HUMORERKENNUNG IM NATURAL LANGUAGE PROCESSING

DBE Projekttag | Projekt 2 | 14.03.2023

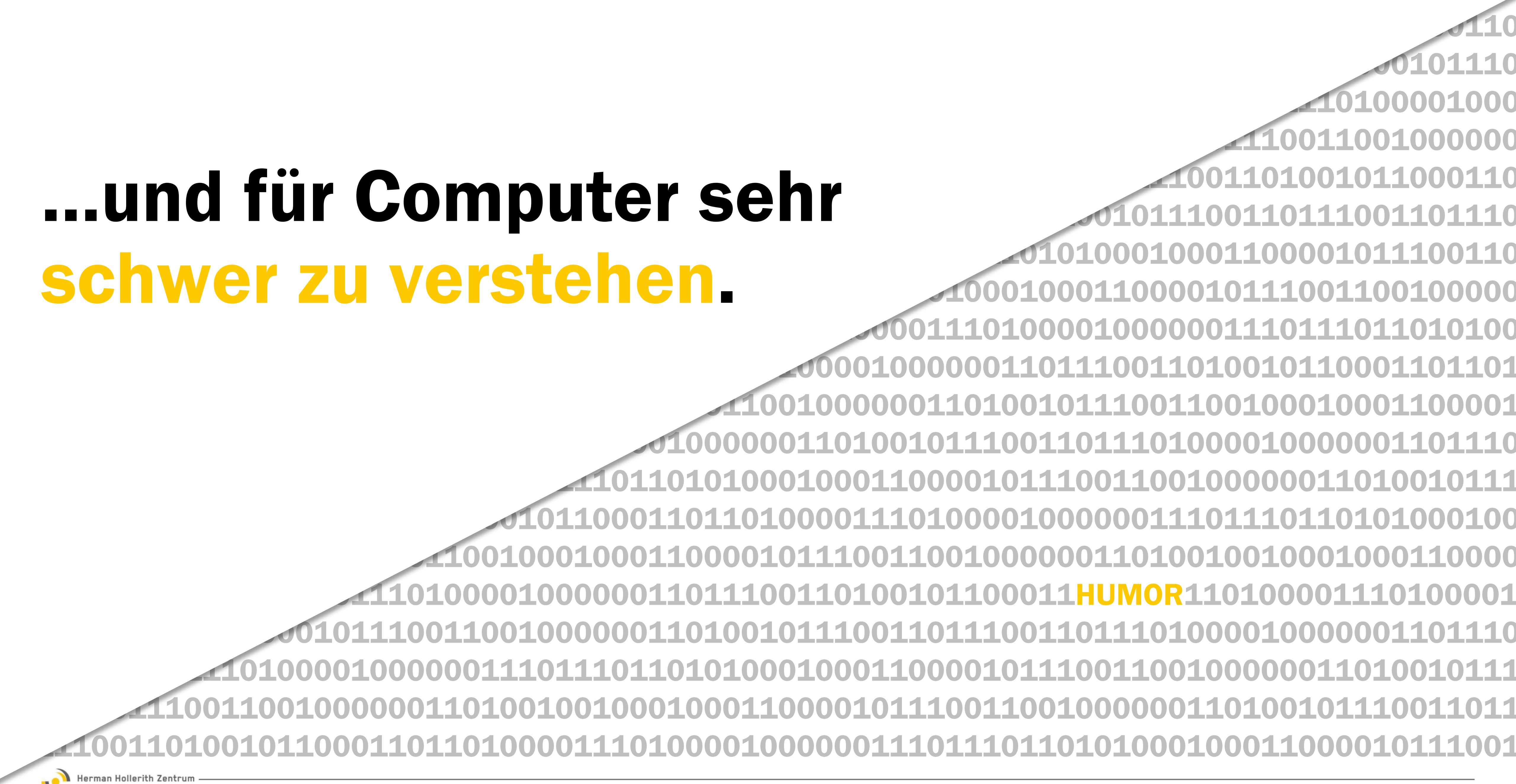
Maria Fürholzer Gowseegan Kanagaratnam Luke Roy Lucas Schweizer

Projekt 1
in a nutshell

Menschlicher Humor ist sehr individuell…

Cattell & Luborsky, 1947

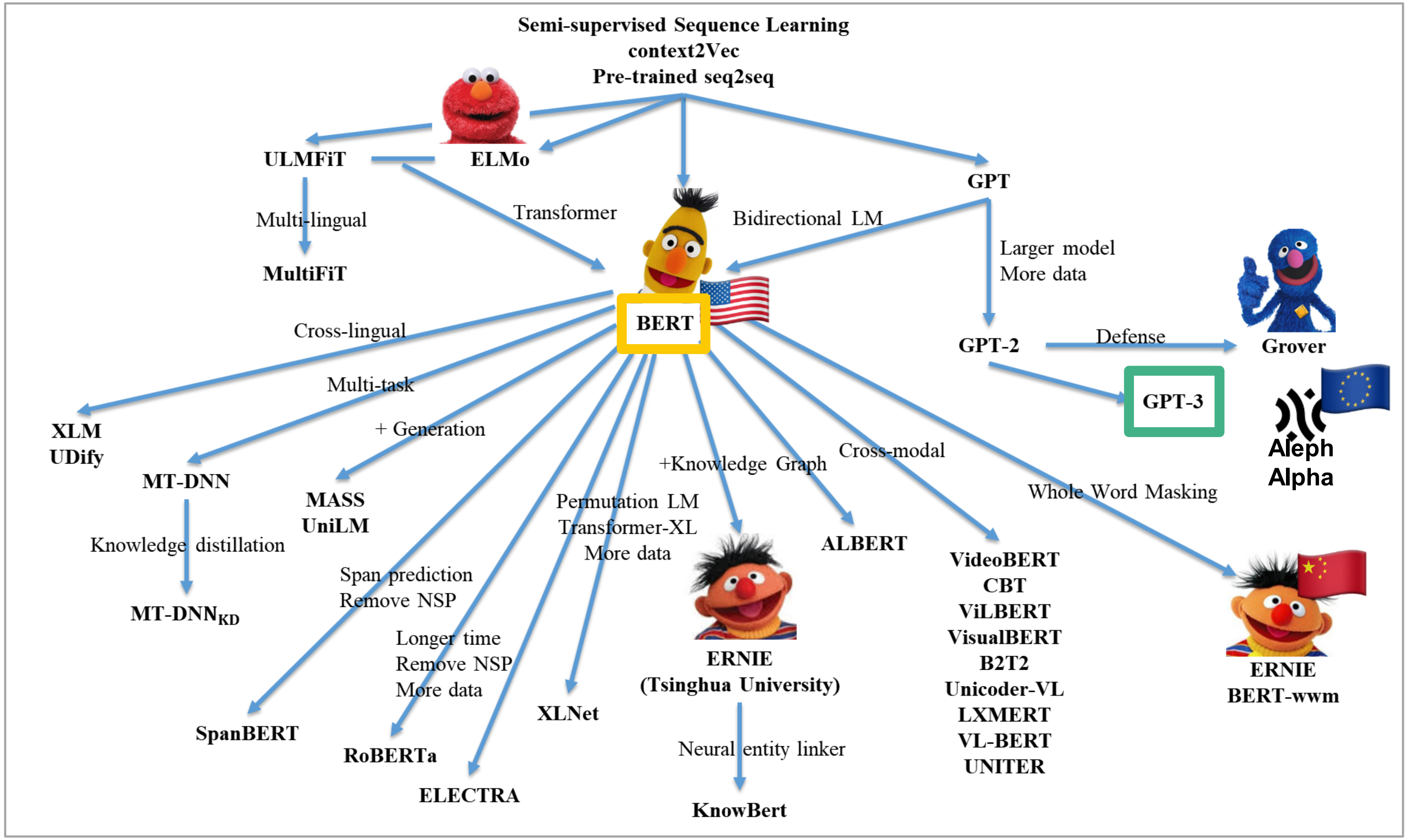
...und für Computer sehr schwer zu verstehen.



Natural **L**anguage **P**rocessing

NLP nutzt verschiedene Ansätze, um **menschliche Sprache verarbeiten und interpretieren** zu können.

Ziel ist, dass ein Computer **Sprache semantisch und kontextuell „verstehen“** kann.

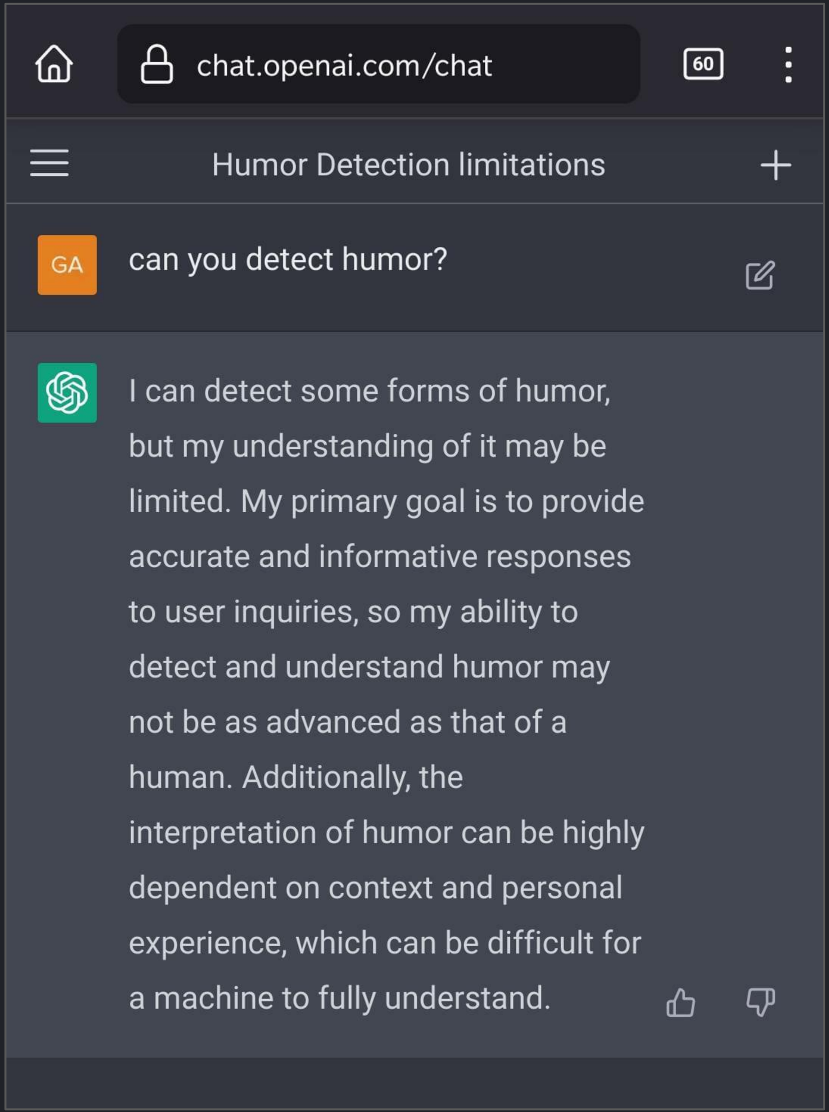


Must-Read Papers on Pre-trained Language Models – Wang & Zhang, 2019; aktualisierte Version

ChatGPT

OpenAI

GPT-3 und *GPT-4-unreleased* sind nicht mehr open source


chat.openai.com/chat
60
Humor Detection limitations
GA
can you detect humor?
I can detect some forms of humor, but my understanding of it may be limited. My primary goal is to provide accurate and informative responses to user inquiries, so my ability to detect and understand humor may not be as advanced as that of a human. Additionally, the interpretation of humor can be highly dependent on context and personal experience, which can be difficult for a machine to fully understand.

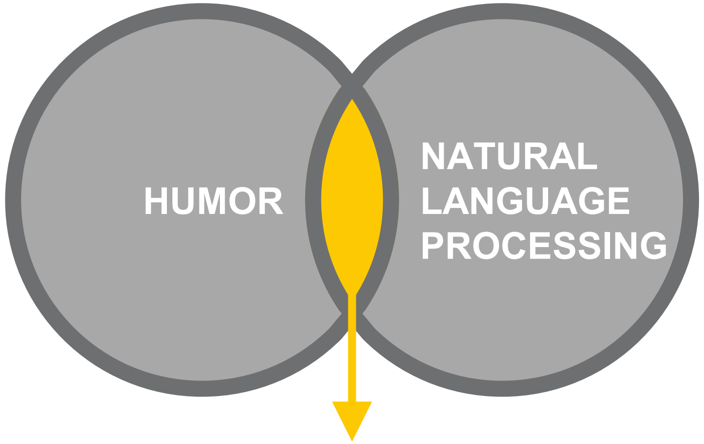
Projektplan

2022

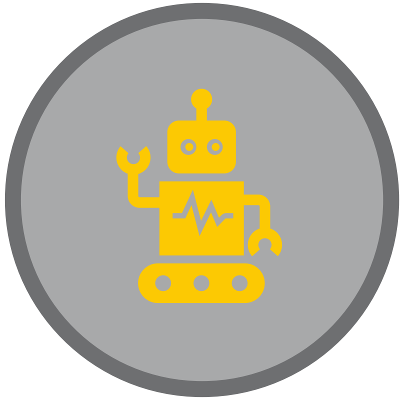
Projekt 1 Theorie verstehen

HEUTE

Projekt 2 Theorie anwenden



Humorerkennung im Natural Language Processing
Humortheorie, NLP-Ansätze, Forschungsstand



„Hey Computer, ist das lustig?"
Prototypische Implementierung einer WebApp

01
Technische Realisierung

02
Evaluation der Ergebnisse

03
Prototypische WebApp

01
Technische Realisierung

02
Evaluation der Ergebnisse

03
Prototypische WebApp

„One Liner“
Gelabelte Daten
99.965 Witze
100.001 Nicht-Witze
Englischsprachig

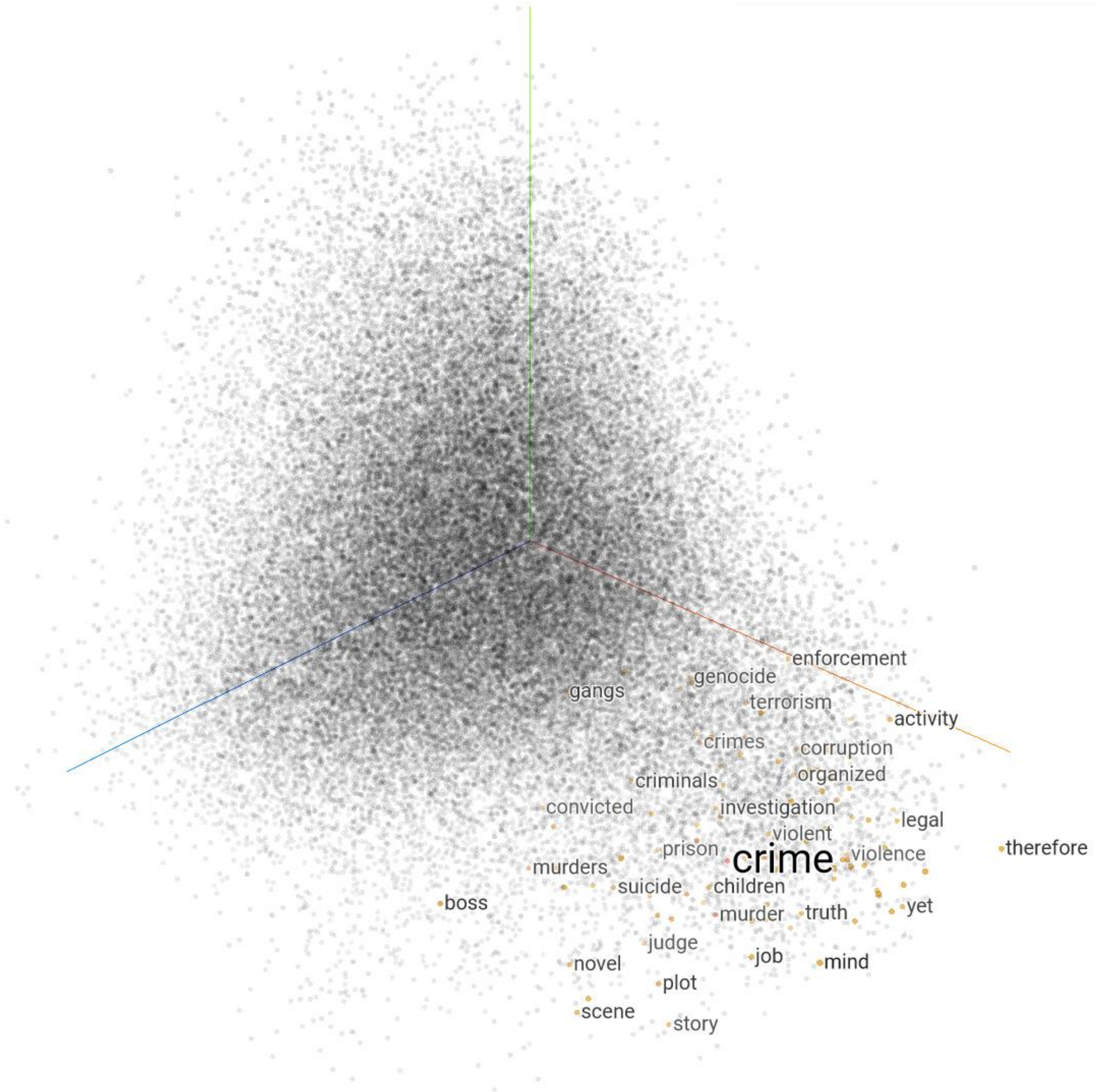
```
text,humor
"Joe biden rules out 2020 bid: 'guys, i'm not running'",False
Watch: darvish gave hitter whiplash with slow pitch,False
What do you call a turtle without its shell? dead.,True
5 reasons the 2016 election feels so personal,False
"Pasco police shot mexican migrant from behind, new autopsy shows",False
"Martha stewart tweets hideous food photo, twitter responds accordingly",False
What is a pokemon master's favorite kind of pasta? wartortellini!,True
Why do native americans hate it when it rains in april? because it brings mayflowers.,True
"Obama's climate change legacy is impressive, imperfect and vulnerable",False
"My family tree is a cactus, we're all pricks.",True
Donald trump has found something mysterious for rudy giuliani to do,False
How donald trump and ted cruz's love affair is all relationships,False
Want to know why athletes chose to #takeaknee? look at our broken justice system,False
How are music and candy similar? we throw away the rappers.,True
Famous couples who help each other stay healthy and fit,False
Study finds strong link between zika and guillain-barre syndrome,False
Alec baldwin and wife hilaria welcome another baby boy,False
"Trump says iran is complying with nuclear deal, but remains a dangerous threat",False
Kim kardashian baby name: reality star discusses the 'k' name possibility (video),False
"I just ended a 5 year relationship i'm fine, it wasn't my relationship :p",True
Here's what the oscar nominations should look like,False
Dating tip: surprise your date! show up a day early.,True
Reflections from davos: leaders deliberate what's next for climate action after paris deal,False
What do you call an explanation of an asian cooking show? a wok-through.,True
Swimming toward a brighter future: how i was introduced to the world of autism,False
Why did little miss muffet have gps on her tuffet? to keep her from losing her whey.,True
The pixelated 'simpsons' should be a real couch gag,False
All pants are breakaway pants if you're angry enough,True
Watch: former british open champ makes embarrassing putting fail,False
Chrissy teigen's 2015 grammy dress is skintight and perfect,False
"Ugh, i just spilled red wine all over the inside of my tummy.",True
The next iphone update will help you save lives,False
Celebrating the fourth of july with airport profiling,False
"The big bend, a u-shaped skyscraper, could become the longest in the world",False
Oscars 2016 red carpet: all the stunning looks from the academy awards,False
Why do jews have big noses? because the air is free,True
Interesting fact: by the year 2020 all actors on american tv shows will be australian.,True
I'd tell you a chemistry joke but i know i won't get a reaction,True
Arkansas approves law to let people carry guns in bars and at public colleges,False
On set with paul mitchell: from our network,False
Did you know diarrhea is genetic? it runs in your jeans,True
"My sons ebola joke what do africans have for breakfast? ebola cereal :) (be kind,he's only 14 lol)",True
What was the sci-fi remake of a streetcar named desire? interstellllllllaaaaaaar,True
What do you call a clan of barbarians you cant see? invisigoths,True
```

github.com/Moradnejad/ColBERT-Using-BERT-Sentence-Embedding-for-Humor-Detection

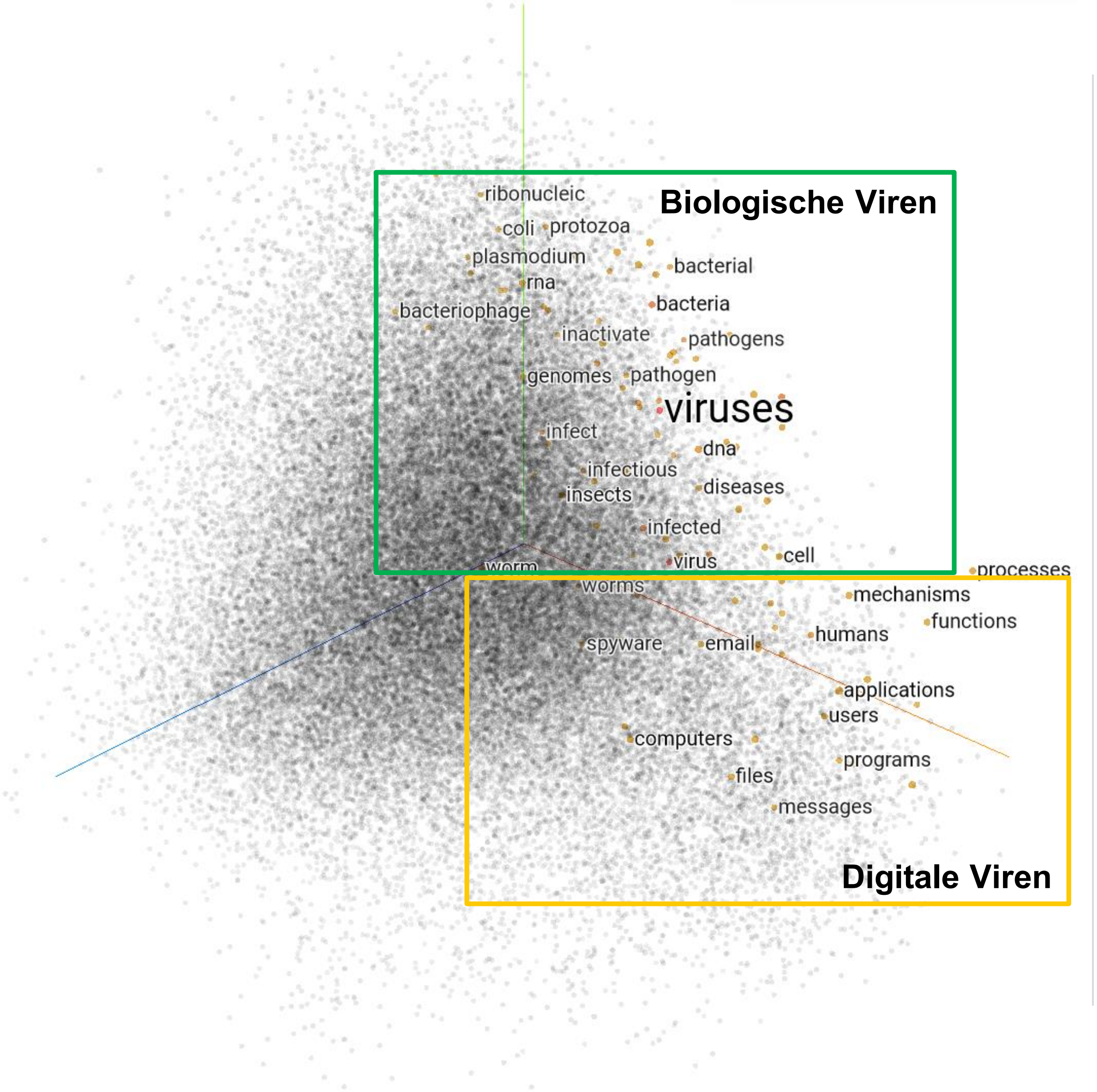
Beispiel

Word2Vec 71k

Umwandlung von 71.291 Wörtern in numerische Repräsentationen (= Vektoren), um mathematisch **Zusammenhänge der Wörter** zueinander berechnen und erkennen zu können

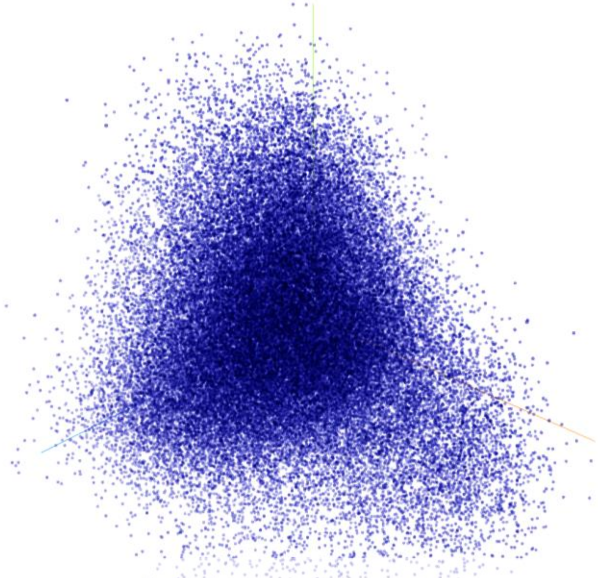


TensorFlow Projector


Biologische Viren
ribonucleic
coli
protozoa
plasmodium
bacterial
rna
bacteriophage
bacteria
inactivate
pathogens
genomes
pathogen
viruses
infect
dna
infectious
insects
diseases
infected
virus
cell
worm
worms
processes
mechanisms
functions
humans
spyware
email
applications
users
computers
programs
files
messages
Digitale Viren

Word2Vec

- **Kontextunabhängig**
  Ein Vektor (Embedding) für jedes Wort
- **Reihenfolgeunabhängig**
  Keine Beachtung der Wortreihenfolge

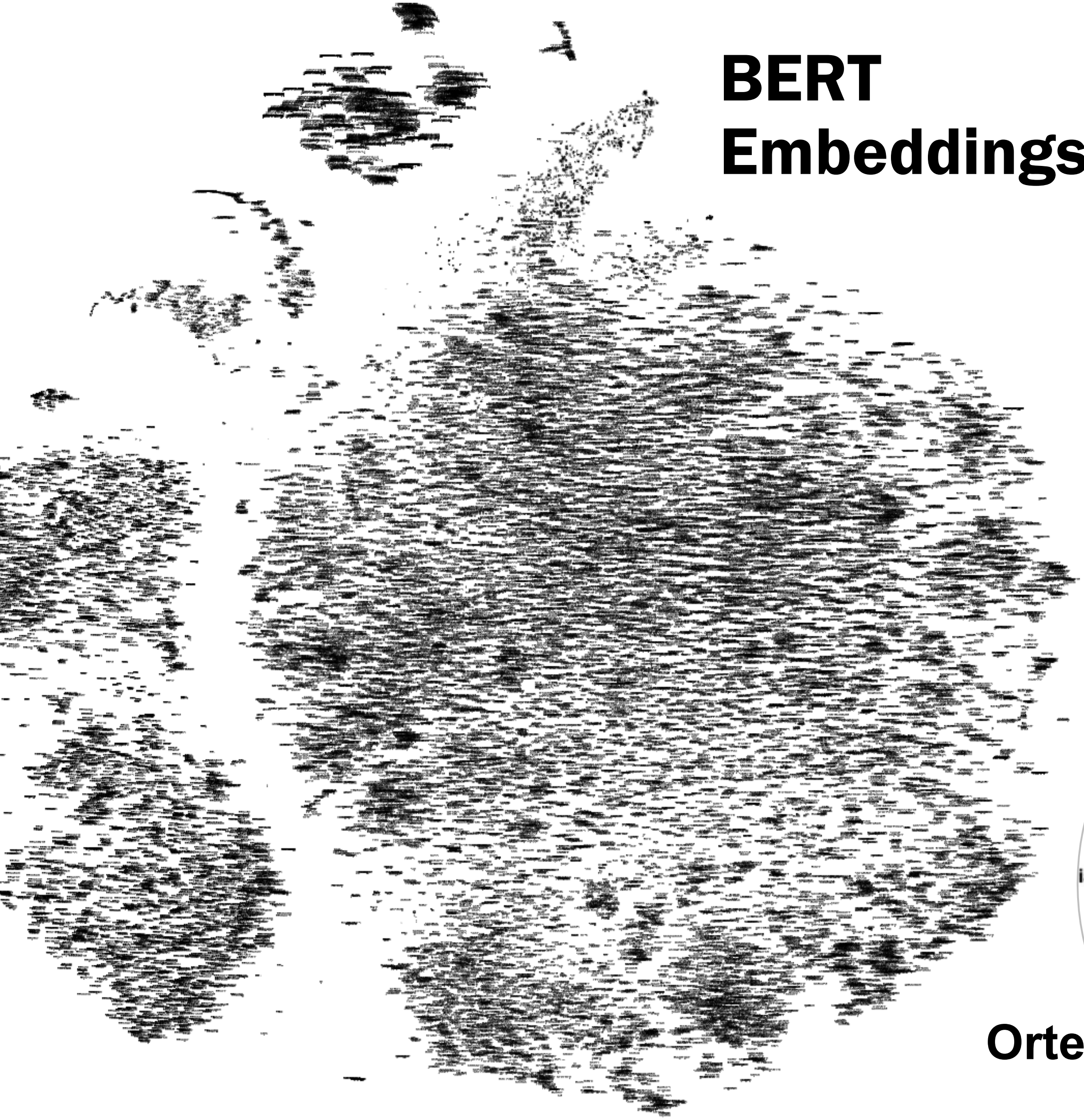


BERT

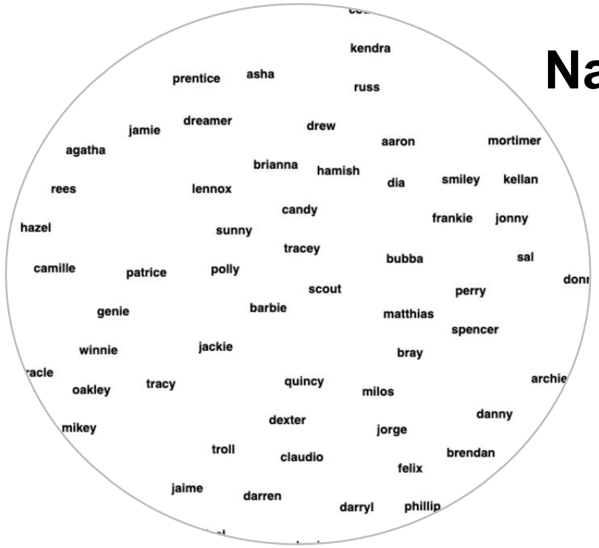
- **Kontextabhängig**
  x Vektoren (Embeddings) für jedes Wort
  Sitzbank vs Geldbank
- **Reihenfolgeabhängig**
  Beachtung der Wortreihenfolge im Satz

Sprachmodelle **berechnen die kontextuellen Beziehungen (Korrelationen)** zwischen Wörtern und **repräsentieren damit die natürliche Sprache als Vektor.** Diese Vektoren sind für den **Einsatz in Deep Learning Methoden** geeignet.

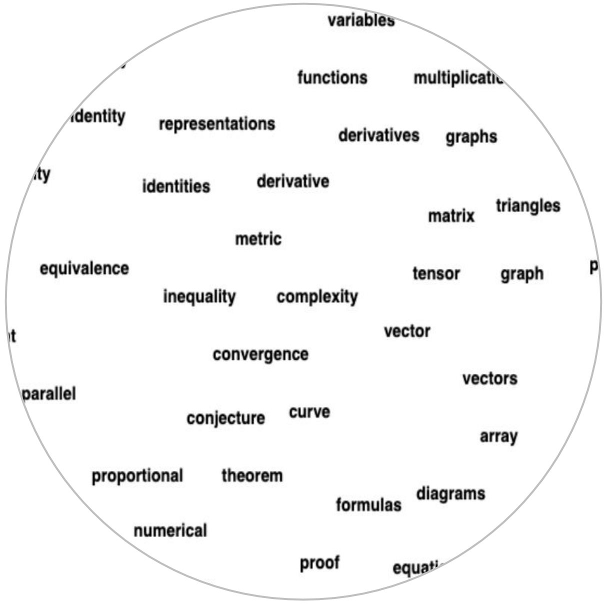
BERT Embeddings



Namen



Mathematik



Orte

Hu, ohne Jahr

Herman Hollerith Zentrum

Technische Realisierung

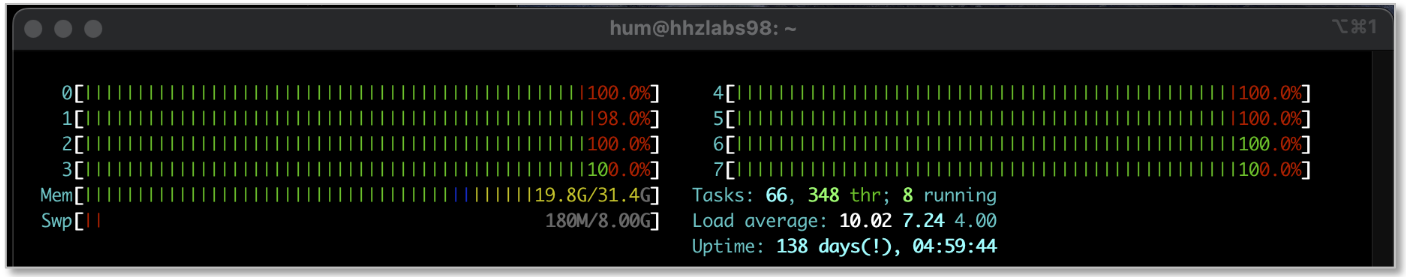
One Liner
Datensatz / Pre-Processing

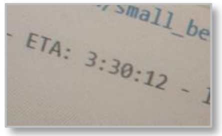
Pre-trained Language Model
smallBERT Konfigurationen

Deep Learning Model
Classifier

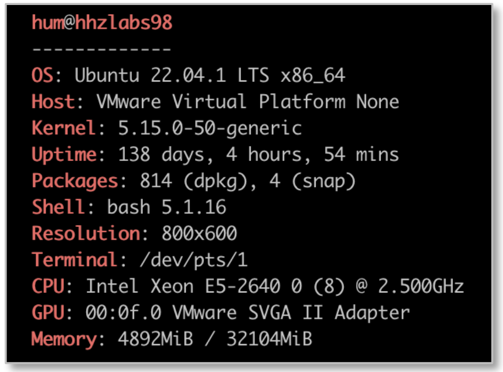
Kritische Reflexion der Umsetzung

Erkenntnis
NLP-Methoden benötigen sehr vieler Ressourcen





Problem & Lösungsversuch
> Limitierte Ressourcen des INF-Servers (Intel Xeon 8-core, 32GB RAM)
> IBM Cloud: leider konnte das Skript nicht adhoc auf den GPUs initialisiert werden



Ergebnis
> Automatisiertes Skript für das Pre-Processing und zum Trainieren sowie Abspeichern der Modelle
> Sieben smallBERT Konfigurationen konnten erfolgreich trainiert werden (26 Modelle in 7 verschiedenen Konfigs)

01
Technische Realisierung

02
Evaluation der Ergebnisse

03
Prototypische WebApp

Evaluation der Modellgüte

Probleme

- **Fehler im Python-Skript**

  Alle Modelle wurden nur mit der standardmäßigen Accuracy-Metrik gespeichert; dass weitere Metriken nicht mitgespeichert werden, war uns nicht bewusst

  → Accuracy = wie wahrscheinlich, dass Humor

  → 1 – Accuracy = wie wahrscheinlich, dass kein Humor

- **Fehler in der Entwicklungsumgebung**

  Zellen-Output ist in remote Jupyter Notebooks nicht mehr aufrufbar; somit sind die Metriken, die während des Trainings dargestellt worden sind, verloren gegangen

Lösung

- **Eigene Evaluationslogik mit Testdatensatz entwickelt**
- **30 Witze und 30 Nicht-Witze in englischer Sprache**
- **Logik Best Case Accuracy**

  → 30 Witze à 100% und 30 Nicht-Witze à 0% Accuracy

  → Optimaler Wert = 30 (ergibt sich aus 30 * 100% + 30 * 0%)

  → Summe aller 60 Accuracy-Kennzahlen pro Modell

  → Ermittlung des Abstands zum Optimum

  → Je kleiner der Abstand, desto höher die Modellgüte

  → Zusätzlich farbliche Codierung der Ergebnisse

=ABS(SUMME(Accuracy**)-30)**

WITZE

GÜTE	MODELL	My wife to	I went to b	I failed ma	Did you kn	Don't you	Two wifi e	What's the	Refusing t	Despite th	The rotatio	A book fell	If you don'	A ghost wa	A blind ma	How do yo	Maybe if w	Build a ma	Why did th	What do y	The troubl	Why don't	I wanted m	What's the	What do y	Don't spel	There's a r	Atheism is	The other	Animal tes	What do yo
9.6	model_small_bert_	98.16%	97.96%	96.07%	96.71%	98.46%	81.65%	98.27%	78.83%	41.96%	91.48%	97.67%	93.08%	97.89%	97.44%	98.26%	97.34%	96.28%	97.44%	98.16%	95.11%	96.66%	98.32%	97.76%	98.16%	93.57%	97.22%	24.28%	97.81%	96.99%	98.26%
9.7	model_small_bert_	98.73%	98.92%	97.36%	97.91%	98.58%	58.44%	98.01%	74.00%	64.25%	90.69%	98.85%	93.95%	98.92%	98.56%	98.27%	97.72%	96.47%	98.39%	98.88%	94.54%	98.26%	99.08%	98.82%	98.83%	96.40%	98.41%	24.61%	98.83%	98.06%	98.35%
11.9	model_albert_en_	99.44%	99.39%	97.93%	98.96%	98.98%	57.14%	97.58%	80.25%	54.95%	88.97%	99.40%	88.17%	99.45%	99.27%	99.09%	97.14%	95.16%	98.49%	99.44%	94.54%	98.84%	99.55%	99.49%	99.53%	92.23%	98.26%	26.47%	99.39%	99.03%	98.47%
12.1	model_small_bert_	98.18%	98.08%	97.15%	98.22%	98.53%	70.87%	97.91%	84.77%	61.65%	91.57%	97.54%	95.23%	98.01%	98.51%	98.41%	96.93%	95.67%	98.01%	98.42%	93.80%	98.01%	98.38%	98.35%	98.66%	95.64%	97.50%	31.64%	98.20%	97.34%	98.61%
13.1	model_small_bert_	97.53%	97.91%	97.15%	98.06%	98.16%	88.05%	97.82%	86.28%	83.98%	93.70%	97.90%	97.01%	97.81%	97.82%	98.40%	97.21%	97.33%	98.15%	98.40%	96.51%	97.89%	98.12%	97.96%	97.72%	97.17%	98.09%	46.50%	97.88%	97.42%	98.18%
14.6	model_small_bert_	97.45%	98.11%	97.29%	97.47%	98.38%	90.63%	98.30%	90.52%	80.12%	94.96%	97.43%	96.96%	98.09%	97.59%	97.85%	98.33%	96.31%	97.35%	97.35%	96.32%	98.36%	98.02%	98.24%	97.95%	97.36%	97.86%	55.80%	97.37%	98.21%	98.10%
15.5	model_small_bert_	98.25%	97.39%	98.01%	98.14%	98.75%	87.20%	97.03%	86.35%	68.86%	91.38%	98.14%	97.56%	97.90%	97.20%	98.46%	98.13%	93.85%	98.27%	98.26%	97.96%	98.53%	98.35%	98.69%	98.16%	93.64%	98.07%	14.80%	98.33%	97.53%	98.52%
17.2	model_small_bert_	99.36%	95.77%	92.27%	99.00%	91.97%	59.63%	73.19%	92.54%	45.68%	91.15%	98.97%	72.29%	98.21%	97.63%	99.07%	93.89%	84.38%	98.48%	99.11%	96.41%	95.68%	99.25%	99.39%	99.21%	36.40%	83.56%	33.09%	99.20%	98.71%	96.72%
17.2	model_small_bert_	99.60%	97.39%	97.74%	99.48%	99.19%	54.57%	98.47%	73.92%	52.30%	64.99%	99.43%	91.89%	99.36%	98.39%	99.72%	98.24%	85.11%	99.66%	99.66%	89.50%	99.55%	99.51%	99.77%	99.76%	38.82%	90.87%	9.68%	99.28%	98.96%	99.50%
19.8	model_small_bert_	99.89%	68.63%	65.50%	99.87%	31.62%	94.54%	39.56%	96.99%	92.86%	97.77%	99.83%	4.14%	98.52%	99.36%	99.97%	19.81%	75.86%	97.97%	99.96%	98.95%	82.78%	99.81%	99.98%	99.98%	3.69%	3.08%	89.51%	99.71%	98.04%	96.99%
20.0	model_small_bert_	99.99%	0.01%	0.00%	100.00%	0.01%	99.97%	0.01%	99.99%	99.72%	99.99%	99.98%	0.00%	0.07%	99.98%	100.00%	0.03%	0.11%	99.31%	100.00%	99.99%	99.26%	99.99%	100.00%	100.00%	0.00%	0.00%	99.98%	99.99%	100.00%	99.74%
20.6	model_small_bert_	99.76%	37.36%	69.29%	99.71%	44.27%	84.69%	38.86%	97.29%	82.11%	93.50%	99.16%	31.07%	98.74%	99.63%	99.77%	71.54%	46.17%	91.67%	99.82%	98.01%	81.51%	99.77%	99.86%	99.84%	4.09%	10.51%	81.32%	99.48%	99.60%	85.67%
20.8	model_small_bert_	99.90%	64.43%	38.13%	99.97%	16.09%	94.31%	15.62%	99.24%	97.83%	99.58%	99.82%	1.89%	97.63%	99.98%	99.95%	7.46%	71.93%	93.82%	99.97%	99.75%	67.44%	99.89%	99.99%	100.00%	1.06%	2.34%	92.74%	99.62%	99.64%	56.97%
21.7	model_small_bert_	99.73%	99.55%	98.82%	99.64%	99.70%	75.75%	99.52%	94.68%	75.97%	98.42%	99.70%	96.89%	99.71%	99.65%	99.78%	99.38%	99.27%	99.62%	99.78%	99.49%	99.43%	99.70%	99.83%	99.79%	95.69%	99.29%	52.21%	99.66%	99.68%	99.69%
23.5	model_small_bert_	99.56%	98.97%	99.45%	99.30%	97.92%	94.54%	99.27%	92.74%	84.87%	96.07%	99.21%	87.94%	99.37%	98.97%	99.41%	98.44%	96.85%	99.60%	99.68%	97.14%	99.52%	99.54%	99.76%	99.67%	92.76%	97.92%	62.23%	99.53%	98.31%	99.57%
23.6	model_small_bert_	99.54%	99.41%	99.47%	99.39%	99.17%	93.15%	99.39%	93.65%	93.06%	97.63%	99.39%	97.11%	99.63%	99.47%	99.71%	99.25%	99.26%	99.58%	99.71%	97.58%	99.55%	99.55%	99.52%	99.74%	97.26%	99.28%	83.40%	99.24%	98.75%	99.70%
23.9	model_small_bert_	99.79%	98.71%	98.35%	99.74%	94.71%	83.36%	90.06%	97.65%	94.37%	97.66%	99.67%	70.10%	99.12%	99.39%	99.84%	98.53%	89.39%	99.69%	99.78%	99.18%	97.55%	99.77%	99.88%	99.86%	42.89%	95.17%	59.40%	99.69%	99.61%	98.63%
24.4	model_small_bert_	99.94%	58.23%	47.98%	99.92%	64.23%	67.13%	73.12%	99.86%	99.15%	99.81%	99.58%	4.50%	99.11%	99.90%	99.95%	45.50%	66.10%	99.68%	99.92%	99.85%	99.67%	99.91%	99.97%	99.96%	12.50%	48.69%	99.67%	99.85%	99.90%	99.13%
24.5	model_small_bert_	99.80%	99.26%	99.46%	99.64%	99.47%	95.82%	99.54%	95.65%	91.88%	96.84%	99.63%	85.35%	99.47%	99.28%	99.84%	99.56%	97.54%	99.74%	99.88%	93.32%	99.56%	99.64%	99.93%	99.89%	83.18%	93.52%	79.45%	99.69%	99.19%	99.72%
24.5	model_small_bert_	99.66%	99.56%	99.58%	99.38%	99.35%	97.52%	99.67%	95.92%	89.66%	98.36%	99.63%	94.29%	99.62%	99.26%	99.68%	99.48%	99.02%	99.62%	99.73%	98.85%	99.59%	99.67%	99.71%	99.70%	98.74%	99.50%	69.35%	99.60%	99.44%	99.69%
24.7	model_small_bert_	99.93%	82.22%	93.94%	99.78%	66.90%	90.70%	75.13%	99.79%	97.09%	99.74%	99.66%	11.08%	94.51%	99.58%	99.97%	83.09%	75.91%	99.46%	99.95%	99.80%	97.62%	99.51%	99.99%	99.98%	1.34%	20.66%	99.08%	99.89%	99.90%	99.21%
25.4	model_small_bert_	99.89%	94.45%	99.49%	99.87%	20.85%	95.90%	89.86%	98.94%	98.16%	99.11%	99.80%	26.36%	99.52%	99.54%	99.92%	97.88%	90.39%	99.79%	99.96%	99.67%	98.35%	99.91%	99.96%	99.97%	22.66%	65.40%	93.77%	99.79%	99.81%	99.79%
25.5	model_small_bert_	99.90%	97.27%	99.17%	99.89%	89.72%	87.04%	97.07%	99.62%	98.16%	99.10%	98.48%	40.95%	99.00%	99.17%	99.92%	97.87%	94.14%	99.73%	99.94%	99.57%	99.72%	99.91%	99.93%	99.94%	48.36%	57.37%	85.92%	99.38%	99.75%	99.66%
26.3	model_small_bert_	99.94%	96.17%	97.22%	99.95%	92.31%	81.85%	97.65%	99.58%	98.99%	99.63%	99.87%	59.14%	99.65%	99.65%	99.96%	92.10%	95.94%	99.92%	99.97%	99.81%	99.83%	99.80%	99.99%	99.98%	43.85%	50.48%	96.13%	99.61%	99.80%	99.90%
26.3	model_small_bert_	99.85%	99.03%	98.89%	99.56%	98.62%	98.50%	98.51%	96.57%	95.25%	97.34%	99.79%	75.69%	99.81%	99.25%	99.83%	99.00%	98.77%	99.57%	99.81%	97.43%	99.40%	99.83%	99.91%	99.81%	95.24%	98.13%	87.89%	99.67%	99.24%	99.39%
27.6	model_small_bert_	99.55%	98.98%	99.15%	99.53%	98.60%	94.42%	98.94%	99.05%	95.87%	97.37%	99.33%	83.97%	99.60%	99.29%	99.62%	98.98%	99.45%	99.63%	99.72%	98.59%	99.61%	99.57%	99.75%	99.73%	95.20%	95.05%	94.55%	99.34%	99.30%	99.59%

NICHT-WITZE

GÜTE	MODELL	Germany i	Cost of livi	Along with	Nowadays	This TF Hu	They threv	Himalayan	The centra	The Chief J	Since 2021	The start o	Julius Caes	NATO forn	Brussels is	The Weser	The town l	pandas is a	Matko Des	In 1349; th	Farming is	The electio	Born into t	However t	In Decemb	It is a conc	He also se	The compa	The gestat	The univer	The city is
9.6	model_small_bert_	2.90%	63.71%	3.04%	4.19%	31.66%	95.08%	40.74%	33.40%	8.43%	27.76%	54.09%	25.37%	16.77%	2.91%	83.68%	38.79%	12.99%	58.72%	12.26%	59.57%	17.01%	89.09%	55.92%	29.83%	93.79%	37.83%	38.07%	32.50%	58.09%	86.85%
9.7	model_small_bert_	5.27%	55.73%	8.84%	9.53%	47.72%	98.19%	21.12%	22.93%	4.97%	17.12%	67.87%	22.71%	10.22%	3.46%	85.81%	45.55%	16.68%	72.68%	34.28%	66.15%	20.38%	88.26%	71.67%	24.20%	89.29%	19.20%	50.10%	34.21%	19.73%	72.34%
11.9	model_albert_en_	10.00%	71.47%	11.12%	20.18%	49.22%	99.29%	38.87%	18.02%	7.88%	31.52%	80.31%	30.28%	18.15%	3.59%	84.28%	68.52%	28.05%	81.92%	44.87%	80.02%	30.25%	94.38%	78.52%	25.92%	94.55%	32.67%	64.10%	31.63%	21.35%	80.12%
12.1	model_small_bert_	17.71%	46.86%	10.03%	17.49%	19.78%	91.73%	64.57%	15.65%	5.77%	32.63%	76.60%	17.84%	21.42%	4.63%	77.95%	75.74%	38.61%	64.82%	21.94%	80.35%	53.66%	86.05%	49.22%	31.16%	91.34%	62.34%	75.61%	48.91%	42.65%	82.15%
13.1	model_small_bert_	13.82%	44.22%	9.46%	10.59%	58.34%	96.67%	18.99%	64.01%	6.15%	29.47%	73.91%	48.20%	18.81%	11.34%	91.70%	41.41%	26.05%	54.66%	53.46%	70.02%	19.63%	85.87%	70.67%	57.83%	92.11%	19.31%	65.99%	71.07%	67.77%	78.15%
14.6	model_small_bert_	6.72%	61.27%	3.62%	4.70%	79.12%	97.38%	31.25%	68.11%	5.40%	50.01%	86.55%	66.77%	32.03%	8.34%	94.29%	63.98%	15.06%	66.01%	34.14%	83.41%	28.97%	89.34%	72.80%	58.13%	92.59%	21.66%	70.29%	73.63%	62.00%	82.00%
15.5	model_small_bert_	43.62%	36.91%	34.76%	19.76%	55.04%	96.96%	40.36%	74.01%	15.82%	60.85%	94.63%	60.27%	49.44%	31.29%	95.87%	78.03%	32.61%	51.09%	50.22%	64.26%	22.53%	89.87%	68.25%	64.67%	92.26%	21.98%	61.04%	74.72%	86.80%	91.89%
17.2	model_small_bert_	66.57%	92.91%	47.02%	64.07%	78.86%	98.58%	80.85%	4.28%	32.37%	47.47%	80.33%	86.01%	66.62%	36.39%	94.30%	85.47%	64.97%	87.31%	87.41%	87.66%	31.30%	95.63%	72.17%	42.85%	96.53%	78.44%	89.35%	61.27%	47.81%	92.99%
17.2	model_small_bert_	71.93%	90.26%	24.39%	80.72%	47.15%	99.10%	76.14%	7.23%	31.38%	73.25%	82.79%	56.34%	46.81%	38.64%	75.28%	85.20%	84.10%	61.51%	83.39%	96.37%	44.44%	98.13%	80.96%	62.22%	99.28%	91.30%	93.07%	41.32%	71.81%	95.79%
19.8	model_small_bert_	96.22%	99.41%	96.88%	98.66%	97.12%	99.54%	88.11%	28.07%	80.29%	99.44%	58.46%	80.41%	65.86%	83.28%	97.12%	91.63%	96.40%	92.91%	74.04%	97.41%	84.85%	96.59%	98.20%	90.15%	99.02%	93.80%	96.93%	76.01%	73.35%	96.08%
20.0	model_small_bert_	99.98%	100.00%	99.88%	99.98%	99.99%	99.99%	99.98%	99.49%	99.98%	99.99%	99.97%	99.99%	99.99%	99.64%	99.99%	99.98%	99.97%	98.72%	99.96%	99.93%	99.95%	100.00%	99.99%	99.96%	99.98%	99.99%	99.99%	99.98%	99.99%	99.99%
20.6	model_small_bert_	95.69%	98.70%	71.25%	94.63%	91.35%	99.62%	98.02%	11.59%	94.51%	91.98%	87.29%	95.20%	91.99%	82.68%	97.59%	98.50%	95.87%	93.23%	97.70%	99.02%	84.53%	99.47%	92.40%	89.41%	99.70%	98.44%	98.19%	92.59%	74.81%	99.40%
20.8	model_small_bert_	98.99%	99.92%	98.36%	99.38%	96.74%	99.93%	96.83%	39.40%	96.29%	99.74%	95.95%	96.77%	86.39%	95.15%	99.09%	98.56%	97.99%	87.54%	96.23%	98.59%	99.36%	99.67%	98.05%	99.23%	99.60%	99.21%	99.32%	97.29%	97.30%	99.74%
21.7	model_small_bert_	66.86%	96.17%	27.34%	73.16%	73.55%	99.51%	96.03%	34.51%	31.57%	72.22%	95.18%	67.43%	52.67%	14.48%	96.62%	91.19%	88.44%	96.64%	77.00%	98.63%	68.37%	99.15%	87.20%	76.29%	99.46%	87.90%	94.49%	55.77%	78.01%	98.48%
23.5	model_small_bert_	87.20%	94.81%	91.18%	74.85%	82.00%	99.41%	91.16%	45.04%	21.16%	96.46%	92.49%	42.54%	76.28%	67.01%	85.69%	94.93%	91.98%	84.00%	93.43%	97.36%	80.86%	95.59%	95.83%	70.49%	99.02%	55.38%	92.65%	77.62%	91.91%	96.50%
23.6	model_small_bert_	82.34%	92.53%	88.66%	70.49%	69.72%	99.00%	84.50%	66.57%	12.37%	86.26%	87.58%	65.89%	44.81%	82.00%	94.01%	79.40%	83.53%	91.52%	93.43%	94.76%	58.33%	92.44%	87.31%	86.71%	99.17%	80.85%	94.04%	74.68%	83.52%	95.24%
23.9	model_small_bert_	96.79%	98.86%	67.29%	97.03%	74.59%	99.44%	94.55%	16.89%	67.55%	86.33%	98.80%	66.24%	86.75%	76.07%	98.23%	94.34%	95.47%	95.22%	98.29%	99.33%	90.06%	98.04%	72.52%	80.96%	99.50%	83.69%	94.84%	91.50%	72.90%	98.40%
24.4	model_small_bert_	99.76%	99.86%	96.65%	99.70%	99.50%	99.91%	99.78%	85.85%	99.04%	99.39%	97.65%	99.55%	99.70%	98.86%	99.86%	99.62%	92.83%	98.57%	96.88%	99.82%	97.73%	99.73%	99.33%	97.71%	99.82%	99.25%	99.56%	99.70%	98.56%	99.84%
24.5	model_small_bert_	87.76%	82.49%	90.77%	94.61%	91.94%	99.24%	93.37%	42.67%	58.06%	96.48%	95.95%	75.94%	35.23%	75.89%	97.45%	95.09%	95.44%	96.14%	85.15%	96.15%	67.71%	92.00%	89.25%	67.69%	99.24%	83.99%	93.60%	93.46%	76.70%	98.55%
24.5	model_small_bert_	78.11%	97.70%	91.16%	53.41%	80.47%	99.50%	91.77%	50.19%	55.85%	84.69%	95.12%	82.61%	62.89%	52.78%	97.25%	94.25%	86.84%	86.03%	92.37%	96.95%	80.87%	98.43%	87.21%	82.52%	99.39%	73.68%	95.52%	80.65%	94.40%	98.48%
24.7	model_small_bert_	99.55%	99.80%	98.00%	99.46%	99.48%	99.86%	98.38%	76.44%	97.38%	99.21%	90.37%	98.75%	95.74%	92.97%	99.42%	96.46%	98.81%	89.45%	97.24%	94.99%	97.03%	99.77%	91.57%	84.08%	99.65%	95.89%	99.48%	98.96%	95.49%	99.75%
25.4	model_small_bert_	97.25%	99.67%	99.39%	98.01%	88.03%	99.94%	98.75%	72.65%	75.85%	98.74%	98.97%	89.97%	86.75%	93.31%	98.41%	98.41%	99.24%	96.65%	99.24%	99.81%	97.67%	99.35%	99.23%	94.91%	99.91%	91.56%	97.33%	96.57%	89.95%	99.74%
25.5	model_small_bert_	94.57%	99.00%	97.89%	99.03%	99.05%	99.55%	99.38%	47.31%	76.12%	97.16%	90.00%	95.83%	84.05%	96.81%	96.98%	97.82%	97.57%	96.49%	91.23%	98.95%	76.17%	95.13%	92.10%	81.81%	99.66%	91.57%	99.29%	93.35%	85.28%	99.59%
26.3	model_small_bert_	99.55%	99.60%	96.16%	99.73%	97.92%	99.86%	99.78%	49.31%	96.74%	99.33%	97.10%	97.77%	78.55%	97.86%	99.10%	98.89%	98.26%	94.21%	92.94%	99.35%	93.17%	94.53%	96.99%	84.16%	99.77%	94.59%	99.62%	97.31%	76.46%	99.57%
26.3	model_small_bert_	96.79%	97.99%	88.53%	97.31%	90.65%	99.68%	98.31%	43.85%	88.22%	95.77%	95.15%	94.97%	72.19%	77.14%	98.51%	98.11%	95.99%	94.81%	82.98%	95.94%	90.55%	98.53%	95.77%	52.90%	99.14%	89.52%	97.40%	89.42%	86.10%	98.33%
27.6	model_small_bert_	97.12%	97.93%	86.54%	96.92%	97.13%	99.49%	96.98%	40.58%	91.82%	96.29%	95.05%	97.26%	97.65%	83.91%	98.90%	98.52%	84.24%	96.28%	95.18%	98.76%	96.29%	99.13%	95.78%	94.64%	99.35%	95.87%	98.31%	98.50%	95.30%	98.62%

Das Modell mit der höchsten Güte

model_small_bert_bert_en_uncased_L-2_H-128_A-2

Kompakte BERT-Konfiguration nach Turc et al., 2019

Modellfamilie

Vortrainiert auf Englisch

Keine Beachtung von Groß-/Kleinschreibung

Hidden Layer

Hidden Size

Attention Heads

Zum Vergleich
BERT base model_bert_en_uncased_**L-12_H-768_A-12**
BERT large model_bert_en_uncased_**L-24_H-1024_A-16**

Erkenntnisse in den Testdaten über alle Modelle hinweg

- **BERT-konfigurationspezifische Eigenheiten**: Englische Sprache wird am besten erkannt; Groß-/Kleinschreibung wird nicht beachtet; Bedeutung mehrdeutiger Wörter wird richtig erkannt
- Einige Modelle scheinen **nicht-englische Eigennamen** (Orte, Flüsse, Personen) nicht richtig einordnen zu können
  Beispiele: Flüsse Fulda und Werra, Olaf Scholz **→ Zusammenhang mit englischer Sprache?**
- **Schwarzer Humor** (Verbrechen, Krankheit, Tod) wird teilweise nicht erkannt; anderseits werden aber auch teilweise **unangemessene Inhalte** als Humor klassifiziert **→ Zusammenhang mit Trainingsdatensatz oder BERT-Entwicklungsethik?**
- Themen wie **Politik, Religion und Ethik** werden von guten Modellen scheinbar **neutral betrachtet** **→ Zusammenhang mit Trainingsdatensatz oder mit BERT-Entwicklungsethik?**
- Einige der umfangreicheren Modell-Konfigurationen neigen zu **starkem Overfitting** **→ Datensatz zu „klein"?**

Vermutung: Antrainierte AI-Neutralität bei spezifischen Themen

GÜTE	MODELL	Atheism is a non-prophet organization.
9.6	model_small_bert_be	24.28%
9.7	model_small_bert_be	24.61%
11.9	model_albert_en_bas	26.47%
12.1	model_small_bert_be	31.64%
13.1	model_small_bert_be	46.50%
14.6	model_small_bert_be	55.80%
15.5	model_small_bert_be	14.80%
17.2	model_small_bert_be	33.09%
17.2	model_small_bert_be	9.68%
19.8	model_small_bert_be	89.51%
20.0	model_small_bert_be	99.98%
20.6	model_small_bert_be	81.32%
20.8	model_small_bert_be	92.74%
21.7	model_small_bert_be	52.21%
23.5	model_small_bert_be	62.23%
23.6	model_small_bert_be	83.40%
23.9	model_small_bert_be	59.40%
24.4	model_small_bert_be	99.67%
24.5	model_small_bert_be	79.45%
24.5	model_small_bert_be	69.35%
24.7	model_small_bert_be	99.08%
25.4	model_small_bert_be	93.77%
25.5	model_small_bert_be	85.92%
26.3	model_small_bert_be	96.13%
26.3	model_small_bert_be	87.89%
27.6	model_small_bert_be	94.55%

Atheism is a non-prophet organization.

Eigentlich zuverlässige Modelle verhalten sich neutral und klassifizieren den religiösen Witz als Nicht-Witz

Lediglich die Overfitting-Modelle klassifizieren den religiösen Witz als Witz

Inhaltliche Aspekte

- NLP ist verdammt **komplex**! Humorerkennung ist **noch komplexer**!
- Unsichere Interpretation der Genauigkeit, da BERT-Embeddings für uns eine nicht zu kontrollierende **Blackbox** sind und es sich um ein sehr **subjektives Thema** handelt
- Durch die Nutzung von BERT ist ein genereller **englischer Sprach- und Kultur-Bias** zu erwarten
- Durch die gelabelten Trainingsdaten (isJoke = TRUE / FALSE) wurde bereits ein **spezifisches, subjektives Humorempfinden** (= Bias) in das Modelltraining eingebracht
- Auch eine **menschliche Testgruppe** (bspw. mittels Amazon Mechanical Turk) hätte **keine 100% Genauigkeit beim Thema Humor**

01
Technische Realisierung

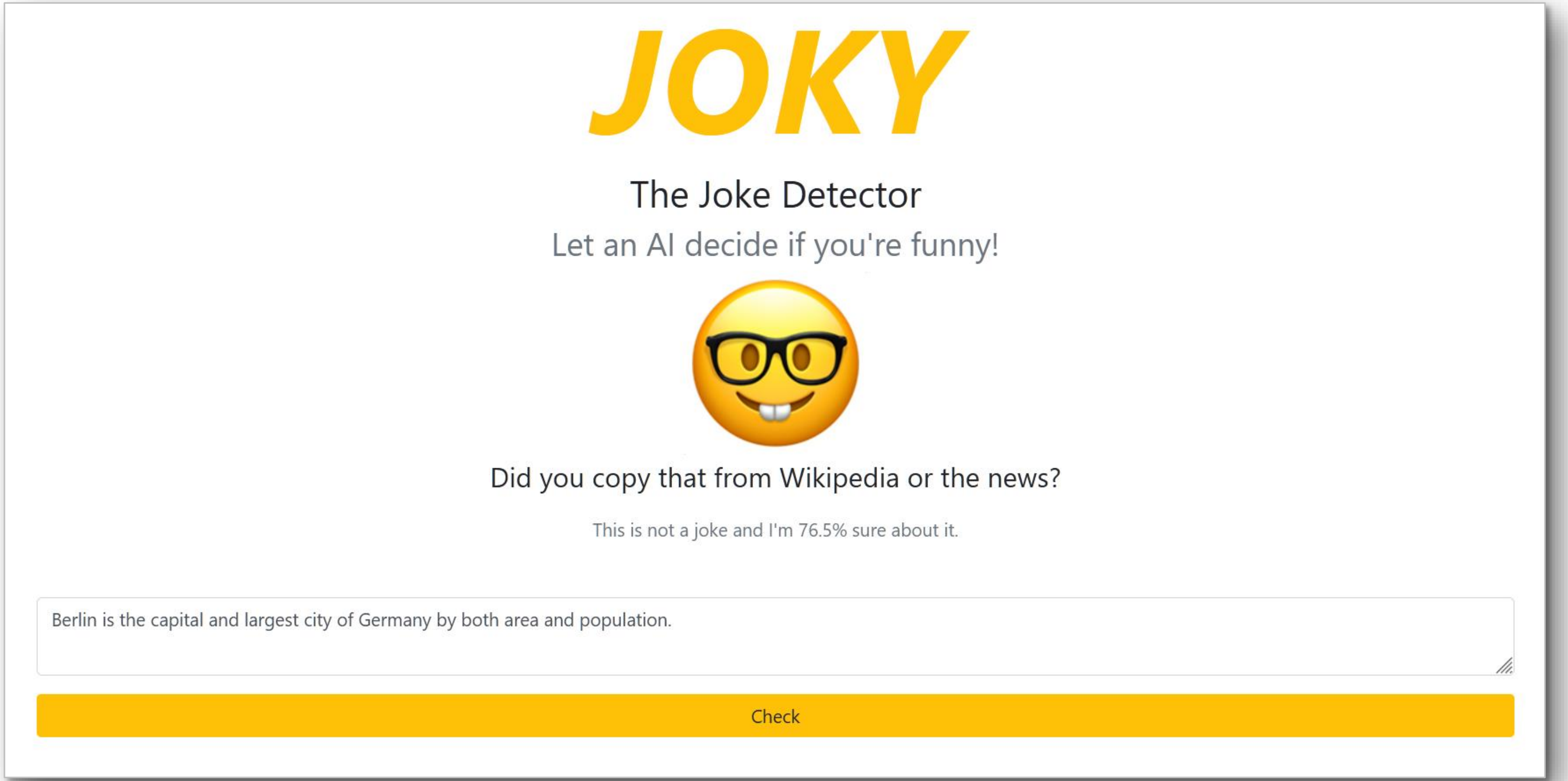
02
Evaluation der Ergebnisse

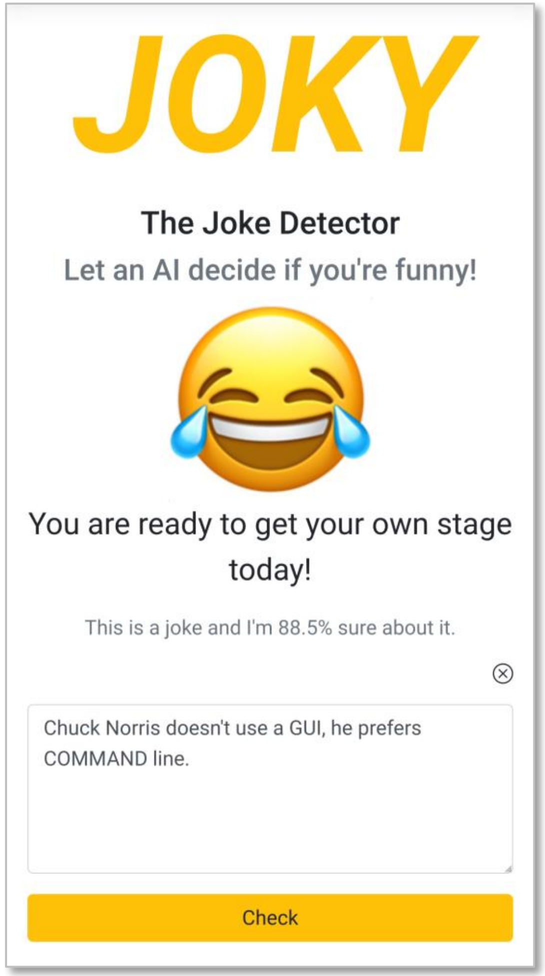
03
Prototypische WebApp

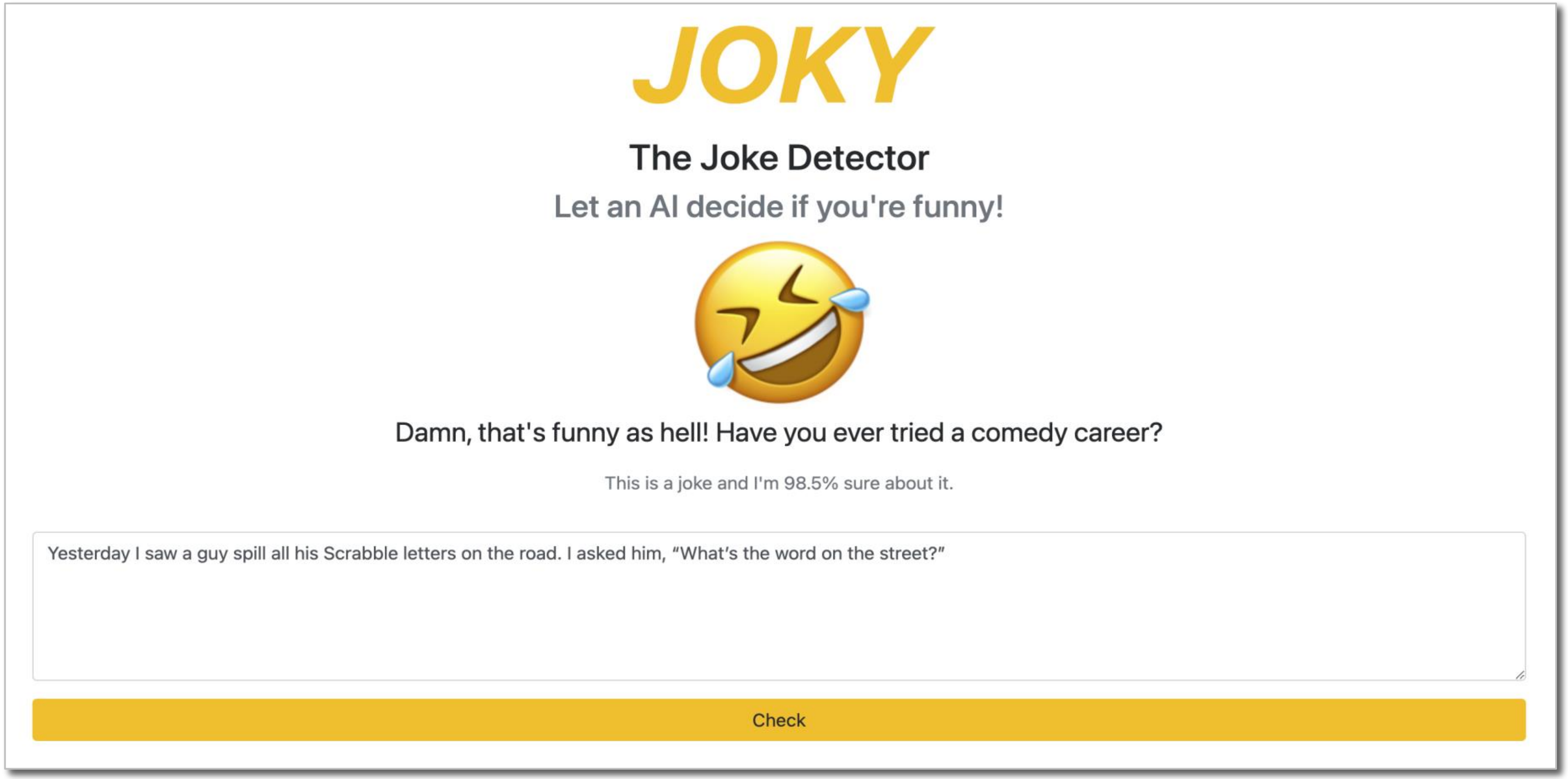
IBM
Cloud Code Engine

Funktionsweise
1. User tippen englischen Satz ein
2. Modell berechnet, mit welcher Wahrscheinlichkeit es sich bei der Eingabe um einen Witz bzw. nicht um einen Witz handelt
3. Direkte Ausgabe der Wahrscheinlichkeit und entsprechende schlagfertige Rückmeldung an die User


JOKY
The Joke Detector
Let an AI decide if you're funny!
"I don't get this joke...Maybe it's because it isn't one?"
This is not a joke and I'm 52.5% sure about it.
It was the second major conflict in the Opium Wars, which were fought over the right to import opium to China, and resulted in a second defeat for the Qing dynasty and the forced legalisation of the opium trade.
Check


JOKY
The Joke Detector
Let an AI decide if you're funny!
Did you copy that from Wikipedia or the news?
This is not a joke and I'm 76.5% sure about it.
Berlin is the capital and largest city of Germany by both area and population.
Check


JOKY
The Joke Detector
Let an AI decide if you're funny!
You are ready to get your own stage today!
This is a joke and I'm 88.5% sure about it.
Chuck Norris doesn't use a GUI, he prefers COMMAND line.
Check


JOKY
The Joke Detector
Let an AI decide if you're funny!
Damn, that's funny as hell! Have you ever tried a comedy career?
This is a joke and I'm 98.5% sure about it.
Yesterday I saw a guy spill all his Scrabble letters on the road. I asked him, "What's the word on the street?"
Check

Technische Aspekte

- Einbindung von NLP-Modellen bedarf **großer Rechenleistung**
- Der **INF-Server** hatte für größere BERT-Modelle zu wenig Leistung
- Auf der **IBM Cloud mit GPUs** (*Danke an Lukes Manager bei IBM*) konnte unser automatisiertes Skript leider nicht adhoc initialisiert werden
- Ein **remote Jupyter Notebook** ist nicht ideal für den Use Case „Automatisiertes Training von ML-Modellen im Hintergrund“

  Zellen-Output geht verloren, siehe Restoring computation output after disconnect in Jupyter notebook
- **Wenig praxisnahe Dokumentation** zur Optimierung von BERT in Zusammenspiel mit Tensor Flow

Ausprobieren und gemeinsam diskutieren.

Vielen Dank!

https://bit.ly/3JAVTrm

⚠ **Aktuell ist kein Filter für unangemessene Inhalte implementiert** ⚠

Quellenverzeichnis und Dokumentation

- **Cattell, R. B. und Luborsky, L. B. 1947: Personality factors in response to humor**, *The Journal of Abnormal and Social Psychology 42(4)*
- **ColBERT 2022:** ColBERT Dataset, *github.com/Moradnejad/ColBERT-Using-BERT-Sentence-Embedding-for-Humor-Detection*
- **Devlin, J., Chang, M.W., Lee, K. und Toutanova, K. 2019:** BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding
- **Google (Herausgeber):** Bidirectional Encoder Representations from Transformers (BERT), unter: tfhub.dev/google/collections/bert/1
- **Hu, Y. *keine Jahresangabe*:** Visualizing BERT Vocab/Token Embeddings, unter: yifanhu.net/BERT/
- **Turc, I., Chang, M.W., Lee, K., Toutanova, K. 2019:** Well-Read Students Learn Better: On the Importance of Pre-training Compact Models
- **Wang, X und Zhang, Z. 2019:** Must-Read Papers on Pre-trained Language Models, *github.com/thunlp/PLMpapers*

JOKY

https://bit.ly/3JAVTrm

GitHub

github.com/skywalkeretw/DBE-Humor-Prototype