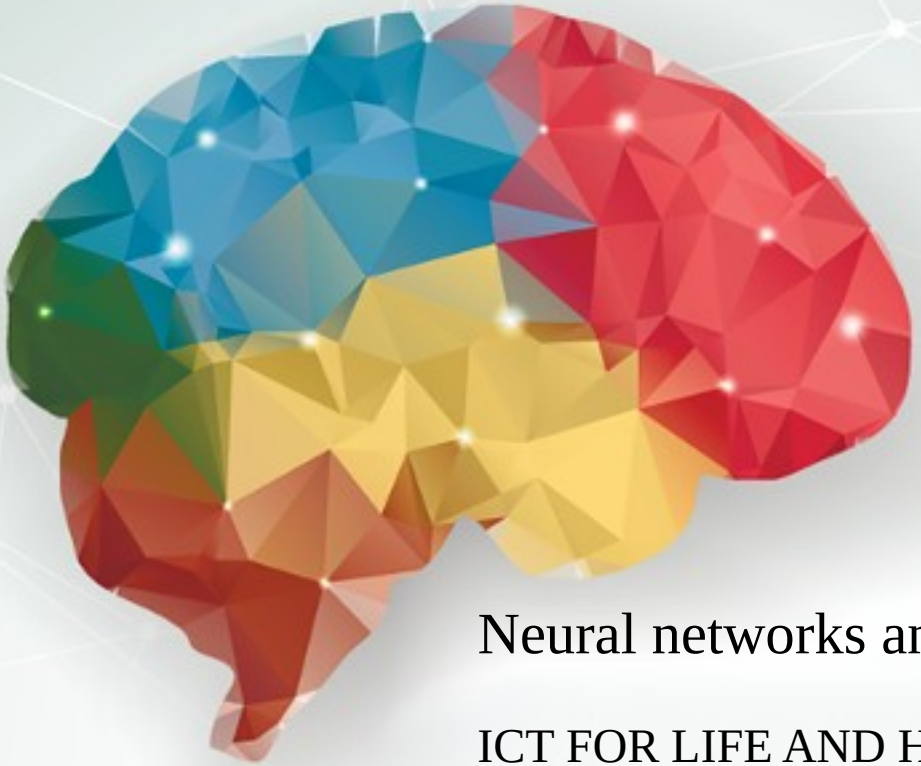

Neural networks and deep learning

ICT FOR LIFE AND HEALTH - Department of Information Engineering

PHYSICS OF DATA - Department of Physics and Astronomy
COGNITIVE NEUROSCIENCE AND CLINICAL NEUROPSYCHOLOGY - Department of Psychology

A.A. 2019/20 (6 CFU)

Dr. Alberto Testolin,
Dr. Federico Chiariotti (chiariot@dei.unipd.it)

The pytorch dataloader

- The torch.utils.DataLoader class can import a dataset
  - Customizable order
  - Automatic batches
  - Multi-process and GPU loading
- Options are configured by the constructor
  - DataLoader(dataset, batch_size=1, shuffle=False, sampler=None,
    batch_sampler=None, num_workers=0,
    collate_fn=None, pin_memory=False,
    drop_last=False, timeout=0, worker_init_fn=None)

Dataset types

1. **Map-style datasets (keys are mapped to data samples)**
2. **Iterable datasets (stream-like)**

Map-style: load an image file depending on the index

Iterable: read a feed in real time or from a remote server

Batching

The dataloader performs automatic batching using the Sampler class, but users can specify a batch_sampler for map-style datasets

Memory pinning

Tensors can be memory pinned by passing pin_memory=True

Memory pinning only works for Tensors and maps and iterables that contain tensors, and makes it a lot faster to transfer data to the GPU