Introduction to Deep Learning (I2DL)

Exercise 3: Datasets

Today's Outline

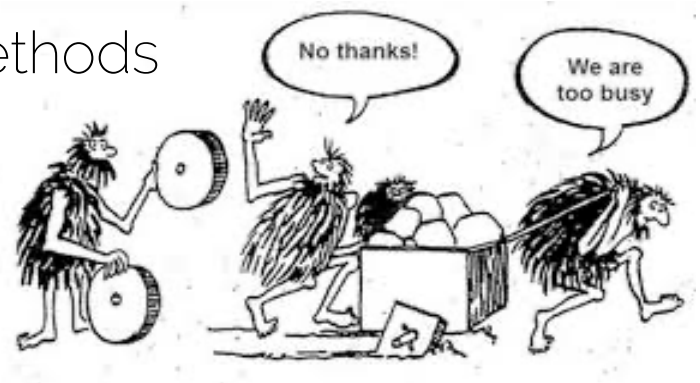
- Exercises outline
  - Reinvent the wheel
  - Pillars of Deep Learning
- Contents of the first python exercise
  - Example Datasets in Machine Learning
  - Dataloader
  - Submission 1
- Outlook exercise 4

Reminder

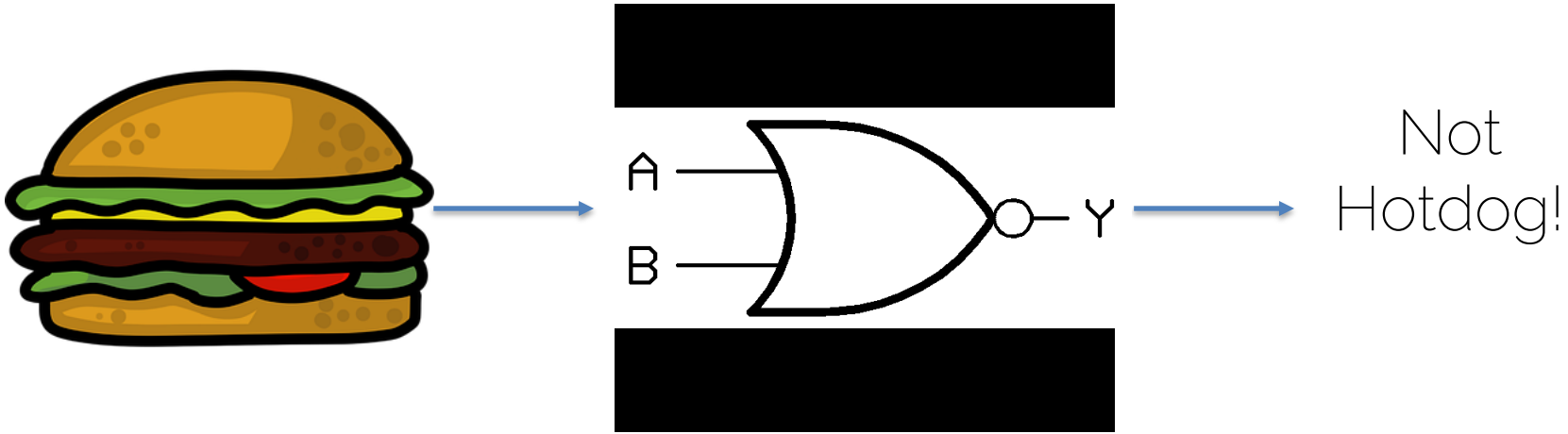
- Unregistered TUM/LMU students
  - Link available to google form available on our website
- Use Piazza for questions and private questions
- Office hours starting this week!
  - Schedule on Piazza
- Solutions
  - will be published together with following exercises

Your task for the exercises 3-5

- Implementation of
  - ~~Classic datasets and data loading~~
  - ~~Classification pipeline~~ using
    - Traditional machine learning methods
    - Neural Networks
      - Layers
      - Optimizers
      - Etc.
- „Reimplement the wheel“



Why spend the effort?



Before lecture

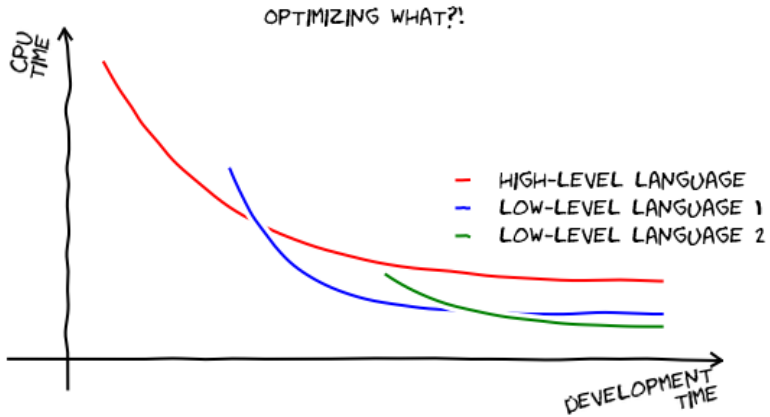
Expectations

Reality ☹

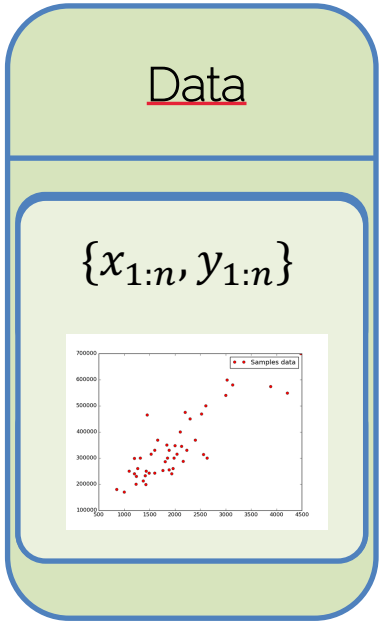
Knowledge + Experience

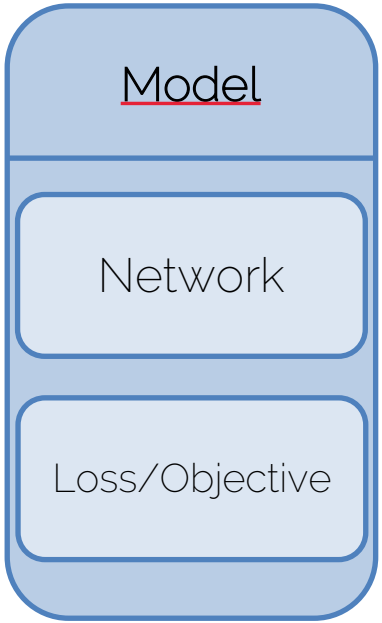
Why Python?

- Why python:
  - Very easy to write development code thanks to an intuitive syntax
  - Biggest language used in deep learning research (and probably production)



The Pillars of Deep Learning

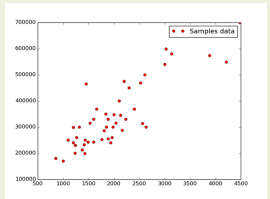




The Pillars of Deep Learning

Data

$$\{x_{1:n}, y_{1:n}\}$$



Model

$$\hat{y}_i = \theta_0 + \sum_{j=1}^{d} x_{ij}\theta_j$$

bias

$$J(\boldsymbol{\theta}) = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2$$

Solver

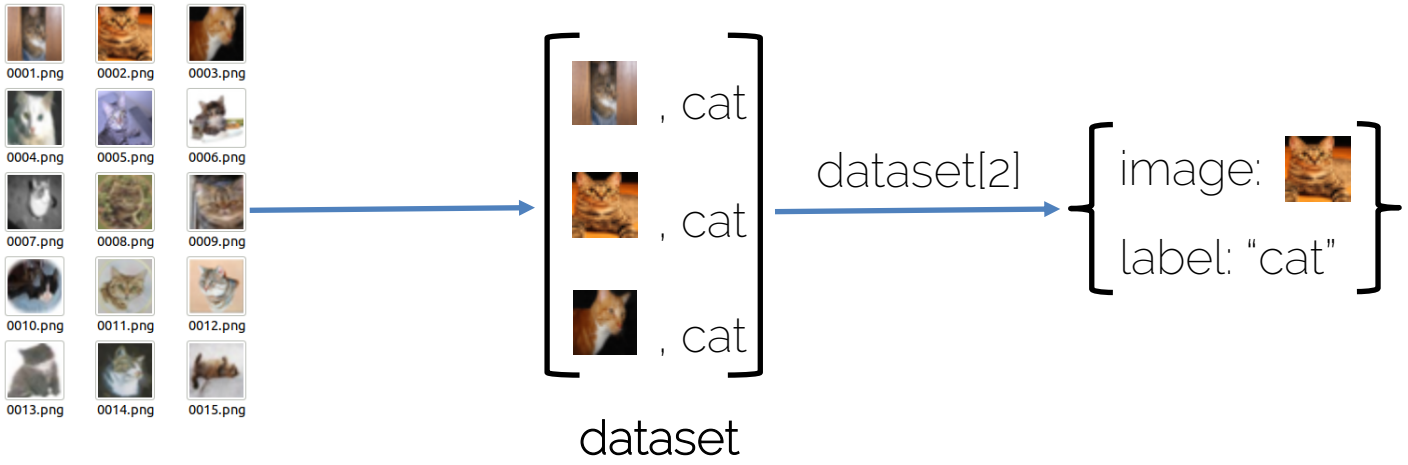
Optimizer

Training Loop

Validation

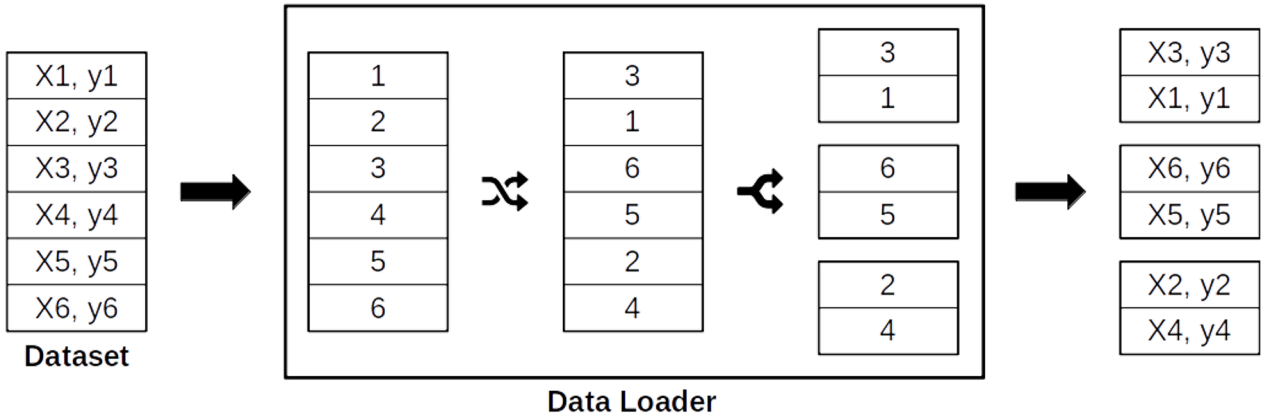
Exercise 3: Dataset

- Stores the data in an efficient, accessible form
- Performs data preprocessing steps using Transforms
- Example: Image Folder Dataset



Exercise 3: Dataloader

- Defines how to load the dataset for model training
- Shuffles the dataset
- Splits the dataset into small subsets



Overview Exercise 3

- Two notebooks
  - Dataset: CIFAR10
  - Dataloader

Fixed Deadline:
Nov 25, 2020 15.59

- Submission 1
  - Have to implement parts of both objects
  - Single submission file creation in Dataloder notebook

Hitchhiker's Guide: Notebooks

1. Run cells from top to bottom
2. Be careful when changing notebook cells
3. Don't code outside our boxes
4. Che
   - G
   - L

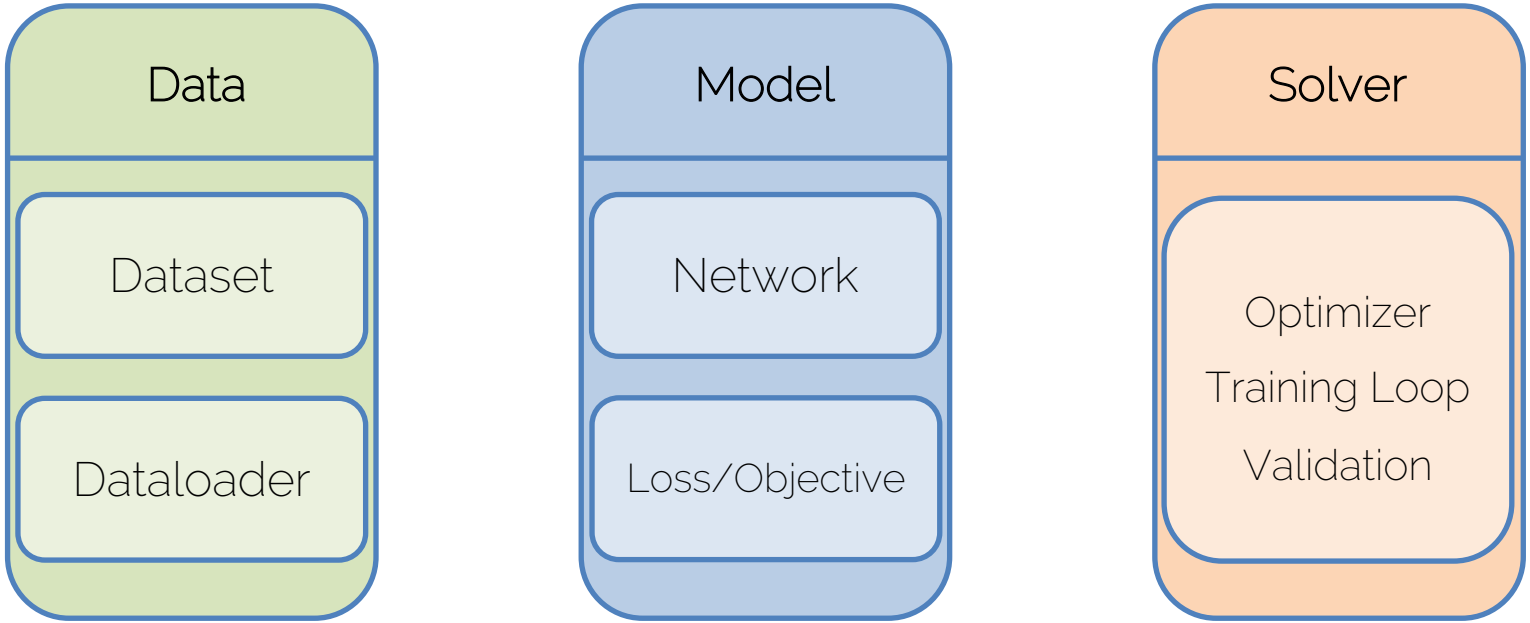
```
########################################################################
# TODO
# Implement the dummy machine functic
#
###################################

pass

###################################
#                              END OF YOUR CODE                        #
########################################################################

return 60
```

Task: Check Code

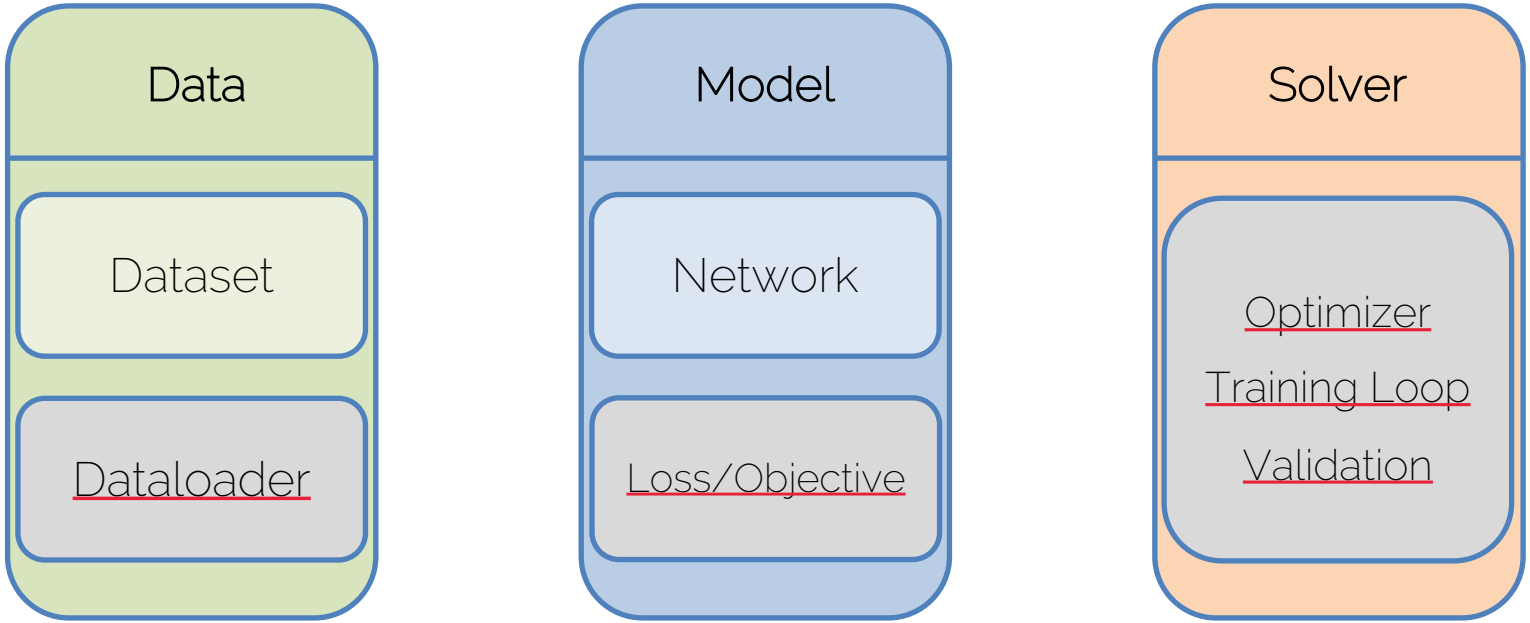
Please read `make_dataset(directory, class`
tasks. Additionally, it would be wise decision to ge
projects. As it is not beginner friendly, we removed i

4. Checking other code
   - Generally optional
   - Look out for green boxes

The Pillars of Deep Learning



The Pillars of Deep Learning

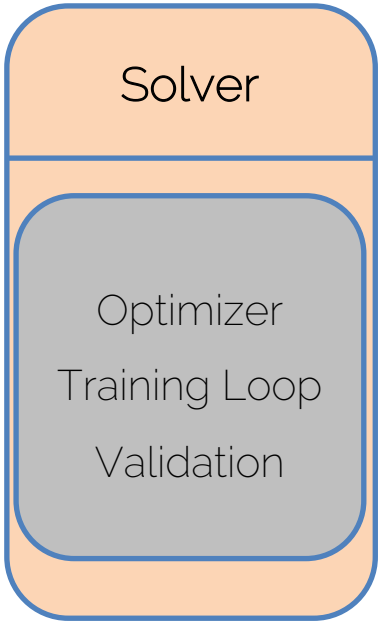


✓ Can be implemented once and used in multiple projects

Upcoming Lectures

- **Next lecture:**
  Lecture 4: Backpropagation

- **Next Thursday:**
  Exercise 4: Solver
  (and first network) with Franziska



See you next week ☺