Artificial Intelligence and Deep Learning

NAE GAFOE Session 2019
AI & Deep Learning Session Chairs

Sujith Ravi
Google Research
http://www.sravi.org/

Stefan Uhlich
Sony Corporate Research
https://uhlich.updog.co/
Artificial Intelligence (AI)

Human Brain

- Language
- Vision
- Speech

Machine Learning
Artificial Intelligence (AI)

Human Brain

- Language
- Vision
- Speech

Machine Learning

- Natural Language Processing
- Computer Vision
- Automatic Speech Processing
Deep Learning
Deep Learning
Neural Networks are good *function approximators*
Neural Networks are good function approximators
Neural Networks that see & understand images

Surpasses human-level performance for recognizing thousands of object categories (ILSVRC 2014 task)
Neural Networks that see & understand images

Image → Concept

Dessert
Neural Networks that *hear* & *understand* speech

Acoustic input

Text output

Words / Sentences

Reduced word errors in speech recognition by >30%
Neural Networks that *read & understand text*

German text

Ich liebe Hamburg

English translation

I love Hamburg

*Neural machine translation achieves state-of-the-art performance*
And applications beyond...

Health

Self-driving / Automotive

Games

Why is Deep Learning Effective?

- Data
- Computing
- Models
- Tools

Large
Fast
Powerful
Easy-to-use
Why now?

- Simple neurons/networks have been around since ~1950s
- Why is deep learning trending now?
Why now?

- Simple neurons/networks have been around since ~1950s
- Why is deep learning trending now?

Why now?

- **Massive parallel processing on GPUs/TPUs**
  - Allows us to train large networks on large datasets

![Graph showing performance normalized to CPU](https://www.nvidia.com/de-de/data-center/tesla-v100/)

**Workload:** ResNet-50  |  **CPU:** 1X Xeon E5-2690v4 @ 2.6 GHz  |  **GPU:** Add 1X Tesla P100 or V100

**Source:** https://www.nvidia.com/de-de/data-center/tesla-v100/

GPU ... Graphics processing unit  
TPU ... Tensor processing unit
Why now?

- **Access to huge datasets**
  - Image, Video, Audio, Text/NLP, ...

1.2M images, 1K classes

5.8 thousand hours of audio, 527 classes

350K hours (= 40 years), ~4K classes

[YouTube-100M: 600 years of video - more than one human will see in his lifetime]
Why now?

- **Access to huge datasets**
  - Deep neural networks (DNNs) nicely scale with more data
Why now?

- Open-source software + ArXiv/ArXiv sanity
  - Open-source DNN toolboxes available: Tensorflow, PyTorch, …
  - Many papers come with source code on Github
  - Timely access to latest results

→ Fast progress
So - everything is solved?

No - there are still many open problems that need to be tackled ...

- Fairness in machine learning (avoid to learn biases from datasets)
- Robustness against adversarial samples/adversarial users
- Explanations for decisions from a deep neural network/ML algorithm
- Learning from noisy labels
- (General) AI still not available
- ...

→ Session presentations
Session Program

Software 2.0: Machine Learning is Changing Software
Chris Re, Stanford University

Teaching Machines to Communicate with People using Deep Learning
Zornitsa Kozareva, Google

Deep Learning for Visual and Virtual Worlds
Eleonora Vig, German Aerospace Center (DLR)

Artificial Intelligence in Cognitive Neural Engineering
Moritz Grosse-Wentrup, University of Vienna