

Neue Herausforderungen
und Chancen
durch künstliche Intelligenz
in der IT-Sicherheit

DI (FH) Georg Knabl, MSc



Page On Stage

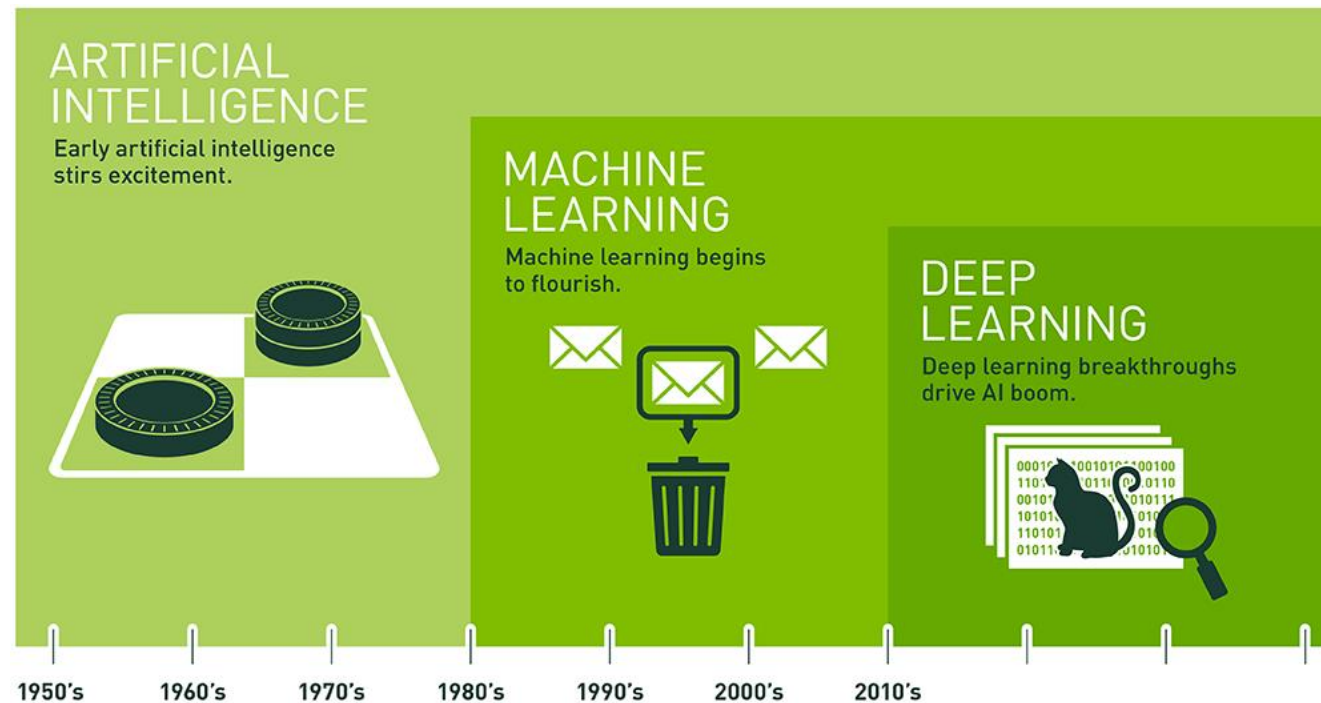
Advanced Web Engineering

DI (FH) Georg Knabl, MSc

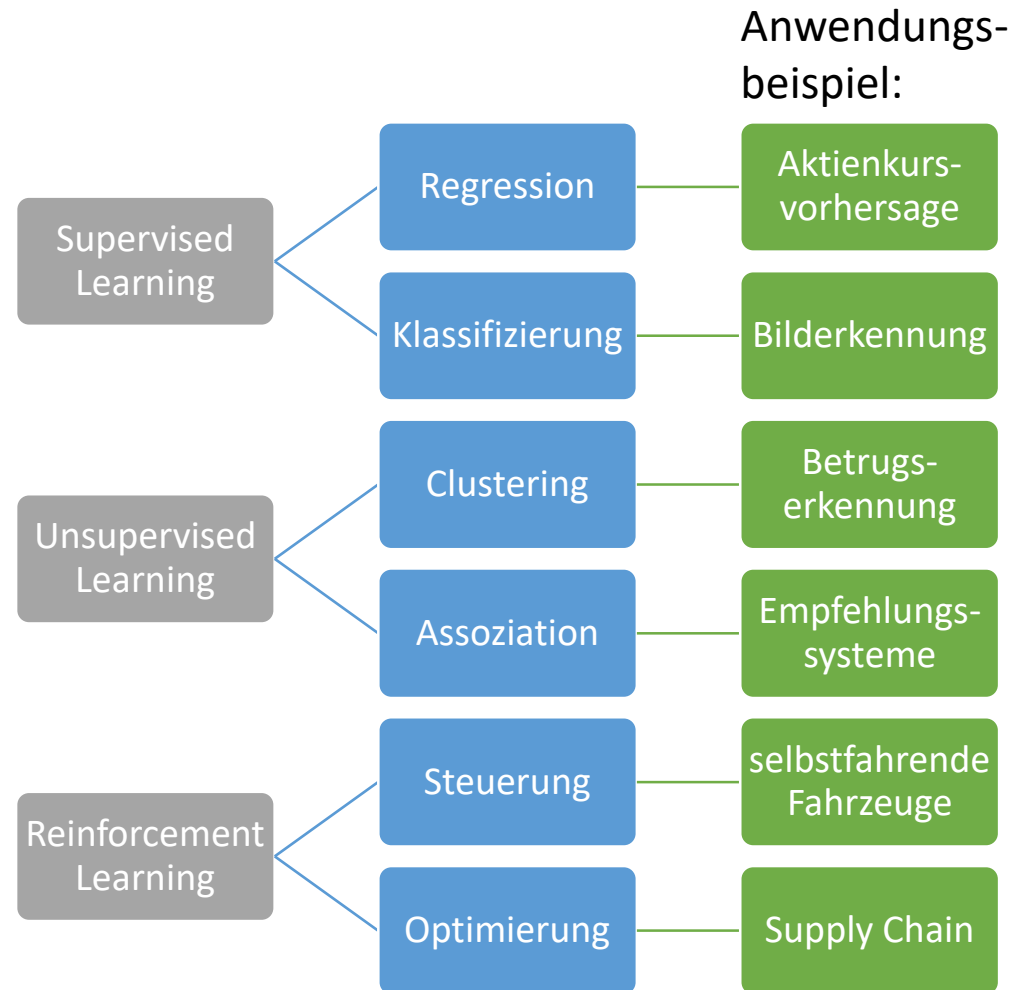
- Selbstständiger IT-Consultant, Software Engineer
- Fokusgebiete
 - Machine Learning Implementierungen
 - Webapplikationen
insb. Informationssysteme & Content Management Systeme
 - IT-Security



Künstliche Intelligenz



Machine Learning Kategorisierung



Allgemeine Möglichkeiten durch Machine Learning

- Klassifizierung
- Vorhersage von Werten
- Identifikation von und in Bild, Video, Audio
- Generierung
- Anomalie-Erkennung
- Steuerung
- ...

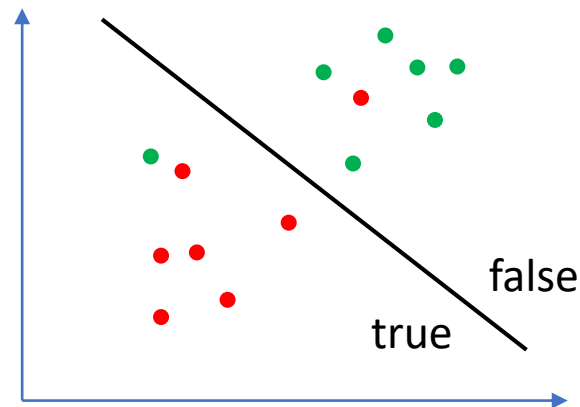


Beispiel: Phishing-Mail Klassifizierung

...Geben Sie Ihre Kontodaten auf folgender Website an...	true
...Hast du morgen Zeit für ein Treffen?...	false
...Um sich vom Newsletter abzumelden...	false
...Dringender Handlungsbedarf! Beantworten Sie...	true



geben	kontodaten	website	zeit	treffen	newsletter	abmelden	dringend	handlungsbedarf	beantworten
1	1	1	0	0	0	0	0	0	0
0	0	0	1	1	0	0	0	0	0
0	0	0	0	0	1	1	0	0	0
0	0	0	0	0	0	0	1	1	1



ähnliche Anwendungsbereiche:

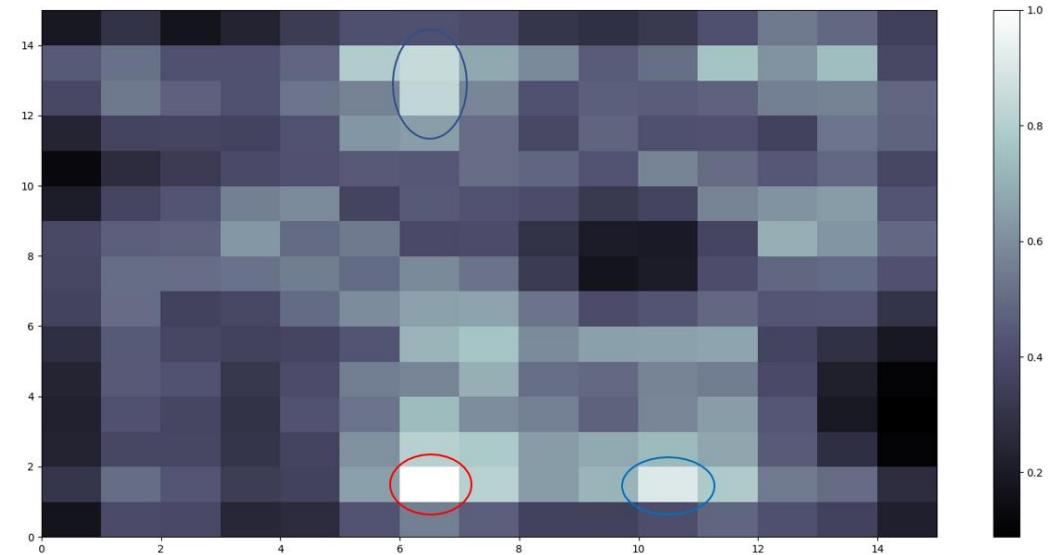
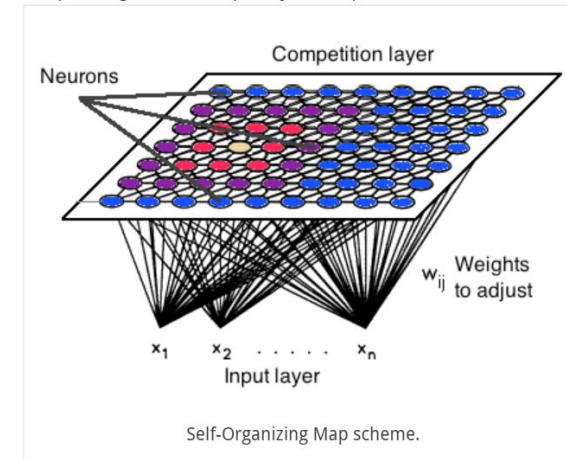
- Malware-Erkennung
- Intrusion Detection
- Botnet Datenverkehr erkennen



Beispiel: Betrugserkennung

dataset - DataFrame

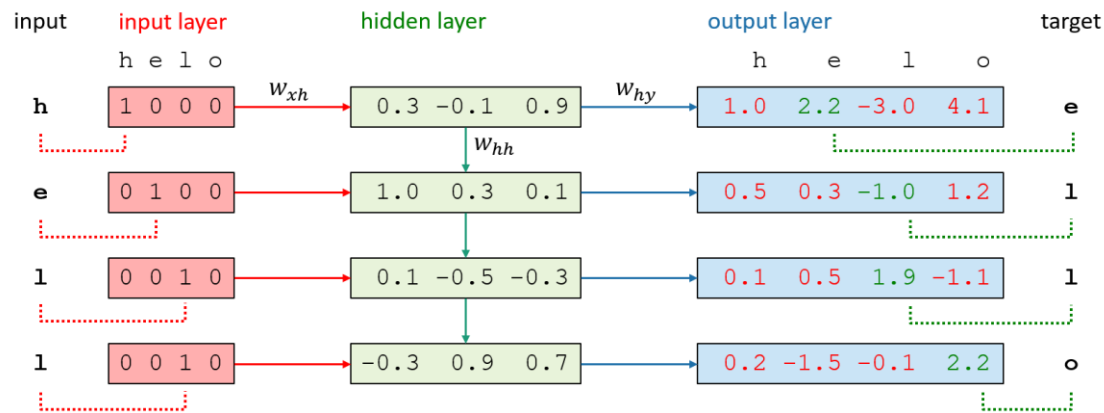
Index	CustomerID	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13	A14	Class
0	15776156	1	22.1	11.5	2	4	4	1.58	0	0	0	1	2	100	1213	0
1	15739548	0	22.7	7	2	8	4	0.165	0	0	0	0	2	160	1	0
2	15662854	0	29.6	1.75	1	4	4	1.25	0	0	0	1	2	280	1	0
3	15687688	0	21.7	11.5	1	5	3	0	1	1	11	1	2	0	1	1
4	15715750	1	20.2	8.17	2	6	4	1.96	1	1	14	0	2	60	159	1
5	15571121	0	15.8	0.585	2	8	8	1.5	1	1	2	0	2	100	1	1
6	15726466	1	17.4	6.5	2	3	4	0.125	0	0	0	0	2	60	101	0
7	15660390	0	58.7	4.46	2	11	8	3.04	1	1	6	0	2	43	561	1
8	15663942	1	27.8	1	1	2	8	3	0	0	0	0	2	176	538	0
9	15638610	0	55.8	7.08	2	4	8	6.75	1	1	3	1	2	100	51	0



Beispiel: Individuelle Passwortlisten-Generierung: Angriff

Gender	Username	First Name	Last Name	Year of Birth	Password
f	margarete	Judy	Wells	1972	Wells106
f	sondra	Lucia	Morrow	1950	cvbnm
f	zakia	Gale	Weiss	1999	syndikat
f	eada	Ana	Ellriott	1994	Ana94
f	karalee	Denise	Hanson	1965	OLIVER
m	agatha	Edmond	Daniels	1956	Agatha
...					

m	john.smith	John	Smith	2050	
---	------------	------	-------	------	--



```
John2050
180374
06031982
John!
john!
j0hn.5mi7
john.smith
Smith866
asdfghj
John50
[...]
```

Quellen: Knabl, Georg (2018). "Machine Learning-driven Password Lists". Master's thesis. FH JOANNEUM, Kapfenberg, Austria.

<http://karpathy.github.io/2015/05/21/rnn-effectiveness/>



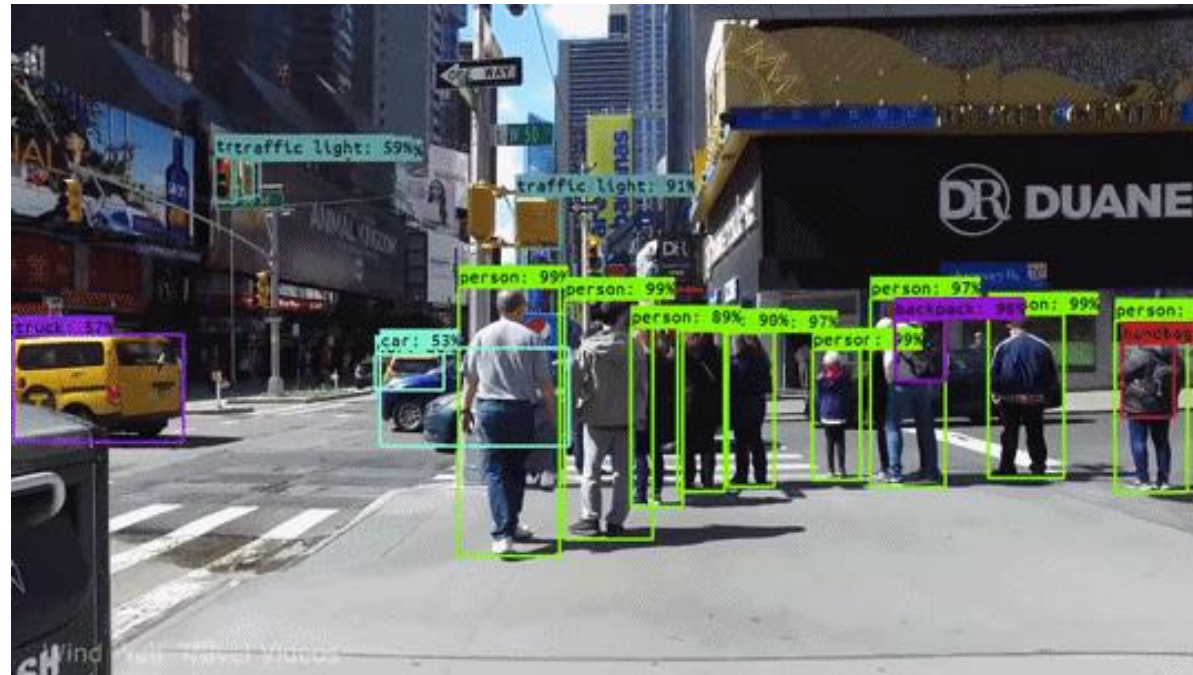
Beispiel: Individuelle Passwortlisten-Generierung: Absicherung

- Lösung: menschlich generierte Passwörter erkennen und Benutzer warnen
- Passwort-Klassifizierer
99% Trefferquote

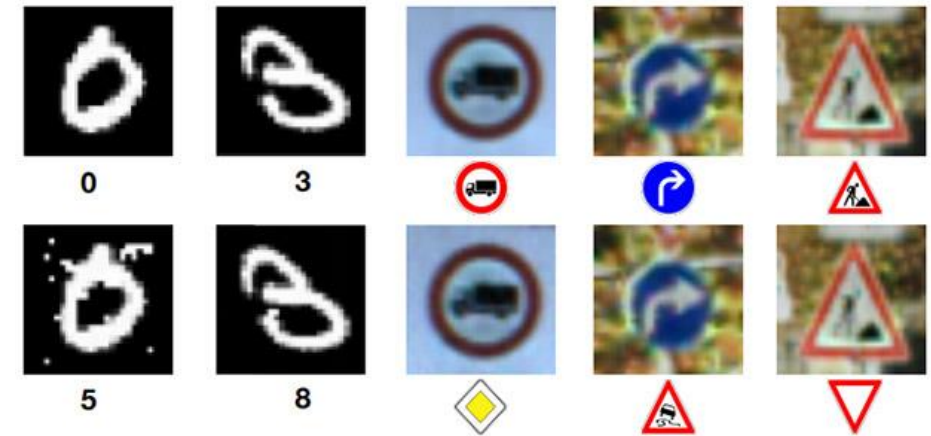
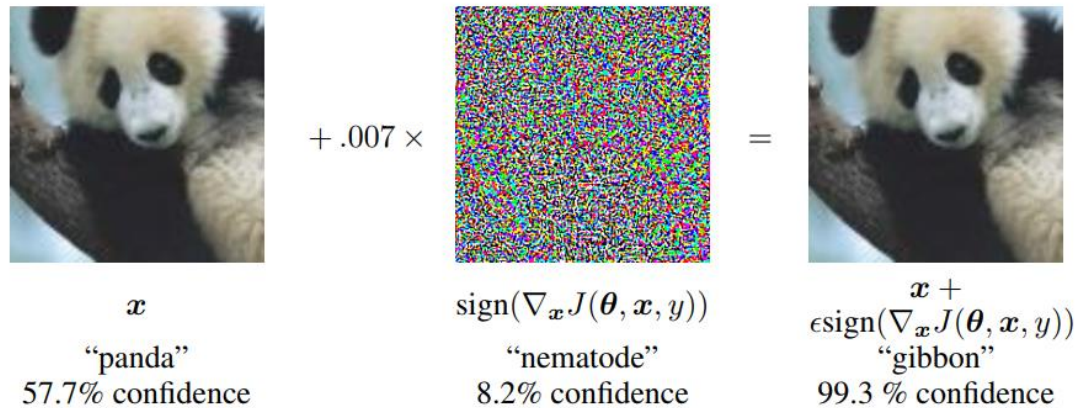
&CtAEaCp?b&v"s%	m	14061966	0.9961306540
-SUuf4TLtF	m	y-JQ6{v;_yb q	0.0000000000
mallrats	h	ZBT4n#z-x	0.0000121259
bP0.}BO/L&{:	m	longball	0.9920406811
^=c.rgH\$z	m	vikings	0.9723564484
boxers	h	gunit	0.9683620674
j&uzHCutff_\A{	m	.XP?]b3\6nP]l	0.0000000000
656565	h	8J9{Bd^	0.0000107884
6>\IB ~@4^n}K	m	123india	0.9986476258
forever1	h	*[qg;t	0.0000058089
...		...	



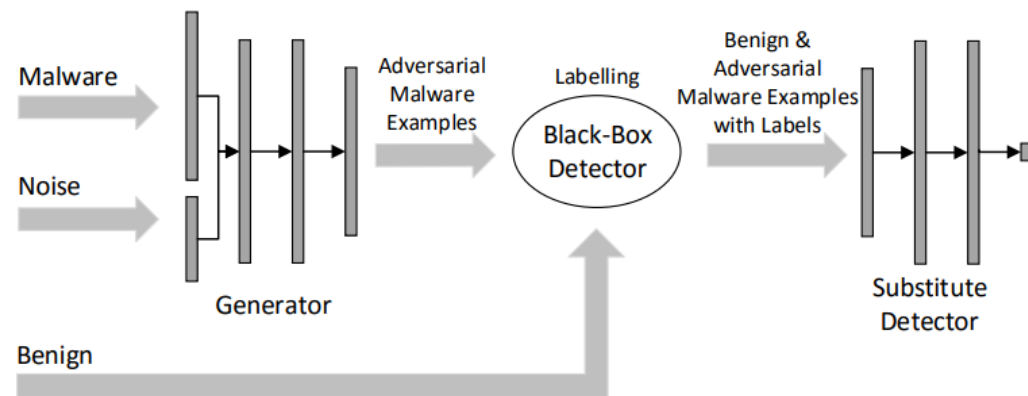
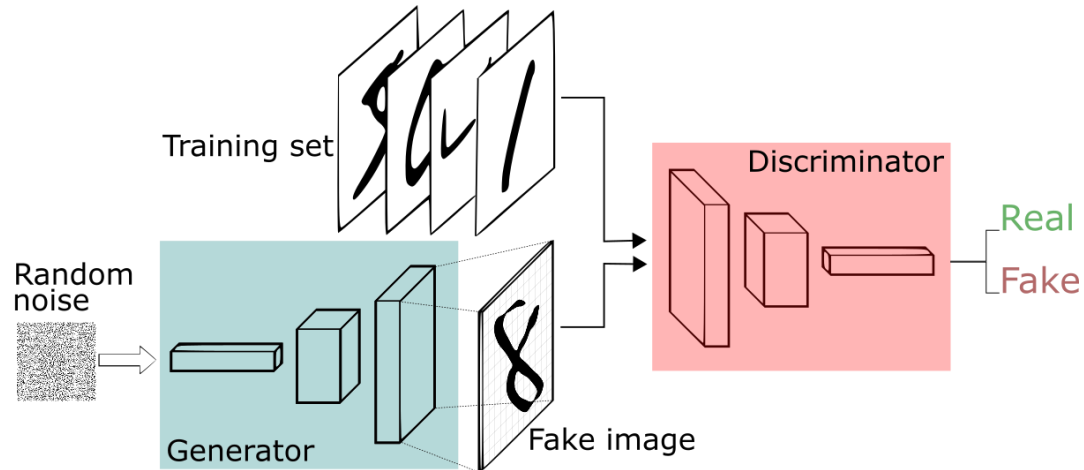
Beispiel: Objekterkennung



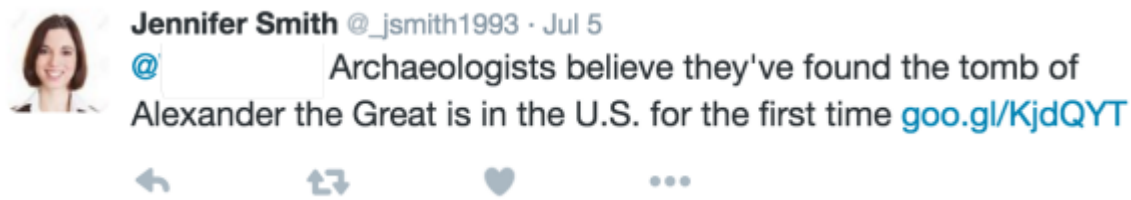
Beispiel: Bilderkennungsalgorithmen stören



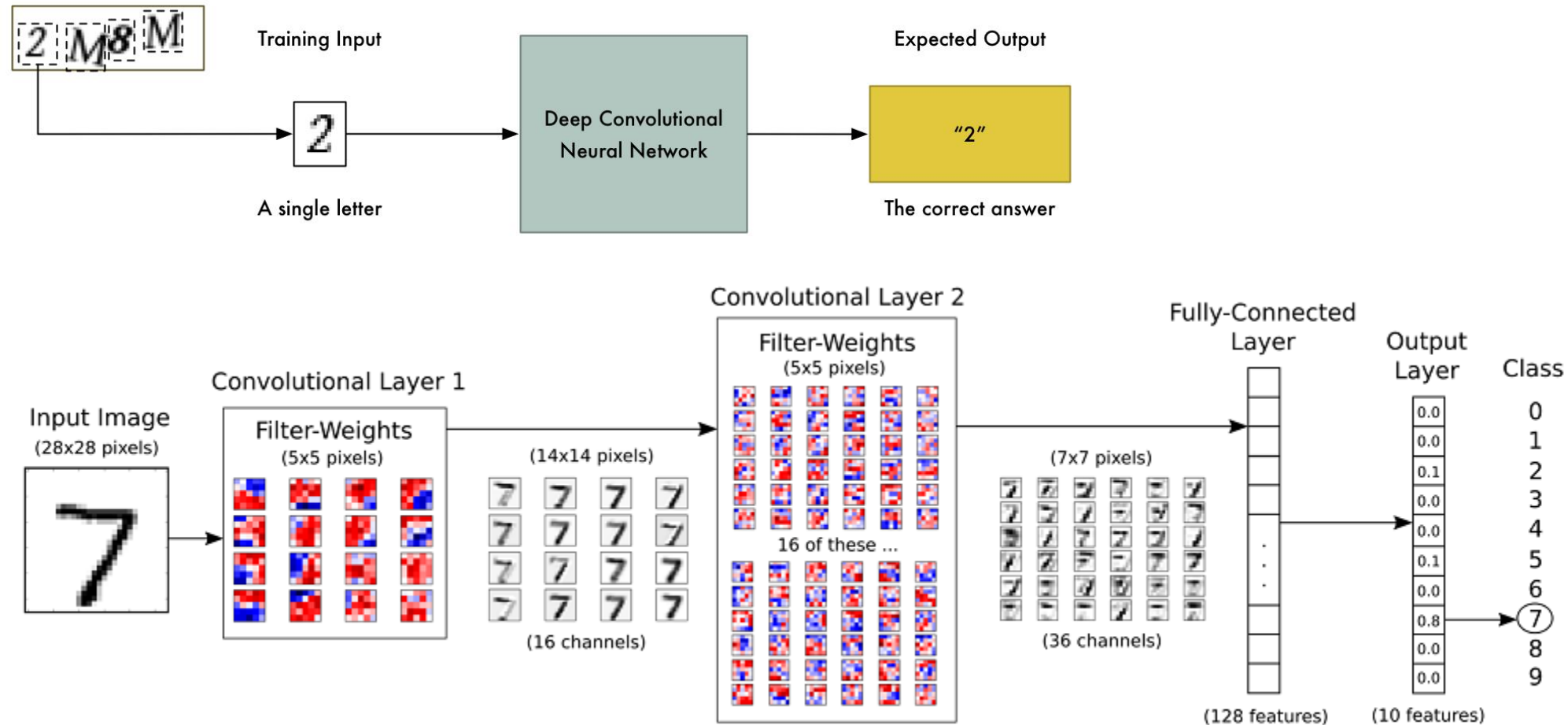
Beispiel: Erschwert erkennbare Malware durch GANs



Beispiel: Spear-Phishing mit Natural Language Processing



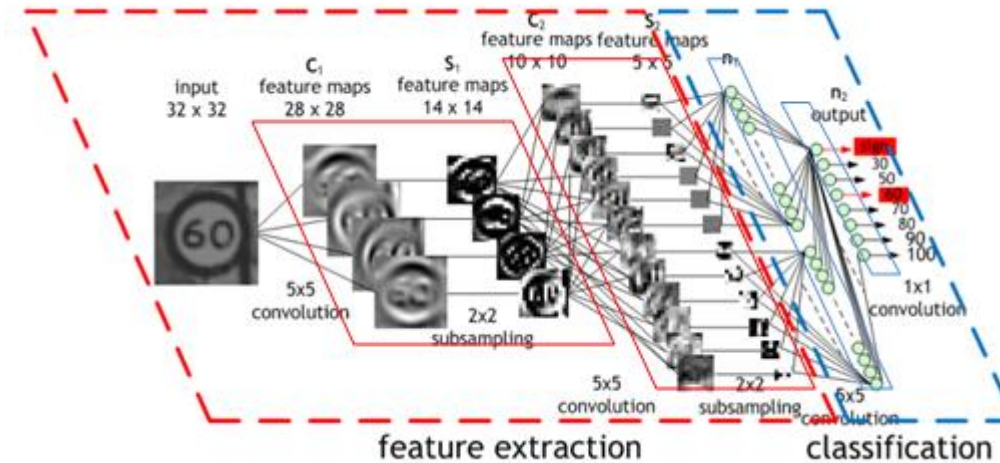
Beispiel: Klassisches CAPTCHA knacken



Beispiel: Google reCAPTCHA knacken



Figure 1: The reCaptcha widget.

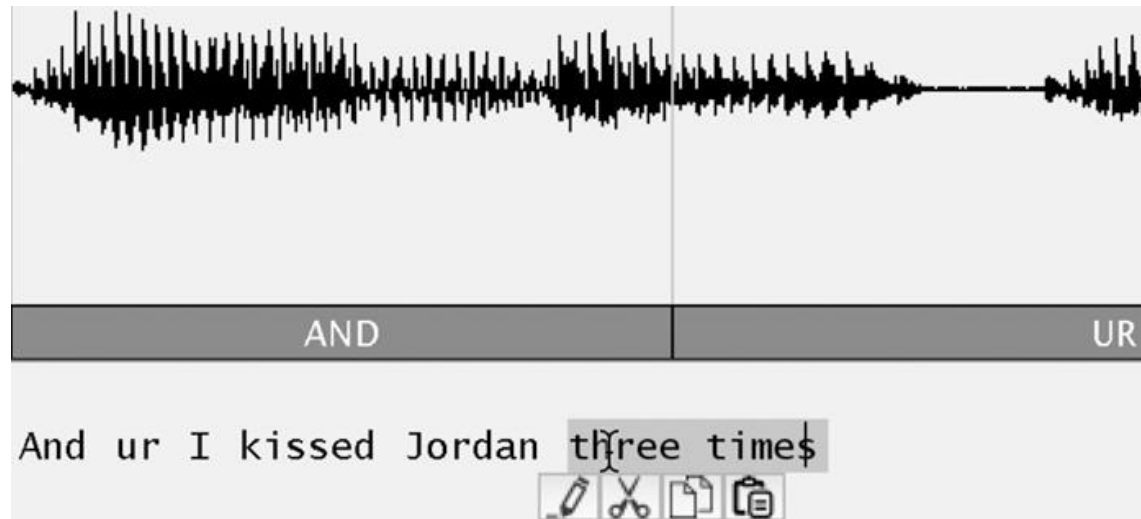


	GRIS	Alchemy	Clarifai	TDL	NeuralTalk	Caffe
	wine and blood	wine, glass	glass, red wine, wine, merlot, liquid, bottle, still, glassware, alcohol, drink, wineglass, beverage, pouring, white wine, cabernet, taste, leaded glass, dining, party, vino	red wine, goblet, wine bottle, punching bag, beer glass, perfume, balloon	a glass of wine sitting on top of a table	red wine, wine, alcohol, drug of abuse, drug, red wine, punching bag, beaker, cocktail shaker, table lamp



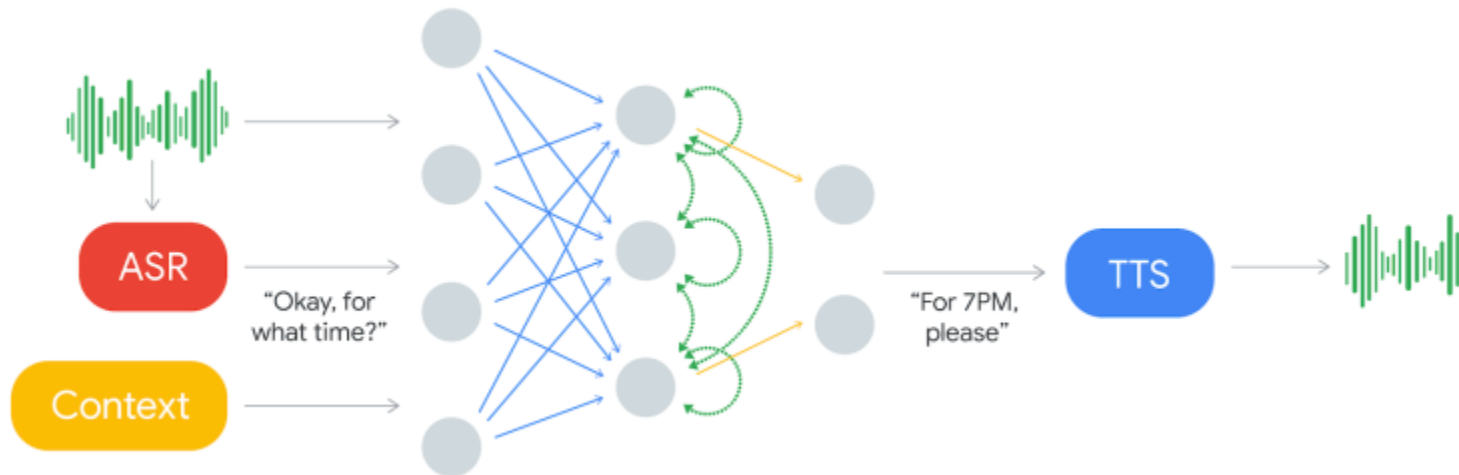
Beispiel: Stimmen immitieren

- Adobe VoCo - [YouTube](#)

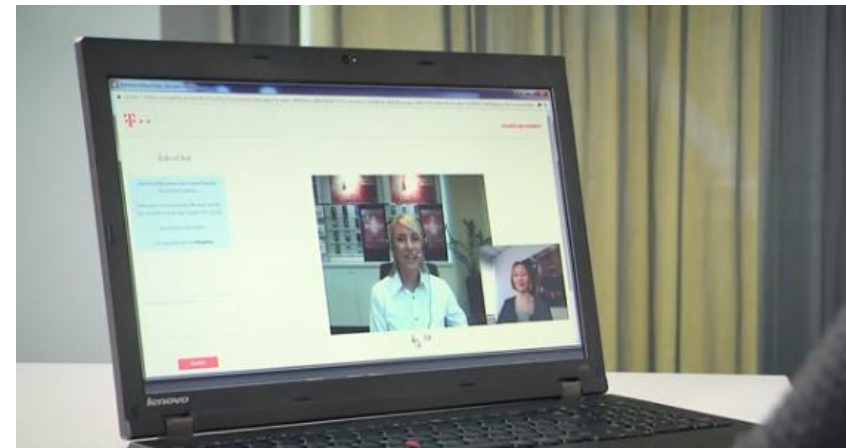


Beispiel: Autonome Chatbots

- Google Duplex - [YouTube](#)



Beispiel: Face2Face & DeepFake



Quelle: <http://hight3ch.com/face2face-real-time-face-capture-and-reenactment/>, https://mashable.com/2018/02/02/what-are-deepfakes/?europe=true#mL_5eGhtbqq3,
<https://boingboing.net/2018/01/25/deepfakes-fakeapp.html>, <https://www.telekom.de/hilfe/vertrag-meine-daten/login-daten-passwoerter/videoident>

Beispiel: Chatbots gegen Scammer

Good Day

Thank you for your response.

I can confirm that you are a listed beneficiary for fund transfer.

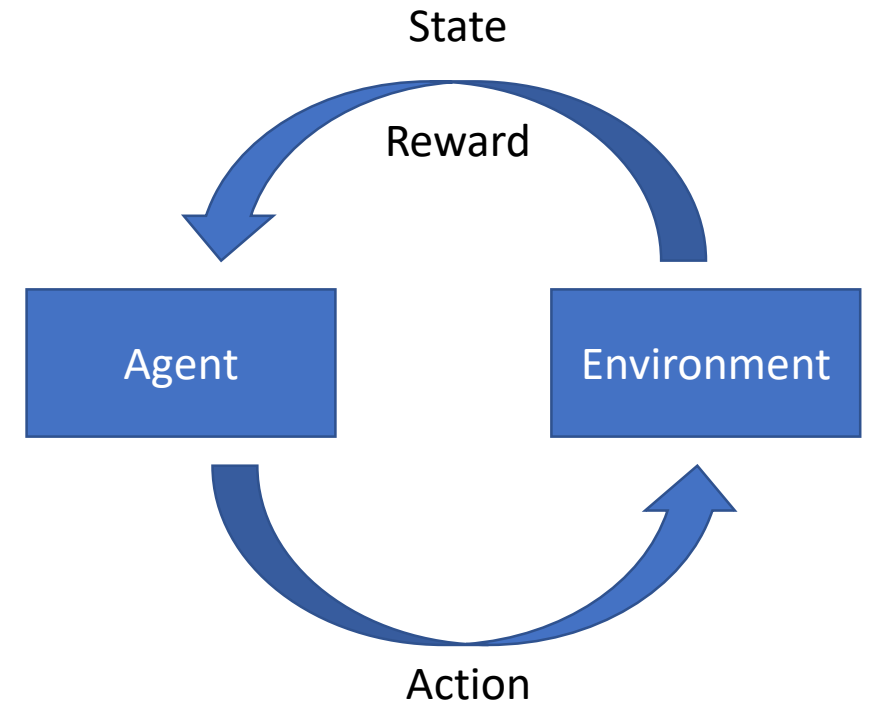
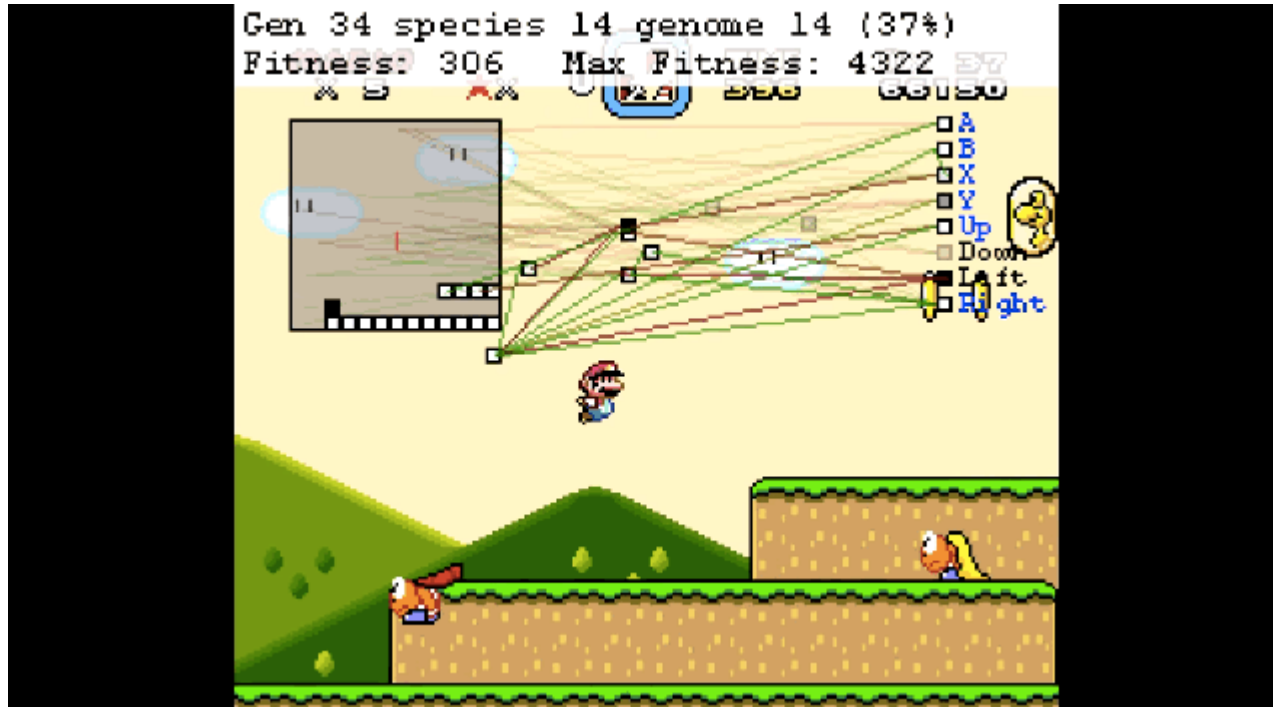
In response to the notification from the office of the Remittance Director on the payment, release and transfer of the fund with Ref. (Specified) to you after reconfirmation from you.

Hello.

Thanks for getting back to me. I'm happy to provide you with the details you requested but I'm a wee bit nervous that I might accidentally send them to someone I can't trust. Can you please confirm with me your email address so I can memorize it?



Beispiel: KI lernt Super Mario



Battle of the Algorithms

- DEF CON 24 (2016): DARPA's Cyber Grand Challenge
 - erstes **Capture-The-Flag** mit autonomen Maschinen
 - Finden & Patchen von Sicherheitslücken in Binaries
 - Angreifen anderer Maschinen
 - innerhalb von 5min:
 1. erstmaliger Kontakt mit Sicherheitslücke
 2. Analyse
 3. Patch-Erstellung
 4. Rollout



Fazit

- Technologie schon heute **einsatzbereit**
- maschineller Input **kaum** noch von menschlichem **unterscheidbar**
- technologisches **Wettrüsten** zwischen Angreifer und Verteidiger
- Machine Learning **mit Machine Learning bekämpfbar**
- Machine Learning im **Sicherheitskonzept** berücksichtigen





DI (FH) Georg Knabl, MSc

georg.knabl@pageonstage.at

+43 664 750 399 21

<https://www.linkedin.com/in/georg-knabl/>

