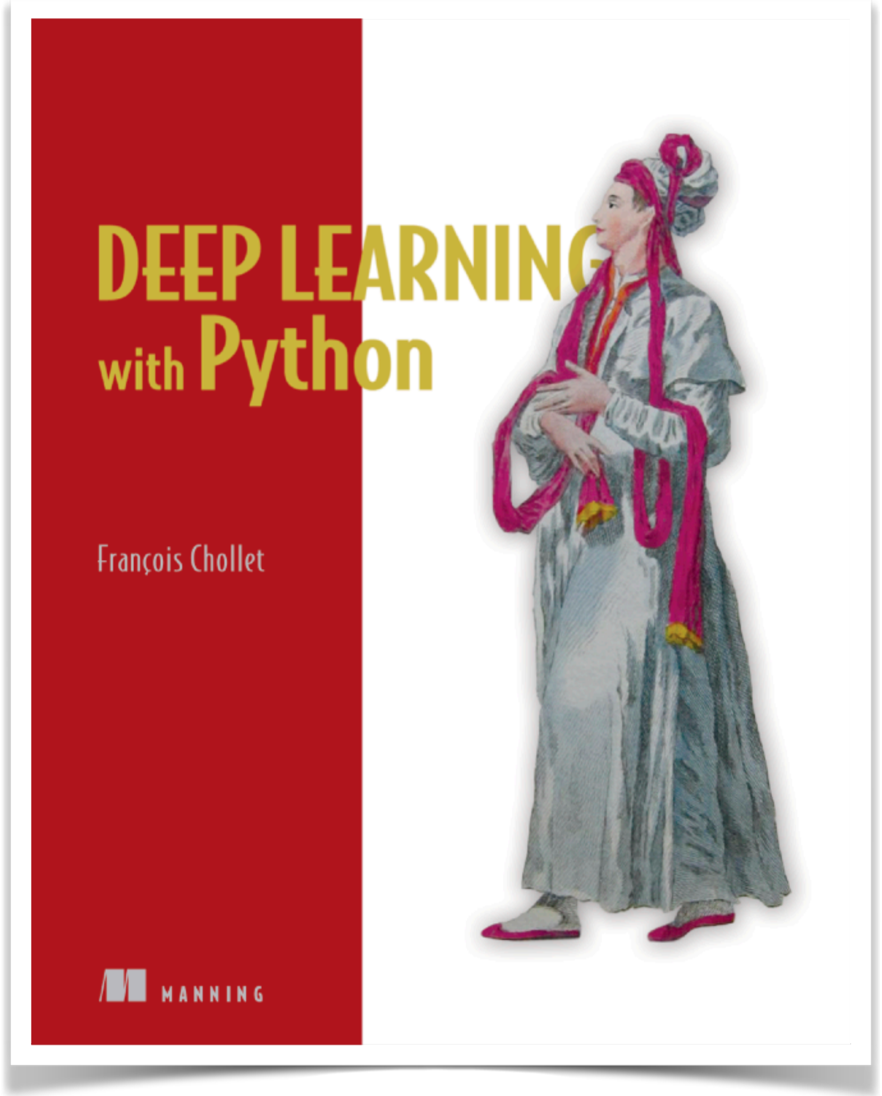
- **Online course by Andrew Ng (Stanford/Baidu)**

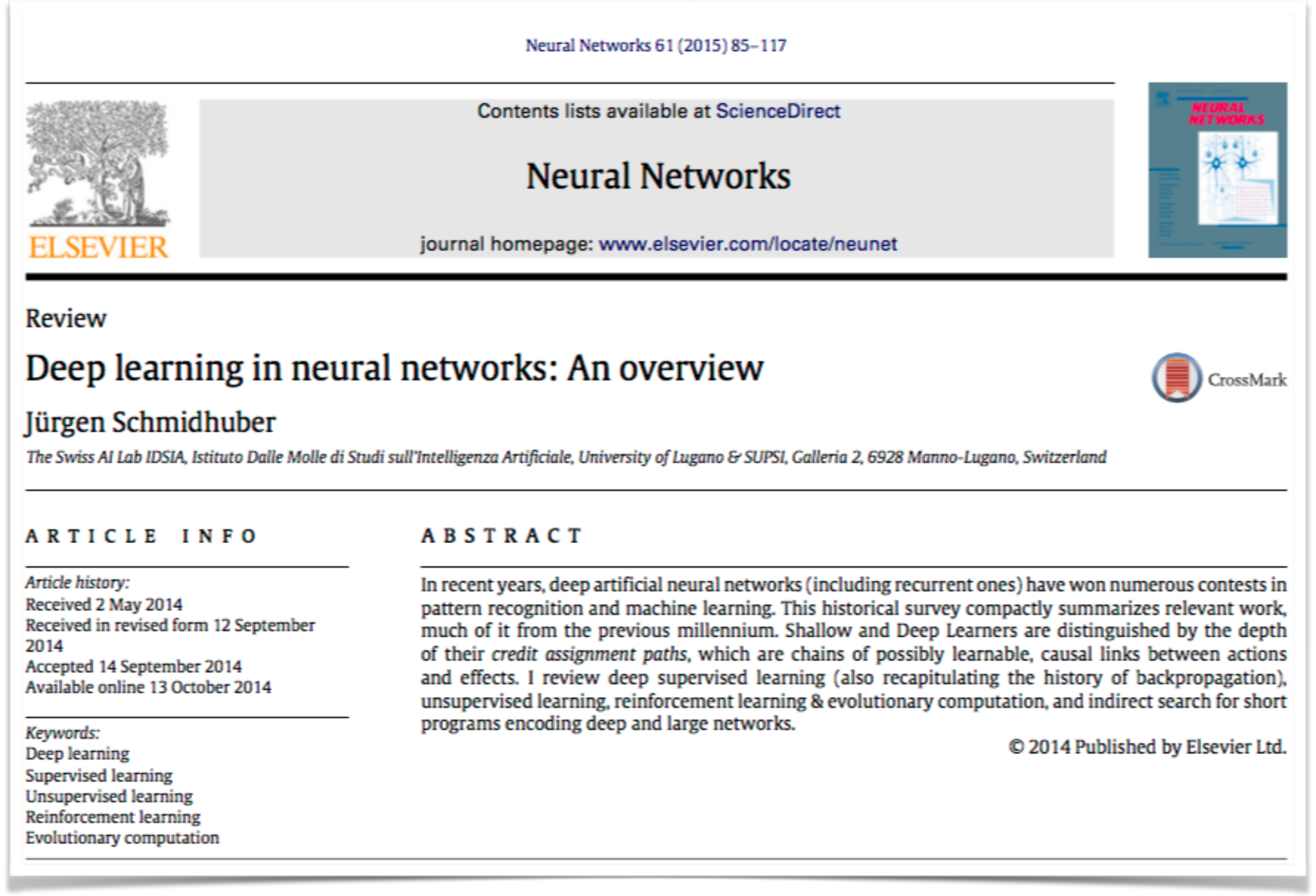
  Youtube: link

  Online course on Coursera: link

- **Review paper: Deep learning in neural networks by J. Schmidhube pdf: link**

- **Github repo: https://github.com/CIA-Oceanix/DLOA2023**




Neural Networks 61 (2015) 85–117

Contents lists available at ScienceDirect

Neural Networks

journal homepage: www.elsevier.com/locate/neunet

ELSEVIER

Review

Deep learning in neural networks: An overview

Jürgen Schmidhuber

The Swiss AI Lab IDSIA, Istituto Dalle Molle di Studi sull'Intelligenza Artificiale, University of Lugano & SUPSI, Galleria 2, 6928 Manno-Lugano, Switzerland

ARTICLE INFO

Article history:
Received 2 May 2014
Received in revised form 12 September 2014
Accepted 14 September 2014
Available online 13 October 2014

Keywords:
Deep learning
Supervised learning
Unsupervised learning
Reinforcement learning
Evolutionary computation

ABSTRACT

In recent years, deep artificial neural networks (including recurrent ones) have won numerous contests in pattern recognition and machine learning. This historical survey compactly summarizes relevant work, much of it from the previous millennium. Shallow and Deep Learners are distinguished by the depth of their *credit assignment paths*, which are chains of possibly learnable, causal links between actions and effects. I review deep supervised learning (also recapitulating the history of backpropagation), unsupervised learning, reinforcement learning & evolutionary computation, and indirect search for short programs encoding deep and large networks.

Requirements

- access to discord server OK?
- google colab account?

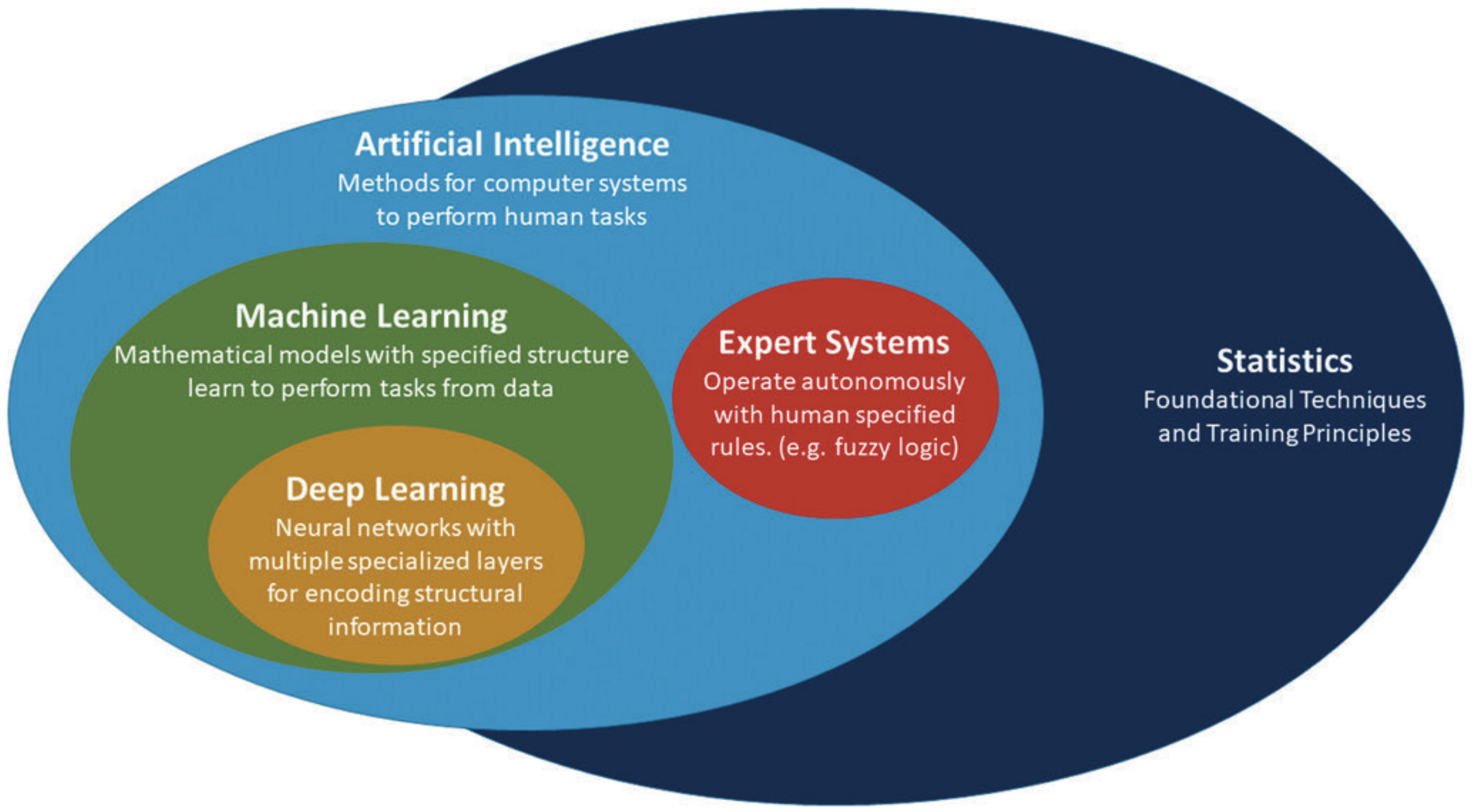
Lecture #1: The basics

- Deep Learning facts
  - definition, emergence, uses, culture, softwares, carbon footprint
- How learning works
  - description, mathematical formulation, basic concepts, vocabulary
- The Multi Layer Perceptron
  - Perceptron, activation functions, feedforward networks, layers…
- The learning process
  - Loss function, Gradient descent, backpropagation
- Introduction to Pytorch
  - Basic tips and practice

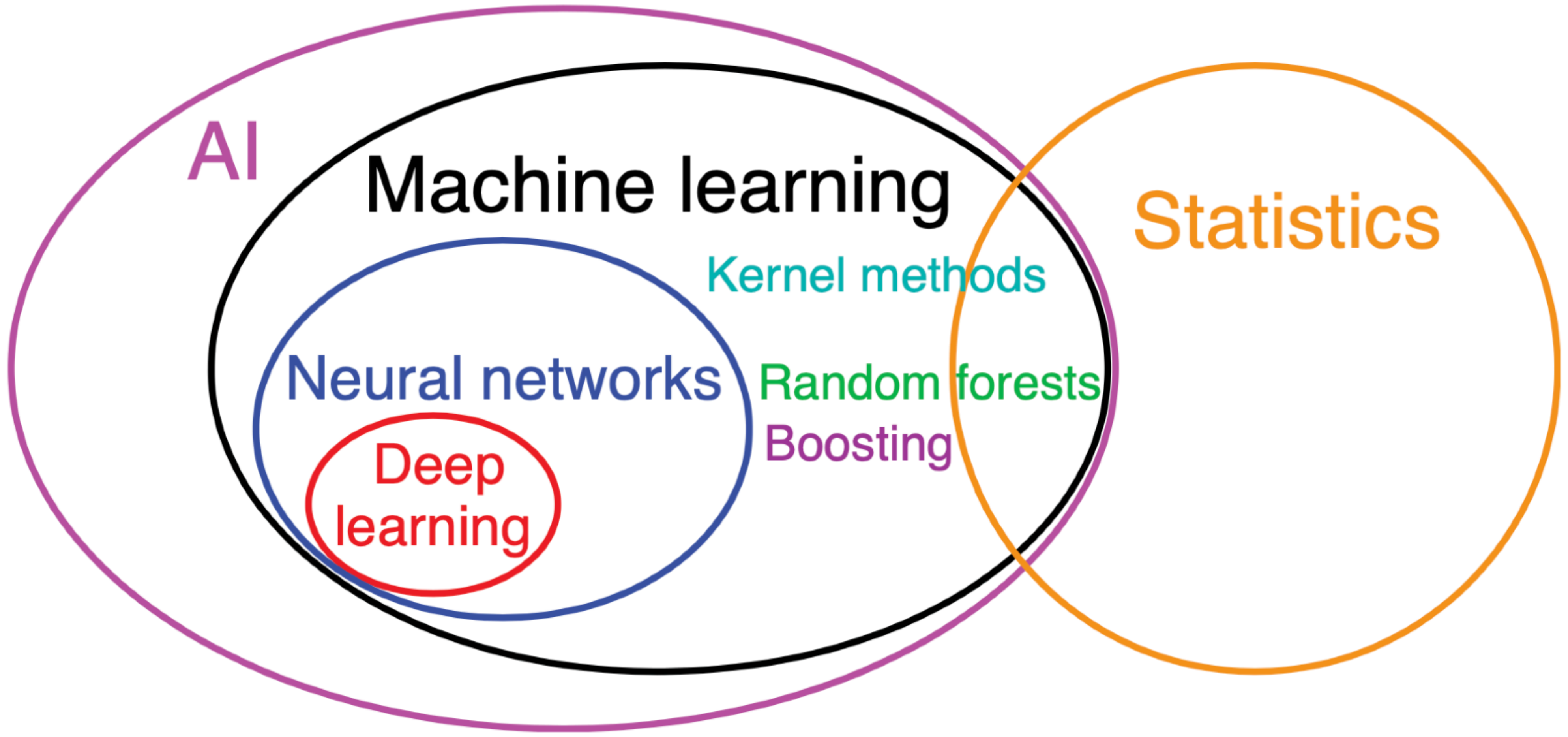
Deep Learning facts

Definition, emergence, uses, culture, softwares, carbon footprint

AI, Machine Learning, Deep Learning, etc



From Haupt et al (2022)



From Hsieh (2022)

Deep learning has become the standard framework for a wide range of applications in computer vision, natural language processing, signal processing…

Machine Learning and "traditional" simulation

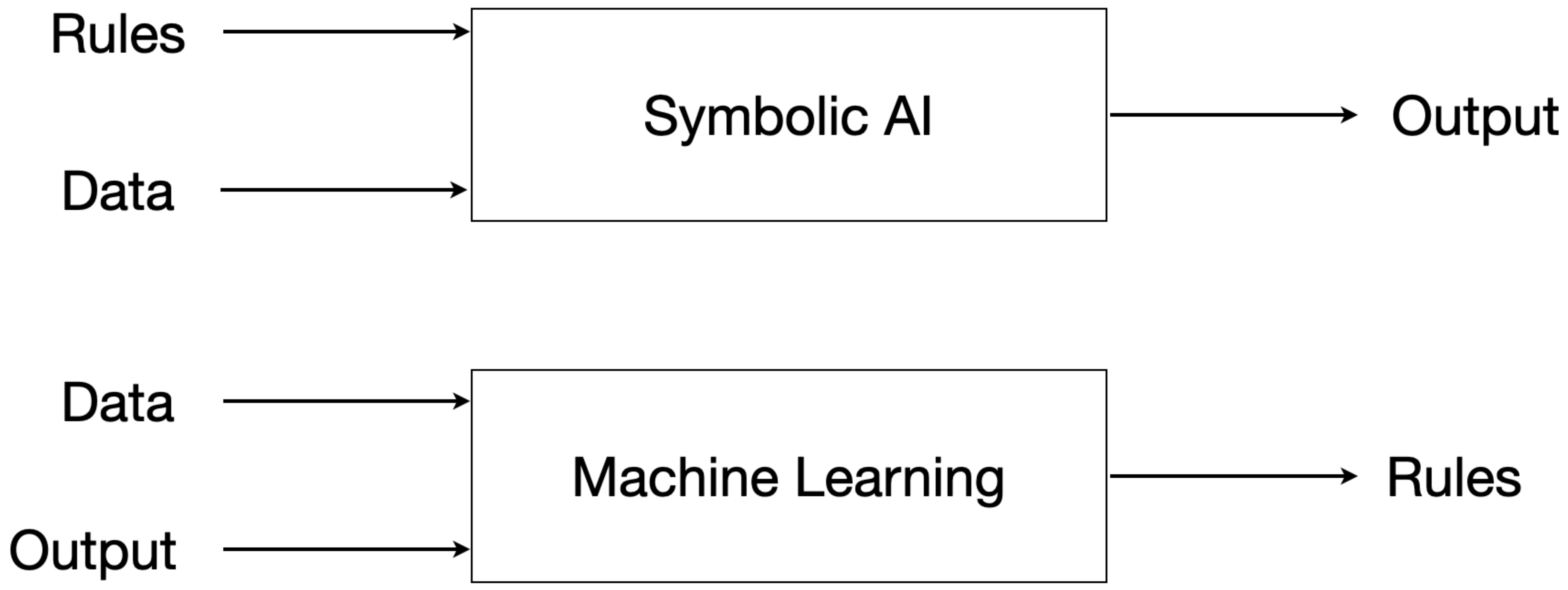
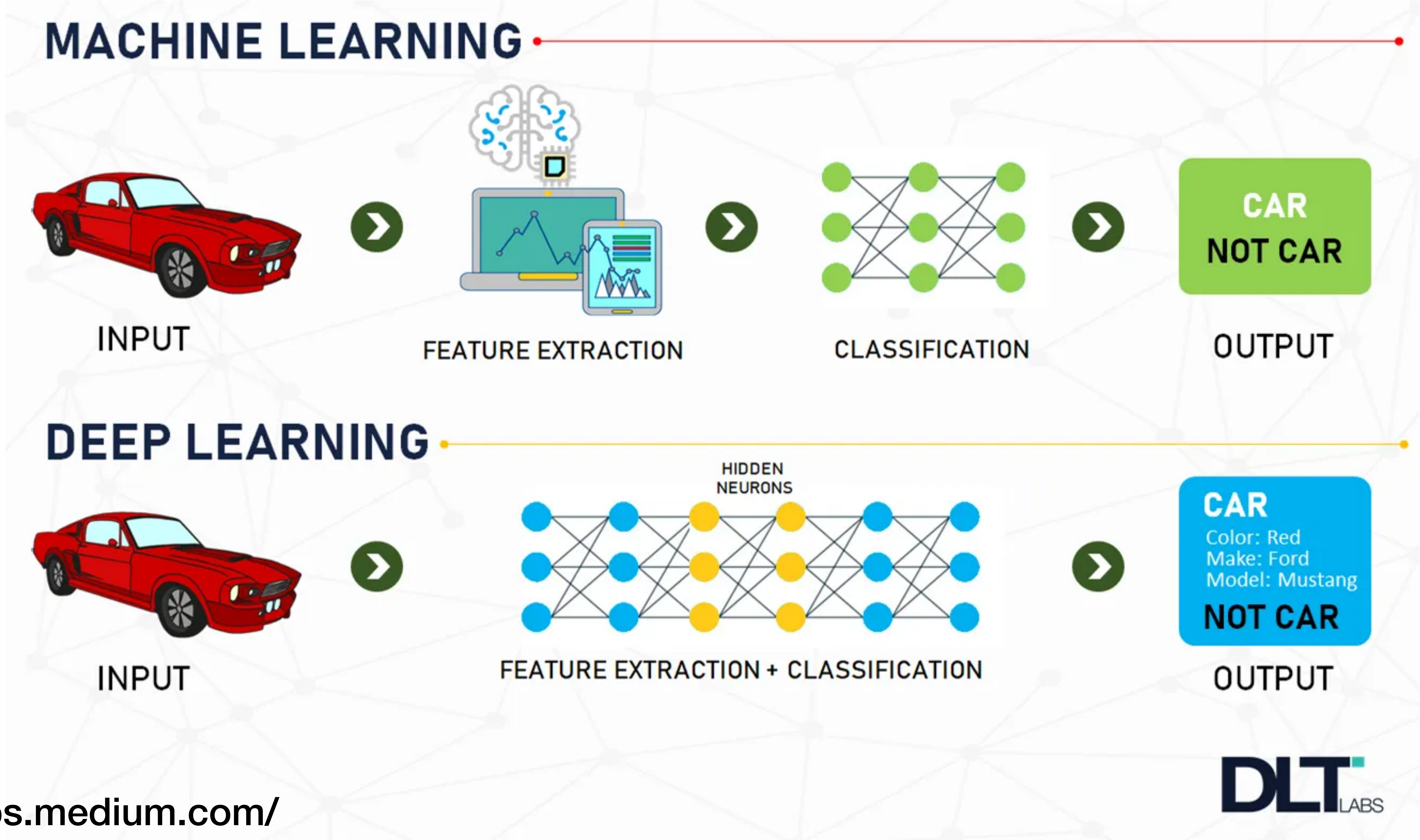


Figure largely inspired by Chollet (2017)

Machine Learning vs Deep Learning



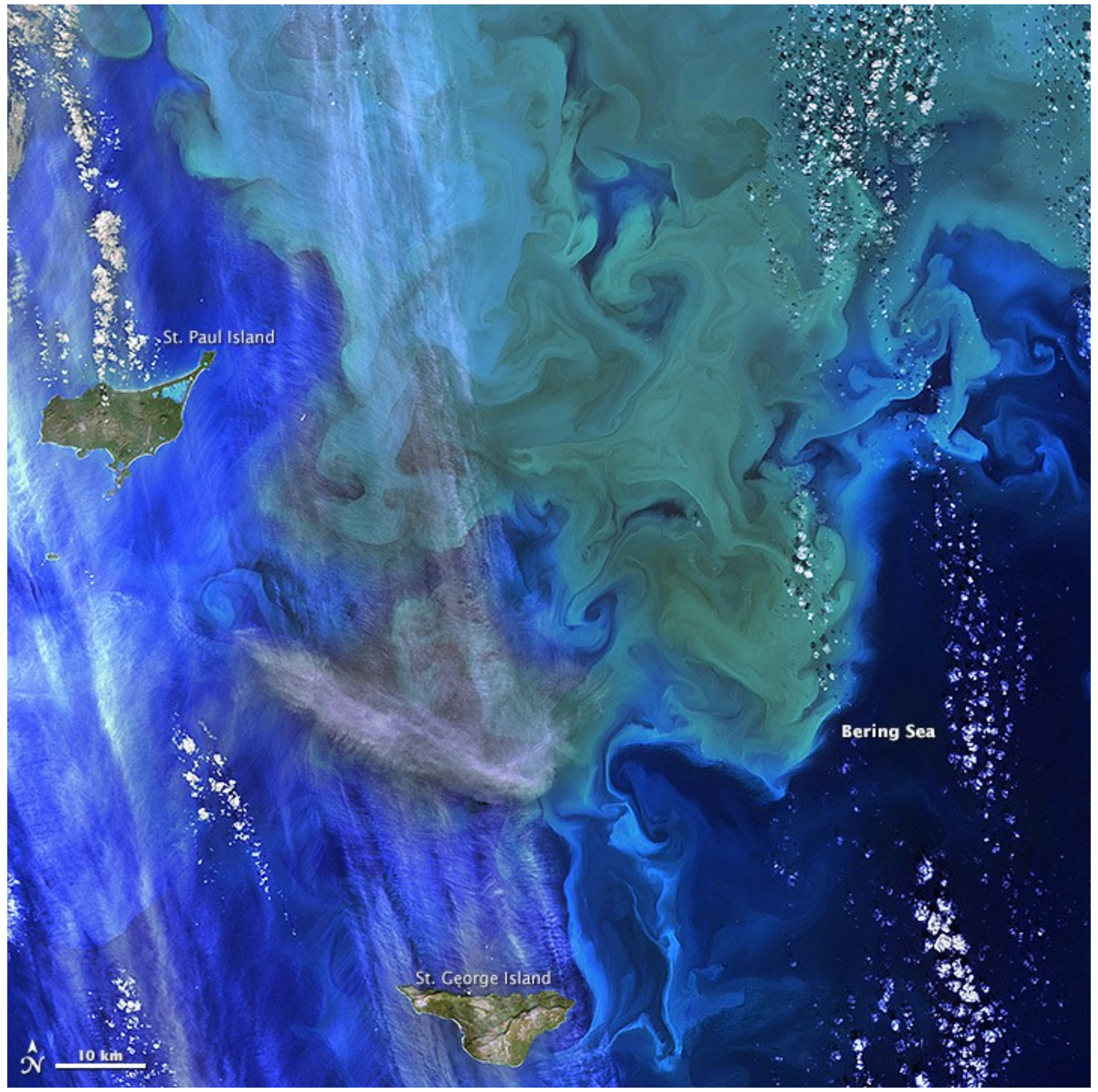
https://dltlabs.medium.com/

Emergence of deep learning

$$
\begin{aligned}
\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} + w\frac{\partial u}{\partial z} &= fv - \frac{1}{\rho}\frac{\partial p}{\partial x} + K_u\frac{\partial^2 u}{\partial z^2} \\
\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} + w\frac{\partial v}{\partial z} &= -fu - \frac{1}{\rho}\frac{\partial p}{\partial y} + K_v\frac{\partial^2 v}{\partial z^2} \\
-\frac{\partial p}{\partial z} &= \rho g
\end{aligned}
$$

Can you solve these equations
to picture the ocean circulation?

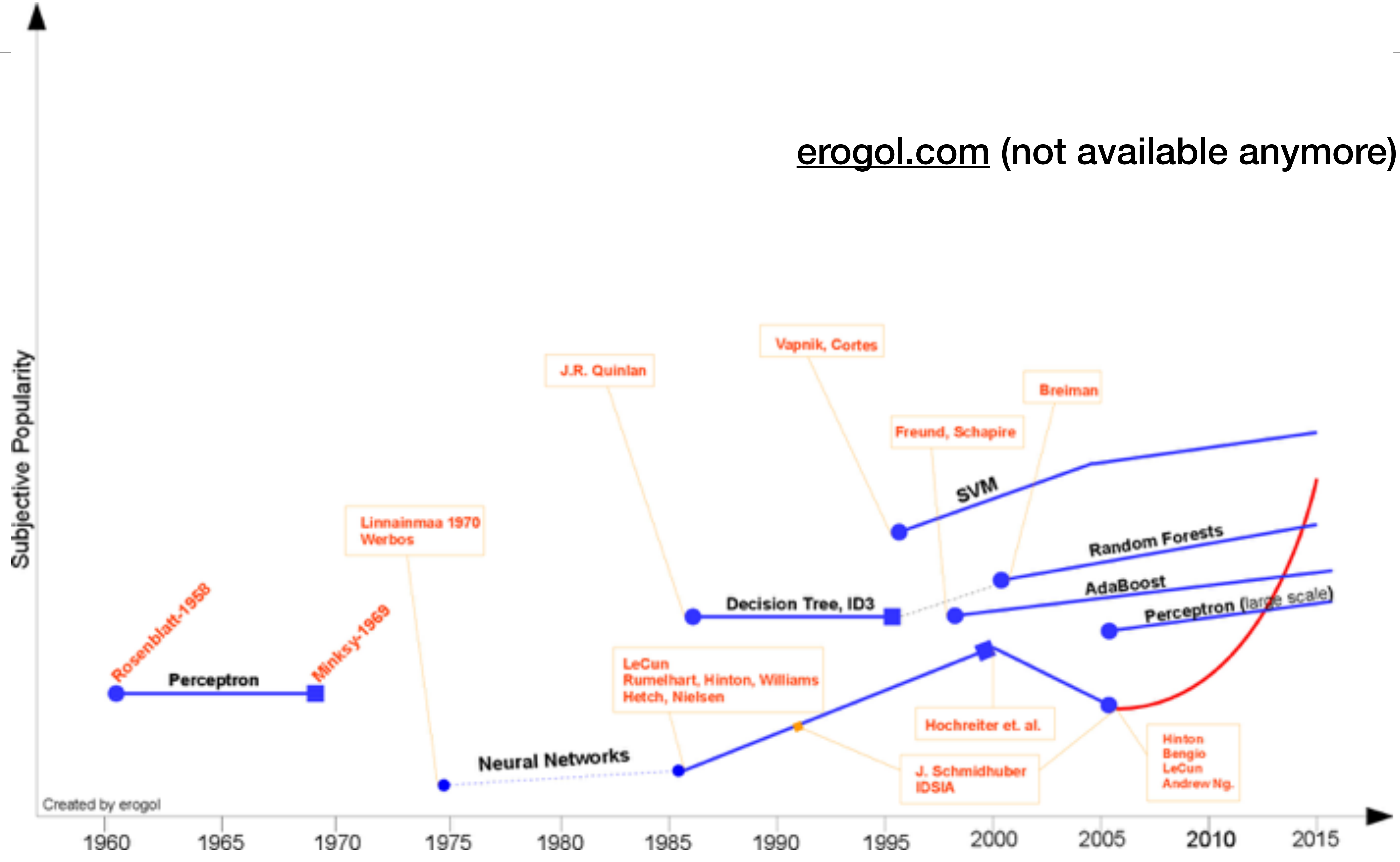
What is on this picture?



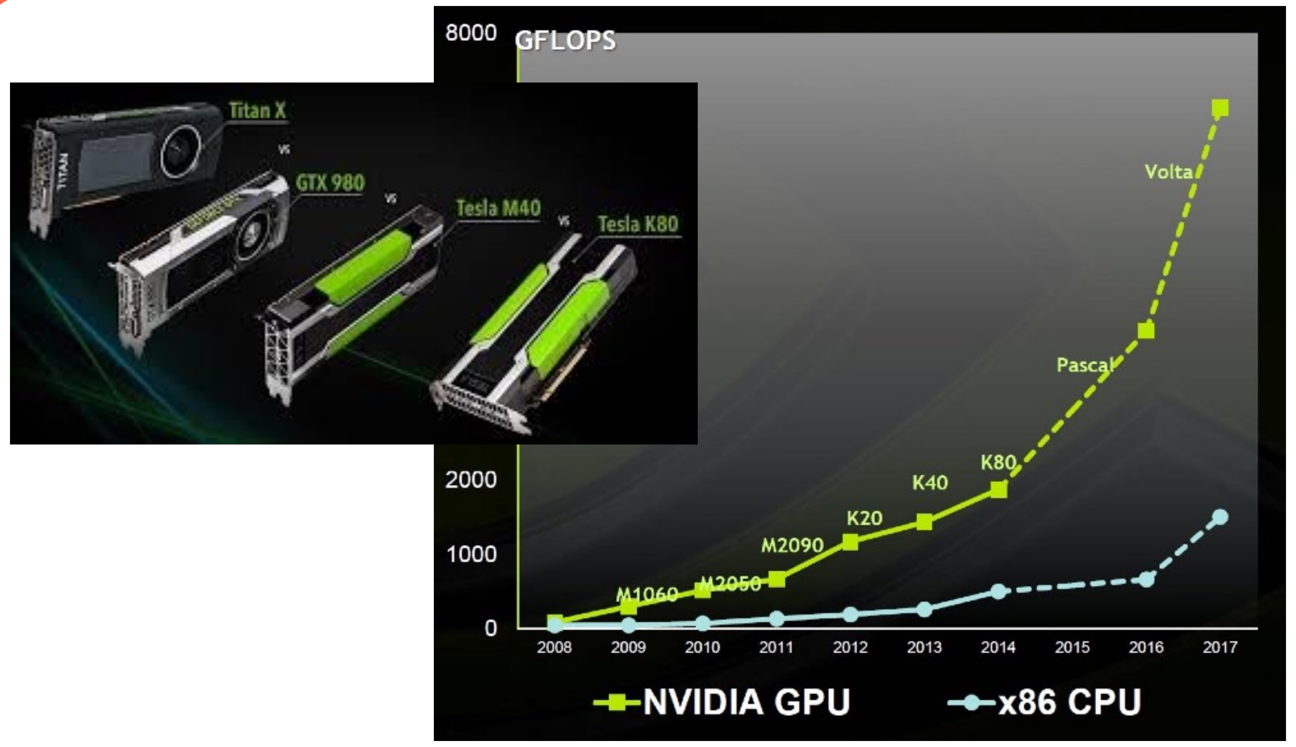
This is a satellite view of phytoplankton at the ocean surface. Any idea of how surface currents are?

Emergence of deep learning

erogol.com (not available anymore)



Emergence of deep learning



High-performance computing (GPU)

Large annotated dataset (> 1M)

+

Keras

TensorFlow

Caffe

PYTORCH

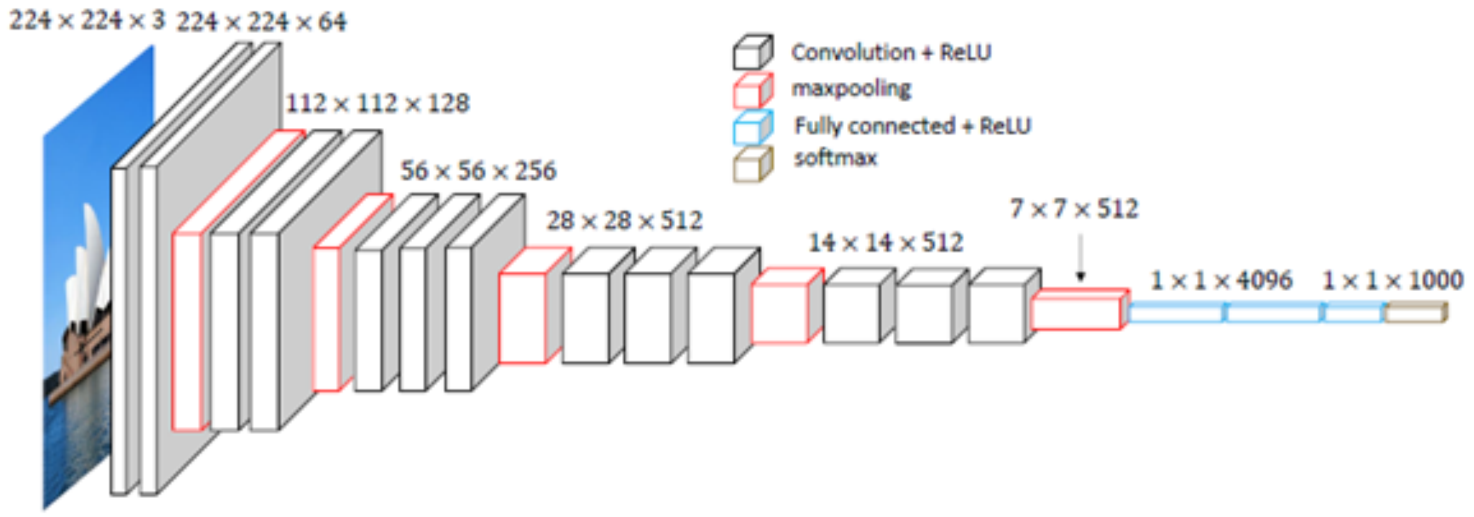


Figure 2: The architecture of VGG16 model .

An history of DL

Neural Networks 61 (2015) 85–117

Contents lists available at ScienceDirect

Neural Networks

journal homepage: www.elsevier.com/locate/neunet

ELSEVIER

Review

Deep learning in neural networks: An overview

Jürgen Schmidhuber

The Swiss AI Lab IDSIA, Istituto Dalle Molle di Studi sull'Intelligenza Artificiale, University of Lugano & SUPSI, Galleria 2, 6928 Manno-Lugano, Switzerland

ARTICLE INFO

Article history:
Received 2 May 2014
Received in revised form 12 September 2014
Accepted 14 September 2014
Available online 13 October 2014

Keywords:
Deep learning
Supervised learning
Unsupervised learning
Reinforcement learning
Evolutionary computation

ABSTRACT

In recent years, deep artificial neural networks (including recurrent ones) have won numerous contests in pattern recognition and machine learning. This historical survey compactly summarizes relevant work, much of it from the previous millennium. Shallow and Deep Learners are distinguished by the depth of their *credit assignment paths*, which are chains of possibly learnable, causal links between actions and effects. I review deep supervised learning (also recapitulating the history of backpropagation), unsupervised learning, reinforcement learning & evolutionary computation, and indirect search for short programs encoding deep and large networks.

Today’s research in AI in France

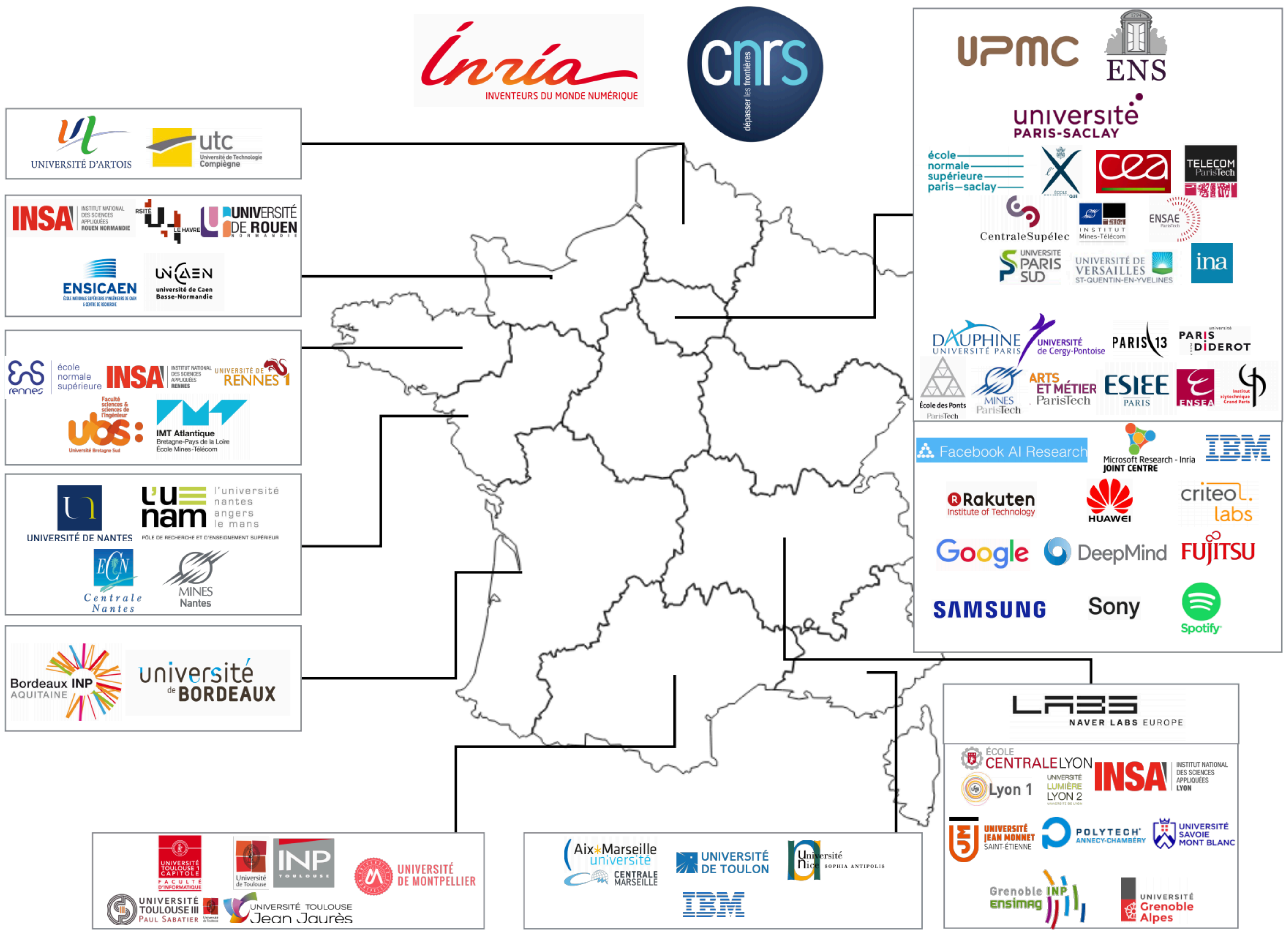


Image captioning (image to text)

A woman is throwing a **frisbee** in a park.

A **dog** is standing on a hardwood floor.

A little **girl** sitting on a bed with a teddy bear.

A group of **people** sitting on a boat in the water.

Vinyals et al., ICML 2014

… and generation (text to image)

Open AI™

Imagine if the best AI models were open and free.

What would you do with them?

Open AI is working to make the world's best AI models open.

In the meantime…

Generate any image imaginable.

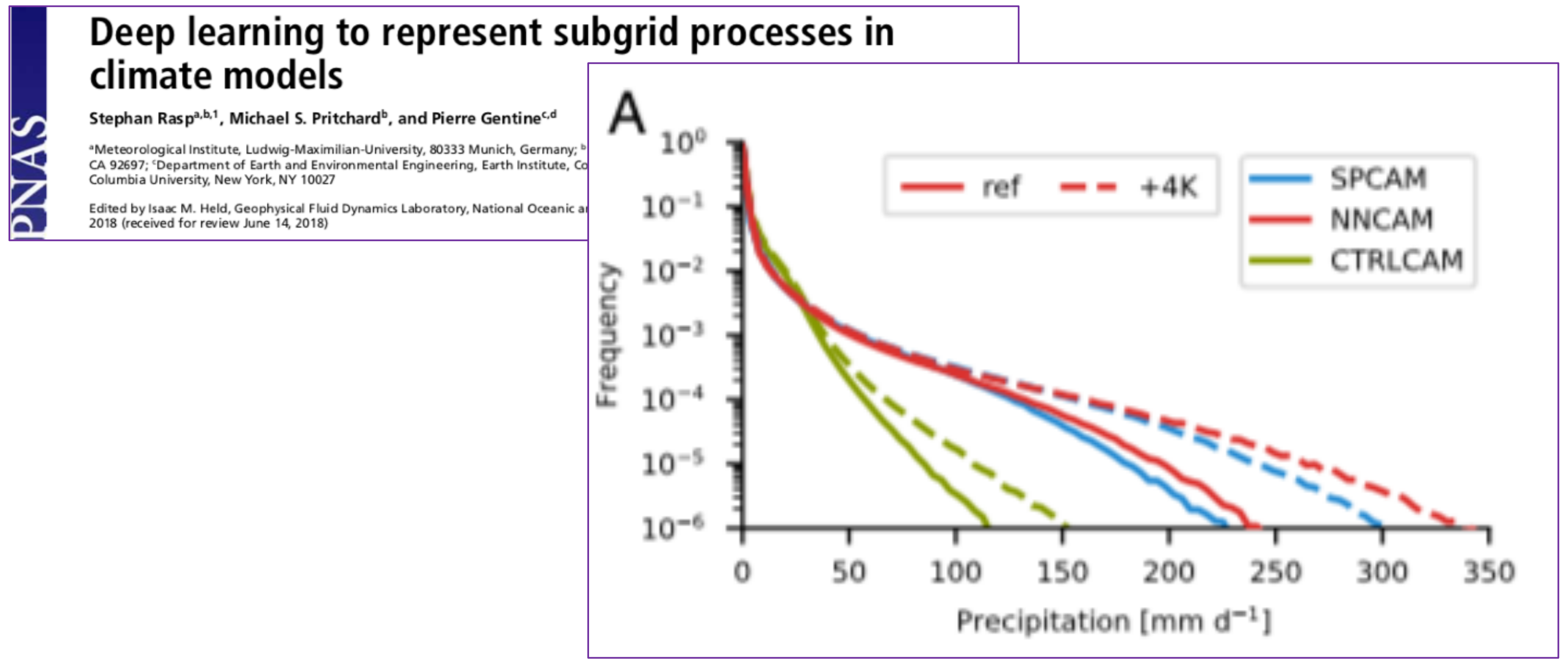
Generate

justin trudeau style, highly detailed, digital painting, artstation, concept art, wallpaper, smooth, sharp focus, illustration, art by artgerm and greg rutkowski and alphonse mucha trending on

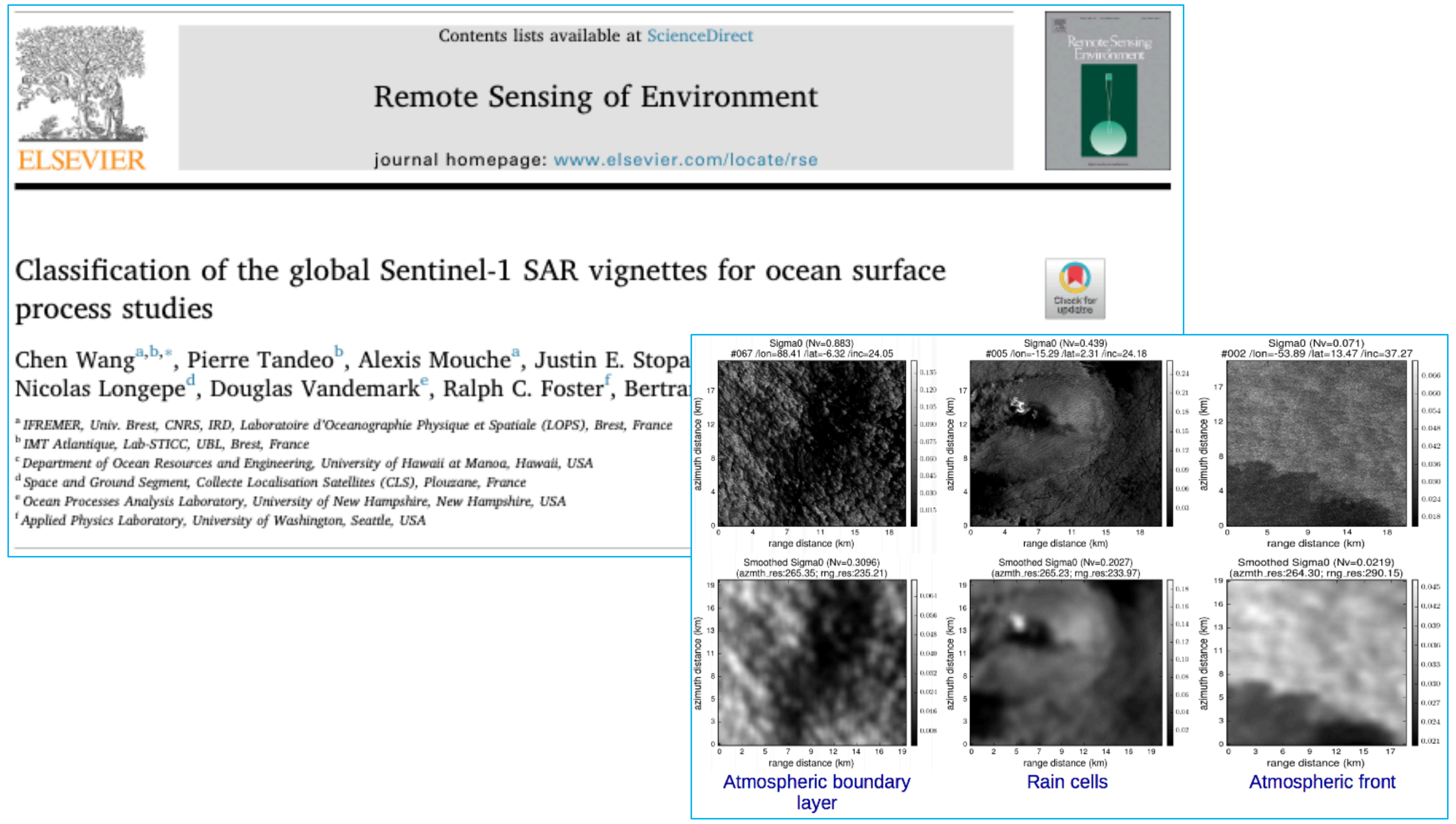
Interactive discussions… (ChatGPT)

- for summaries
- brainstorming
- get a quick answer to a question
- exploring
- wasting your time…

In OA sciences


Deep learning to represent subgrid processes in climate models

Stephan Rasp, Michael S. Pritchard, and Pierre Gentine

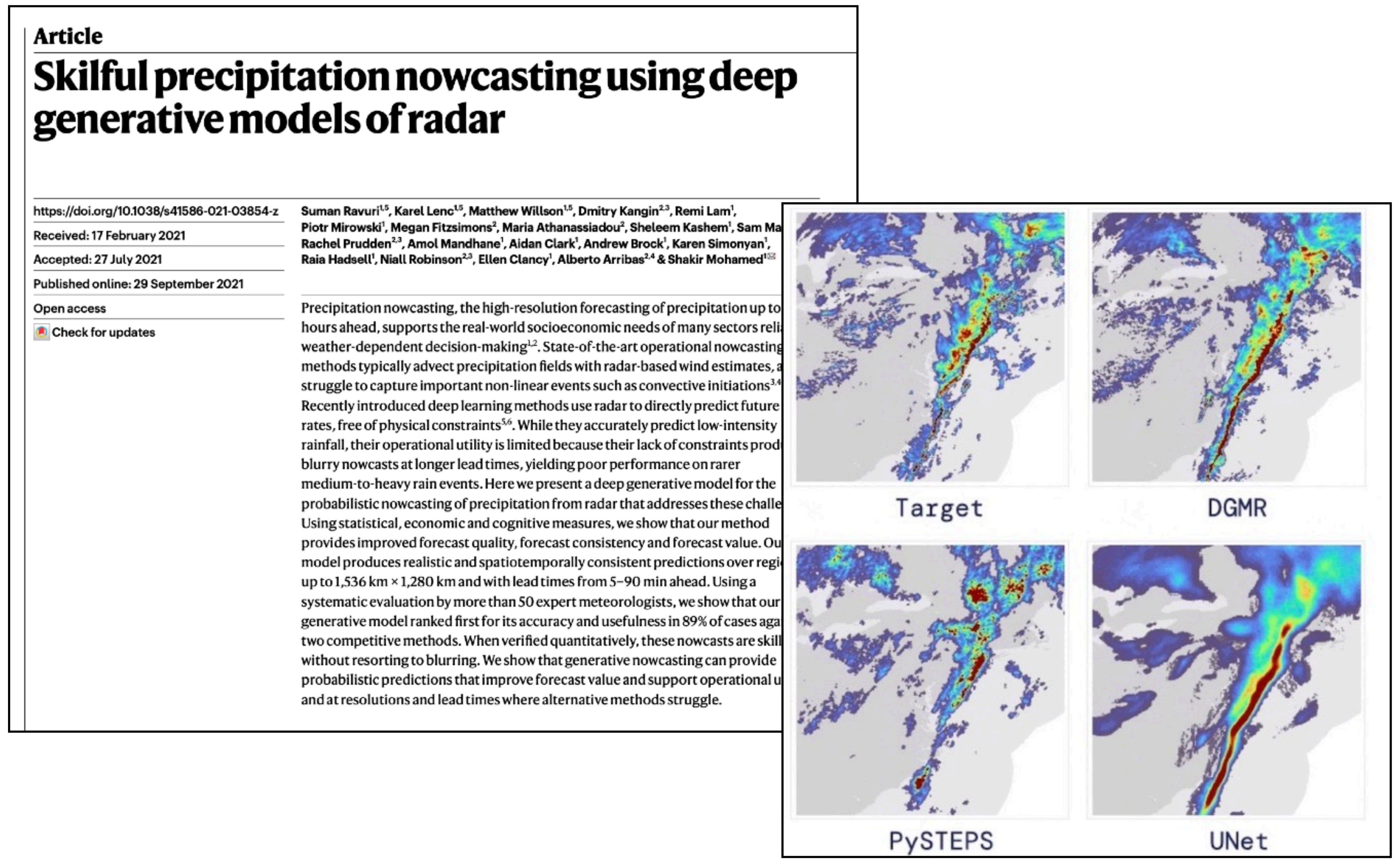

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Classification of the global Sentinel-1 SAR vignettes for ocean surface process studies

Chen Wang, Pierre Tandeo, Alexis Mouche, Justin E. Stopa, Nicolas Longepe, Douglas Vandemark, Ralph C. Foster, Bertra


Article

Skilful precipitation nowcasting using deep generative models of radar

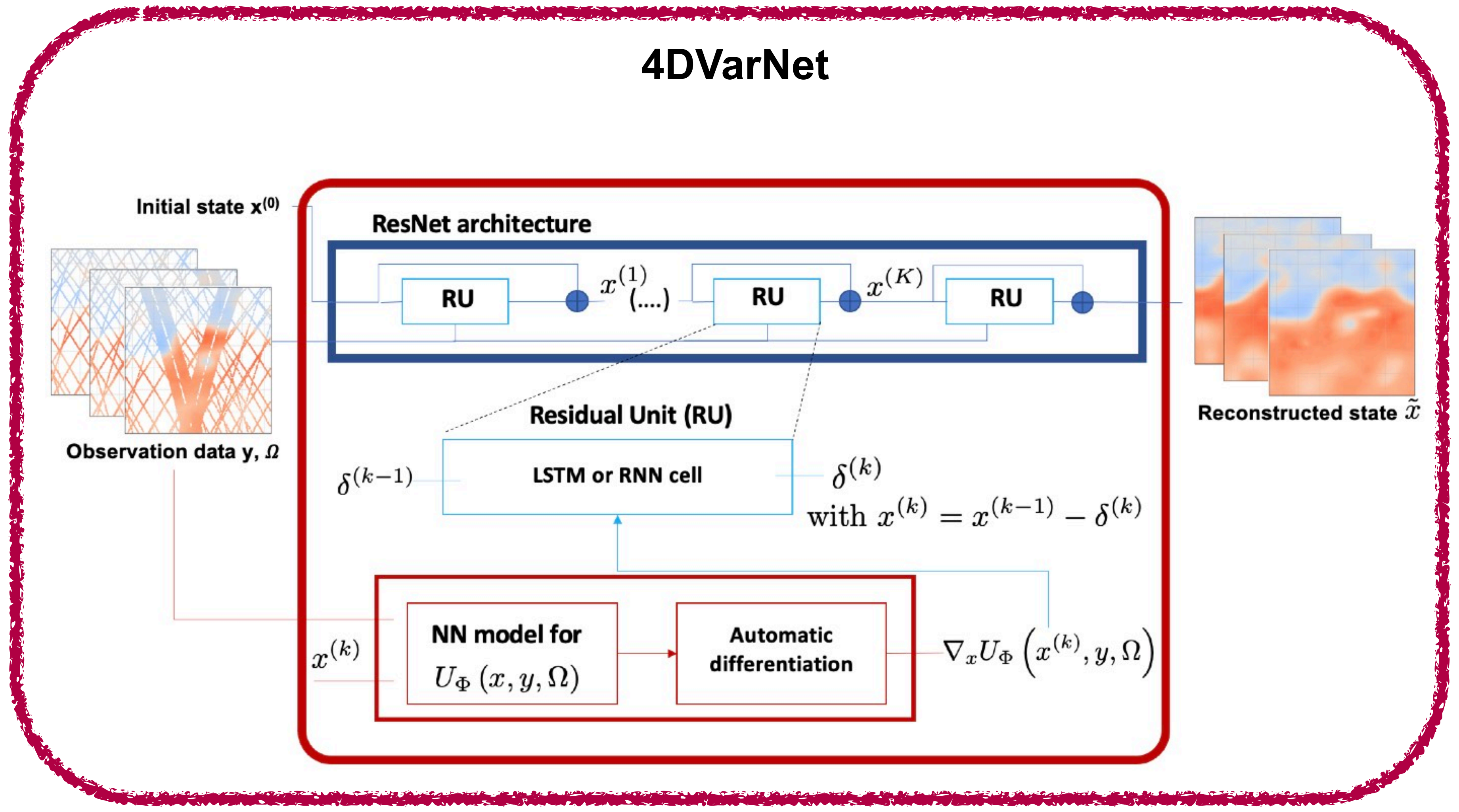
https://doi.org/10.1038/s41586-021-03854-z

Received: 17 February 2021

Accepted: 27 July 2021

Published online: 29 September 2021

Open access



Classical datasets

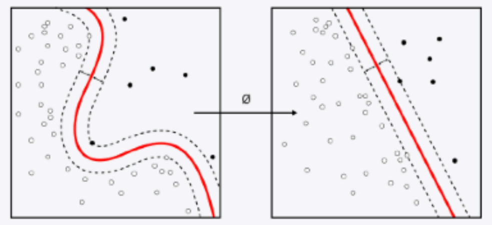
List of datasets for machine-learning research

From Wikipedia, the free encyclopedia

These datasets are applied for machine learning research and have been cited in peer-reviewed academic journals. Datasets are an integral part of the field of machine learning. Major advances in this field can result from advances in learning algorithms (such as deep learning), computer hardware, and, less-intuitively, the availability of high-quality training datasets.[1] High-quality labeled training datasets for supervised and semi-supervised machine learning algorithms are usually difficult and expensive to produce because of the large amount of time needed to label the data. Although they do not need to be labeled, high-quality datasets for unsupervised learning can also be difficult and costly to produce.[2][3][4][5]

Contents [hide]

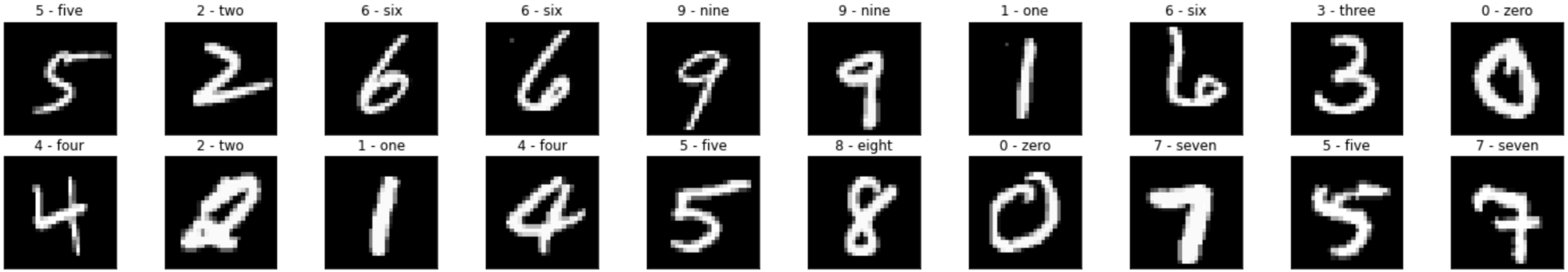
Part of a series on
Machine learning and data mining



- Problems [show]
- Supervised learning (classification • regression) [show]
- Clustering [show]
- Dimensionality reduction [show]
- Structured prediction [show]
- Anomaly detection [show]
- Artificial neural network [show]
- Reinforcement learning [show]
- Learning with humans [show]
- Model diagnostics [show]
- Theory [show]
- Machine-learning venues [show]
- Related articles [show]

V • T • E

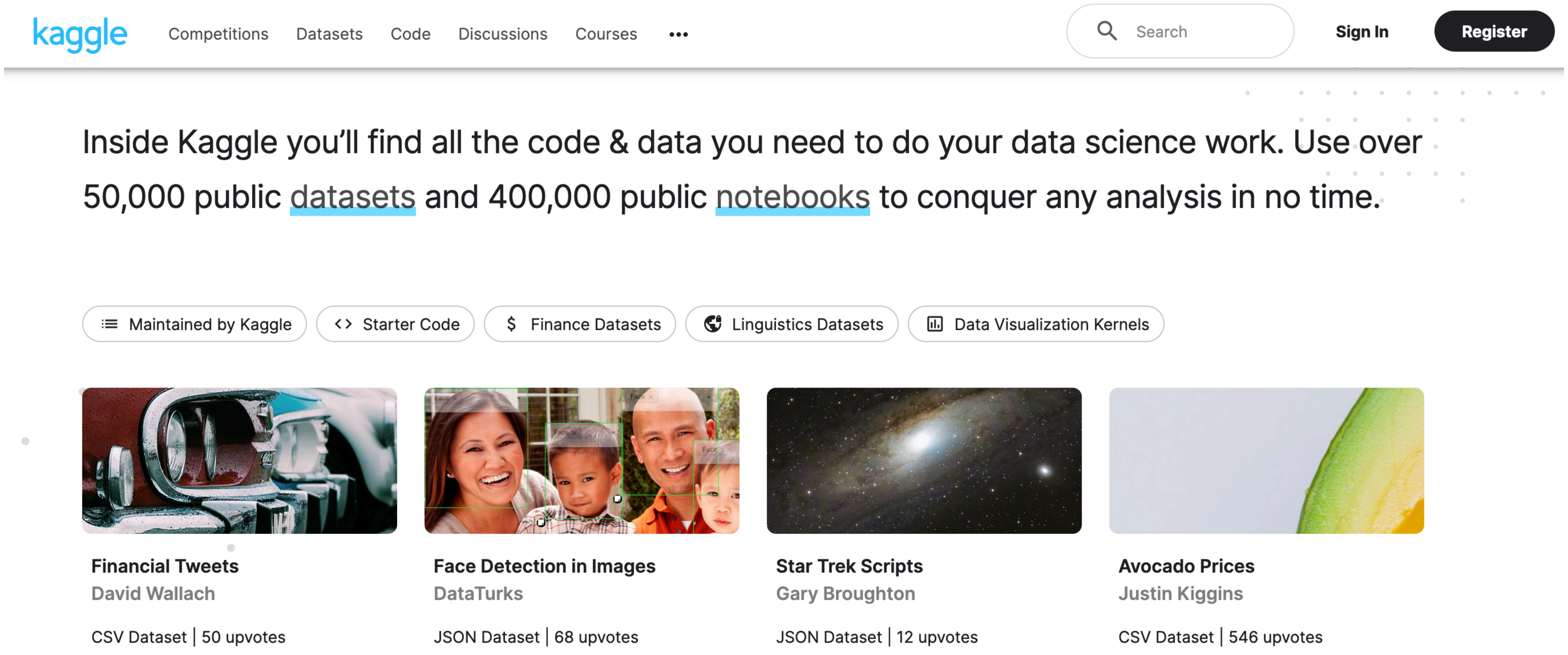
Classical datasets: MNIST



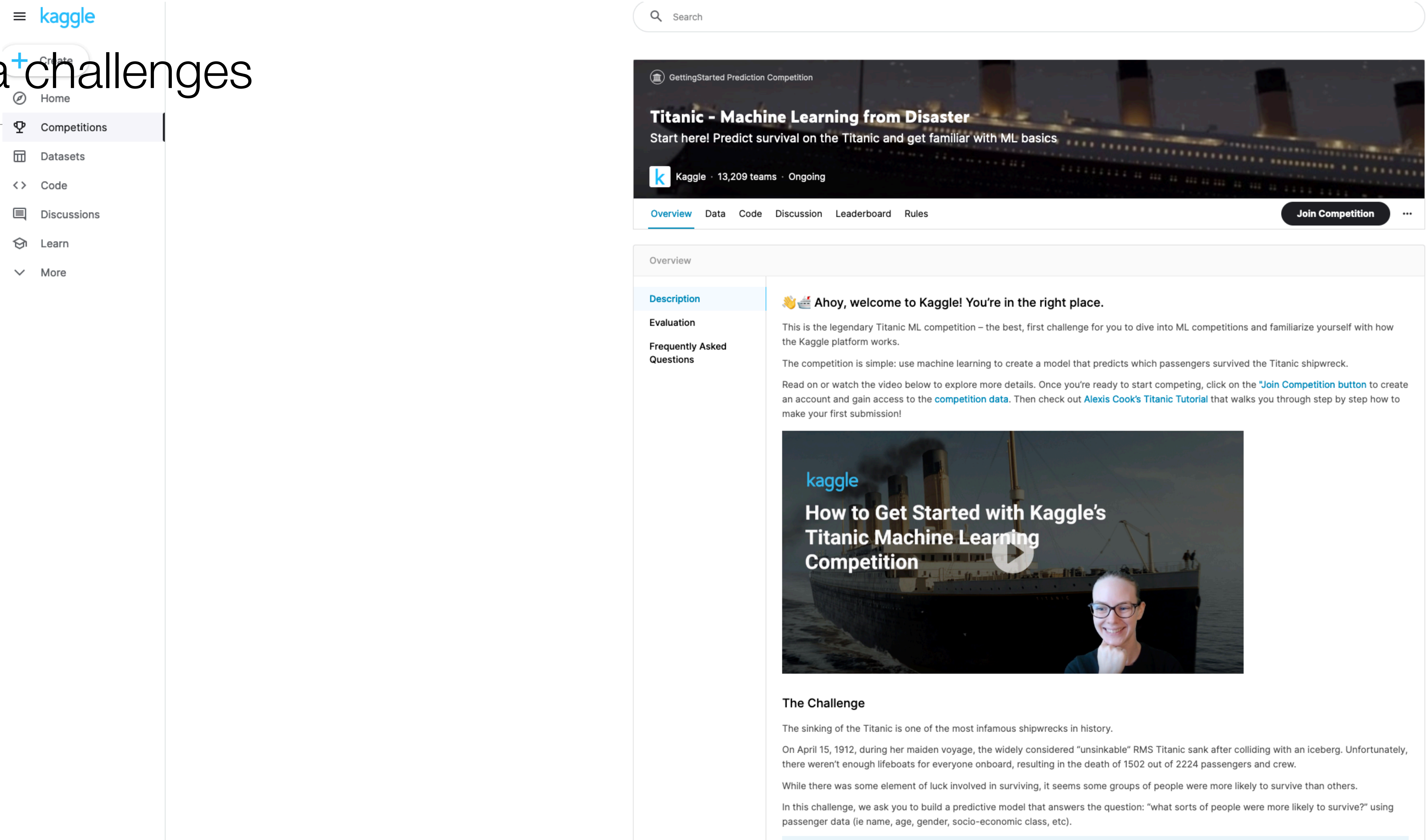
Softwares

- Python
- numpy
- matplotlib
- pandas
- tensorflow/keras
- PyTorch, pytorch-lightning
- Scikit-learn
- xarray
- and many others

Data challenges



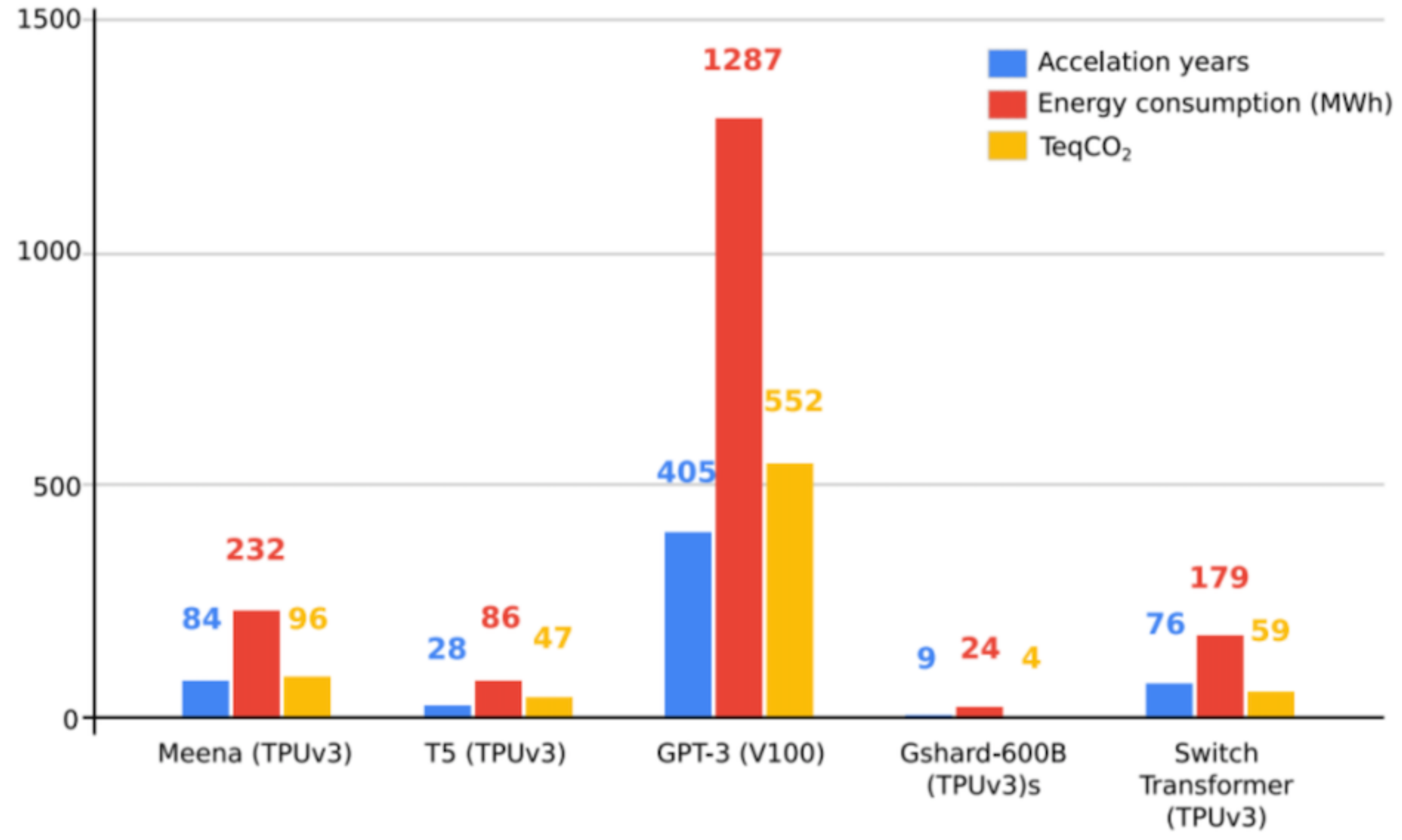
Data challenges



https://github.com/ocean-data-challenges

Data challenges

Energy consumption of AI



Quelques facteurs d'énergie utilisés pour l'entraînement de plusieurs réseaux de neurones modernes (ici pour le traitement du langage). Nombre d'années équivalent-GPU (bleu), facture électrique en MWh (pour information, 1MWh équivaut environ à 100 heures) en rouge, et **tonnage équivalent en CO_2 (pour information, 1TeqCO2 équivaut à 1 A/R Paris-New York), en jaune**. Le simple coût d'entraînement d'un réseau de neurones sur une application ciblée est 250 fois supérieur au maximum annuel autorisé à chaque Européen (2TeqCO2) pour atteindre l'équilibre carbone en 2050.

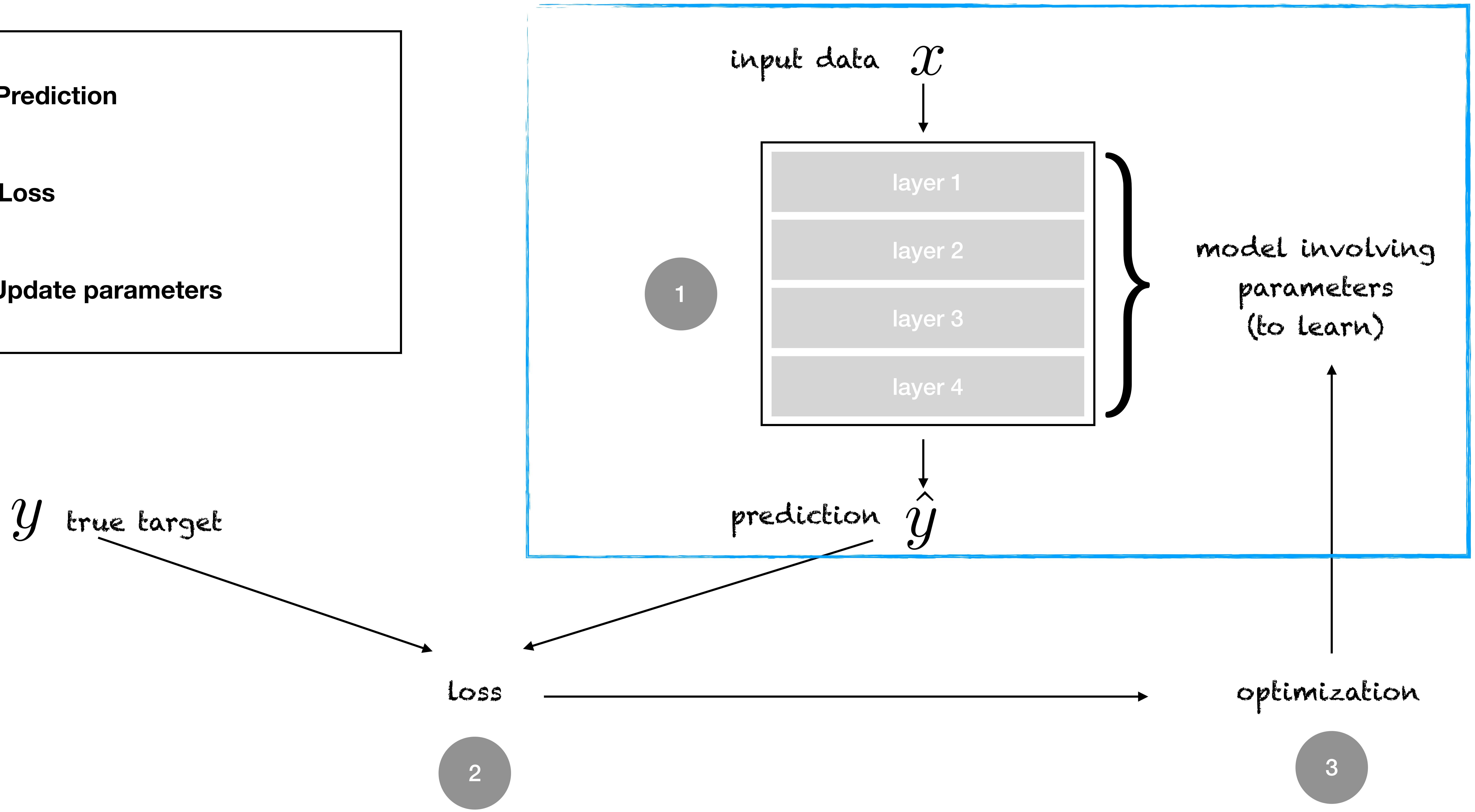
Figure from Patterson et al, 2021, https://arxiv.org/abs/2104.10350 ; Caption from Trystram et al, 2021 https://theconversation.com/apprentissage-profond-et-consommation-energetique-la-partie-immergee-de-lia-ceberg-172341

How learning works

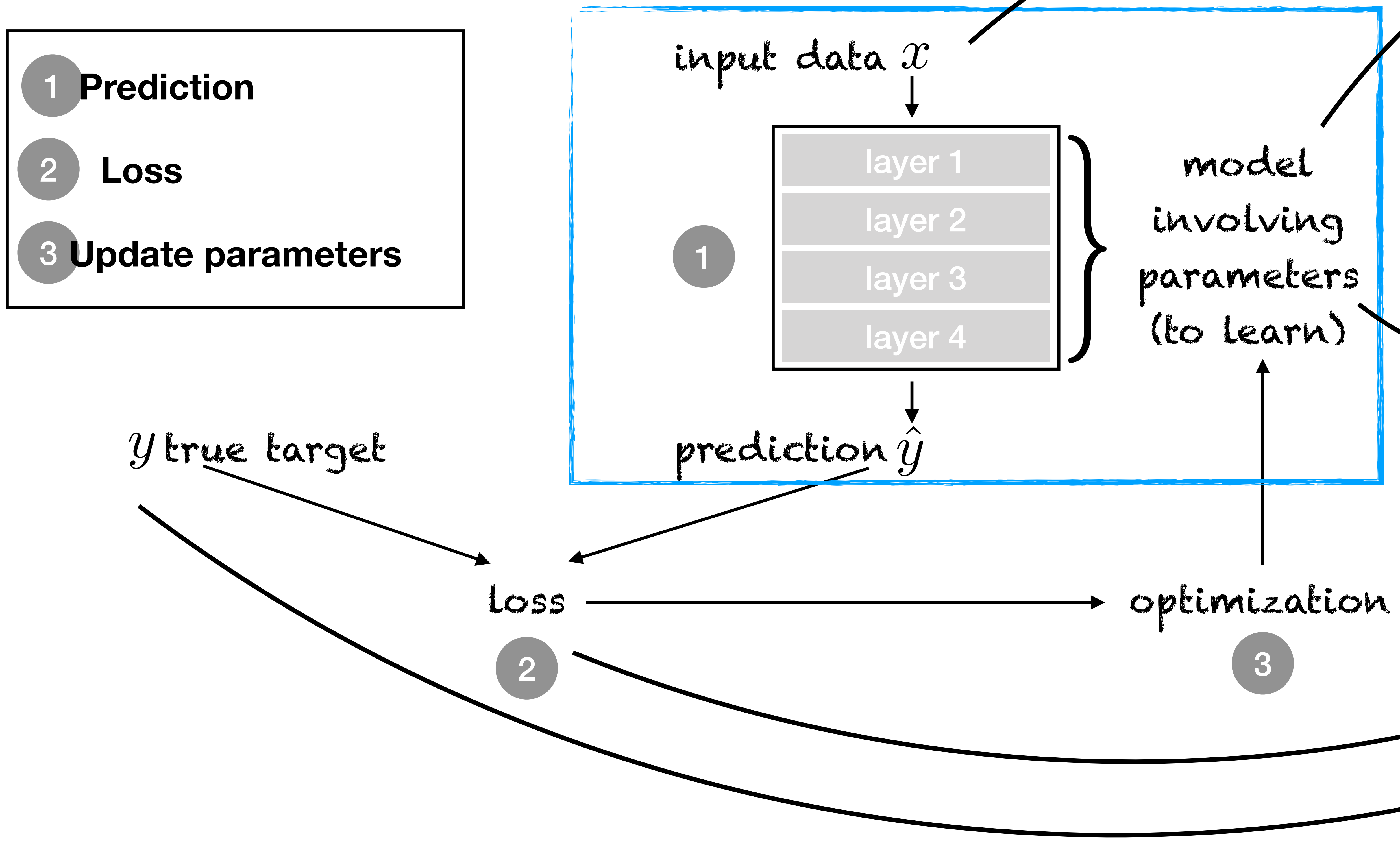
Description, mathematical formulation, basic concepts, vocabulary

Machine/Deep Learning

1. **Prediction**
2. **Loss**
3. **Update parameters**

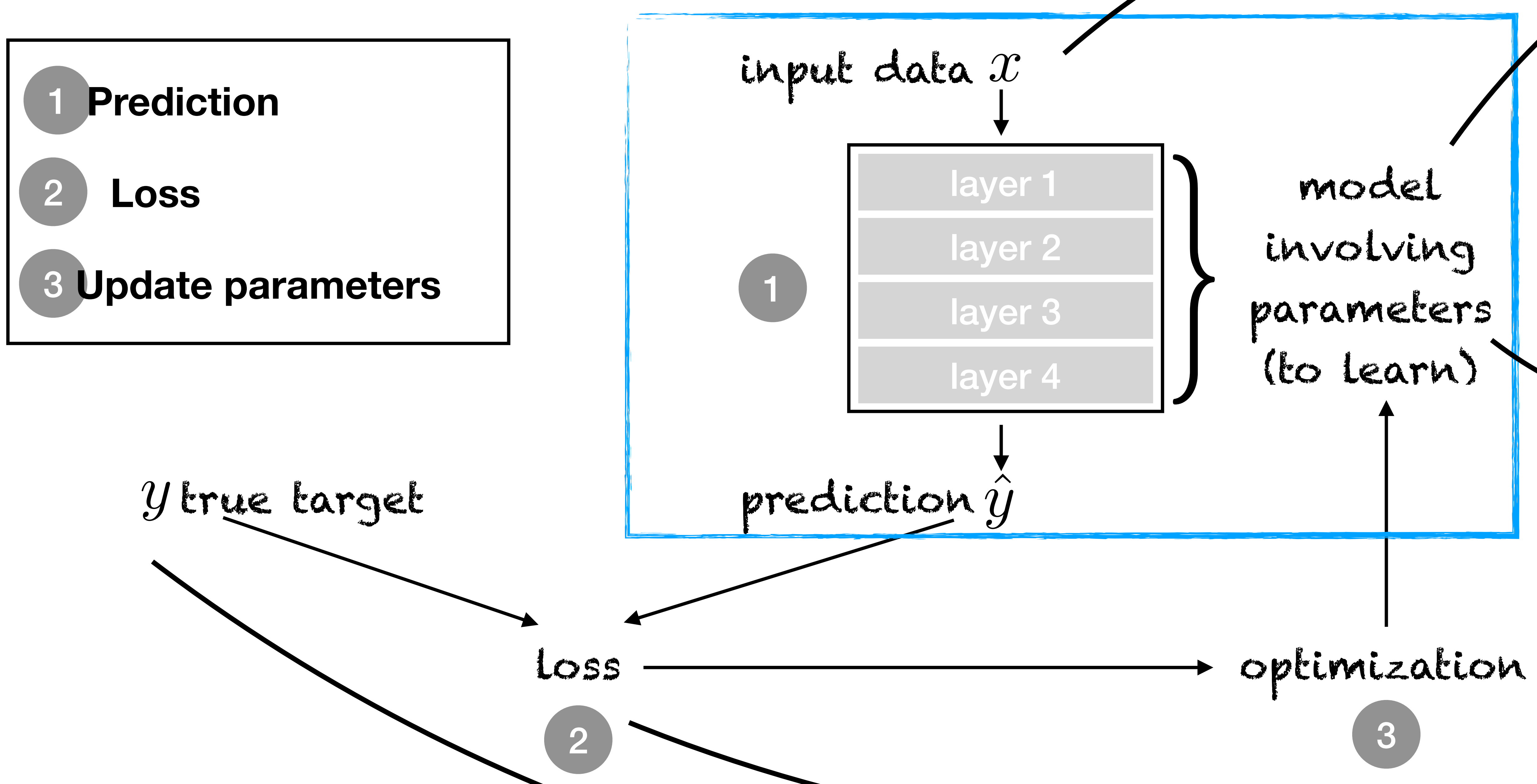


Machine/Deep Learning



$$\widehat{\theta} = \arg\min_{\theta} \mathcal{L}\left(\{x_i, y_i\}_{i \in \{1,..,N\}}; f_\theta\right)$$

Machine/Deep Learning



$$\widehat{\theta} = \arg\min_{\theta} \mathcal{L}\left(\{x_i, y_i\}_{i\in\{1,..,N\}}; f_\theta\right)$$

Key questions:

- how to design the model?
- How to choose the loss function?

Machine Learning terminology

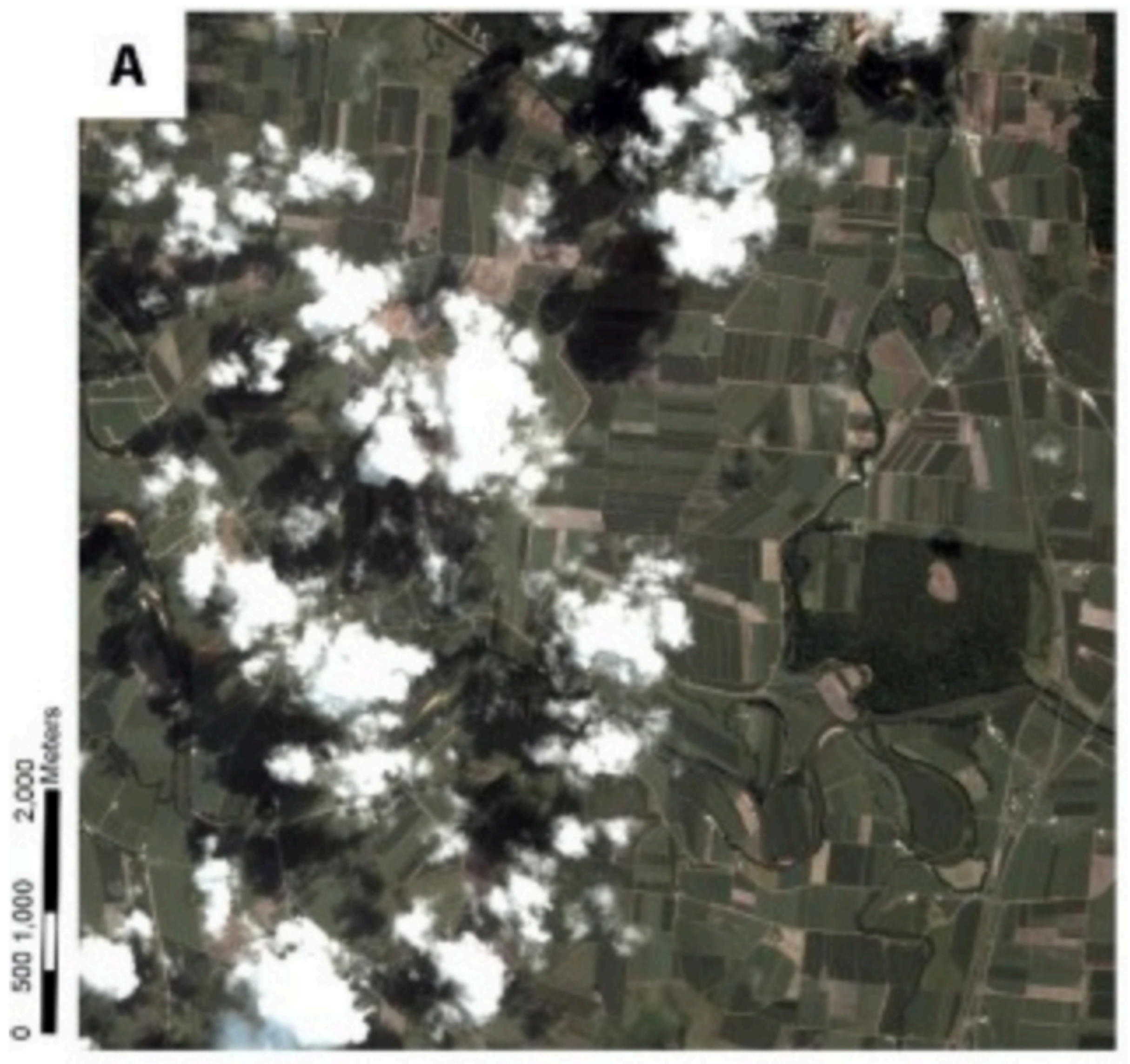
- Training and test dataset (Jeux de données d’apprentissage et test)
- Training loss (Coût d’apprentissage)
- Model (Modèle)
- Supervised learning (Apprentissage supervisé)
- Unsupervised learning (Apprentissage non-supervisé)
- Regression
- Classification

Types of Machine Learning problems

Binary classification

Classification answers to: what (pre-defined) class does this item belong to?

Supervised learning.



?

Cloudy

Not Cloudy

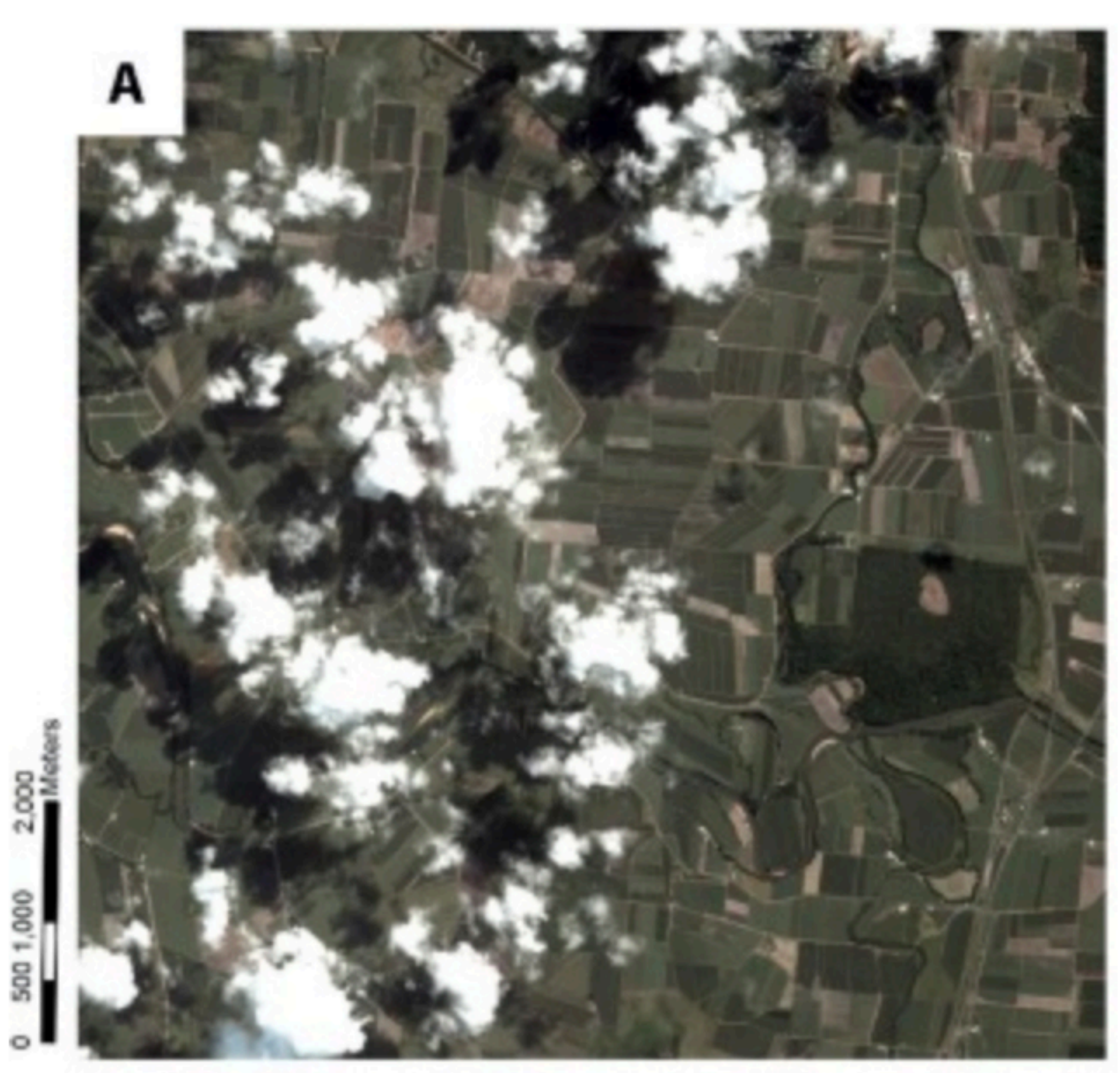
image from Shendryk et al (2019). https://www.sciencedirect.com/science/article/pii/S0924271619302023

Types of Machine Learning problems

Classification answers to: what (pre-defined) class does this item belong to?

Supervised learning.

Multi-class classification



?

Cloudy

Partially Cloudy

Not Cloudy

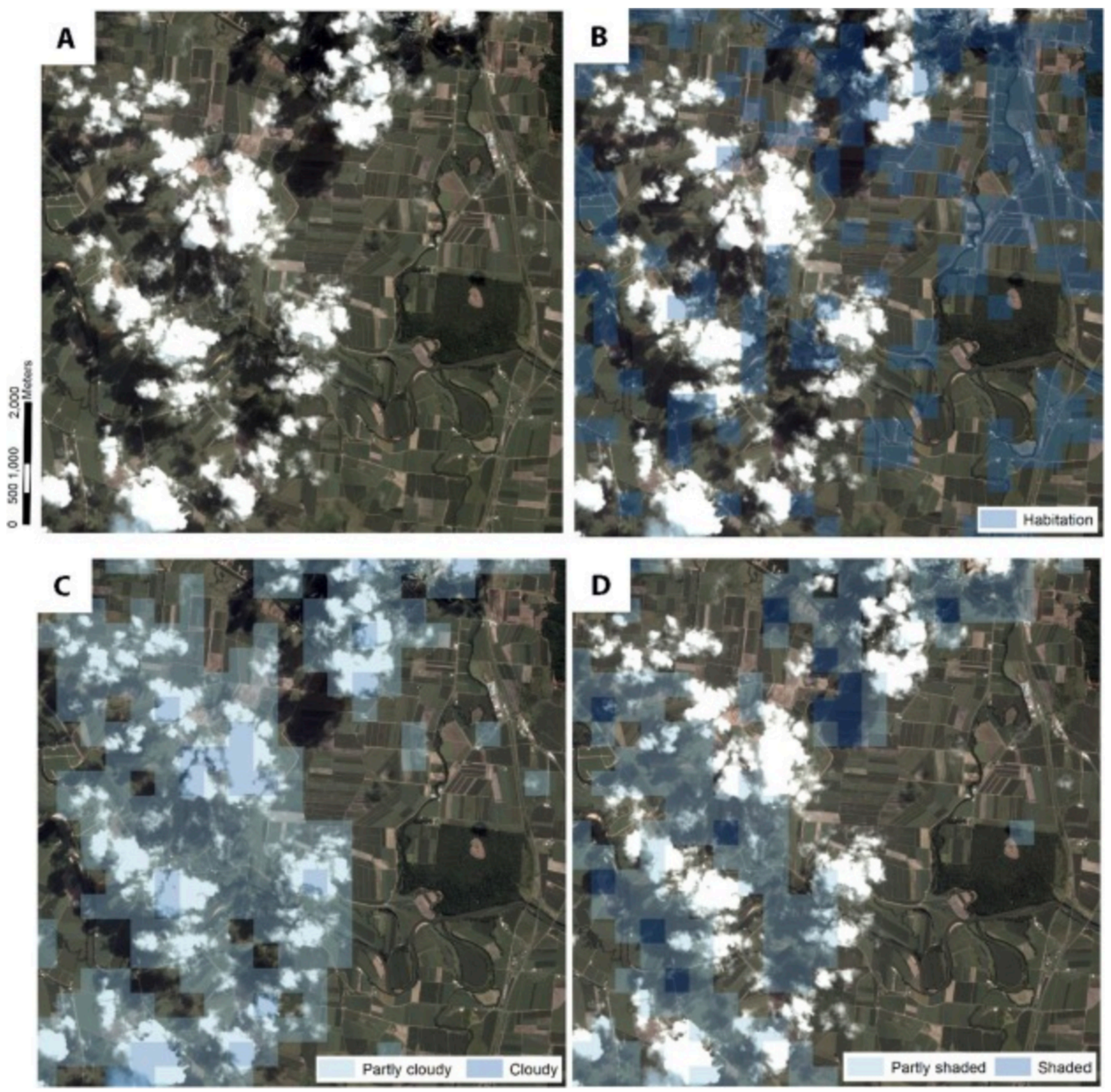
image from Shendryk et al (2019). https://www.sciencedirect.com/science/article/pii/S0924271619302023

Types of Machine Learning problems

Classification answers to: what (pre-defined) class does this item belong to?

Supervised learning.

Multi-label classification: "contain habitation", "partly cloudy", "cloudy", "partly shaded", "shaded"



Download : Download high-res image (1MB)   Download : Download full-size image

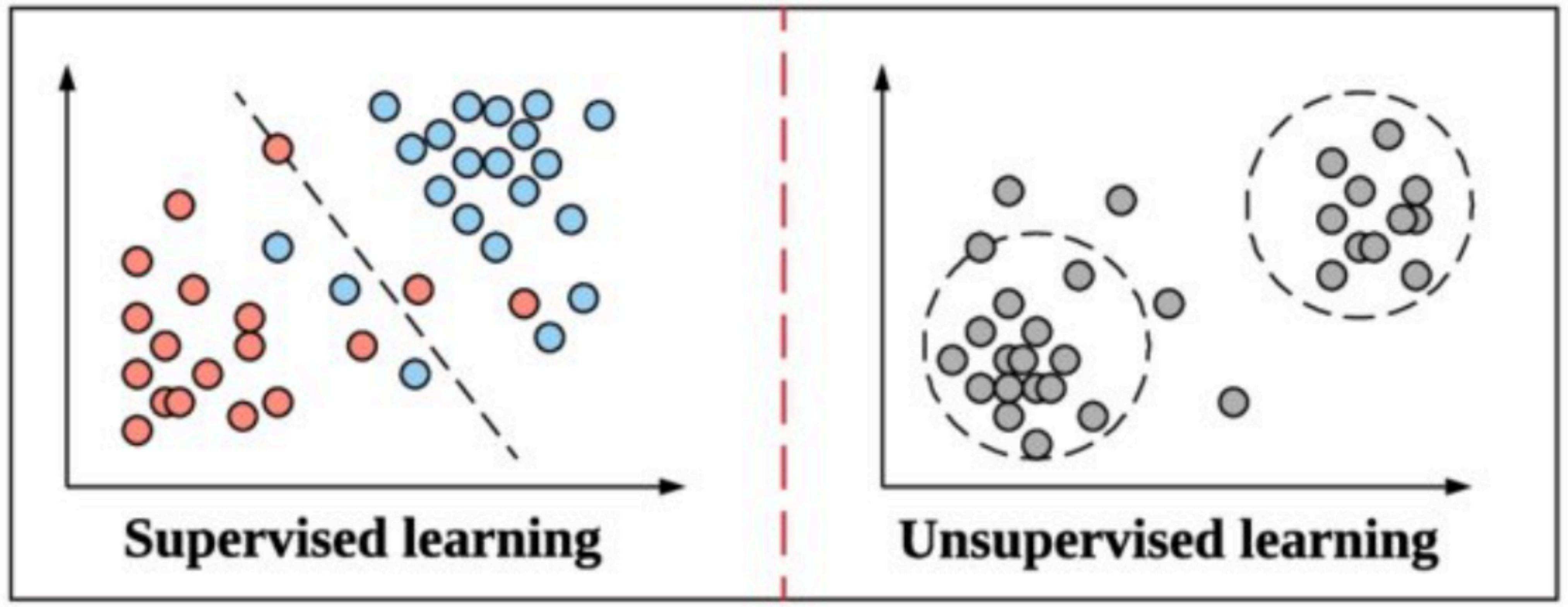
Fig. 4. An example of a multi-label scene classification using our ensemble CNN model of a PlanetScope image collected over the Wet Tropics of Australia on May 9, 2017: (A) original RGB image, (B) scenes identified to contain *'habitation'* labels (as an example of a land cover label group), (C) scenes identified to contain *'partly cloudy'* and *'cloudy'* labels, (D) scenes identified to contain *'partly shaded'* and *'shaded'* labels. Each scene is 128 × 128 pixels (i.e. 400 × 400 m) in size.

Shendryk et al (2019). https://www.sciencedirect.com/science/article/pii/S0924271619302023

Types of Machine Learning problems

Clustering answers to: how can I group these items? No pre-defined classes.

Unsupervised learning.

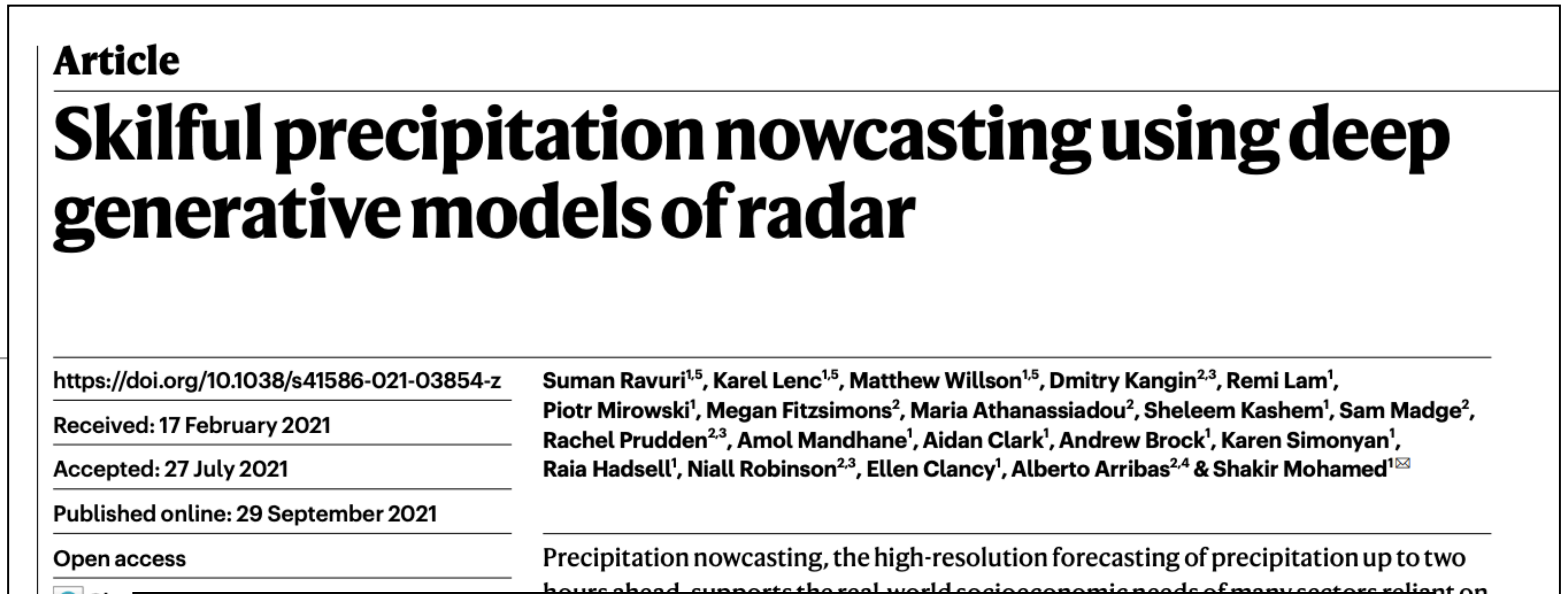


Qian et al, 2019

Types of Machine Learning problems

Regression refers to the case where the task is to provide a real-valued quantity.

Supervised learning.

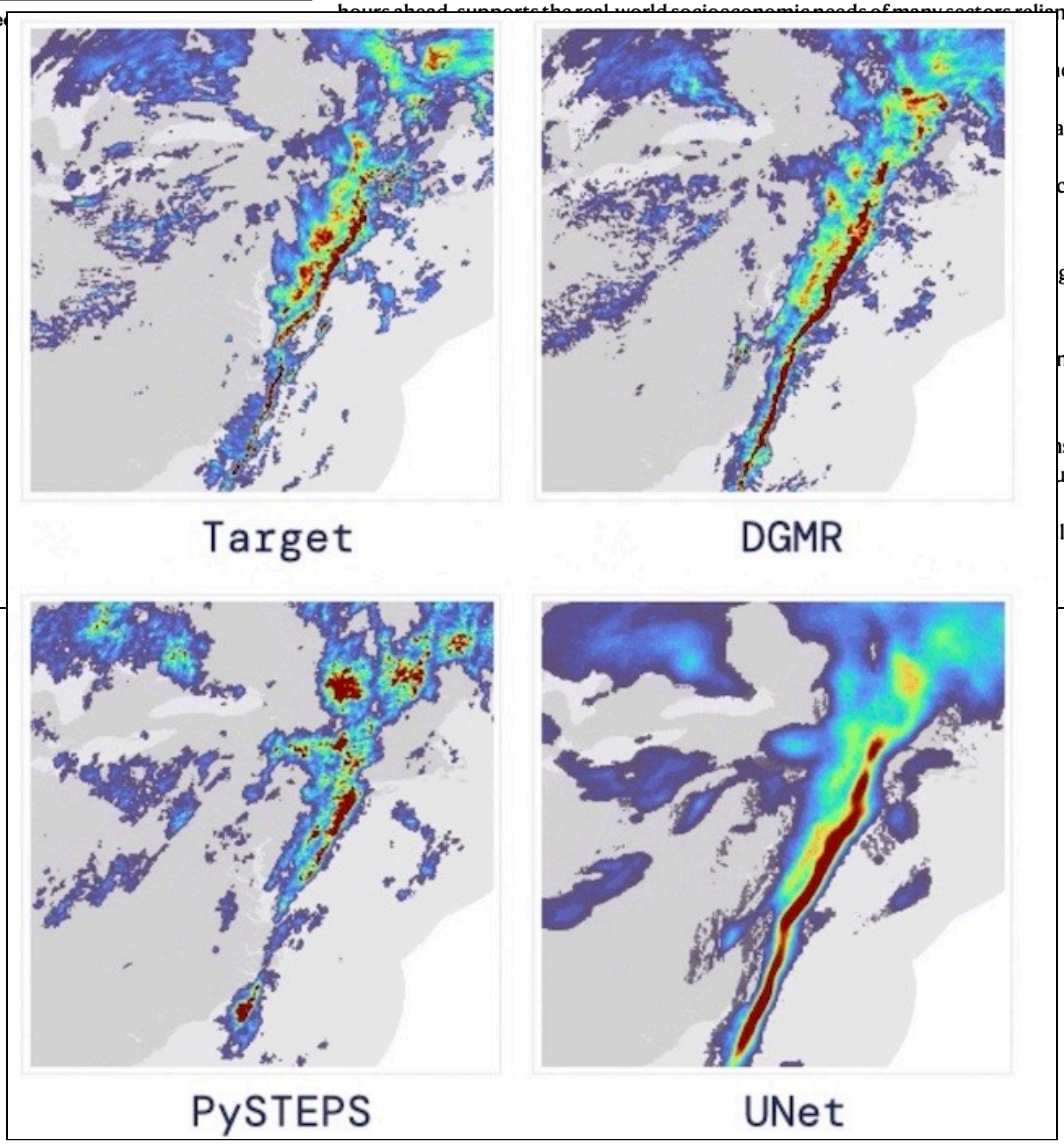

Article

Skilful precipitation nowcasting using deep generative models of radar

https://doi.org/10.1038/s41586-021-03854-z
Received: 17 February 2021
Accepted: 27 July 2021
Published online: 29 September 2021
Open access

Suman Ravuri[1,5], Karel Lenc[1,5], Matthew Willson[1,5], Dmitry Kangin[2,3], Remi Lam[1], Piotr Mirowski[1], Megan Fitzsimons[2], Maria Athanassiadou[2], Sheleem Kashem[1], Sam Madge[2], Rachel Prudden[2,3], Amol Mandhane[1], Aidan Clark[1], Andrew Brock[1], Karen Simonyan[1], Raia Hadsell[1], Niall Robinson[2,3], Ellen Clancy[1], Alberto Arribas[2,4] & Shakir Mohamed[1]✉

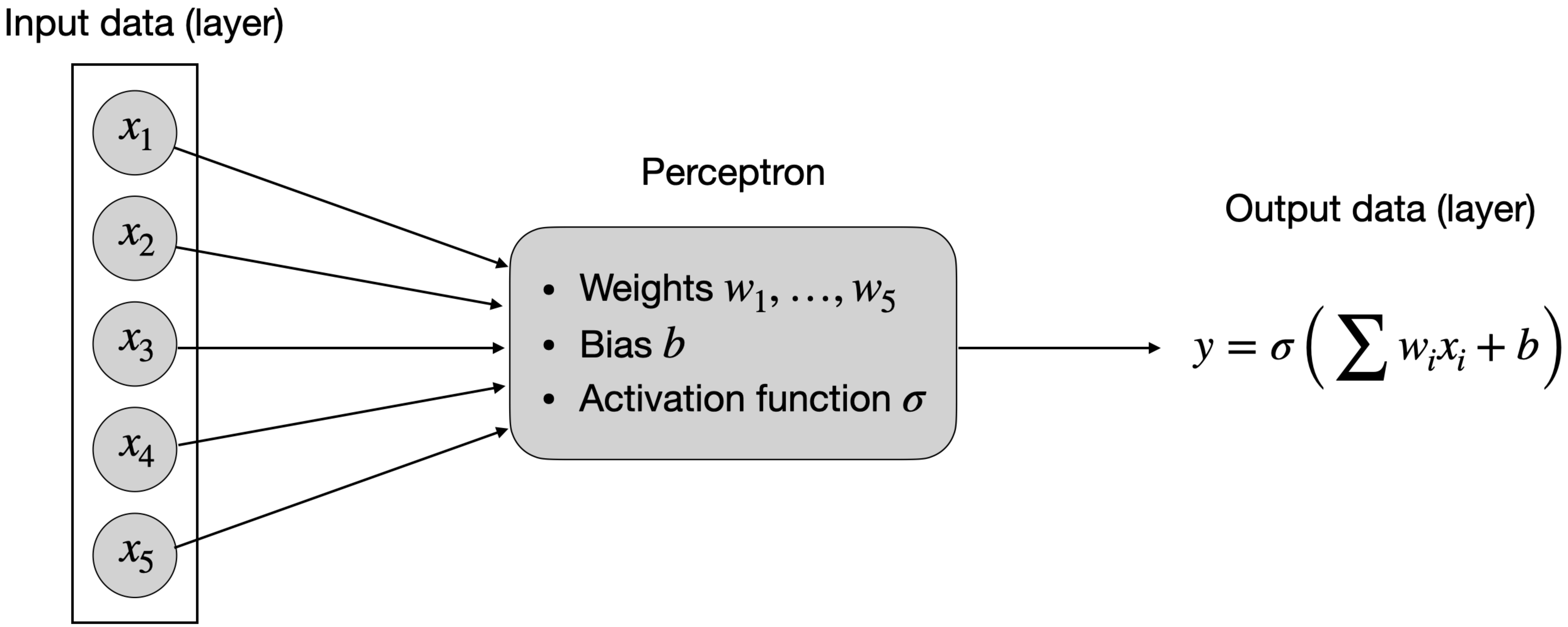
Precipitation nowcasting, the high-resolution forecasting of precipitation up to two



The Multi-Layer Perceptron

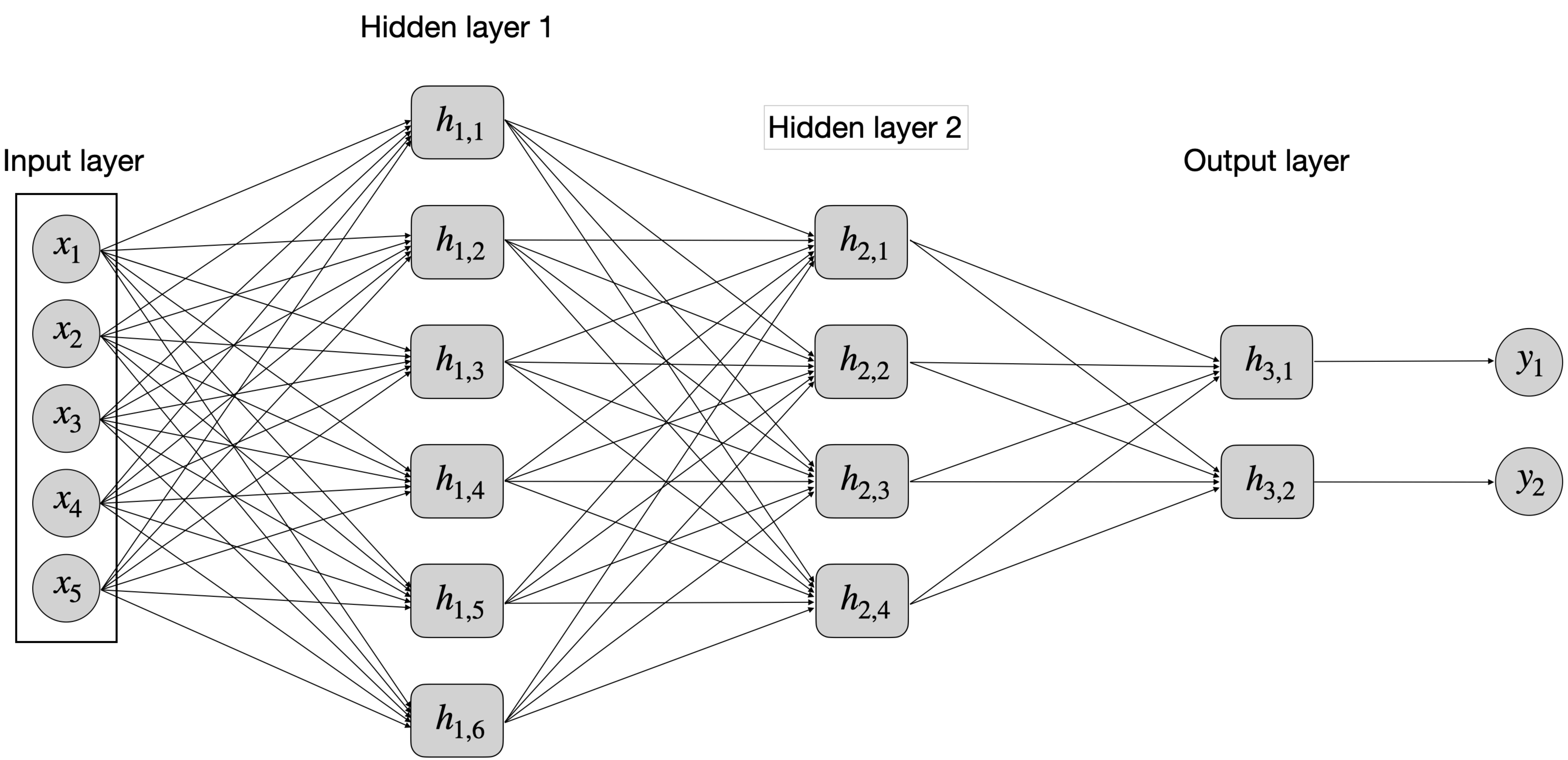
Perceptron, activation functions, feedforward networks, layers…

The artificial neuron (perceptron)

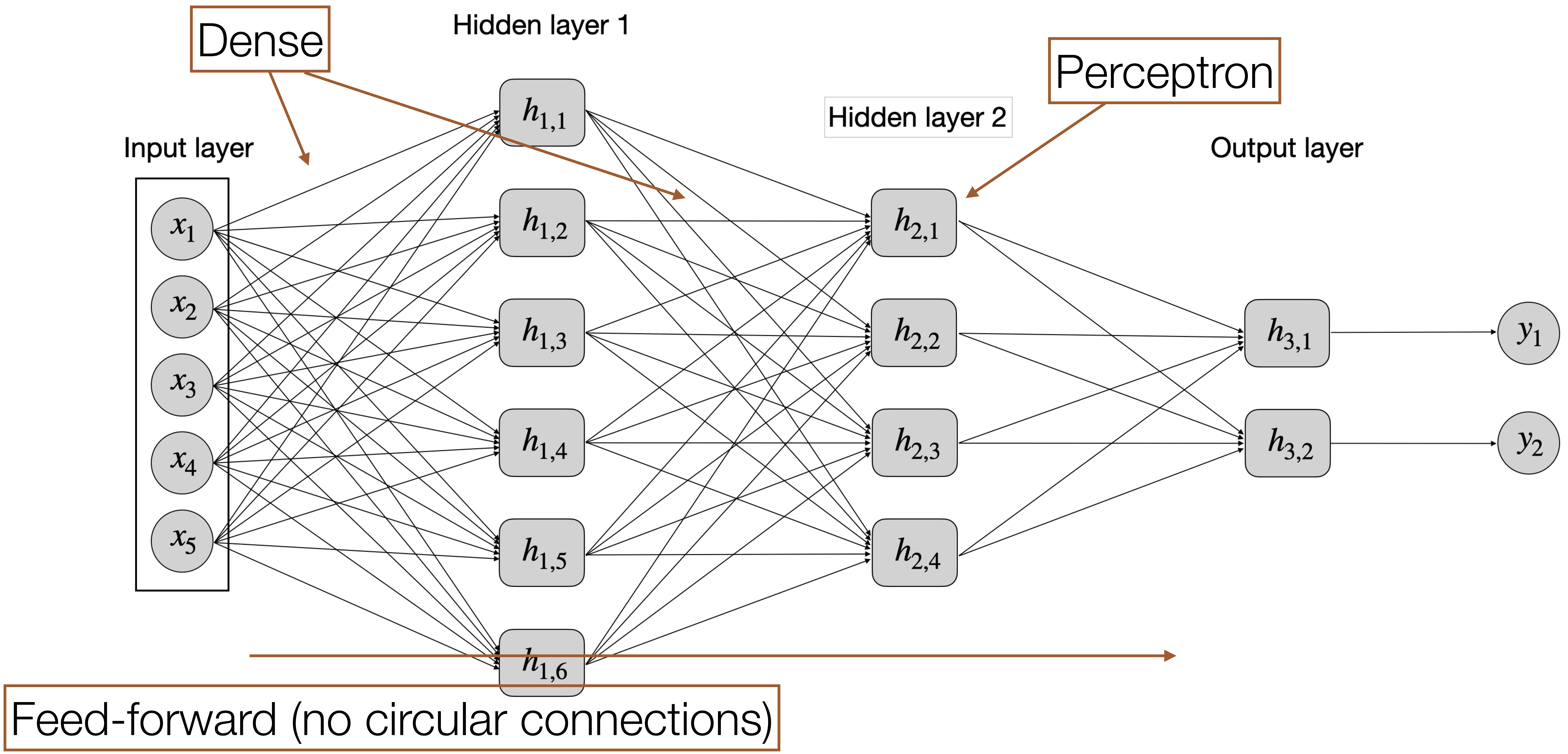


Ground for linear regression

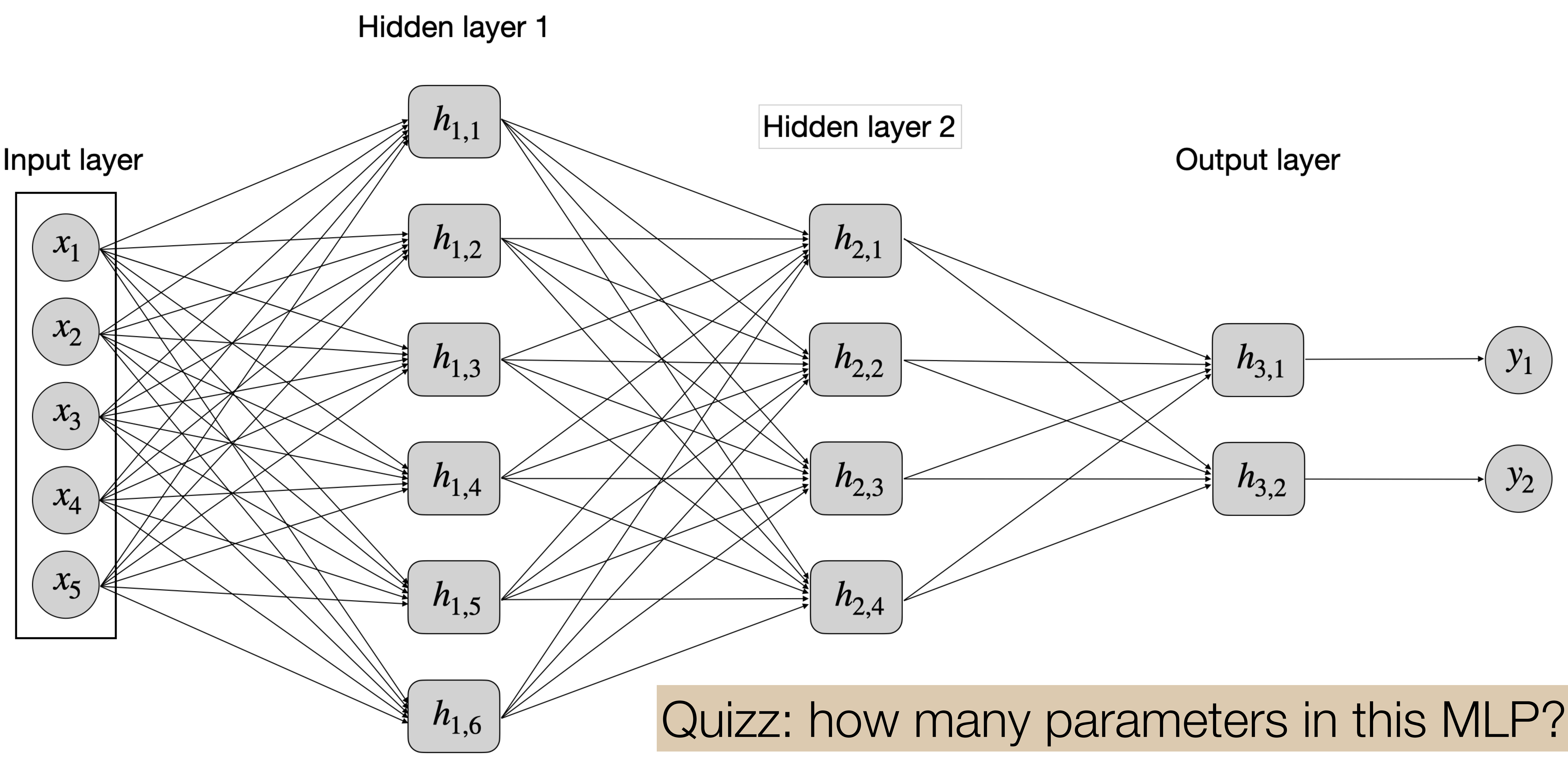
Multi Layer Perceptron (MLP) or Dense Neural Network (DNN)



Multi Layer Perceptron (MLP) or Dense Neural Network (DNN)



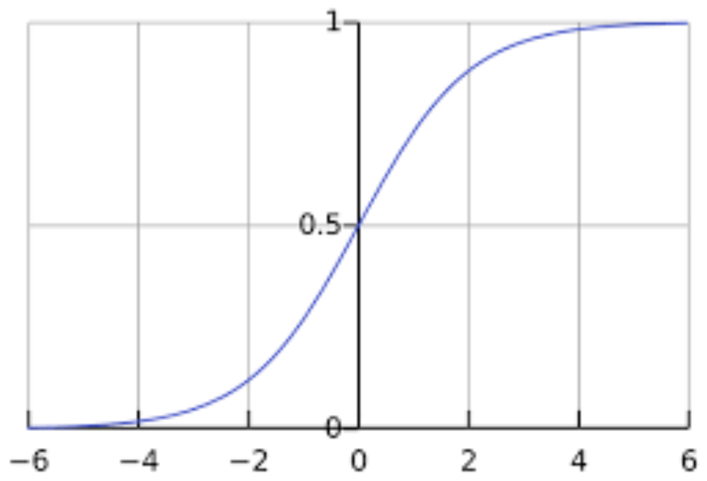
Multi Layer Perceptron (MLP) or Dense Neural Network (DNN)



Quizz: how many parameters in this MLP?

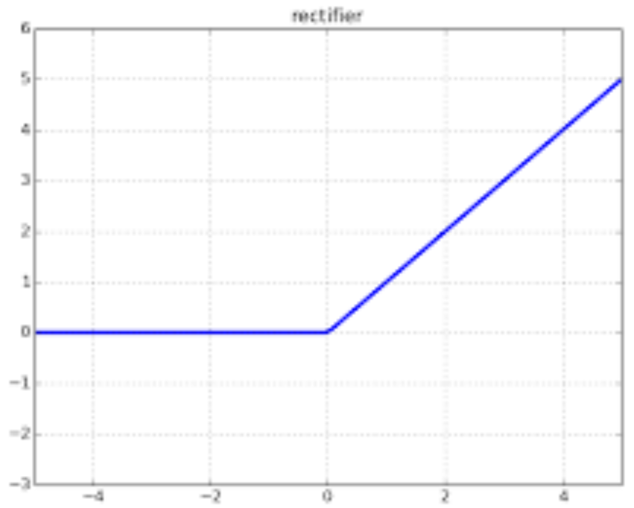
Activation functions

softmax



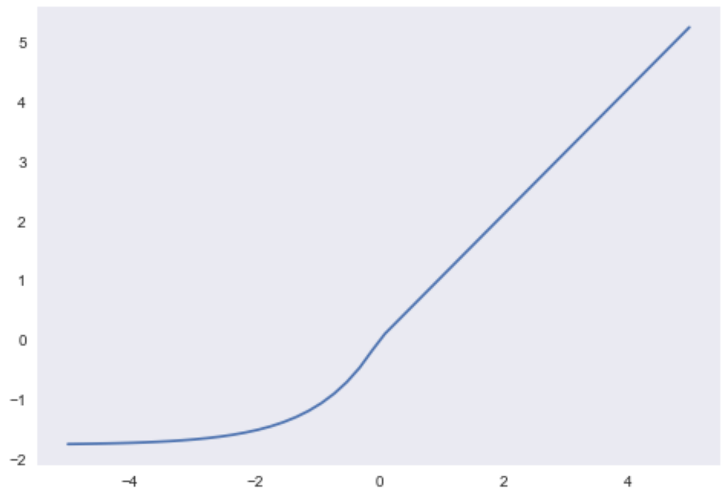
$$\sigma : \mathbb{R}^K \to (0,1)^K$$
$$\sigma(\mathbf{z})_j = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}} \quad \text{for } j = 1, \ldots, K.$$

ReLU



$$f(x) = x^+ = \max(0, x)$$

selu



$$\mathrm{selu}(x) = \lambda \begin{cases} x & \text{if } x > 0 \\ \alpha e^x - \alpha & \text{if } x \leqslant 0 \end{cases}$$

etc.

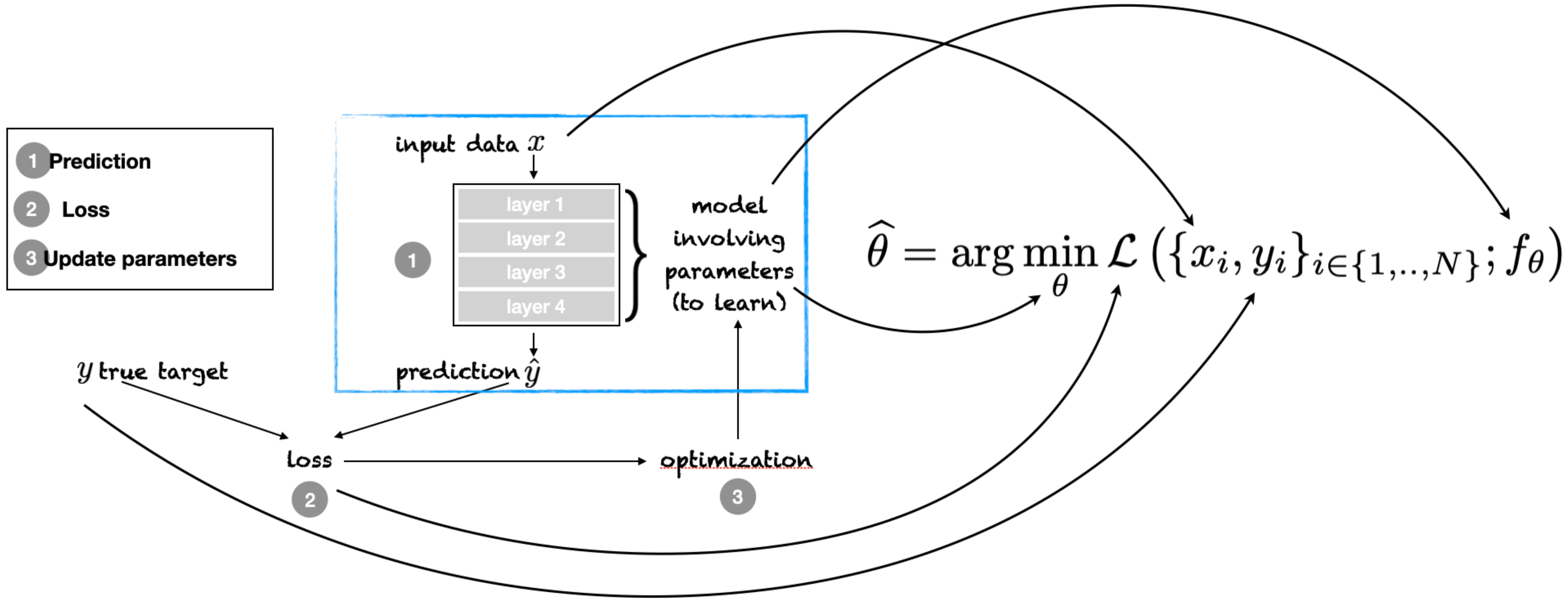
Let’s play with MLPs

- https://playground.tensorflow.org

The learning process

Loss function, Gradient descent, backpropagation

Loss functions



$$J(\theta) = \frac{1}{m}\sum_{i=1}^{m} |y^{(i)} - f(x^{(i)};\theta)|$$

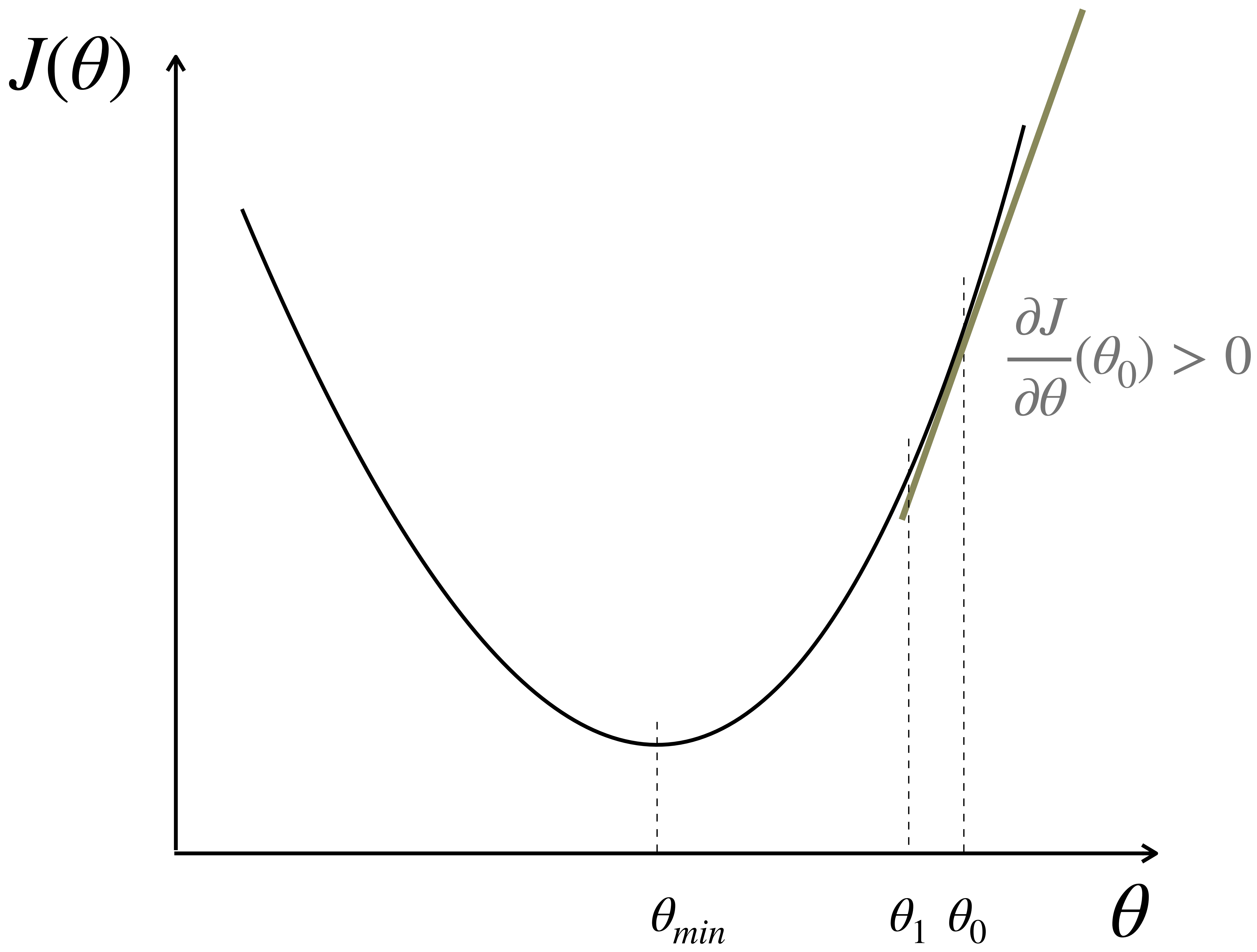
$$J(\theta) = \frac{1}{m}\sum_{i=1}^{m} \left(y^{(i)} - f(x^{(i)};\theta)\right)^2$$

$$J(\theta) = -\frac{1}{m}\sum_{i=1}^{m} y^{(i)} \log f(x^{(i)};\theta)$$

etc.

($\mathscr{L}$ has become J)

Gradient descent



Gradient descent:

$$\theta_{n+1} = \theta_n - \lambda \nabla J(\theta_N)$$

Learning rate

Calculating the gradient: backpropagation

- on the black board

Introduction to pytorch

Basic tips and practice

Basic tips for pytorch

- High-level python library (platform-independent, CPU/GPU, distributed computing,…)

- Very similar to numpy syntax but different (eg, torch.tensor vs. numpy.array)

- Specific types and classes to implement forward and backward passes through a network

- Object-based coding through classes for neural models

- Among the two main state-of-the-art DL frameworks with Tensorflow (but there are others, JuliaDiff, Jax, Caffe,…)

Practice

- https://github.com/CIA-Oceanix/DLOA2023/blob/main/lectures/notebooks/intro_pytorch_learning.ipynb