

VAN RISICO NAAR KWALITEIT VAN AI

Sander Mol

Werkgroep Testen en AI

Waarschuwing: deze presentatie gaat over AI
en bevat dus kattenplaatjes en Skynet.



Deze presentatie op YouTube:

<https://www.youtube.com/watch?v=yuhDuCxZMuA>

Onze vorige presentatie op YouTube:

https://www.youtube.com/watch?v=iHG64_Czmps

Al ons andere werk:

<https://www.testnet.org/testnet/p000610/werkgroepen/werkgroep-testen-en-ai>





Hoe gaan testrobots ons testers helpen?



Kunnen testtools ook
explorerend testen en
resultaten beoordelen?

Rik Marselis



Sander Mol

SAUVES

Onze tweede lente **Waar staat AI in het testvak nu?**

Gerald de Vrieze
Hannie van Kooten
Marco Verhoeven
Okto Wahjuwibowo
Sander Mol



Testen van AI

Wat maakt AI testen anders



Werkgroep Testen en AI
• Peter Collewyn
• Sander Mol – Werkgroep Voorzitter

Onze werkgroep

Rik
Marselis

Martin
van Helden

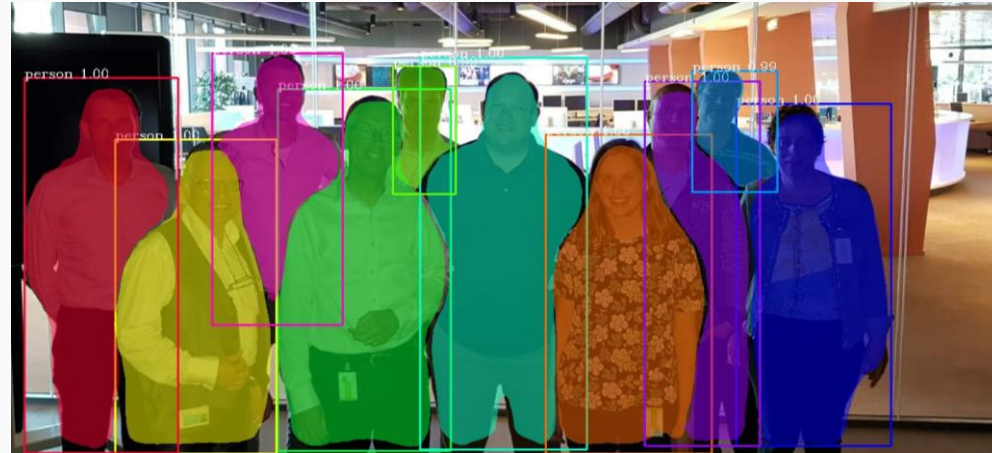
Marco
Verhoeven

Richard
van Emmerik

Mariëlle
van der Sluys

Johannes
Sim

Peter
Collewijn



Gerline
van Lieburg

Hannie
van Kooten

Sander
Mol

Leden

Agenda

De basis: Machine Learning en Deep Learning

Vijf risico's van AI

Hoe test je AI

De basis: Machine Learning

Ons startpunt: een vast algoritme

De voorspelde huizenprijs in euro's =

$700 * \text{woonoppervlak} +$
 $500 * \text{perceeloppervlak} +$
 $8.000 * \text{aantal slaapkamers} +$
score op basis van de postcode



De basis: Machine Learning

Een nieuw startpunt: data

Woon-oppervlak	Perceel-oppervlak	Slaap-kamers	Postcode	Prijs
110	173	4	3311 AA	400.000
104	145	6	3351 ES	375.000
122	211	5	3352 VA	450.000
...	...	...	...	...

De basis: Machine Learning

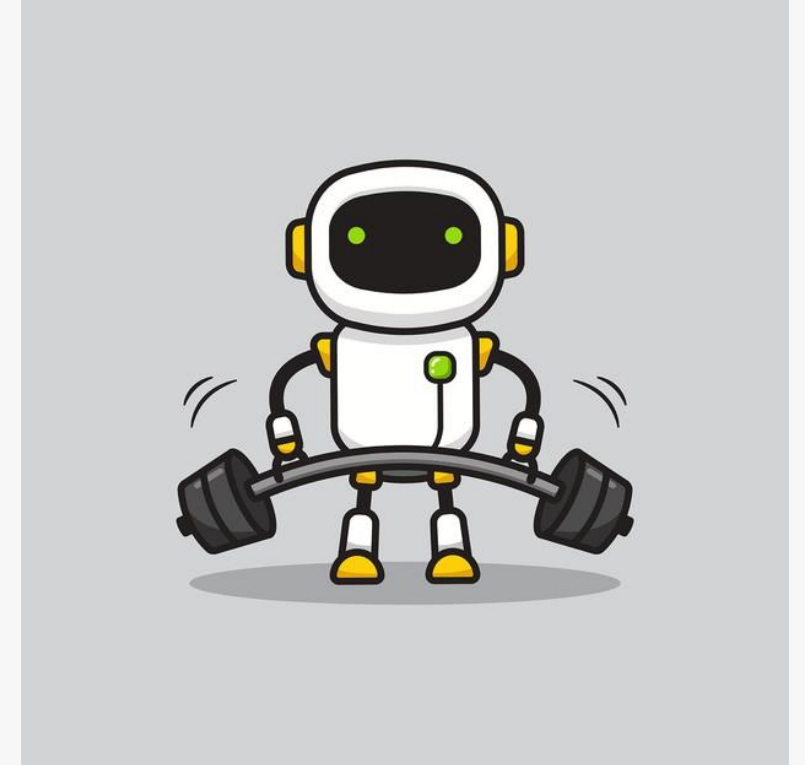
Een lerend algoritme



De basis: Machine Learning

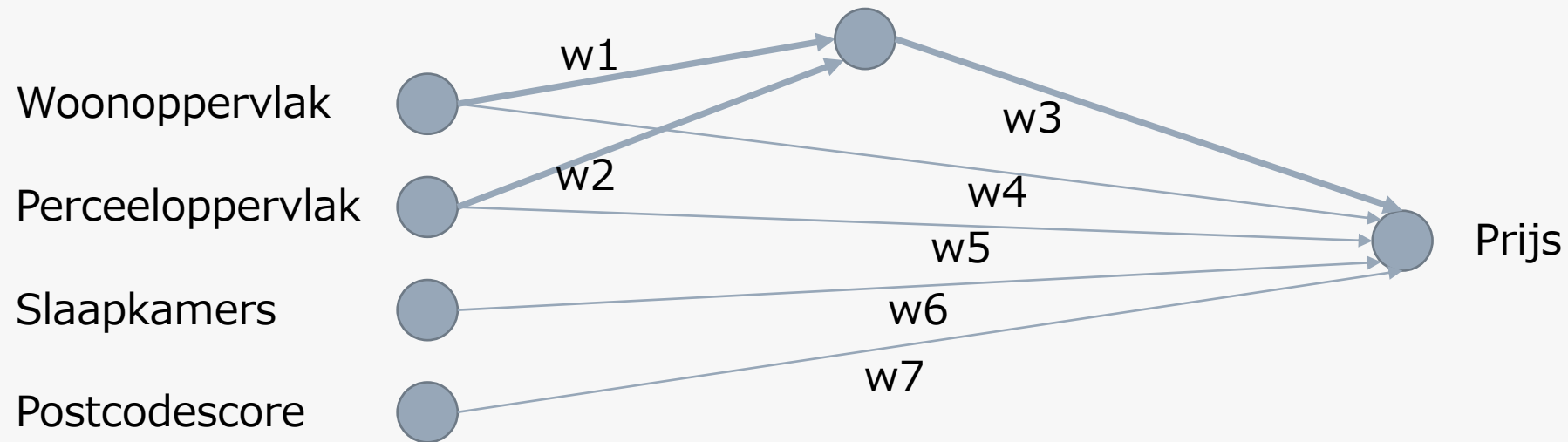
De juiste gewichten leren

- Vul willekeurige gewichten in
- Herhaal het volgende
 - Bereken de prijs op basis van deze gewichten
 - Vergelijk de berekening met de juiste prijs
 - Stel de gewichten een klein beetje bij
 - Kijk of de voorspelde prijs iets beter wordt
 - Bewaar de beste gewichten



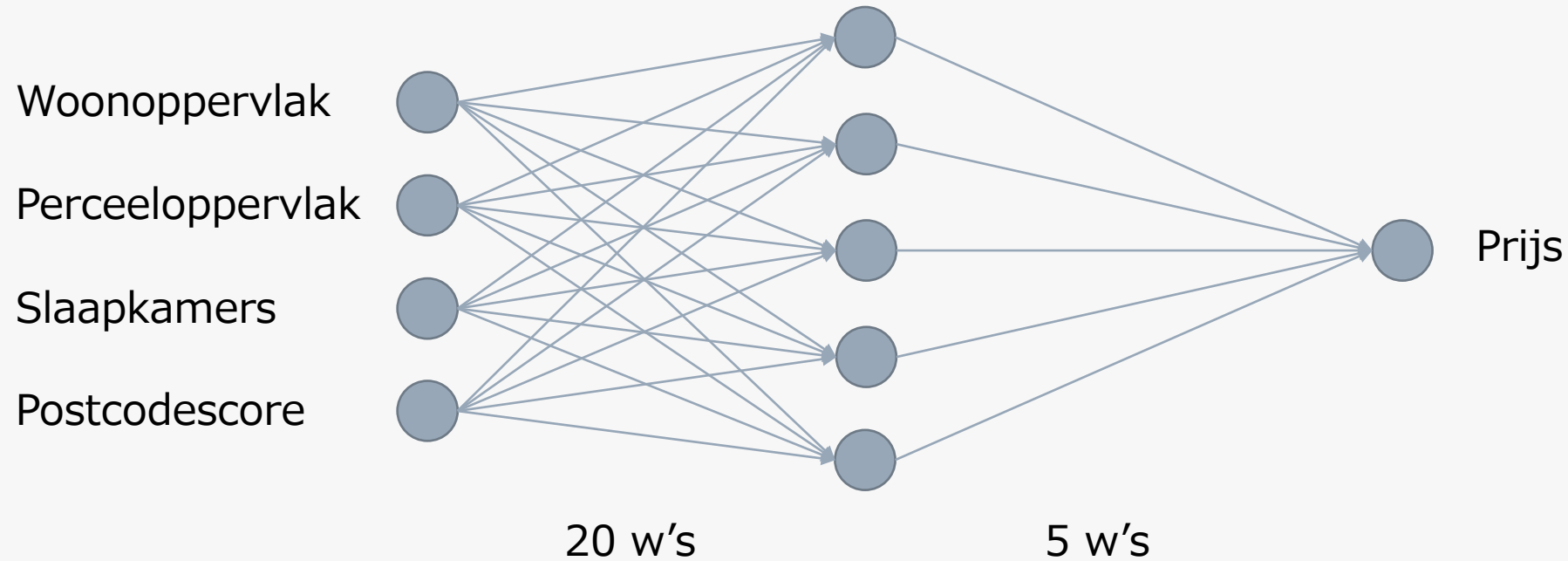
De basis: Deep Learning

Een lerend algoritme



De basis: Deep Learning

Een lerend algoritme



Vijf risico's van AI (ML)

Risico 1: onzekere uitkomsten

Bij huizenprijzen (dus 'regressie'): de variabelen gaan een model in en er komt ongeveer een logische prijs uit. Maar geen exacte verwachte waarde.

Bij kattenplaatjes: het plaatje gaat het model in en er komt een zekerheidspercentage uit. Dit is nooit 100%! Dus wanneer is het goed genoeg?



96% kat
42% hond
12% vis
2% auto

Vijf risico's van AI

Risico 2: afhankelijkheid van data

Data is vaak:

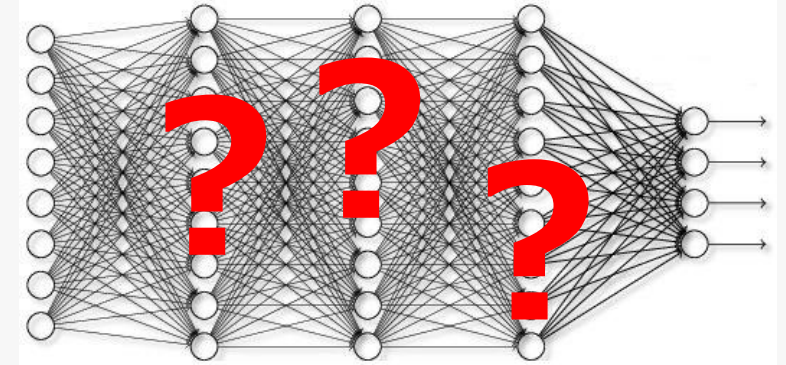
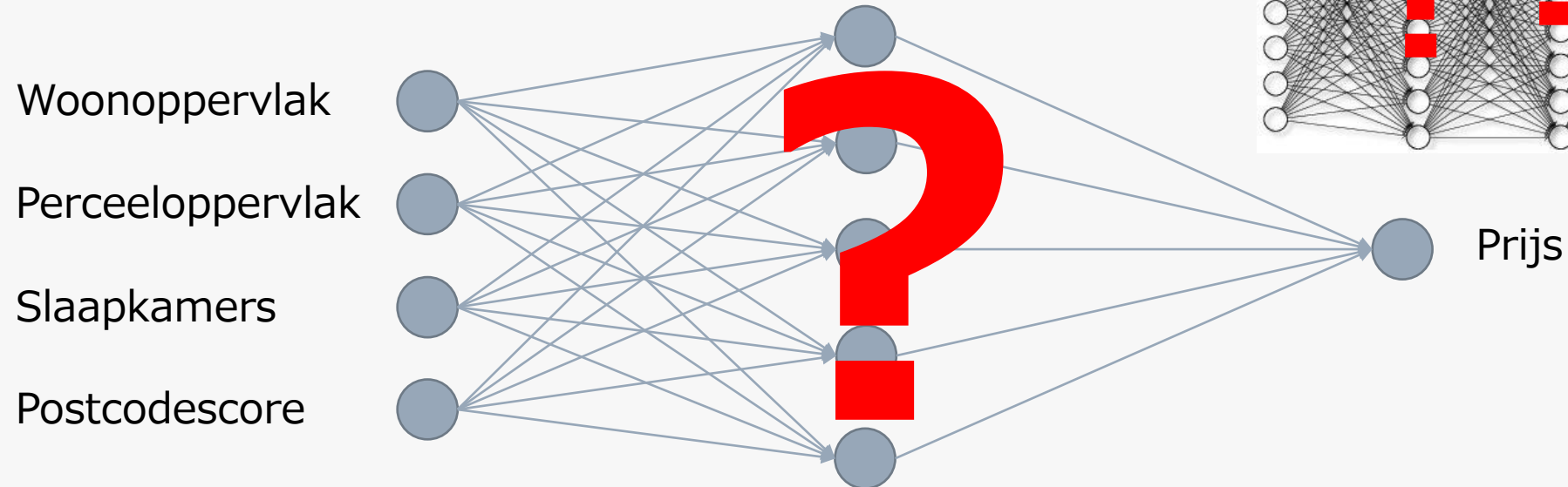
- verouderd
- incompleet of onjuist
- selectief verzameld
- vatbaar voor interpretatie
- en het toont soms patronen aan die we niet zo leuk vinden



64% kat ?

Vijf risico's van AI

Risico 3: beperkte uitlegbaarheid



Vijf risico's van AI

Risico 4: veranderende werkelijkheid of behoefte

Veranderende werkelijkheid: huizenprijzen! Maar ook wetgeving, trends (EK?), taalontwikkeling ("skeer"?), verandering in assortiment, enzovoort.

Veranderende behoefte: "Handig die herkenning van autoruitschade, doe dat ook maar voor bussen, woningen en kassen."



Vijf risico's van AI

Risico 5: algemene angst voor AI

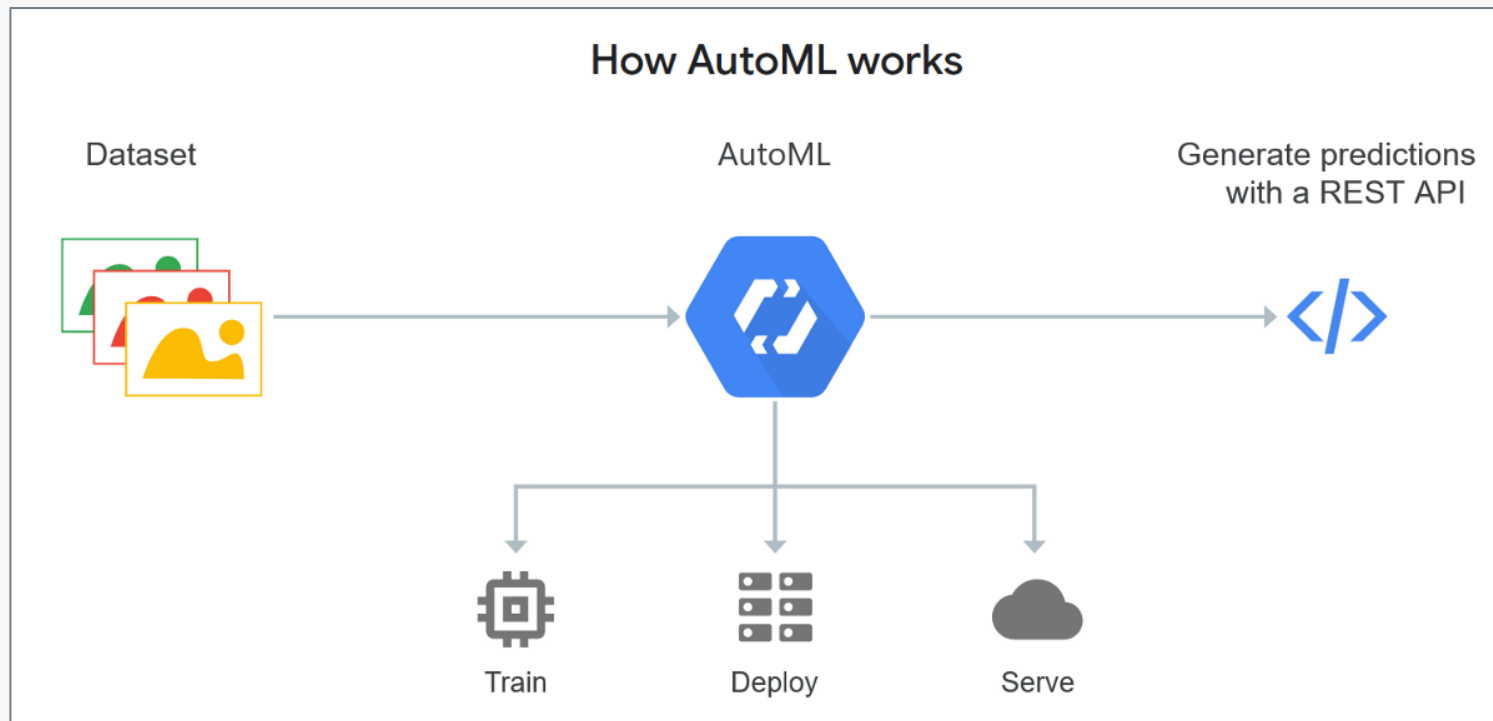
Soms door eerdere **slechte ervaringen** met te snelle AI implementaties, of door kennis van wat **ongeremde toepassing** kan doen.

Soms op basis van **onwetendheid**, eventueel gecombineerd met **doemscenario's**: stél dat de AI niet alleen kan leren, maar ook kan redeneren, context kan begrijpen en zelf normen ontwikkelt en doelen kan stellen.



Vijf risico's van AI

Bonusrisico: snel te maken, lastig te testen



Hoe test je AI (ML)

Test je data!

Stel alle belangrijke vragen over data:

- Is het actueel en representatief?
- Interpreteert iedereen de data hetzelfde?
- Zijn de bewerkingen logisch en consequent?
- Enzovoort.

Welk type muis zoekt u?

- Maak een keuze - ▲

Rechtshandig
Linkshandig
Neutraal

- Maak een keuze - ▼

Rechtshandig ▼

Hoe test je AI

Hoe test je representatie:

Controleer of alle veldwaarden (in- en uitvoer) goed vertegenwoordigd zijn. Controleer ook de verschillen tussen test- en trainingset.

Ga eventueel een stap dieper, bijvoorbeeld met pairwise testen.

Je kunt ook in gesprek met 'de business' over wat een goede dataverdeling is. Dat kan met de Data Combinatie Test!



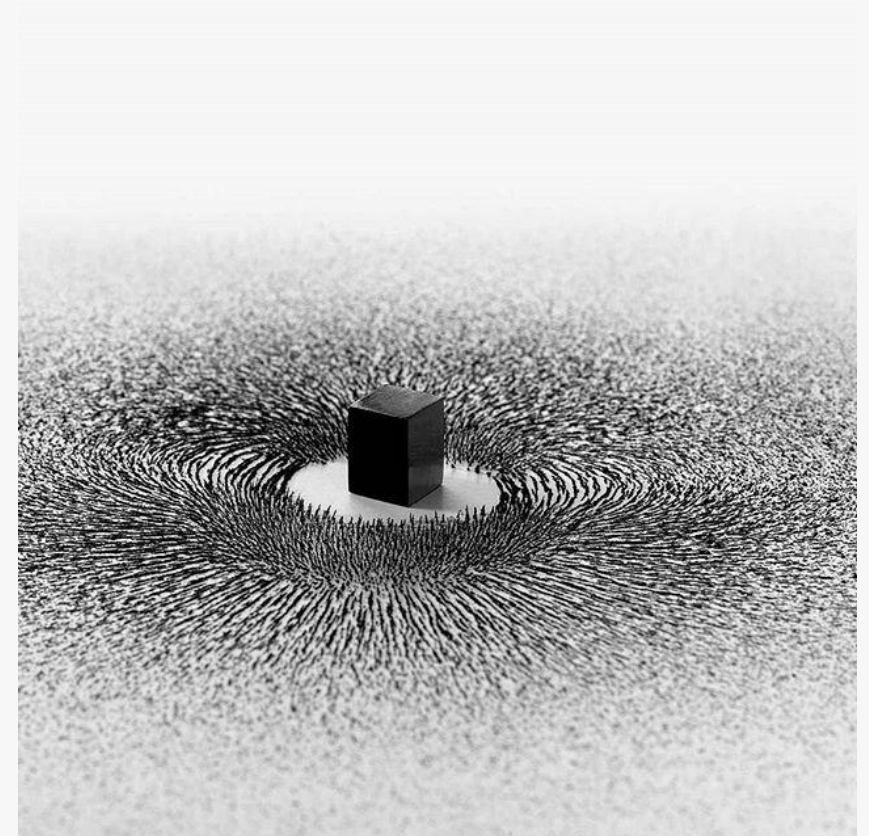
(een representatief businesspaar)

Hoe test je AI

Test geen uitkomsten, maar patronen

Pas de invoerdata een klein beetje aan en kijk of de uitvoer goed mee verandert. Dus bij huizenprijzen: bij een **groter perceeloppervlak** zou je een **hogere prijs** verwachten.

Metamorphic testing: het testen van verbanden.



Hoe test je AI

Test op robuustheid

Afhandelen van extreme waarden. Ofwel 'safety nets'.

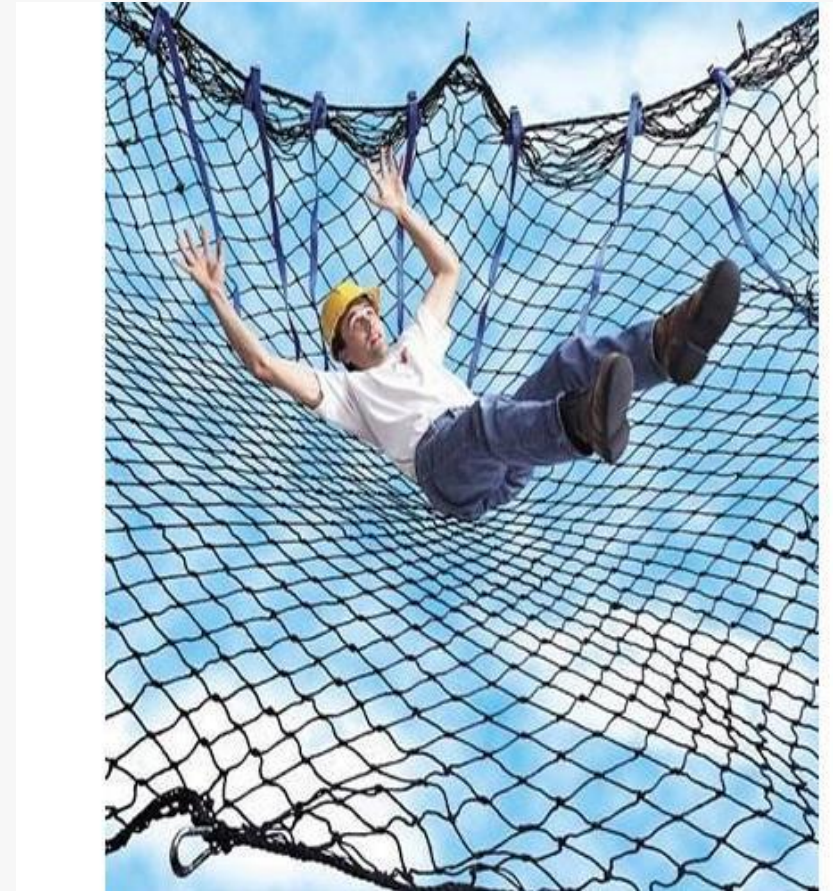
Herkenning als iemand de AI probeert te bespelen.

Test AI in de keten: worden alle in- en uitvoervarianten goed afgehandeld?

invoer: alle minimale gegevens ingevuld?

Logische combinatie van gegevens?

uitvoer: bevestigingsmail bij vergunning-aanvraag, melding bij geen passend product om aan te bevelen, enzovoort.



Hoe test je AI

Test op belangen

De AI kan alle mogelijke uitkomsten geven, zorg dat verschillende doelgroepen goed bediend worden op basis van voorbeelden.

Stel persona's op en koppel die aan use cases!



Hoe test je AI

Laat mensen meekijken!

Graven in data is heel leuk, maar uiteindelijk moeten er mensen mee werken. Laat eventuele gebruikers en hun managers het ontwikkelproces zien.

Zorg dat alle expertise betrokken is. Neem de ideeën van experts mee en adresseer hun zorgen.



Hoe test je AI

Neem de richtlijnen door

Er wordt al lang nagedacht over de voorwaarden om AI goed bruikbaar en eerlijk te maken.

- Binnen de overheid (EU, Nederland)
- Binnen branches en beroepsverenigingen
- Binnen organisaties

Ja, er is zeker uniformiteit nodig, maar er zijn ongetwijfeld nuttige tips uit te halen! Bovendien zullen de overheidsrichtlijnen ooit de basis voor wetten en regels worden.



Einde: nog even alles op één slide



Risico's

- Onzekere uitkomsten.
- Afhankelijk van data.
- Beperkte uitlegbaarheid.
- Veranderende werkelijkheid of behoefte.
- Algemene angst voor AI.
- (Snel te maken, lastig te testen.)

Hoe test je AI

- Test geen uitkomsten, maar patronen.
- Test je data.
- Test op robuustheid.
- Test op belangen.
- Laat mensen meekijken.
- Neem de richtlijnen door.