Workshop: Building your first Convolutional Neural Network

Domen Rački, Andraž Mehle | Ljubljana, December 2018

Foreword

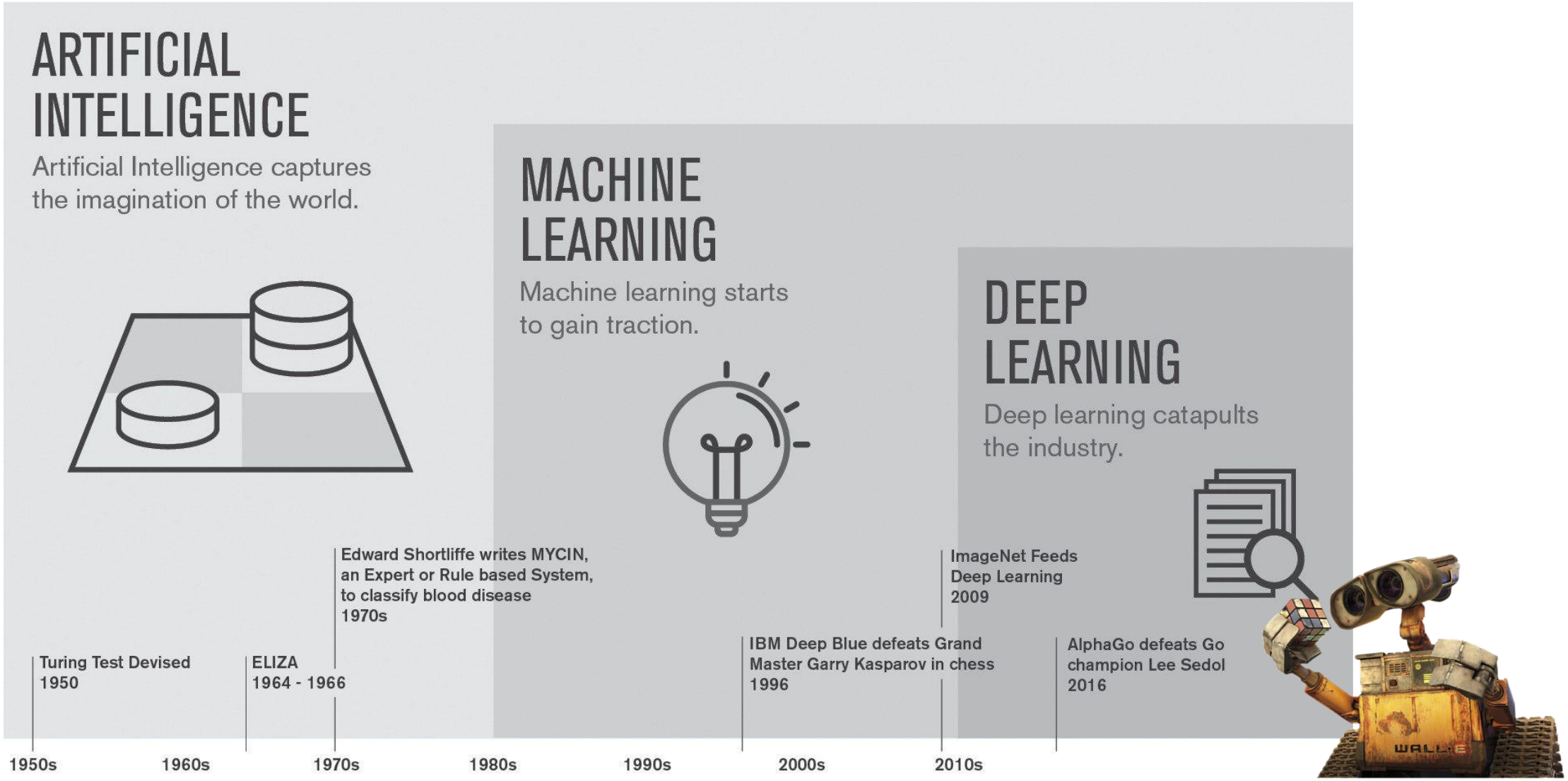
- The workshop is split into two parts
- Part 1: Intuitive deep-learning crash course
  - Domen Rački
- Part 2: Implementation example in Python with Keras and Tensorflow
  - Andraž Mehle
  - https://github.com/MehoMehoMeho/Fashion_MNIST_example

- The workshop's sponsor is
  Sensum, Computer Vision Systems

SENSUM

What is deep learning?

[Image credit: Kingman Tang, Apttus]



What is deep learning?

- Artificial Intelligence (AI)
  - Any algorithm or process that enables machines to mimic, develop and demonstrate human 'cognitive' functions such as 'learning' and 'problem solving'
  - You would need a lot of experience to devise some type of algorithm
- Machine Learning (ML)
  - Algorithms and statistical methods that give machines the ability to 'learn' from some form of data representation without being explicitly programmed
  - You would need formatted data and relevant statistical knowledge
- Deep Learning (DL)
  - Algorithms and methods based on learning data representations, opposed to task-specific algorithms
  - You will learn how to classify fashion objects in less than one hour

[Slide credit: Wikipedia]

Why is deep learning useful?

- Deep learning removes the manual identification of features in data
- It relies on a training process in order to discover useful patterns
- This makes training a neural network easier and faster
- Solves problems that were very hard to solve beforehand

- Deep learning is used for:
  - Image object detection and classification
  - Voice recognition and language translation
  - Text generation and image captioning
  - Autonomous vehicles and pedestrian detection
  - Coloring black-and-white images and videos

[Slide credit: Techopedia, Forbes]

Deep learning is a mapping

- A mapping from some space A to some space B
- Via some mapping function F (deep learning model)

Fashion images
(space A)

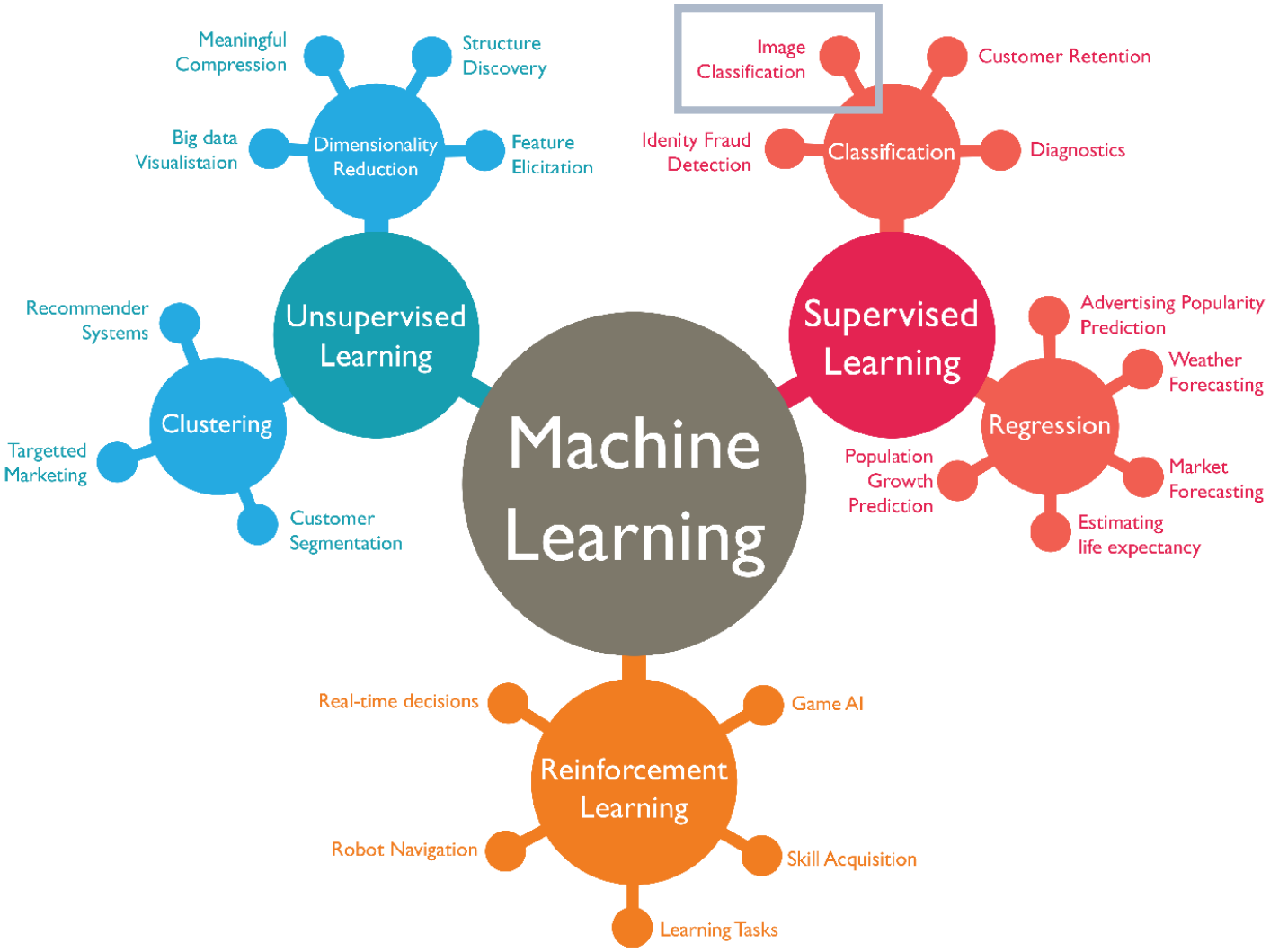
Deep learning model
(mapping function F)

Clothing category
(space B)

T-shirt/top
Trouser
Pullover
Dress
Coat
Sandal
Shirt
Sneaker
Bag
Ankle boot

[Han Xiao et. al., Fashion-MNIST]

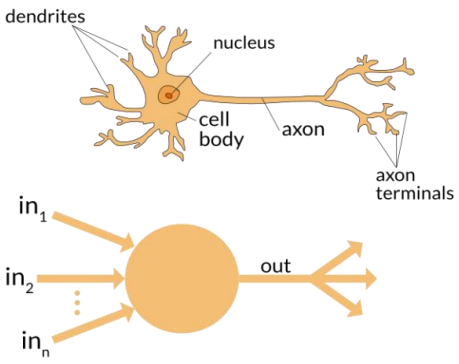
Different types of learning

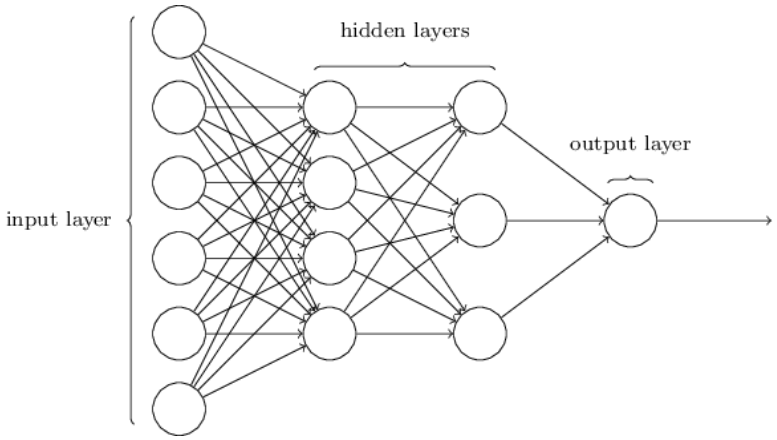


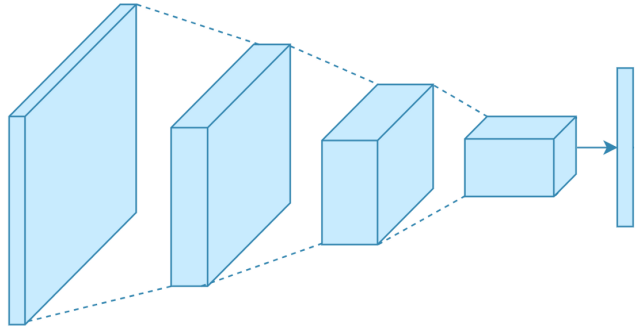
[Image credit: kisspng.com]

What is a Neural Network?

- Mathematical model vaguely inspired by biological neural networks in the brain
- Why is it called deep learning?
  - Because learning is done with a grater number of (hidden) layers
- What is a Convolutional Neural Network (CNN)?
  - A class of neural networks commonly applied to analysing images
  - More precisely, analysing image pixels via the convolution operation



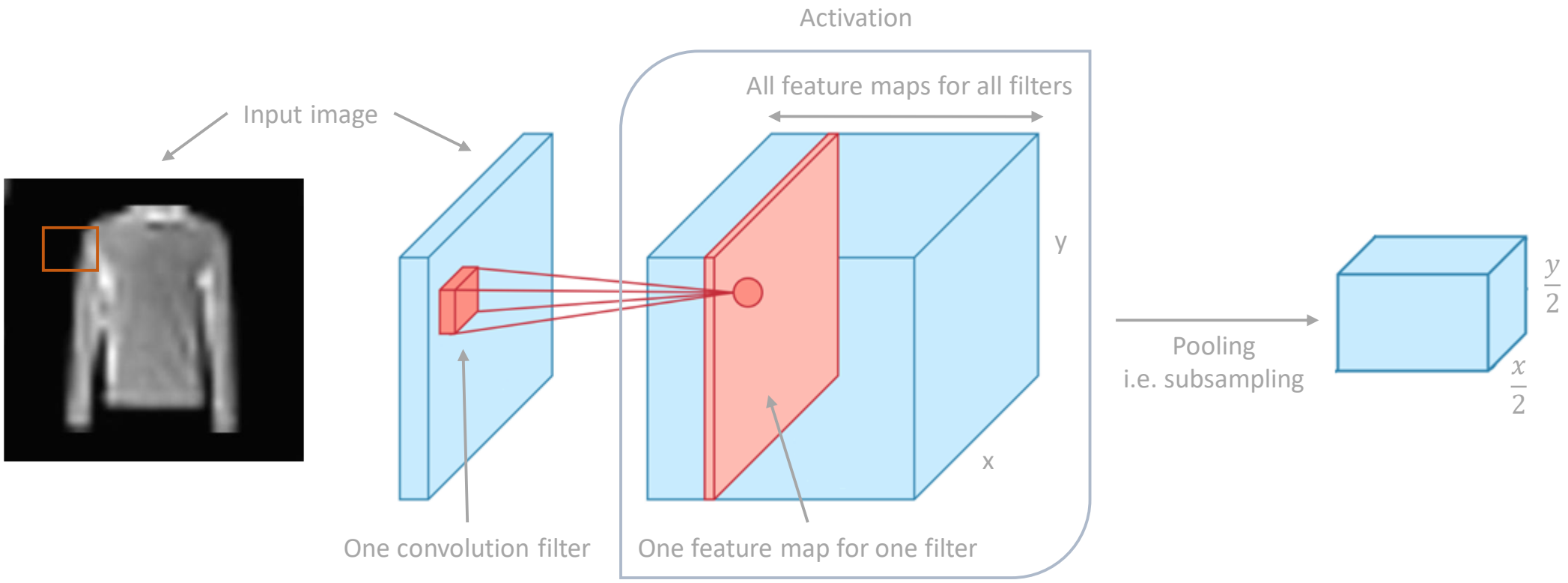




[Image credit: analyticsvidhya.com, neuralnetworksanddeeplearning.com, Applied Deep Learning]

How does a CNN work?

- Convolution, activation, pooling, repeat



[Image credit: A. Dertat, Applied Deep Learning]

How do I do deep learning?

- Step 0: Set up a DL environment
- Step 1: Split your data
- Step 2: Construct a model
- Step 3: Choose a loss function
- Step 4: Choose an optimization algorithm
- Step 5: Train and evaluate your network

Setting up a DL environment

- Step 1: Download and Install Anaconda 3.7
  - `https://www.anaconda.com/download/`
- Step 2: Set up Anaconda environment
  - Open console and create anaconda environment: `conda create -n tfenv python=3.5 anaconda`
- Step 3: Install Tensorflow
  - Activate anaconda environment: `activate tfenv`
  - Install tensorflow: `pip install --ignore-installed --upgrade https://storage.googleapis.com/tensorflow/windows/cpu/tensorflow-1.12.0-cp35-cp35m-win_amd64.whl`
- Step 4: Check Tensorflow version:
  - `ipython`
  - `import tensorflow as tf`
  - `tf.__version__`

Split your data

- Split your data into a training and testing set
  - To see if your algorithm really learned anything
- Make sure that:
  - Your training and test sets do not overlap
  - Your sets are large enough to yield meaningful results
  - Your sets are representative of your problem
- A less biased approach is k-fold cross validation

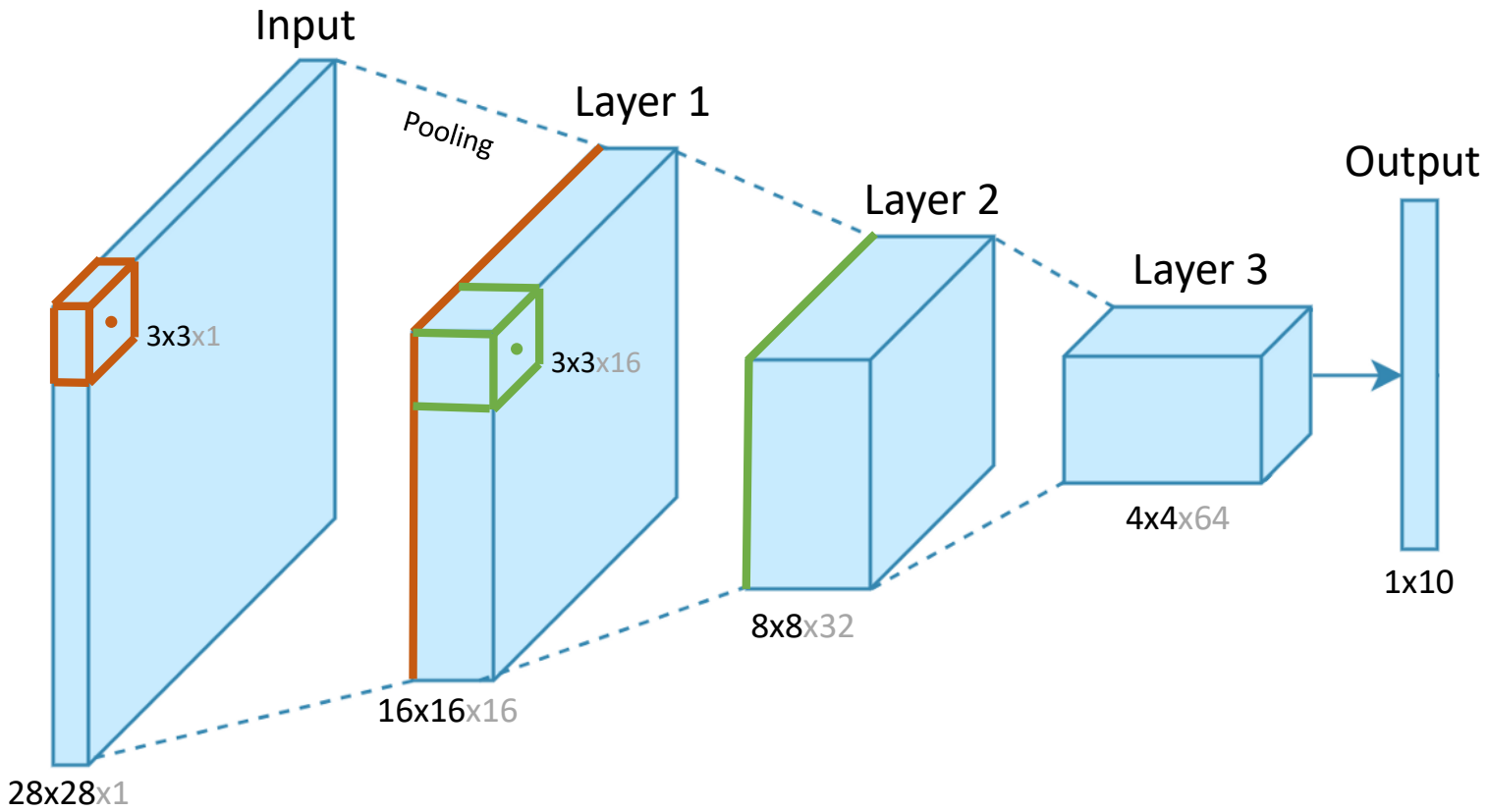
Training set

Testing set

[Image credit: DanB, kaggle.com]

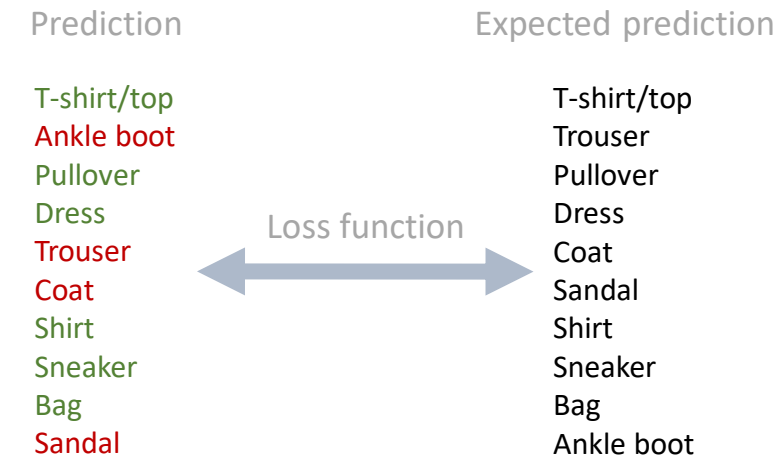
Construct a model

- My inputs are grayscale images of size 28x28x1 pixels
- My model will have 3 layers with 16, 32 and 64 filters
- My filters will be 3x3 pixels and I will subsample by a factor of 2 in each layer
- I want to classify my image into one of 10 categories



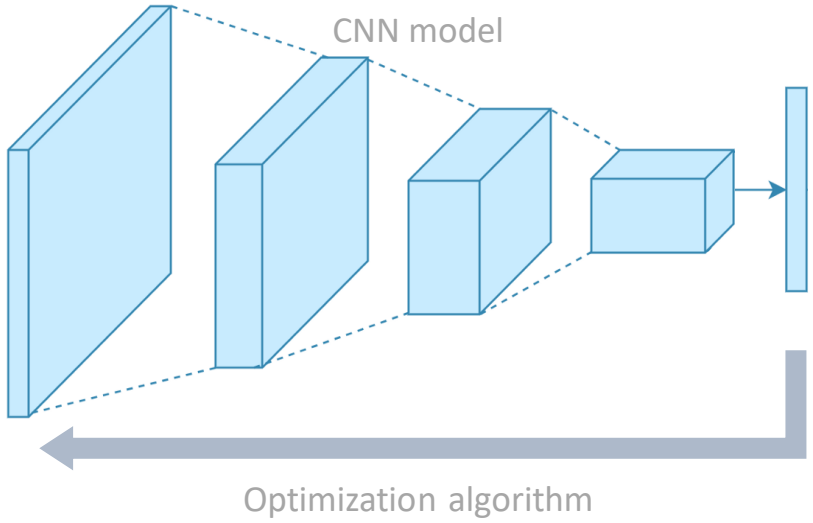
Choose a loss function

- What does a loss function do?
  - It shows you how 'wrong' your predictions are form your expected predictions
- Which loss function should I pick?
  - This depends on your specific problem
  - Generally the one that works best
- Classification problem
  - You want to predict a discrete category
  - Cross-entropy is preferred
- Regression problem
  - You want to predict a continuous number
  - Mean squared error is preferred



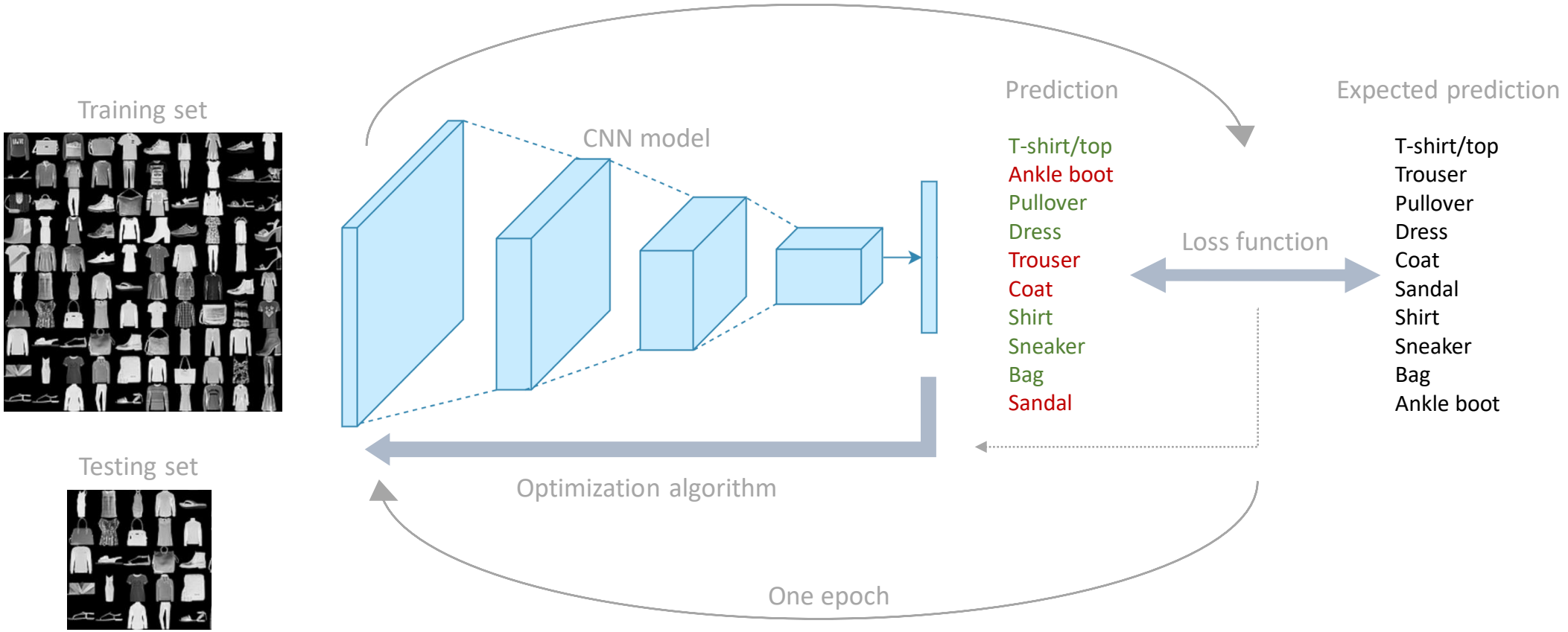
Choose an optimization algorithm

- What does an optimization algorithm do?
  - It 'adjusts' your model so that next time around the predicted accuracy improves
- Which optimization algorithm should I pick?
  - This depends on your specific problem
  - Generally the one that works best
  - Usually the Adam optimizer is used
- What does a learning rate do?
  - It controls how much your model is 'adjusted'
  - How do I choose a learning?
  - Generally by trial-and-error
  - Usually a learning rate around 0.001 is a good choice



Train and evaluate your network

- Show the training set to the network 100 times and show the testing set only once
- Train my network for 100 epochs on the training set and evaluate it on the test set



Takeaways

- Deep learning removes the manual identification of features in data
- Split your data to ensure unbiased model performance evaluation
- Choosing a suitable loss function is crucial
- Choosing a suitable optimization algorithm and learning rate is crucial

- Want more? Just google:
  - Stanford University CS231n
  - MIT Deep Learning Book