

# Algoritmes 28 november 2019 HCC

Waar vind ik zelflerende algoritmes?

- => Helsinki: [elementsofai.com](http://elementsofai.com)
  - nu ook in NL: [app.ai-cursus.nl/home](http://app.ai-cursus.nl/home)
- Dan naar Silicon Valley, Stanford University:

Coursera

- [coursera.org/learn/machine-learning](http://coursera.org/learn/machine-learning)
- Nu ook: [coursera.org/learn/ai-for-everyone](http://coursera.org/learn/ai-for-everyone)
- 

Ongelofelijk...

- Algoritme zoekt zelf uit wat de correlatie is..
- Door met alle voorbeelden veel te rekenen
- Resultaat: voorspelling die 90, of 95, of 99%
- En elk nieuw voorbeeld verbetert de uitkomst

Tekst als input,

bijvoorbeeld spam

- Vroeger: zelf woorden kiezen zoals Viagra, sex, engelse tekst, etc.
- Nu: voorbeelden spam berichten
  - geen motivatie nodig –

Je hoeft zelf niet te zeggen waarom je een mail verdacht vindt....

Hoe werkt het spam algoritme?

- In elke mail is elk woord een kenmerk, correlatie met elk ander woord wordt uitgerekend.
  - bijv. 2000 meest gebruikte woorden
  - dat zijn 4 miljoen kruisverbanden
- 5000 voorbeelden nodig “to train the algorithm”
- Nauwkeurigheid wordt 95% (beter dan mensen)

Meer voorbeelden met tekst als input

- Medisch
  - Psychiatrie: voorspellen agressie, suicide
  - Intensive Care: risico heropname, gebruik antibiotica
  - Kans op zwangerschap diabetes
- Rechtspraak (?), politie
- Wetenschappelijke artikelen
- Fraude detectie
- Solliciteren
  - welke voorbeeld brieven zijn gebruikt om het algoritme te trainen?
- Beurs: basis = persberichten

*Byass = vooringenomenheid*

## Voorbeelden patroonherkenning

- Zelfrijdende auto's
- CT en Röntgen scans
- Foto's van tumors
- Gezichtsherkenning
- Ongewone bewegingen – bewaking
- Instagram

## Voice - stemherkenning

- Natural Language Processing zit in de lift
  - bedienen van smartphones
  - vertalen op de smartphone
  - vragen van patiënten
  - optische waarnemingen (bijv. liplezen)
  - Digitale butlers: Alexa (Amazon), Siri (Apple)
    - emotie voorspelling
  - callcenters: training van algoritmes
    - Ademhaling, stress, toonhoogte, inhoud

## Facebook en Google

- Hebben van iedereen veel informatie
  - Wat komen ze niet te weten?
- Kunnen dan onbeperkt voorspellen
  - Zoals huisarts of politie functies
- Wat heb je nodig, wie zijn je vrienden
- Combinatie van gegevens, foto's en stem

## SyRI: bij voorbaat verdacht.nl

## Integrated Joint Operations Platform

### China (Xinjiang) stopt Oeigoeren in kampen (Volkskrant)

#### *Op basis van:*

- Klagen over religie of ethnische afkomst
- Oproep om te bidden
- In buitenland geweest
- Een baard laten staan
- Een bepaalde app op de telefoon
- Vaak de achterdeur van hun huis gebruiken
- Computer voorspelt 24.412 namen
- waarvan er 15.683 zijn opgepakt

## 6 (USA) dominante spelers

- Google/Alphabet: Google Brain, Magenta, Deepdream
  - Youtube, android apps, "stay on the website"
- Apple: appstore, hardware, butler Siri
- Amazon: cloud, butler Alexa, retail
  - Gaat leveren zonder te bestellen
- IBM: Deepmind, Watson for oncologie, debating

- Microsoft: cloud
- Facebook: Oculus Rift, face recognition, deeptext
  - Whatsapp, Instagram

*China heeft de USA ingehaald, met eigen spelers*

*Europa is nergens te bekennen..... behalve bij regelgeving*

### Stap voor stap Machine Learning

- Bedenk label (wat zoek ik) en kenmerken
- Bouw een algoritme (wiskunde)
- Geef véél voorbeelden: big data (incl. resultaat)
- Algoritme berekent in veel stappen de
  - Correlatie tussen alle kenmerken (m2, postcode, etc.)
  - wordt geleerd : *zelf lerend.....*
- En gebruikt dat om nauwkeurig te voorspellen
  
- Labels zoals intelligentie, politiek voorkeur, lievelingskleur..

### Wat had ik niet verwacht?

- Overall aanwezig
- 95% voorspellend
- Geen kennis van correlaties nodig
- Discriminatie
  - vooroordeel – bias
- duizenden kenmerken
- Netflix advies

### Wat hoop ik?

- Betere zorg in de Gezondheid
  - diagnose, voorspelling
- Veiligheid
  - bijv. kans op recidive
- False positive, false negative: wat is erger?

### Waar ben ik bang voor?

- Geen transparantie:
  - Wie? Welke voorbeelden? Waarom?
  - Koppelingen?
- Voorsprong van China en USA (big 5)
- Autonoom gebruik
- Sociale controle

### ML Leren met de toppers

- Amazon: Mlearning, AWS
  - 10.000+ klanten
- Google: crash course + app: tensorflow
- IBM: Watson Studio + ML for dummies
- Microsoft: Azure

- Apple: Core ML 2
  - natural language, vision
- Facebook: heeft veel te verbergen

Wat is wat?

