

8 Machine Learning III

- Specialized Areas in Machine Learning

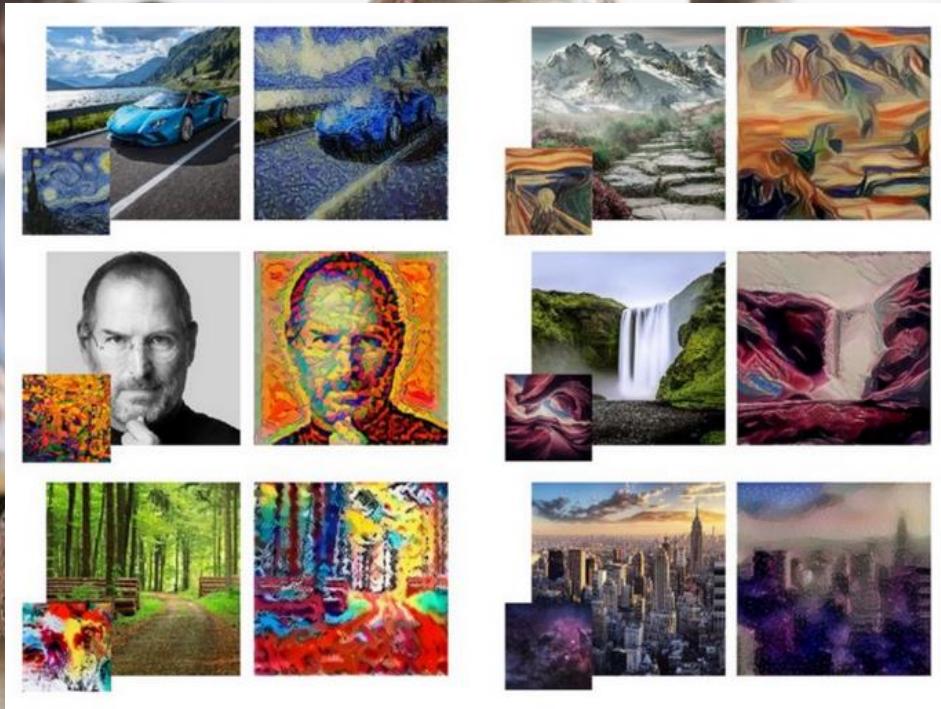
Content:

1. Transfer Learning & Teachable Machine
2. YOLO & Real-Time Object Detection
3. Autoencoder & Super Sampling
4. Generative Adversarial Networks (GAN)
5. Reinforcement Learning
6. The Human Role in Machine Learning **Live**
7. Summary



Which machine learning technique is used in this case (and in Deepfakes)?

Schwierigkeitsgrad	Art des Wissens Abfragewissen (Vorlesung)	Anwendungswissen (Literatur)
Einfach	Green	Yellow
Mittel	Yellow	Red
Schwierig	Red	Red



- a) Convolutional Neural Networks
- b) Transfer Learning
- c) Generative Adversarial Networks
- d) Reinforcement Learning
- e) Transformer

8 Machine Learning III

- Specialized Areas in Machine Learning

(1) Transfer Learning

Transfer Learning: When to apply

When to apply Transfer Learning:

- Data: When there is **huge amount of similar Data** in another Domain, but **less Data** in current Domain.
- Process: Have **enough Data** for Training, but don't have enough **Computational Resources** to train it.

Deep Learning - using ResNets for Transfer Learning, 2019
madhuramiah.medium.com/deep-learning-using-resnets-for-transfer-learning-d7f4799fa863

Transfer Learning: Benefits

Benefits from Transfer Learning:

- Related to **Data**: Train with **smaller amount of Data**.
- Related to **Process**: **Faster training progress** (higher start, slope, asymptote).

A Gentle Introduction to Transfer Learning for Deep Learning
machinelearningmastery.com/transfer-learning-for-deep-learning/

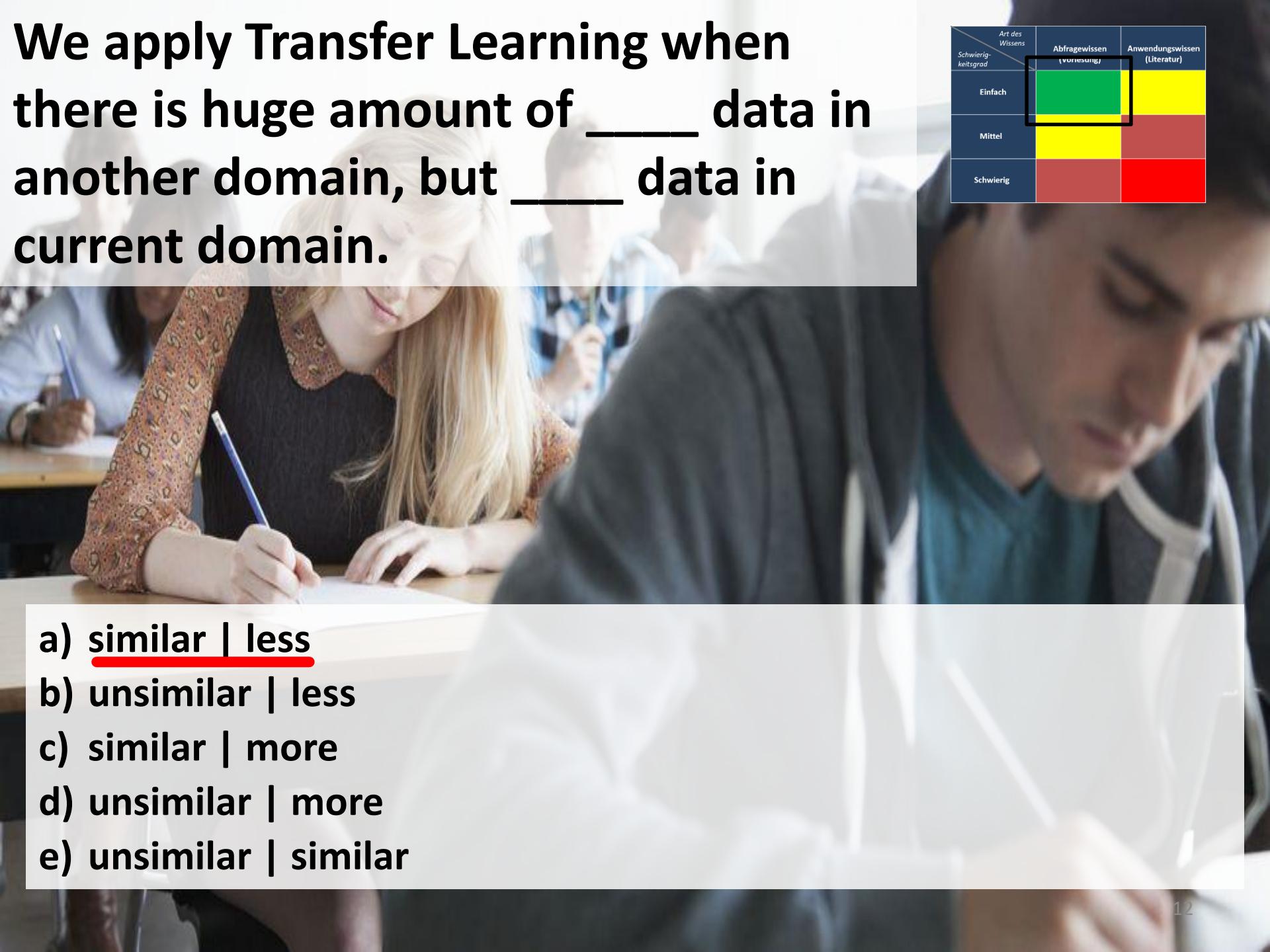
Which statement about Transfer Learning is FALSE?

Schwierigkeitsgrad	Art des Wissens	Abfragewissen (Vorlesung)	Anwendungswissen (Literatur)
Einfach		Green	Yellow
Mittel		Yellow	Red
Schwierig		Red	Red

- a) Learning of a new **task** relies on the previous learned tasks.
- b) Learning reuses abstract knowledge, retrains specific knowledge.
- c) Learning is initialized from pre-trained weights and biases.
- d) Learning process can be faster at the cost of more training data.
- e) It has the potential to significantly improve the sample efficiency of a reinforcement learning agent.

We apply Transfer Learning when there is huge amount of _____ data in another domain, but _____ data in current domain.

		Art des Wissens	Abfragewissen (Vorlesung)	Anwendungswissen (Literatur)	
		Schwierigkeitsgrad	Einfach	Mittel	Schwierig
Schwierigkeitsgrad	Einfach				
	Mittel				
Schwierig					

- 
- A blurred background image shows several students in a classroom setting, focused on writing in their notebooks. The lighting is soft, creating a professional and academic atmosphere.
- a) similar | less
 - b) unsimilar | less
 - c) similar | more
 - d) unsimilar | more
 - e) unsimilar | similar

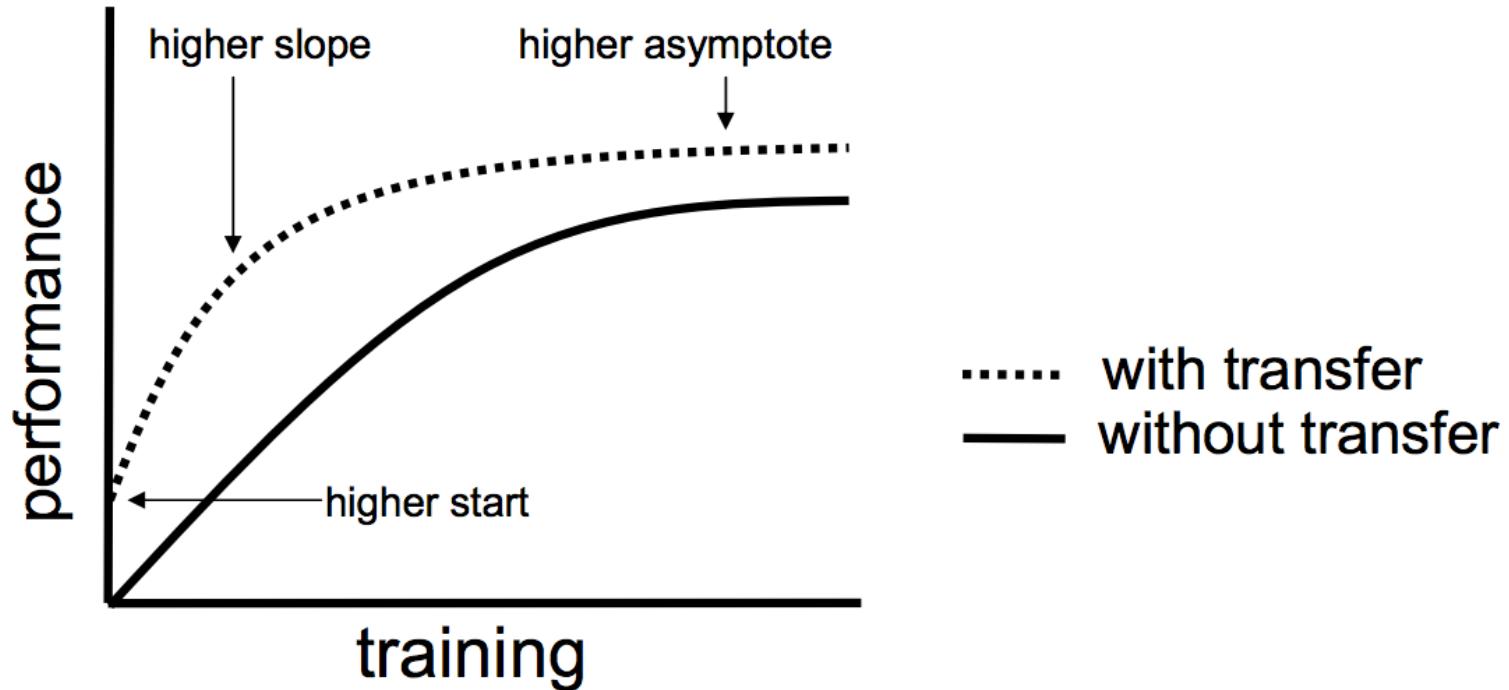
We apply Transfer Learning when
there are enough _____ for training
but do not have enough _____ to train
it.

		Art des Wissens	Abfragewissen (Vorlesung)	Anwendungswissen (Literatur)	
		Schwierigkeitsgrad	Einfach	Mittel	Schwierig
Schwierigkeitsgrad	Einfach				
	Mittel				
	Schwierig				



- a) Domain | Data
- b) Data | Domain
- c) Process | Computational Resources
- d) Data | Computational Resources
- e) Process | Data

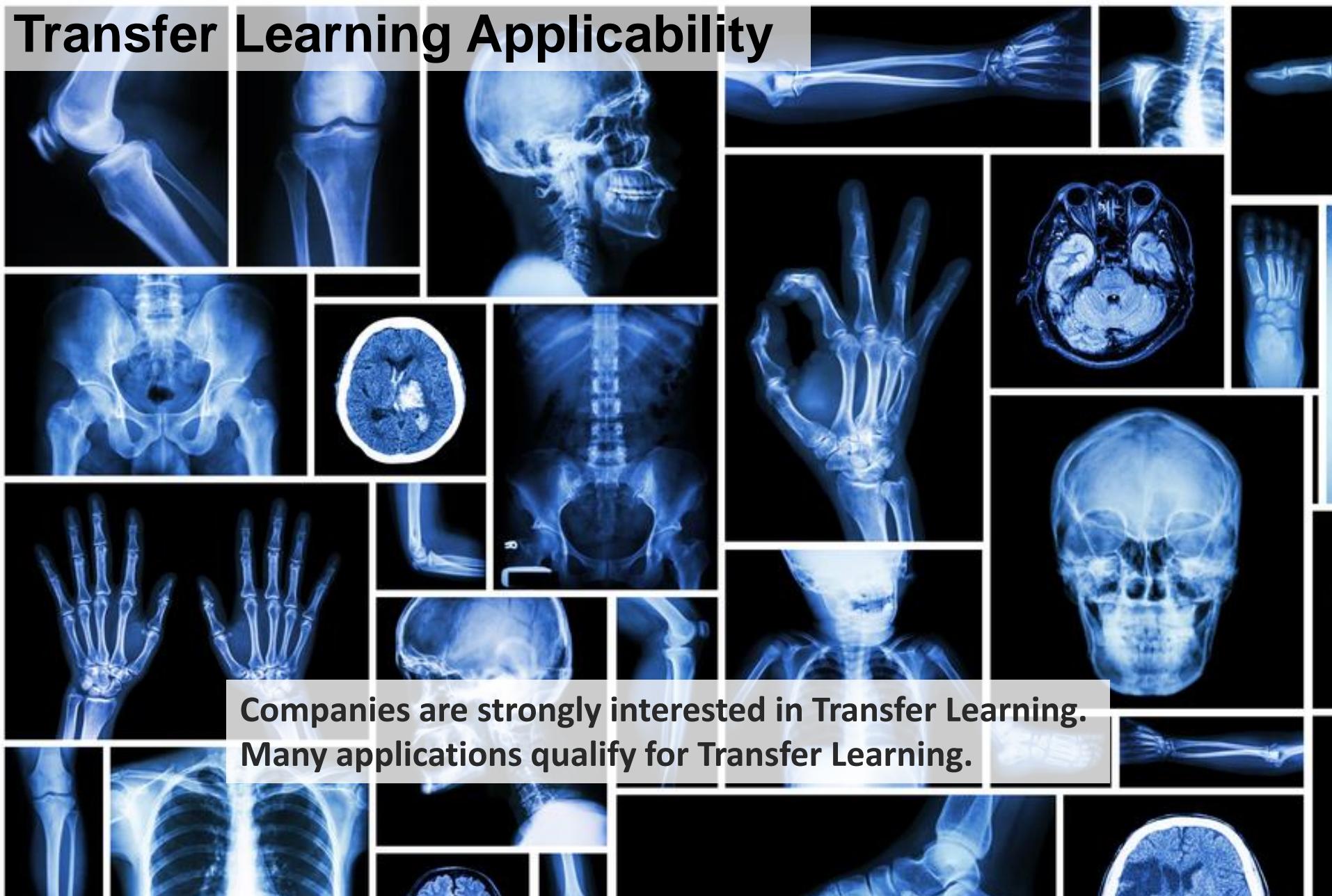
Transfer Learning: Benefits



Benefits from Transfer Learning:

- Related to **Data**: Train with **smaller amount of Data**.
- Related to **Process**: **Faster training progress** (higher start, slope, asymptote).

Transfer Learning Applicability

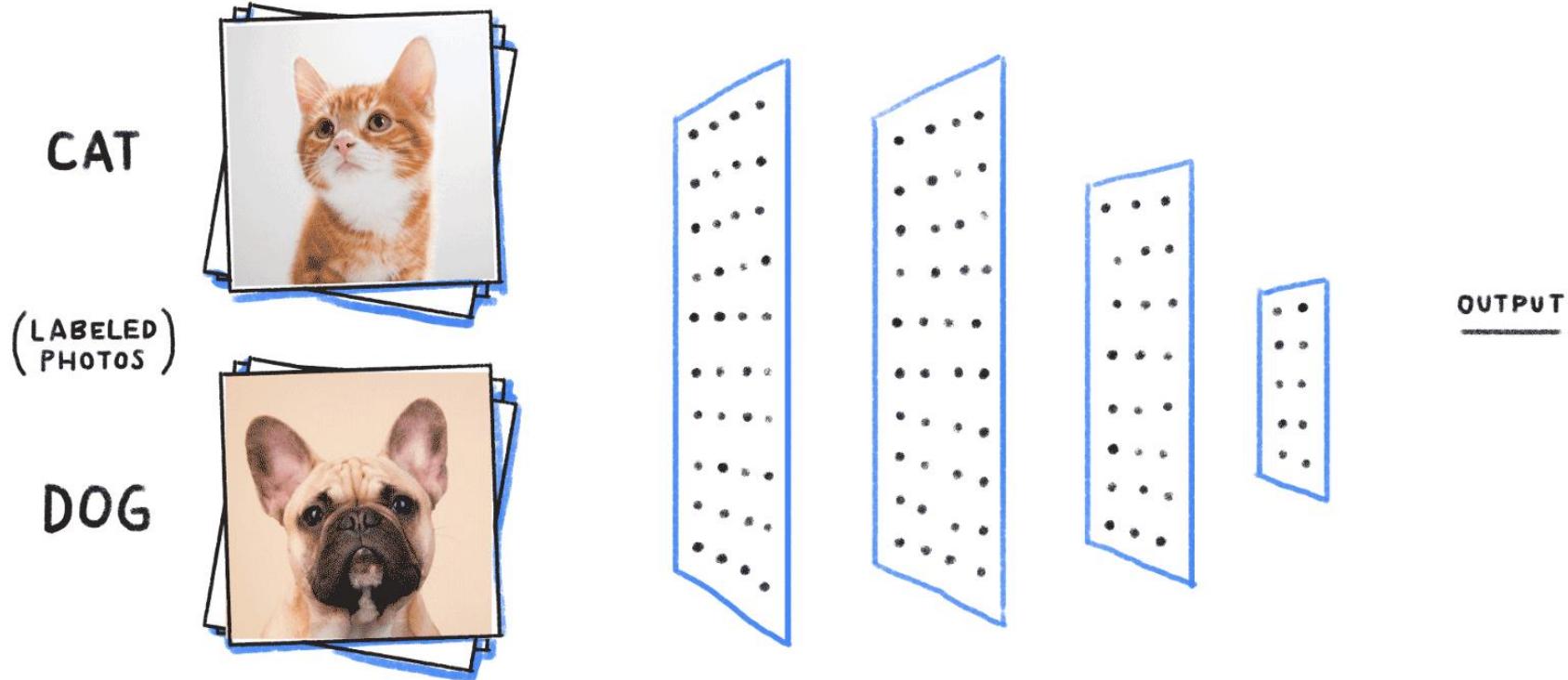


Companies are strongly interested in Transfer Learning.
Many applications qualify for Transfer Learning.

Understanding Transfer Learning For Medical Applications

analyticsindiamag.com/understanding-transfer-learning-for-medical-applications...

Transfer Learning Playground



Michael Amberg

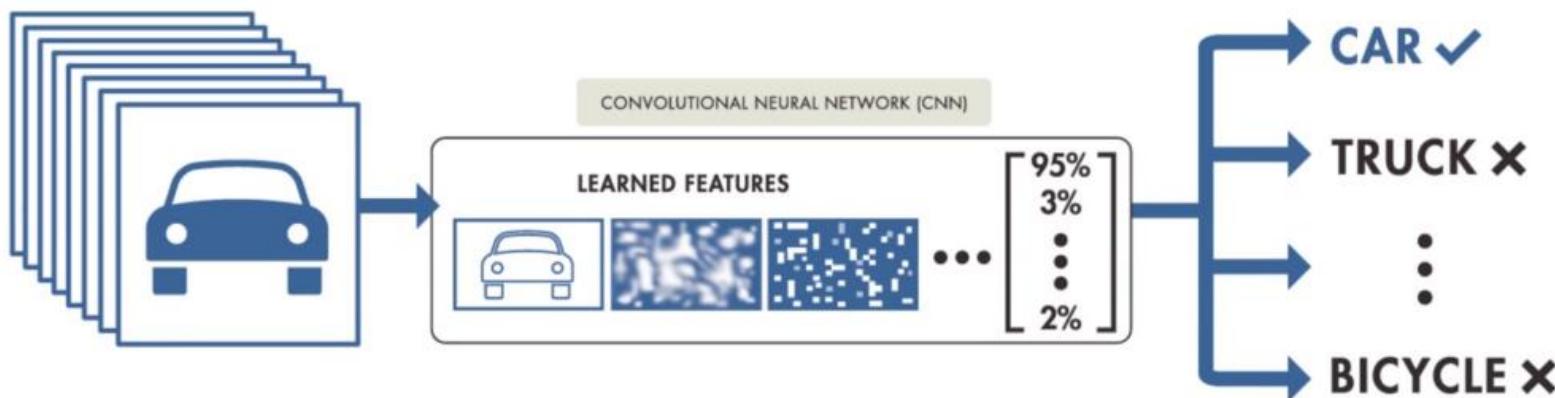
Todays Content:

1. Transfer Learning & Teachable Machine
2. YOLO & Real-Time Object Detection
3. Autoencoder & Super Sampling
4. Generative Adversarial Networks (GAN)
5. Reinforcement Learning
6. The Human Role in Machine Learning **Live**
7. Summary

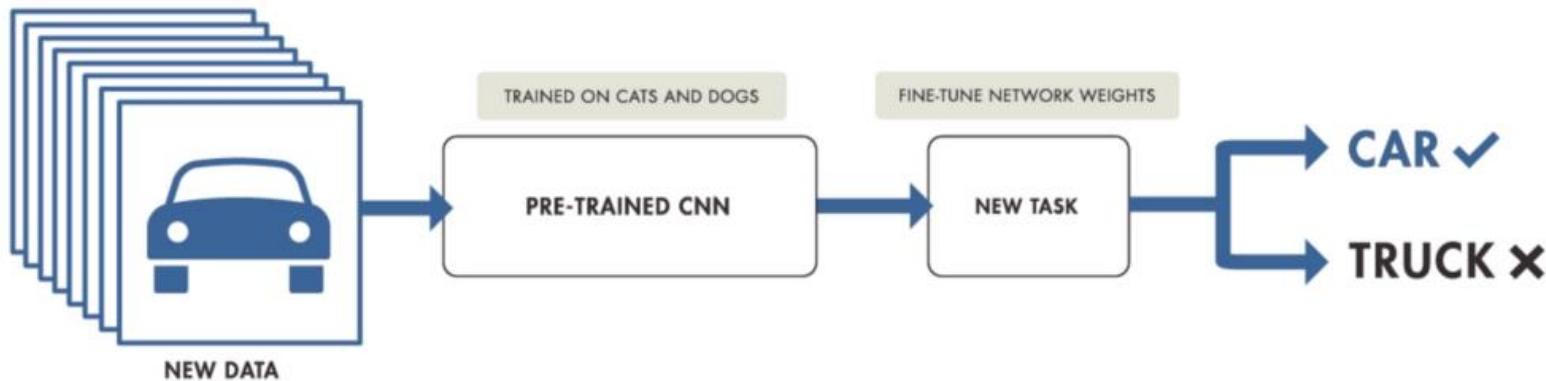


1. Transfer Learning & Teachable Machine

TRAINING FROM SCRATCH



TRANSFER LEARNING



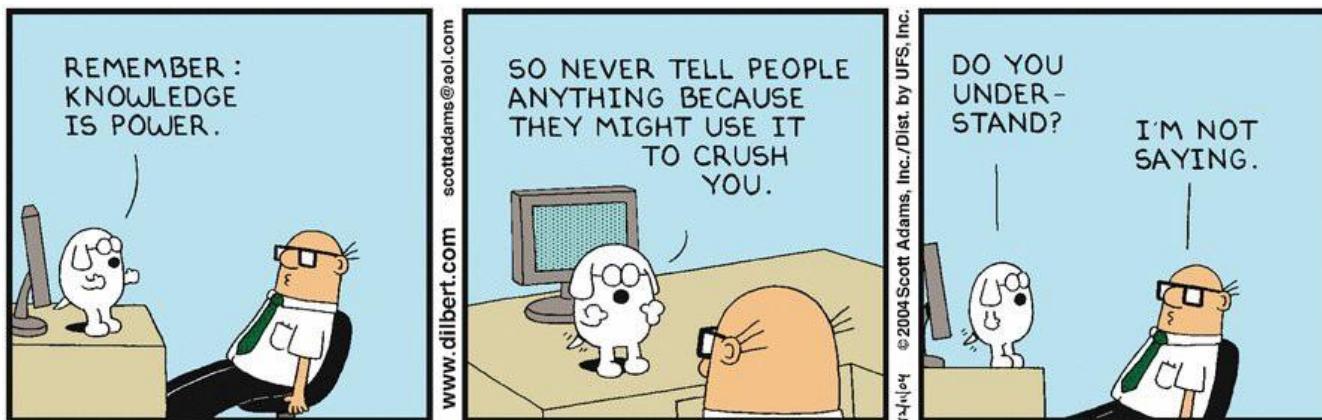
Transfer Learning

Transfer learning

From Wikipedia, the free encyclopedia

Transfer learning (TL) is a research problem in machine learning (ML) that focuses on storing knowledge gained while solving one problem and applying it to a different but related problem.^[1] For example, knowledge gained while learning to recognize cars could apply when trying to recognize trucks.

This area of research bears some relation to the long history of psychological literature on [transfer of learning](#), although formal ties between the two fields are limited. From the practical standpoint, reusing or transferring information from previously learned tasks for the learning of new tasks has the potential to significantly improve the sample efficiency of a [reinforcement learning](#) agent.^[2]



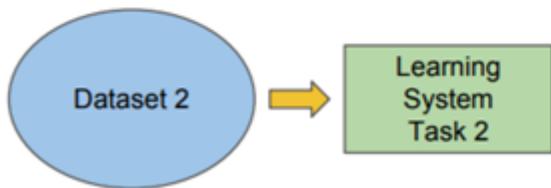
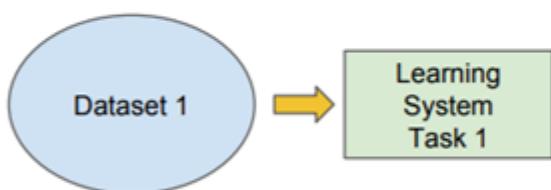
Transfer Learning vs Traditional Machine Learning

Traditional ML

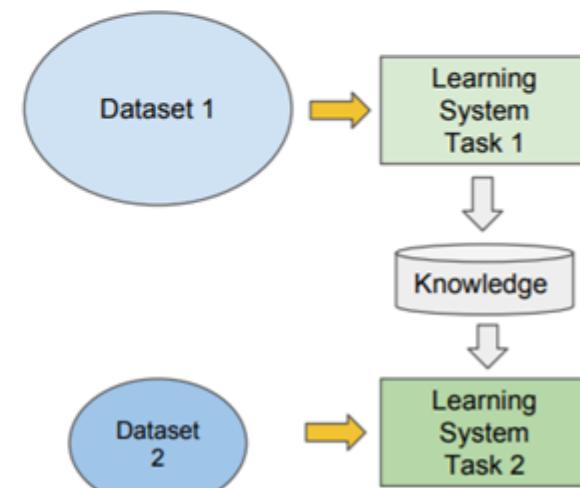
vs

Transfer Learning

- Isolated, single task learning:
 - Knowledge is not retained or accumulated. Learning is performed w.o. considering past learned knowledge in other tasks



- Learning of a new tasks relies on the previous learned tasks:
 - Learning process can be faster, more accurate and/or need less training data



Transfer Learning with CNNs (Convolutional NN)

OpenAI

OpenAI Microscope

We're introducing a collection of visualizations for layer and neuron of "neuron organisms" which increase interpretability. Microscope analyzes the features learned by neural networks, and is a research community for understanding them.

April 14, 2020
2 minute read

Mit OpenAI Microscope sieht man, dass frühe Schichten eines Modells abstrakte, späte Schichten konkrete Muster lernen. Kernelement beim Transfer Learning: Abstraktes Wissen weiterverwenden, spezifisches Wissen nachtrainieren.

Microscope

MODELS ABOUT

Models

AlexNet

The OpenAI Microscope is a collection of visualizations for significant neurons in neural vision models.

Inception v1

Inception v1 (Places)

VGG 19

Inception v3

Inception v4

ResNet v2 50

Inception v1

mixed4c

Unit 447

FEATURE VISUALIZATION

An artificial, optimized image that maximizes activations of the given unit. [Read more](#).

CHANNEL OPTIMIZATION objective results in a repeating pattern.

NEURON OPTIMIZATION objective shows spatial preferences.

DATASET SAMPLES

Pieces of images from the training dataset that result in the largest activations from the given unit.

These images are cropped and downsize samples from the [ImageNet](#) research dataset. Unlike our other visualizations, they are not CC-BY-SA because they are derived from ImageNet.

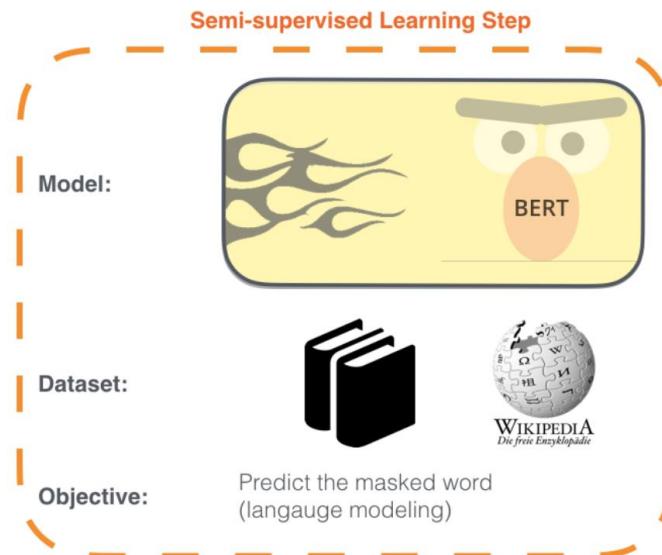
OpenAI Microscope

microscope.openai.com/models

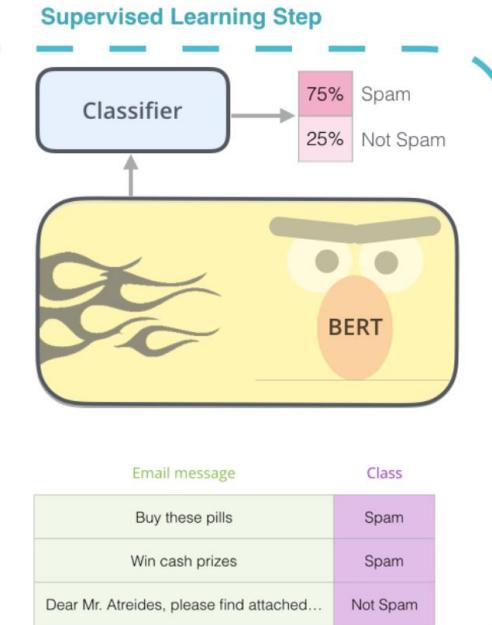
Transfer Learning with BERT Transformer

1 - **Semi-supervised** training on large amounts of text (books, wikipedia..etc).

The model is trained on a certain task that enables it to grasp patterns in language. By the end of the training process, BERT has language-processing abilities capable of empowering many models we later need to build and train in a supervised way.



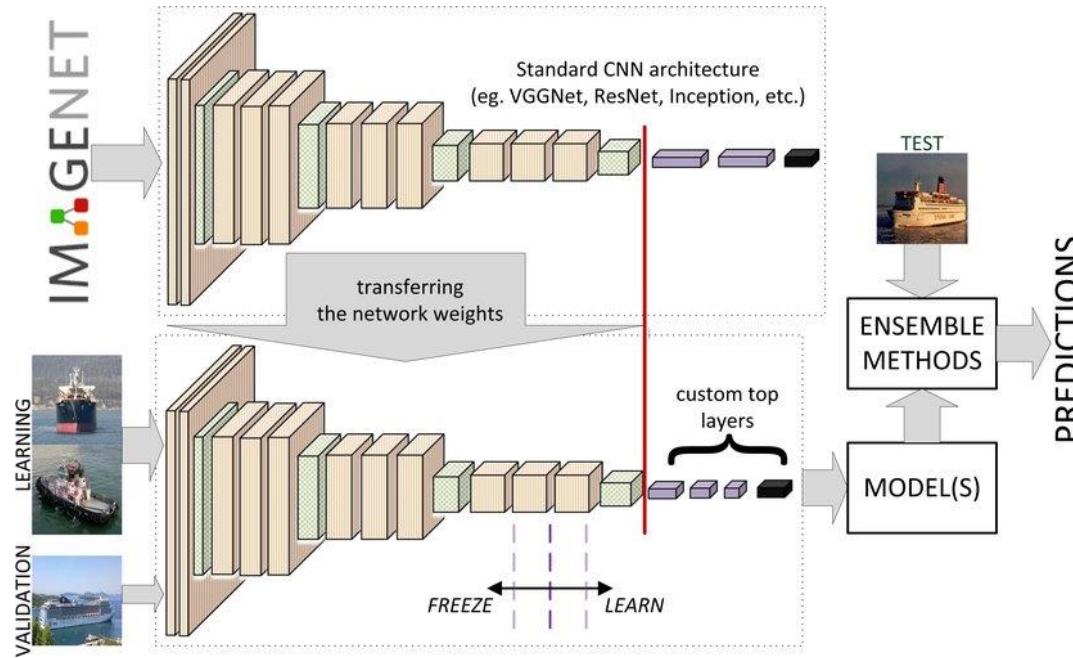
2 - **Supervised** training on a specific task with a labeled dataset.



The two steps of how BERT is developed. You can download the model pre-trained in step 1 (trained on un-annotated data), and only worry about fine-tuning it for step 2. [Source for book icon].

Eine Form von **Transfer Learning**:
Fokus auf **Fine-Tuning, Pre-Training** übernehmen

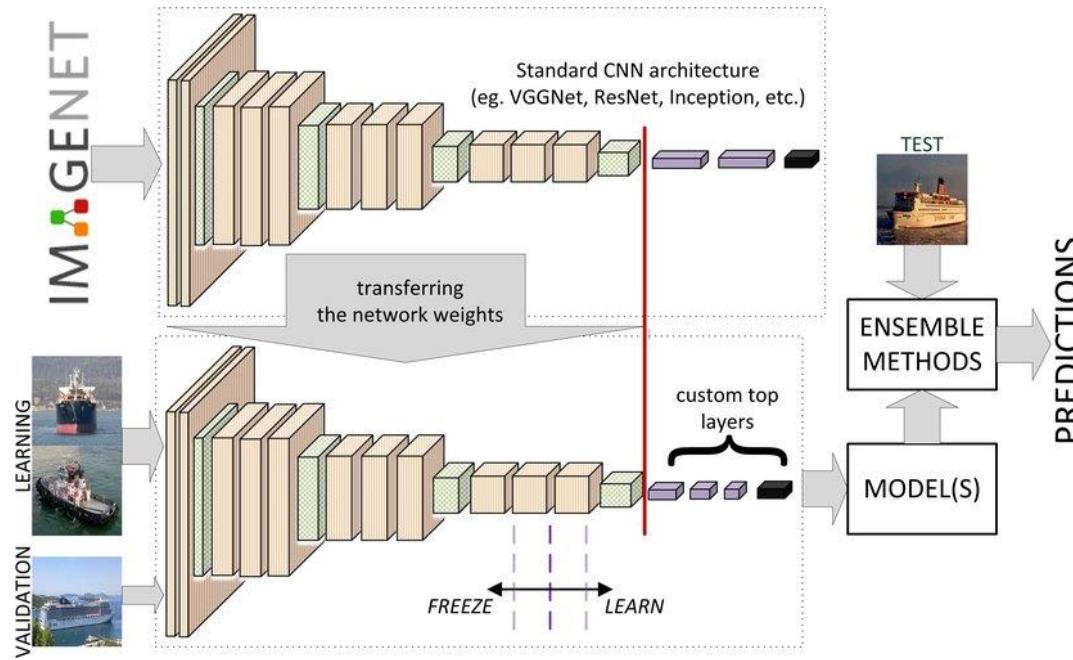
Transfer Learning: When to apply



When to apply Transfer Learning:

- **Data:** When there is **huge amount of similar Data** in another Domain, but **less Data in current Domain**.
- **Process:** Have **enough Data** for Training, but don't have enough **Computational Resources** to train it.

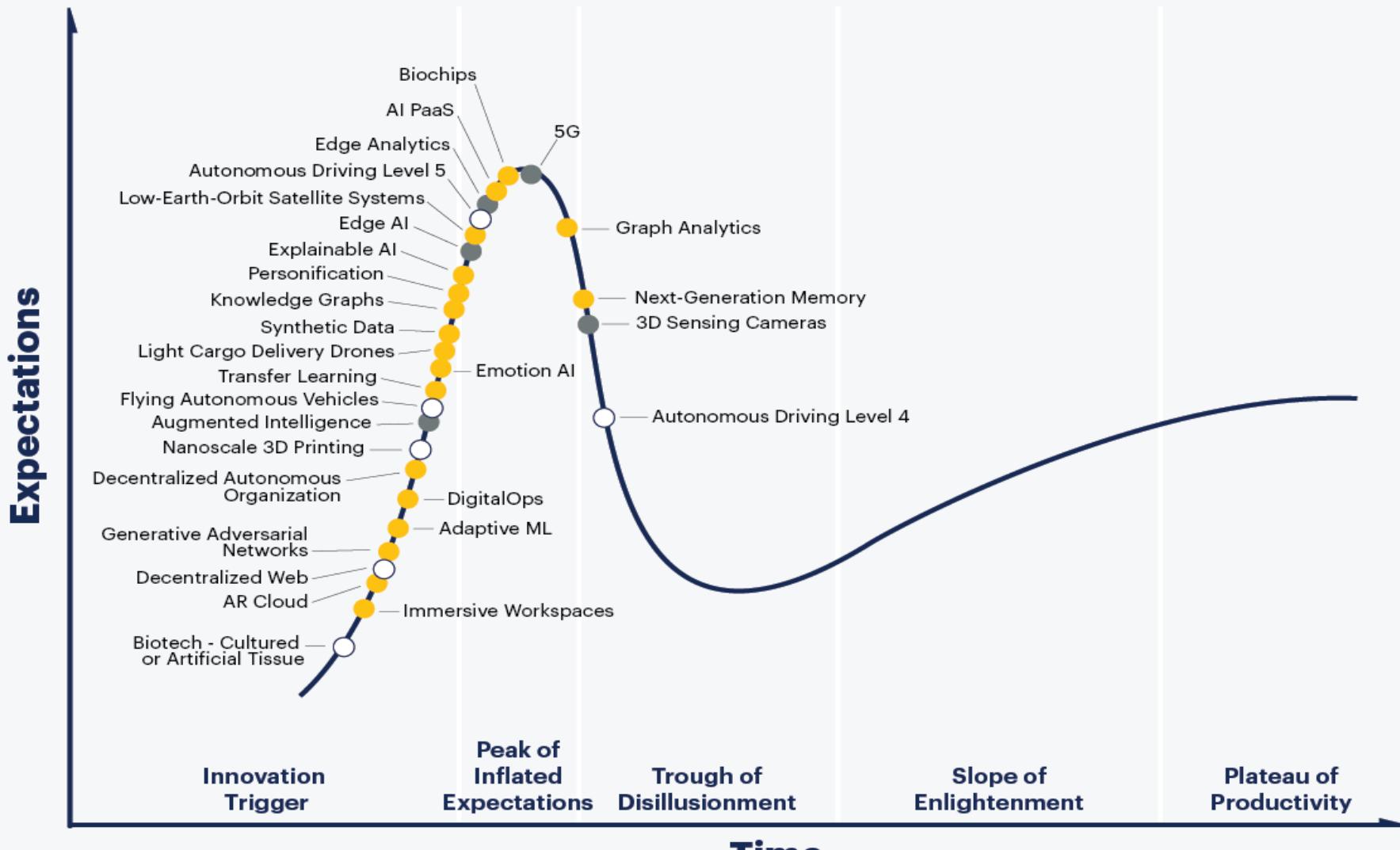
Transfer Learning: Options when applying



2 Options When applying Transfer Learning in Neural Networks:

- **Freeze the Weights and Bias on Initial Few Layers** and train only the **Last Few Layers** and **Fully Connected**.
In this case you **don't** need to **re-train** the **whole Network Model** again.
- **Re-train the whole Network**, initializing from the **learned Weights and Bias**.
Keep the Learning Rate very low so that the **original Weights** don't deviate drastically.

Gartner Hype Cycle for Emerging Technologies, 2019



Plateau will be reached:

Less than 2 years

2 to 5 years

5 to 10 years

More than 10 years

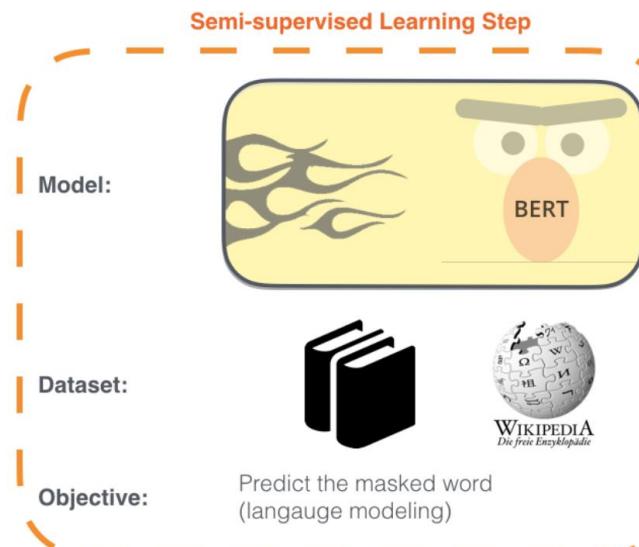
Obsolete before plateau

As of August 2019

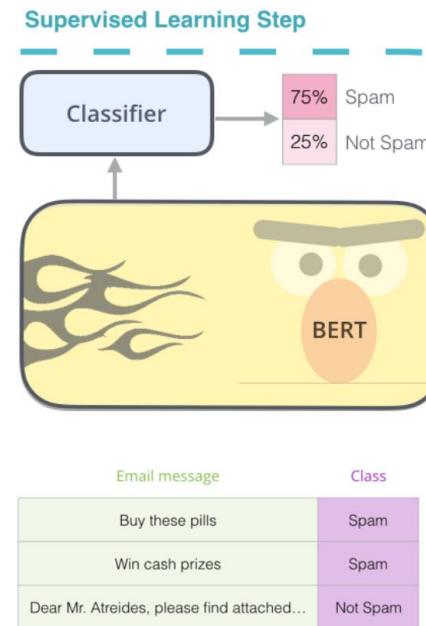
BERT Transformer, Google 2018

1 - Semi-supervised training on large amounts of text (books, wikipedia..etc).

The model is trained on a certain task that enables it to grasp patterns in language. By the end of the training process, BERT has language-processing abilities capable of empowering many models we later need to build and train in a supervised way.



2 - Supervised training on a specific task with a labeled dataset.



The two steps of how BERT is developed. You can download the model pre-trained in step 1 (trained on un-annotated data), and only worry about fine-tuning it for step 2. [Source for book icon].

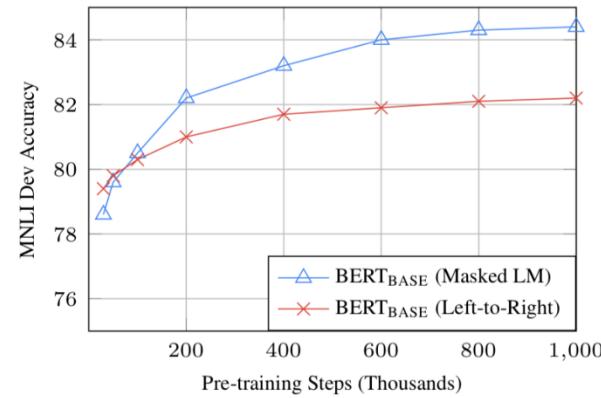
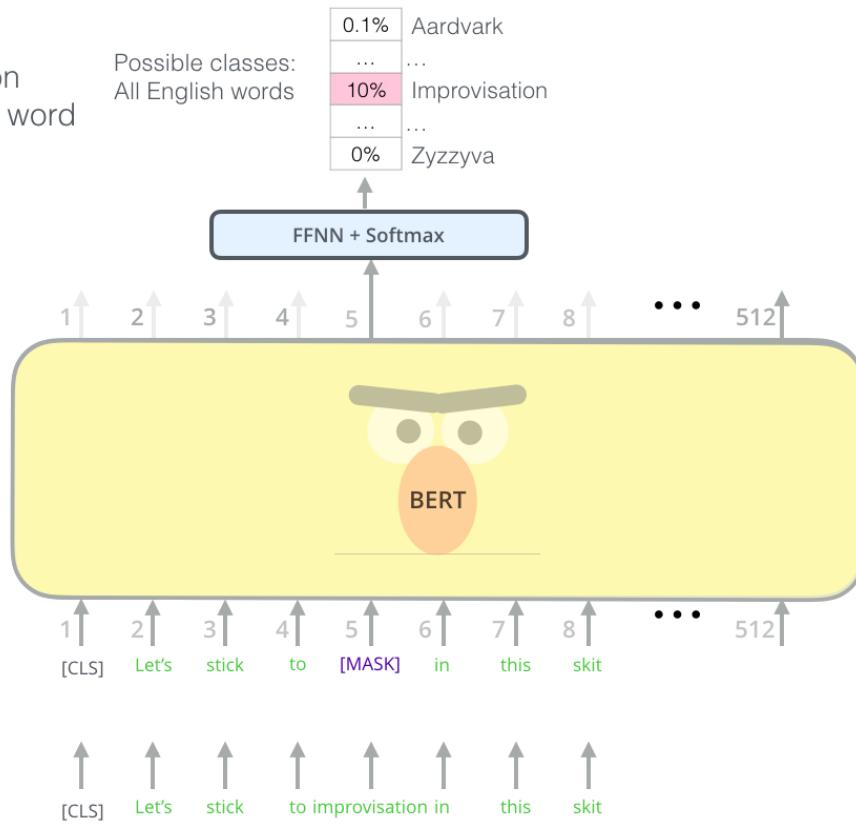
Das Training läuft in zwei Schritten: Pre-Training und Fine-Tuning (eine Form von Transfer Learning)

BERT Training using Masked Language Model

Use the output of the masked word's position to predict the masked word

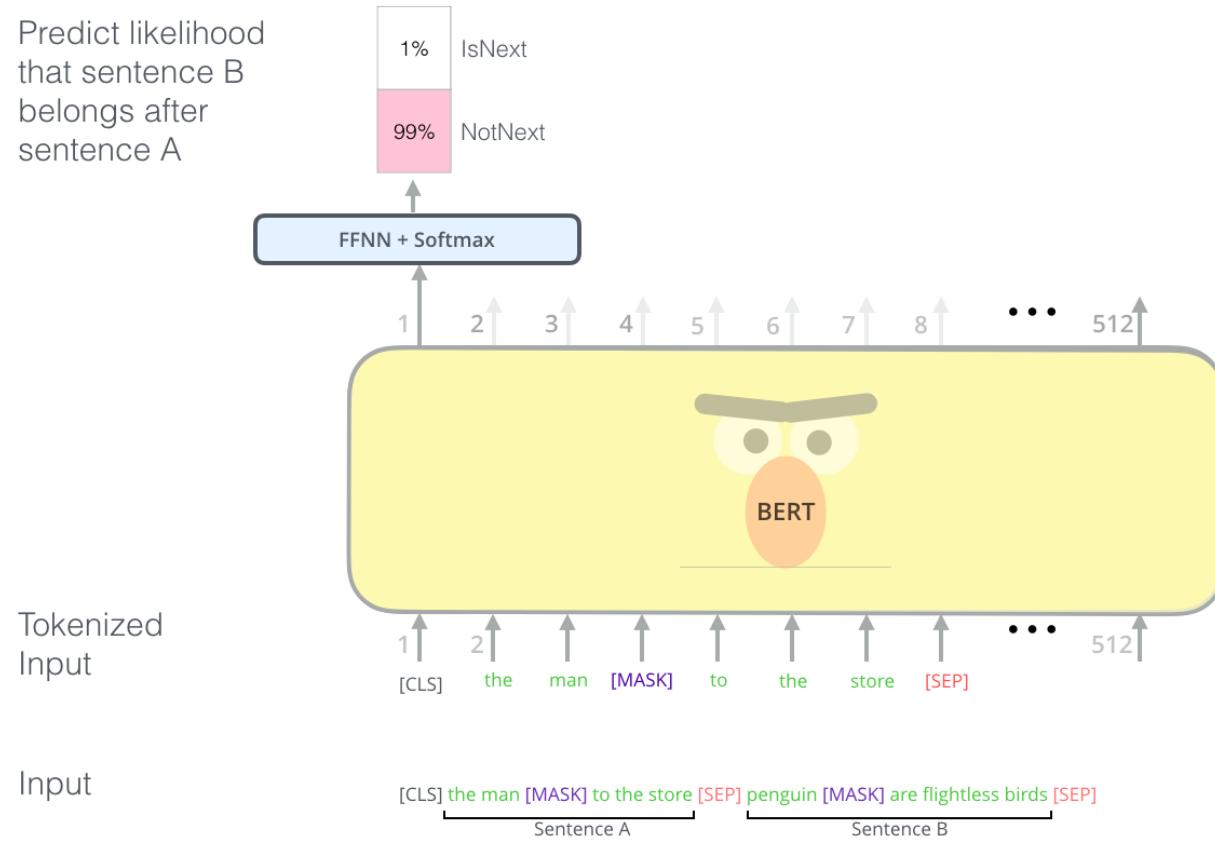
Randomly mask 15% of tokens

Input



Masked Language Model: BERT maskiert **15% der Worte im Input** und lässt das fehlende Wort **vorhersagen (Bidirektionale Suche)**.

BERT Training using Next Sentence Prediction

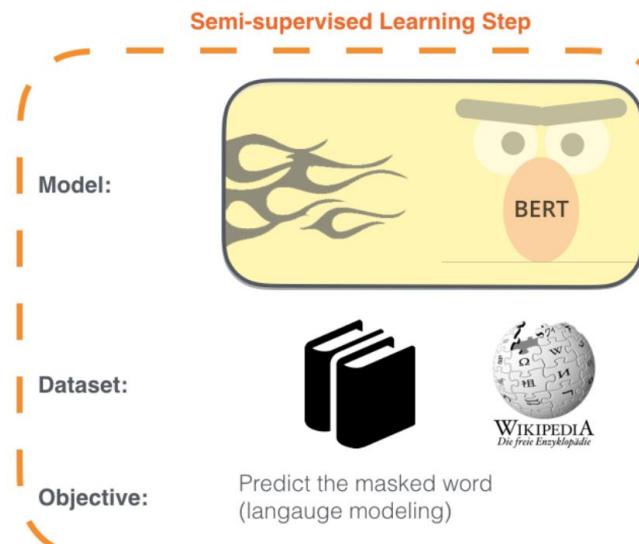


Next Sentence Prediction: BERT bekommt 2 Sätze im Input und sagt die Reihenfolge (ja/nein) voraus.

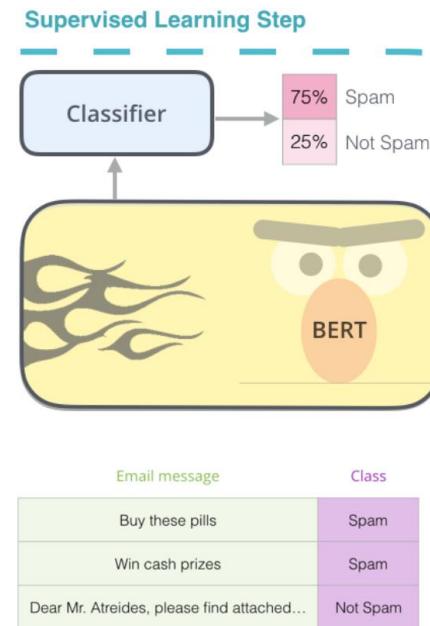
BERT Transformer, Google 2018

1 - Semi-supervised training on large amounts of text (books, wikipedia..etc).

The model is trained on a certain task that enables it to grasp patterns in language. By the end of the training process, BERT has language-processing abilities capable of empowering many models we later need to build and train in a supervised way.



2 - Supervised training on a specific task with a labeled dataset.

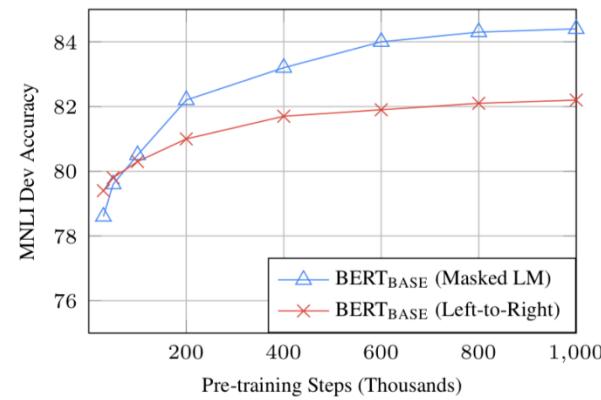
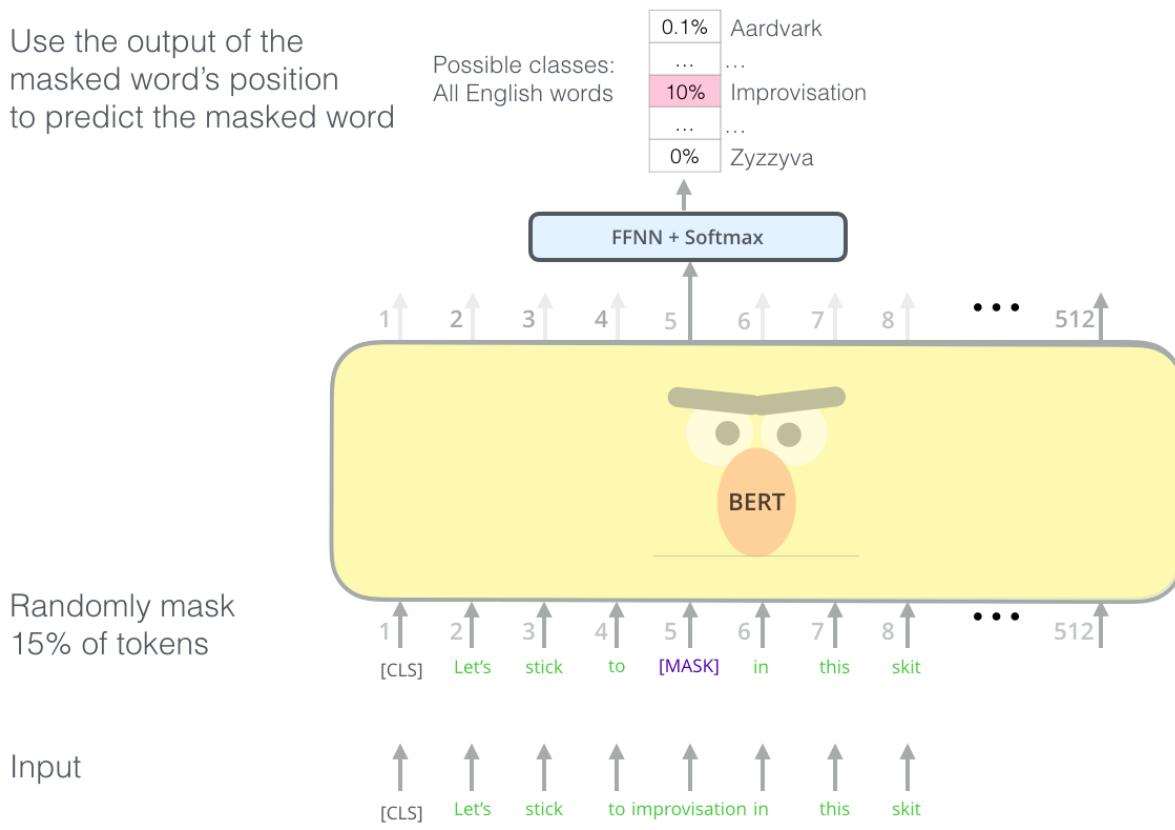


The two steps of how BERT is developed. You can download the model pre-trained in step 1 (trained on un-annotated data), and only worry about fine-tuning it for step 2. [Source for book icon].

Das **Training** läuft in zwei Schritten: **Pre-Training** und **Fine-Tuning** (eine Form von **Transfer Learning**)

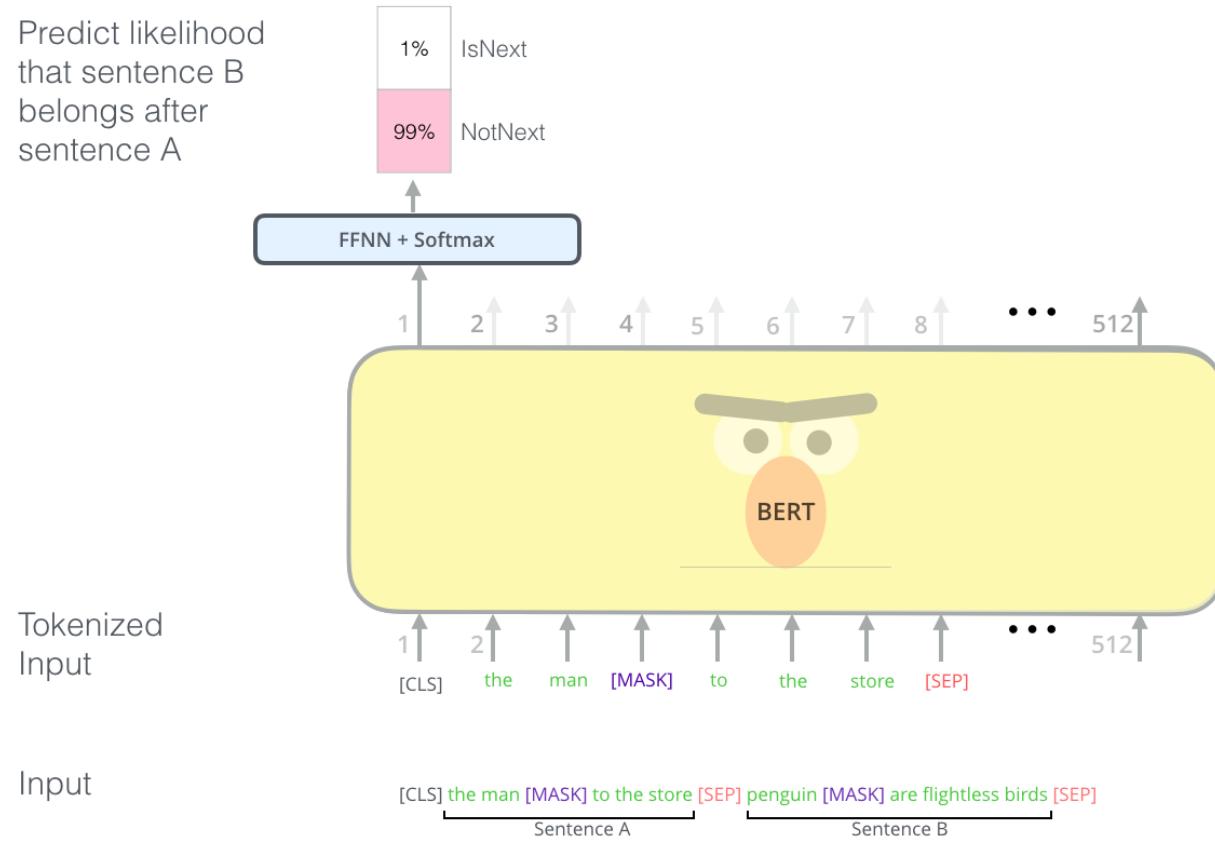
BERT Training using Masked Language Model

Use the output of the masked word's position to predict the masked word



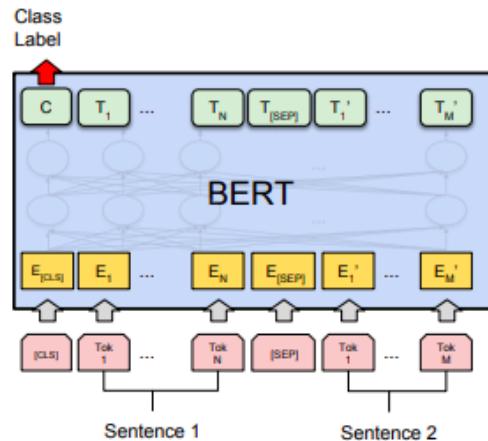
Masked Language Model: BERT maskiert **15% der Worte im Input** und lässt das fehlende Wort **vorhersagen (Bidirektionale Suche)**.

BERT Training using Next Sentence Prediction

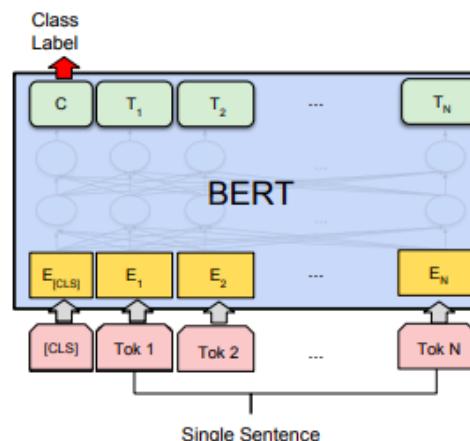


Next Sentence Prediction: BERT bekommt 2 Sätze im Input und sagt die Reihenfolge (ja/nein) voraus.

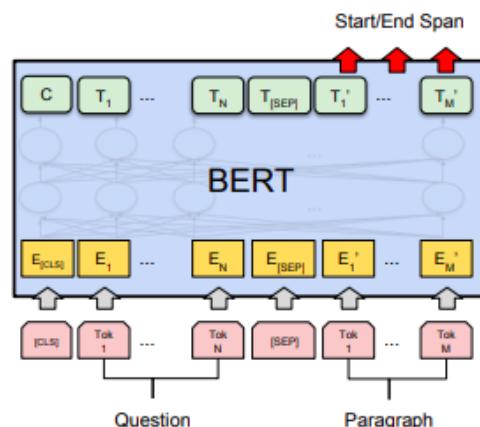
BERT Fine-Tuning Tasks (bzgl. GLUE Benchmark)



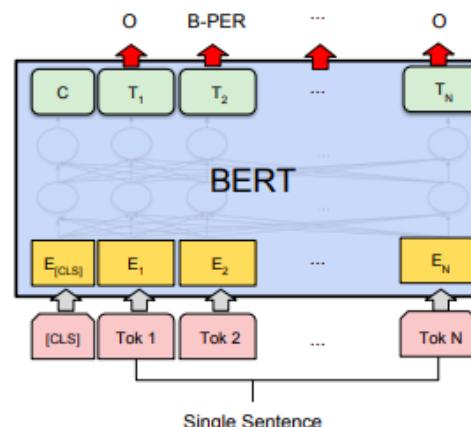
(a) Sentence Pair Classification Tasks:
MNLI, QQP, QNLI, STS-B, MRPC,
RTE, SWAG



(b) Single Sentence Classification Tasks:
SST-2, CoLA



(c) Question Answering Tasks:
SQuAD v1.1



(d) Single Sentence Tagging Tasks:
CoNLL-2003 NER