

BASIS VOOR BELEID

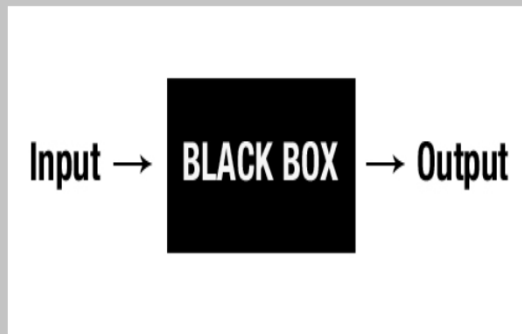


Narly Rambharos

www.os-groningen.nl

Impact door samenwerking

- Δ **Rol O&S**
- **Onderzoeksfunctie \approx / \neq Informatiefunctie?**
- **Big data – Data Science – ML – algoritme**



- **Data driven → policy making → beslissing burger !**

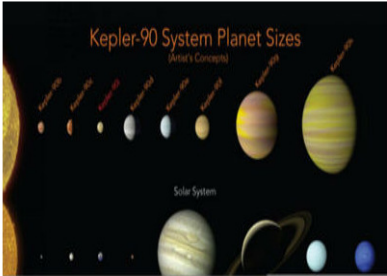
Data-driven policy making

- **Wiskunde + statistiek → non-profit beleid maken?
Hoe? Nut?**

News • Artificial intelligence finds solar system with 8 planets like ours

Artificial intelligence finds solar system with 8 planets like ours

AFP | Updated: Dec 15, 2017, 12:22 IST



Kepler-90 System Planet Sizes
(Artist's Conception)

Solar System

HIGHLIGHTS

- The eight-planet system — the largest known outside of ours — orbits a star called Kepler 90 some 2,545 light-years away.
- However, none of the planets are expected to be hospitable to life.
- NASA calculated its average temperature at about 800 degrees Fahrenheit (426 Celsius) — as hot as Mercury, the closest planet to the Sun.

This illustration made available by NASA shows a comparison of the planets in the solar system and those or... Read More

MIAMI: A solar system with as many planets as our own has been discovered with the help of NASA's [Kepler space telescope](#) and Artificial Intelligence (AI), the US space agency said Thursday.

Hoe werkt big data-technologie bij de politie?

© VRUJDAG, 18:40 AANGEPAST VRUJDAG, 18:40 POLITIEK



Risicomodel Sociale Zekerheidsfraude (SZ)

- **Voorspellingsmodel: opsporen SZ-fraude**
- **Waarom?**
 - **'Traditionele' vorm fraude opsporen**
 - **Handhaving 'richting geven'**
 - **Objectief oordeel**
 - **Sneller, slimmer, nieuwe kennis**

Risicomodel Sociale Zekerheidsfraude (SZ)

- **Samenstelling analysebestand**
- **Operationalisatie begrip fraude**
- **Betrokkenheid handhaving**
- **'Statistisch profileren'**

Risicomodel Sociale Zekerheidsfraude (SZ)

- **Samenstelling analysebestand**
 - **Fraude-administratie, BRP, uitkeringsgegevens**
 - **Databronnen: ontsluiten en koppeling**
 - **80 procent totale duur project**
 - **Pre-processing fase**

Risicomodel Sociale Zekerheidsfraude (SZ)

➤ Operationalisatie begrip fraude

▪ Wel

- schending inlichtingenplicht
- verzwijgen werkzaamheden en/of inkomsten

▪ Niet

- 'zware' vorm van fraude: leefvorm, identiteit

Risicomodel Sociale Zekerheidsfraude (SZ)

➤ **Betrokkenheid handhaving**

- **Reden**
 - **resultaten duiden**
 - **gezamenlijk draagvlak**

- **Brainstormsessie (begin)**

- **Periodiek overleg (tussentijds)**

Risicomodel Sociale Zekerheidsfraude (SZ)

- **‘Statistisch profileren’**
 - **Gemeenschappelijke ‘features’ van fraudeurs en niet-fraudeurs**
- **Model; beschrijving complexe werkelijkheid**
- **Actueel bestand, objectief inschatten fraude/geen fraude**

Samenwerking: effect !

➤ **Bevindingen**

- **Nieuwe kennis (statushouders)**
- **Rest: 'was iets mee aan de hand'**
 - **beëindiging uitkering**
 - **waarschuwing**

Samenwerking: effect !

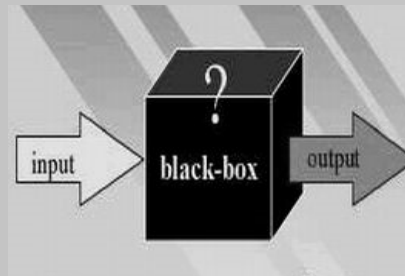
- **'Data-driven policy making' alleen o.b.v. 'berekeningen'?**
- **Extractie kennis = wiskunde + statistiek + domeinkennis**
- **Technologie: menselijke rol niet te vervangen!**

Aanbevelingen meer effect?

➤ Doorontwikkeling voorspelmodel SZ

- **Verrijken databestand**
- **Fraude-begrip herdefiniëren**
- **Andere algoritmen (toepassing)**

➤ Inzicht 'black box'

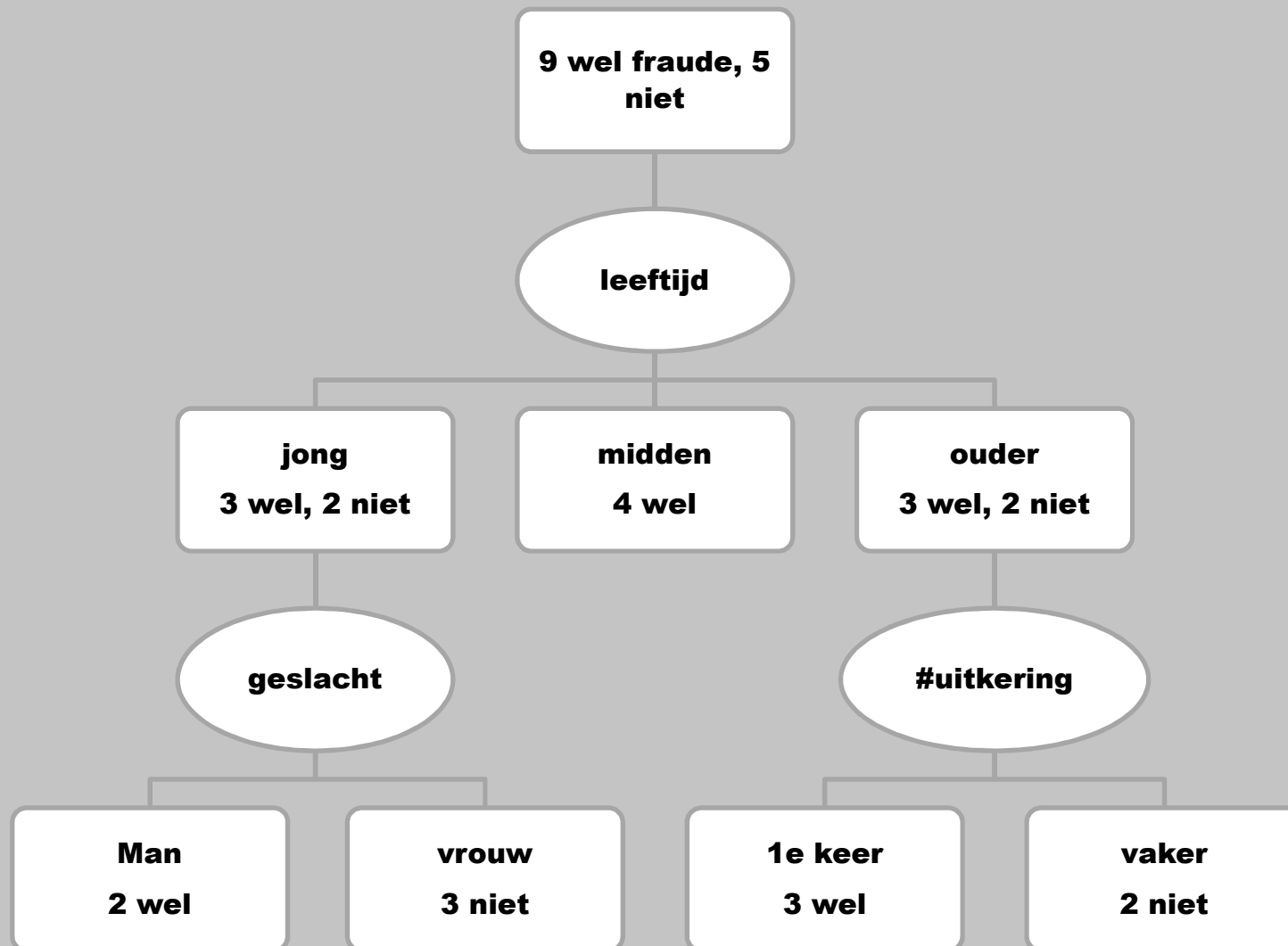


Algoritme 'Decision tree'

➤ Algoritme

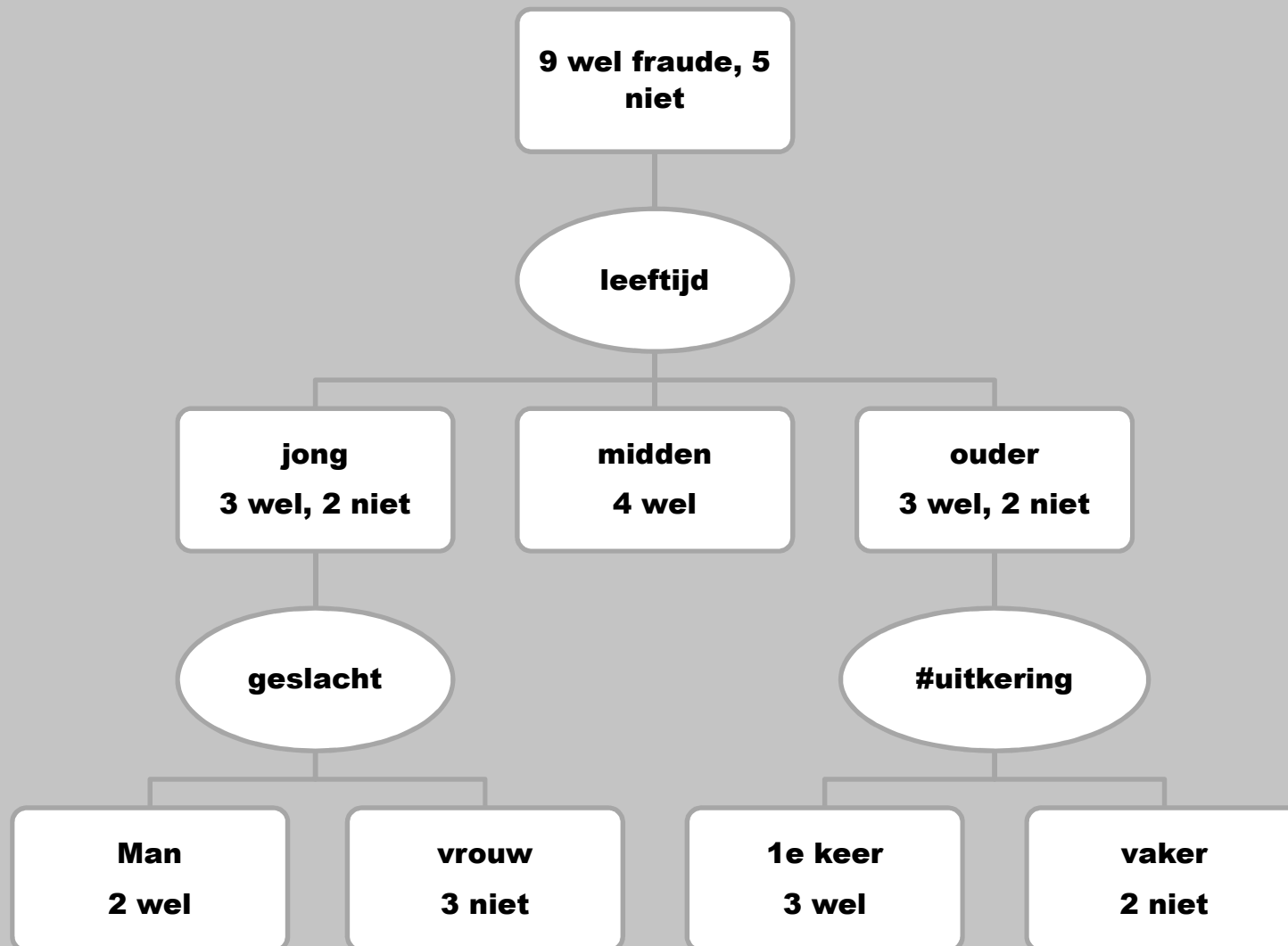
- 'reeks van instructies'
- voorbeeld Decision Tree

Klant	Leeftijd	Geslacht	# Keren in de uitkering	Wel/geen fraude
1	jong	vrouw	1	nee
2	jong	vrouw	> 1	nee
3	midden	vrouw	1	ja
4	ouder	vrouw	1	ja
5	ouder	man	1	ja
6	ouder	man	> 1	nee
7	midden	man	> 1	ja
8	jong	vrouw	1	nee
9	jong	man	1	ja
10	ouder	man	1	ja
11	jong	man	> 1	ja
12	midden	vrouw	> 1	ja
13	midden	man	1	ja
14	ouder	vrouw	> 1	nee



Decision tree; samenvattend

- **Telt # fraudeurs en # niet-fraudeurs**
- **Dataset: verdelen in deelsets**
- **Zuivere set (allemaal 'ja' of 'nee'): stop**
- **Niet-zuivere set: doorgaan**
- **Nieuwe cliënt: classificatie?**



Decision tree

➤ **Complex**

- **Unbalanced data**
- **Majority class: beter voorspellen**
- **Overfitting; 'snoeien tree'**
- **Model performance (uitleg)**



university of
 groningen

Narly Rambharos
PhD-student

BASIS VOOR BELEID



Vragen ?