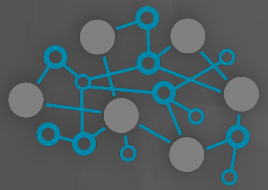


Wanneer moet je geen Deep Learning gebruiken?

25 september 2019, Gouda

Lisa Tostrams

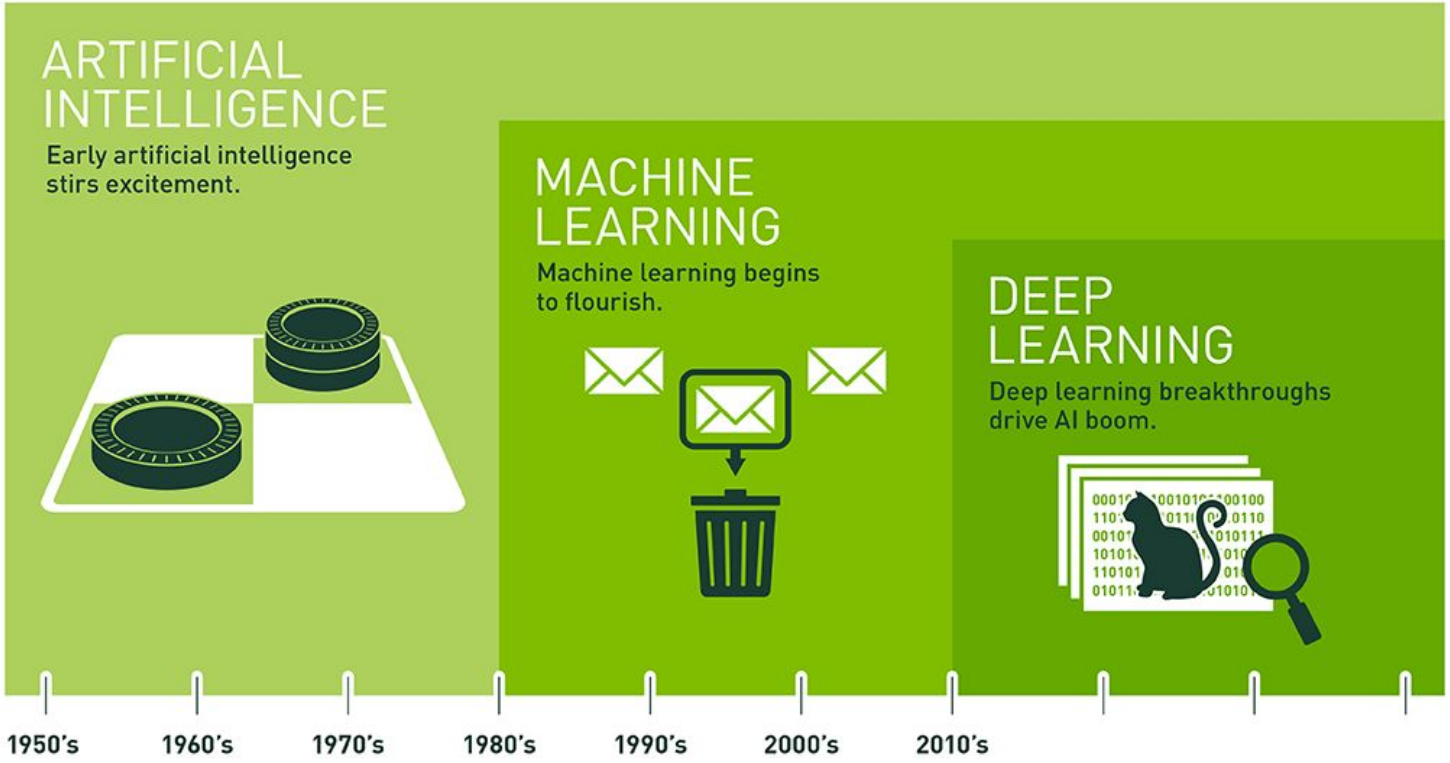
Outline

- Traditional ML
- Statistical ML
- Deep learning
- Waarom maakt het uit?
- ML in bedrijven
- Voorbeelden uit het nieuws

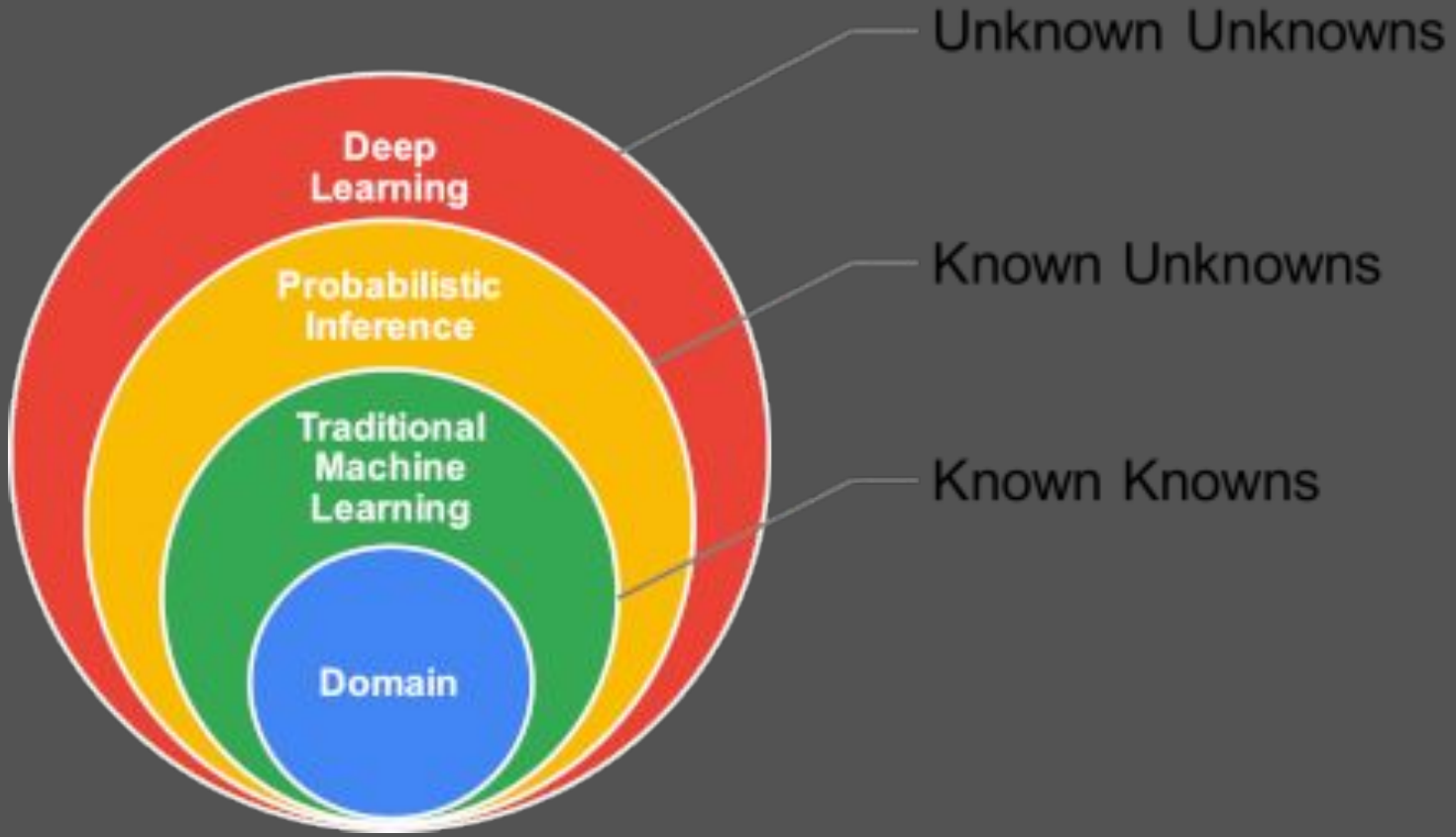
Whoami

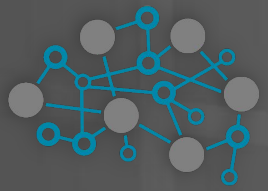
Lisa Tostrams

- Assistent bij data mining, bayesian statistics, etc. bij de Radboud Universiteit
- Machine Learning Engineer bij Machine2Learn



Since an early flush of optimism in the 1950s, smaller subsets of artificial intelligence – first machine learning, then deep learning, a subset of machine learning – have created ever larger disruptions.


Deep Learning
Probabilistic Inference
Traditional Machine Learning
Domain
Unknown Unknowns
Known Unknowns
Known Knowns

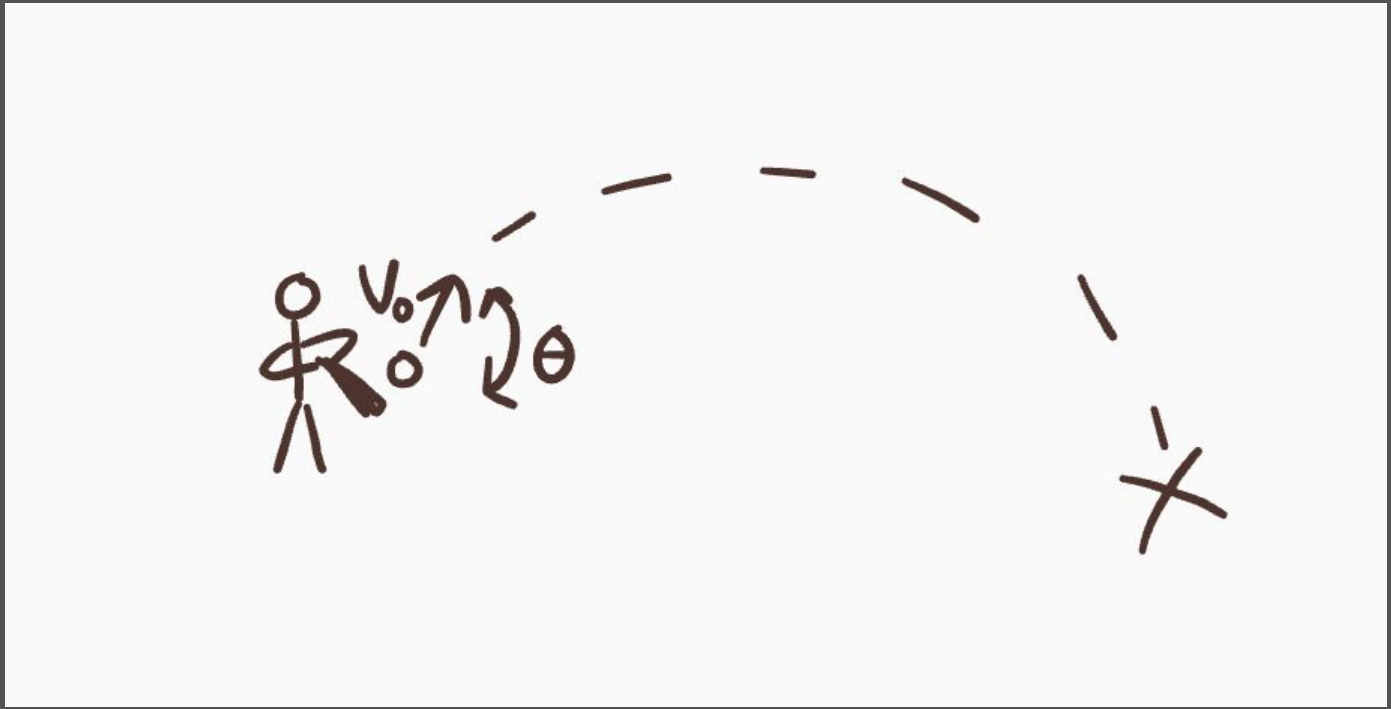


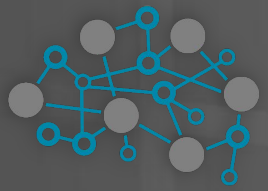
Traditional Machine Learning

- ‘Known knowns’

Trajectory of a baseball

MACHINE2LEARN

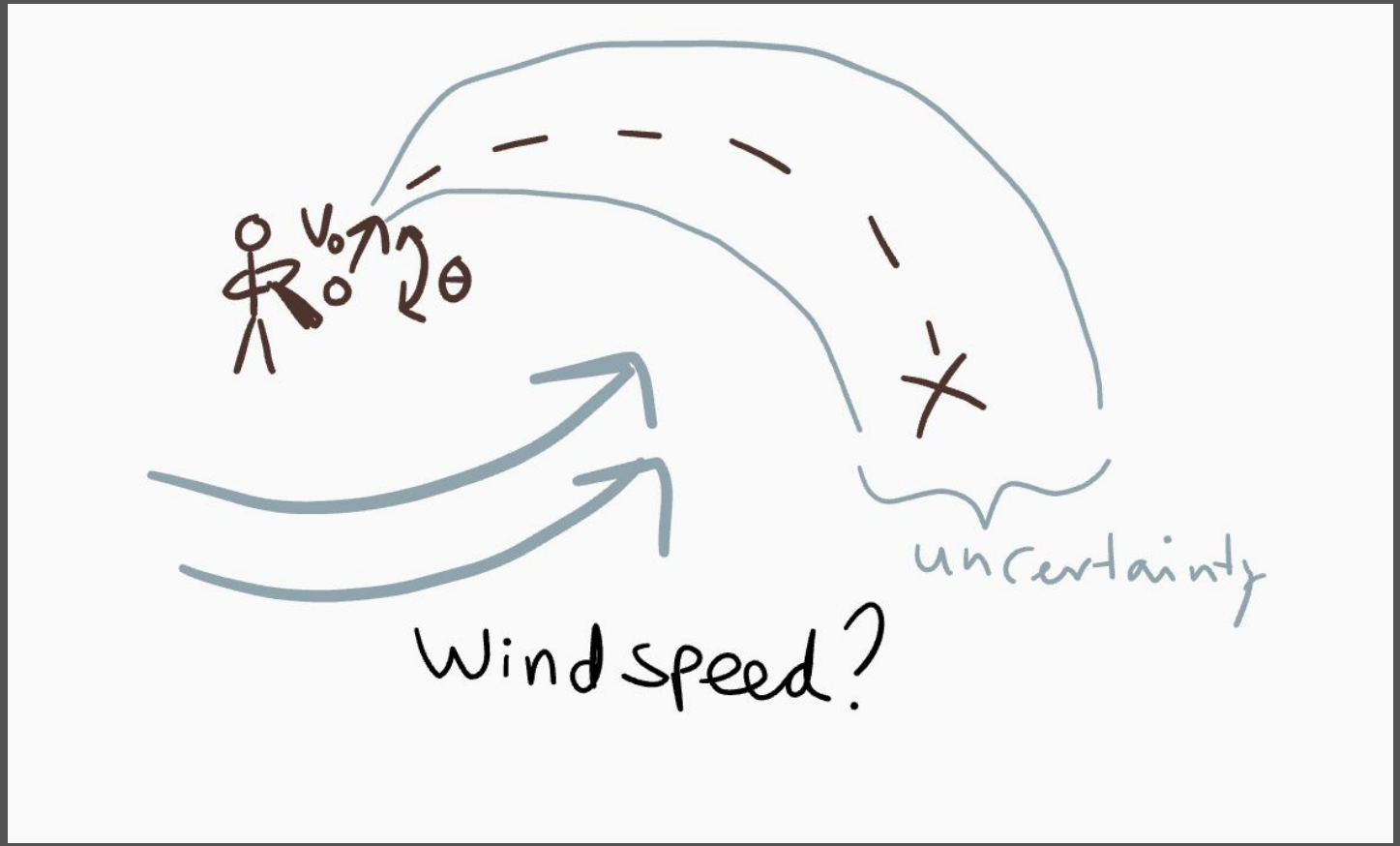


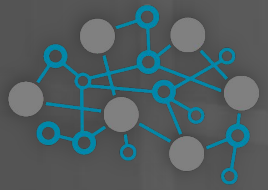


Statistische Machine Learning

- ‘Known unknowns’

Trajectory of a baseball with uncertainty



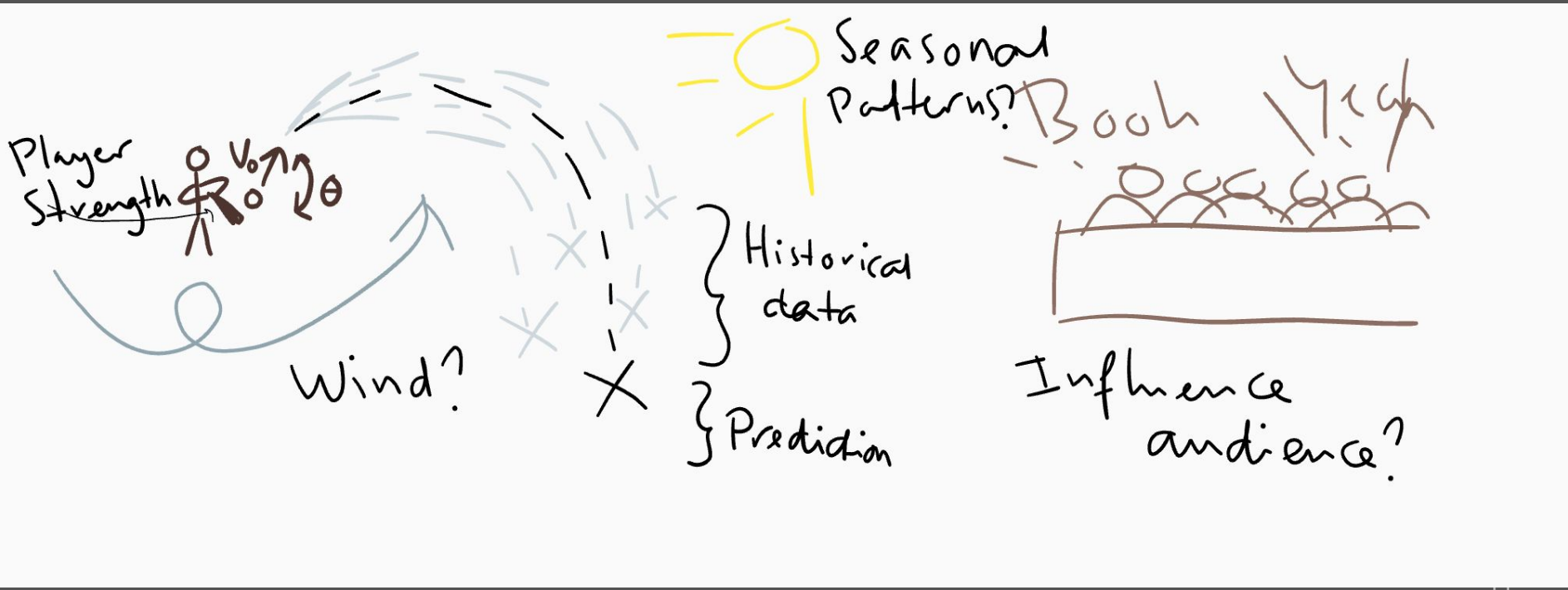


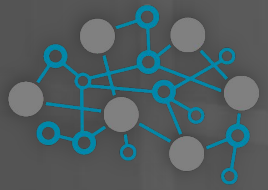
Deep learning

- ‘Unknown unknowns’

Trajectory of a baseball with unknown patterns

MACHINE2LEARN





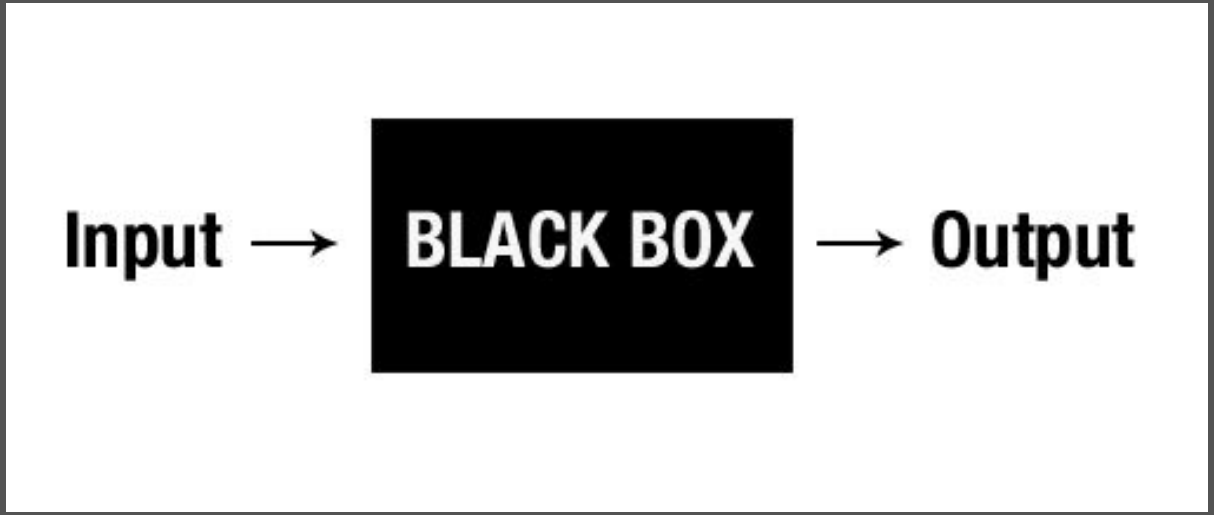
Waarom maakt dit uit?

- Black box
- ‘Defendable science’
- Niet elke vraag kan beantwoord worden door elk type ML

Black box AI

MACHINE2LEARN

‘We are drowning in information and starving for knowledge’ – John Naisbitt



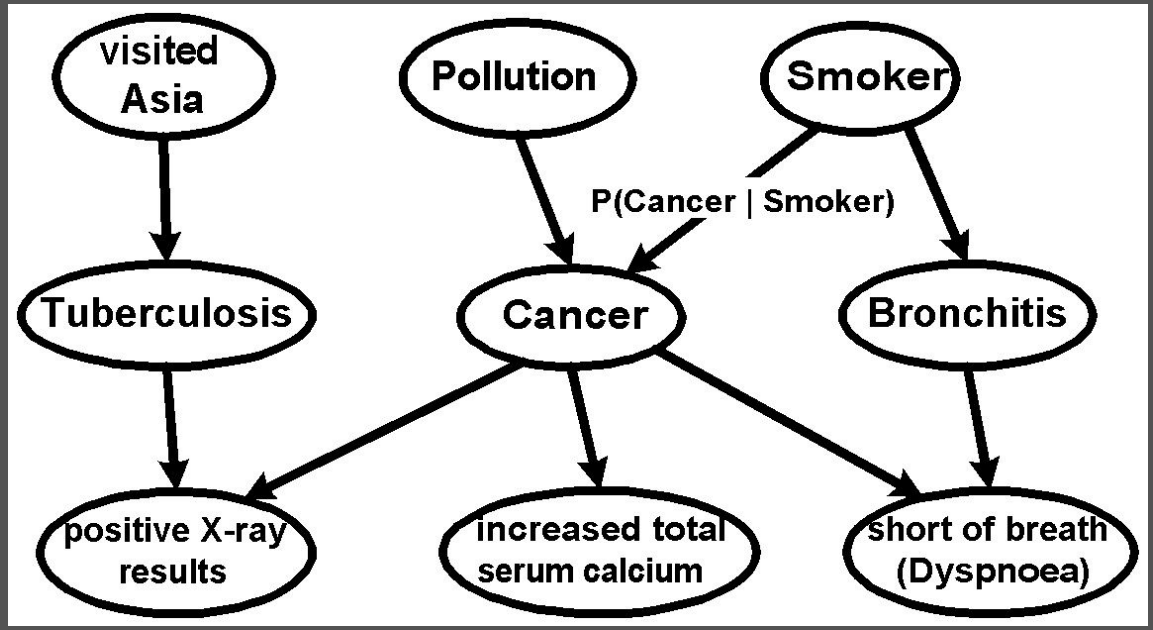
Defendable science

Welke vraag wil je beantwoord hebben?

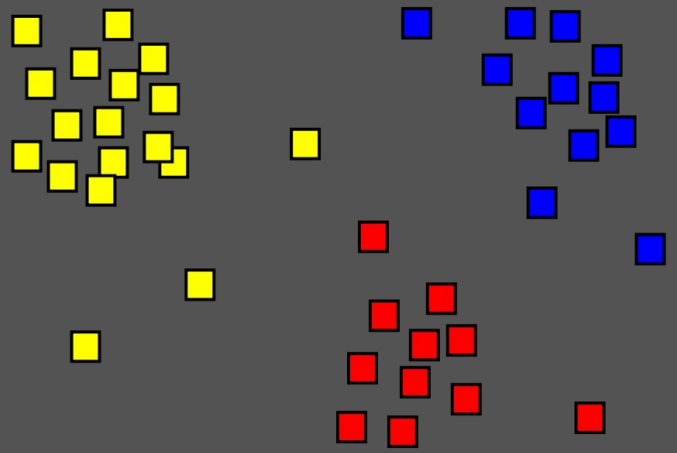
- Wil je de beste voorspelling doen op basis van alle mogelijke beschikbare data?
- Wil je weten welke factoren een process beinvloeden?

- GDPR
- Krediet scores

Example 1: Causal modelling in medical diagnoses

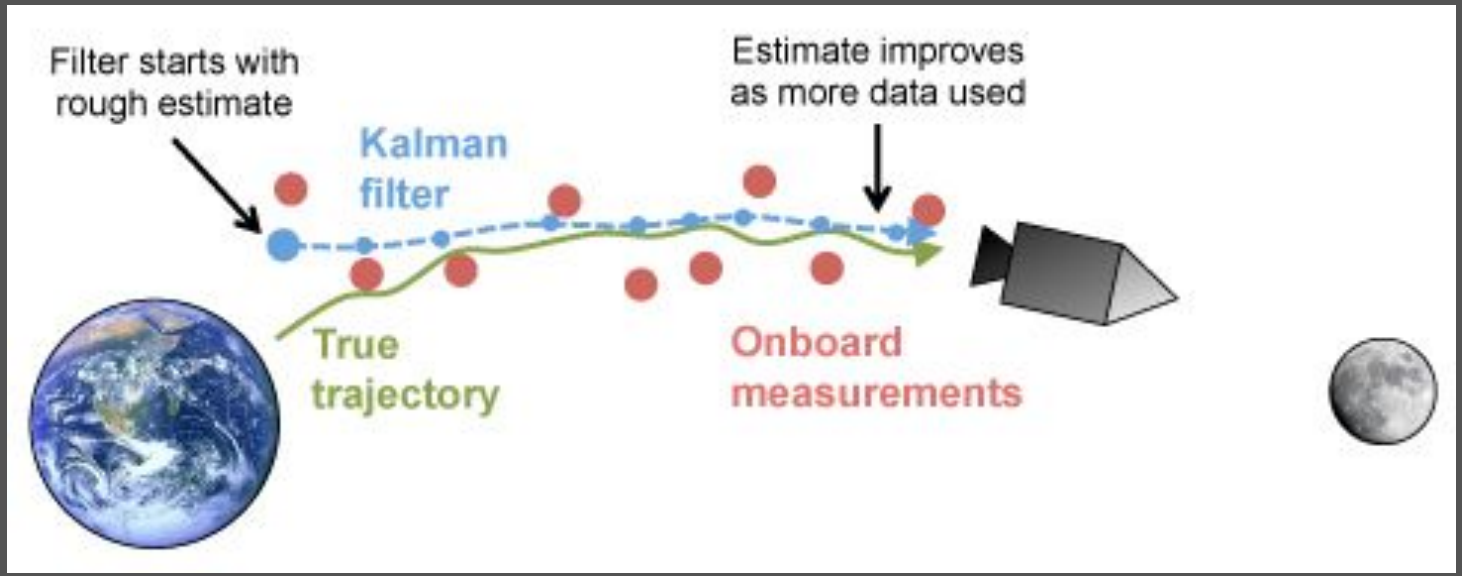


Example 2: Bayesian inference in clustering



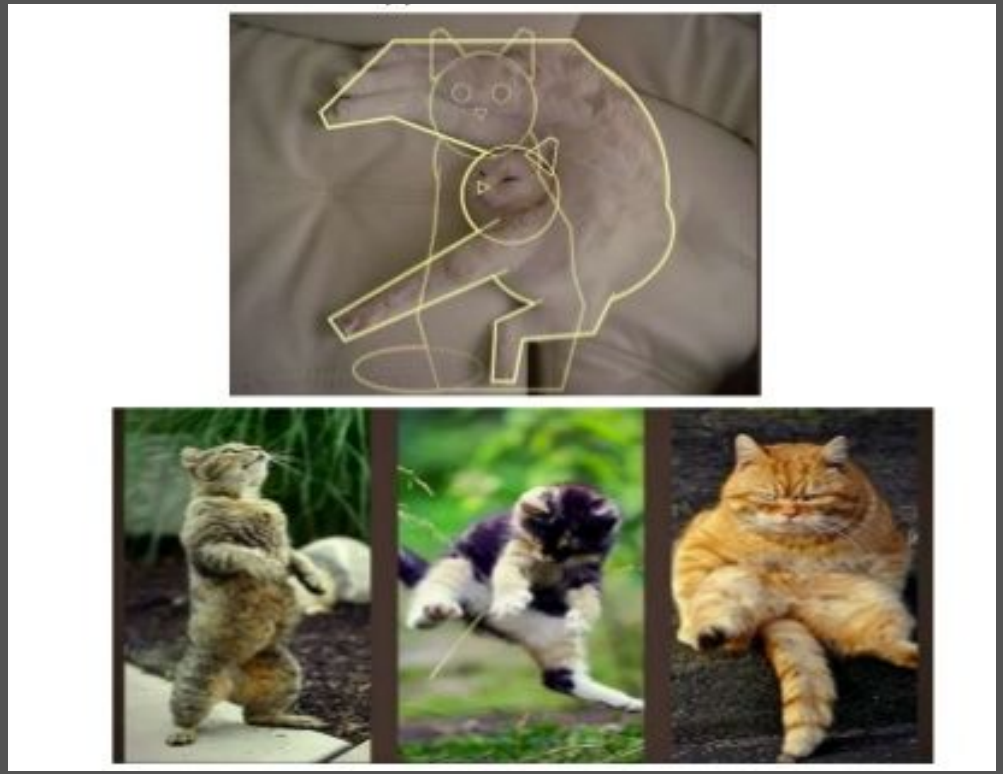
Example 3: Physics/inference in state-space

MACHINE2LEARN



Example 4: CNN's in computer vision

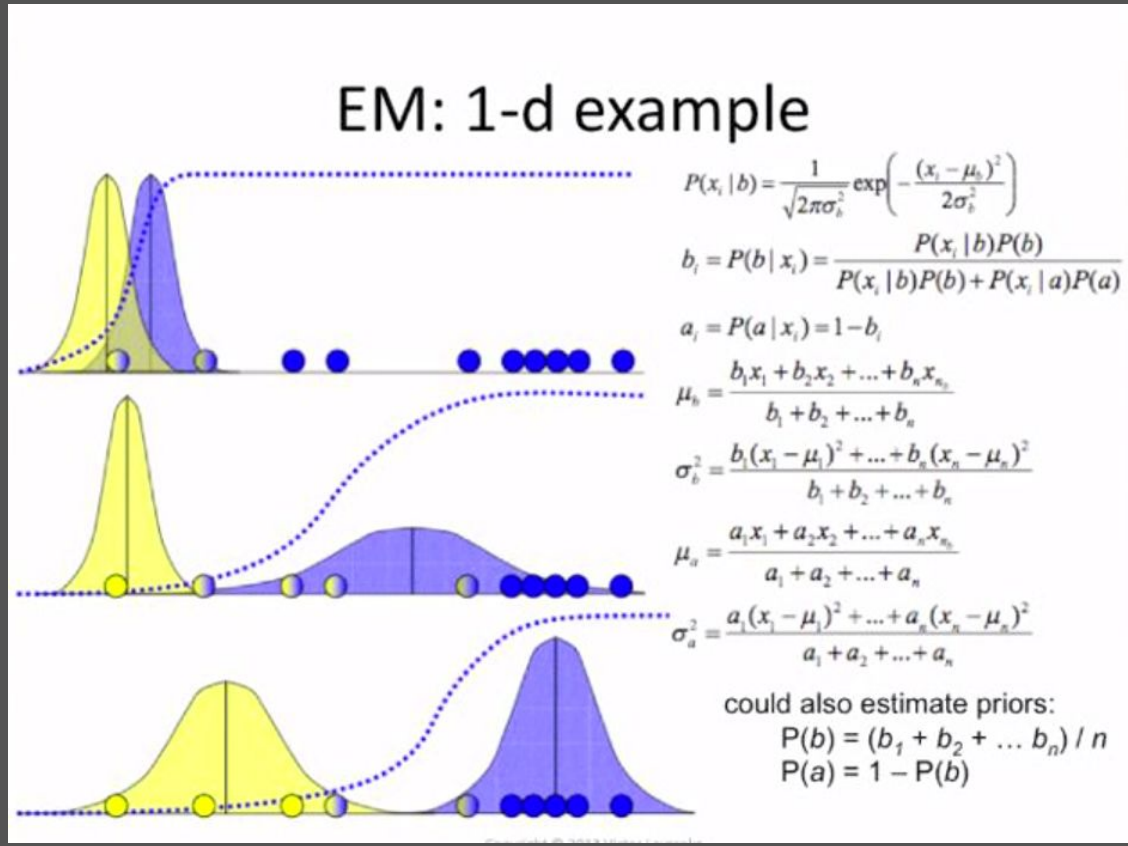
MACHINE2LEARN

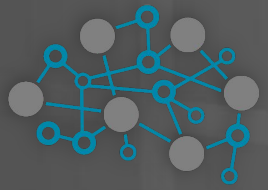


Example 5: Hybrid learning

MACHINE2LEARN

- Variational Autoencoders
- Deep reinforcement learning





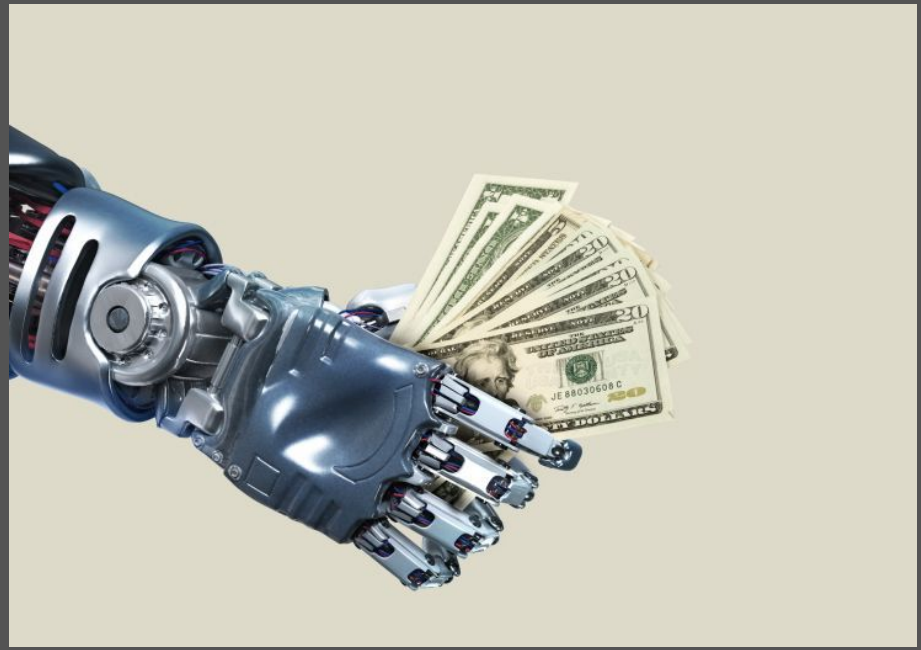
Machine learning in bedrijven

- Waar begin je?
- Wat kan jij doen?
- Wat kan jouw bedrijf doen?

Zo veel AI apps om te maken, zo weinig tijd..

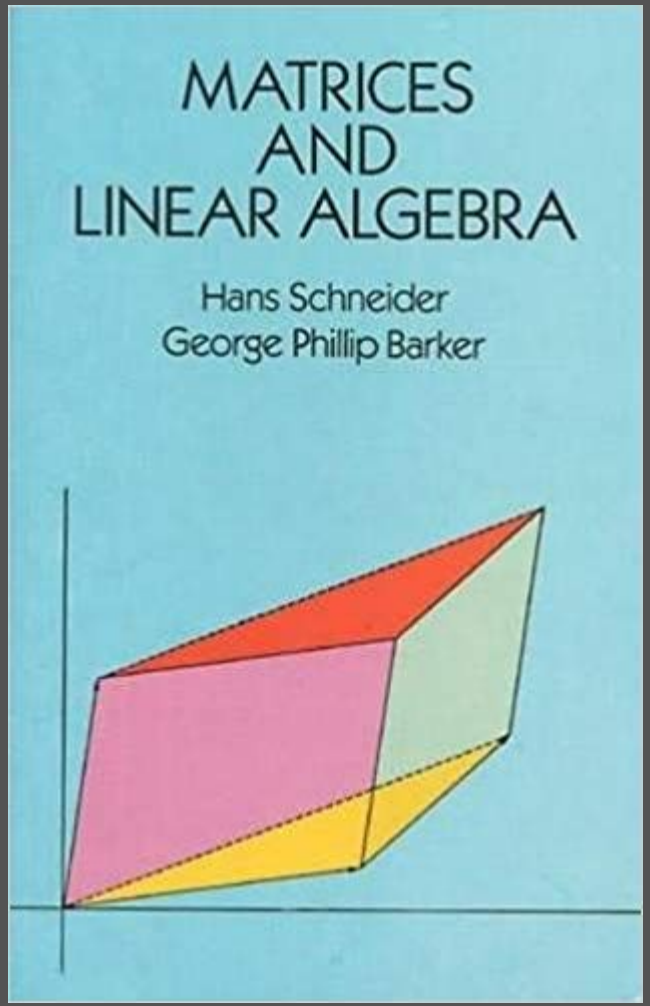
MACHINE2LEARN

- Everybody has heard fantastic predictions about machine learning and AI adding billions in economic value



Wat kan jij doen?

https://www.coursera.org/learn/machine-learning

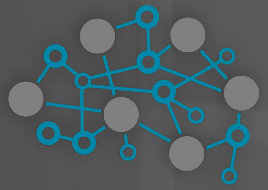


Wat kan jouw bedrijf doen?

MACHINE2LEARN

https://landing.ai/ai-transformation-playbook/

Voorbeelden uit het nieuws



MACHINE2LEARN

Computer says no

MACHINE2LEARN

- In China krijg je automatisch een boete voor jaywalking
- Je gezicht verschijnt daarnaast op een scherm als vorm van public shaming
- Wel handig als systeem het personen van billboards onderscheidt…

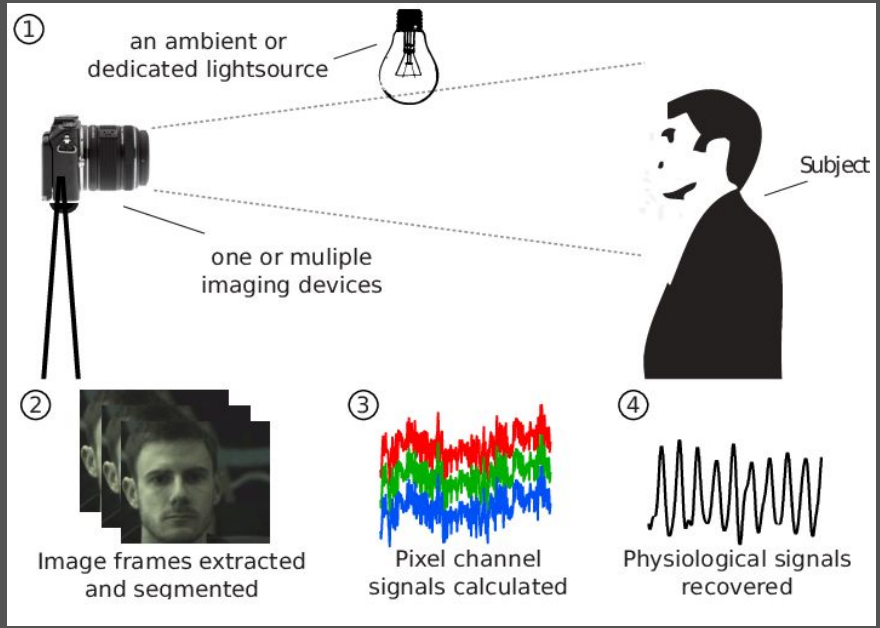
Lessons

MACHINE2LEARN

- Hoe kan je aanwezige technologie gebruiken om het systeem accurater te maken?
- Bij biometrische classificatie -> “proof of life”
- “PPG” is haalbaar met een smartphone sensor

https://www.extremetech.com/computing/159309-mit-researchers-measure-your-pulse-detect-heart-abnormalities-with-smartphone-camera

https://www.youtube.com/watch?v=Q9MK-vtWzUM



Be careful what you break

MACHINE2LEARN

“Uber’s autonomous mode disables Volvo’s factory-installed automatic emergency braking system”

Lessons

MACHINE2LEARN

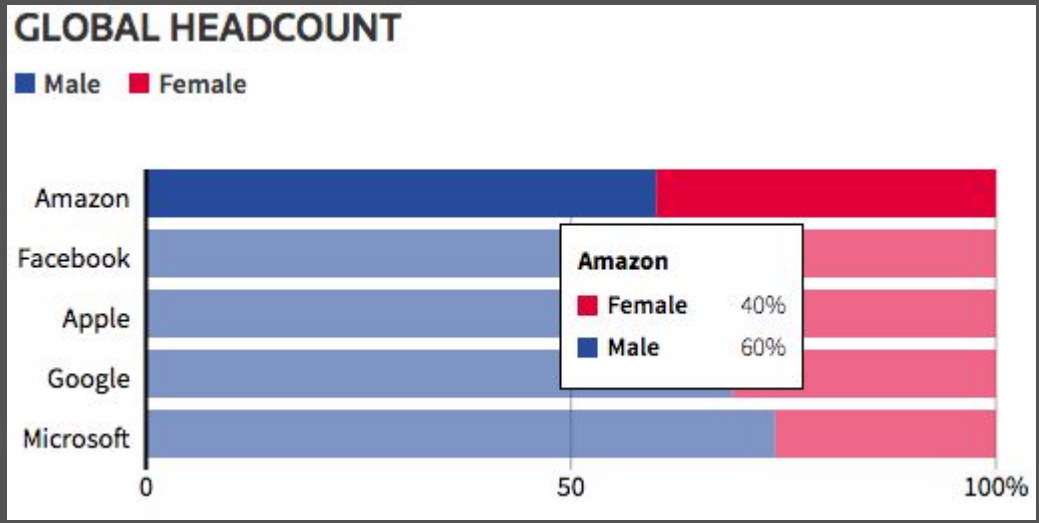
- Hoe werkt je systeem samen met niet-intelligente technologie?
- Wat zijn de consequenties van een enkele fout? Er zat een “human safety driver” achter het stuur
- Bijna een jaar gestopt met zelfrijdende auto’s op publieke wegen
- Perceptie van mensen die schadelijk fouten maken is compleet anders

Amazon recruitment tool

MACHINE2LEARN

...the technology favored candidates who described themselves using verbs more commonly found on male engineers' resumes, such as "executed" and "captured," one person said.

Amazon's computer models were trained to vet applicants by observing patterns in resumes submitted to the company over a 10-year period. Most came from men, a reflection of male dominance across the tech industry.



Elementary dear Watson

- IBM Watson is ontwikkeld om patiënten te helpen bij Sloan Kettering
- Interacties tussen verschillende medicijnen moest nog een beetje worden afgesteld

> In one instance during testing, a 65-year-old patient with severe bleeding was diagnosed with a drug that may have led to a “severe or fatal haemorrhage”.

Lessons

- Modellen leren de biases van de dataset. Duh.
- Domeinkennis van experts is cruciaal. Diversifieer je team van data scientists waar mogelijk.
- Voor grote beslissingen, laat machines voorlopig mensen informeren i.p.v. vervangen

Amazon edited the programs to make them neutral to these particular terms. **But that was no guarantee that the machines would not devise other ways of sorting candidates that could prove discriminatory, the people said.**

dus…

- Modelleren helpt. Voor NLP en andere perceptie van patronen is dit echter lastig, grondige validatie is essentieel.

Alexa likes to party

MACHINE2LEARN

- Oliver Haberstroh gaat stappen in Hamburg
- Om 01:50 zet Alexa hard muziek aan
- Buren bellen politie om het “feestje” te stoppen
- Politie krijgt geen gehoor en forceert de deur
- Oliver moet zijn nieuwe sleutels op het politiebureau ophalen
- ...en de slotenmaker betalen

Lessons

MACHINE2LEARN

- Wat is het maximale volume waarop de muziek mag staan zonder extra input?
- Aangesloten speakers variëren in volume -> gebruik de microfoon voor normalisatie
- Hoe zeker moet Alexa er van zijn dat er überhaupt iemand thuis is?
- Wie is juridisch aansprakelijk voor de financiële schade?

Have the robot watch the kids

MACHINE2LEARN

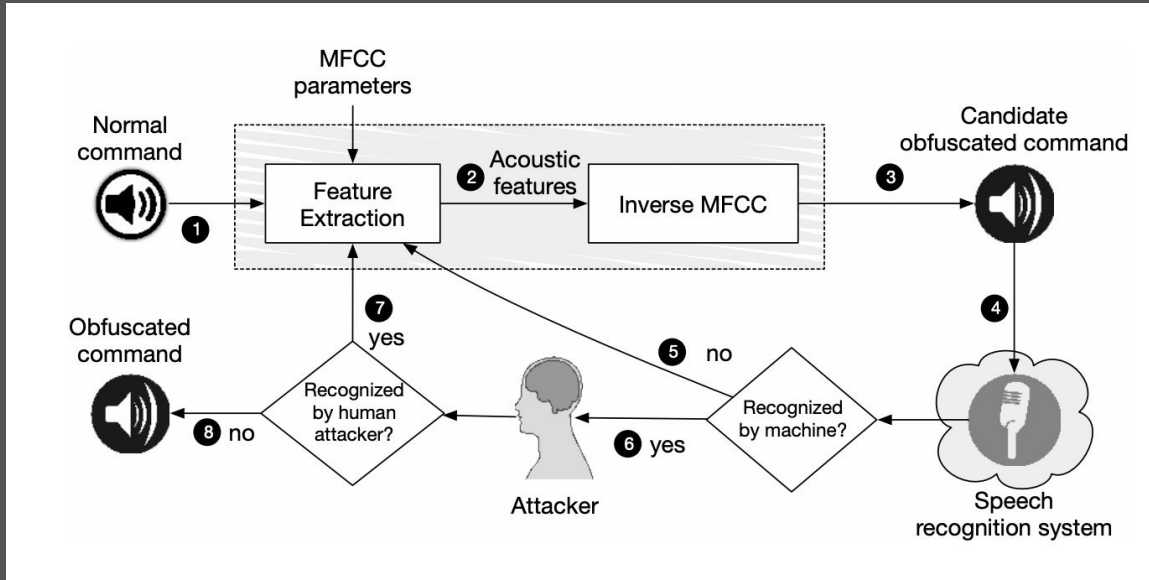
- 6 jaar oud meisje bestelt met Alexa een poppenhuis van $170 en vier kilo koekjes
- Moeder doneert het poppenhuis en zet een kinderslot op Alexa
- Nieuwslezer Jim Patton: “I love the little girl saying, ‘Alexa ordered me a dollhouse,’”
- Meerdere Alexa’s in de huizen van kijkers proberen een poppenhuis te bestellen

Lessons

MACHINE2LEARN

- Hoeveel mag het systeem zelf doen zonder rigoureuze verificatie?
- Wordt er bij financieel belang goed threat modelling gedaan voor adversarial input?
- Wanneer tot de eerste Facebook video's verschijnen die bij iedereen producten bestellen?

https://nicholas.carlini.com/papers/2016_usenix_hiddenvoicecommands.pdf



Be careful what you learn from

MACHINE2LEARN

- Microsoft publiceert een Twitter bot die leert van reacties van andere Twitteraars
- 4Chan komt er achter en gaat haatdragende gesprekken met de bot houden

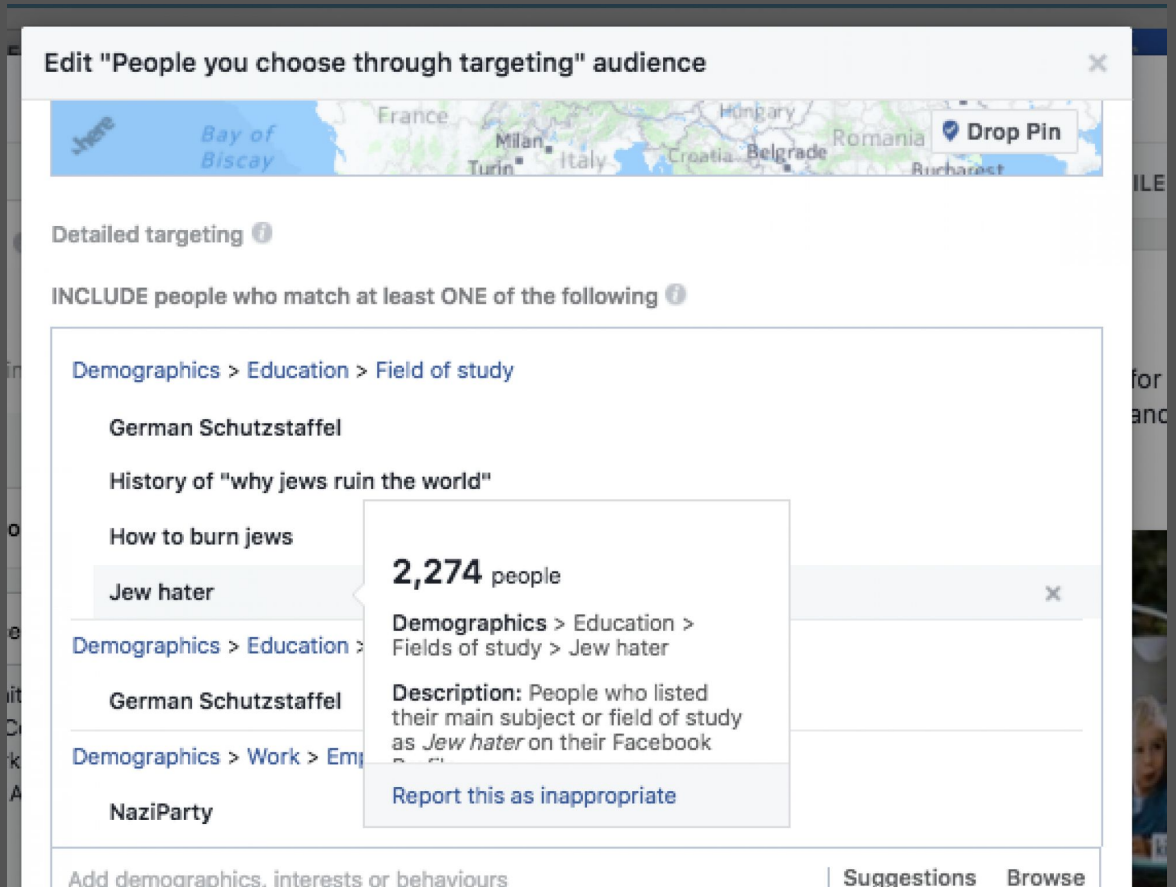
Lessons

MACHINE2LEARN

- Threat model: kan het grote boze internet vertrouwd worden om nette input te geven?
- Tuurlijk niet!
- Wat is de laatste keer dat communities zoals 4Chan een kans voorbij heeft laten gaan om iets te verpesten met “humor”?
- Some people just want to watch the world burn

Facebook Ads allow me to target what?

MACHINE2LEARN



Lessons

- Tot systemen zelf weten wat politiek correct is, output door mensen laten verifiëren

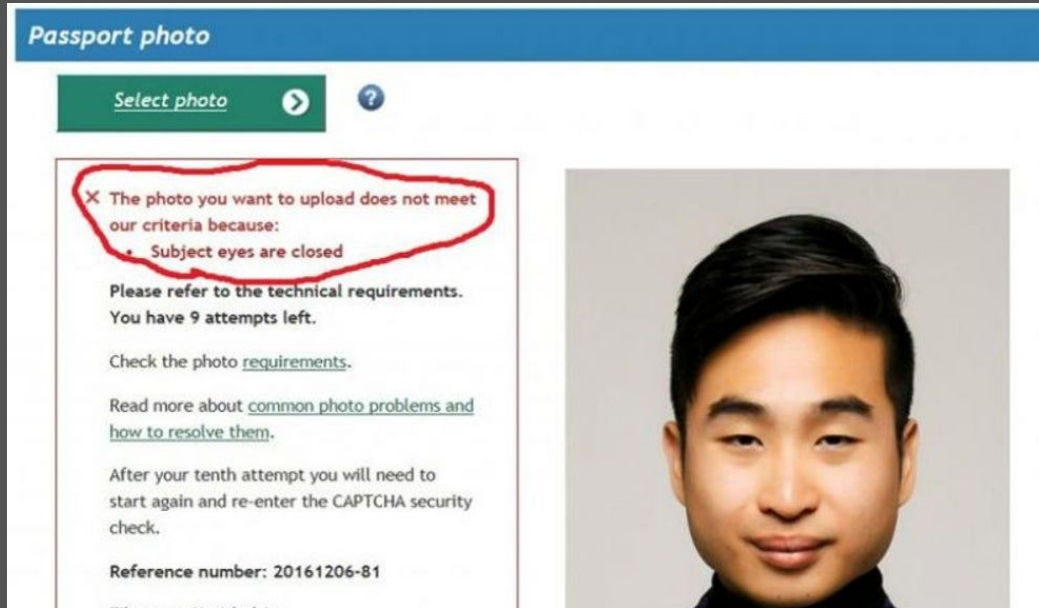
- Machine learning kan het gros van het werk wegnemen
- ...maar sommige dingen blijven mensenwerk

- Dat het technisch een geldige categorie is, betekent niet dat het iets is waar je bedrijf voor staat

I think it’s the model that needs to open its eyes

MACHINE2LEARN

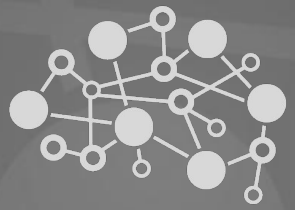
- Uploaden van pasfoto’s of maken van kiekjes; veel fotosystemen hebben ML ingebouwd
- Verificatie kan kritiek hebben die nogal racistisch klinkt



Lessons

MACHINE2LEARN

- Verificatie gebeurt door een model wat aannemelijk Landmark Detection doet
- Threshold voor landmarks van ogen is waarschijnlijk niet gekalibreerd voor verschillende etniciteiten
- Ga brainstormen over hoe output ongevoelig over kan komen
- *Systemen staan nooit los van de sociale en culturele omgeving waarin ze worden gedeployed*



MACHINE2LEARN

Vragen?