

Introduction to Deep Learning

Chapter 1: Introduction

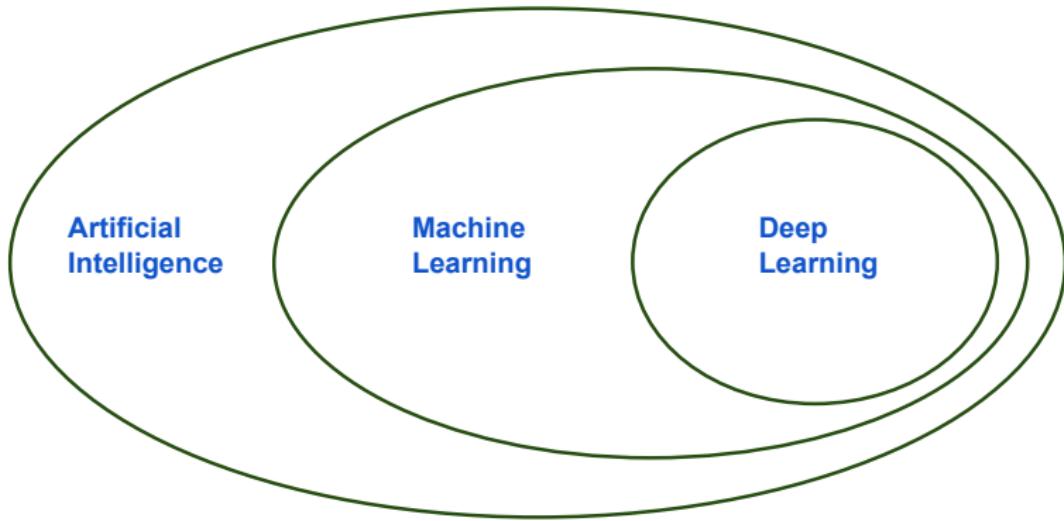
Bernd Bischl

Department of Statistics – LMU Munich

Winter term 2021



WHAT IS DEEP LEARNING



- Deep learning is the use of artificial neural networks to construct models on large amounts of (unstructured) data.

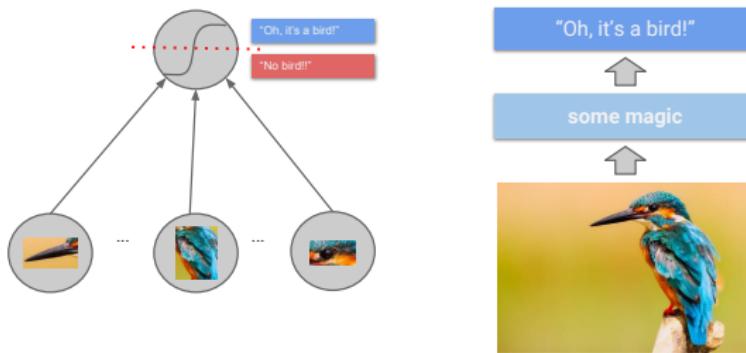
DEEP LEARNING AND NEURAL NETWORKS

- Deep learning and neural networks are mostly equivalent.
- Deep learning itself is not *new*:
 - Neural networks have been around since the 70s
 - *Deep* neural networks, i.e., networks with multiple hidden layers, are not much younger.
- Why everybody is talking about deep learning now:
 - ❶ Specialized, powerful hardware allows training of huge neural networks to push the state-of-the-art on difficult problems.
 - ❷ Large amount of data is available.
 - ❸ Special network architectures for image/text data.
 - ❹ Better optimization and regularization strategies.

IMAGE CLASSIFICATION WITH NEURAL NETWORKS

“Machine learning algorithms, inspired by the brain, based on learning multiple levels of representation/abstraction.”

Y. Bengio

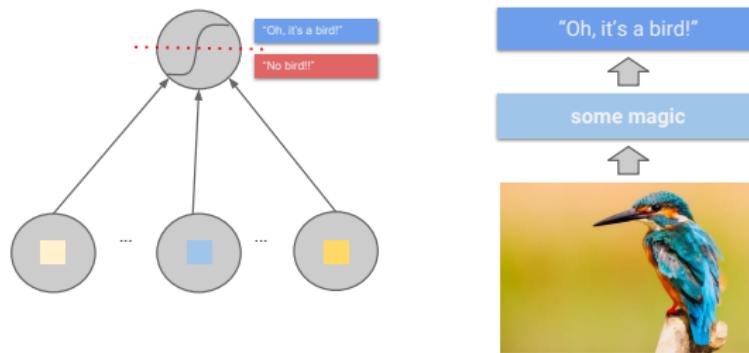


Caption 1

IMAGE CLASSIFICATION WITH NEURAL NETWORKS

“Machine learning algorithms, inspired by the brain, based on learning multiple levels of representation/abstraction.”

Y. Bengio

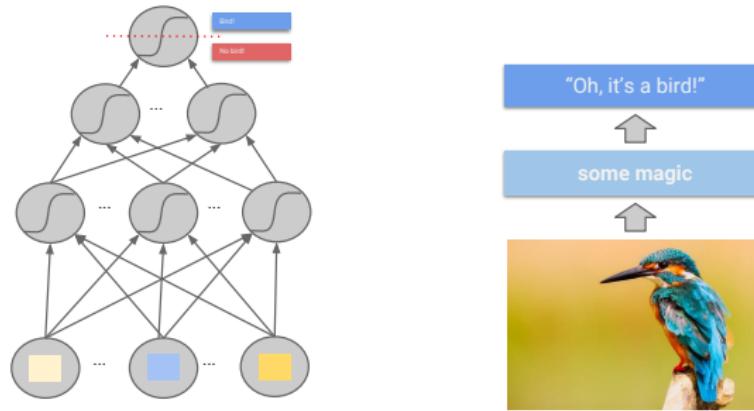


Caption 2

IMAGE CLASSIFICATION WITH NEURAL NETWORKS

“Machine learning algorithms, inspired by the brain, based on learning multiple levels of representation/abstraction.”

Y. Bengio

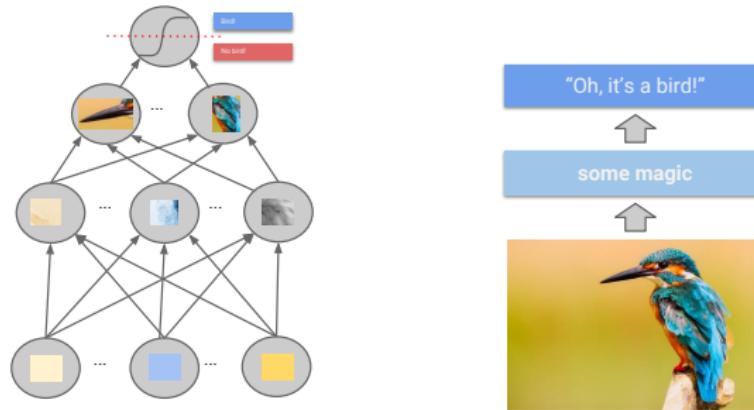


Caption 3

IMAGE CLASSIFICATION WITH NEURAL NETWORKS

“Machine learning algorithms, inspired by the brain, based on learning multiple levels of representation/abstraction.”

Y. Bengio



Caption 4

POSSIBLE USE-CASES

Deep learning can be extremely valuable if the data has these properties:

- It is high dimensional.
- Each single feature itself is not very informative but only a combination of them might be.
- There is a large amount of training data.

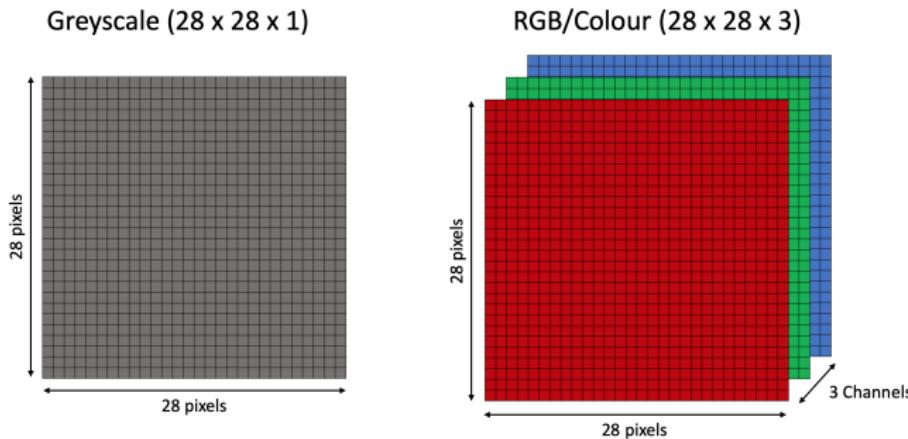
This implies that for tabular data, deep learning is almost never the correct model choice.

- Models like random forests or gradient boosting will outperform deep learning most of the time.
- One exception is data with categorical features with many levels.

POSSIBLE USE-CASE: IMAGES

- **High Dimensional:** A color image with 255×255 (3 Colors) pixels already has 195075 features.
- **Informative:** A single pixel is not meaningful in itself.
- **Training Data:** Depending on applications huge amounts of data are available.

Architecture: **Convolutional Neural Networks (CNN)**



POSSIBLE USE-CASE: IMAGES

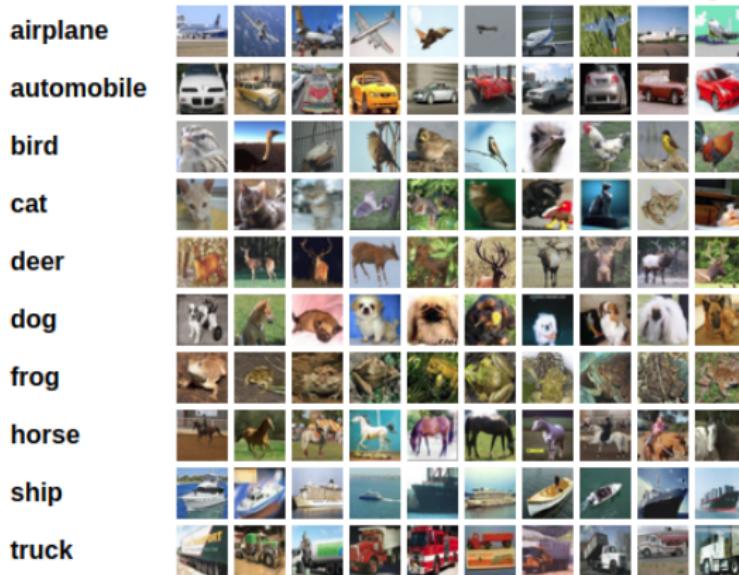
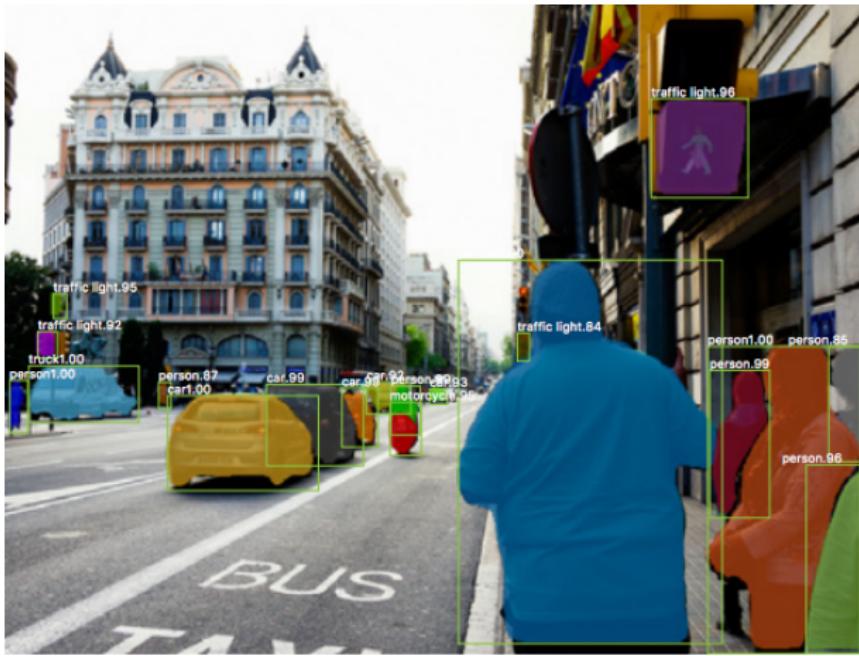


Image classification tries to predict a single label for each image (Alex Krizhevsky (2009)) CIFAR-10 is a well-known dataset used for image classification. It consists of 60,000 32x32 color images containing one of 10 object classes, with 6000 images per class.

POSSIBLE USE-CASE: IMAGES



Object Detection (Kaiming He (2017)) Mask R-CNN is a general framework for instance segmentation, which efficiently detects objects in an image while simultaneously generating a high-quality segmentation mask for each instance.

POSSIBLE USE-CASE: IMAGES

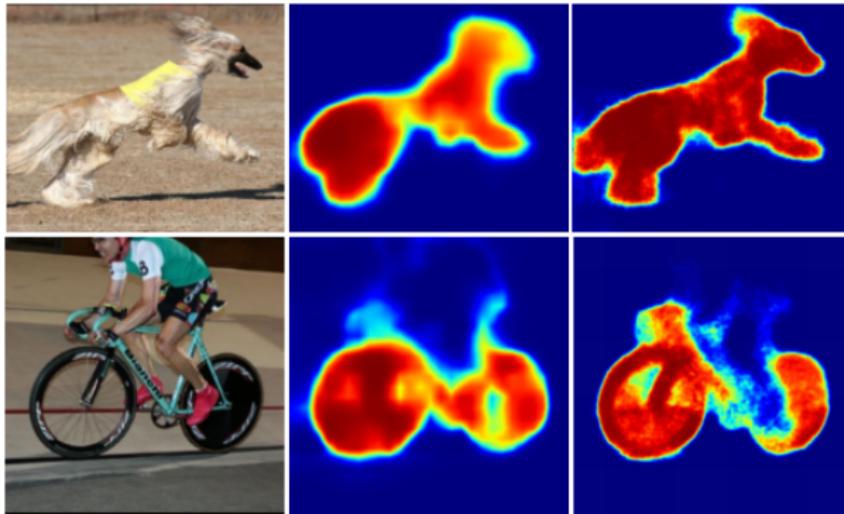
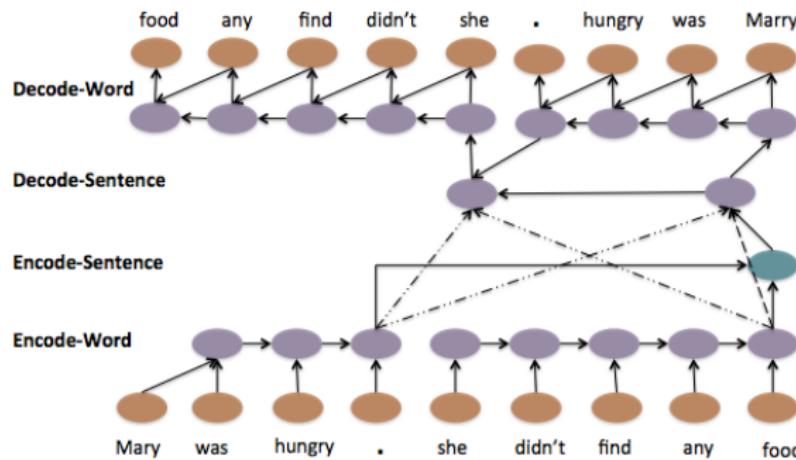


Image segmentation partitions the image into (multiple) segments
(Hyeonwoo Noh (2015))

POSSIBLE USE-CASE: TEXT

- **High Dimensional:** Each word can be a single feature (300000 words in the German language).
- **Informative:** A single word does not provide much context.
- **Training Data:** Huge amounts of text data available.

Architecture: Recurrent Neural Networks (RNN)



POSSIBLE USE-CASE: TEXT

Applications:

- Natural Language Processing, e.g.,
 - Sentiment Analysis
 - Email Classification
 - Chat-bots
 - ...
- Modeling Sequential Data (Time-Series, Speech)

POSSIBLE USE-CASE: TEXT

The image displays two separate translation interface boxes, likely from a tool like Google Translate. Each box has a source language dropdown, a target language dropdown, and a central text area with edit and audio controls.

Top Box (English to German):

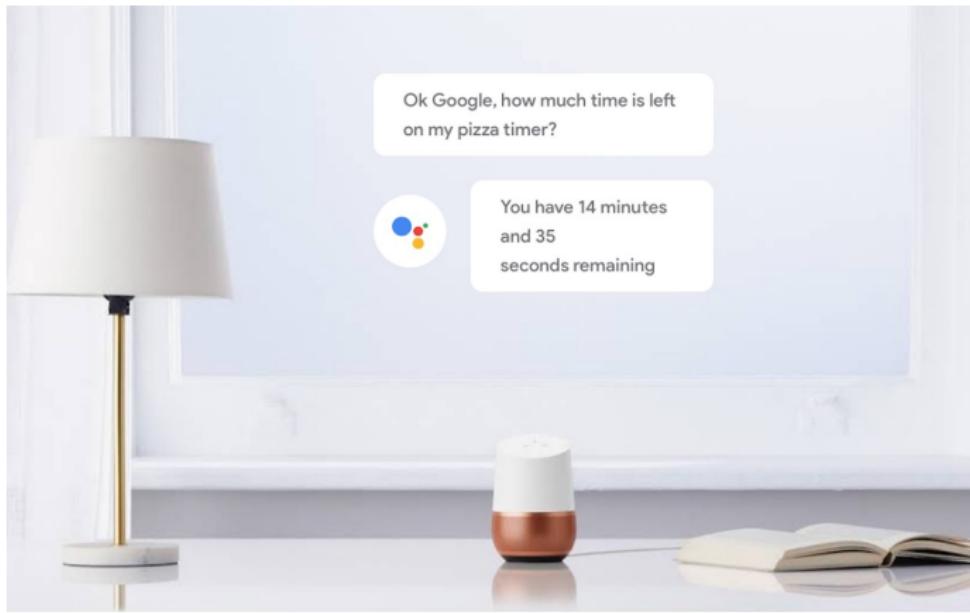
- Source: English – detected
- Target: German
- Text: He loves to eat Edit
- Output: Er liebt es zu essen

Bottom Box (Norwegian to English):

- Source: Norwegian
- Target: English
- Text: Butikken er stengt Edit
- Output: The store is closed

Machine Translation (e.g. google translate) Neural machine translation exploits neural networks to predict the likelihood of a sequence of words, typically modeling entire sentences in a single integrated model.

APPLICATIONS OF DEEP LEARNING: SPEECH



Speech Recognition and Generation (e.g. google assistant) Neural network extracts features from audio data in order to classify emotions in speech.