Machine Learning

Session 3 - Neural networks


hadrien.salem@centralelille.fr


introduction-to-data-science

Introduction

What did we do last time?

Course outline

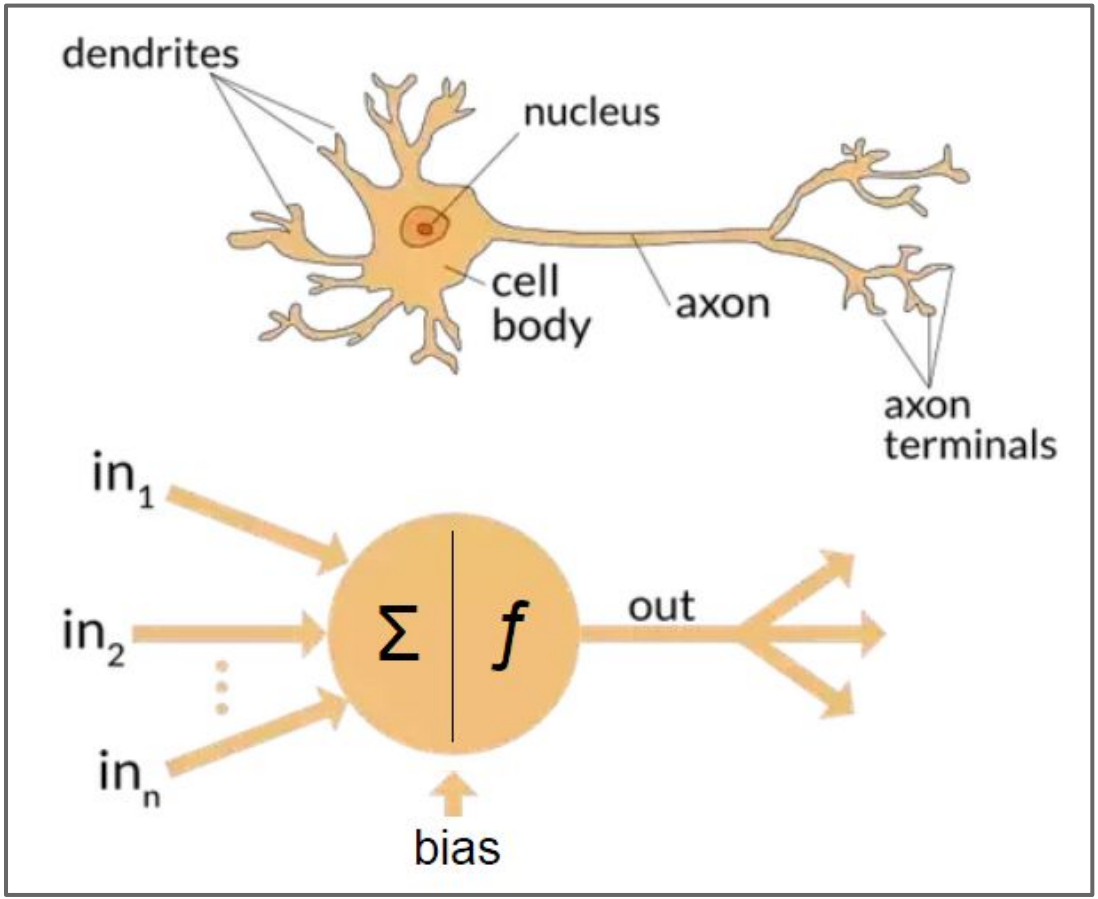
Intro to ML course

Session 1: Introduction to ML & Regression

Session 2: Supervised classification

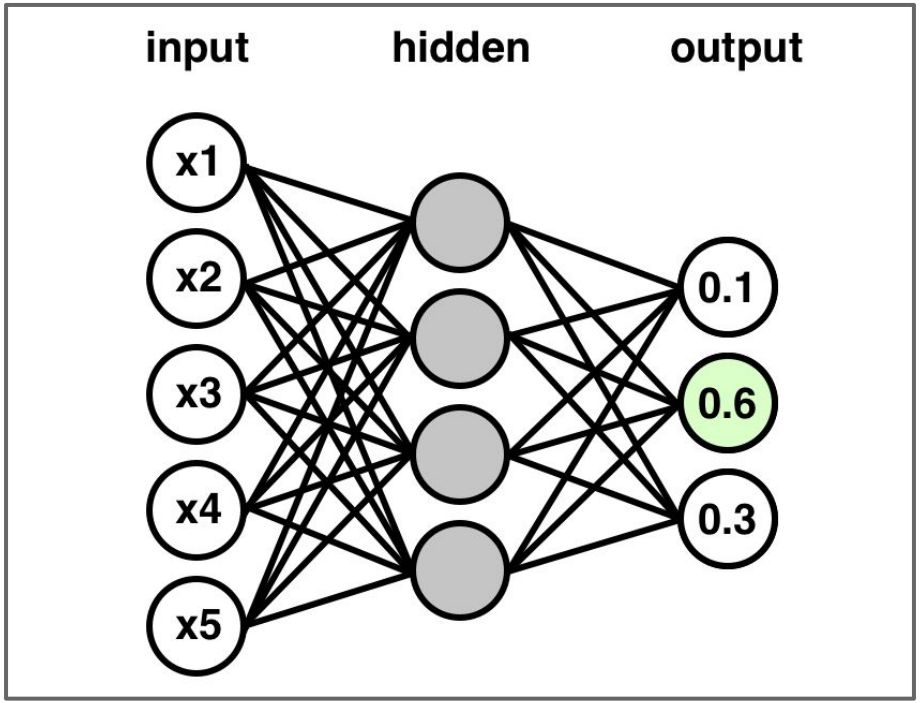
Session 3: Neural networks

What are artificial neural networks?

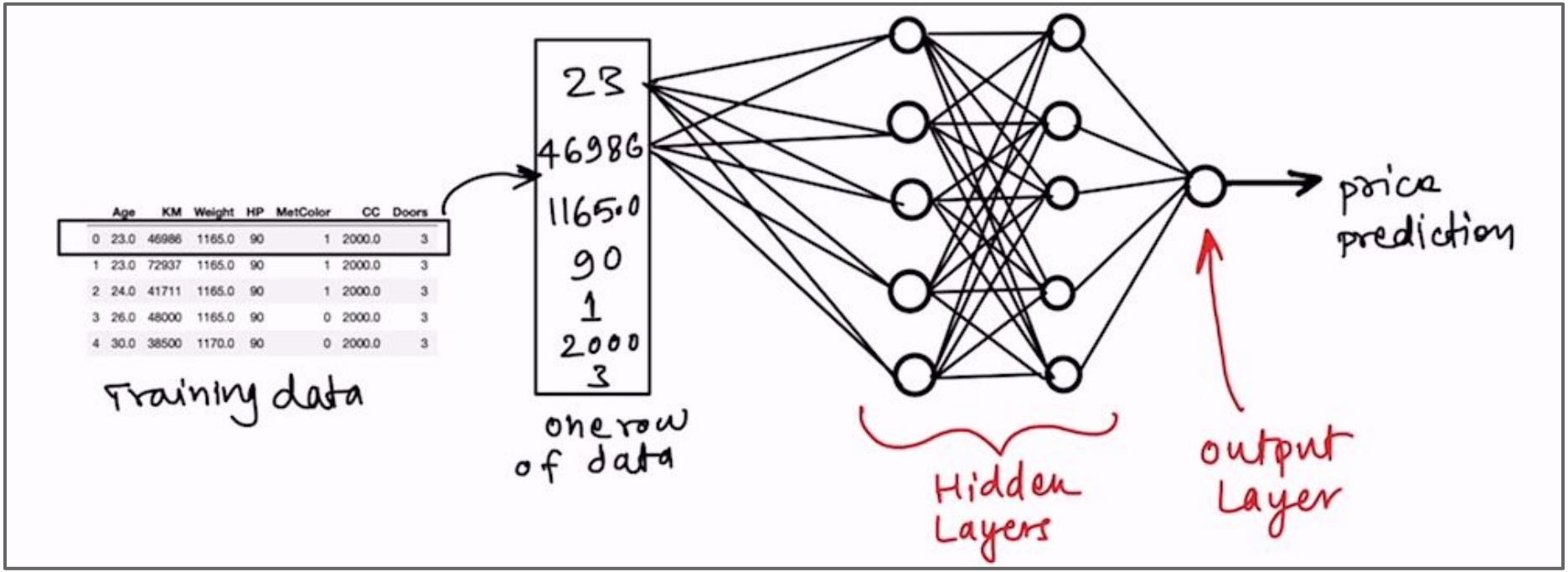


Neural networks are named after an analogy with the human brain

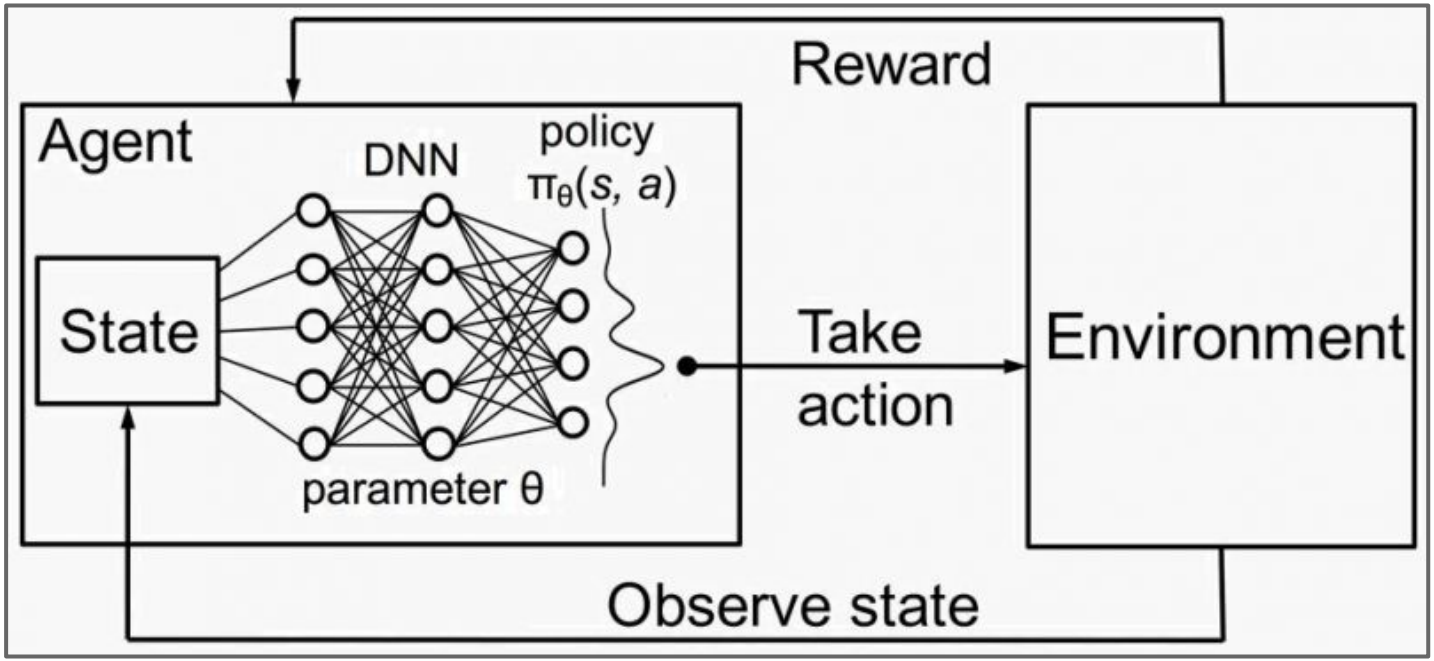
Image source



The network outputs the probability for the sample to belong to each class



The network outputs the predicted value



The network learns the best action to execute in a certain state

Image source

Strength and weaknesses of neural networks

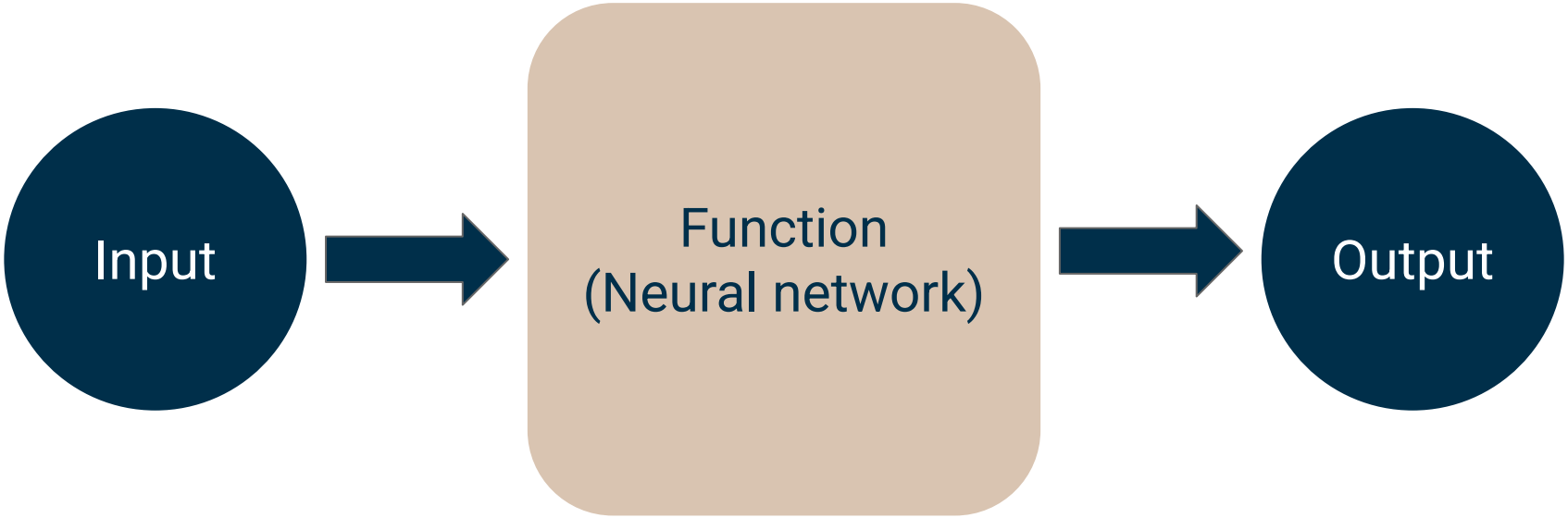
Strengths

- Extremely flexible
- Learns nonlinear functions
- Fast predictions once trained
- Scalable to large datasets

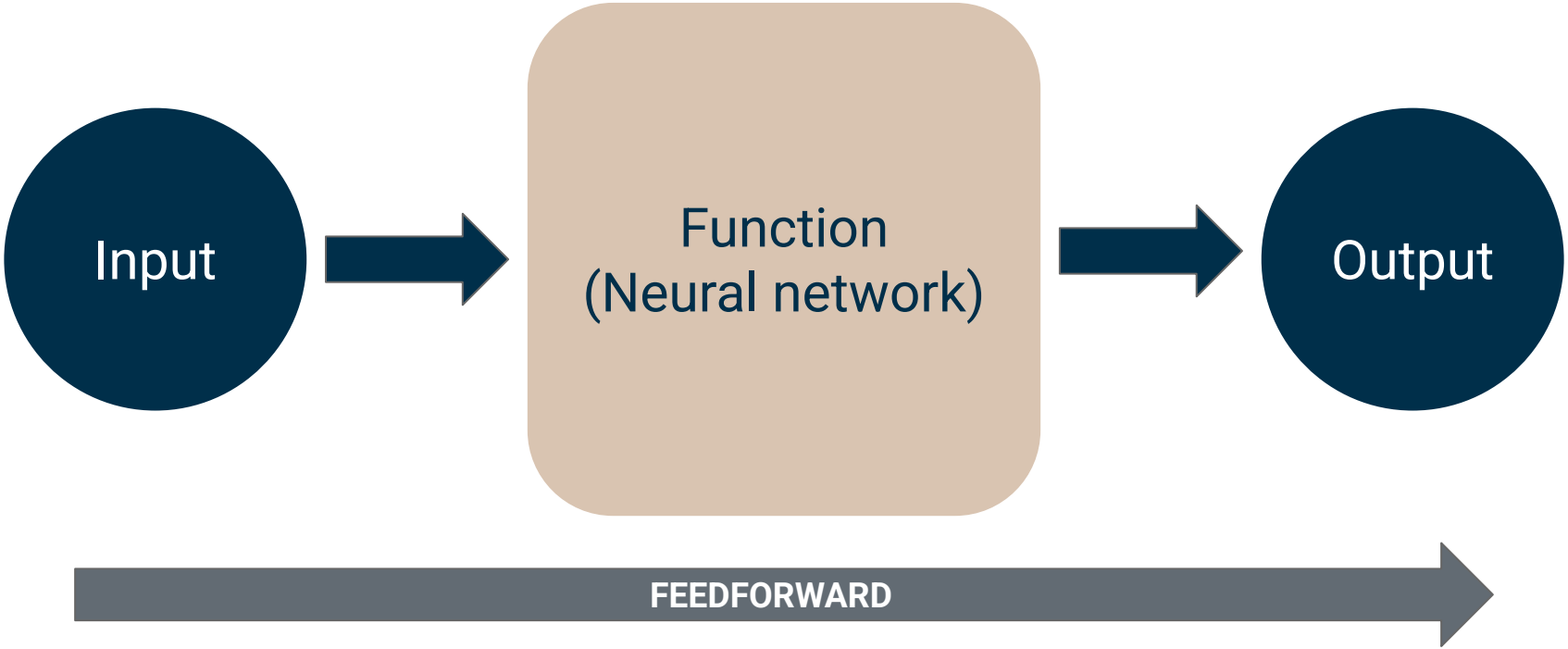
Weaknesses

- Very difficult to interpret
- Computationally expensive
- Very slow training without GPUs

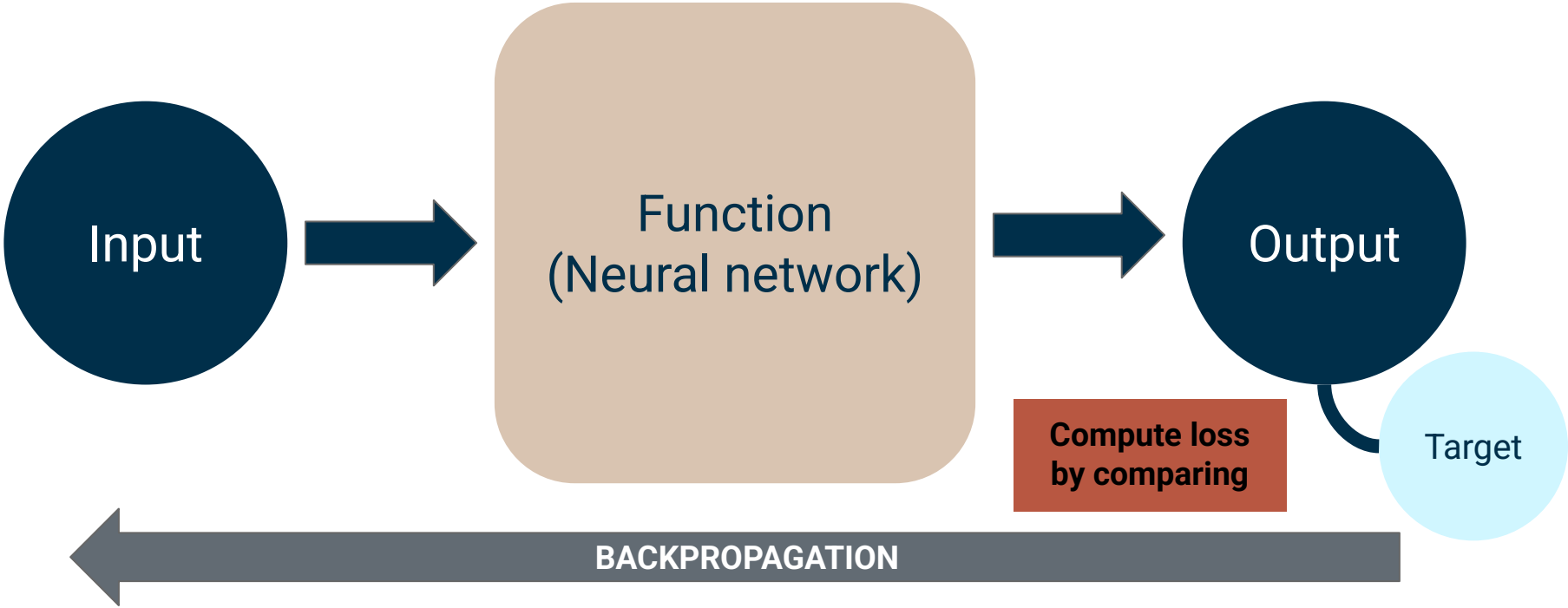
How do ANNs work?



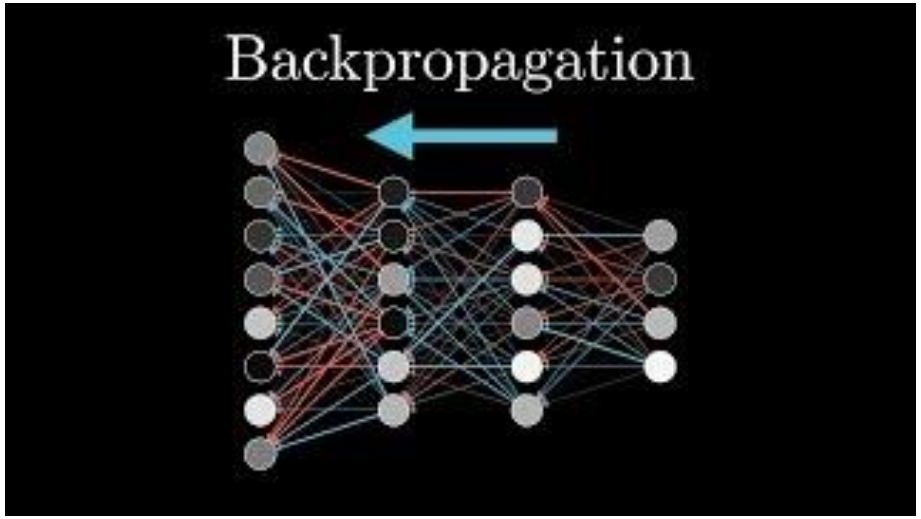
Neural networks are algorithms that learn a function linking input and output

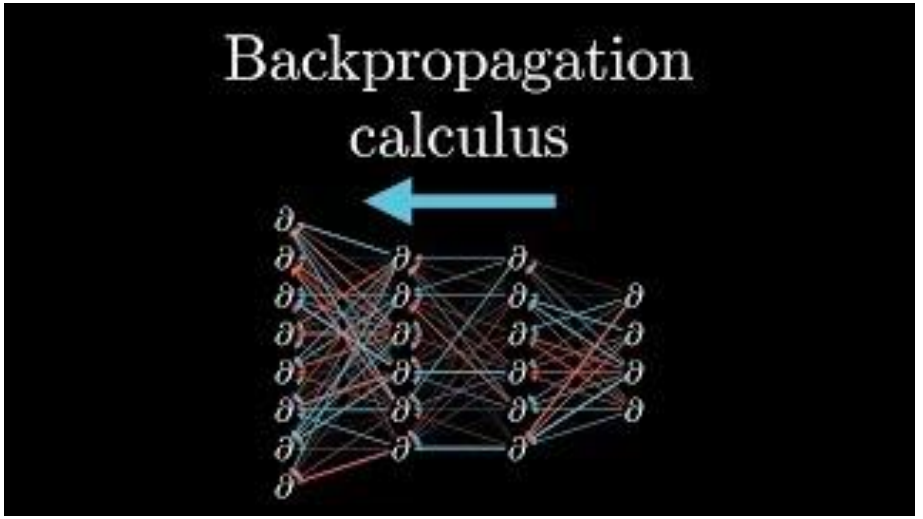


In order to train the model, inputs are “fed forward” into the network, and outputs are computed



The loss is computed, and the error is propagated back into the network to adjust weights





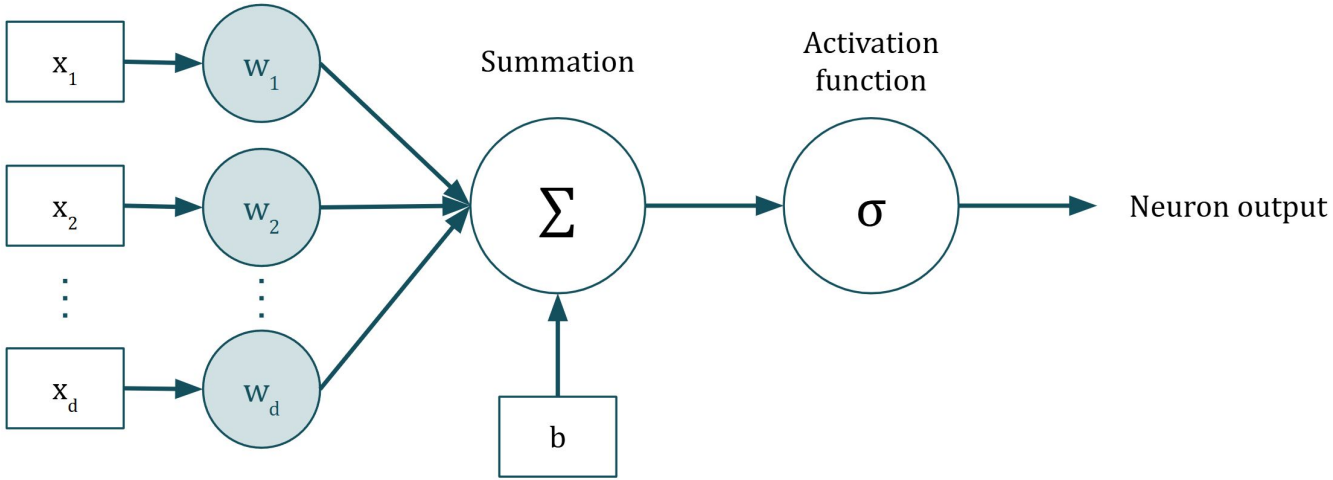
https://www.youtube.com/watch?v=Ilg3gGewQ5U

https://www.youtube.com/watch?v=tIeHLnjs5U8

If you are interested in the math behind it, check out these excellent videos on the subject!
In short: backpropagation is adjusting the network's weights to reduce the loss

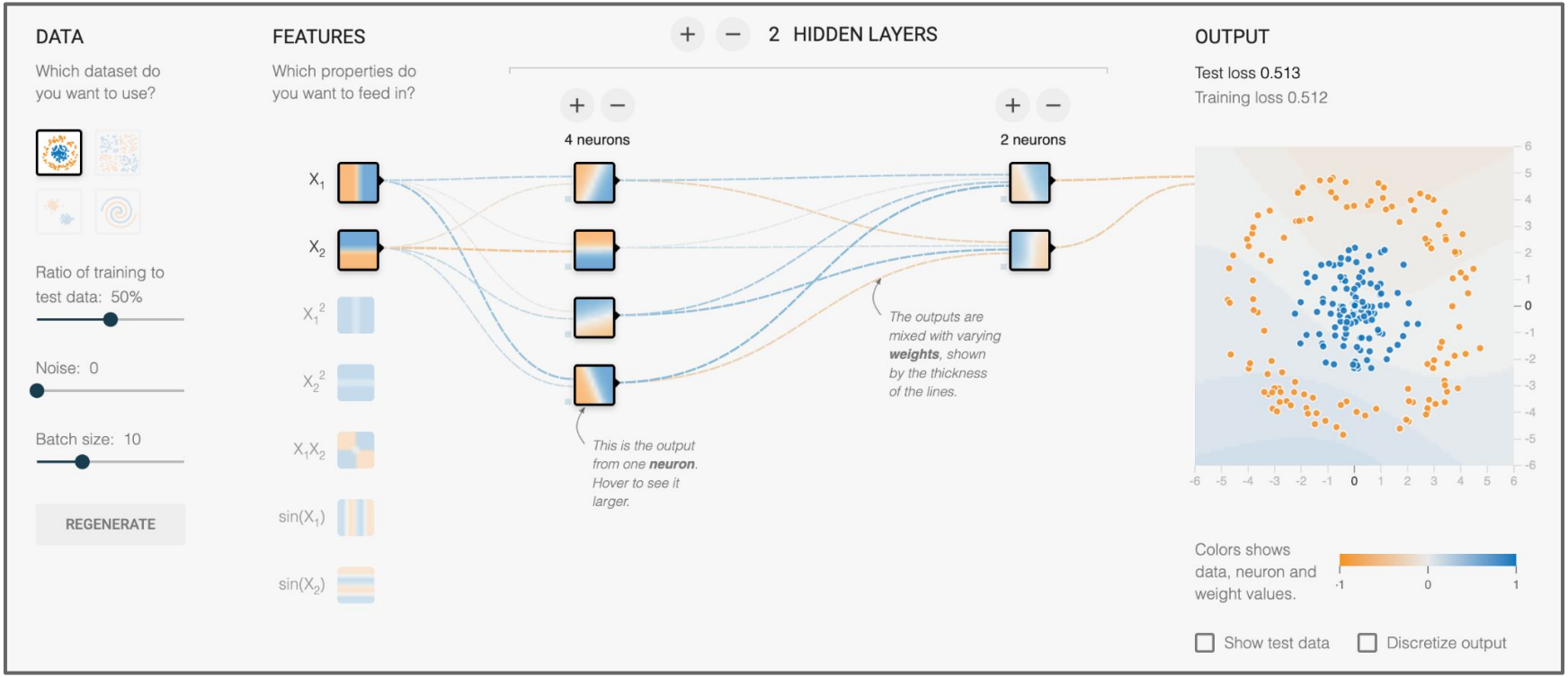
Backpropagation is performed with stochastic gradient descent

What are ANNs made of?



Weights are the **parameters** of the algorithm. They are adjusted during the learning process.

The **activation function** introduces nonlinearity in the network. This means the model can learn complex functions.

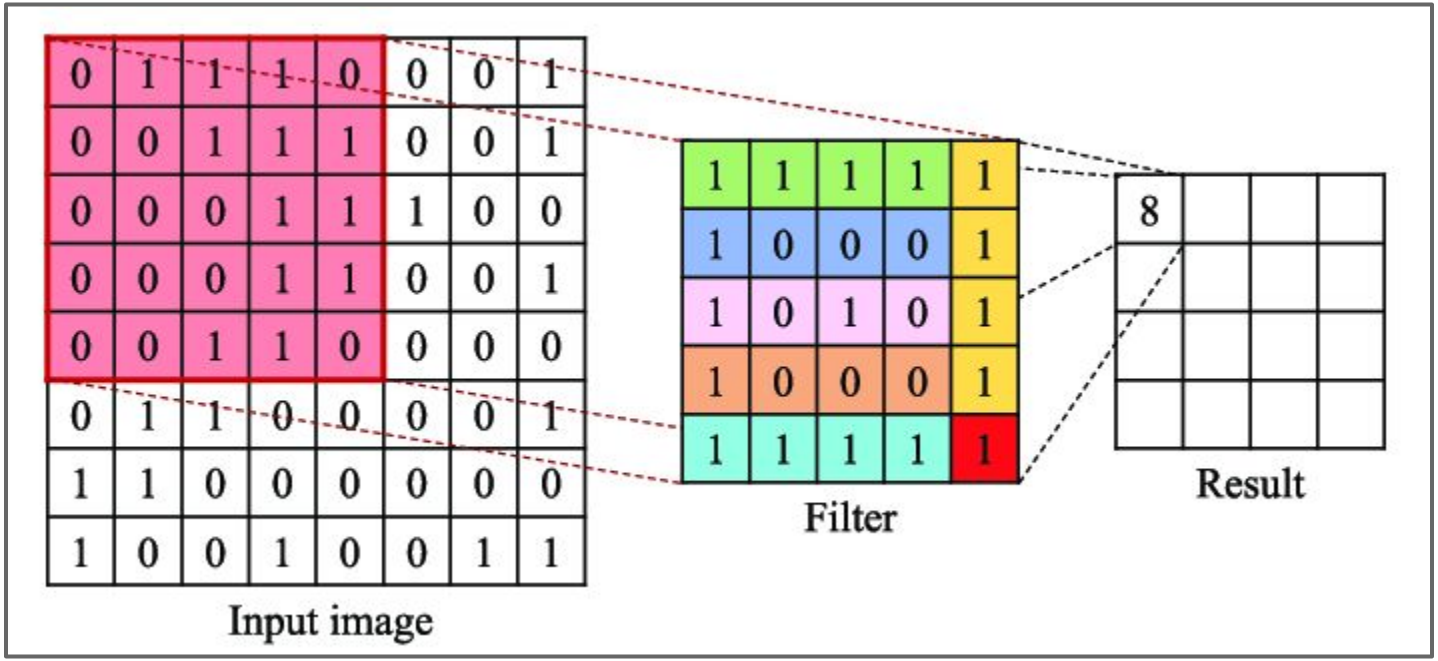


https://playground.tensorflow.org/

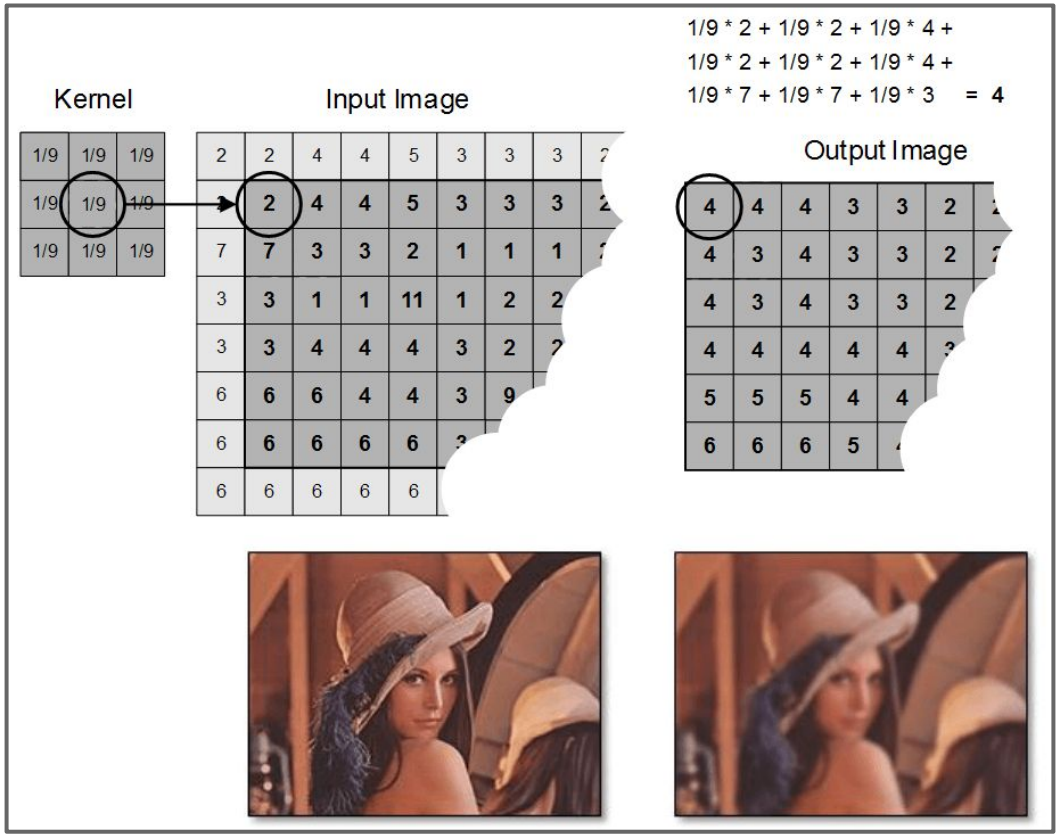
Input, output and hidden layers

Convolutional Neural Networks

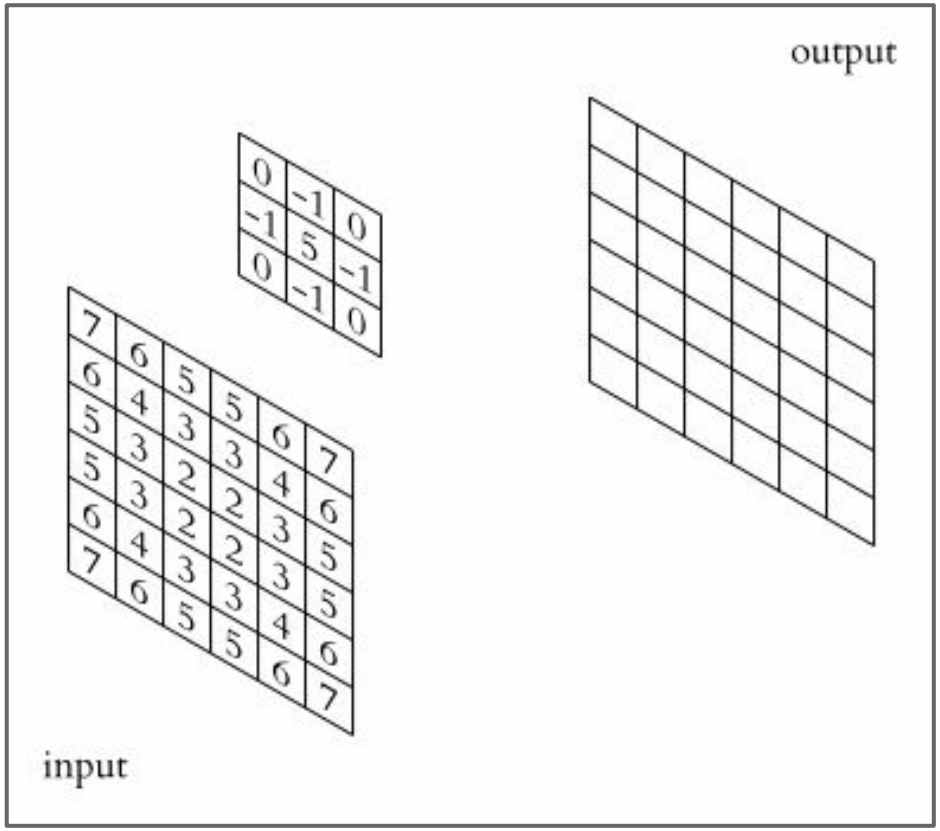
Common ANN for image processing

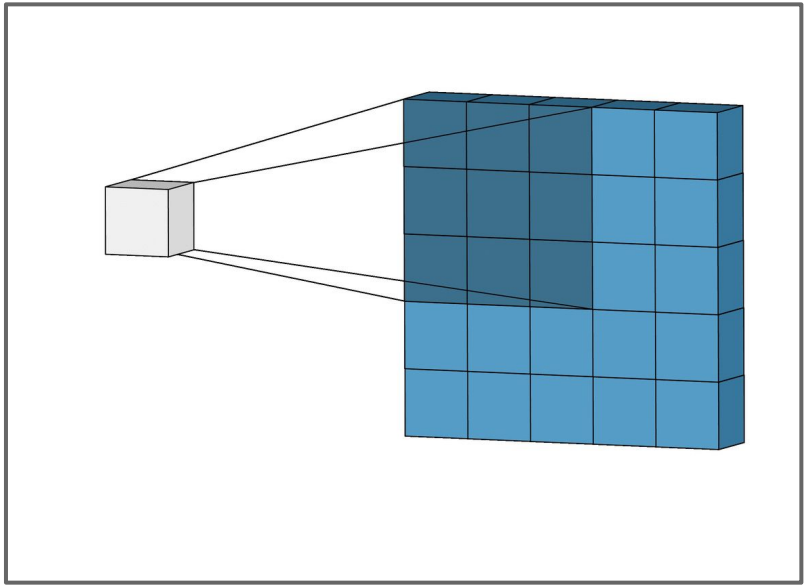


What is a filter?

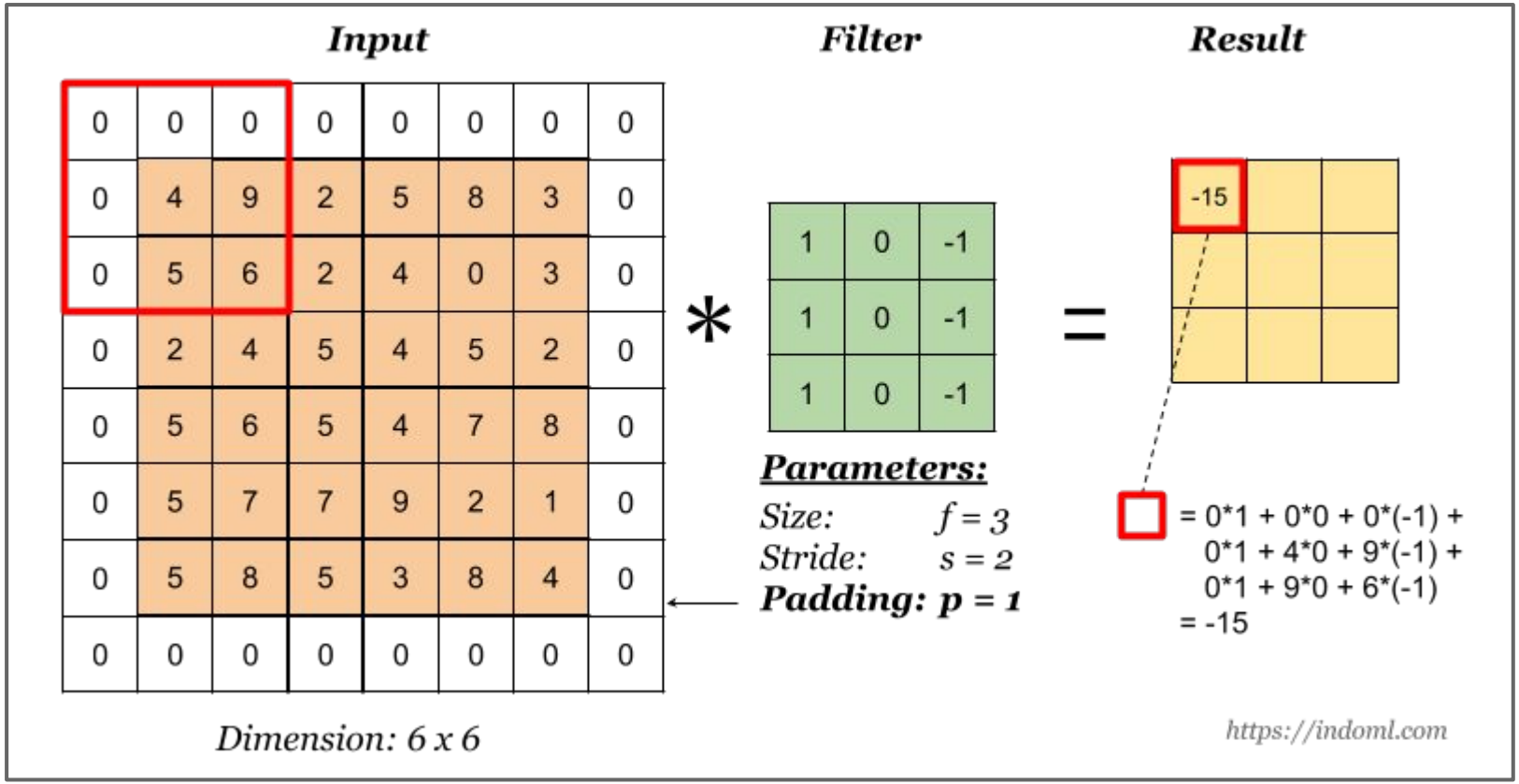


Filters create an output image with different properties (capturing certain characteristics)

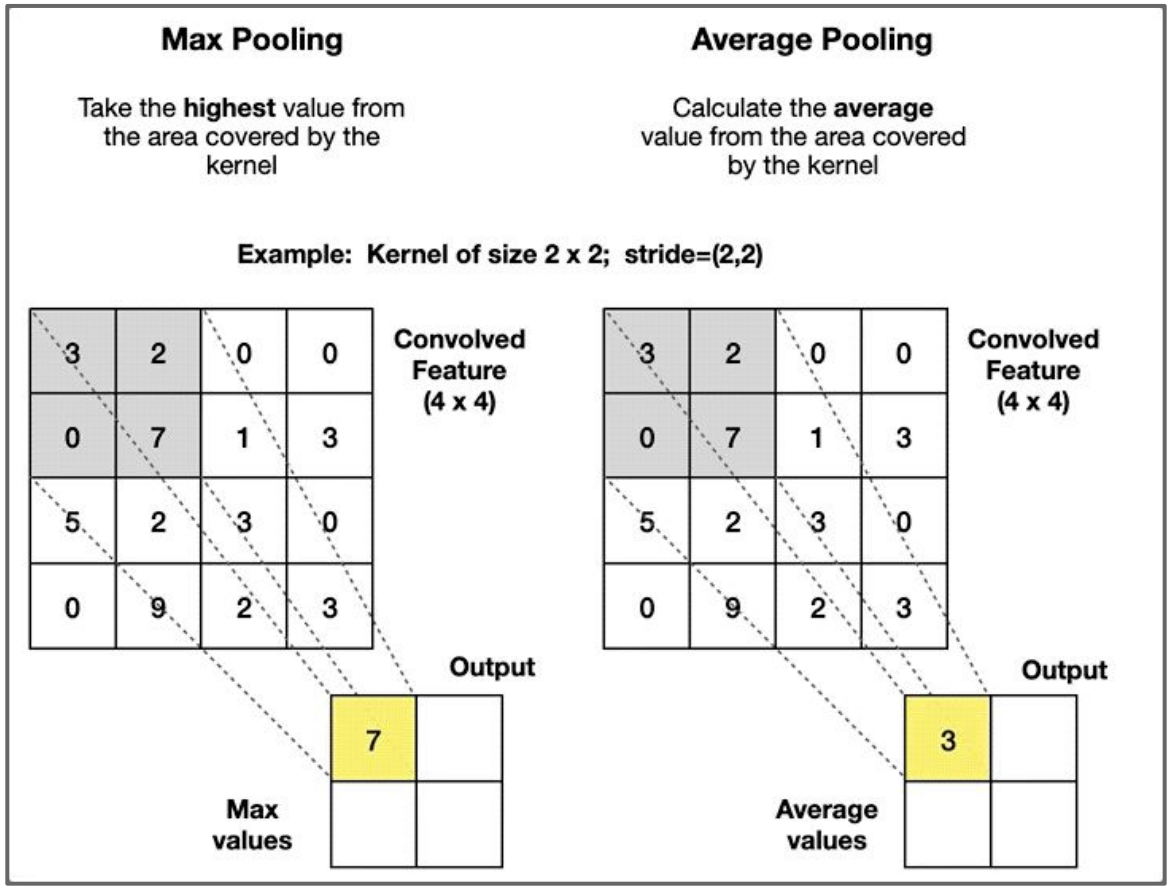




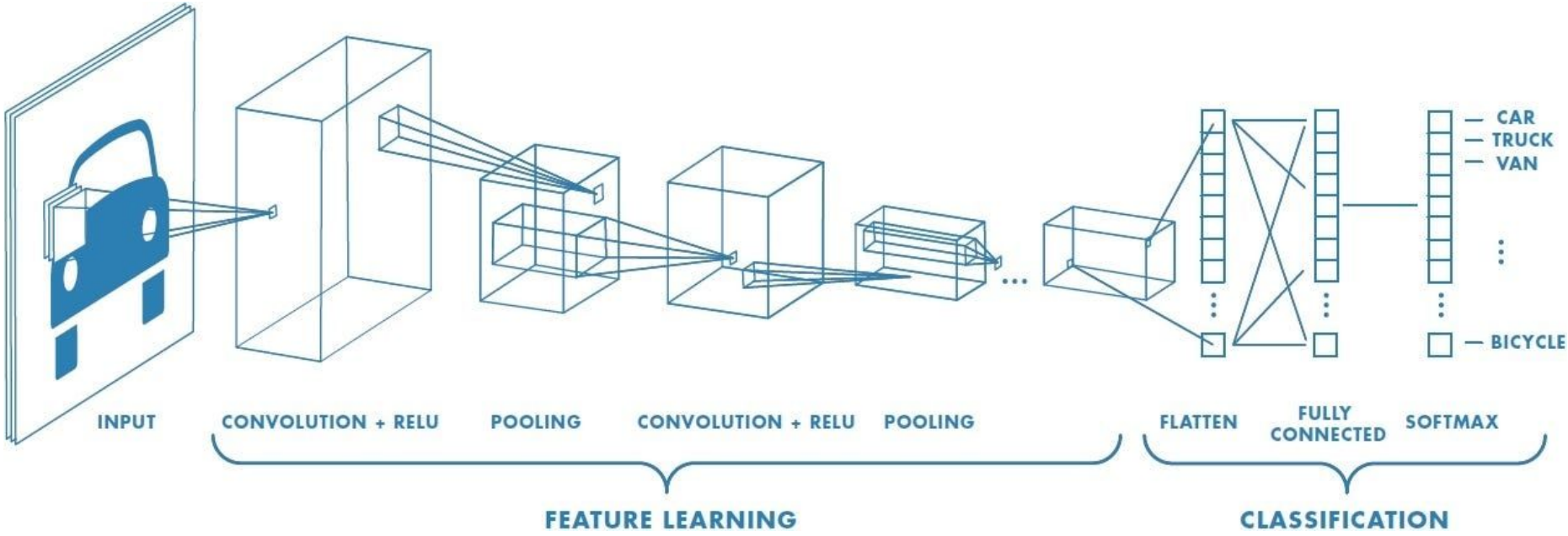
Convolutional product



Parameters for the convolutional product



Pooling as a dimensionality reduction technique



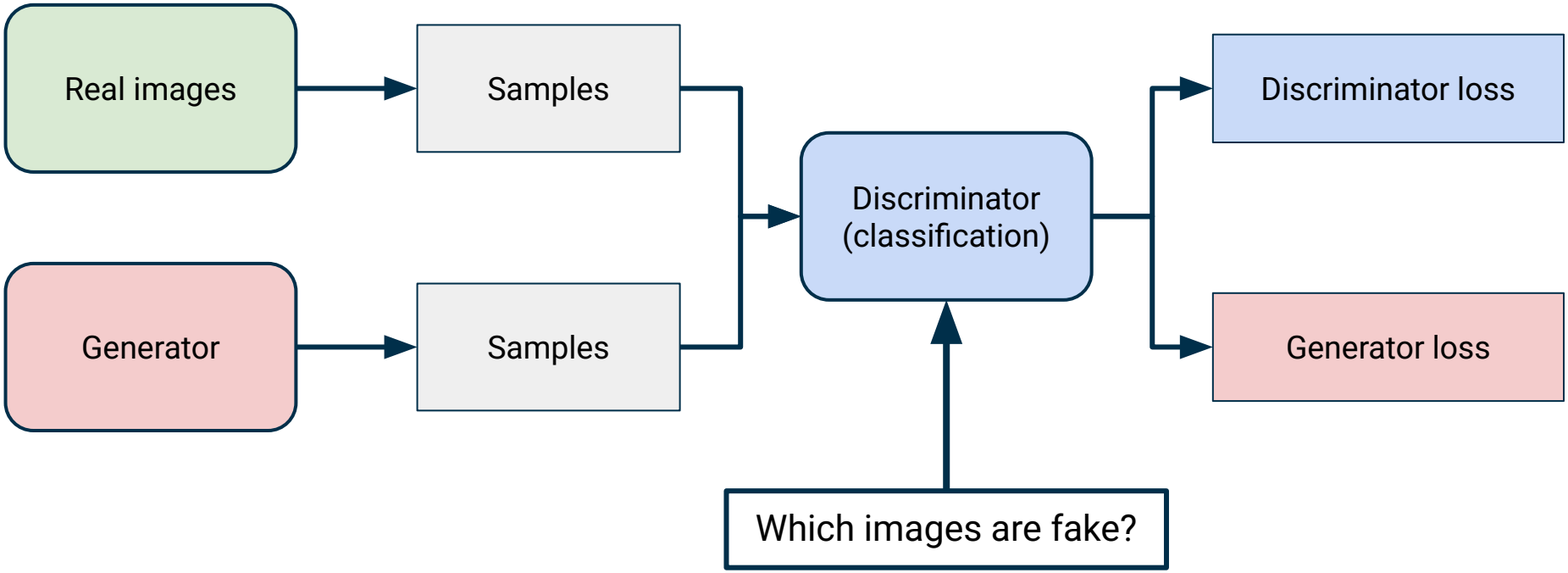
Convolutional neural networks: learning filters (learning a representation of your data)

CNN - Summary

CNNs are powerful networks used mainly for image processing

- **CNNs learn filters = kernels**
  - Filters are operations applied locally to an matrix
- **During feedforward, kernels slide across the picture to generate the output**
  - “Picture” can be generalized to any matrix (e.g. time series)
- **Convolutional layers are usually coupled with auxiliary operations**
  - Maxpooling favors strong signals
  - Average pooling is more balanced
  - Feedforward networks let CNNs have the desired output shape

Generative Adversarial Networks



GANs rely on a generator and a discriminator, trained simultaneously

GANs let you generate artificial data

When to use GANs?

Sample generation in imbalanced contexts

Text or image generation

Other state-of-the-art networks

Other neural network architectures

In recent years, several ANN architectures have been invented to solve a wide variety of problems

Recurrent Neural Network (RNN)

- Designed for processing **sequences** of data
- Connections that loop back on themselves to capture information from previous steps
- Used for NLP and time series processing

Long Short-Term Memory (LSTM)

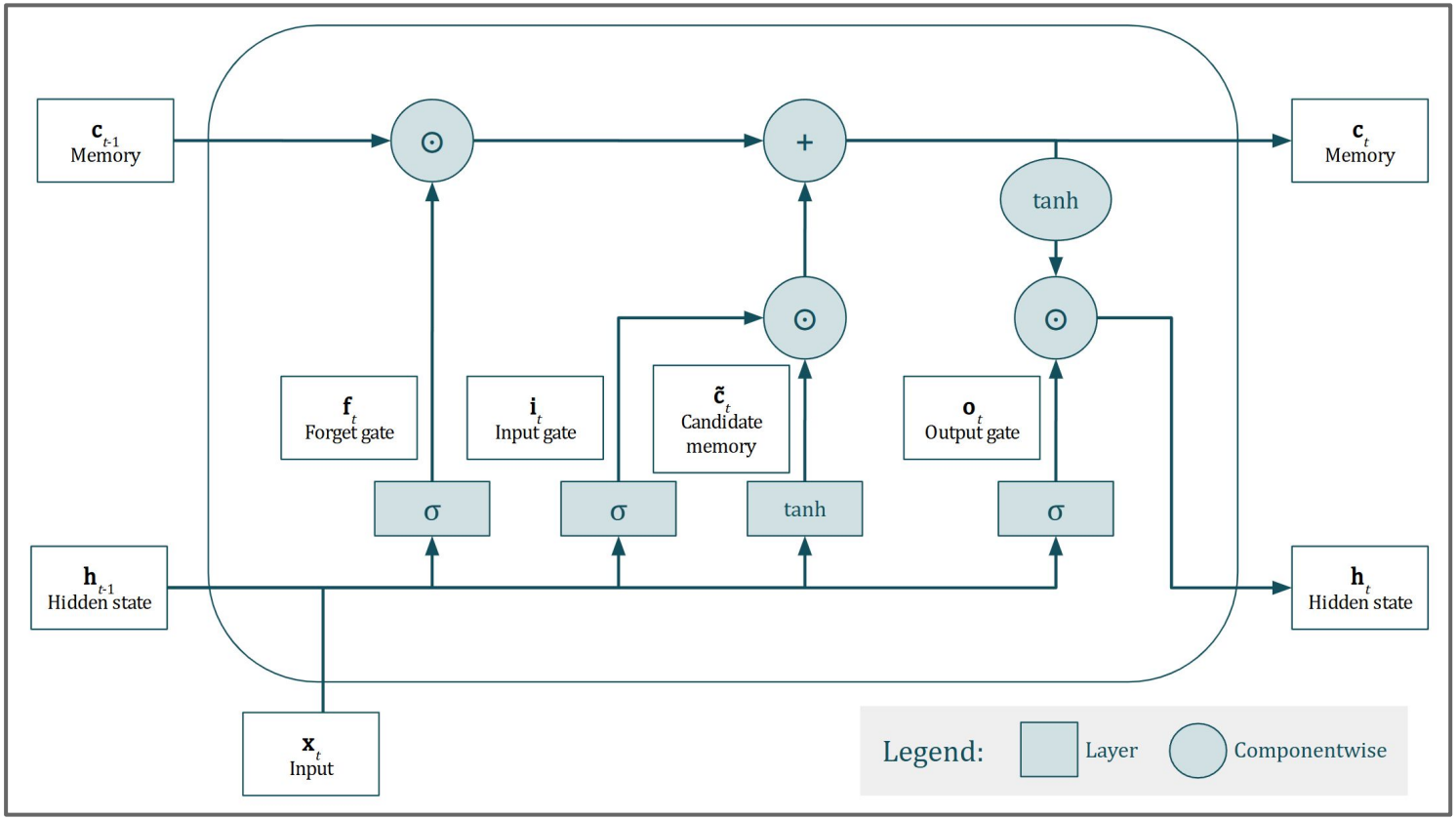
- Designed as an **improvement to RNNs** (addresses the vanishing gradient problem)
- Capture long-range dependencies (e.g. in NLP)

Transformer Networks

- Initially introduced for NLP
- Rely on the attention mechanism, flexible

Generative Pre-trained Transfomer (GPT)

- Based on Transformer networks
- Trained on **large** amounts of data



An example of ANN designed for sequential data: LSTM

And more!

Thousands of new ANN architectures and pre-trained models come out every year. The best way to prepare is to know where to look for the tools you need.

Other models

- **Image segmentation**: U-net, YOLO, etc.
- **Recommender systems**: LightFM, etc.
- **Reinforcement learning**

Where to find models

- **Science articles**: New networks usually get published in papers, but it can be difficult to find what you need in the ocean of articles that come out every year.
- **HuggingFace**: The largest repository for pre-trained models, usually has better models than what you can implement yourself.
- **Open-source community**: Machine Learning has a very active and open community. A lot of models are available on GitHub with their source code and usage instructions.

Practical work

The notebook contains all the necessary instructions

Debrief

Debrief

What did we learn today?

What could we have done better?

What are we doing next time?

Machine Learning

Session 3 - Neural networks


hadrien.salem@centralelille.fr


introduction-to-data-science