

**Aan** 5.1.2.e  
**Cc** 5.1.2.e  
**Van** 5.1.2.e  
**Onderwerp** Beoordeling algoritmes, AI systeem AI Verordening  
**Datum** 16 januari 2025

## Inleiding

In het programma Zicht op Ondernijning wordt met behulp van data-analyse inzicht verkregen in de onderliggende patronen van ondernijnende criminaliteit.

In het kader van het programma Zicht op Ondernijning worden 2 algoritmes gebruikt die binnen het programma ontwikkeld zijn. Het betreft een algoritme dat een inschatting maakt van welke jongeren mogelijk risico lopen om in de drugscriminaliteit terecht te komen en een algoritme dat een inschatting maakt van welke woningen een verhoogd risico lopen om misbruikt te worden als hennepkwekerij. De uitkomsten van deze algoritmes worden gebruikt voor het inzichtelijk maken van patronen die kunnen duiden op crimineel ondernijnd gedrag. Deze data-analyses kunnen gemeenten gebruiken voor het ontwikkelen van preventiebeleid. Voor deze analyses wordt gebruik gemaakt van data van het CBS. Eén van de voorwaarden voor het gebruik van deze data is dat de analyses alleen gebruikt worden voor preventie en niet voor opsporing en strafvervolgning. Verder zijn de inzichten van één van de twee modellen (risicojongeren) ook gebruikt om een aanvraag voor het programma Preventie met Gezag te doen. Gebruikers van de data-analyses vragen zich af hoe deze algoritmes gezien moeten worden in relatie tot de eisen van de AI act.

### 1. Vraag

Met de vraag die de gebruikers gesteld hebben willen zij het volgende weten.

1. Zijn de 2 algoritmes te zien als een AI systeem in de zin van de AI act?
  - a. Zo ja, in welke risico categorie vallen de 2 algoritmes?

### 2. Antwoord

- a. De algoritmes zijn te zien als AI systeem in de zin van de AI act (artikel 3 onder a AI act)
- b. De algoritmes vallen niet onder verboden praktijken. (artikel 5 AI act)
- c. De algoritmes vallen niet onder hoog risico (artikel 6 AI act)
- d. De algoritmes vallen niet onder de transparantieplichtingen (artikel 50 AI act)
- e. De algoritmes vallen niet onder AI voor algemene doeleinden (artikel 51 AI Act General purpose AI)

### 3. Aanbevelingen

- Hoewel de algoritmes gebruikt worden om tot een goede informatiepositie te komen, is het aan te raden te blijven uitleggen wat de waarde is van de informatie die je met de algoritmes verkrijgt. Hoe je deze informatie kan gebruiken en hoe beslist niet.
- Blijf in de gaten houden dat het algoritme nog steeds doet wat het moet doen.

### 4. Toelichting

De toelichting is als volgt opgebouwd

- a. Definitie AI systeem

- b. Werking algoritmes
- c. Beoordeling AI systeem

#### *Ad 4a Definitie AI systeem*

“AI-systeem”: een op een machine gebaseerd systeem dat is ontworpen om met verschillende niveaus van autonomie te werken en dat na het inzetten ervan aanpassingsvermogen kan vertonen, en dat, voor expliciete of impliciete doelstellingen, uit de ontvangen input afleidt hoe output te genereren zoals voorspellingen, inhoud, aanbevelingen of beslissingen die van invloed kunnen zijn op fysieke of virtuele omgevingen. (artikel 3, onder 1 AI Act)

#### *Ad 4b Werking algoritmes*

De algoritmes waarover vragen zijn gesteld zijn random forest algoritmes. Random Forest-algoritme is een tree learning-techniek in Machine Learning. Het werkt door een aantal beslisbomen te creëren tijdens de trainingsfase. Elke beslisboom wordt geconstrueerd met behulp van een willekeurige subset van de dataset om een willekeurige subset van features in elke partitie te meten. Deze willekeurigheid introduceert variabiliteit tussen individuele beslisbomen, waardoor het risico op overfitting wordt verminderd en de algehele voorspellingsprestaties worden verbeterd.

Bij voorspellingen verzamelt het algoritme de resultaten van alle beslisbomen, hetzij door te stemmen (voor classificatietaken) of door te middelen (voor regressietaken). Dit collaboratieve besluitvormingsproces, ondersteund door meerdere beslisbomen met hun inzichten, biedt een voorbeeld van stabiele en nauwkeurige resultaten. Random forests worden veel gebruikt voor classificatie- en regressiefuncties, die bekend staan om hun vermogen om complexe gegevens te verwerken, overfitting te verminderen en betrouwbare voorspellingen te bieden in verschillende omgevingen. Gedetailleerder beschreven werkt een random forest algoritme als volgt.

**Ensemble van beslissingsbomen:** Random Forest maakt gebruik van de kracht van ensemble learning door een leger van beslissingsbomen te construeren. Deze beslisbomen zijn als individuele experts, die zich elk specialiseren in een bepaald aspect van de data. Belangrijk is dat ze onafhankelijk opereren, waardoor het risico dat het model te veel wordt beïnvloed door de nuances van een enkele beslisboom, wordt geminimaliseerd.

**Random Feature Selection:** Om ervoor te zorgen dat elke beslissingsboom in het ensemble een uniek perspectief biedt, maakt Random Forest gebruik van random feature selection. Tijdens de training van elke beslisboom wordt een willekeurige subset van kenmerken gekozen. Deze willekeurigheid zorgt ervoor dat elke boom zich richt op verschillende aspecten van de data, wat een diverse set van voorspellers binnen het ensemble bevordert.

**Bootstrap Aggregating of Bagging:** De techniek van bagging is een hoeksteen van de trainingsstrategie van Random Forest, die het maken van meerdere bootstrap-samples van de originele dataset omvat, waardoor instanties met vervanging kunnen worden bemonsterd. Dit resulteert in verschillende subsets van data voor elke beslissingsboom, wat variabiliteit in het trainingsproces introduceert en het model robuuster maakt.

**Besluitvorming en stemmen:** Als het gaat om het maken van voorspellingen, brengt elke beslisboom in het Random Forest zijn stem uit. Voor classificatietaken wordt de uiteindelijke voorspelling bepaald door de modus (meest voorkomende voorspelling) in alle beslisbomen. Bij regressietaken wordt het gemiddelde van de individuele boomvoorspellingen genomen. Dit interne stemmechanisme zorgt voor een evenwichtig en collectief besluitvormingsproces.<sup>1</sup>

Bij de algoritmes van Zicht op Ondernijning gaat het om classificatie.

---

<sup>1</sup> De uitleg van de werking van het algoritme is een vertaling van de tekst die te vinden is op <https://www.geeksforgeeks.org/random-forest-algorithm-in-machine-learning/>



#### Ad 4c Beoordeling AI systeem

De random forest algoritmes zijn aan te merken als AI systeem, omdat zij een zeker niveau van autonomie hebben en met de resultaten van de algoritmes de fysieke omgeving beïnvloed wordt. De elementen van de definitie zijn als volgt beoordeeld.

- a. De algoritmes zijn machinegebaseerd, want de algoritmes draaien op computerprogramma's op een computer/server
- b. *De algoritmes functioneren in enige mate autonoom. Want het algoritme bepaalt de rekenregels van het model. Er wordt gebruik gemaakt van een machine learning techniek waarbij het algoritme (o.a. de beslisbomen) wordt gegenereerd op basis van willekeurige features en datasets in de trainingsdata. Hier heeft de mens geen invloed op. De datascientist kan alleen bepaalde parameters meegeven tijdens het trainen het algoritme*
- c. *Het ontwikkelde algoritme kan zichzelf niet aanpassen. Het algoritme wordt eenmalig getraind en is niet in staat zichzelf aan te passen op basis van nieuwe input.*
- d. *De doelstellingen van de algoritmes zijn hetzelfde, want de programma's waarin de algoritmes ontwikkeld en toegepast worden zien beiden toe op het voorkomen en signaleren van ondermijning*
- e. *De output van de algoritmes wordt gebruikt om beleid te maken dat bedoeld is om de fysieke omgeving te beïnvloeden*

*De elementen autonomie en beïnvloeding van de omgeving worden gezien als de belangrijkste kenmerken van een AI systeem.*

*Op basis van de toelichting van de werking van de algoritmes (4b) en de beoordeling AI systeem (4c) concludeer ik dat de algoritmes die gebruikt worden bij Zicht op Ondermijning vallen onder de definitie AI systeem van de AI act.*

#### 5. Beoordeling risico categorie

Nu beargumenteerd is dat de algoritmes als AI systemen in de zin van de AI act te zien is, beoordeel ik in welke risico categorie de algoritmes vallen.

Hiervoor heb ik gekeken naar verboden praktijken, hoog risico, transparantievereisten en AI-modellen met algemene doeleinden.

#### Verboden praktijken

Van de verboden praktijken heb ik specifiek gekeken naar social scoring, als zijnde praktijk die wellicht van toepassing zou kunnen zijn.

Artikel 5 eerste lid, aanhef en onder c van de AI act is als volgt.

De volgende AI praktijken zijn verboden:

(...)

c het in de handel brengen, het in gebruik stellen of het gebruiken van AI-systemen voor de evaluatie of classificatie van natuurlijke personen of groepen personen gedurende een bepaalde periode *op basis van hun sociale gedrag of bekende, afgeleide of voorspelde persoonlijke of persoonlijkheidskenmerken, waarbij de sociale score een of beide van de volgende gevolgen heeft:*

- I. de nadelige of ongunstige behandeling van bepaalde natuurlijke personen of groepen personen in een sociale context die *geen verband* houdt met de context waarin de data oorspronkelijk werden gegenereerd of verzameld;
- II. de nadelige of ongunstige behandeling van bepaalde natuurlijke personen of groepen personen die *ongerechtvaardigd of onevenredig* met hun sociale gedrag of de ernst hiervan is;

De output van de algoritmes is nooit te herleiden naar natuurlijke personen omdat de resultaten altijd geaggregeerd worden op gemeente-, wijk- en buurtniveau. Verder worden er enkel gegevens gepubliceerd wanneer er in het betreffende gebied minstens 10 jongeren/woningen als risicocategorie worden aangemerkt. Ook is de output van de algoritmes nooit te herleiden naar groepen mensen waarvan je weet dat ze dezelfde eigenschappen delen.

Daarom vallen de algoritmes naar mijn mening niet onder verboden AI praktijken.

### *Hoog risico*

De algoritmes zijn geen veiligheidscomponent of product waarvan EU wetgeving geldt. (artikel 6, eerste lid AI act)

Van de hoog risico op grond van artikel 6 tweede lid, heb ik gekeken naar rechtshandhaving.

AI-systemen met een hoog risico op grond van artikel 6, lid 2, zijn de vermelde AI-systemen op de volgende gebieden:

- 6      Rechtshandhaving, voor zover het gebruik ervan is toegestaan op grond van het toepasselijke Unierecht of nationale recht:
  - (...)
  - d  
AI-systemen die bedoeld zijn om door of namens rechtshandhavingsinstanties of door instellingen, organen of instanties van de Unie ter ondersteuning van rechtshandhavingsinstanties te worden gebruikt om te beoordelen hoe groot het risico is dat een natuurlijke persoon (opnieuw) een strafbaar feit zal plegen, niet uitsluitend op basis van profilering van natuurlijke personen als bedoeld in artikel 3, punt 4, van Richtlijn (EU) 2016/680, of om persoonlijkheidskenmerken en eigenschappen of eerder crimineel gedrag van natuurlijke personen of groepen te beoordelen;

De output van de algoritmes wordt gebruikt voor preventie en niet voor rechtshandhaving. Ook wordt de output van de algoritmes niet gebruikt ten aanzien van natuurlijke personen.

Daarom vallen de algoritmes naar mijn mening niet onder hoog risico.

### *Transparantieverplichtingen (artikel 50 AI act)*

De output van de algoritmes worden niet gebruikt voor het aanbieden van een vorm van content. De mededeling dat content met AI gemaakt is, hoeft om die reden dan ook niet gedaan te worden. (artikel 50, tweede lid AI act)

De output van de algoritmes worden niet gebruikt:

1. In de directe interactie met natuurlijke personen (artikel 50, eerste lid AI act)
2. Voor het herkennen van emoties of biometrische categorisering (artikel 50, derde lid AI act)
3. Voor deep-fake uitingen (artikel 50, vierde lid AI act)

De output van de algoritmes worden niet gebruikt op de manier zoals beschreven in artikel 50 AI act. Er hoeft om die reden niet voldaan te worden aan artikel 50 van de AI act.

### *AI modellen voor algemene doeleinden (artikel 51 AI act)*

De algoritmes vallen op geen enkele manier onder de definitie AI model voor algemene doeleinden.

Definitie artikel 3 onder 1 AI act

“AI-systeem”: een op een machine gebaseerd systeem dat is ontworpen om met verschillende niveaus van autonomie te werken en dat na het inzetten ervan aanpassingsvermogen kan vertonen, en dat, voor expliciete of impliciete doelstellingen, uit de ontvangen input afleidt hoe output te genereren zoals voorspellingen, inhoud, aanbevelingen of beslissingen die van invloed kunnen zijn op fysieke of virtuele omgevingen.

De volgende onderdelen van de definitie moet je beoordelen om tot de conclusie te komen of iets een AI systeem is of niet. De tekst onder a is het relevante gedeelte van overweging 12 AI act.

1. Machine gebaseerd
  - a. De term “op een machine gebaseerd” verwijst naar het feit dat AI-systemen op machines draaien.
2. Verschillende niveaus van autonomie
  - a. AI-systemen worden zodanig ontworpen dat zij in verschillende mate autonoom kunnen functioneren, wat betekent dat zij een zekere mate van onafhankelijkheid van menselijke betrokkenheid bezitten en zonder menselijke tussenkomst kunnen functioneren.
  - b. Autonoom: Een onderdeel van een systeem dat in staat is het oorspronkelijke domein, doel, of gebruik te veranderen zonder externe interventie, controle of toezicht.  
ISO/IEC 22989:2022(en) onderdeel 3.1.5
3. Aanpassingsvermogen
  - a. Het aanpassingsvermogen dat een AI-systeem na het inzetten ervan kan vertonen, heeft betrekking op zelflerende capaciteiten, waardoor het systeem tijdens het gebruik kan veranderen.
4. Expliciete of impliciete doelstellingen
  - a. De verwijzing naar expliciete of impliciete doelstellingen onderstreept dat AI-systemen kunnen functioneren volgens expliciete, gedefinieerde doelstellingen, of volgens impliciete doelstellingen. De doelstellingen van een AI-systeem kunnen verschillen van het beoogde doel van het AI-systeem in een specifieke context.
    - i. Context: Voor de toepassing van deze verordening moeten onder omgevingen de contexten worden verstaan waarin de AI-systemen werken, terwijl de output die door het AI-systeem wordt gegenereerd een uiting is van de verschillende functies van AI-systemen en de vorm kan aannemen van voorspellingen, content, aanbevelingen of besluiten.
5. Invloed op fysieke of virtuele omgevingen
  - a. Dit inferentievermogen slaat op het proces waarbij output, zoals voorspellingen, content, aanbevelingen of besluiten, wordt verkregen waarmee fysieke en virtuele omgevingen kunnen worden beïnvloed, en op het vermogen van AI-systemen om modellen of algoritmen, of beide, af te leiden uit input of data. De technieken die inferentie mogelijk maken bij de opbouw van een AI-systeem, omvatten benaderingen op basis van machinaal leren waarbij aan de hand van data wordt geleerd hoe bepaalde doelstellingen kunnen worden bereikt, alsook op logica en kennis gebaseerde benaderingen waarbij iets wordt geïnfereerd uit gecodeerde kennis of uit een symbolische weergave van de op te lossen taak. Het inferentievermogen van een AI-systeem overstijgt de elementaire verwerking van data door leren, redeneren of modelleren mogelijk te maken.